Deciphering the domain-specific functions of polymerase θ during interstrand crosslink repair in *Drosophila melanogaster*

AN HONORS THESIS FOR THE DEPARTMENT OF BIOLOGY

Robin Armstrong

Tufts University, 2014

Table of Contents

Abstract

DNA polymerase θ (Pol θ) is upregulated in many human cancers which is associated with a poor prognosis for these cancer patients. The *D. melanogaster* homologue of Pol θ is encoded by *mus308* and contains two domains, an N-terminal helicase-like domain and a C-terminal polymerase domain, connected by a central, non-conserved linker region. Unlike human Pol θ mutants, which are extremely radiosensitive, *D. melanogaster mus308* mutants demonstrate extreme sensitivity to interstrand crosslinking (ICL) agents, such as nitrogen mustard (HN2), but are not radiosensitive. In order to investigate the domain-specific functions of Pol θ during ICL repair, we generated a host of Pol θ domain-specific mutants. Our HN2 sensitivity studies with ATPase-dead and polymerase-dead transgenic flies have demonstrated that the helicase-like domain of Pol θ is essential for the repair of HN2 induced ICLs while polymerase-dead flies were only moderately sensitive to ICL agents. Flies which express truncated Pol θ protein expressing part of the linker region and either only the N-terminus or C-terminus are equally extremely sensitive to HN2. These suggests that the C-terminus of Pol θ could have some function in recruitment of other proteins to ICLs outside of its catalytic function. It has been shown that the sensitivity to HN2 is partially rescued in flies which express the N- and C-termini of the protein in trans, suggesting that there is not an absolute requirement that the two catalytic domains be part of the same polypeptide. Our data support a model in which the helicase-like domain of Pol θ functions upstream of the polymerase domain and is crucial for repair of HN2 induced ICLs. Furthermore, we propose that the presence, but not the catalytic function, of the polymerase domain is also required for proper repair of ICLs.

Introduction

I. DNA Repair

The cell is constantly experiencing DNA damage, including single-strand breaks, double-strand breaks (DSBs), thymine dimers, and mismatches (Lindahl, 1993; Figure 1). These various types of DNA lesions occur endogenously on a regular basis and are readily repaired by the cell. A few examples of sources of endogenous damage include oxidative free radicals, which are often derived from oxygen consumed during metabolism in the mitochondria, topoisomerase failures, and mechanical stress (Lieber, 2010; Handy and Loscalzo, 2012). Lesions can also be induced in a laboratory setting using DNA damaging agents, including chemicals and radiation. Each specific DNA damaging agent can help to pinpoint whether or not a protein of interest is involved in the particular repair process being targeted. For example, hydroxyurea (HU) partially depletes the pool of dNTPS in a cell resulting in stalled replication forks (Slater, 1973). Treatment of a genetic mutant with HU, and subsequent analysis of survival, can help determine if the mutated protein is involved in fork restart mechanisms. It is critical that the repair mechanisms programmed by the cell are working properly, for without these repair mechanisms in place, genomic instability results in the form of insertions, deletions, chromosomal translocations, and mutations. As mutations are passed on through successive cycles of cellular division, this genomic instability could result in dangerous extremes for the cell and organism, such as cellular apoptosis and cancer.

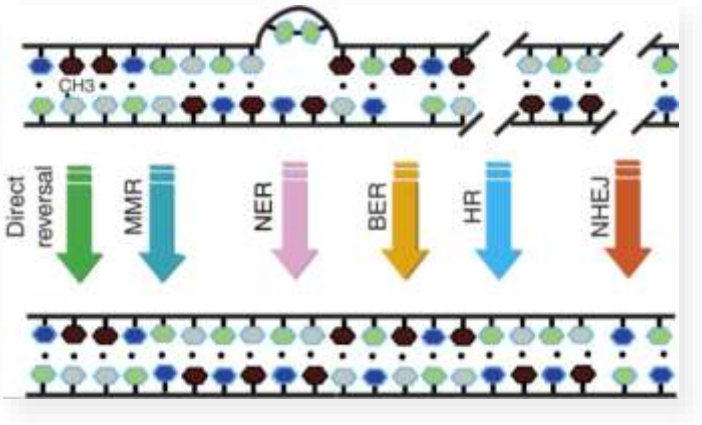


Figure 1. Representation of the various types of DNA damage experienced and repair mechanisms employed by cells. Repair mechanisms programmed by the cell such as direct reversal, mismatch repair (MMR), nucleotide excision repair (NER), base excision repair (BER), homologous recombination (HR) and non-homologous end-joining (NHEJ) are specific for particular types of damage and are constantly working to maintain genomic stability (Modified from Hakem 2008).

A. Double-strand break (DSB) repair

DSBs serve as one of the most severe types of DNA damage as genetic information on both strands of the double helix is damaged, therefore, genetic sequences flanking the break are often lost during DSB repair events. Consistent with other mechanisms of DNA repair, inaccurate DSB repair can contribute greatly to genomic instability; however, defective DSB repair results in cellular toxicity at a much higher scale than the repair of mismatches, pyrimidine dimers, abasic sites, and single-stranded breaks. A single DSB in a molecule of DNA, if repaired inaccurately, can result in the deletion of millions of base pairs, a consequence which the cell cannot afford (Helleday et al., 2007). DSBs can exist in three different forms: two-ended DNA DSBs, one-ended DNA DSBs, and daughter strand gaps (Figure 2).

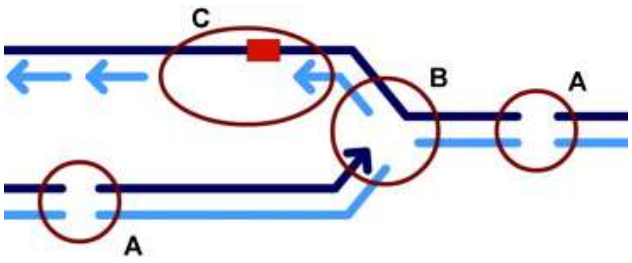


Figure 2. Classifications of DSBs. A) A two-ended DSB created via a direct fracture in the helix. B) A one-ended DSB created when a replication fork encounters a single-stranded break. C) A daughter strand gap created when a lesion is encountered during replication. (Modified from Helleday et al., 2007).

Although a DSB is often visualized as a direct fracture of the double helix, one-ended DNA DSBs and daughter strand gaps, derived from lesions encountered during the DNA replication processes on both the leading and lagging strands, make up a significant portion of the total number of DSBs generated in the cell (Helleday et al., 2007). When small lesions or single-strand breaks are encountered by the replisome during DNA replication, these lesions are often converted to highly toxic DSBs due to fork stalling and collapse at these lesions (Strumberg et al., 2000). Organisms manage the potentially toxic effects of DSBs with a host of DSB repair mechanisms with ranging degrees of mutagenicity (Figure 3). The mutagenic nature of DSB repair has been exploited by a variety of carefully regulated physiological processes, in addition to repair of lesions, such as genetic recombination during meiosis and immunoglobulin class switching (Radding, 1973; Stavnezer, 1996). Deficiencies in DSB repair genes, therefore, can result in sterility, immunodeficiency, and neurodegenerative disorders in addition to the adverse effects experienced with inaccurate repair of lesions. The cell utilizes different mechanisms for the repair of DSBs, including homologous recombination (HR) and non-homologous end-joining (NHEJ). These two repair process are the main mechanisms of DSB repair employed by the cell, although both HR and NHEJ encompass a broad number of alternative pathways.

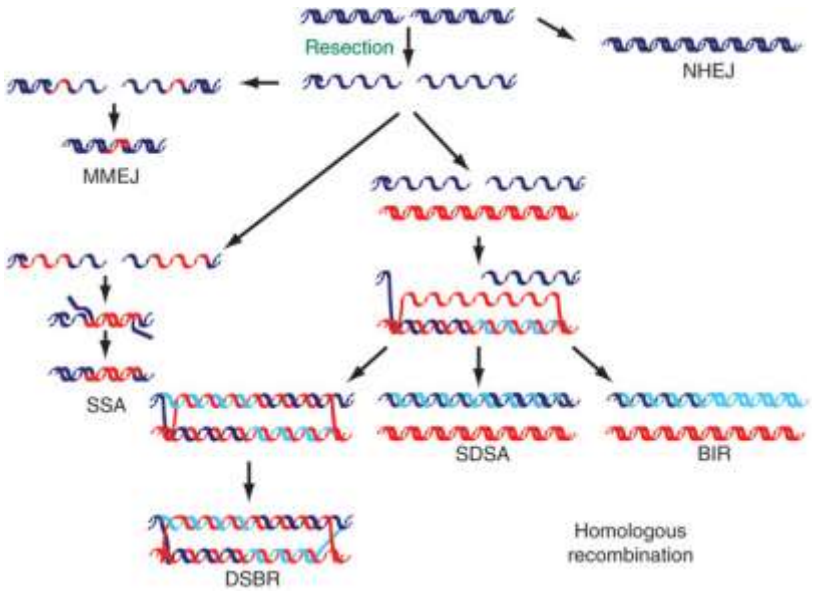


Figure 3. Mechanisms of repair of DNA double-strand breaks. The two main DSB pathways are end-joining (EJ) and homologous recombination (HR). EJ includes both non-homologous end-joining (NHEJ), microhomology-mediated end-joining (MMEJ), and alternative end-joining (alt-EJ, not pictured). HR includes single-strand annealing (SSA), double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), and break-induced replication (BIR). (Modified from Huertas, 2010)

Homologous Recombination

In homologous recombination, error-free repair of the DSB can be accomplished as long as the homologous template strand of DNA is undamaged (Li and Heyer, 2008). HR is a conserved mechanism of DSB repair, although it has been suggested that its primary function is to resolve stalled or collapsed replication forks (Aguilera and Gómez-González, 2008). All of the sub-pathways of HR in mammals begin with a MRN complex-mediated 5' to 3' degradation of DNA flanking the DSB which yields flanking single-stranded 3' overhangs, although additional proteins are also involved in this resection; the MRN complex is composed of Mre11, Rad50, and Nbs1 (Williams et al., 2007). The 3'single-stranded overhang of the double-stranded DNA is bound by the heterotrimeric replication protein A (RPA) which, with the help of mediator proteins Rad55, Rad57, and Rad 52 in yeast, is replaced by the protein Rad51 (Heyer, 2007; Gasior et al., 1998). RPA binds ssDNA before replacement by Rad51 as RPA displays a higher

binding affinity and specificity for ssDNA than Rad51 (Wold, 1997). Following these presynapsis steps, the sub-pathways of HR diverge to follow separate mechanisms (Figure 3).

DSBR, SDSA, and BIR: Mechanisms of HR which invade a homologous template

In three of the sub-pathways of HR, the Rad51-DNA filament complex invades its homologous template strand in search of homology (Heyer, 2007). After invasion into a homologous template, the classical double-strand break repair (DSBR) route involves formation of a small displacement loop (D-loop) which enlarges until significant complementarity between the invading strand and the template is achieved (Szostak et al., 1983). The complementary regions anneal to each other and synthesis off of the template initiates; concurrently, a second round of repair synthesis initiates from the second 3' overhang catalyzed by Rad52 in yeast (BRCA2 in mammals) (Szostak et al., 1983; Sugiyama et al., 2006). Completed synthesis from the invading 3' overhang initiates branch migration and, thus, formation of the second Holliday junction (Szostak et al., 1983). Together, two rounds of single-stranded repair synthesis repair the double-strand gap, and the resolution of the two Holliday junctions can result in various recombination gene products with the end goal of two separate, undamaged homologous chromosomes, although DSBR can also occur between sister chromatids (Heyer, 2007). Dissolution of the double Holliday junction by mammalian BLM-TOPOIIIα would result in the formation of non-crossover products while structure-specific endonuclease cleaving of the double Holliday junction would yield both crossover and non-crossover products (Plank et al., 2006; Liu et al., 2004). This is likely the mechanism for DSB repair during meiotic recombination.

Synthesis-dependent strand annealing (SDSA) repairs DSBs differently than classical DSB repair with the avoidance of cross-over products as the most striking difference, a

characteristic which makes SDSA more attractive for DSB repair in somatic cells than classical DSBR (Li and Heyer, 2008). During SDSA repair of a DSB, the two single-stranded overhangs invade the homologous template independently of each other, either as one-end or two-end invasion events, forming local bubbles which synthesize off of the homologous template as opposed to large, expanding D-loops (Nassif et al., 1994). Human RTEL1 has been postulated to disrupt D-loop formation, promoting SDSA over classical DSBR (Barber et al., 2008). After BLM further dissolves double Holliday junctions of the D-loop, it is believed that RECQ5 helicase prevents the reformation of the Rad51 filament (Paliwal et al., 2014). Synthesis off of the homologous template is followed by rapid collapse of the local invasion bubble and subsequent displacement of the newly synthesized DNA (Nassif et al., 1994). The partially synthesized strand is free to reanneal, and extension and repair through the break at the original duplex occurs (Nassif et al., 1993).

When only one end of a DSB shares homology with its template, as is the case with a stalled or nicked replication fork, break-induced replication (BIR) can serve to reinitiate DNA replication (Morrow et al., 1997). It is believed that BIR repairs DSBs in a fashion similar to replication fork restart with resection, coating by Rad51, and subsequent single Holliday junction formation (Li and Heyer, 2008). Because BIR repairs at a one-sided DSB during replication, the second end of the DSB is never engaged due to lack of template homology, resulting in the loss of all genetic information on that fragment (Malkova et al., 1996). Additional factors to those involved in canonical DSB repair are required for BIR repair of a one-sided DSB at the replication fork including the Pol32 subunit of polymerase δ and Pif1 helicase (Wilson et al., 2013). Studies in budding yeast have demonstrated that the Shu complex, which is comprised of Shu1-Psy3-Shu2-Csm2 and shows both structural and functional homology to mammalian

Rad51 paralogs, is necessary for the formation of recombination intermediates which are subsequently processed by Sgs1-Top3-Rmi1; this suggests that the Shu complex may be involved in Rad51 filament formation during BIR repair of one-sided breaks in the replication fork (Mankouri et al., 2007; Martin et al., 2006)

SSA: Repair of a DSB in repeat sequences without a homologous template

One error-prone sub-pathway of HR, single-strand annealing (SSA), does not require a homologous template during repair of DSBs located between two repeated or homologous sequences; the neighboring homologous sequences can be separated by as much as 15 kb (Lin et al., 1985; Paques and Haber, 1999). As SSA does not involve the invasion of a homologous template, it functions in DSB repair independent of Rad51 (Tutt et al., 2001). Long single-stranded overhangs generated via resection of DNA flanking a DSB, when exposed to neighboring complimentary regions, can re-anneal and copy off of these homologous sequences (Paques and Haber, 1999). Repair in this fashion often results in deletions flanking the DSB, a repair product characteristic to SSA (Paques and Haber, 1999).

In the cell, a DNA template for repair of a DSB is not always readily available; therefore, error-prone mechanisms of DSB repair outside of HR such as classical non-homologous end-joining (C-NHEJ) and alternative end-joining (alt-EJ), two genetically distinct mechanisms of NHEJ, are used.

<u>End-Joining</u>

End-joining repair of a DSB is characterized by direct ligation of the two ends of the DSB which, due to lack of homology search characteristic of HR, can result in a variety of mutation-ridden repair products. Chromosomal translocations serve as the predominant driver for approximately 20% of cancer cases (Mitelman et al., 2007). C-NHEJ and alternative pathways

of end-joining have been shown to readily induce genomic instability in the form of chromosomal translocations in the absence of key repair factors such as BRCA1, which is required for HR, and the FA complex, which is required for interstrand crosslink (ICL) repair (Pace et al., 2010; Bunting et al., 2010). When end-joining repair processes function as the dominant DSB repair pathway, direct ligation of DNA fragments from separate chromosomes or separate locations on the same chromosome largely contributes to genomic instability. Understanding C-NHEJ and its alternative sub-pathways allows for a greater comprehension of how error-prone DNA repair pathways shape the genome of cancer cells.

C-NHEJ: Repair of a DSB through direct ligation broken ends

C-NHEJ repair of a DSB is a lower fidelity repair process than the HR sub-pathways as small insertions, deletions, or genomic translocations can accompany repair of DSBs via C-NHEJ (Chan et al., 2010). C-NHEJ utilizes two complexes of proteins: one complex nonspecifically binds and protects the blunt ends of the DSB from degradation followed by the recruiting of necessary processing enzymes, and a second core complex of proteins catalyzes the ligation of the processed ends to repair the DSB; small deletions or insertions can result from this repair process (Lieber et al., 2008). In C-NHEJ, the Ku70-Ku80 (XRCC6-XRCC5) heterodimer binds the damaged DNA ends, and the Ku heterodimer serves as a docking station for the complex of DNA-dependent protein kinase catalytic-subunit (DNA-PKcs) and processing nuclease (Artemis), polymerases (mu and lambda), and ligase (DNA ligase IV) proteins (Lieber, 2008; Lieber, 2010). Processing of each DNA strand via Artemis continues until some small amount of annealing between strands is possible which allows the two strands to be ligated together by DNA ligase IV and repaired (Lieber, 2010).

Alt-EJ: Repair of a DSB in the absence of core C-NHEJ components

In the absence of one or more C-NHEJ components, alt-EJ takes over as the end-joining mechanism of choice, relying on microhomologies for the repair of DSBs (McVey et al., 2008). As this subset of end-joining is approximately 20-fold less efficient than C-NHEJ in yeast, alt-EJ almost always results in deletions or chromosomal translocations featuring microhomologies, and, thus, may promote chromosome instability (Boulton and Jackson, 1996; McVey et al., 2008; Difilippantonio et al., 2002). Resection by CtIP of the flanking ends of the DSB produces single-strand overhangs allowing subsequent unwinding of a longer stretch of double-stranded DNA until pre-existing microhomologies are exposed (Zhang and Jasin, 2011; Chan et al., 2010). The Ku70-Ku80 complex is thought to inhibit the onset of HR in DSB repair; therefore, it is possible that, in the absence of Ku70-Ku80, alt-EJ is able to search for microhomologies in a similar mechanism to how a damaged chromosome searches for homologies during HR (Boulton and Jackson, 1996). Annealing at microhomologies allows for synthesis which can result in repair products with either insertions or deletions of genetic material (Figure 4). Final steps involving cleavage of generated flaps is uncertain; however, there is evidence for involvement of both Ligase I and Ligase III during final ligation steps of alt-EJ, although the exact mechanism of action remains uncharacterized (McVey, 2008).

