ALTERNATIVE POLYMERASES IN THE MAINTENANCE OF GENOME STABILITY IN *C. ELEGANS*

Sophie Roerink

Layout: Hugo Roerink
Cover design: Sophie Roerink
Cartoon back page: Chris Madden
Printing: GVO Drukkers en vormgevers B.V., Ede, the Netherlands

Printing of this thesis was financially supported by Greiner Bio-One B.V.

ISBN: 978-90-6464-725-3

ALTERNATIVE POLYMERASES IN THE MAINTENANCE OF GENOME STABILITY IN *C. ELEGANS*

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 9 januari 2014
klokke 11.15

door

Sophie Frederique Roerink

geboren te Almelo
in 1982

PROMOTIECOMMISSIE

Promotor:	Prof. dr. L.H.F.Mullenders	Leiden University Medical Center
Co-promotor:	Dr. M. Tijsterman	
Overige leden:	Prof. dr. R. Kanaar	Erasmus Medical Center, Rotterdam
	Prof. dr. S.J.L. van den Heuvel	Utrecht University
	Prof. dr. J.H. den Dunnen	Leiden University Medical Center

TABLE OF CONTENTS

1

GENERAL INTRODUCTION

GENOME INSTABILITY: IMPLICATIONS FOR EVOLUTION AND CANCER

DNA - the carrier of genetic information in the cell - is a chemically reactive molecule. It is constantly attacked by various sources of damage, from endogenous as well as exogenous origin. Replication of damaged bases can result in the incorporation of incorrect bases, introducing mutations in the genome. In addition, the replication process can be disturbed by damage in the template strand eventually leading to a double strand breaks (DSBs) in the DNA that can give rise to chromosomal aberrations and the loss of genetic information.

Both processes contribute to the accumulation of mutations in the genome over time. A limited source of genetic variability is desirable and even necessary. In higher eukaryotes for example, DSB repair and mutagenesis are indispensible for the generation of a diverse pool of antibodies that can attack a wide range of antigens. From an evolutionary point of view, an absolutely stable genome would be incompatible with development of higher order life, since it would not allow for evolutionary adaptation to take place.

Nevertheless, the same process that drives evolution at the population level - natural selection - forms a threat at the level of the organism by causing cancer development (Stratton *et al.* 2009). This can be explained by considering a cell in a multicellular environment as an individual in a population. Mutations that confer the ability to a cell to proliferate and survive more effectively than its neighbours will promote its growth. The transformation of a normal human cell into a cancer cell is caused by a succession of genetic alterations, each conferring a certain type of growth advantage. As stated by the landmark reviews by Hanahan and Weinberg (Hanahan and Weinberg 2000; 2011), almost each cancer cell genome is characterized by mutations leading to eight essential alterations in cell physiology: 1) self-sufficiency in growth signals 2) insensitivity to antigrowth signals 3) evasion of apoptosis 4) limitless replicative potential 5) sustained angiogenesis 6) tissue invasion and metastasis 7) altered energy metabolism and 8) evasion of the immune response. The acquisition of these distinctive and complementary characteristics via mutations in the responsible tumor suppressor genes and oncogenes, causes a cell to grow out into a cancer cell (Figure 1).

Genome instability and cancer: cause or consequence?

DNA replication is a remarkably accurate process, displaying an estimated overall error rate of only one misincorporation in every billion nucleotides (Kunkel and Bebenek 2000). The accuracy in replication is ensured by the proofreading activity of replicative polymerases. In addition, an elaborate set of genome maintenance mechanisms removes base damage and incorrectly incorporated nucleotides (Hoeijmakers 2001). By the combined activity of proofreading and repair mechanisms, mutation levels are kept low.

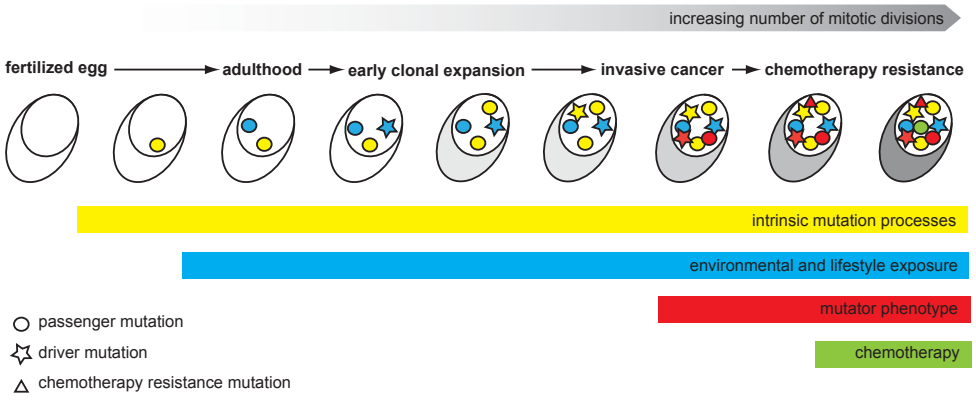


FIGURE 1. Mutational evolution during the outgrowth of a fertilized egg into a cancerous cell. (Stratton *et al.* 2009). Intrinsic mutation processes due to endogenous and exogenous damage are operating during normal cell divisions. The occurrence of driver mutations causes clonal expansion, while passenger mutations do not have any effect on the cancer cell. Rapid cell cycling and additional mutations in repair genes may further contribute to the mutational burden.

In contrast, almost all tumour cells display genomic instability, leading to high mutation levels and chromosomal aberrations (Negrini *et al.* 2010). In fact, genome instability may be considered as another cancer hallmark and act as a driving force in a cell for the stepwise acquisition of the other cancer hallmarks (Negrini *et al.* 2010; Hanahan and Weinberg 2011).

The central role of genome instability in cancer processes is clearly demonstrated by hereditary cancers or cancer susceptibility syndromes that are characterized by mutations in DNA repair genes. Well studied examples include mutations in repair genes that deal with UV-induced damage, causing the skin cancer predisposition disease Xeroderma pigmentosum or mutations in mismatch repair genes, resulting in the colon cancer predisposing Lynch Syndrome.

At the same time, genome instability is also an acquired characteristic of most sporadic tumours, causing accumulation of mutations due to their rapid cell cycles. While driver mutations cause changes that overcome barriers in cell growth and proliferation and actively drive tumor growth, passenger mutations are an essentially neutral byproduct of mutagenesis in cancer cells (Stratton *et al.* 2009).

Characterization of cancer genomes

During recent years, the development of large scale sequencing technology prompted several consortia to sequence the genomes of various cancers and thereby create catalogues of somatic mutations in specific cancers (Pleasance *et al.* 2010; Berger

et al. 2011; Chapman *et al.* 2011; Nik-Zainal *et al.* 2012). Further expansion of these mutation catalogues might lead to the identification of yet unknown driver genes in tumorigenesis.
Furthermore, a full description of cancer genomes compared to matching DNA samples from healthy tissue from the same individuals will help to reconstruct the mutational processes that occurred during tumorigenesis (Nik-Zainal *et al.* 2012).
Although sequencing techniques are rapidly improving and becoming more widely available, sequencing of a human cancer genome is still a costly undertaking, due to its complexity.
In the research described in this thesis, I will describe mutational processes in the nematode *C. elegans*, whose genome is about 32 times as small as the human genome, due to its compact organization. An additional advantage of using *C. elegans* is the possibility to follow mutagenesis in a very controlled fashion.

C. elegans as a model to study genome (in)stability processes

C. elegans has been introduced as a model organism widely before genome sequencing became common practice, in 1974, by Sydney Brenner in order to describe the function of genes in a developing multicellular environment (Brenner 1974). He was looking for a model organism that was simple enough to tackle the most basic questions and easy to maintain under laboratory conditions. The nematode *C. elegans* fulfilled his requirements: it has only 959 cells, it is translucent, fast growing and can be maintained easily on *E. coli* food. Populations consist naturally mainly of hermaphrodites (XX) which can self-fertilize, but males (X) occur spontaneously at a low frequency and enable the combination of genetic characteristics by crossing. Forward genetics by mutagenesis and reverse genetics by use of RNAi resulted in functional characterization of many novel genes, which share homology with their human counterparts (Kamath *et al.* 2003). One of the features of *C. elegans* that makes it an attractive model for developmental biologists is its spatial organization of the different temporal stages of germ cell maturation (Figure 2).
Genome stability mechanisms in *C. elegans* bear resemblance to higher eukaryotes: for almost all known DNA repair routes homologs have been identified in the *C. elegans* genome (O'Neil and Rose 2006). Many genes have been identified by mutagenesis and RNAi screens for hypersensitivity of germ cells to UV-irradiation or X-rays (Hartman and Herman 1982; van Haaften *et al.* 2006).
In the next sections, I will give an overview of the different sources of genome instability, and the processes operating in *C. elegans* and higher eukaryotes to counteract genome instability due to these sources.

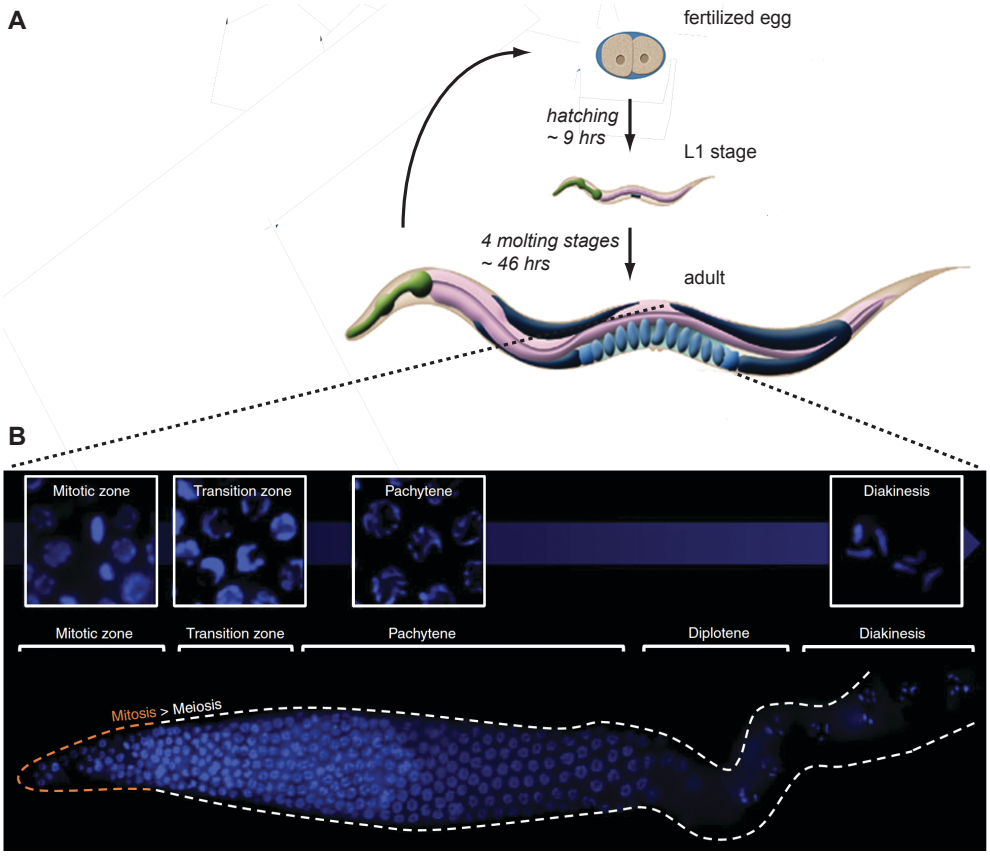


FIGURE 2. *C. elegans* as a model organism.
A. Life cycle of a *C. elegans* hermaphrodite. The two gonad arms that end in the vulva are indicated in blue. Developing germ cells progress gradually through the germline and are fertilized in the spermatheca at the proximal end of the germline. Eggs develop until about 30 cells in utero and then develop further till they hatch at L1 stage. Via four molting stages, animals grow into adults. **B.** DAPI staining of an adult gonad arm, stretched out horizontally (Lemmens and Tijsterman 2010). Mitotic cells at the distal tip behave as stem cells, and produce germ cell precursors. Developing germ cells move through the germline and enter meiosis in the transition zone, which is characterized by the crescent-shaped appearance of highly compact chromatin. In the pachytene zone the clustered chromosomes are redispersed over the nucleus due to alignment of homologous chromosomes. During this stage programmed DSBs trigger meiotic crossovers that physically connect the homologs. During diakinesis, pairs of homologous chromosome can easily be determined in the maturing oocyte as six bivalents.

SOURCES OF GENOME INSTABILITY

Spontaneous instability

Although very infrequently, polymerases can incidentally insert an incorrect nucleotide, resulting in a mismatched base pair. Furthermore, polymerases may induce errors when replicating microsatellites - regions containing repetitive stretches of one to four nucleotides. Slippage of the replication fork on these stretches cause small loops of extrahelical nucleotides that lead to insertions or deletions in the newly generated DNA.

Certain sequences may also interfere with efficient replication due to their ability to form stable secondary structures. An example extensively studied in our group is DNA capable of forming guanine quadruplex (G4) strucures, where guanines are arranged in planar squares stabilized by Hoogsteen bonding (Kruisselbrink *et al.* 2008). This very stable secondary structure is thought to interfere with replication and transcription (Bochman *et al.* 2012).

Some chemical bonds in DNA can be subject to spontaneous desintegration (Barnes and Lindahl 2004). The most frequent event is deamination of cytosine, leading to a miscoding uracil base - a mechanism which is exploited in nature by immune cells for diversification of antibody genes (Barnes and Lindahl 2004).

Base damage

In addition to spontaneous instability, base modifications by endogenous and exogenous reactive compounds limit the coding capacity of DNA and threaten genome stability, if not repaired. An overview of the most common sources of base damage is shown in figure 3.

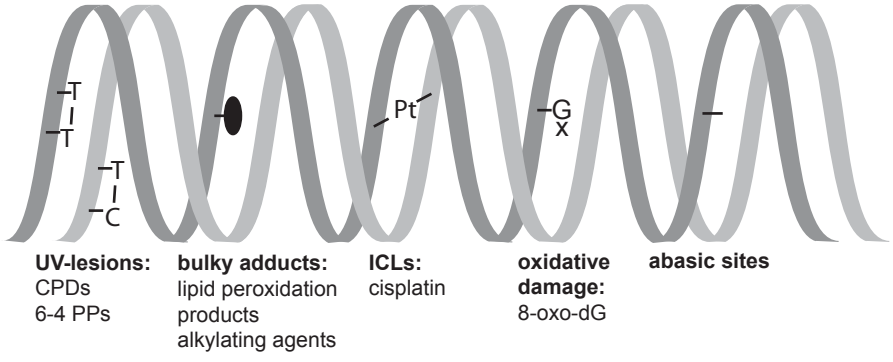


FIGURE 3. Common types of base damage requiring bypass or repair.

Endogenous metabolites

Metabolites formed within the cell can be highly reactive towards DNA. Reactive oxygen species such as superoxide anions and hydroxyl radicals can cause a plethora of base modifications in the DNA (Barnes and Lindahl 2004; Cadet *et al.* 2010). One of the most ubiquitous mutagenic lesion types caused by oxidative stress are 8-oxo-dGs (Barnes and Lindahl 2004). Other oxidative lesions such as thymine glycols are not miscoding, but harmful as they block the progressing replication fork.
Furthermore, lipid peroxidation may cause bulky adducts on DNA bases interfering with replication (Cadet *et al.* 2010). Examples are adducts that result from conjugation of lipid peroxidation-derived aldehydes with the N^2 group of guanine residues.

Environmental factors

Cells are also continuously exposed to environmental agents that cause alterations in DNA structure. One of the most abundant sources of DNA damage is the UV component of sunlight, causing mainly two types of dipyrimidine photoproducts in exposed cells: cyclobutane pyrimidine dimers (CPDs) and 6-4-photoproducts (6-4 PPs). Both lesion types crosslink two adjacent pyrimidines by covalent bonds, resulting in helix distortion and miscoding properties at that position.
In addition, various toxic compounds present in food or in the environment may covalently bind to DNA bases and thereby form so-called bulky adducts. Examples include benzo[a]pyrene present in cigarette smoke, mold-produced aflatoxin, and heterocyclic amines in burnt meat (Wogan *et al.* 2004). Most of these compounds require metabolic activation by cytochrome P450 enzymes to be transformed in electrophiles reacting with DNA.

REPAIR AND BYPASS MECHANISMS

Mismatch repair

Although very infrequently, polymerases may leave errors by misinsertions or slippage of the replication fork. DNA mismatch repair (MMR) is the main repair mechanism to restore these errors (Kunz *et al.* 2009; Peña-Diaz *et al.* 2012). The central role of MMR for genome stability is illustrated by the high cancer predisposition in Lynch syndrome patients, who carry defects in essential mismatch repair genes and develop early onset colon cancer. Key factors in eukaryotic mismatch repair are MutS homologs (MSH2, MSH3 and MSH6 in mammals) which recognize mismatched base pairs, MutL homologs (MLH1 and PMS2 in mammals), which nicks the newly synthesized strand and EXO1 which generates single stranded gaps. Subsequent resynthesis by Pol δ and ligation by Ligase I ensure error-free restoration of the original sequence. In addition to mismatched base pairs, MSH proteins also recognize loops of extrahelical nucleotides that arise during slippage of the replication fork to prevent microsatellite instability.

C. elegans contains homologs for all mentioned mismatch repair genes except for MSH3 (Tijsterman *et al.* 2002).

Repair of base damage

Multiple pathways have evolved that deal with damaged template DNA. Processes that restore the original sequence information are error-free. In some cases, however, only error-prone repair pathways are capable of repairing the damage, which may result in mutations at the lesion site.

Direct reversal

Only a few types of base damage can be completely removed in an error-free fashion by the activity of a single enzyme in a process called direct reversal (Eker *et al.* 2009). O6-meG-DNA-methyltransferase (MGMT) is able to transfer the methyl group to a cysteine residue of the protein, thereby removing O6-methyl-guanyl adducts, produced by alkylating agents. Specialized photolyases can remove CPDs and (6-4) PPs induced by UV light. Although these photolyases are widespread in bacteria and various eukaryotes, *C. elegans* and placental mammals do not contain photolyases.

Base excision repair

Base excision repair (BER) involves the excision of damaged bases, followed by repair of the resulting abasic site by specialized polymerases (Krokan *et al.* 2000; Robertson *et al.* 2009). Most lesion sites targeted by BER are small base alterations due to endogenous damage such as oxygen radicals. Lesion-specific DNA-glycosylases recognize and remove the damaged base. For example 8-oxoguanine-glycosylase (OGG) remove 8-oxo-dG bases. The resulting abasic site (AP site) is processed into a single nucleotide gap by an AP-endonuclease. DNA-polymerases can restore this gap in an error-free manner, either by DNA synthesis of a single nucleotide to bridge the gap (short-patch BER) or by synthesis of about six nucleotides (long-patch BER). The core polymerase acting in BER is Pol β, although also other polymerases have been implicated.
C. elegans extracts do display BER activity in biochemical assays and contain homologs for DNA glycosylases and AP endonucleases operating in BER (Zakaria *et al.* 2010; Asagoshi *et al.* 2012; Hunter *et al.* 2012). However, a homolog for Pol β is lacking in *C. elegans* (Asagoshi *et al.* 2012).

Nucleotide excision repair

Bulky and helix-distorting lesions, such as UV-induced pyrimidine dimers, that interfere with replication and transcription, are substrates for nucleotide excision repair (NER) (Nouspikel 2009). Two subpathways exist: global genome NER (GG-NER) deals with DNA damage in the whole genome while transcription-coupled repair (TCR) only repairs DNA-lesions from the transcribed strand of active genes. Damage recognition is

different in the two subpathways, while the subsequent steps in removal of the lesion are performed by the same set of proteins. In the case of TCR, DNA damage recognition occurs by stalling of RNA polymerase II (RNAPII) at the site of a lesion, which triggers recruitment of Cockayne Syndrome protein B (CSB). The CSB/RNAPII complex facilitates the assembly of a complex consisting of Cockayne Syndrome protein A (CSA) and DNA damage binding protein 1 (DDB1). In the case of GG-NER, lesion recognition is performed by a complex containing the Xeroderma pigmentosum group C protein (XPC), the human homolog of *S. cerevisiae* RAD23 (H23B) and centrin 2. Upon lesion detection, transcription factor IIH (TFIIH) generates a denaturation bubble around the lesion. Subsequent steps in both pathways are binding of the XPA-RPA complex to the DNA, followed by incision of the damaged strand at the 3' end by XPG and at the 5' end by ERCC-1-XPF. Filling of the gap is predominantly performed by replicative polymerases δ and ε, resulting in error-free repair - although activity of the error-prone polymerase κ has also been reported (Ogi *et al.* 2010). Finally, ligases I and III operate to seal the gap.

The two subpathways GG-NER and TCR are conserved in *C. elegans* (Lee *et al.* 2002; Park *et al.* 2002; 2004) and appear to have differential functions during development (Lans *et al.* 2010). While GG-NER is mainly active in germ cells and early embryos, TCR is predominantly engaged in later stages (Lans *et al.* 2010). This study stresses the importance of studying the DNA damage response in a developing organism.

Crosslink repair

Interstrand crosslinks (ICLs) are extremely toxic to the cell as they covalently link the two DNA strands and therefore prevent separation of the two strands as is required during replication and transcription. The repair route to remove these lesions encompasses factors from almost all repair pathways (McVey 2010; Deans and West 2011).

Studies with cell-free Xenopus extracts suggest a model for ICL repair that starts with two incisions in the template strand around the crosslink site, resulting in a DSB in one molecule and an adducted nucleotide opposite a gap in the other homolog (Knipscheer *et al.* 2009). The adducted nucleotide site is bypassed by specific polymerases (Ho and Schärer 2010). After bypass, the lesion can be removed by NER. DSBs generated during crosslink repair are repaired by homologous repair.

Specific for ICL repair in higher eukaryotes are the Fanconi Anemia (FA) proteins, mutated in the cancer predisposition disease Fanconi Anemia. Fourteen complementation groups have been described so far, representing genes encoding fourteen different FA proteins, from which seven factors function together in the FA core complex. The precise function of all individual FA proteins has not been elucidated yet, but it is thought that FA proteins play essential roles in recognition of ICLs and recruitment of necessary repair proteins.

The FA pathway appears to be functionally conserved in *C. elegans*, although it may be represent a simplified model. Homologs for at least five FA components are present, but

not all components of the multiprotein core complex have been identified in *C. elegans* (Youds *et al.* 2009).

Replication fork collapse: a source of double strand breaks

Prolonged stalling of the replication fork due to base damage which has not been repaired in time can eventually lead to replication fork collapse, resulting in a double strand break (DSB).
DSBs are the most genotoxic lesions in the cell, since they can lead to sequence loss or genomic rearrangements. In addition to replication fork collapse, damage by X-rays, crosslinking agents or endogenous sources such as transposition can cause DSBs. If the sister chromatid or homologous chromosome is available as a repair template, DSBs can be repaired in an error-free manner by the process of homologous recombination (HR). Alternatively, the error-prone pathway of non-homologous end joining (NHEJ) can ligate the broken ends of a chromosome (reviewed by Wyman and Kanaar 2006; Chapman *et al.* 2012).