It has been postulated that DNA polymerase θ (Pol θ) performs dual functions in alt-EJ repair of a DSB, both the helicase activity necessary for initial unwinding of double-stranded DNA and the translesion polymerase activity necessary for synthesis at annealed microhomologies; these findings in *Drosophila*, coupled with studies in *C. elegans*, have led to the coining of alt-EJ in *C. elegans* as theta-mediated end joining (TMEJ) (Chan et al., 2010; Roerink et al., 2014). During alt-EJ repair of a DSB, Pol θ binds double-stranded DNA and

initially unwinds short stretches of the duplex to expose pre-existing microhomologous sequences; alignment between the two strands of damaged DNA is mediated by these microhomologies, and the polymerase activity of Pol θ is free to begin synthesis (Figure 4).

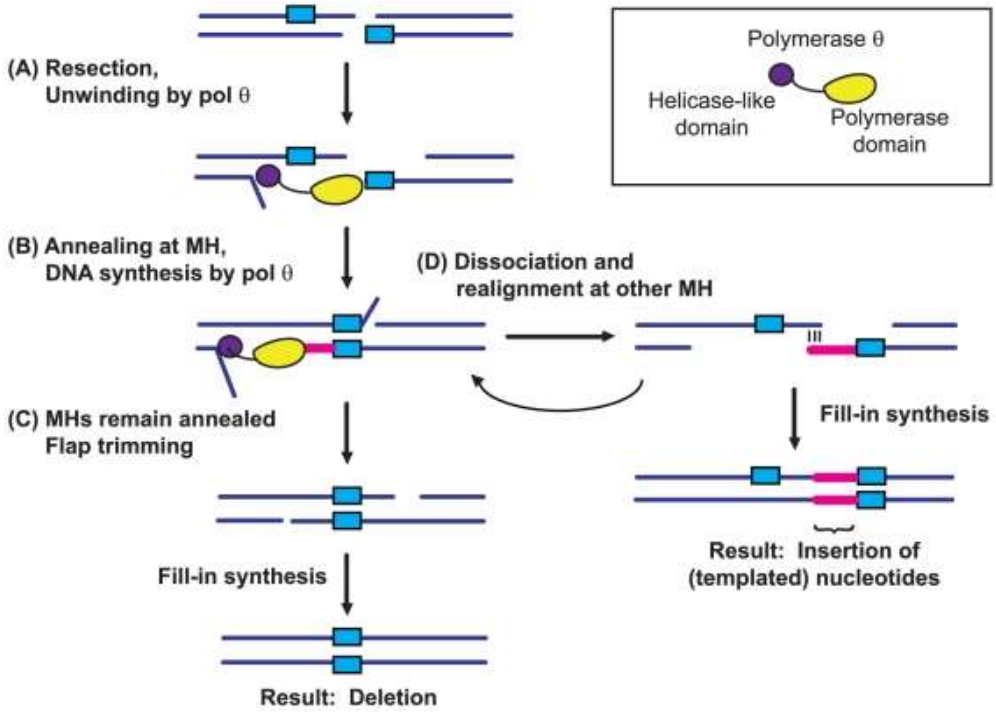


Figure 4. Mechanism of DNA polymerase θ-dependent alternative end-joining (alt-EJ). In the absence of one or more C-NHEJ proteins (Ku70-Ku80, DNA ligase IV, Artemis, or DNA-PKcs) alt-EJ predominates as the end-joining mechanism of choice. Resection flanking the DSB is followed by Pol θ mediated unwinding of double-stranded DNA. Unwinding continues until microhomologies are exposed; annealing at microhomologies and subsequent Pol θ –dependent synthesis can result in deletion products (C) or templated insertion products (D) (Modified from Chan et al., 2010).

B. Interstrand Crosslink (ICL) Repair

ICLs are a particularly harmful form of DNA damage as the two strands of DNA become covalently bound to one another, preventing transcription as well as strand separation and polymerase progression during replication. A single unrepaired ICL in either bacteria or yeast can result in cell death, whereas approximately 40 unrepaired ICLs in repair-deficient conditions can result in mammalian cell death (Lawley and Brookes, 1968; Grossmann et al., 2001; Lawley

and Phillips, 1996). Natural sources of ICLs include by-products of lipid peroxidation which increase from alcoholism or a high-fat diet, abasic sites, natural psoralens, and oestrogens (Stone et al., 2008; Brooks and Theruvathu, 2005; Dutta et al., 2007; Bennets et al., 2008). ICL agents such as nitrogen mustards, mitomycin C (MMC), platinum compounds, and psoralens have been used at varying times as chemotherapeutic drugs since the mid-20th century; however their exact mechanisms of action did not begin to be understood until well after their induction of use. ICL repair appears to involve a coordination of various repair mechanisms including nucleotide excision repair (NER), translesion synthesis (TLS), and HR; understanding the exact mechanisms and key players involved in ICL repair is essential to inform potential cancer therapeutics.

ICL Repair: An overview of mechanism during S-phase with a single replication fork

During S-phase of the cell cycle, duplication of the genome takes place. If a replicative polymerase encounters an ICL upon progressing along the template, replication fork stalling results (Figure 5). Stalling of the replication fork at the ICL and subsequent nicking on either side of the crosslink by NER proteins XPF-ERCC1 leads to the formation of a one-ended DSB. The Mus81-Eme1 complex may be involved in the conversion of an ICL to a one-ended DSB (Hanada et al., 2006). Unhooking of the ICL allows for TLS across the gap of the excised crosslink; the adduct on the other strand can be removed by NER proteins and replicative polymerases are used to restore strand integrity (Niedernhofer et al., 2005). The MRN complex initiates end resection to generate a 3' single-stranded overhang necessary for strand invasion of the sister chromatid; it has been postulated that the exonuclease function of WRN may contribute to end resection prior to strand invasion (Cheng et al., 2004). HR is initiated following end

resection in a Rad51and BRCA2-dependent manner to not only repair the DSB generated from unhooking the ICL, but to also restart the replication fork (Niedernhofer et al., 2005).

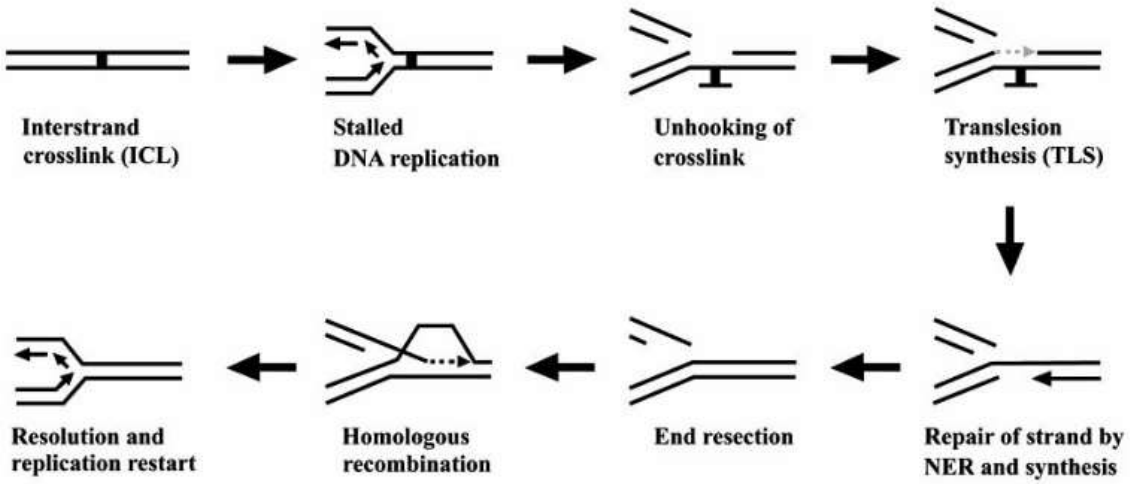


Figure 5. Outline of ICL repair during replication. When a replication fork approaches an ICL, the replication fork stalls. The ICL must be unhooked using nicking enzymes and TLS fills in the gap left by the unhooked crosslink. Removal of the ICL and gap-filling leaves a one-ended DSB that must be repaired via HR mechanisms (Modified from Andreassen and Ren, 2009).

Fanconi anemia (FA): A key to decoding the players involved in ICL

FA is a rare, autosomal recessive syndrome characterized by a predisposition for haematological and squamous cell carcinomas, low-fertility, congenital abnormalities, and progressive bone marrow failure (Fanconi, 1967). To date, fifteen complementation groups have been identified as functioning in the same FA pathway although eight appear to function as the monoubiquitin E3 ligase FA core complex (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM) (Alpi and Patel, 2009). Mutations in any of the fifteen complementation groups can result in FA although different mutations often result in differences in cancer susceptibility and disease onset. An overwhelming characteristic of FA cells is that they demonstrate an extreme sensitivity to ICL agents to the extent that FA is diagnosed through treatment of cells with ICL agents and subsequent analyses of those cells (Auerbach, 1988).

Additionally, in FA cells, it has been demonstrated that specific oncogenic genomic translocations are favored and that a single, unrepaired ICL could be sufficient for these translocations to take place (Jonnalagadda et al., 2005). FANC genes, therefore, are not only necessary for maintenance of genomic integrity but are also essential for repair of ICLs.

The FA complex recognizes an ICL via the heterodimeric FANCM-FAAP24 complex, and this recognition initiates FA core complex-mediated monoubiquitylation of FANCD2 and FANCI (Andreassen and Ren, 2009). Mutations in any of the FA core complex components disrupts the monoubiquitylation of FANCD2 and FANCI by destabilizing other FA core complex components (Andreassen and Ren, 2009). Additionally, mutations in FAAP24 and FAAP100, which associate with the FA core complex, result in the failure to monoubiquitylate FANCD2 and FANCI (Ciccia et al., 2007; Ling et al., 2007). Relocalization of the FANCD2-FANCI heterodimeric complex to DNA ICLs allows for well-coordinated downstream ICL repair (Kim and D'Andrea, 2012). Monoubiquitylated FANCD2 recruits SLX4 which serves as a scaffold for nucleases MUS81-EME1 and XPF-ERCC1 to unhook the ICL (Kim and D'Andrea, 2012; Ciccia et al., 2008). Unhooking of the ICL allows for TLS to restore the nascent strand, which is believed to be mediated by translesion polymerases Rev1 and polymerase ζ, although the specific polymerases involved have not been concretely identified (Kim and D'Andrea, 2012). The DSB which results from unhooking of the ICL is repaired via HR mechanisms; the FA complex functions to not only initiate HR, but to also suppress NHEJ (Bunting and Nussenzweig, 2010). BRCA2 supports RAD51 loading onto the RPA coated sing-stranded DNA overhang; RAD51 also prevents extensive Mre11 resection characteristic of a two-ended DSB as DSBs in ICL repair are one-ended (Moynahan et al., 2001; Hashimoto et al., 2010). Strand invasion allows for successful HR-mediated repair of this one-ended DSB.

Despite the insight that the FA complex has provided in terms of understanding ICL repair, there are still many holes in the ICL repair mechanism. Many of the roles of the FANC genes are uncertain, and the exact function of the FA core complex itself remains unclear. Therefore, further research involving the understanding of the protein players involved in ICL repair is necessary in order to generate therapeutic strategies which take advantage of ICL repair.

II. DNA Polymerases

DNA polymerases function to replicate DNA and are important during both DNA replication in S-phase of the cell cycle and during various DNA repair processes. There are fifteen well-established polymerases in eukaryotes to date, some of which demonstrate redundancy of function with one another (Hubscher et al., 2002). The eukaryotic family X polymerases are comprised of polymerases β, μ, σ, and λ and are involved in a wide range of DNA repair processes including base excision repair (BER) and NHEJ (Belousova and Lavrik, 2010). The eukaryotic B-family of polymerases include the high fidelity replicative polymerases, α, δ, and ε, which perform the bulk of DNA replication, and polymerase ζ, which is involved in TLS (Garg and Burgers, 2005). The Y-family eukaryotic polymerases include η, κ, Rev1, and ι which are characterized by their translesion synthesis activities. The A-family polymerases, including polymerases γ, ν, and θ, which show homology to *E. coli* DNA polymerase I, demonstrate significantly different functions from one another, and their functions are not completely understood. Of these polymerases, there is a variety in both context of function and fidelity of polymerase DNA synthesis (Table 1).

Polymerase	Family	Fidelity	Processivity	Replicative/ Translesion	Primary physiological functions
Pol γ	A	High	High	Replicative	Mitochondrial DNA replication BER
Pol ν	A	Low	Low	Translesion	TLS ICL repair
Pol θ	A	Low	Low	Translesion	TLS, ICL repair? BER, Alt-EJ
Pol α	B	High	Low	Replicative	Replication primase HR S-phase chkpoint
Pol δ	B	High	High	Replicative	Lagging strand synthesis MMR, BER, NER, DSBR
Pol ε	B	High	High	Replicative	Leading strand synthesis BER, NER, HR, S-phase chkpoint
Pol ζ	B	Low	High	Replicative	TLS ICL repair, HR?
Pol β	X	Low	Low	Translesion	BER
Pol μ	X	Low	Low	Translesion	V(D)J recombination NHEJ
Pol σ	X	Low	Low	Translesion	Sister-chromatid cohesion
Pol λ	X	Low	Low	Translesion	NHEJ, BER V(D)J recombination
Pol η	Y	Low	Low	Translesion	Bypass UV lesions HR?
Pol κ	Y	Low	Medium	Translesion	TLS NER
Rev1	Y	Low	Low	Translesion	Bypass abasic sites
Pol ι	Y	Low	Low	Translesion	TLS

Table 1. Functions and enzymatic activities of human DNA polymerases. Each of the fifteen human DNA polymerases is listed with corresponding family, fidelity, processivity, characterization, and primary function data.

A. Replicative DNA Polymerases

Replicative DNA polymerases are essential for duplication of the genomes of all living organisms and are also required for proper repair of DNA damage. These polymerases are constitutively expressed and are regulated via specific cell cycle signals in response to variable growth conditions. Replicative polymerases are categorized as high fidelity, which can be

attributed to their 3'—5' exonuclease proofreading activity and their active site specificity to choose the correct nucleotide, dNTPs over rNTPs (Loeb and Monnat 2008; Arana and Kunkel 2010). Successful duplication of the genome is attained via continuous leading strand synthesis and fragmented lagging strand synthesis. The major replicative polymerases employed by the cell during replication are polymerases α (alpha), δ (delta), and ε (epsilon); polymerase α serves as the primase, polymerase δ carries out lagging strand synthesis, and polymerase ε carries out leading strand synthesis (Nick McElhinny et al., 2010).

B. Translesion DNA Polymerases

Maintaining genomic integrity is critical for cell viability; however, the characteristics of replicative polymerases which enable them to be high fidelity also cause them to stall when encountering various forms of DNA lesions, necessitating a different type of polymerase. Therefore, upon encountering a bulky lesion, replicative polymerases are replaced temporarily by translesion polymerases which are capable of synthesizing DNA across lesions in an error-prone fashion due to their much larger active site pocket which can accommodate damaged DNA. Following a short stretch of TLS, the replicative polymerase can reload past the lesion to continue synthesizing off of the template strand. Incorrectly incorporated bases by the translesion polymerase can later be repaired through mismatch repair mechanisms as translesion polymerases involved in replication have an incorrect incorporation error rate of up to one in ten bases (McCulloch and Kunkel, 2008).

The primary translesion polymerases involved in TLS during replication are members of the Y-family of DNA polymerases, including polymerase η (eta), Rev1, polymerase κ (kappa), and polymerase ι (iota), in addition to polymerase ζ (zeta) of the B-family of DNA polymerases. Polymerase η is primarily responsible for synthesizing across thymine dimers, a DNA lesion

induced by UV radiation (Johnson et al., 1999). The active site of polymerase η can accommodate the UV-induced dinucleotide lesion, thus, allowing the correct dNTP, dA, to bind with both thymines of the dimer (Biertumpfel et al., 2010). Rev 1 incorporates a dC across from an abasic site and is believed to do so in a template-independent manner; the template dG is looped out and replaced by Rev1's own Arg324 residue for replication across specific bulky dG adducts (Nair et al., 2008). Polymerase κ is the most accurate of the Y-family polymerases on undamaged DNA and can replicate across many lesions but is blocked by dinucleotide lesions as its active site can only accommodate a single Watson-Crick base pair (Johnson et al., 2000; Bavoux et al., 2005; Lone et al., 2007). Polymerase ι has been implicated to function in promoting Hoogsteen base pairing, repair of pyrimidine dimers and base excision repair although its exact functions have yet to be elucidated (Nair et al., 2004; Vaisman et al., 2006; Petta et al., 2008). Polymerase ζ is largely considered an extender polymerase and is involved in a wide array of replication and repair processes (Shachar et al., 2009). Despite the seemingly one-step nature to TLS, many believe that TLS is truly a two-step process where translesion polymerases coordinate synthesis, both incorporation across the lesion and extension beyond the lesion, with one other (Shachar et al., 2009).

III. DNA Polymerase θ

DNA polymerase θ (Pol θ) is described as a proofreading deficient A-family DNA polymerase, which demonstrates homology to *Escherichia coli* pol I, and is believed to function in TLS (Sharief et al., 1999). However, Pol θ tends to incorporate single base errors at a frequency 10-100 times higher than other A-family DNA polymerases, characterizing the polymerase function of Pol θ as low fidelity, a characteristic more comparable to that of Y-family polymerases (Seki et al., 2004). This divergence from A-family polymerase activity can

most likely be attributed to the three unique sequence insert regions within the DNA polymerase sequence of Pol θ in some organisms, but not all (Seki et al., 2004). Pol θ can insert a nucleotide across from an abasic site, a thymine glycol lesion, and can extend off of a primer generated by another polymerase in certain instances if it cannot bypass the specific lesion itself (Arana et al., 2008). In fact, Pol θ is the only known enzyme which can carry out both the insertion and extension steps across from an abasic site (Seki et al., 2004). The polymerase domain of Pol θ is located in the C-terminus of the protein while the N-terminus has been demonstrated to contain seven motifs characteristic of RNA and DNA helicases, although no helicase activity has been reported to date (Harris et al., 1996). The N-terminal helicase-like domain and the C-terminal polymerase domain of Pol θ are separated by a large central domain which is less-conserved across species (Seki et al., 2003).