Homologous recombination

In the case of HR, repair starts with detection, binding and processing of broken ends by the MRN complex consisting of MRE11, RAD50 and NBS1. Subsequently, CtIP and Exo1 mediate end resection, generating two 3'OH overhangs that are coated by the single strand binding protein RPA. RPA is then displaced by the recombination factor RAD51, aided by BRCA1, BRCA2 and RAD52. RAD51 catalyzes strand invasion of the sister chromatid, allowing recombinational repair. As a subpathway of HR, RAD51-independent annealing at repeat sequences may take place, followed by flap removal by activity of XPF. This is an error-prone process known as single strand annealing (SSA), which results in deletions.
DSB repair by HR is well conserved in *C. elegans* and plays also a crucial role during meiosis; many HR genes are therefore essential for viability (Lemmens and Tijsterman 2010).

Non-homologous end-joining

In the case of NHEJ, ends are protected from resection by binding of the KU70/80 heterodimer, followed by XRCC 4 - assisted recruitment of Ligase IV, which joins the breaks. In higher eukaryotes, the DNA-PK complex assists end joining, and end processing may require the exonuclease Artemis and polymeraseses λ and μ. NHEJ is characterized by small deletions or insertions at the break site and is thus an error-prone repair mechanism.
Although factors mentioned above are lacking in *C. elegans*, the core components of NHEJ Ku70/80 and Ligase IV are present and crucial for protection against DSBs during development.

The choice between error-free HR and error-prone NHEJ is largely depending on cell cycle phase in somatic cells: during S-phase the availability of a sister chromatid enables HR, while in other stages NHEJ may be the preferred pathway. In *C. elegans* these two repair modes are active during different developmental stages: HR operates mainly in dividing germ cells and early embryos, while NHEJ is active during later somatic development (Clejan *et al.* 2006).

Alternative break repair mechanisms

In addition to these two canonical pathways, numerous lines of evidence proof the existence of a backup 'alternative' end joining pathway, which is not dependent on the canonical NHEJ factors such as Ligase IV and Ku (Audebert *et al.* 2004; Wang *et al.* 2006; Corneo *et al.* 2007; Guirouilh-Barbat *et al.* 2007). At least a subset of these repair products is characterized by deletion of microhomologous sequences around the break site and therefore denoted as microhomology-mediated end-joining (McVey and Lee 2008).

In *C. elegans*, a functional knockout of both HR and NHEJ genes displayed residual repair activity of a transgenic reporter system, inducing somatic DSBs, which suggests the presence of a backup repair pathway (Pontier and Tijsterman 2009).

DNA damage tolerance mechanisms

To avoid genotoxic DSBs due to replication fork collapse, organisms have evolved mechanisms to ensure continuation of DNA synthesis in the presence of unrepaired base damage. These mechanisms are jointly being referred to as DNA damage tolerance (DDT).

The most studied mechanism of DDT is translesion synthesis (TLS) by specialized polymerases, which will be described in detail in the next section. As TLS polymerases are intrinsically error-prone, this process comes at the cost of mutation induction.

Alternatively, error-free DDT may take place, by use of the undamaged sister chromatid as a template. Although error-free DDT is well established as a central mechanism in DNA damage tolerance in *S. cerevisiae*, much less is known about its importance in higher eukaryotes. In yeast, the choice between error prone TLS and error-free damage tolerance, is mediated by modifications at the K164 position of PCNA (Lee and Myung 2008). Rad6/Rad18 mediated ubiquitination triggers recruitment of TLS polymerases, while addition of polyubiquitin chains at this position by MMS2/Rad5/Ubc13 results in error-free damage avoidance. The proteins HLTF and SHPRH have been identified as the mammalian orthologs of Rad5 and are capable of polyubiquitinating PCNA (Motegi *et al.* 2008). However, their functional role appears to be different from their yeast counterpart: they merely play a role in suppressing mutagenesis by TLS in helping recruitment of the correct polymerase (Lin *et al.* 2011).

TLS is functionally conserved in *C. elegans* and will be discussed in more detail in the

next section. Analogously to the situation in mammals, not much is known about error-free DDT mechanisms in *C. elegans*. Rad5, Rad6 and Rad18 genes are conserved, but their functional role is yet unclear.

POLYMERASES IN BYPASS AND REPAIR

The mammalian genome encodes fifteen different DNA polymerases (table 1). Replication of undamaged DNA is carried out by a set of highly efficient polymerases from the B-family: pol α, pol δ and pol ε (Hübscher *et al.* 2002). Pol α functions in a complex with a primase which is able to start DNA synthesis *de novo* by generating an RNA primer of about 10 nucleotides. Pol α extends this DNA-RNA hybrid by about 20 to 30 DNA nucleotides. Subsequently, pol δ and pol ε elongate the DNA in the lagging and leading strands of the replication fork.

Insertion of the correct nucleotides by these polymerases is mainly ensured by two characteristics. First, the binding pocket is shaped in such a way that only correct Watson - Crick basepairs fit without steric clashes (McCulloch and Kunkel 2008; Schmitt *et al.* 2010). Second, pol δ and pol ε contain 3' → 5' proofreading exonuclease domains that are able to remove misincorporated nucleotides. Mice carrying mutant alleles for the exonuclease domains of pol δ and pol ε were shown to die prematurely from carcinogenesis demonstrating the vital role of replication fidelity (Albertson *et al.* 2009).

TABLE 1. Mammalian DNA polymerases and their *C. elegans* homologs

Family	Mammalian homolog	(putative) *C. elegans* homolog	Main function
A	Pol γ	*polg-1*	Mitochondrial DNA replication
	Pol ν	-	
	Pol θ	*polq-1*	Alternative end-joining
B	Pol α	*div-1*	DNA replication priming
	Pol δ	F10C2.4	DNA replication, NER and MMR
	Pol ε	F33H2.5	DNA replication, NER and MMR
	Pol ζ/REV3	Y37B11A.2	TLS
X	Pol β	-	BER
	Pol λ	-	Ig diversification, NHEJ?
	Pol μ	-	Ig diversification, NHEJ?
	TdT	-	Ig diversification
Y	Pol η	*polh-1*	TLS, HR?
	Pol κ	*polk-1*	TLS, NER?
	Pol ι	-	TLS, BER?
	REV1	*rev-1*	TLS

TLS polymerases are characterized by a more flexible binding pocket, allowing for synthesis of DNA across damaged bases (Prakash *et al.* 2005). At least seven of the fifteen DNA polymerases encoded by the mammalian genome have been associated with translesion synthesis activity in *in vitro* or *in vivo* assays: four Y-family members (Pol η, Pol κ, Pol ι and REV1), one B-family member (Pol ζ) and two A-family members (Pol θ and Pol ν) (Lange *et al.* 2011). The more flexible binding site as well as lack of a proofreading exonuclease domain result in a much lower fidelity as compared to the replicative polymerases. However, these structural features are the basis for the unique property of bypass polymerases to ensure continuation of the replication process in spite of base damage in the template strand.

The Y-family of polymerases

The most extensively studied bypass polymerases in eukaryotes are the Y-family polymerases that share their structure with the DinB polymerase, originally identified in *E. coli*. Four family members are known in mammals: Pol η, Pol κ, Pol ι and REV1. *S. cerevisiae* contains only Pol η and REV1, while *C. elegans* contains Pol η, Pol κ and REV1. Here, I will give a more detailed description of the contribution of each Y-family member to mutagenesis and repair processes.

Polymerase η

DNA Polymerase η efficiently and accurately bypasses CPDs generated by exposure to UV light (Johnson *et al.* 1999b). In the absence of Pol η, more error-prone polymerases may take over, resulting in increased mutagenesis (Dumstorf *et al.* 2006; Yoon *et al.* 2009). Its role in bypass of CPD lesion sites most likely also explains the disease phenotype in patients with mutations in the Polymerase η gene. These patients suffer from the variant form of the cancer predisposition disease Xeroderma pigmentosum XP: they are very sensitive to UV-exposure and develop malignant skin neoplasia at young age. (Masutani *et al.* 1999; Johnson *et al.* 1999a). While most XP causing mutations are in core NER genes, patients from complementation group XP-V harbour mutations in the gene encoding Polymerase η. The molecular consequence and the clinical outcome in patients with NER defects and patients with XP-V mutations is largely the same: rapid accumulation of UV-induced mutations, leading to early onset cancer in sunlight-exposed skin tissue (DiGiovanna and Kraemer 2012).

In addition to its *in vivo* role in bypass of CPDs, *in vitro* experiments identified Pol η as an efficient bypasser of various other DNA adducts such as O^6-methylguanines caused by methylating agents (Haracska *et al.* 2000a) and 1,2 d(GpG) intrastrand crosslinks caused by the anticancer drug cisplatin (Alt *et al.* 2007). Endogenous lesions that are effectively bypassed *in vitro* by Pol η include thymine glycols (Kusumoto *et al.* 2002) and 8-oxo-dGs (Haracska *et al.* 2000b).

In contrast to its error-free character in replicating past CPDs or certain other helix

distorting lesions, Pol η is highly error-prone when replicating undamaged templates (Matsuda *et al.* 2000). Nevertheless, Pol η seems to have a protective role in spontaneous mutagenesis, as knockdown in mammalian cells resulted in instability of common fragile sites during unperturbed replication (Rey *et al.* 2009).
Studies in chicken DT40 cells have suggested an additional role for Pol η in repair of DSBs by HR (Kawamoto *et al.* 2005; McIlwraith *et al.* 2005) although the functional implications from this observation are less clear and no HR defects have been observed in mammalian systems.
RNAi studies in *C. elegans* demonstrate a functional role for Pol η: knockdown resulted in increased sensitivity of embryos against the alkylating agent MMS and irradiation by UV(Holway *et al.* 2006; Kim and Michael 2008). Furthermore, knockdown of both Pol η and Pol κ was found to increase instability of stretches capable of forming G-quadruplex structures in the absence of the helicase *dog-1* (Youds *et al.* 2009). Genetic knockouts of both genes will be discussed in detail in this thesis; they did not affect G-quadruplex instability in various reporter systems in *C. elegans* (W. Koole et al, unpublished data).

Polymerase κ

Eukaryotic polymerase κ displays most similarity with bacterial DinB. In spite of its evolutionary conservation in many species studied - except notably in *S. cerevisiae* - its *in vivo* function is not yet fully understood. Purified polymerase κ protein can bypass various N^2-dG adducted sites (Avkin *et al.* 2004; Yuan *et al.* 2008) and also efficiently extends mismatched primer termini (Haracska *et al.* 2002; Carlson *et al.* 2006). Extension by Pol κ of mismatched primer termini resulted in -1 bp frameshifts in biochemical assays, suspective of mutagenic activity of Pol κ (Wolfle *et al.* 2003; Lone *et al.* 2007).
In line with a role for Pol κ in bypassing N^2-dG adducted sites, Pol κ deficient cells were hypersensitive to the smoke carcinogen benzo[a]pyrene (Ogi *et al.* 2002). Sterols and peptides may be an endogenous source of N^2-dG adducted lesion sites that require Pol κ for replication (Mizutani *et al.* 2004; Suzuki *et al.* 2004; Bavoux *et al.* 2005; Waters *et al.* 2009). Other endogenous lesion sites such as thymine glycols, 8-oxo-dG sites and lipid peroxidation-derived adducts may require Pol κ as an extender after the insertion step has been performed by another TLS polymerase (Fischhaber *et al.* 2002; Haracska *et al.* 2002; Wolfle *et al.* 2006).
Compared to the other polymerases, Pol κ has a very low fidelity while replicating undamaged DNA, but a moderate processivity, suggestive of a role in spontaneous mutagenesis (Ohashi *et al.* 2000). In bacteria overexpressing Pol κ homolog DinB, spontaneous mutagenesis due to base substitutions was increased (Kim *et al.* 1997). In eukaryotes, though, Pol κ appears to be protective against spontaneous mutagenesis as suggested by studies in mice (Stancel *et al.* 2009).
As mentioned earlier, Pol κ is also one of the polymerases involved in repair synthesis during NER, as reflected by increased sensitivity of Pol κ deficient cells to UV, which

could not be explained by any direct function in bypass of UV-induced photolesions (Ogi and Lehmann 2006; Ogi *et al.* 2010).
RNAi experiments knocking down Pol κ in *C. elegans* showed increased sensitivity against the alkylating agent MMS. Characterization of a genetic knockout in Pol κ will be described in this thesis.