A. DNA Polymerase θ Across Species

Human *POLQ* encodes a protein of 2590 amino acids with a molecular mass of 290 kDa and was identified as the eighth human DNA polymerase after an extensive homology search against the *Drosophila melanogaster* gene encoding Pol θ, *mus308* (Sharief et al., 1999; Seki et al., 2003). Studies in *D. melanogaster* first isolated the *mus308* gene through a screen which identified genes that conferred resistance against various DNA damaging agents; *mus308* mutants were characterized as hypersensitive to the ICL agent nitrogen mustard (HN2), although not sensitive to monofunctional alkylating agents, implicating a role in ICL repair for Pol θ in *D. melanogaster* but not a role in base excision repair (BER) (Boyd et al., 1990). Identification of *D. melanogaster mus308* prompted the search for similar protein factors to each of the catalytic domains of Pol θ. The polymerase domain of Pol θ displays significant homology to human polymerase ν (*POLN*); interestingly, the C-terminal polymerase domain of *D. melanogaster*

mus308 displays greater homology to human *POLN* than to the C-terminal polymerase domain of human *POLQ* (Marini et al., 2003). Additionally, the N-terminal helicase-like domain of Pol θ displays significant homology to human *HEL308* although, dissimilar to the relationship between *mus308* and *POLN*, the helicase-like domain of *mus308* displays slightly lesser homology to human *HEL308* than to human *POLQ* (Tafel et al., 2011). A yeast *mus308* homolog has not been identified, however, an ortholog of the *D. melanogaster* gene encoding Pol θ was identified in *C. elegans, polq-1*, in *M. musculus, chaos1* (not pictured), and in the plant species *Arabidopsis, tebichi* (Muzzini et al., 2008; Inagaki et al., 2009) (Figure 6).

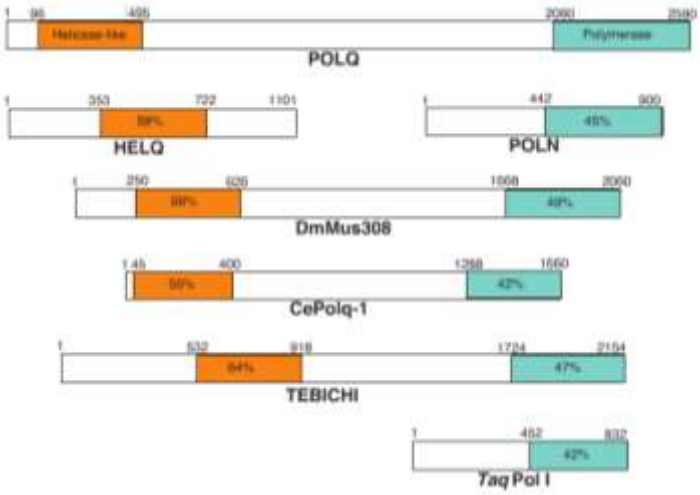


Figure 6. Relation among gene family members of Pol θ across species. Relationships between Pol θ homologs are represented based on percent similarity. The orange regions and blue regions of Pol θ represent the helicase-like domain and polymerase domain respectively (Modified from Yousefzadeh and Wood, 2013).

Functional divergence of Pol θ across species

Invertebrate Pol θ: Hypersensitive to ICL agents but not radiosensitive

D. melanogaster mus308 mutants, *C. elegans polq-1* mutants, and the plant species, *Arabidopsis, tebichi* mutants are hypersensitive to ICL agents but do not demonstrate the radiosensitivity observed in mammals lacking Pol θ (Leonhardt et al., 1992; Muzzini et al., 2008;

Inagaki et al., 2009). Pol θ has also been suggested to function during alt-EJ repair of DSBs in invertebrates. In a 2010 study in *Drosophila melanogaster* by Chan et al., *mus308* mutants demonstrated ineffective alt-EJ repair of DSBs; Pol θ was hypothesized to serve dual functions in this process, both in unwinding double-stranded DNA in search of microhomologies and in synthesis off of these annealed microhomologous sequences (Chan et al., 2010). *D. melanogaster* lacks several C-NHEJ components found in mammals, potentially preferentially predisposing them to non-C-NHEJ repair mechanisms, such as alt-EJ (Chan et al., 2010; Decottignies, 2013). As ICL repair includes a DSB intermediate, the sensitivity of Pol θ mutants to ICL agents, in combination with ineffective alt-EJ repair in Pol θ mutants in *D. melanogaster*, may implicate a Pol θ-dependent role for alt-EJ during ICL repair (Chan et al., 2010).

Studies in *C. elegans* have suggested that *polq-1*, a Pol θ ortholog, and *hel308*, a protein orthologous to the N-terminal helicase-like domain of Pol θ, function in two distinct ICL repair pathways; *polq-1* appears to function in a BRCA1-dependent ICL repair pathway while *hel308* seems to function in a FA-dependent manner (Muzzini et al., 2008). Additionally, Pol θ has been implicated to function as a backup polymerase in the absence of polβ during BER in *C. elegans* although an alternate polymerase may also perform this role (Aasagoshi et al., 2012). In each of these processes in which Pol θ is believed to function, the specific function of each of the two catalytic domains is uncertain and further research is necessary to ascertain their novel functions.

Mammalian Pol θ: Radiosensitive but not sensitive to ICL agents

Mammalian Pol θ has been implicated in maintaining chromosome stability, as a backup enzyme for BER, and in somatic hypermutation although the exact role, if any, has yet to be elucidated; however, the strongest role for Pol θ *in vivo* is believed to be in the repair of

radiation-induced damage as mammalian Pol θ mutants are extremely radiosensitive (Goff et al., 2009). This role for Pol θ during the repair of DSBs in combination with a postulated role for Pol θ during alt-EJ in *D. melanogaster* suggests that Pol θ may function in a similar HR-independent microhomology-mediated EJ process in mammals (Yu and McVey, 2010). Although human Pol θ mutants do not demonstrate sensitivity to ICL agents, the human ortholog of the polymerase domain of Pol θ, POLN, demonstrates sensitivity to ICL agents (Zietlow et al., 2009). Additionlly, human HEL308 has been shown to colocalize with both Rad51 and FANCD2, two proteins involved in FA-mediated repair of ICLs, implicating a role for HEL308 during ICL repair as well (Tafel et al., 2011).

Potential sources for functional separation

Divergence of eukaryotic Pol θ orthologs in the form of functional separation may be due, in some part, to the shortening of unique sequence inserts and nonconserved central domains of these invertebrate Pol θ members compared to mammalian family members (Yousefzadeh and Wood, 2013). Additionally, there is a higher redundancy in function of proteins in the more complex genomes of mammalian family members than invertebrate family members, which could account for variance in sensitivity to specific DNA damaging agents in different species. Human POLN and HEL308 demonstrate similar expression patterns both in tissues and cancer cell lines indicating that these proteins, with sequence and suggested functional similarity to the two catalytic domains of human Pol θ, may be working in a redundant fashion to *POLQ*, as they function at similar locations despite the lack of a physical connection (Marini et al., 2003).

Role of POLQ in human cancers

Pol θ is overexpressed in many human cancers, including breast cancer, oral squamous cell carcinoma, lung cancer, stomach cancer, and colorectal cancer, and conveys a poor prognosis in these cancer patients, which makes the characterization of the novel functions of Pol θ essential to direct the development of potential cancer therapeutics (Higgins et al., 2010; Lessa et al., 2013; Pillaire et al., 2010; Kawamura *et al.* 2004). Upregulation of Pol θ in specific cancers could additionally serve as a useful marker for early diagnosis and disease prognosis. It is unclear whether or not the upregulation of Pol θ is causative of carcinogenesis or whether cancer cells themselves upregulate the expression of Pol θ. As Pol θ displays TLS functions, it is evident that the upregulation of an error-prone translesion polymerase would be beneficial for propagation and drug resistance of a cancerous tumor as mutations drive cancer propagation.

B. A Study in *D. melanogaster*: *mus308*, one protein with two functionally distinct catalytic domains

Given the apparent functional divergence of Pol θ across species, studying the role of Pol θ in a non-mammalian model organism may give rise to novel functions masked by redundancy of protein function found in higher eukaryotes. This study investigates the role of Pol θ in *D. melanogaster* during ICL repair, specifically seeking to identify separate novel functions of the helicase-like and polymerase domains during this repair process. Our data support a model in which the catalytic function of the helicase-like domain of Pol θ is crucial for repair of HN2-induced ICLs, and that the presence, more than that of the catalytic function, of the polymerase domain is also required for proper repair of ICLs.

Methods

D. melanogaster stocks

Mus308 alleles which were not generated, but were used, in this study include $mus308^{\Delta N}$, $mus308^{\Delta C}$, $mus308^{poldead}$, and $mus308^{helded}$. The $mus308^{\Delta N}$ allele is a transgenic allele located on the third chromosome in a $mus308^{null}$ background and includes the endogenous *mus308* promoter followed by the C-terminal portion of the *mus308* gene, nucleotide positions 4029 to the end of the coding sequence. The $mus308^{\Delta C}$ allele is located at the endogenous *mus308* location and is a deletion mutant of the C-terminal region of *mus308*, from nucleotide position 4288 to 6043; there is a small insertion which results in the addition of a premature stop codon. The $mus308^{heldead}$ allele (full name $mus308^{K262A}$) is a transgenic allele which is located on the second chromosome in a $mus308^{null}$ background and has an introduced point mutation which changes a lysine residue at amino acid position 262 to an alanine in the Walker-A box of the ATPase domain of the N-terminal helicase-like domain. The $mus308^{poldead}$ (full name $mus308^{E2009A}$) allele is a transgenic allele which is located on the second chromosome in a $mus308^{null}$ background and has an introduced point mutation which alters a glutamic acid residue at amino acid position 2009 to an alanine residue in the active site of the C-terminal polymerase domain.

Fly DNA Preparation

A single fly from each sample was placed in a 0.5 mL tube and mashed with a pipette tip that contained 50 µL of squishing buffer (SB) and 1 µL of Proteinase K at concentration 10 mg/mL. SB is 10 mM Tris-Cl with a pH of 8.2, 1mM EDTA, and 25mM NaCl. The fly prep was then placed in a thermal cycler where the sample was incubated at 25-37ºC for 20-30 minutes followed by the inactivation of Proteinase K by heating to 95ºC for 1-2 minutes.

mus308^{null} deletion generation and characterization

The *mus308*^{null} allele was generated through an imprecise P-element excision screen in which virgin female flies with a large transposable P-element inserted at position 6042 in the *D. melanogaster* gene *mus308* were crossed to male flies which contained a transposase. After screening through a series of excision events of the P-element by the transposase in the F_1 generation via walk-out PCR analysis, a deletion span of 14.25 kB was detected using a forward primer in the neighboring *men* gene (*men*2016F), 5'-GGAACGTGGAGTGTGTACG-3', and a reverse primer in the *mus308* gene (*mus308*7005R), 5'-GAGAAGTACCTGTTCGATGTGC-3'.

Hatching Frequency Assays (*mus308*^{null}, *mus308*$^{\Delta C}$, *mus308*^{Cterm}, *hPOLQ*)

Hatch rate assays were performed with approximately 20-30 virgin female flies crossed to 10-20 male flies. Virgin females were crossed to corresponding males less than 24-hours post-eclosure in a cage. Fertilized eggs were collected on a grape agar plate at time points 24, 48, and 72 hours. Eggs were scored two days after removal from fly cage as either hatched or unhatched and percent hatching frequencies were determined at percent hatched divided by the total number of eggs multiplied by 100%.

Crosses: ***mus308***^{null}**:** $mus308^{null}$ x $mus308^{null}$, $mus308^{null}$ x w^{1118} (w^{1118} males)

mus308$^{\Delta C}$**:** $mus308^{\Delta C}$ x $mus308^{\Delta C}$, $mus308^{\Delta C}$ x w^{1118} (w^{1118} males)

mus308^{Cterm}: $mus308^{Cterm}$ x $mus308^{Cterm}$, $mus308^{Cterm}$ x w^{1118} (w^{1118} males)

hPOLQ: $\frac{hPOLQ}{cyo}$; $mus308^{null}$ $x\frac{hPOLQ}{cyo}$; $mus308^{null}$

Embryo Fixation Protocol

Using a paint brush, the embryos were moved from the grape agar plate to a collection basket. Embryos were washed with embryo wash solution (7% NaCl and 0.7% Triton X-100). The embryos were then dechorionated by washing them in 50% bleach for 1.5-2 minutes. The

embryos were then washed again with embryo wash for 5 minutes and transferred to an eppendorf tube with 500µL of PBS. After centrifugation, PBS was pipetted off. The embryos were then fixed in a 1:1 mixture of heptane and 37% formaldehyde (500µL of each) for 5 minutes. The 37% formaldehyde was pipetted from the bottom leaving heptane and the embryos in the eppendorf tube. After adding 500µL of MeOH, the eppendorf tube was shaken vigorously for two minutes to devittelinize the embryos. The devittelinized embryos sank to the bottom and the remaining solution was pipetted off. The embryos were washed with MeOH twice and followed by once with PBS.

DAPI Staining

The embryos were stained with a 1 µg/mL DAPI for 5 minutes. The embryos were then washed twice with PBS for 2 minutes each time. Following the washes, the embryos were suspended in 100 µL of 50% glycerol and mounted on to slide with FluoroMount for visualization using fluorescence microscopy.

HN2 Sensitivity Assays

Sensitivity to the ICL agent HN2 for *mus308* mutant flies was determined by crossing balanced, heterozygous parents to each other. Three days after the initial parental cross was set up, the parental generation was flipped into a corresponding set of control vials. The experimental vials were treated with concentrations of HN2 which ranged from 0.001%-0.005% HN2. Control vials were treated with distilled water. Relative present survival was calculated as the percent of homozygotes in the resulting F_1 generation of the treated vials divided by the percent of resulting homozygotes from the control vials.

Crosses: ***mus308^{null}***: $\frac{\text{mus308}^{\text{null}}}{\text{TM6B}}$ x $\frac{\text{mus308}^{\text{null}}}{\text{TM6B}}$

mus308$^{\Delta C}$: $\frac{mus308^{\Delta C}}{TM3}$ x $\frac{mus308^{\Delta C}}{TM3}$

mus308^{Cterm}: $\frac{\text{mus308}^{\text{Cterm}}}{\text{TM6B}}$ x $\frac{\text{mus308}^{\text{Cterm}}}{\text{TM6B}}$

RNA isolation and RT-PCR

RNA isolation from 30 *mus308^{Cterm}* male flies was treated using the Zymo Research Corporation RNA isolation protocol. Reverse transcription of the isolated RNA was performed using 2 µg of RNA in a solution of oligo(DT) and nuclease free water and incubated at 70-85°C for 3 minutes. Tubes were removed to ice. RT buffer, dNTP mix, RNase inhibitor, and MMLV-RT were added to the reaction tubes and 42-44°C for one hour. Afterwards, the reaction mixture was incubated at 92°C for 10 minutes. Subsequent PCR was performed using primers internal to the coding sequence of the *mus308^{Cterm}* to test for the presence of RNA transcript.

Results

I. Characterization of the *mus308*^{null} allele

Pol θ is unique among proteins in that it contains a helicase-like domain in addition to its A-family polymerase domain. It is believed that these two catalytic domains demonstrate functional separation during ICL repair. In order to identify the novel functions of Pol θ during ICL repair, we initially sought out to create and characterize a null Pol θ mutant. Previous studies in our lab have used *mus308* alleles which induce point mutations in the gene which alter Pol θ expression, but do not completely disrupt expression. The *mus308*^{2003} allele has an introduced nonsense mutation upstream of the polymerase domain; RT-PCR analysis has detected mRNA transcript rendering the helicase-like domain potentially functional (Figure 1). The *mus308*^{D5} allele has an introduced missense mutation which alters a conserved proline in the N-terminal portion of the central linker region to a lysine, the *mus308*^{3294} allele alters a glycine residue in the helicase-like domain to a serine, while the *mus308*^{D2} allele does not have a mutation in the coding region of the gene but, rather an essential regulatory region, and is classified as a genetic hypomorph (Figure 7).

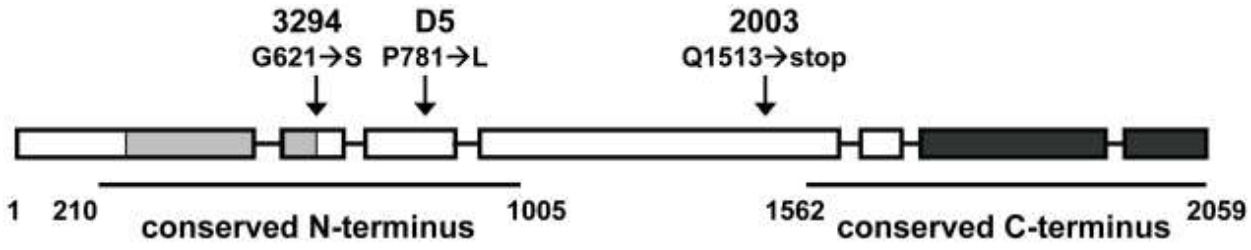


Figure 7. Mutations in *mus308* gene prior to this study. The *mus308*^{D2} allele (not pictured) is considered a genetic hypomorph and is believed to have a mutation in an essential regulatory region of the *mus308* gene. The *mus308*^{3294} gene has a guanine to serine mutation in the helicase-like domain. The *mus308*^{D5} allele has a proline to lysine mutation in the central linker region. The *mus308*^{2003} allele introduces a premature stop codon upstream of the polymerase domain (Modified from Chan et al., 2010).

A. Mapping the *mus308*^{null} deletion allele

The deletion of *mus308* in *Drosophila melanogaster* was generated through an imprecise P-element excision screen of a P-element inserted at nucleotide position 6042 in the *mus308* gene (Renedo, unpublished data). Excision of the P-element was achieved by crossing virgin female flies containing the transposable P-element insertion in the *mus308* gene with male flies which contained a transposase. After genetic fly crosses performed by Renedo were complete, subsequent screening was performed by the author of this study. After screening through a series of excision events of the P-element by the transposase in the F_1 generation via walk-out PCR analysis, the deletion span was characterized as spanning from the active site of the polymerase domain (nucleotide position 6043 in *mus308* (8531.978k)), through the promoter region of *mus308*, to nucleotide position 2434 in *men* (8545.730k) (Figure 8). There is a 30 base pair insertion between the C-terminal portion of the *men* gene and the N-terminal portion of *mus308.* It is important to note that the *men* gene has been classified as a nonessential gene as flies deficient in *men* are both viable and fertile, therefore, the partial deletion of the *men* gene should not affect our *mus308*^{null} allele (Voelker et al., 1981).