Polymerase ι

Polymerase ι is an evolutionary more recent member of the Y-family, being present in higher eukaryotes but not in yeast or *C. elegans*. It is reasonably accurate while replicating dA templates but highly error-prone while replicating dT templates (Sale *et al.* 2012).
It can insert the correct nucleotide across an 8-oxo-guanine, thereby preventing mutagenesis at this most abundant oxidative DNA lesion (Kirouac and Ling 2011).
In addition, Pol ι may function as a gap-filling polymerase during BER, further protecting the genome against the toxic effects of oxidative damage (Petta *et al.* 2008).

REV1

In contrast to other Y-family polymerases, REV1's catalytic function is limited to incorporation of dC residues due to the structure of its active site (Nelson *et al.* 1996; Nair *et al.* 2005). In addition to dC incorporation opposite template dG, the unique structure of the catalytic site also allows efficient dC incorporation opposite bulky N^2-dG adducts (Zhang *et al.* 2002; Washington *et al.* 2004) and abasic sites (Nair *et al.* 2011).
S. cerevisiae REV1 mutants were originally characterized as 'reversionless', being impaired in phenotypic reversion of a locus for arginine requirement, due to impaired spontaneous or UV-induced mutagenesis (Lemontt 1971). Multiple experiments showed involvement of mammalian and yeast REV1 in induction of mutagenesis and prevention of cytotoxicity after UV exposure (Gibbs *et al.* 2000; Jansen 2005; Otsuka *et al.* 2005). In addition, REV1 deficient cells or organisms are hypersensitive to various other DNA damaging agents such as the crosslinker cisplatin and the alkylating agent methyl methane sulfonate (Simpson and Sale 2003; Ross *et al.* 2005; D'Souza *et al.* 2008). However, these phenotypes cannot be explained solely by the catalytic activity of REV1, since for most of these functions, the catalytic domain is dispensable. Furthermore, REV1-dependent spontaneous or induced mutagenesis is not characterized by exclusive incorporation of dC residues, which would be the result of its catalytic activity.
Two other functional domains were identified that mediate a more general role in damage bypass: the N-terminal BRCT domain and a C-terminus, that contains ubiquitin binding motifs (UBMs) and interacts with other polymerases. The BRCT domain is essential for resistance to and mutagenesis by various DNA damaging agents and may interact with the DNA sliding clamp PCNA (Jansen 2005; Guo *et al.* 2006). Interaction of the C-terminus with other Y-family polymerases suggests a model where REV1 regulates

the switch to TLS polymerases (Jansen *et al.* 2009b).
The *C. elegans* genome contains a REV1 homolog; this study provides its first functional analysis (Chapter 5).

The B-family member Pol ζ

The Y-family members perform the process of TLS often in cooperation with B-family member Pol ζ. Pol ζ is a heterodimer consisting of the catalytic REV3 subunit and the accessory REV7 subunit. Similarly to REV1, REV3 and REV7 were originally identified as essential factors for induced or spontaneous mutagenesis in *S. cerevisiae* (Lemontt 1971).
The B-family of polymerases further consists of the high-fidelity replicative polymerases α, δ and ε.
Compared to the Y-family polymerases, REV3/Pol ζ is relatively accurate in replicating undamaged templates (Prakash *et al.* 2005). Analogously to Pol κ, Pol ζ functions mainly as an extender, after other TLS polymerases have bypassed the lesion. (Prakash and Prakash 2002). Notably, Pol ζ is proficient in extension from a much wider range of lesion sites than Pol κ, including extension from across CPDs, 6-4 PPs, thymine glycols and N^2dG adducts (Prakash *et al.* 2005).
Its broad function in damage bypass during embryonic development may account for the embryonic lethality of *Rev3-/-* mice (Esposito *et al.* 2000; Wittschieben *et al.* 2000; Gan *et al.* 2008), although yeast *rev3* mutants are viable. Studies in chicken DT40 cells and yeast showed spontaneous chromosomal instability in the absence of REV1 and REV3 (Okada *et al.* 2005).
REV3 is conserved in *C. elegans* but no functional data are available; recent identification of a stop mutant will facilitate further functional analysis of TLS in *C. elegans* in the near future.

Regulation of translesion synthesis

Because of their mutagenic potential, the access of TLS polymerases to DNA is tightly controlled (Jansen *et al.* 2007). A key event for the recruitment of Y-family polymerases in *S. cerevisiae* and mammals is monoubiquitination of replication factor PCNA at lysine 164 (Stelter and Ulrich 2003; Kannouche *et al.* 2004; Bienko *et al.* 2005). This modification is mediated by the Rad6/Rad18 ubiquitin ligase, which responds to the presence of stretches of single stranded DNA, coated with RPA, that are generated by replication fork stalling at lesion sites (Davies *et al.* 2008). Y-family members contain PIP-box motifs that interact directly with ubiquitinated PCNA (Naryzhny 2008) and UBM and UBZ ubiquitin-binding domains that also stimulate interaction with ubiquitinated PCNA (Bienko *et al.* 2005).
Nevertheless, in higher eukaryotes mutations in PCNA K164 are not fully epistatic

with TLS deficiencies, suggesting that also a PCNA K164 - independent route exists for recruitment of TLS polymerases (Edmunds *et al.* 2008; Hendel *et al.* 2011). The non-catalytic function of Y-family member REV1 may also be critical for regulation of TLS by interactions of REV1's C-terminus with the other TLS polymerases (Guo *et al.* 2003; Ohashi *et al.* 2004; 2009).

It remains an intriguing question how the most suitable polymerase is selected to bypass a lesion. Current knowledge has lead to an affinity-based model, where all TLS polymerases associate simultaneously with ubiquitinated PCNA, and trial and error decides which polymerase is most suitable to bypass the lesion (Jansen *et al.* 2007).

Studies in chicken cells suggest two distinct phases of lesion bypass: direct bypass of the DNA lesion at the replication fork which may be independent of PCNA K164 modifications and postreplicative filling of single stranded gaps, for which PCNA K164 is critical (Edmunds *et al.* 2008). The two modes of action of TLS - bypass 'on the fly' versus postreplicative gap-filling - have also been suggested by studies in mammalian cells and *S. cerevisiae* (Jansen *et al.* 2009a; Karras and Jentsch 2010; Daigaku *et al.* 2010).

Finally, upon bypass, ongoing DNA replication needs to switch back to the more precise

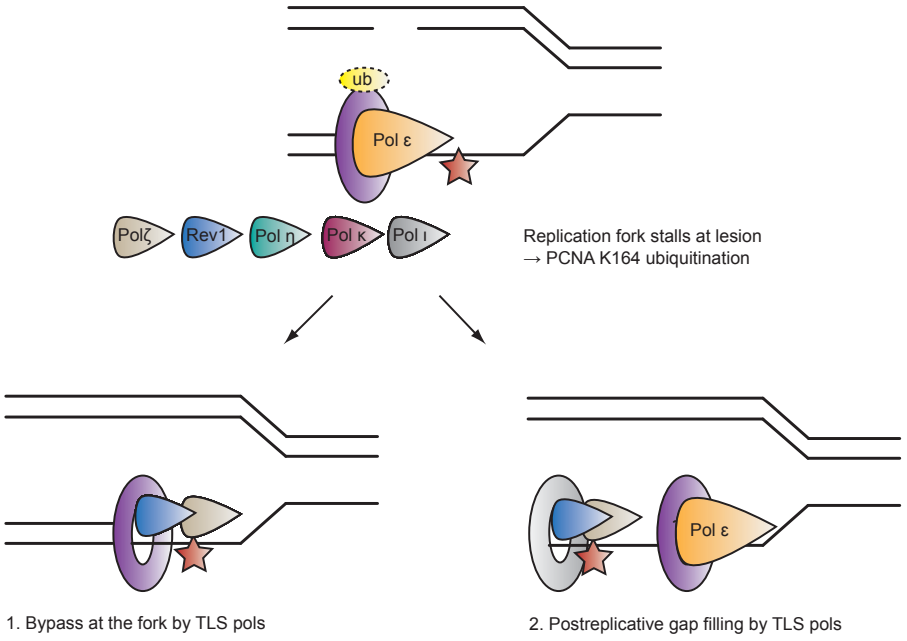


FIGURE 4. Mechanism of translesion synthesis by Y-family polymerases and Pol ζ.
Damaged bases cause stalling of replicative polymerases, resulting in PCNA K164 ubiquitination. TLS polymerases are recruited to the replication fork to overcome the replication barrier by two different modes of action: 1) bypass directly at the fork by polymerase switching 2) postreplicative gap filling.

replicative polymerases. Mutagenesis by activity of error-prone polymerases on undamaged templates is restricted in general by their limited processivity (Jansen *et al.* 2007). Another level of regulation may be deubiquitination of PCNA (Huang *et al.* 2006) or direct modifications of TLS proteins leading to their degradation (Kim and Michael 2008; Bienko *et al.* 2010).
Studies in *C. elegans* embryos showed DNA damage dependent sumoylation and stabilization of Pol η, and controlled degradation of the protein in the absence of these modifications (Kim and Michael 2008).

Other alternative polymerases

Although the term TLS is generally used in reference to the combined activity of the Y-family polymerases and Pol ζ, also polymerases from the X-family and A-family have been shown to possess bypass or extension activity *in vitro*.
While in less complex organisms only one X-family member is present, higher eukaryotes contain four members: Pol β, Pol λ, Pol μ and TdT. Notably, no X-family member has been identified in *C. elegans* or *D. melanogaster* (Uchiyama *et al.* 2009). Pol β is essential for mammalian development, possibly due to its key role in base excision repair (Sobol *et al.* 1996; Sugo *et al.* 2000). Its 5′-deoxyribose lyase domain is able to remove the sugar–phosphate residue at an abasic site, while its polymerase domain can subsequently synthesize DNA to fill the gap (Matsumoto and Kim 1995). Pol λ, which also contains both a 5′-deoxyribose lyase domain and a polymerase domain may act as a backup polymerase in BER (Braithwaite *et al.* 2005). Eukaryotic Pol μ and TdT are mainly expressed in immune cells. Their capacity to extend primer templates with random nucleotides suggests a role in hypermutation of immunoglobulin genes (Domínguez *et al.* 2000).
The A-family of polymerases comprises Pol γ which is the mitochondrial replicase and the only member that has proofreading ability, and the error-prone polymerases Pol θ and Pol ν. Analogously to Pol μ and TdT, A-family Pol θ and Pol ν may contribute to diversification of immunoglobulin genes, although the evidence is conflicting (Kohzaki *et al.* 2010). The structure of Pol θ is unusual in having an N-terminal ATPase-helicase like domain in addition to its C-terminal polymerase domain (Seki *et al.* 2003).
Biochemical assays have demonstrated bypass activity on specific lesions by various A- and X-family members: Pol β is able to bypass certain cisplatin adducts (Hoffmann *et al.* 1995), Pol ν and Pol θ can successfully bypass thymine glycols (Seki *et al.* 2004; Takata 2006) and Pol θ is also an efficient bypasser of abasic sites (Seki *et al.* 2004). However, the *in vivo* relevance of these enzymatic activities for TLS processes is unknown. Furthermore Pol λ, Pol μ and Pol θ are efficient extenders from mismatched primer termini or single stranded ends (Zhang *et al.* 2001; Picher *et al.* 2006; Seki and Wood 2008) albeit in the case of Pol μ being accompanied with frameshifts (Zhang *et al.* 2001).

Polymerases in break repair

Their ability to synthesize DNA on gapped templates may be the basis for a suggested function for Pol λ, Pol μ and Pol θ in break repair (Daley *et al.* 2005; Ramsden and Asagoshi 2012). Biochemical assays showed interaction of the X-family polymerases Pol λ and Pol μ with the Ku complex - a core factor in the process of non-homologous end joining (Ma *et al.* 2004).
In vivo studies demonstrated that Pol λ and Pol μ deficient mice were impaired in recombination of immunoglobulin loci, a process that joins V, D and J elements in an error-prone manner to contribute to antibody gene diversification (Bertocci *et al.* 2003; 2006). The significance of X-family polymerases for efficient break repair in contexts other than V(D)J recombination is unclear: Pol μ deficiency may mildly sensitize cells to irradiation, but the evidence is conflicting (Bertocci *et al.* 2006; Chayot *et al.* 2010).
Pol θ is implicated in tolerance against several endogenous and exogenous sources of DNA damage in mammalian cells, *D. melanogaster* and *C. elegans* (Boyd *et al.* 1990; Shima *et al.* 2004; Ukai *et al.* 2006; Muzzini *et al.* 2008; Goff *et al.* 2009; Li *et al.* 2011) including a profound effect on crosslinking agents as cisplatin (Boyd *et al.* 1990; Muzzini *et al.* 2008; Li *et al.* 2011), which are unlikely to be attributed to its limited capacity in bypassing damaged templates. Studies in Drosophila propose a role for the Pol θ homolog in break repair mediated by microhomology (Chan *et al.* 2010). *C. elegans* Pol θ mutants were found to be highly sensitive to cross linking agents (Muzzini *et al.* 2008).

Polymerases and cancer

As the knowledge on the molecular function of the various polymerases is evolving, the question rises to which extent error-prone polymerases contribute to the genome instability phenotype seen in almost all cancers.
As mentioned previously, the only polymerase gene for which genetic mutations have been identified that are clearly associated with cancer predisposition, is the Pol η or XP-V gene.
Mouse models with a hypomorphic mutation in *Rev1* showed an elevation of UV-induced skin mutations in a NER deficient background, but this phenotype appears to be mediated by inflammation processes and not by increased genome instability (Tsaalbi-Shtylik *et al.* 2009)

Large-scale sequencing techniques that have been developed in the last decade vastly increase the amount of information available on somatic mutation profiles of tumours. In the near future, a more detailed view may be obtained on the presence of somatic mutations in genes encoding bypass polymerases in tumor tissue.
So far, mutations in Pol η and Pol κ have been found in prostate cancer, possibly leading

to hypermutation (Makridakis *et al.* 2009); furthermore Pol η mutations have been identified in breast cancer and melanomas, while Pol κ mutations were found in ovarian and kidney cancers (Makridakis and Reichardt 2012). The base excision repair polymerase Pol β is mutated in about thirty percent of tumors from various tissue origin, and therefore a highly suspected candidate for contributing to genome instability in tumours (Starcevic *et al.* 2004; Donigan *et al.* 2012).

As opposed to inactivating mutations, overexpression or aberrant expression of alternative polymerases may also increase genome instability in cancer cells (Hoffmann and Cazaux 2010). The Y-family member Pol κ was found to be overexpressed in certain lung cancers (O-Wang *et al.* 2001). A study by Lemée and coworkers evaluated expression levels of all known polymerases in breast cancer (Lemée *et al.* 2010). They identified a highly significant upregulation of Pol θ, which together with overexpression studies in tissue culture argue for a causal relation of increased Pol θ levels with tumorigenesis (Lemée *et al.* 2010).

Furthermore, bypass polymerases may play an important role in tolerance to anticancer drugs. Acquired resistance to chemotherapeutic agents that induce DNA lesions - such as cisplatin - is a major therapeutic problem in anticancer therapy. Clinical data showed that expression levels of Pol η - which efficiently bypasses cisplatin induced adducts in cellular assays - were inversely correlated with efficacy of cisplatin therapy in lung cancer (Ceppi *et al.* 2009). Furthermore, inhibitors of Pol β - another efficient bypasser of cisplatin induced lesions - are being tested in preclinical studies to enhance the effect of cisplatin and other DNA damaging agents (Boudsocq *et al.* 2005; Wilson *et al.* 2010).

As more information on somatic mutations and gene expression levels in tumours will be available, bypass polymerases may become new targets for improving anticancer therapy.

AIM OF THIS STUDY

Research described in this thesis focuses on the role of four alternative polymerases in the maintenance of genome stability in *C. elegans*: Pol η, Pol κ, REV1 and Pol θ. This study provides the first functional description of TLS knockout mutants in *C. elegans*. The compact genome of the nematode *C. elegans* allowed for in-depth analysis of genetic changes over time in these mutant backgrounds by whole genome sequencing.

By studying mutagenesis processes in *C. elegans* polymerase mutants, I hope to provide new insights on the role of these polymerases in genome stability processes during evolution and cancer development.

OUTLINE OF THIS THESIS

Chapter 2 focuses on the role of the *C. elegans* homologs of the Y-family polymerases Pol η and Pol κ in protection against DNA damage. I will describe the developmentally regulated role of Pol η and Pol κ in protection against different of DNA damage. Pol η and Pol κ play an important role especially during early embryonic development in *C. elegans*. An RNAi screen identified possible components of this pathway.
Chapter 3 addresses the role of Pol η and Pol κ in genome maintenance under normal growth conditions – i.e. without exposure to exogenous damage. In the absence of functional TLS, endogenous damage causes DSBs in the germline, resulting in spontaneous chromosomal instability. This instability is mediated by error-prone break repair dependent on A-family Polymerase θ.
Chapter 4 presents a transposition-based assay to study the molecular function of Pol θ in the repair of DSBs in the *C. elegans* germline.
Chapter 5 describes a third Y-family polymerase in *C. elegans*: REV1. This polymerase appears to have a broader role in development and genome maintenance, as we observed severe growth problems in *rev-1(null)* mutants. Analogously to *polh-1polk-1* deficient strains, *rev-1* mutants accumulate deletions that may interfere with normal growth and development.
In Chapter 6, I will discuss the main conclusions from this thesis in the context of current literature on the contribution of polymerases on the maintenance of genome stability.

REFERENCES

Albertson T. M., Ogawa M., Bugni J. M., Hays L. E., Chen Y., Wang Y., Treuting P. M., Heddle J. A., Goldsby R. E., Preston B. D., 2009 DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice. Proceedings of the National Academy of Sciences **106**: 17101–17104.

Alt A., Lammens K., Chiocchini C., Lammens A., Pieck J. C., Kuch D., Hopfner K.-P., Carell T., 2007 Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. Science **318**: 967–970.

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

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

Avkin S., Goldsmith M., Velasco-Miguel S., Geacintov N., Friedberg E. C., Livneh Z., 2004 Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa. J Biol Chem **279**: 53298–53305.

Barnes D. E., Lindahl T., 2004 Repair and genetic

consequences of endogenous DNA base damage in mammalian cells. Annu Rev Genet **38**: 445–476.

Bavoux C., Hoffmann J.-S., Cazaux C., 2005 Adaptation to DNA damage and stimulation of genetic instability: the double-edged sword mammalian DNA polymerase kappa. Biochimie **87**: 637–646.

Berger M. F., Lawrence M. S., Demichelis F., Drier Y., Cibulskis K., Sivachenko A. Y., Sboner A., Esgueva R., Pflueger D., Sougnez C., Onofrio R., Carter S. L., Park K., Habegger L., Ambrogio L., Fennell T., Parkin M., Saksena G., Voet D., Ramos A. H., Pugh T. J., Wilkinson J., Fisher S., Winckler W., Mahan S., Ardlie K., Baldwin J., Simons J. W., Kitabayashi N., MacDonald T. Y., Kantoff P. W., Chin L., Gabriel S. B., Gerstein M. B., Golub T. R., Meyerson M., Tewari A., Lander E. S., Getz G., Rubin M. A., Garraway L. A., 2011 The genomic complexity of primary human prostate cancer. Nature **470**: 214–220.

Bertocci B., De Smet A., Berek C., Weill J.-C., Reynaud C.-A., 2003 Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. Immunity **19**: 203–211.

Bertocci B., De Smet A., Weill J.-C., Reynaud C.-A., 2006 Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo. Immunity **25**: 31–41.

Bienko M., Green C. M., Crosetto N., Rudolf F., Zapart G., Coull B., Kannouche P., Wider G., Peter M., Lehmann A. R., Hofmann K., Dikic I., 2005 Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science **310**: 1821–1824.

Bienko M., Green C. M., Sabbioneda S., Crosetto N., Matic I., Hibbert R. G., Begovic T., Niimi A., Mann M., Lehmann A. R., Dikic I., 2010 Regulation of translesion synthesis DNA polymerase eta by monoubiquitination. Mol Cell **37**: 396–407.

Bochman M. L., Paeschke K., Zakian V. A., 2012 DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet. **13**: 770–780.

Boudsocq F., Benaim P., Canitrot Y., Knibiehler M., Ausseil F., Capp J. P., Bieth A., Long C., David B., Shevelev I., Frierich-Heinecken E., Hubscher U., Amalric F., Massiot G., Hoffmann J.-S., Cazaux C., 2005 Modulation of cellular response to cisplatin by a novel inhibitor of DNA polymerase beta. Mol. Pharmacol. **67**: 1485–1492.

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

Braithwaite E. K., Prasad R., Shock D. D., Hou E. W., Beard W. A., Wilson S. H., 2005 DNA polymerase lambda mediates a back-up base excision repair activity in extracts of mouse embryonic fibroblasts. J Biol Chem **280**: 18469–18475.

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics **77**: 71–94.

Cadet J., Douki T., Ravanat J.-L., 2010 Oxidatively generated base damage to cellular DNA. Free Radic. Biol. Med. **49**: 9–21.

Carlson K. D., Johnson R. E., Prakash L., Prakash S., Washington M. T., 2006 Human DNA polymerase kappa forms nonproductive complexes with matched primer termini but not with mismatched primer termini. Proc Natl Acad Sci USA **103**: 15776–15781.

Ceppi P., Novello S., Cambieri A., Longo M., Monica V., Iacono Lo M., Giaj-Levra M., Saviozzi S., Volante M., Papotti M., Scagliotti G., 2009 Polymerase eta mRNA expression predicts survival of non-small cell lung cancer patients treated with platinum-based chemotherapy. Clin. Cancer Res. **15**: 1039–1045.