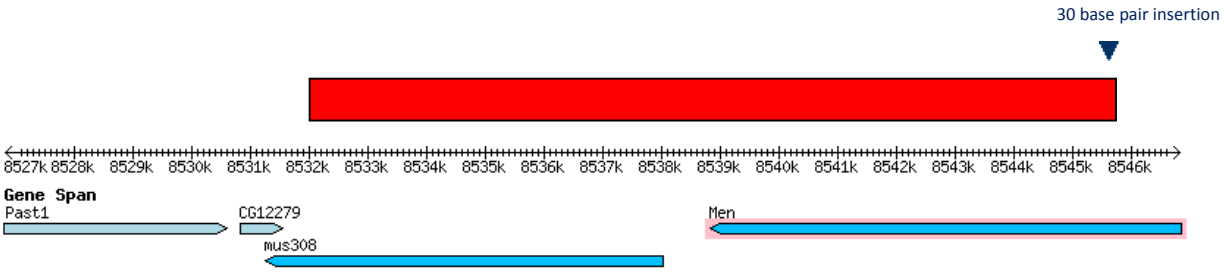


Figure 8. Deletion span for the mus308^{null} allele. The deletion spans from position 6043 in mus308 (8531.978k) to position 2434 in men (8545.730k). There is a 30 base pair insertion between the N-terminal portion of the Men gene and the C-terminal portion of mus308.

B. *mus308*^{null} flies demonstrate a more severe hatching frequency defect than previously characterized *mus308* alleles

It has been observed that *mus308*^{D2} mutants display developmental abnormalities including erect wings, reduction in flight capacity, low fertility, abnormal chromosome segregation, and fragility of the protective membrane layers surrounding embryos (Leonhardt et al., 1992). It is possible that many of these developmental abnormalities stem from misregulation of vital processes during early development. The most direct method we utilize for detection of abnormalities during embryonic development is hatching frequency, the percentage of eggs laid during a given time period which hatch. A low hatching frequency is indicative of failure to perform essential developmental processes, including DNA repair processes and replication.

Mus308^{null} mutants have a severe defect in hatching frequency, with only 3.56% of eggs laid hatching (Figure 9). The extreme hatching frequency defect observed in flies deficient in Pol θ indicates that Pol θ function is essential during embryonic development. This defect in hatching frequency can be rescued, with a hatching frequency of 71.1%, when virgin female *mus308*^{null} flies are crossed to wild type male flies. During early embryogenesis, the first two hours of development, the embryo only has access to maternally loaded DNA. After the first two hours of development, zygotic transcription begins, resulting in access to a functional copy of Pol θ. A rescue of hatching frequency by crossing *mus308*^{null} virgin females to wild type males indicates that Pol θ does not primarily function during the first two hours of development. However, it is important to note that the rescue of hatching frequency does not reach hatching frequency levels observed in wild type flies, 88.6%. This suggests that Pol θ may play some less significant role during early embryogenesis although it functions most significantly during the later stages of embryogenesis.

This theory is supported by the complete rescue of hatching frequency when *mus308*^{D2} and *mus308*^{2003} virgin females were crossed to wild type males. The *mus308*^{D2} allele displays very low levels of Pol θ transcript, but, given the proposed small role for Pol θ during early embryogenesis, this level of transcript could be enough to allow developing embryos to survive through these early developmental stages. As the *mus308*^{2003} allele has an introduced premature stop codon in the central linker region just upstream of the C-terminal polymerase domain, it is possible that the transcript produced is not translated into functional protein. However, if the truncated mRNA transcript for the *mus308*^{2003} allele is translated into a functional N-terminal helicase-like domain of Pol θ, this would suggest that the polymerase domain of Pol θ does not perform the potential function during early embryogenesis as complete rescue is observed upon gain of Pol θ polymerase function at the zygotic transition.

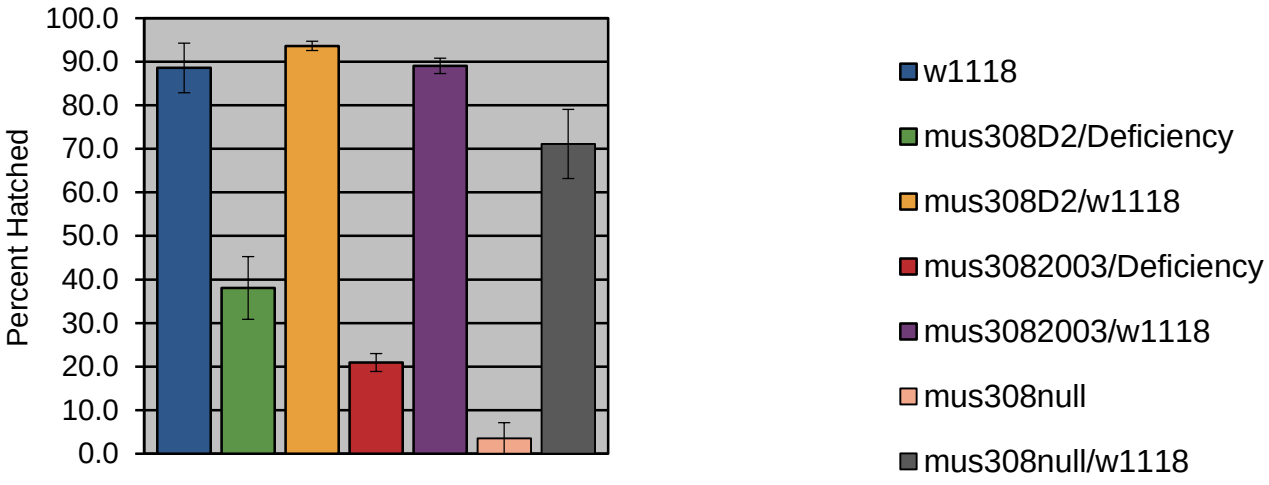


Figure 9. Hatching frequencies of *mus308* mutant alleles. Wild type hatching frequency is at 88.6%. *Mus308*^{D2} flies have a hatching frequency of 38.1%. This hatching frequency is rescued to 93.6% when virgin female *mus308*^{D2} flies are crossed to wild type males. *Mus308*^{2003} flies have a hatching frequency of 21.0% while crossing virgin female *mus308*^{2003} flies to wild type males rescues this hatching frequency to 89.0%. *Mus308*^{null} flies have a hatching frequency of 3.56% and this hatching frequency is partially rescued when *mus308*^{null} virgin female flies are crossed to wild type males to 71.1% indicating that the most significant role for Pol θ during embryonic development is post-early embryogenesis. First genotype listed is female; second genotype listed is male.

C. In the absence of Pol θ, *Drosophila* embryos demonstrate developmental arrest at the single nuclear stage

A phenotype often seen in conjunction with an extremely low hatching frequency is abnormal development during embryogenesis. During early embryogenesis, the developing embryo undergoes fourteen, synchronous, syncytial nuclear divisions without undergoing cellular division. If Pol θ functions during early embryonic stages of development as prior studies have implicated, we would expect *mus308*^{null} embryos to display defects associated with nuclear segregation (Appendix I). To investigate whether or not *mus308*^{null} embryos experience developmental defects during early embryogenesis, we performed nuclear staining of *mus308*^{null} embryos during the first two hours of development.

In a previous study performed by the author, severe nuclear abnormalities were observed in embryos which expressed the *mus308*$^{\Delta C}$ allele, a Pol θ mutant allele which expresses functional helicase-like Pol θ protein but not a functional polymerase domain (Appendix I). In these embryos, nuclear abnormalities included nuclear clumping, anaphase bridges, and asynchronous nuclear divisions; however embryonic development progressed through embryogenesis only approximately 50% of the time (Appendix I). In this study, all *mus308*^{null} embryos were arrested at a single nuclear stage even when given up to eight hours to develop (Figure 10). This early embryonic phenotype of *mus308*^{null} embryos suggests a role for Pol θ during the first two hours of development although it does not completely reflect upon whether or not the observed defect is a result of a failure to initiate early embryogenesis in the absence of Pol θ or an accumulation of defects prior to embryogenesis in the absence of Pol θ which retards initiation of embryogenesis. Hatching frequency data would suggest that Pol θ does not primarily function during the first two hours of development, which, in combination with the apparent failure to initiate nuclear division during embryogenesis, would implicate a prominent

function of Pol θ prior to early embryogenesis. Pol θ has been proposed to function during oogenesis which could possibly account for the lack of embryonic development observed in *mus308*^{null} embryos (Beagan, unpublished data).

The developmental arrest phenotype also raises the question of how embryos with *mus308*^{null} mothers reinitiate embryogenesis with wild type fathers as paternal genes should not function until after the first fourteen nuclear divisions. It is possible that *mus308*^{null} embryos arrest at the single nuclear stage until another protein of redundant function can help the developing embryo reinitiate nuclear division, although the mechanism of hatching frequency rescue with wild type fathers remains unclear. It is clear, however, that Pol θ serves an essential function during early development through both hatching frequency analysis and nuclear staining of *mus308*^{null} embryos.

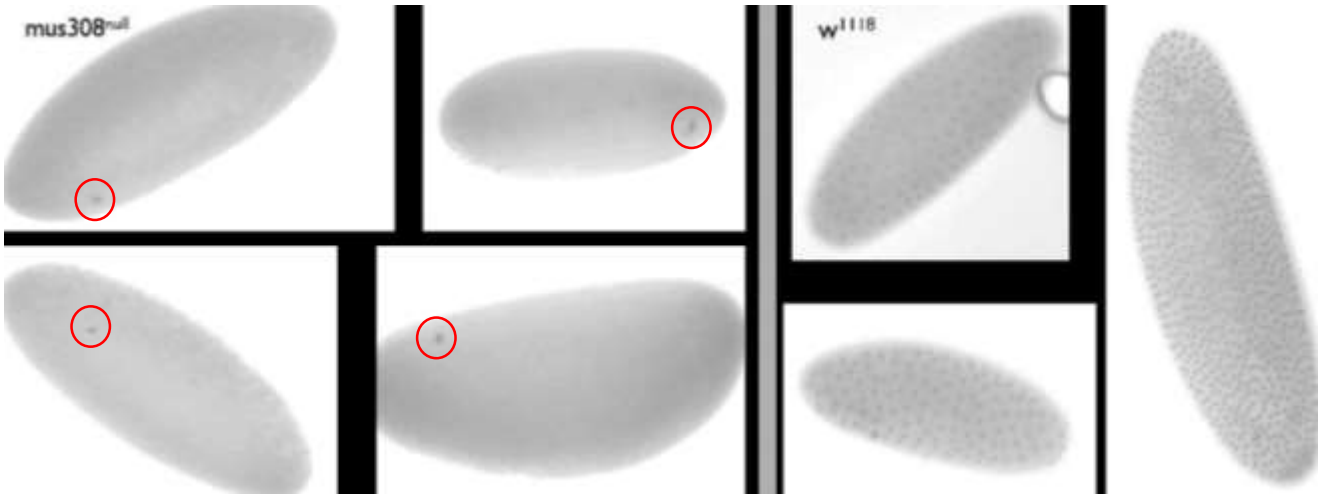


Figure 10. Nuclear arrest in *mus308*^{null} embryos suggests a role for Pol θ prior to embryogenesis. Embryos were stained during the first two hours of development using the nuclear stain, DAPI. *Mus308*^{null} embryos demonstrate developmental arrest at the single nuclear stage while wild type embryos show proper nuclear division and equal nuclear segregation throughout the embryo.

D. *Mus308*^{null} flies display extreme sensitivity to the ICL agent HN2

Prior studies with the *mus308*^{D2} allele have shown that Pol θ mutants demonstrate an extreme sensitivity to the ICL agent HN2 although the exact role of Pol θ during this repair process has yet to be elucidated. The *mus308*^{D2} allele, a Pol θ genetic hypomorph, demonstrates a 15% survival when treated with 0.005% HN2, a moderate dose; we were interested in determining whether or not our established *mus308*^{null} allele would show sensitivity to ICL agents which mimicked that of the *mus308*^{D2} allele or a sensitivity of greater severity. Additionally, determining the sensitivity of our *mus308*^{null} mutants was essential in order to later investigate the sensitivity of Pol θ single-domain mutants in hopes of exacting the novel functions of each catalytic domain during ICL repair.

The *mus308*^{nulll} allele displays extreme sensitivity to relatively low doses of the ICL agent, HN2. When treated with 0.003% HN2, *mus308*^{null} flies demonstrate a 43.6% survival rate relative to a heterozygous control; this survival rate decreases to 15.4% when *mus308*^{null} flies were treated with the 0.005% HN2 (Figure 11). The survival curve observed in this study is similar to that observed with the *mus308*^{D2} allele, validating the essential role for Pol θ during repair of an induced ICL.

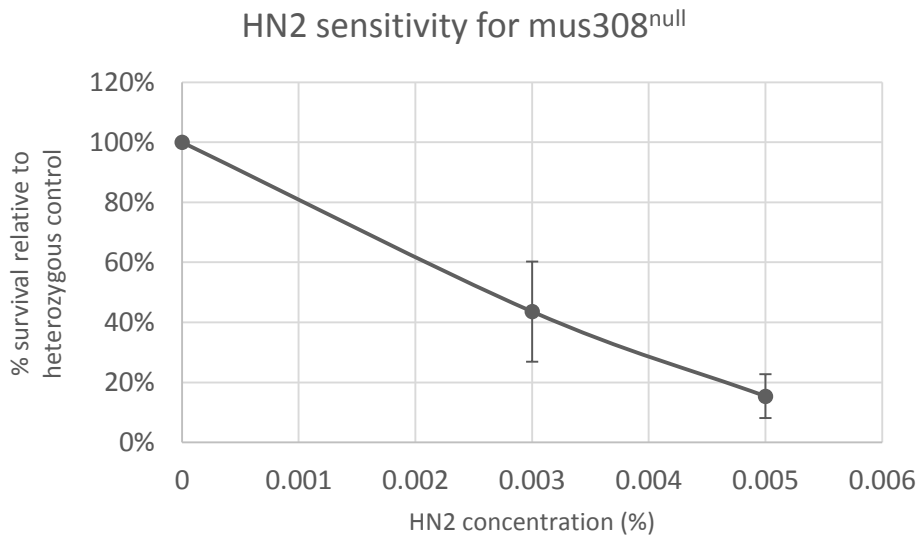


Figure 11. Flies deficient in Pol θ display an extreme inability to repair HN2 induced ICLs. *Mus308*^{null} flies exhibit a 43.6% survival rate relative to a heterozygous control when treated with 0.003% HN2 and this survival rate drops to 15.4% when treated with 0.005% HN2. The survival curve of *mus308*^{null} flies mimics that of *mus308*^{D2} flies. Three trials were performed at each data point.

II. Deciphering the novel functions of Pol θ during ICL repair

Where exactly in ICL repair Pol θ functions and the relative importance of each of the two catalytic domains during this process remains unclear. The N-terminal helicase-like domain of Pol θ is believed to act as a helicase. Despite the fact that unwinding capabilities characteristic of a helicase have yet to be displayed in Pol θ, ATPase activity and helicase sequence homology has been demonstrated across Pol θ orthologs (McVey, 2010). Although Pol θ has been definitively shown to function during ICL repair, this study sought to identify the individual roles of the helicase-like and polymerase domains of Pol θ. To address whether one catalytic function of Pol θ serves a more important function than the other, we generated ATPase-dead, *mus308*^{K262A} (referred to as *mus308*^{heldead} in this study), and polymerase-dead, *mus308*^{E2009A} (referred to as *mus308*^{poldead} in this study), Pol θ mutants and observed their relative survivals when treated with HN2. Additionally, in order to investigate the possibility of non-catalytic functions of both the N-terminus and C-terminus of Pol θ during ICL repair, we created truncated Pol θ deletion mutants of either the entire conserved N-terminus and part of the central linker region, *mus308*$^{\Delta N}$, or part of the C-terminus of Pol θ, *mus308*$^{\Delta C}$. Variation in survival between the domain-dead and truncated Pol θ mutants would suggest a possible role in recruitment of other proteins for that particular domain. (Figure 12).

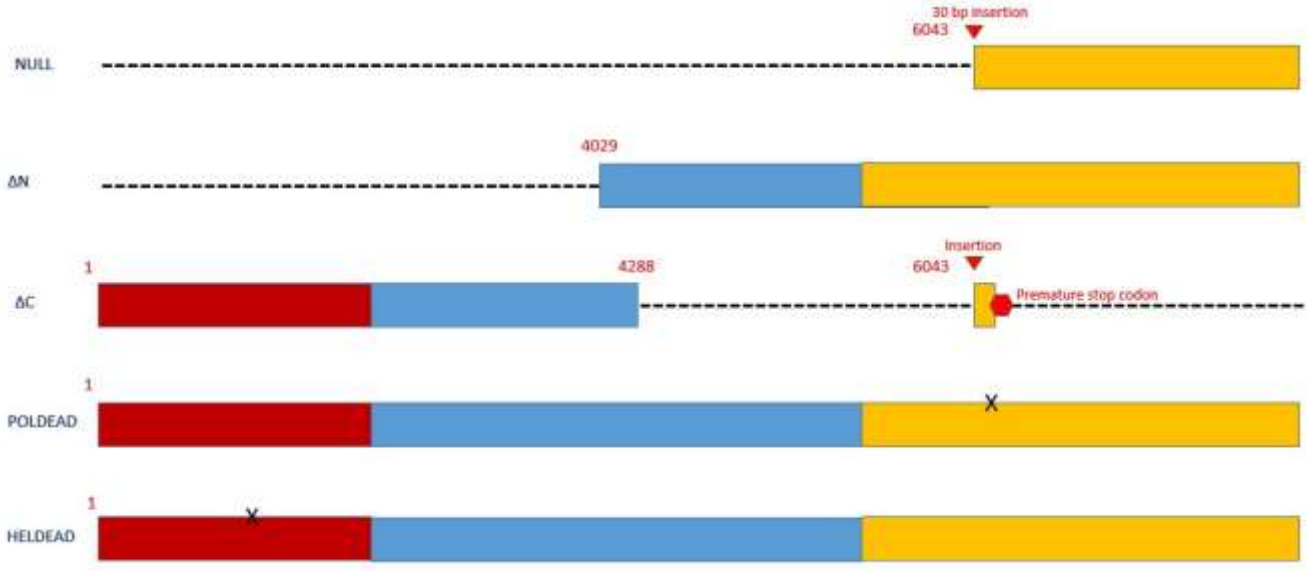


Figure 12. *Mus308* mutant alleles used to determine novel functions of Pol θ during ICL repair. The *mus308*^{null} allele is missing the *mus308* promoter and does not have a functional helicase-like domain or polymerase domain. The *mus308*$^{\Delta N}$ allele is a transgenic allele under control of the endogenous *mus308* promoter located on the third chromosome in a *mus308*^{null} background and is missing the N-terminal portion of *mus308* until nucleotide position 4029 which is located in the central, non-conserved linker region. The *mus308*$^{\Delta C}$ allele is a deletion mutant between nucleotide positions 4288 and 6043, followed by a small insertion which induces a premature stop codon and only contains a functional helicase-like domain, not a functional polymerase domain. The *mus308*^{poldead} has an introduced point mutation which changes a conserved glutamic acid residue at amino acid position 2009 to an alanine residue in the active site of the C-terminal polymerase domain, rendering the polymerase domain inactive. The *mus308*^{heldead} allele has an introduced point mutation which changes a lysine residue at amino acid position 262 to an alanine in the Walker-A box of the ATPase domain of the N-terminal helicase-like domain, rendering the helicase-like domain inactive.