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

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

Chapman M. A., Lawrence M. S., Keats J. J., Cibulskis K., Sougnez C., Schinzel A. C., Harview C. L., Brunet J.-P., Ahmann G. J., Adli M., Anderson K. C., Ardlie K. G., Auclair D., Baker A., Bergsagel P. L., Bernstein B. E., Drier Y., Fonseca R., Gabriel S. B., Hofmeister C. C., Jagannath S., Jakubowiak A. J., Krishnan A., Levy J., Liefeld T., Lonial S., Mahan S., Mfuko B., Monti S., Perkins L. M., Onofrio R., Pugh T. J., Rajkumar S. V., Ramos A. H., Siegel D. S., Sivachenko A., Stewart A. K., Trudel S., Vij R., Voet D., Winckler W., Zimmerman T., Carpten J., Trent J., Hahn W. C., Garraway L. A., Meyerson M., Lander E. S., Getz G., Golub T. R., 2011 Initial genome sequencing and analysis of multiple myeloma. Nature **471**: 467–472.

Chayot R., Danckaert A., Montagne B., Ricchetti M., 2010 Lack of DNA polymerase μ affects the kinetics of DNA double-strand break repair and impacts on cellular senescence. DNA Repair **9**: 1187–1199.

Clejan I., Boerckel J., Ahmed S., 2006 Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics **173**: 1301–1317.

Corneo B., Wendland R. L., Deriano L., Cui X., Klein I. A., Wong S.-Y., Arnal S., Holub A. J., Weller G. R., Pancake B. A., Shah S., Brandt V. L., Meek K., Roth D. B., 2007 Rag mutations reveal robust alternative end joining. Nature **449**: 483–486.

D'Souza S., Waters L. S., Walker G. C., 2008 Novel conserved motifs in Rev1 C-terminus are required for mutagenic DNA damage tolerance. DNA Repair **7**: 1455–1470.

Daigaku Y., Davies A. A., Ulrich H. D., 2010 Ubiquitin-dependent DNA damage bypass is separable from genome replication. Nature **465**: 951–955.

Daley J. M., Laan R. L. V., Suresh A., Wilson T. E., 2005 DNA joint dependence of pol X family polymerase action in nonhomologous end joining. J Biol Chem **280**: 29030–29037.

Davies A. A., Huttner D., Daigaku Y., Chen S., Ulrich H. D., 2008 Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. Mol Cell **29**: 625–636.

Deans A. J., West S. C., 2011 DNA interstrand crosslink repair and cancer. Nat Rev Cancer **11**: 467–480.

DiGiovanna J. J., Kraemer K. H., 2012 Shining a light on xeroderma pigmentosum. J. Invest. Dermatol. **132**: 785–796.

Domínguez O., Ruiz J. F., Laín de Lera T., García-Díaz M., González M. A., Kirchhoff T., Martínez-A C., Bernad A., Blanco L., 2000 DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. EMBO J **19**: 1731–1742.

Donigan K. A., Sun K.-W., Nemec A. A., Murphy D. L., Cong X., Northrup V., Zelterman D., Sweasy J. B., 2012 Human POLB gene is mutated in high percentage of colorectal tumors. Journal of Biological Chemistry **287**: 23830–23839.

Dumstorf C. A., Clark A. B., Lin Q., Kissling G. E., Yuan T., Kucherlapati R., McGregor W. G., Kunkel T. A., 2006 Participation of mouse DNA polymerase iota in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. Proc Natl Acad Sci USA **103**: 18083–18088.

Edmunds C. E., Simpson L. J., Sale J. E., 2008 PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell **30**: 519–529.

Eker A. P. M., Quayle C., Chaves I., van der Horst G. T. J., 2009 DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems. Cell Mol Life Sci **66**: 968–980.

Esposito G., Godindagger I., Klein U., Yaspo M. L., Cumano A., Rajewsky K., 2000 Disruption of the Rev3l-encoded catalytic subunit of

polymerase zeta in mice results in early embryonic lethality. Curr Biol **10**: 1221–1224.

Fischhaber P. L., Gerlach V. L., Feaver W. J., Hatahet Z., Wallace S. S., Friedberg E. C., 2002 Human DNA polymerase kappa bypasses and extends beyond thymine glycols during translesion synthesis in vitro, preferentially incorporating correct nucleotides. J Biol Chem **277**: 37604–37611.

Gan G. N., Wittschieben J. P., Wittschieben B. Ø., Wood R. D., 2008 DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res **18**: 174–183.

Gibbs P. E., Wang X. D., Li Z., McManus T. P., McGregor W. G., Lawrence C. W., Maher V. M., 2000 The function of the human homolog of Saccharomyces cerevisiae REV1 is required for mutagenesis induced by UV light. Proc Natl Acad Sci USA **97**: 4186–4191.

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

Guirouilh-Barbat J., Rass E., Plo I., Bertrand P., Lopez B. S., 2007 Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. Proceedings of the National Academy of Sciences **104**: 20902–20907.

Guo C., Fischhaber P. L., Luk-Paszyc M. J., Masuda Y., Zhou J., Kamiya K., Kisker C., Friedberg E. C., 2003 Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. EMBO J **22**: 6621–6630.

Guo C., Sonoda E., Tang T.-S., Parker J. L., Bielen A. B., Takeda S., Ulrich H. D., Friedberg E. C., 2006 REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. Mol Cell **23**: 265–271.

Hanahan D., Weinberg R. A., 2000 The hallmarks of cancer. Cell **100**: 57–70.

Hanahan D., Weinberg R. A., 2011 Hallmarks of cancer: the next generation. Cell **144**: 646–674.

Haracska L., Prakash L., Prakash S., 2002 Role of human DNA polymerase kappa as an extender in translesion synthesis. Proc Natl Acad Sci USA **99**: 16000–16005.

Haracska L., Prakash S., Prakash L., 2000a Replication past O(6)-methylguanine by yeast and human DNA polymerase eta. Mol Cell Biol **20**: 8001–8007.

Haracska L., Yu S. L., Johnson R. E., Prakash L., Prakash S., 2000b Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat Genet **25**: 458–461.

Hartman P. S., Herman R. K., 1982 Radiation-sensitive mutants of Caenorhabditis elegans. Genetics **102**: 159–178.

Hendel A., Krijger P. H. L., Diamant N., Goren Z., Langerak P., Kim J., Reissner T., Lee K.-Y., Geacintov N. E., Carell T., Myung K., Tateishi S., D'Andrea A., Jacobs H., Livneh Z., 2011 PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLoS Genet **7**: e1002262.

Ho T. V., Schärer O. D., 2010 Translesion DNA synthesis polymerases in DNA interstrand crosslink repair. Environmental and molecular mutagenesis **51**: 552–566.

Hoeijmakers J. H. J., 2001 Genome maintenance mechanisms for preventing cancer. Nature **411**: 366–374.

Hoffmann J.-S., Cazaux C., 2010 Aberrant expression of alternative DNA polymerases: a source of mutator phenotype as well as replicative stress in cancer. Semin. Cancer Biol. **20**: 312–319.

Hoffmann J.-S., Pillaire M. J., Maga G., Podust V., Hubscher U., Villani G., 1995 DNA polymerase beta bypasses in vitro a single d(GpG)-cisplatin

adduct placed on codon 13 of the HRAS gene. Proc Natl Acad Sci USA **92**: 5356–5360.

Holway A. H., Kim S.-H., La Volpe A., Michael W. M., 2006 Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos. The Journal of Cell Biology **172**: 999–1008.

Huang T. T., Nijman S. M. B., Mirchandani K. D., Galardy P. J., Cohn M. A., Haas W., Gygi S. P., Ploegh H. L., Bernards R., D'Andrea A. D., 2006 Regulation of monoubiquitinated PCNA by DUB autocleavage. Nat Cell Biol **8**: 339–347.

Hunter S. E., Gustafson M. A., Margillo K. M., Lee S. A., Ryde I. T., Meyer J. N., 2012 In vivo repair of alkylating and oxidative DNA damage in the mitochondrial and nuclear genomes of wild-type and glycosylase-deficient Caenorhabditis elegans. DNA Repair **11**: 857–863.

Hübscher U., Maga G., Spadari S., 2002 Eukaryotic DNA polymerases. Annu Rev Biochem **71**: 133–163.

Jansen J. G., 2005 The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Research **33**: 356–365.

Jansen J. G., Fousteri M. I., de Wind N., 2007 Send in the clamps: control of DNA translesion synthesis in eukaryotes. Mol Cell **28**: 522–529.

Jansen J. G., Tsaalbi-Shtylik A., Hendriks G., Gali H., Hendel A., Johansson F., Erixon K., Livneh Z., Mullenders L. H. F., Haracska L., de Wind N., 2009a Separate domains of Rev1 mediate two modes of DNA damage bypass in mammalian cells. Mol Cell Biol **29**: 3113–3123.

Jansen J., Tsaalbi-Shtylik A., Hendriks G., Verspuy J., Gali H., Haracska L., de Wind N., 2009b Mammalian polymerase zeta is essential for post-replication repair of UV-induced DNA lesions. DNA Repair.

Johnson R. E., Kondratick C. M., Prakash S., Prakash L., 1999a hRAD30 mutations in the variant form of xeroderma pigmentosum. Science **285**: 263–265.

Johnson R. E., Prakash S., Prakash L., 1999b Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science **283**: 1001–1004.

Kamath R. S., Fraser A. G., Dong Y., Poulin G., Durbin R., Gotta M., Kanapin A., Le Bot N., Moreno S., Sohrmann M., Welchman D. P., Zipperlen P., Ahringer J., 2003 Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature **421**: 231–237.

Kannouche P. L., Wing J., Lehmann A. R., 2004 Interaction of human DNA polymerase eta with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. Mol Cell **14**: 491–500.

Karras G. I., Jentsch S., 2010 The RAD6 DNA damage tolerance pathway operates uncoupled from the replication fork and is functional beyond S phase. Cell **141**: 255–267.

Kawamoto T., Araki K., Sonoda E., Yamashita Y. M., Harada K., Kikuchi K., Masutani C., Hanaoka F., Nozaki K., Hashimoto N., Takeda S., 2005 Dual roles for DNA polymerase eta in homologous DNA recombination and translesion DNA synthesis. Mol Cell **20**: 793–799.

Kim S. R., Maenhaut-Michel G., Yamada M., Yamamoto Y., Matsui K., Sofuni T., Nohmi T., Ohmori H., 1997 Multiple pathways for SOS-induced mutagenesis in Escherichia coli: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. Proc Natl Acad Sci USA **94**: 13792–13797.

Kim S.-H., Michael W. M., 2008 Regulated proteolysis of DNA polymerase eta during the DNA-damage response in C. elegans. Mol Cell **32**: 757–766.

Kirouac K. N., Ling H., 2011 Unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase iota. Proceedings of the National Academy of Sciences **108**: 3210–3215.

Knipscheer P., Räschle M., Smogorzewska A., Enoiu M., Ho T. V., Schärer O. D., Elledge S. J., Walter J. C., 2009 The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science **326**: 1698–1701.

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

Krokan H. E., Nilsen H., Skorpen F., Otterlei M., Slupphaug G., 2000 Base excision repair of DNA in mammalian cells. FEBS Lett **476**: 73–77.

Kruisselbrink E., Guryev V., Brouwer K., Pontier D. B., Cuppen E., Tijsterman M., 2008 Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ-defective C. elegans. Curr Biol **18**: 900–905.

Kunkel T. A., Bebenek K., 2000 DNA REPLICATION FIDELITY1. Annu Rev Biochem **69**: 497–529.

Kunz C., Saito Y., Schär P., 2009 DNA Repair in mammalian cells: Mismatched repair: variations on a theme. Cell Mol Life Sci **66**: 1021–1038.

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

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

Lans H., Marteijn J. A., Schumacher B., Hoeijmakers J. H. J., Jansen G., Vermeulen W., 2010 Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development. PLoS Genet **6**: e1000941.

Lee K.-Y., Myung K., 2008 PCNA modifications for regulation of post-replication repair pathways. Mol. Cells **26**: 5–11.

Lee M. H., Ahn B., Choi I. S., Koo H.-S., 2002 The gene expression and deficiency phenotypes of Cockayne syndrome B protein in Caenorhabditis elegans. FEBS Lett **522**: 47–51.

Lemée F., Bergoglio V., Fernandez-Vidal A., Machado-Silva A., Pillaire M.-J., Bieth A., Gentil C., Baker L., Martin A.-L., Leduc C., Lam E., Magdeleine E., Filleron T., Oumouhou N., Kaina B., Seki M., Grimal F., Lacroix-Triki M., Thompson A., Roché H., Bourdon J.-C., Wood R. D., Hoffmann J.-S., Cazaux C., 2010 DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proceedings of the National Academy of Sciences **107**: 13390–13395.

Lemmens B. B. L. G., Tijsterman M., 2010 DNA double-strand break repair in Caenorhabditis elegans. Chromosoma **120**: 1–21.

Lemontt J. F., 1971 Mutants of yeast defective in mutation induced by ultraviolet light. Genetics **68**: 21–33.

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

Lin J.-R., Zeman M. K., Chen J.-Y., Yee M.-C., Cimprich K. A., 2011 SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. Mol Cell **42**: 237–249.

Lone S., Townson S. A., Uljon S. N., Johnson R. E., Brahma A., Nair D. T., Prakash S., Prakash L., Aggarwal A. K., 2007 Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. Mol Cell **25**: 601–614.

Ma Y., Lu H., Tippin B., Goodman M. F., Shimazaki N., Koiwai O., Hsieh C.-L., Schwarz K., Lieber M. R., 2004 A biochemically defined system for mammalian nonhomologous DNA end joining.

Mol Cell **16**: 701–713.

Makridakis N. M., Reichardt J. K. V., 2012 Translesion DNA polymerases and cancer. Front Genet **3**: 174.

Makridakis N. M., Caldas Ferraz L. F., Reichardt J. K. V., 2009 Genomic analysis of cancer tissue reveals that somatic mutations commonly occur in a specific motif. Hum. Mutat. **30**: 39–48.

Masutani C., Kusumoto R., Yamada A., Dohmae N., Yokoi M., Yuasa M., Araki M., Iwai S., Takio K., Hanaoka F., 1999 The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. Nature **399**: 700–704.

Matsuda T., Bebenek K., Masutani C., Hanaoka F., Kunkel T. A., 2000 Low fidelity DNA synthesis by human DNA polymerase-eta. Nature **404**: 1011–1013.

Matsumoto Y., Kim K., 1995 Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. Science **269**: 699–702.

McCulloch S. D., Kunkel T. A., 2008 The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res **18**: 148–161.

McIlwraith M. J., Mcllwraith M. J., Vaisman A., Liu Y., Fanning E., Woodgate R., West S. C., 2005 Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination. Mol Cell **20**: 783–792.

McVey M., 2010 Strategies for DNA interstrand crosslink repair: Insights from worms, flies, frogs, and slime molds. Environmental and molecular mutagenesis.

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

Mizutani A., Okada T., Shibutani S., Sonoda E., Hochegger H., Nishigori C., Miyachi Y., Takeda S., Yamazoe M., 2004 Extensive chromosomal breaks are induced by tamoxifen and estrogen in DNA repair-deficient cells. Cancer Res **64**: 3144–3147.

Motegi A., Liaw H.-J., Lee K.-Y., Roest H. P., Maas A., Wu X., Moinova H., Markowitz S. D., Ding H., Hoeijmakers J. H. J., Myung K., 2008 Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. Proceedings of the National Academy of Sciences **105**: 12411–12416.

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

Nair D. T., Johnson R. E., Prakash L., Prakash S., Aggarwal A. K., 2005 Rev1 employs a novel mechanism of DNA synthesis using a protein template. Science **309**: 2219–2222.

Nair D. T., Johnson R. E., Prakash L., Prakash S., Aggarwal A. K., 2011 DNA synthesis across an abasic lesion by yeast REV1 DNA polymerase. Journal of Molecular Biology **406**: 18–28.

Naryzhny S. N., 2008 Proliferating cell nuclear antigen: a proteomics view. Cell Mol Life Sci **65**: 3789–3808.

Negrini S., Gorgoulis V. G., Halazonetis T. D., 2010 Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol **11**: 220–228.

Nelson J. R., Lawrence C. W., Hinkle D. C., 1996 Deoxycytidyl transferase activity of yeast REV1 protein. Nature **382**: 729–731.

Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., Raine K., Jones D., Hinton J., Marshall J., Stebbings L. A., Menzies A., Martin S., Leung K., Chen L., Leroy C., Ramakrishna M., Rance R., Lau K. W., Mudie L. J., Varela I., McBride D. J., Bignell G. R., Cooke S. L., Shlien A., Gamble J., Whitmore I., Maddison M., Tarpey P. S., Davies H. R., Papaemmanuil E., Stephens P. J., McLaren S., Butler A. P., Teague J. W., Jönsson G., Garber J. E., Silver D., Miron P.,

Fatima A., Boyault S., Langerød A., Tutt A., Martens J. W. M., Aparicio S. A. J. R., Borg Å., Salomon A. V., Thomas G., Børresen-Dale A.-L., Richardson A. L., Neuberger M. S., Futreal P. A., Campbell P. J., Stratton M. R., the Breast Cancer Working Group of the International Cancer Genome Consortium, 2012 Mutational Processes Molding the Genomes of 21 Breast Cancers. Cell **149**: 979–993.

Nouspikel T., 2009 DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. Cell Mol Life Sci **66**: 994–1009.

O'Neil N., Rose A., 2006 DNA repair. WormBook : the online review of C elegans biology: 1–12.

O-Wang J., Kawamura K., Tada Y., Ohmori H., Kimura H., Sakiyama S., Tagawa M., 2001 DNA polymerase kappa, implicated in spontaneous and DNA damage-induced mutagenesis, is overexpressed in lung cancer. Cancer Res **61**: 5366–5369.

Ogi T., Lehmann A. R., 2006 The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. Nat Cell Biol **8**: 640–642.

Ogi T., Limsirichaikul S., Overmeer R. M., Volker M., Takenaka K., Cloney R., Nakazawa Y., Niimi A., Miki Y., Jaspers N. G., Mullenders L. H. F., Yamashita S., Fousteri M. I., Lehmann A. R., 2010 Three DNA Polymerases, Recruited by Different Mechanisms, Carry Out NER Repair Synthesis in Human Cells. Mol Cell **37**: 714–727.

Ogi T., Shinkai Y., Tanaka K., Ohmori H., 2002 Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. Proc Natl Acad Sci USA **99**: 15548–15553.

Ohashi E., Bebenek K., Matsuda T., Feaver W. J., Gerlach V. L., Friedberg E. C., Ohmori H., Kunkel T. A., 2000 Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. J Biol Chem **275**: 39678–39684.

Ohashi E., Hanafusa T., Kamei K., Song I., Tomida J., Hashimoto H., Vaziri C., Ohmori H., 2009 Identification of a novel REV1-interacting motif necessary for DNA polymerase kappa function. Genes Cells **14**: 101–111.

Ohashi E., Murakumo Y., Kanjo N., Akagi J.-I., Masutani C., Hanaoka F., Ohmori H., 2004 Interaction of hREV1 with three human Y-family DNA polymerases. Genes Cells **9**: 523–531.

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

Otsuka C., Kunitomi N., Iwai S., Loakes D., Negishi K., 2005 Roles of the polymerase and BRCT domains of Rev1 protein in translesion DNA synthesis in yeast in vivo. Mutat Res **578**: 79–87.

Park H. K., Suh D., Hyun M., Koo H.-S., Ahn B., 2004 A DNA repair gene of Caenorhabditis elegans: a homolog of human XPF. DNA Repair **3**: 1375–1383.

Park H. K., Yook J.-S., Koo H.-S., Choi I. S., Ahn B., 2002 The Caenorhabditis elegans XPA homolog of human XPA. Mol. Cells **14**: 50–55.

Peña-Diaz J., Bregenhorn S., Ghodgaonkar M., Follonier C., Artola-Borán M., Castor D., Lopes M., Sartori A. A., Jiricny J., 2012 Noncanonical Mismatch Repair as a Source of Genomic Instability in Human Cells. Mol Cell.

Petta T. B., Nakajima S., Zlatanou A., Despras E., Couve-Privat S., Ishchenko A., Sarasin A., Yasui A., Kannouche P., 2008 Human DNA polymerase iota protects cells against oxidative stress. EMBO J **27**: 2883–2895.

Picher A. J., Garcia-Diaz M., Bebenek K., Pedersen L. C., Kunkel T. A., Blanco L., 2006 Promiscuous mismatch extension by human DNA polymerase lambda. Nucleic Acids Research **34**: 3259–3266.

Pleasance E. D., Cheetham R. K., Stephens P. J., McBride D. J., Humphray S. J., Greenman C. D., Varela I., Lin M.-L., Ordóñez G. R., Bignell G. R., Ye K., Alipaz J., Bauer M. J., Beare D.,

Butler A., Carter R. J., Chen L., Cox A. J., Edkins S., Kokko-Gonzales P. I., Gormley N. A., Grocock R. J., Haudenschild C. D., Hims M. M., James T., Jia M., Kingsbury Z., Leroy C., Marshall J., Menzies A., Mudie L. J., Ning Z., Royce T., Schulz-Trieglaff O. B., Spiridou A., Stebbings L. A., Szajkowski L., Teague J., Williamson D., Chin L., Ross M. T., Campbell P. J., Bentley D. R., Futreal P. A., Stratton M. R., 2010 A comprehensive catalogue of somatic mutations from a human cancer genome. Nature **463**: 191–196.

Pontier D. B., Tijsterman M., 2009 A robust network of double-strand break repair pathways governs genome integrity during C. elegans development. Curr Biol **19**: 1384–1388.

Prakash S., Prakash L., 2002 Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. Genes Dev **16**: 1872–1883.

Prakash S., Johnson R. E., Prakash L., 2005 EUKARYOTIC TRANSLESION SYNTHESIS DNA POLYMERASES: Specificity of Structure and Function. Annu Rev Biochem **74**: 317–353.

Ramsden D. A., Asagoshi K., 2012 DNA polymerases in nonhomologous end joining: Are there any benefits to standing out from the crowd? Environmental and molecular mutagenesis **53**: 741–751.

Rey L., Sidorova J. M., Puget N., Boudsocq F., Biard D. S. F., Monnat R. J., Cazaux C., Hoffmann J.-S., 2009 Human DNA Polymerase Is Required for Common Fragile Site Stability during Unperturbed DNA Replication. Mol Cell Biol **29**: 3344–3354.

Robertson A. B., Klungland A., Rognes T., Leiros I., 2009 DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell Mol Life Sci **66**: 981–993.

Ross A.-L., Simpson L. J., Sale J. E., 2005 Vertebrate DNA damage tolerance requires the C-terminus but not BRCT or transferase domains of REV1. Nucleic Acids Research **33**: 1280–1289.

Sale J. E., Lehmann A. R., Woodgate R., 2012 Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat Rev Mol Cell Biol **13**: 141–152.

Schmitt M. W., Venkatesan R. N., Pillaire M.-J., Hoffmann J.-S., Sidorova J. M., Loeb L. A., 2010 Active site mutations in mammalian DNA polymerase delta alter accuracy and replication fork progression. J Biol Chem **285**: 32264–32272.

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

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

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

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

Simpson L. J., Sale J. E., 2003 Rev1 is essential for DNA damage tolerance and non-templated immunoglobulin gene mutation in a vertebrate cell line. EMBO J **22**: 1654–1664.