A. The catalytic function of the helicase-like domain of Pol θ is essential during ICL repair

Although it is clear that Pol θ plays an essential role during ICL repair, it is unclear how each of the two catalytic domains of Pol θ contribute to this repair process. In order to address this question, we created domain-dead Pol θ transgenic mutants which express the entire *mus308* gene under the control of the endogenous *mus308* promoter on the second chromosome and the *mus308*^{null} allele on the third chromosome. These transgenic constructs have a single, induced

point mutation which alters a critical amino acid for catalytic function. The *mus308*^{heldead} allele has an introduced point mutation in the Walker-A Box of the ATPase domain which changes an essential lysine residue to an alanine residue, thus rendering the ATPase function of the N-terminal helicase-like domain inactive. The *mus308*^{poldead} allele has an introduced point mutation which alters a conserved glutamic acid residue to an alanine residue in the active site of the A-family polymerase domain of Pol θ, rendering the translesion polymerase activity of Pol θ inactive. These mutants, therefore, lose function in one catalytic domain but maintain all other structural attributes. We do not expect misfolding issues in our catalytically-dead mutants based on previous findings from other groups which mutated the same series of amino acids (Yoon et al., 2014). Identifying whether or not one catalytic domain functions with greater importance than the other could contribute towards informing a proposed mechanism of Pol θ function during ICL repair.

Sensitivity studies with the ICL agent HN2 involving our transgenic, domain-dead Pol θ mutants demonstrate that the helicase function of Pol θ is more important during the repair of a HN2-induced ICL than the polymerase domain of Pol θ (Figure 13). At a treatment dose of 0.003% HN2, *mus308*^{heldead} flies show 42% survival relative to a heterozygous control while *mus308*^{poldead} flies show 74.9% survival relative to a heterozygous control; this sensitivity displayed in the domain-dead Pol θ mutants is not due to the location of the transgenic constructs on the second chromosome as the control transgene, *mus308*^{ctrl}, displays a survival rate comparable to wild type (Figure 13). The survival rate of both *mus308*^{heldead} and *mus308*^{poldead} declines when treated with 0.005% HN2, although the *mus308*^{poldead} flies show only moderate sensitivity at this dose with 56.9% survival relative to a heterozygous control. On the contrary, the survival rate observed in the *mus308*^{heldead} flies drops to levels observed in the *mus308*^{null}

flies; *mus308*^{heldead} flies display a survival rate of 13.8% while *mus308*^{null} flies display a survival rate of 15.4% at 0.005% HN2.

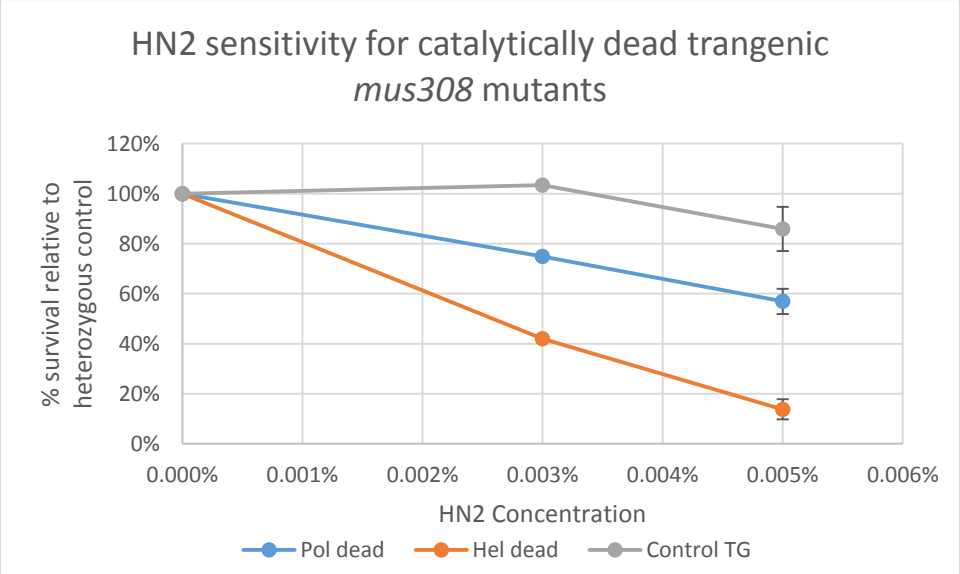


Figure 13. The helicase-like domain of Pol θ plays a more significant role during ICL repair than the polymerase domain. At 0.003% HN2, *mus308*^{ctrl} flies show a 103.4% relative survival rate, *mus308*^{heldead} flies show a 42.00% relative survival rate, and *mus308*^{poldead} show a 74.9% relative survival rate. At 0.005% HN2, *mus308*^{ctrl} flies show a 85.9% relative survival rate, *mus308*^{heldead} flies show a 13.8% relative survival rate, and *mus308*^{poldead} show a 56.9% relative survival rate. Two trials were performed for each transgene at 0.003% HN2 while three trials were performed for each transgene at 0.005% HN2. ($P<0.001$ for 0.003% and 0.005% HN2 for *mus308*^{heldead} and *mus308*^{poldead})

B. The C-terminus of Pol θ demonstrates an essential function in ICL repair outside of its catalytic function

In order to further investigate novel functions of the N- and C-termini of Pol θ outside of their catalytic functions, we sought to generate truncated versions of Pol θ. The *mus308*$^{\Delta C}$ allele is a deletion mutant of the C-terminal portion of the central linker region through nucleotide position 6043 in the polymerase domain. Additionally, a short insertion at nucleotide position 6043 results in a premature stop codon which results in loss of most of the C-terminus of Pol θ. The *mus308*$^{\Delta N}$ allele is a transgenic mutant located on the third chromosome under control of the

endogenous *mus308* promoter which removes the N-terminus of the *mus308* gene through nucleotide position 4029; this transgene is in a $mus308^{null}$ background.

The $mus308^{\Delta C}$ allele exhibits the most severe sensitivity to the ICL agent HN2 of all Pol θ mutants observed in both previous studies and this study with 0% relative survival to a heterozygous control at both 0.003% HN2 and 0.005% HN2 (Figure 14). When compared to the HN2 sensitivity observed in the $mus308^{poldead}$ mutant, 74.9% survival at 0.003% HN2 and 56.9% survival at 0.005% HN2, the HN2 sensitivity observed in the $mus308^{\Delta C}$ mutant implicates an essential non-catalytic function for the C-terminus of Pol θ (Figure 14). Where the polymerase domain of Pol θ does not have an essential catalytic function for ICL repair, we propose that the primary function of the C-terminus is in the recruitment of essential proteins to the ICL repair junction as the physical presence of the C-terminus of Pol θ results in a significant increase in resistance to ICL agents. The $mus308^{\Delta N}$ allele did not display a significant change in sensitivity to the ICL agent HN2 at 0.003% HN2 (11.2% survival) compared to the $mus308^{heldead}$ mutant (42% survival) and this trend continued at 0.005% HN2 which suggests that the primary function of the N-terminus during ICL repair is its helicase activity and does not play a role in protein recruitment.

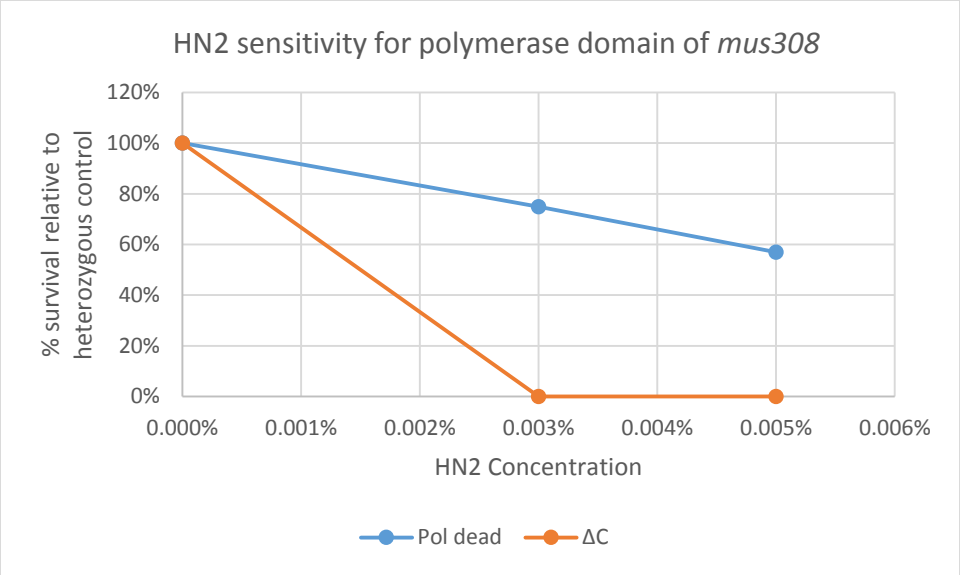


Figure 14. The C-terminus of Pol θ has an essential function during ICL repair outside its catalytic function. The *mus308*$^{\Delta C}$ allele, which is a truncated version of Pol θ, demonstrates extreme sensitivity to HN2 with 0% relative survival at 0.003% HN2 and 0.005% HN2. The *mus308*^{poldead} allele demonstrates moderate sensitivity to HN2 with 74.9% relative survival at 0.003% HN2 and 56.9% relative survival at 0.005% HN2. This implicates a role for the C-terminus of Pol θ in recruitment of essential proteins to the ICL repair junction. Three trials were performed for all data points except *mus308*^{poldead} in which two trials were performed. ($P<0.001$ for 0.003% and 0.005% HN2).

C. **There is not an absolute requirement that the two catalytic domains of Pol θ be a part of the same polypeptide**

Pol θ is unique among polymerases in that its translesion polymerase activity is coupled with ATPase activity and helicase motif sequence homology. Evolutionarily, these two catalytic domains of Pol θ have remained coupled across species which raises the question of whether these two domains must be physically connected for proper function during ICL repair. By expressing the *mus308*$^{\Delta C}$ and *mus308*$^{\Delta N}$ alleles in trans and observing subsequent HN2 sensitivity, conclusions regarding whether or not connectivity of the helicase-like domain and the polymerase domain is required for proper function during ICL repair can be made.

When the *mus308*$^{\Delta C}$ and *mus308*$^{\Delta N}$ alleles are expressed in trans, there is a significant increase in relative survival compared to the HN2 sensitivities of both the *mus308*$^{\Delta N}$ and

mus308$^{\Delta C}$ alleles (Figure 15). At 0.003% HN2, *mus308*$^{\Delta N}$/*mus308*$^{\Delta C}$ flies show 60.3% survival relative to a heterozygous control while *mus308*^{Cterm} flies show 11.2% relative survival and the *mus308*$^{\Delta C}$ flies show 0% relative survival (Figure 15). As the dose of HN2 increases to 0.005%, the significant increase in survival is maintained with *mus308*$^{\Delta N}$/*mus308*$^{\Delta C}$ flies displaying 36.2% relative survival compared to 9.3% relative survival in *mus308*$^{\Delta N}$ flies and 0% relative survival in *mus308*$^{\Delta C}$ flies (Figure 15). The increased resistance to the ICL agent HN2 in *mus308*$^{\Delta N}$/*mus308*$^{\Delta C}$ flies implicates that there is not an absolute requirement for the helicase-like domain and the polymerase domains of Pol θ to be a part of the same polypeptide in order to function during ICL repair, although connectivity of the two domains does improve Pol θ function during this repair process.

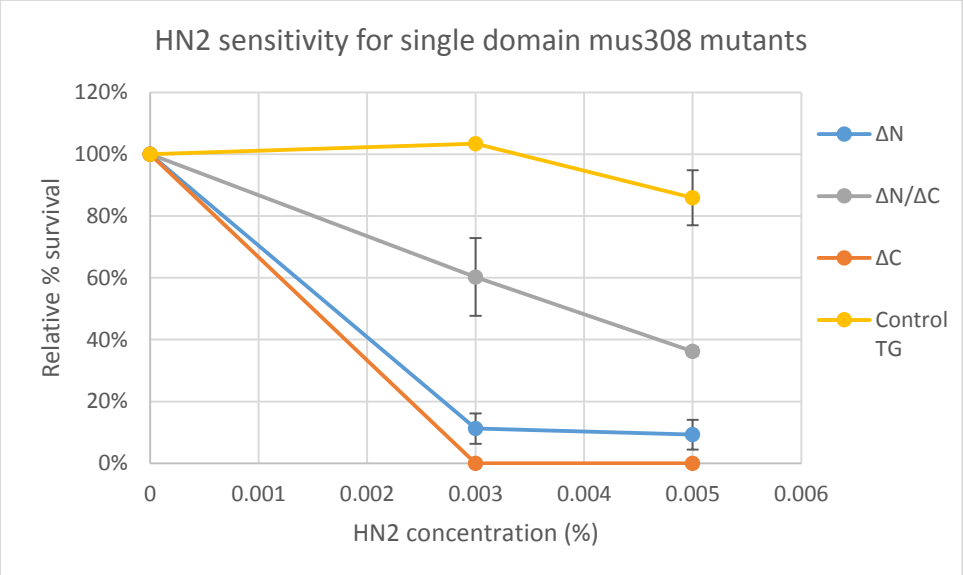


Figure 15. Expressing the *mus308$^{\Delta N}$* and *mus308$^{\Delta C}$* alleles in trans increases resistance to the ICL agent, HN2, compared to expressing each Pol θ mutant separately. At 0.003% HN2, *mus308$^{\Delta N}$* /*mus308$^{\Delta C}$* flies show 60.3% survival relative to a heterozygous control while *mus308$^{\Delta N}$* flies show 11.2% relative survival and the *mus308$^{\Delta C}$* flies show 0% relative survival. At 0.005% HN2, *mus308$^{\Delta N}$* /*mus308$^{\Delta C}$* flies show 36.2% relative survival compared to 9.3% relative survival in *mus308$^{\Delta N}$* flies and 0% relative survival in *mus308$^{\Delta C}$* flies. Three trials were performed for all concentrations of HN2 for *mus308$^{\Delta N}$* and *mus308$^{\Delta C}$* alleles while two trials were performed for HN2 sensitivity assays with the mutants in trans. ($P<0.001$ for 0.003% and 0.005% HN2 between *mus308$^{\Delta N}$* /*mus308$^{\Delta C}$* and *mus308$^{\Delta C}$*; $P<0.01$ for 0.003% and 0.005% HN2 between *mus308$^{\Delta N}$* /*mus308$^{\Delta C}$* and *mus308$^{\Delta N}$*).

Discussion

Characterization of a *mus308*^{null} allele

The *mus308*^{null} allele generated in this study was generated through an imprecise P-element excision screen and resulted in the deletion of the *mus308* promoter, the entire N-terminal helicase-like domain, the central linker region, and part of the C-terminal polymerase domain; additionally, a portion of the neighboring *men* gene, which encodes malic enzyme, was deleted, but this does not affect our *mus308*^{null} allele (Voelker et al., 1981). The *mus308*^{null} allele demonstrates an extremely severe hatching frequency defect with only 3.56% of eggs laid hatching. This hatching frequency defect was rescued when *mus308*^{null} females were crossed to wild type males as active Pol θ was accessible to the developing embryo at the zygotic transition. The rescue in hatching frequency observed when *mus308*^{null} females were crossed to wild type males suggests that Pol θ primarily acts during later stages of embryogenesis, although we also propose a less significant function for Pol θ during early embryogenesis as the observed hatching frequency rescue did not reach the hatching frequency observed in wild type flies. Determining when in development Pol θ acts could provide insight as to whether or not the error-prone DNA repair processes in which Pol θ is believed to act function during early embryonic development. Additionally, a role for Pol θ outside of DNA repair, potentially during replication, may be considered, since the primary function of Pol θ during embryogenesis does not seem to be during early developmental stages.

Preliminary data suggests that Pol θ may function during the multiple rounds of gene amplification during oogenesis. It is possible that, in the absence of Pol θ, genes essential for development during embryogenesis are not properly produced, thus resulting in arrest at the single nuclear stage of embryogenesis. However, this theory fails to answer the question of why

the hatching frequency of *mus308*^{null} flies is rescued when *mus308*^{null} virgin females were crossed to wild type males.

To further investigate the role of Pol θ during embryogenesis and attempt to address the observed rescue in hatching frequency, staining for γH2aV, a signal of DSB formation, in *mus308*^{null} embryos should be performed during the later stages of embryogenesis. If *mus308*^{null} embryos exhibit a greater frequency of DSBs during the later stages of embryogenesis than wild type embryos, it would validate that the primary function of Pol θ during embryogenesis is in the later stages and that this role is repair oriented.

Identifying the domain-specific catalytic functions of Pol θ during ICL repair

This study sought out to identify the novel functions of the helicase-like and polymerase domains of Pol θ during HN2-induced ICL repair in *D. melanogaster*; however, the sensitivities of Pol θ to ICL agents is not specific to the lesions induced using HN2 (Appendix II). Sensitivity studies with the ICL agent HN2 identified that the catalytic activity of the helicase-like domain of Pol θ is essential during ICL repair and that the catalytic activity of the polymerase domain of Pol θ plays a role during ICL repair, although this role appears to be non-essential. Because the *mus308*^{heldead} and *mus308*^{null} alleles are equally sensitive at both 0.003% and 0.005% HN2 and the catalytic activity of the helicase-like domain is more important during ICL repair than the polymerase domain, we suggest that the helicase-like domain of Pol θ acts upstream of the polymerase domain during ICL repair. In the absence of helicase activity, the ICL repair machinery stalls upstream of downstream repair steps where the polymerase domain functions; thus, the *mus308*^{heldead} flies act phenotypically identical to *mus308*^{null} flies in ICL sensitivity studies because the helicase domain is inactive and the polymerase domain does not

have an opportunity to perform its function during ICL repair due to upstream stalling of ICL repair machinery.

The *mus308*^{poldead} flies demonstrate low to moderate sensitivity to ICL agents since the polymerase domain of Pol θ functions downstream of the essential helicase-like domain. In *mus308*^{poldead} flies, it is possible that other translesion polymerases can perform a redundant function during ICL repair to Pol θ, rendering the polymerase activity of Pol θ less significant during ICL repair than the helicase function. We believe that the helicase-like domain functions to unhook the ICL after XPF-ERCC1 creates the necessary nicks for unhooking to occur. Following unhooking of the HN2-induced ICL, we propose that the polymerase domain of Pol θ performs error-prone translesion synthesis across the helicase-like domain-dependent unhooked ICL.

A role for the C-terminus of Pol θ in interacting with FA proteins during ICL repair

The differential sensitivities observed between the *mus308*$^{\Delta C}$ and *mus308*^{poldead} mutants can be explained by an essential non-catalytic function of the C-terminal domain of Pol θ. It has been demonstrated that the polymerase domain of Pol θ displays greater percent identity and similarity to human POLN than to the C-terminal polymerase domain of human Pol θ (Marini et al., 2003). The amino acid sequence alignment between human POLN and *D. melanogaster* Pol θ extends upstream of the conserved polymerase motif through amino acid position 1563, the same amino acid position to which conservation between human Pol θ and *Drosophila* Pol θ extends through. POLN is required for ICL repair in mammals and cells deficient in POLN demonstrates phenotypes similar to that observed in FA cells (Moldovan et al., 2010). Additionally, POLN has been shown to interact during S-phase, both physically and genetically, with FA proteins including activated FANCD2, FANCI, and PCNA (Moldovan et al., 2010). Of

the fifteen FA genes, *Drosophila* only have homologs of FANCD2, FANCL, FANCM, FANCJ, and FANCI. It is possible that upon FANCM-mediated ICL detection and subsequent activation of FANCL-dependent FANCD2 monoubiquitylation, activated FANCD2 recruits Pol θ to the specific ICL lesion via a direct interaction with the C-terminal polymerase domain of Pol θ. Once Pol θ has been recruited to the ICL site by activated FANCD2, unhooking of the nicked lesion can be performed through the helicase-like domain of Pol θ. Following unhooking of the ICL, the translesion polymerase activity of Pol θ can synthesize across the lesion in an error-prone fashion. Gap-filling translesion synthesis across from the unhooked ICL allows for excision of the ICL and template strand gap-filling by a replicative polymerase. The resulting DSB can then be repaired through HR mechanisms.

The two catalytic domains of Pol θ can act independently of each other

We were surprised to find that there is not an absolute requirement that the two catalytic domains of Pol θ be a part of the same polypeptide during HN2-induced ICL repair given the widespread conservation across species of the connectivity of the helicase-like and polymerase domains of Pol θ. We believe that the partial rescue in survival observed in *mus308*$^{\Delta N}$/*mus308*$^{\Delta C}$ flies when treated with HN2 compared to the respective single-domain mutants can be explained by the ability of the helicase-like and polymerase domain catalytic functions to act independently, although the efficiency of these catalytic functions is improved upon tethering together of the two domains. The independent nature of these two domains is reflected in higher eukaryotes through the presence of *HEL308* and *POLN*, two genes homologous to that of the helicase-like and polymerase domains of *POLQ*. One hypothesis for this apparent divergence would be that *HEL308* and *POLN* serve redundant functions to *POLQ* and in situations where Pol θ is not expressed for cannot function properly, HEL308 and POLN perform the necessary

redundant functions, but in an independent manner of each other. Pol θ remains conserved in these higher eukaryotes because, as identified in our HN2 studies, the two catalytic functions are more efficient when tethered together. To establish whether or not the presence of redundancy in higher eukaryotes is abrogating potential roles of mammalian Pol θ in damage repair pathways, mutagen sensitivity studies should be performed with POLN and/or HEL308/Pol θ double mutants.

A potential role for Pol θ traversal of ICLs

Given that, recently, many proteins in mammals have been identified as capable of bypassing various types of lesions and continuing synthesis in lieu of stalling replication upon encountering a lesion, we believe that it is possible that Pol θ could potentially traverse ICLs instead of acting to repair them. Wang et al. recently demonstrated that FANCM/MHF coordinates the traverse of ICLs to avoid replication stalling at this specific lesion while Mouron et al. demonstrated that PrimPol, a new human primase and DNA translesion polymerase, is capable of bypassing UV-induced photolesions, repriming downstream of the lesion and subsequently reinitiating synthesis off of the template. The incredibly harmful nature of ICLs in preventing replication can be avoided through bypass activity of FANCM/MHF and the photolesion bypass activity of PrimPol prevents fork collapse, in both cases, increasing the cell's tolerance to DNA damage. Without an ortholog of PrimPol in *Drosophila* and given the newly identified potential for traversal of ICLs, it is possible that Pol θ may traverse ICLs instead of unhooking them.

Investigating FA components of ICL repair in *Drosophila*

Moving forward, we plan to further investigate the proposed Pol θ-dependent ICL repair model through studying the activity of FANCD2 during ICL repair in *Drosophila.* Since

FANCD2 null mutations are lethal in *Drosophila*, prospective studies necessitate tissue-specific knockdown alleles of FANCD2 (Marek and Bale, 2006). Epistasis observed between FANCD2 mutants and Pol θ mutants would support our proposed mechanism of activated FANCD2-mediated recruitment of Pol θ during ICL repair. Double mutant studies would also address the possibility of the C-terminus of Pol θ recruiting activated FANCD2, or other FA complex members, to the ICL lesion, a mechanism converse to our proposed mechanism. Successful production of an antibody specific for *Drosophila* Pol θ and for *Drosophila* FANCD2 would, moreover, allow us to identify if FANCD2 and Pol θ co-localize at ICLs, which would classify Pol θ as functioning prior to HR-mediated DSB resolution steps of ICL repair.

Identifying the domain-specific roles of Pol θ during alt-EJ

In *D. melanogaster,* Pol θ has been shown to function during alt-EJ repair of a DSB, although the specific functions of the helicase-like and polymerase domains during this repair process have not been elucidated (Chan et al., 2010). After identifying that the helicase-like and polymerase domains of Pol θ serve different functions during ICL repair, we plan on subjugating the various Pol θ alleles from this study to the P{w^a} assay, a DSB repair assay common to our lab (Adams et al., 2004). Here, we will be able to determine which domain of Pol θ is more important during alt-EJ repair of a DSB by analyzing the relative ratios of repair events between Pol θ mutants.

Understanding the functions of human Pol θ

Our studies in *D. melanogaster* of *mus308* serve great importance in identifying the novel functions of Pol θ during both ICL repair and alt-EJ repair; however, we are equally interested in ascertaining the roles of Pol θ in *D. melanogaster* which are also present in mammals. To address this question, we have successfully cloned the human copy of Pol θ into *D. melanogaster*

in a *mus308*^{null} background (Appendix III). Human Pol θ-dependent rescue of specific phenotypes observed in our *mus308*^{null} flies could serve to identify homologous Pol θ functions between vertebrates and invertebrates. A distinct functional overlap between human and *Drosophila* Pol θ may serve to inform cancer therapeutics which target Pol θ, based on functions identified in flies.

Preliminary data suggests that human Pol θ does not rescue the hatching frequency defect observed in *mus308*^{null} flies; however, the presence of human Pol θ in the absence of *D. melanogaster* Pol θ appears to fully rescue a patchy eggshell phenotype observed in *mus308*^{null} flies (unpublished data). These preliminary data suggest that human Pol θ may rescue some, but not all, of the phenotypes observed in our *mus308*^{null} flies. We plan to investigate whether or not human Pol θ, in the absence of the redundancy found in higher eukaryotes, rescues the observed sensitivities to HN2 observed in Pol θ mutants in this study. Additionally, following analysis of domain specific Pol θ mutants in the P{w^{a}} assay, subsequent analyses of flies with incorporated human Pol θ will help to inform whether or not human Pol θ plays a role in alt-EJ DNA repair of a DSB.

With Pol θ expression levels highly upregulated in many human cancers, Pol θ could serve as a potential therapeutic target. As Pol θ-deficient cells are radiosensitive, one could envision a therapeutic approach coupling radiation therapy with Pol θ-inhibiting drugs (Higgins et al., 2010). Further identification of the domain specific roles of Pol θ during both ICL repair and alt-EJ repair is essential to further mechanistic understanding of both of these repair processes, and, thus, further understanding of how Pol θ can be targeted with cancer therapies.

Literature Cited

1. Lindahl, T. Instability and decay of the primary structure of DNA, Nature 362 709-715, (1993).
2. Handy, D. E. & Loscalzo, J. Redox regulation of mitochondrial function. Antioxid. Redox Signal. 16, 1323–1367, (2012).
3. Slater, M. L. *J. Bacteriol.* 113, 263–270, (1973).
4. Lieber, M.R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end joining pathway. Annu. Rev. Biochem., 79; 181-211, (2010).
5. Hakem, R. DNA-damage repair; the good, the bad, and the ugly. The EMBO Journal. 27, 589-605, (2008).
6. Helleday, T., Lo, J., van Gent, D. C., and Engelward, B. P. DNA double-strand break repair: Fro mechanistic understanding to cancer treatment. DNA Repair. 923-935, (2007).
7. Strumberg, D., Pilon, A. A., Smith, M., Hickey, R., Malkas, L., and Pommier, Y. 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, (2000).
8. Stavnezer, J. Antibody class switching. Adv. Immunol. 61:79-146 (1996).
9. Radding, C. Ann. Rev. Genet. 4. 7: 87-1 11 (1973).
10. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nature Structural & Molecular Biology* **17,** 11–16 (2010).
11. Li, X. and Heyer, W. D. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res 18, 99-113 (2008).
12. Aguilera, A., and Gómez-González, B. Genome instability: a mechanistic view of its causes and consequences. Nat. Rev. Genet. 9, 204-217 (2008).
13. Williams, R.S., Williams, J. S., and Tainer, J. A. Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. Biochem. Cel Biol. 85(4): 509-20, (2007).
14. Heyer, W.-D. Biochemistry of eukaryotic homologous recombination. In: Topics in Current Genetics Vol. 17. Recombination. Edited by A. Aguilera and R. Rothstein. (2007).
15. Plank, J. T., Wu, J. H., and Hsieh, T. S. Topoisomerase III alpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. Proc. Natl. Acad. Sci. USA. 103:11118-11123, (2006).
16. Liu, Y. L., Masson, J. Y., Shah, R., O'Regan, P., and West, S. C. Rad51C is required for Holliday junction processing in mammalian cells. Science. 303:243-246, (2004).
17. Gasior, S. L., Wng, A. K., Kora, Y., Shinohara, A., and Bishop, D.K. Rad52 associates with RPA and functions with Rad55 and Rad57 to assemble meiotic recombination complexes. Genes Dev. 12:2208-2221, (1998).
18. Wold, M.S. Replication protein A: heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. 66:61-92, (1997).
19. Szostak, J. W. Orr-Weaver, T. L., and Rothstein, R. J. The double-strand break repair model for recombination. Cell. 33:25-35, (1983).
20. Sugiyama T., Lantake, N., Wu, Y., and Kowalczykowski, S. C. Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. EMBO J. 25:5539-5548, (2006).

21. Li, X. and Heyer, W.D. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 18(1):99-113, (2008).
22. Nassif, N., Penney, J., Pal, S., Engels, W. R., and Gloor, G.B. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol. Cell Biol. 14:1613-1625, (1994).
23. Barber, L. J., Youds, J. L., Ward, J. D., McIlwraith, M. J., O'Neil, N. J., Petalcorin, M. I., Martin, J. S., Collis, S. J., Cantor, S. B., Auclair, M., et al. RTEL1 maintains genomic stability by suppressing homologous recombination. Cell. 135:261-271, (2008).
24. Paliwal, S., Kanagaraj, R., Sturzenegger, A., Burdova, K., and Janscak, P. Human RECQ5 helicase promotes repair of DNA double-strand breaks by synthesis-dependent strand annealing. Nucleic Acids Research. 42(4):2380-2390, (2014).
25. Morrow, D. M. Connelly, C. and Hieter, P. "Break-copy" duplication: a model for chromosome fragment formation in *Saccaromyces cerevisiae.* Genetics. 147, 371-382, (1997).
26. Malkova, A., Ivankov, E. L., and Haber, J. E. Double-strand break repair in the absence of Rad51 in yeast: a possible role for break-induced DNA replication. Proc. Ntl. Acad. Sci. USA. 93: 7131-7136, (1996).
27. Wilson, M. A., Kwon, Y., Xu, Y., Chung, X., Chi, P., Niu, H., Mayle, R., Chen, X., Malkova, A., Sung, P., and Ira G. Pif1 helicase and Polδ promote recombination-coupled DNA synthesis via bubble migration.
28. Mankouri, H. W., Ngo, H. P., and Hickson, I. D. Shu proteins promote the formation of homologous recombination intermediates that are processed by sgs1-rmi1-top3. Mol. Biol. Cell. 18:4062-4073, (2007).
29. Martin, V., Chahwan, C., Gao, H., Blais, V., Wohlschlegel, J., Yates III, J. R., McGowan, C. H. and Russell, P. Sws1 is a conserved regulator of homologous recombination in eukaryotic cells. EMBO J. 25(11):2564-2574, (2006).
30. Lin, F.-L., Sperle, K., and Sternberg, N. Recombination in mouse L cells between DNA introduced into cells and homologous chromosomal sequences. Proc. Natl. Acad. Sci. USA. 82:1391-1395, (1985).
31. Paques, F. and Jaber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. 63:349-404, (1999).
32. Tutt, A., Bertwistle, D., Valentine, J., Gabriel, A., Swift, S., Ross, G., Griffin, C., Thacker, J., and Ashworth, A. Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. EMBO J. 20:4704-4716, (2001).
33. Mitelman, F., Johansson, B., and Mertens, F. The impact of translocations and gene fusions on cancer causation. Nature Rev. Cancer. 7: 233–245 (2007).
34. Pace, P., Mosedale, G., Hodskinson, M. R., Rosado, I. V., Sivasubramaniam, M, and Patel, K. J. Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. Science. 329:219–223, (2010).
35. Bunting, S. F., Callén, E., Wong, N., Chen, H. T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo,O., Cao, L., Xu, X., Deng, C. X., Finkel, T., Nussenzweig, M., Stark, J. M., and Nussenzweig, A. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell 141, 243–254 (2010)

36. Chan, S. H., Yu, A. M., and McVey, M. Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. PLoS Genetics 6 (7): E100, (2010).
37. Lieber, M. R., Lu, H., Gu, J., and Schwarz, K. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. Cell Res 18:125-133, (2008).
38. Lieber, M.R. The mechanism of human nonhomologous DNA end joining. Journal of Biological Chemistry, 283(1):1-5, (2008).
39. Difilippantonio, M. J., Petersen, S., Chen, H. T.,Johnson, R., Jasin, M., Kanaar, R., Ried, T., and Nussenzweig, A. Evidence for replicative repair of DNA double-strand breaks leading to oncogenic translocation and gene amplification. J. Exp. Med. 196:469–480, (2002).
40. Zhang, Y. & Jasin, M. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. Nature Struct. Mol. Biol. 18:80–84, (2011).
41. Boulton S., Jackson S. P. *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.* 15:5083–5103, (1996).
42. McVey, M., Lee, S.E. MMEJ repair of double strand-breaks (director's cut): deleted sequences and alternative endings. Trends in Genetics. 24:529-538, (2008).
43. Roerink, S. F., van Schendel, R., and Tijsterman, M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Res. 10.1101/gr.170431.113, (2014).
44. Lawley, P.D., and Brookes, P. Cytotoxicity of alkylating agents towards sensitive and resistant strains of Escherichia coli in relation to extent and mode of alkylation of cellular macromolecules and repair of alkylation lesions in deoxyribonucleic acids. Biochem. J. 109:433–447, (1968).
45. Grossmann, K.F., Ward, A.M., Matkovic, M.E., Folias, A.E., and Moses, R.E. *S. cerevisiae* has three pathways for DNA interstrand crosslink repair. Mutat. Res. 487:73–83, (2001).
46. Lawley, P.D., Phillips, D.H. DNA adducts from chemotherapeutic agents. Mutat. Res. 355:13–40, (1996).
47. Stone, M. P., Cho, Y. J., Huang, H., Kim, H. Y., Kozekov, I. D., Kozekova, A., Wang, H., Minko, I. G., Lloyd, R. S., Harris, T. S., and Rizzo, C. J. Interstrand DNA cross-links induced by α,β-unsaturated aldehydes derived from lipid peroxidation and environmental sources. Acc. Chem. Res. 41:793-804, (2008).
48. Brooks, P. J. and Theruvathu, J. A. DNA adducts from acetaldehyde: Implications for alcohol-related carcinogenesis. Alcohol. 35:187-193, (2005).
49. Dutta, S., Chowdhury, G., and Gates, K. S. Interstrand cross-links generated by abasic sites in duplex DNA. J. Am. Chem. Soc. 129:1852-1853, (2007).
50. Bennetts, L. E., De Iuliis, G. N., Nixon, B., Kime,M., Zelski, K., McVicar, C. M., Lewis, S. E., and Aitken, R. J. Impact of estrogenic compounds on DNA integrity in human spermatozoa: evidence for cross-linking and redox cycling activities. Mutat. Res. 641:1-11, (2008).

51. Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar, R. The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinksinto double-strand breaks. EMBO J. 25:4921-4932, (2006).
52. Niedernhofer, L. J., Lalai, A. S., and Hoeijmakers, J. H. Fanconi anemia (cross)linked to DNA repair. Cell. 123:1191-1198, (2005).
53. Fanconi, G. Familial constitutional panmyelocytopathy, Fanconi's anemia (F. A.). I. Clinical aspects. Semin. Hematol. 4:233–240, (1967)
54. Andreassen, P. R. and Ren, Keqin. Fanconi Anemia proteins, DNA interstrand crosslink repair pathways, and cancer therapy. Current Cancer Drug Targets. 9:101-117, (2009).
55. Alpi, A. F., and Patel, K. J. Monoubiquitylation in the Fanconi anemia DNA damage response pathway. DNA Repair. 8:430–435, (2009).
56. Auerbach, A. D. A test for Fanconi anemia. Blood. 72:366-367, (1988).
57. Jonnalagadda, V. S., Matsuguchi, T., and Engelward, B. P. Interstrand crosslink-induced homologous recombination carries an increased risk of deletions and insertions. DNA Repair. 4:594–605, (2005).
58. Ciccia, A., Ling, C., Coulthard, R., Yan, Z., Xue, Y., Meetei, A. R., Laghmani el, H., Joenje, H., McDonald, N., de Winter, J. P., Wang, W., and West, S. C. Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. Mol. Cell. 25:331-343, (2007).
59. Ling, C., Ishiai, M., Ali, A. M., Medhurst, A. L., Neveling, K., Kalb, R., Yan, Z., Xue, Y., Oostra, A. B., Auerbach, A. D., Hoat-lin, M. E., Schindler, D., Joenje, H., de Winter, J. P., Takata, M., and Meetei, A. R. Wang, W. FAAP100 is essential for activation of the Fanconi anemia-associated DNA damage response pathway. EMBO J. 2007, 26, 2104-2114.
60. Kim, H., and D'Andea, A. D. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev. 26:1393-1408, (2012).
61. Ciccia, A., McDonald, N., and West, S.C. Structural and functional relationships of the XPF/MUS81 family of proteins. Annu. Rev. Biochem. 77:259–287, (2008).
62. Bunting, S.F. and Nussenzweig, A. Dangerous liaisons: Fanconi anemia and toxic nonhomologous end joining in DNA crosslink repair. Mol. Cell 39:164–166, (2010).
63. Moynahan, M.E., Pierce, A.J., and Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. Mol. Cell 7:263–272, (2001).
64. Hubscher, U., Maga, G., and Spadari, S. Eukaryotic DNA polymerases. Annu. Rev. Biochem. 71:133–163, (2002).
65. Garg, P. and Burgers, P. M. DNA polymerases that propagate the eukaryotic DNA replication fork. Crit. Rev. Biochem. Mol. Biol. 40:112-128, (2005).
66. Belousova, E. A. and Lavrik, O. I. DNA polymerases beta and lambda, and their roles in the DNA replication and repair. Mol. Biol. (Mosk). 44(6):947-965, (2010).
67. Loeb, L.A., Monnat, R.J. Jr. DNA polymerases and human disease. Nat. Rev. Genet. 9:594–604, (2008).
68. Arana, M.E., Kunkel, T.A. Mutator phenotypes due to DNA replication infidelity. Semin. Cancer Biol. 20:304–311, (2010).
69. Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M.,, and Kunkel, T.A. Division of labor at the eukaryotic replication fork. Mol. Cell 30:137–144, (2008).
70. McCulloch, S. D. and Kunkel, T. A. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res. 18:148-161, (2008).

71. Johnson, R. E., Prakash, S., and Prakash, L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase Polη. Science. 283:1001-1004, (1999).
72. Biertumpfel, C., Zhao, Y., Kondo, Y., Ramon-Maique, S., Gregory,M., Lee, J. Y. Masutani, C., Lehmann, A. R., Hanaoka, F., and Yang, W. Structure and mechanism of human DNA polymerase η. Nature. 465:1044-1048, (2010).
73. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A.K. Protein-template-directed synthesis across an acrolein-derived DNA adduct by yeast Rev1 DNA polymerase. Structure. 16:239-245, (2008).
74. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. Nature. 406:1015-1019, (2000).
75. Bavoux, C., Hoffman, J. S., and Cazaux, C. Adaptation to DNA damage and stimulation of genetic instability: The double-edged sword of mammalian DNA polymerase κ. Biochimie. 87: 637-646, (2005).
76. Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A. Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. Human DNA polymerase κ encircles DNA: Implications for mismatch extension and lesion bypass. Mol. Cell. 25:601-614, (2007).
77. Vaisman, A., Takasawa, K., Iwai, S., and Woodgate, R. DNA polymerase iota-dependent translesion replication of uracil containing cyclobutane pyrimdine dimers. DNA Repair (Amst.). 5(2):210-218, (2006)
78. Petta, T. B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A., and Kannouche, P. Human DNA polymerase iota protects cells against oxidative stress. EMBO J. 27(21):2883-2895, (2008).
79. Shachar, S., Ziv, O., Avkin, S., Adar, S., Wittscieben, J., Reissner, T., Chaney, S., Friedberg, E. C., Wang, Z., Carrell, T., Geacintov, N., and Livneh, Z. Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. EMBO J. 28(4):383-393, (2009).
80. Hashimoto, Y., Chaudhuri, A.R., Lopes, M., and Costanzo, V. Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. Nat Struct. Mol. Biol. 17:1305–1311, (2010).
81. Sharief, F. S., Vojta, P. J., Ropp, P. A., and Copeland, W. C. Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase. Genomics. 59:90-96, (1999).
82. Seki, M., Masutani, C., Yang, L. W., Schuffert, A., Iwai, S., Bahar, I., and Wood, R. D. High-efficiency bypass of DNA damage by human DNA polymerase Q. EMBO J. 23:4484-4494, (2004).
83. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I.B., and Kunkel, T. A. Low-fidelity DNA synthesis by human DNA polymerase theta. Nucleic Acids Res. 36(11):3847-3856, (2008).
84. Harris, P. V., Mazina, O. M., Leonhardt, E. A., Case, R. B., Boyd, J. B., and Burtis, K. C. Molecular cloning of Drosophila mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. Molecular and Cellular Biology. 16(10):5764-5771, (1996).
85. Seki, M., Marini, F., and Wood, R. D. POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. Nucleic Acids Res. 31:6117-6129, (2003).

86. Boyd, J. B., Sakaguchi, K., and Harris, P. V. mus308 mutants of *Drosophila* exhibit hypersensitivity to DNA crosslinking agents and are defective in a deoxyribonuclease. Genetics. 125:813-819, (1990).
87. Marini, F., Kim,N., Schuffert, A., and Wood, R. D. POLN, a nuclear POLA family DNA polymerase homologous to the DNA cross-link sensitivity protein *mus308*. J. Biol. Chem. 278:32014-32019, (2003).
88. Muzzini, D. M., Plevani, P., Boulton, S. J., Cassata, G., and Marini, F. *Caenorhabditis elegans* POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. DNA Repair (Amst.). 7:941-950, (2008).
89. Inagazi, S., Nakamura, K., and Morikami, A. A link among DNA replication, recombination, and gene expression revealed by genetic and genomic analysis of TEBICHI gene of *Arabidopsis thaliana*. PLoS Genet. 5, (2009).
90. Yousefzadeh, M. J. and Wood, R. D. DNA polymerase POLQ and cellular defense against DNA damage. DNA Repair. 12:1-9, (2013).
91. Leonhardt, E. A., Henderson, D. S., Rinehart , J. E., and Boyd, J. B. Characterization of the *mus308* Gene in *Drosophila melanogaster*. Genetics 133:87-96, (1992).
92. Decottihnies, A. Alternative end-joining mechanisms: a historical perspective. Front. Genetics. 4:48, (2013).
93. Asagoshi, K., Lehmann, W., Braithwaite, E. K., Santana-Santos, L., Prasad, R., Freedman, J. H., Van Houten, B., and Wilson, S. H. Single-nucleotide base excision repair DNA polymerase activity in *C. elegans* in the absence of DNA polymerase beta. Nucleic Acids Res. 40:670-681, (2012).
94. Goff, J. P., Shields, D. S., Seki, M., Choi, S., Epperly, M. W., Dixon, T., Wamg, H., Bakkenist, C. J., Dertinger, S. D., Torous, D. K., Wittschieben, J., Wood, R. D., and Greenberger, J. S. Lack of DNA polymerase theta (POLQ) radiosensitizes bone marrow stromal cells in vitro and increases reticuolocyte micronuclei after total-body irradiation. Radiat. Res. 172:165-174, (2009).
95. Yu, A. M. and McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. Nucleic Acids Res. 38:5706-5717, (2010).
96. Zietlow, L., Smith, L. A., Bessho, M., and Bessho, T. Evidence for the involvement of human DNA polymerase N in the repair of DNA interstrand cross-links. Biochemistry. 48(49):11817-11824, (2009).
97. Higgins, G.S., Harris, A.L., Prevo, R., Helleday, T., McKenna, W.G., and Buffa, F.M. Overexpression of POLQ confers a poor prognosis in early breast cancer patients. Oncotarget. 1:175-184, (2010).
98. Lessa, R. C., Campos, A. H., Freita, C. E., Silva, F. R., Kowalski, L. P., Carvalho, A. L., and Vettore, A. L. Identification of upregulated genes in oral squamous cell carcinomas. Head and neck. 35(10):1475-1481, (2013).
99. Pillaire, M. J., Selves, J., Gordien, K., Gorraud, P. A., Gentil, C., Danjoux, M., Do, C., Negre, V., Bieth, A., Guimbaud, R., Trouche, D. Pasero, P., Mechali, M., Hoffman, J. S., and Cazaux, C. A DNA replication signature of progression and negative outcome in colorectal cancer. Oncogene. 29:876-887, (2010).
100. Kawamura, K., Bahar, R., Seimiya, M., Chiyo, M., and Wada, A. et al. DNA polymerase is preferentially expressed in lymphoid tissues and upregulated in human cancers. International Journal of Cancer. 109(1):9-16, (2004).

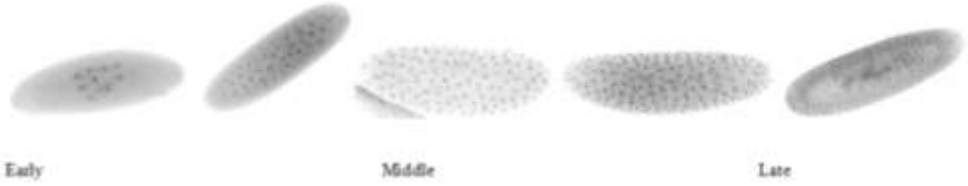
101. Yoon, J. H., Choudhury, J. R., Park, J., Prakash, S., and Prakash, L. A role for DNA polymerase θ in promoting replication through oxidative DNA lesion, thymine glycol, in human cells. J. Biol. Chem. 10.1074/jbc.M114.556977.
102. Voelker, R. A., Ohnishi, S., Langley, C. H., Gausz, J., and Gyurkovics, H. Genetic and Cytogenetic Studies of Malic Enzyme in *Drosophila melanogaster*. Biochemical Genetics 19(5):525-534, (1981).
103. McVey, Mitch. Strategies for DNA interstrand crosslink repair: insights from worms, flies, frogs, and slime molds. Environmental and molecular mutagenesis. 51)6):646-658, (2010).
104. Moldovon, G. L., Madhavan, M. V., Mirchandani, K. D., McCaffrey, R. M., Vinciguerra, P., and D'Andrea, A. D. DNA polymerase POLN participates in cross-link repair and homologous recombination. Mol. Cell. Biol. 30(4):1088-1096, (2010).
105. Huang, J., Liu, S., Bellani, M. A., Thazhathveetil, A. K., Ling, C., de Winter, J. P., Wang, Y., Wang, W., and Seidman M. M. The DNA Translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. Mole. Cell. 52(3):434-446, (2013).
106. Mourón, S., Rodriguez-Acebes, S., Martínez-Jiménez, M. I., García-Gómez, S., Chocrón, S., Blanco, L., and Méndez, J. Repriming of DNA synthesis at stalled replication forks by human PrimPol. Nat. Struct. Mol. Biol. 20(12):1383-1389, (2013).
107. Marek, L. R. and Bale, A. E. *Drosophila* homologs of FANCD2 and FANCL function in DNA repair. DNA Repair. 5:1317-1326, (2006).
108. Adams, M. D., McVey, M., and Sekelsky, J. J. *Drosophila* BLM in double-strand break repair by synthesis-dependent strand annealing. Science. 299(5604):265-267.

Appendix I

$mus308^{\Delta C}$ mutants display developmental abnormalities during early embryogenesis:

Upon fertilization in *D. melanogaster*, the embryo has access to only the maternally loaded RNA, protein, and nutrients, and this preclusion of paternal factors persists for the first two hours of development (Sauer et al., 1996). In these first two hours of embryonic development, nuclei undergo thirteen syncytial nuclear divisions before cellularization of the embryo occurs, and the nuclear divisions occur in a synchronous wave which progresses from one side of the embryo towards the opposite side (Sibon *et al.*, 1997). In wildtype embryos, the distribution of nuclear divisions occurs evenly and synchronously without cellularization throughout the syncytial embryo from stages one to fourteen of early embryogenesis (Figure 1). In contrast to wildtype embryos, the $mus308^{\Delta C}$ embryos, which express the helicase-like domain and part of the linker region of Pol θ but not the polymerase domain, displayed asynchronous divisions, characterized by nuclei in different stages of mitosis randomly dispersed throughout the embryo, and anucleated gaps accompanied by clumps of highly condensed nuclear matter (Figure 3). As the $mus308^{\Delta C}$ embryos aged, the DNA clumping worsened, forming increasingly large aggregates of nuclear matter. Additionally, a significant portion of the embryos DAPI stained demonstrated a single nucleus phenotype. Despite the age of the embryos stained, a substantial proportion of embryos retained the single nucleus phenotype.

w^{1118}



$mus308^{\Delta C}$

Figure 1. Fluorescence image progression of DAPI stains during the first two hours of embryogenesis of $mus308^{\Delta C}$ embryos and w^{1118}embryos. w^{1118} embryos show even distribution of nuclei, equal nuclear density throughout the embryo, a lack of anaphase bridges, and synchronous nuclear divisions during early, middle, and late stages of early embryonic development. In contrast, the $mus308^{\Delta C}$ embryos show gaps throughout the embryo, DNA clumping, anaphase bridges, and asynchronous divisions which persist during early, middle, and late stages of early embryonic development. The extreme developmental defects become apparent in the $mus308^{\Delta C}$ embryos during the middle to late phase of embryogenesis, although developmental defects are present less severely in the early stages as well.

Development and application of the early embryogenesis mutation scoring system:

To quantify any abnormalities present in each embryo, we developed the early embryogenesis mutation scoring system where scores range from zero to six, with a score of six representing the most severe embryonic mutations. In the early embryogenesis mutation scoring system, there are three categories of developmental abnormalities: 1. gaps in the embryo which are indicative of missing nuclei, 2. anaphase bridges, and 3. asynchronous nuclear divisions. Each category could receive a score of zero, one, or two with a score of zero representing no abnormalities, a score of one representing slight abnormalities, and a score of two representing severe abnormalities in that developmental category.

The fourteen stages of early embryonic development in *D. melanogaster* were reduced to three collective phases of early embryogenesis in the early embryogenesis mutation scoring system, early stages (stages one through five), middle stages (stages six through ten), and late stages (stages eleven to cellularization) (Figure 2).

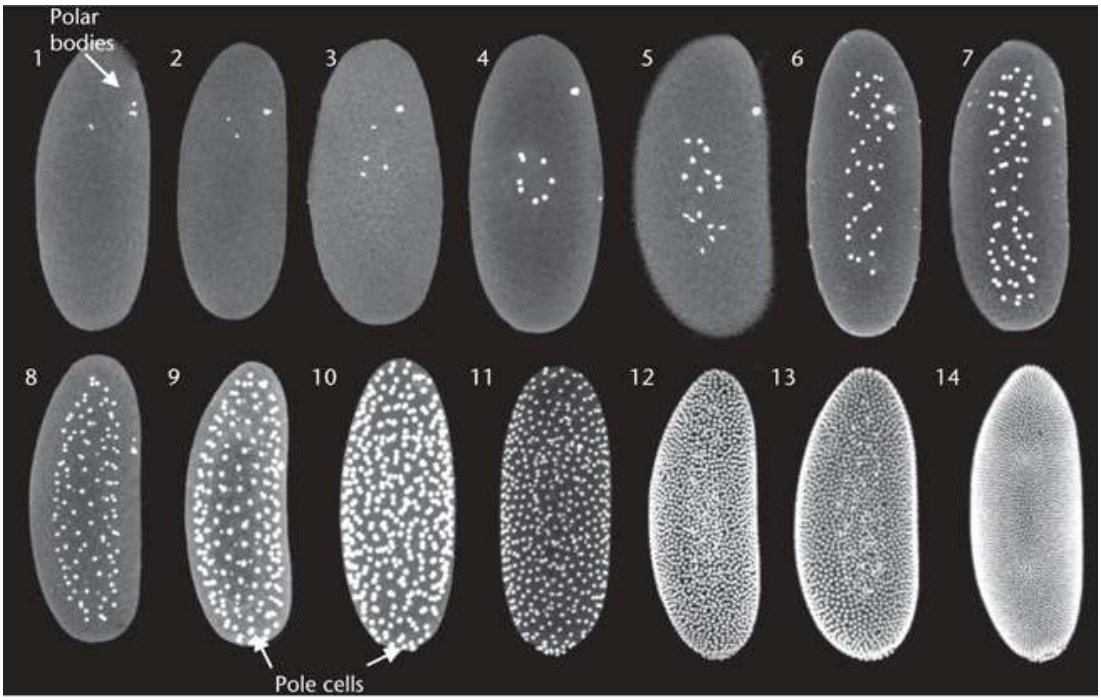


Figure 2. Representation of the fourteen stages of early embryonic development in *D. melanogaster*. Each stage represents a cycle of nuclear division without cellular division. Cellularization occurs at stage fourteen. (Modified from Kotadia et al., 2010).

As the $mus308^{\Delta C}$ embryos progressed developmentally from the early stages of early embryogenesis to the middle stages of early embryonic development, the embryos experienced an increased rate of developmental defects, from an average early embryogenesis mutation score of 0.67 to an average early embryogenesis mutation score of 2.58; in comparison, w^{1118} embryos progressed from an average early embryogenesis mutation score of zero to an average early embryogenesis mutation score of 0.25 (Figure 3). The abnormality scores of the $mus308^{\Delta C}$ embryos in the middle stages of early embryonic development ranged from a score of

zero, representing no developmental defects, to a score of five, which was the most severe score observed in this study. As the embryos progress to the late stages of early embryonic development, the average early embryogenesis mutation score decreases to 1.63 in the *mus308*$^{\Delta C}$ embryos and increases in the w^{1118} embryos to an average early embryogenesis mutation score of 0.53. This could serve to support a time point during the middle stages of early embryonic development where the developing mutations have accumulated to a point where the DNA damage is too severe for embryonic survival.

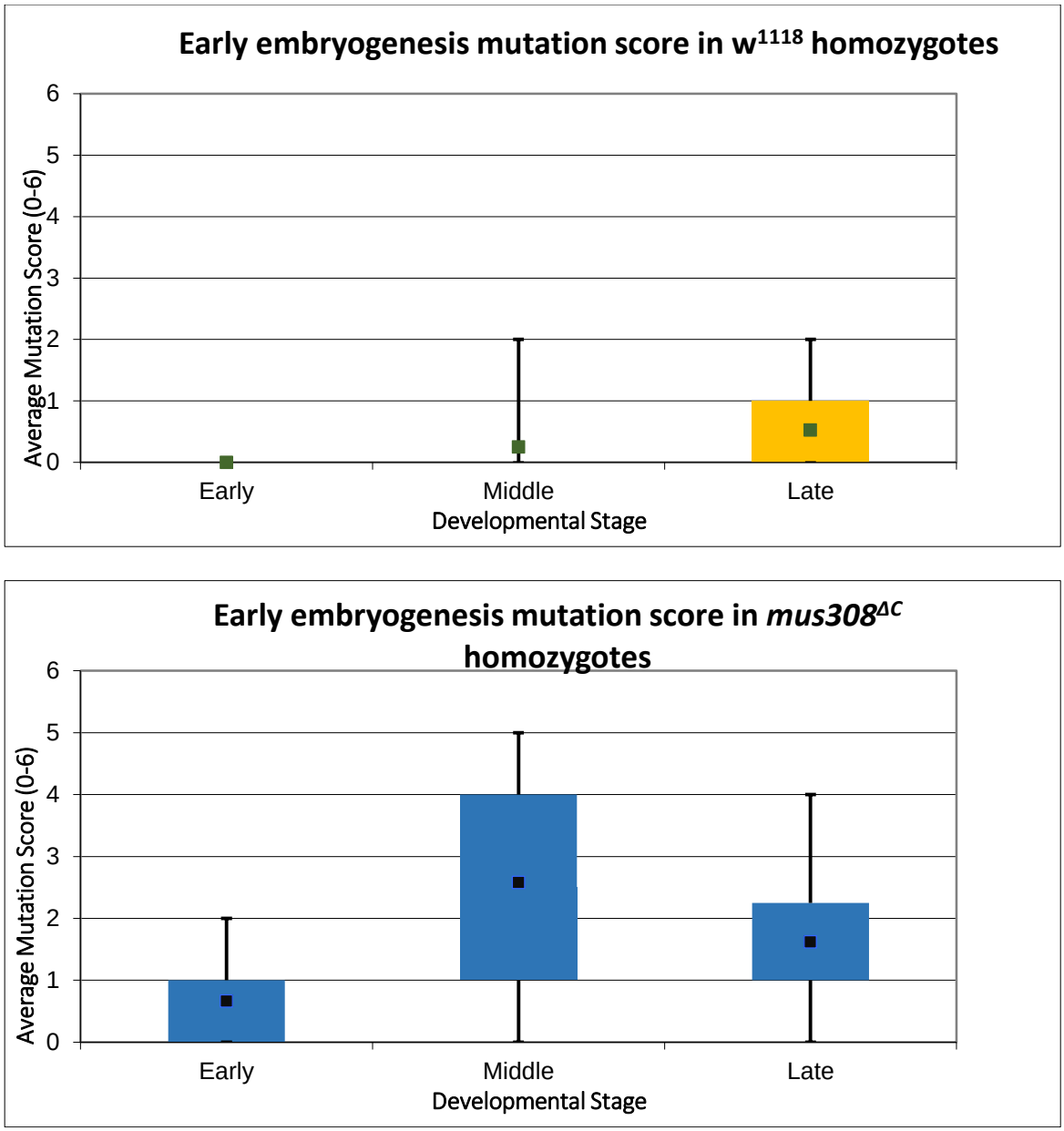


Figure 3. Box and whisker plot of early embryogenesis mutation scores in $mus308^{\Delta C}$ homozygous flies and w^{1118}flies. The w^{1118} flies show an average early embryogenesis mutation score below 1 for all categories, while the $mus308^{\Delta C}$ flies show mean early embryogenesis mutation scores of 2.58 for the middle stages of early embryonic development and 1.63 for the late phases of early embryogenesis. In the middle stages of early embryogenesis, the 75th percentile extends to an early embryogenesis mutation score of 4, which demonstrates that a large percentage of the middle stage embryos were well above the mean score for that particular developmental stage.

Approximately 65% of *mus308*$^{\Delta C}$ embryos scored fell into the lowest mutation severity category (Table 1). This large proportion of embryos with low severity mutations in combination with the fact that the $mus308^{\Delta C}$ mutants have a functional helicase domain of Pol θ but are missing the polymerase domain suggests that the polymerase domain may not always be essential prior to early embryogenesis. Other translesion polymerases may be able to perform a catalytic function redundant to that of the polymerase domain of Pol θ during early embryogenesis. Because extreme developmental defects are capable of arising during early embryogenesis, as demonstrated by the approximately 35% of *mus308*$^{\Delta C}$ embryos with moderate to severe mutations, this may indicate that the polymerase domain of Pol θ could be better suited to function during early embryogenesis than other proteins with relatively redundant functions. When the polymerase domain of Pol θ is not present, severe mutations are capable of developing but do not occur at a high percentage.

Although some $mus308^{\Delta C}$ mutants demonstrate early embryonic developmental defects, the defects were not found to be consistent amongst all the embryos scored. Interestingly, we observed that the majority of the $mus308^{\Delta C}$ embryos visualized were only a single nucleus and, thus, were not counted or scored in this study. If these embryos experienced severe mutations prior to early embryogenesis, then the helicase-like domain of Pol θ could not be enough for the developing fly to overcome damage. If many of the embryos were not developing during early embryogenesis, then the mutations those embryos were experiencing would not have been quantified using this system because the embryos potentially faced extreme DNA damage prior to the nuclear divisions characteristic of early embryogenesis. They were arrested at a single nuclear state and did not further develop due to the truncated form of Pol θ. Perhaps the polymerase domain serves a more critical role than observed in this system.

	Mus308$^{\Delta C}$	w^{1118}
Low (0-2)	**64.70%**	**100%**
Moderate (3-4)	**29.40%**	**0%**
Severe (5-6)	**5.90%**	**0%**

Table 1. Quantification of embryos by mutation severity (on a scale of 0-6) of early embryonic mutations in $mus308^{\Delta C}$ embryos in comparison to w^{1118} embryos.
For mus308$^{\Delta C}$ homozygous embryos, n=51, and for w^{1118} homozygous embryos, n=53. The scoring was based on a 0-2 scale for each of the following defects: asynchronous nuclear divisions, gaps, and anaphase bridges. The highest early embryogenesis mutation score an individual embryo could obtain is a 6. There are no w^{1118} flies with a early embryogenesis mutation score above 2 as all of the embryos visualized fell within the least severe mutation category. In contrast, 64.70% mus308$^{\Delta C}$ fly embryos were in the least severe category, with 29.40% of the mus308$^{\Delta C}$embryos falling into the middle severity category and 5.90% of the embryos in the most severe mutations category.

Literature Cited

1. Sauer, F., Rivera-Pomar, R., Hoch, H., and Jäckle, H. Gene regulation in the *Drosophila* embryo. Philos. Trans. R. Soc. Lond. B Biol. Sci. 351:579-587, (1996).
2. Sibon, O. C. M., Stevenson, V. A., and Theurkauf, W. E. DNA-replication checkpoint control at the *Drosophila* midblastula transition. Nature. 388:93-97, (1997).
3. Kotadia, S., Crest, J., Tram, U., Riggs, B., and Sullivan, W. Blastoderm Formation and Cellularisation in *Drosophila melanogaster*. In: Encyclopedia of Life Sciences (ELS). JohnWiley & Sons, Ltd: Chichester.DOI: 10.1002/9780470015902.a0001071.pub2, (2010).

Appendix II

Pol θ mutants in this study have demonstrated varying degrees of sensitivity to the ICL agent HN2 (nitrogen mustard). However, many of the ICL agents used as chemotherapeutic agents in the clinic for a wide range of carcinomas introduce ICLs of varying structures (Figure 1). In order to clarify that the sensitivities to HN2 in Pol θ mutants observed in this study were not specific to the smaller ICL lesion induced by HN2, we subjected out most severe Pol θ mutant, the *mus308*$^{\Delta C}$ allele, to varying concentrations of cisplatin, an ICL agent which induces a bulkier ICL lesion.

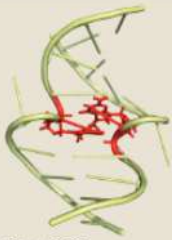
Drug	Clinical application	Dose-limiting toxicity	Example solution structure
Platinums			
Cisplatin	Testicular, ovarian and non-small-cell lung cancer	Central nervous system, renal and gastrointestinal toxicity[153]	 Cisplatin[152]
Carboplatin	Ovarian cancer	Neutropenia[154]	
Oxaliplatin	Colorectal cancer	Neuropathy[155,156]	
Satraplatin	Prostate and breast cancer	Thrombocytopenia and neutropenia[150,157]	
Picoplatin	Phase II trials for relapsed lung cancer and Phase I and II trials for the treatment of solid tumours, and prostate, colorectal and small-cell lung cancer	Thrombocytopenia and neutropenia[158]	
Nitrogen mustards			
Cyclophosphamide	Lymphoma	Neutropenia[100]	 Carmustine[159]
Melphalan	Multiple myeloma, melanoma and ovarian cancer	Leukopenia and thrombocytopenia[101]	
Chlorambucil	Chronic lymphocytic leukaemia	Pancytopenia and neurotoxicity[102]	
Ifosfamide	Non-small-cell lung cancer	Leukopenia, thrombocytopenia and renal toxicity[103]	
Others			
Mitomycin C	Oesophageal and bladder cancer	Leukopenia and thrombocytopenia[105,106]	 Psoralen[104]
Psoralen plus ultraviolet A radiation	Cutaneous T cell lymphoma	Dermatitis[107,108]	
Pyrrolobenzodiazepines	Phase II trial for solid tumours	Fatigue and thrombocytopenia[109]	

Figure 1. Clinically relevant ICL agents used during chemotherapy for various carcinomas. Each chemotherapeutic ICL agent induces ICLs of slightly different structures. The ICL lesion induced by cisplatin, for example, is a much bulkier lesion than that induced by nitrogen mustards. (Modified from Deans and West, 2011).

Because, prior to this study, cisplatin has not been used in treatment of flies in our lab, a wide range of concentrations were used, from 0.01mM-1mM cisplatin. Severe sensitivity in *mus308*$^{\Delta C}$ flies was observed at a concentration of 0.25 mM cisplatin with 18.6% survival relative to a heterozygous control, and complete lethality was observed in *mus308*$^{\Delta C}$ flies at a concentration of 0.5mM cisplatin (Figure 2). The cisplatin sensitivity observed in *mus308*$^{\Delta C}$ flies confirms that the sensitivities of Pol θ mutants in this study are not HN2 specific.

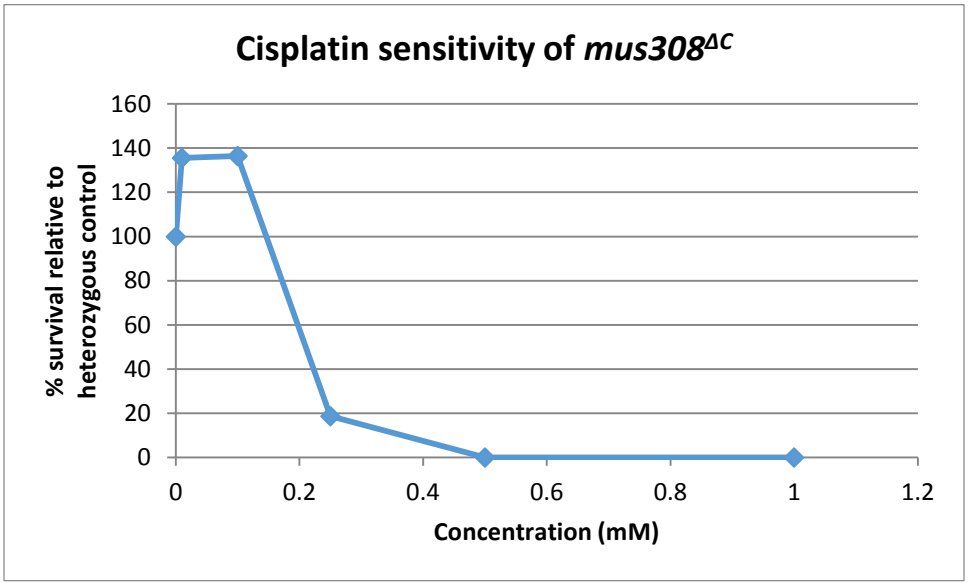


Figure 2. The *mus308*$^{\Delta C}$ allele demonstrates extreme sensitivity to the ICL agent, cisplatin. At a concentration of 0.25mM cisplatin, *mus308*$^{\Delta C}$ flies show 18.6% survival relative to a heterozygous control while complete lethality is observed at concentrations at and above 0.5mM cisplatin. One trial was performed at each concentration.

Literature cited

1. Deans, A. J. and West, S. C. DNA interstrand crosslink repair and cancer. Nature reviews. 11:467-480, (2011).

Appendix III

Human Pol θ was successfully cloned into *Drosophila melanogaster* through ΦC31-transgenesis (Figure 1). Human Pol θ cDNA was cloned onto the second chromosome at cytological position 51C under the control of the endogenous *mus308* promoter in a *mus308*^{null} background. Preliminary data suggests that flies expressing human Pol θ have a hatching frequency similar to that of *mus308*^{null} flies. However, a thin and patchy eggshell phenotype observed in our *mus308*^{null} flies appears to be fully rescued in flies which express the human Pol θ construct. This suggests that human Pol θ performs some, but not all, of the functions of *D. melanogaster* Pol θ. Future studies will investigate whether rescue of ICL sensitivity observed in *Drosophila* Pol θ mutants can be rescued in flies expressing human Pol θ.

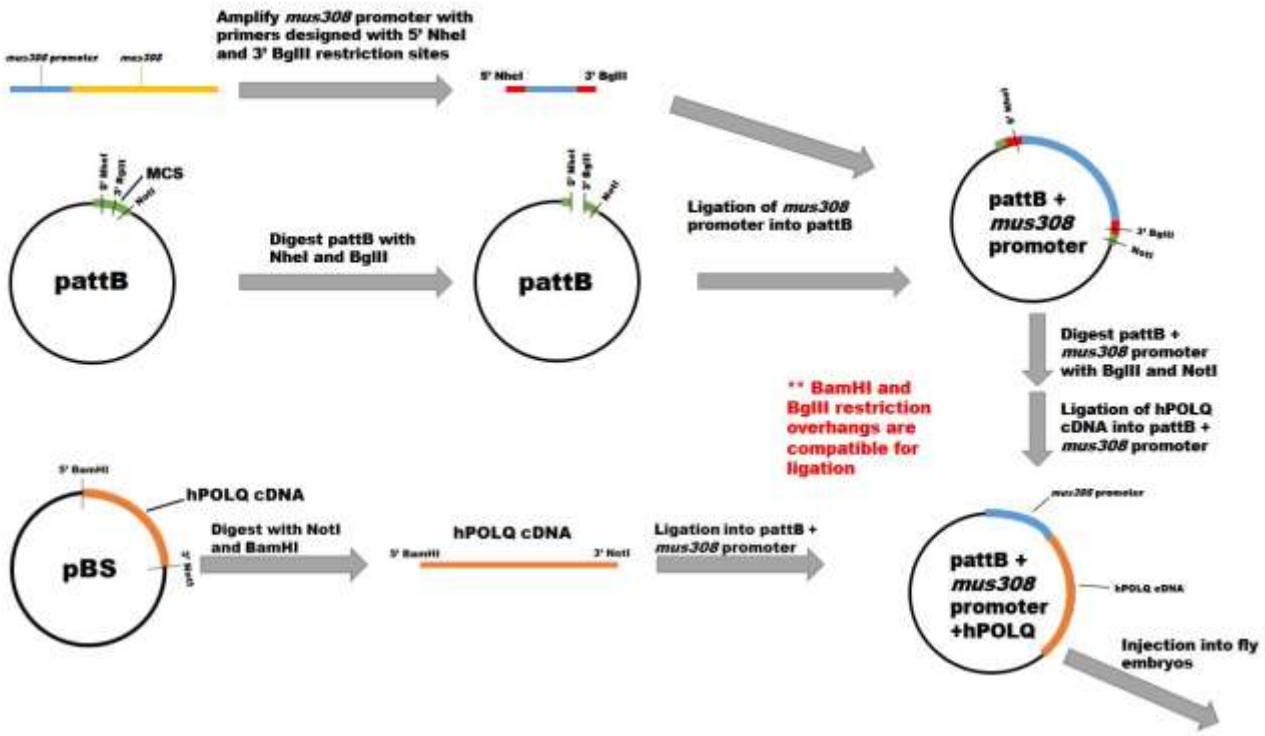


Figure 1. Schematic for cloning of human Pol θ cDNA into *D. melanogaster*. The promoter of *mus308* was PCR amplified with a forward primer with a 5' NheI cut site, 5'-ACCGACGCTAGCATGATGGCCAGATG-3', and a reverse primer with a 3'BglII cut site,

5’-ACCGACAGATCTTGCGCCGGTGCGTTT-3’. The amplified *mus308* promoter, with appropriate restriction enzyme recognition sites, was ligated into the cleaved pattB vector. The pattB vector was cut with NheI in NEB buffer 2; after two hours of digestion with NheI, 0.5µL of 2.5M NaCl, 0.2µL of 1M tris, and BglII were added to the reaction mixture to change NEB buffer 2 to NEB buffer 3. The ligated attB plasmid containing the *mus308* promoter was transformed into XL1-Blue competent cells and plated on amp^{+} LB agar plates; incorporation of the *mus308* promoter into pattB was verified via digestion of mini-preps of colonies with NheI and XhoI. In order to directionally clone the hPOLQ cDNA downstream of the *mus308* promoter, hPOLQ cDNA was digested out of pBS with BamHI and NotI in NEB buffer 3. The sticky overhang generated from cleavage with BamHI anneals to sticky overhangs generated from cleavage with BglII. Therefore, pattB with the *mus308* promoter was digested with BglII and NotI in buffer 3. Following ligation of the hPOLQ cDNA downstream of the *mus308* promoter in pattB, the plasmid was transform into XL1-Blue competent cells and colonies were screened for proper incorporation of the hPOLQ cDNA downstream of the *mus308* promoter in pattB. After successful clone generation, pattB + *mus308* promoter + hPOLQ cDNA was injected into *D. melanogaster* fly embryos. The attB plasmid recombined with attP sites on the second chromosome of *D. melanogaster* at location 51C introducing the cloned construct into the fly genome. Flies with hPOLQ cDNA under the control of the endogenous *mus308* promoter were crossed into a $mus308^{null}$ background for subsequent analyses.