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

Stancel J., McDaniel L., Velasco S., Richardson J., Guo C., Friedberg E., 2009 Polk mutant mice have a spontaneous mutator phenotype. DNA Repair.

Starcevic D., Dalal S., Sweasy J. B., 2004 Is there a link between DNA polymerase beta and cancer? Cell Cycle **3**: 998–1001.

Stelter P., Ulrich H. D., 2003 Control of spontaneous and damage-induced mutagenesis by SUMO

and ubiquitin conjugation. Nature **425**: 188–191.

Stratton M. R., Campbell P. J., Futreal P. A., 2009 The cancer genome. Nature **458**: 719–724.

Sugo N., Aratani Y., Nagashima Y., Kubota Y., Koyama H., 2000 Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta. EMBO J **19**: 1397–1404.

Suzuki N., Itoh S., Poon K., Masutani C., Hanaoka F., Ohmori H., Yoshizawa I., Shibutani S., 2004 Translesion synthesis past estrogen-derived DNA adducts by human DNA polymerases eta and kappa. Biochemistry **43**: 6304–6311.

Takata K. I., 2006 Human DNA Polymerase N (POLN) Is a Low Fidelity Enzyme Capable of Error-free Bypass of 5S-Thymine Glycol. Journal of Biological Chemistry **281**: 23445–23455.

Tijsterman M., Pothof J., Plasterk R. H. A., 2002 Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics **161**: 651–660.

Tsaalbi-Shtylik A., Verspuy J. W. A., Jansen J. G., Rebel H., Carlée L. M., van der Valk M. A., Jonkers J., de Gruijl F. R., de Wind N., 2009 Error-prone translesion replication of damaged DNA suppresses skin carcinogenesis by controlling inflammatory hyperplasia. Proceedings of the National Academy of Sciences **106**: 21836–21841.

Uchiyama Y., Takeuchi R., Kodera H., Sakaguchi K., 2009 Distribution and roles of X-family DNA polymerases in eukaryotes. Biochimie **91**: 165–170.

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

van Haaften G., Romeijn R., Pothof J., Koole W., Mullenders L. H. F., Pastink A., Plasterk R. H. A., Tijsterman M., 2006 Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. Curr Biol **16**: 1344–1350.

Wang M., Wu W., Wu W., Rosidi B., Zhang L., Wang H., Iliakis G., 2006 PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic Acids Research **34**: 6170–6182.

Washington M. T., Minko I. G., Johnson R. E., Haracska L., Harris T. M., Lloyd R. S., Prakash S., Prakash L., 2004 Efficient and error-free replication past a minor-groove N2-guanine adduct by the sequential action of yeast Rev1 and DNA polymerase zeta. Mol Cell Biol **24**: 6900–6906.

Waters L. S., Minesinger B. K., Wiltrout M. E., D'Souza S., Woodruff R. V., Walker G. C., 2009 Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiol Mol Biol Rev **73**: 134–154.

Wilson S. H., Beard W. A., Shock D. D., Batra V. K., Cavanaugh N. A., Prasad R., Hou E. W., Liu Y., Asagoshi K., Horton J. K., Stefanick D. F., Kedar P. S., Carrozza M. J., Masaoka A., Heacock M. L., 2010 Base excision repair and design of small molecule inhibitors of human DNA polymerase β. Cell Mol Life Sci **67**: 3633–3647.

Wittschieben J., Shivji M. K., Lalani E., Jacobs M. A., Marini F., Gearhart P. J., Rosewell I., Stamp G., Wood R. D., 2000 Disruption of the developmentally regulated Rev3l gene causes embryonic lethality. Curr Biol **10**: 1217–1220.

Wogan G. N., Hecht S. S., Felton J. S., Conney A. H., Loeb L. A., 2004 Environmental and chemical carcinogenesis. Semin. Cancer Biol. **14**: 473–486.

Wolfle W. T., Johnson R. E., Minko I. G., Lloyd R. S., Prakash S., Prakash L., 2006 Replication past a trans-4-hydroxynonenal minor-groove adduct by the sequential action of human DNA polymerases iota and kappa. Mol Cell Biol **26**: 381–386.

Wolfle W. T., Washington M. T., Prakash L., Prakash S., 2003 Human DNA polymerase kappa uses template-primer misalignment as a novel means for extending mispaired termini and for generating single-base deletions. Genes Dev **17**: 2191–2199.

Wyman C., Kanaar R., 2006 DNA double-strand break repair: all's well that ends well. Annu Rev Genet **40**: 363–383.

Yoon J.-H., Prakash L., Prakash S., 2009 Highly error-free role of DNA polymerase eta in the replicative bypass of UV-induced pyrimidine dimers in mouse and human cells. Proc Natl Acad Sci USA **106**: 18219–18224.

Youds J. L., Barber L. J., Boulton S. J., 2009 C. elegans: a model of Fanconi anemia and ICL repair. Mutat Res **668**: 103–116.

Yuan B., Cao H., Jiang Y., Hong H., Wang Y., 2008 Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase in vitro and in vivo. Proceedings of the National Academy of Sciences **105**: 8679–8684.

Zakaria C., Kassahun H., Yang X., Labbé J.-C., Nilsen H., Ramotar D., 2010 Caenorhabditis elegans APN-1 plays a vital role in maintaining genome stability. DNA Repair **9**: 169–176.

Zhang Y., Wu X., Rechkoblit O., Geacintov N. E., Taylor J.-S., Wang Z., 2002 Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Research **30**: 1630–1638.

Zhang Y., Wu X., Yuan F., Xie Z., Wang Z., 2001 Highly frequent frameshift DNA synthesis by human DNA polymerase mu. Mol Cell Biol **21**: 7995–8006.

A BROAD REQUIREMENT FOR TLS POLYMERASES η AND κ, AND INTERACTING SUMOYLATION AND NUCLEAR PORE PROTEINS DURING *C. ELEGANS* EMBRYOGENESIS

2

Sophie F. Roerink, Wouter Koole, L. Carine Stapel,
Ron J Romeijn, Marcel Tijsterman

Leiden University Medical Center,
Department of Toxicogenetics

An adapted version of this manuscript has been
published in Plos Genet 2012 vol 8 (6)

ABSTRACT

Translesion synthesis (TLS) polymerases are specialized DNA polymerases capable of inserting nucleotides opposite DNA lesions that escape removal by dedicated DNA repair pathways. TLS polymerases allow cells to complete DNA replication in the presence of damage, thereby preventing checkpoint activation, genome instability and cell death. Here, we characterize functional knockouts for *polh-1* and *polk-1*, encoding the *C. elegans* homologs of the Y-family TLS polymerases η and κ. POLH-1 acts at many different DNA lesions as it protects cells against a wide range of DNA damaging agents, including UV, γ-irradiation, cisplatin and methyl methane sulphonate (MMS). POLK-1 acts specifically but redundantly with POLH-1 in protection against methylation damage. Importantly, both polymerases play a role specifically early in embryonic development to allow fast replication of damaged genomes. Contrary to observations in mammalian cells, we show that neither POLH-1 nor POLK-1 is required for homologous recombination (HR) repair of DNA double-strand breaks. A genome wide RNAi screen for genes that protect the *C. elegans* genome against MMS-induced DNA damage identified novel components in DNA damage bypass in the early embryo. Our data suggest SUMO-mediated regulation of both POLH-1 and POLK-1, and point towards a previously unrecognized role of the nuclear pore in regulating TLS.

INTRODUCTION

DNA damaging agents from both endogenous and exogenous sources can induce replication-blocking DNA lesions that threaten cell cycle progression and, consequently, cell viability. To remove these DNA lesions cells are equipped with various specialized repair mechanisms (Ciccia and Elledge 2010), including nucleotide excision repair (NER) that deals with helix-distorting obstructions (Nouspikel 2009). However, during embryogenesis, which entails phases of rapid cell division, only a limited time window is available for repair processes (O'Farrell *et al.* 2004). Consequently, unrepaired DNA damage may delay replication and cell cycle progression. In the nematode *C. elegans* a delay in replication is detrimental for the developmental program; timing of cell division is fixed and strictly regulated by the homologues of the checkpoint genes CHK1 and ATR (Encalada *et al.* 2000). Indeed, replication stress caused by depletion of nucleotide pools causes fatal errors in the correct timing of the first asynchronous divisions (Brauchle *et al.* 2003a). However, early embryos of *C. elegans* appear to be remarkably resistant to DNA damaging agents, suggesting an efficient way to prevent the induction of replication stress by DNA damage (Holway *et al.* 2006).

To be able to deal with replication obstructions, organisms evolved ways that allow bypass of the damaged template, thus ensuring continuity of the replication process (Andersen *et al.* 2008). Specialized TLS polymerases are capable of direct bypass of DNA lesions in an error-free or error-prone fashion, depending on their affinities for the specific lesion site. In eukaryotes TLS is mediated by the DNA polymerases of the Y-family, Polη, Polκ, Polι, Rev1, and the B-family member Polζ. All members of the Y-family polymerases lack proofreading activity and share a conserved active site, which is different from the high-fidelity polymerases in its open and less sterically constrained structure. It allows for accommodation of a DNA lesion, but is also the basis for reduced fidelity (Prakash *et al.* 2005). The functional specificities of TLS polymerases are due to minor differences in the structural features of the active site.

The *C. elegans* genome encodes several Y-family TLS proteins, including POLH-1 and POLK-1, homologs of mammalian Polη and Polκ, respectively. Purified Polη of yeast and vertebrates is capable of replicating across a wide variety of DNA damages, including UV-induced cyclobutane pyrimidine dimers (CPDs), 7,8-dihydro-8-oxoguanine, O^6-methylguanine, thymine glycol, cisplatin-induced intrastrand crosslinks, acetylaminofluorene-adducted guanine and benzo[*a*]pyrene-N^2-guanine (Waters *et al.* 2009b). In humans, defective Polη has clinical implications. Polη is the product of the gene mutated in Xeroderma Pigmentosum complementation group "Variant" (XPV), a syndrome that is associated with a high predisposition towards developing skin cancers (Johnson *et al.* 1999a). In addition to a role in damage bypass some studies have suggested a role for Polη in homologous recombination, as the polymerase responsible for extension of the invading strand in the D-loop recombination intermediate (Kawamoto *et al.* 2005; McIlwraith *et al.* 2005). Recently, it was reported that Polη

plays a prominent role in early stages of nematode embryogenesis in *C. elegans* (Holway *et al.* 2006; Kim and Michael 2008b). Polκ displays structural similarity to Polη but is considered to be the most evolutionarily conserved member of the Y-family showing homology to prokaryotic DinB (Prakash *et al.* 2005; Waters *et al.* 2009a). Its substrate specificity *in vitro* is limited, although Polκ is an efficient extender of mispaired primer termini and some guanyl adducted lesion sites (Carlson *et al.* 2006; Haracska *et al.* 2002). Furthermore, Polκ has been suggested as one of the gap-filling polymerases in NER, explaining a moderate sensitivity of Polκ-deficient mammalian cells to UV (Ogi and Lehmann 2006; Ogi *et al.* 2010).

Here, we characterize the involvement of Polη and Polκ in various aspects of genome protection during animal development, using the model organism *C. elegans.* The advantages of this animal model are its spatial and temporal organization of gametogenesis and its rapid growth properties that allow monitoring DNA repair or lesion bypass during different developmental stages. We found that POLH-1 is involved in protection against a surprisingly wide range of DNA lesions, whereas the substrate specificity of POLK-1 is much more restricted. Both proteins can act redundantly on some lesions, since double mutants were extremely sensitive to the alkylating agent MMS, whereas both single mutants displayed profoundly less sensitivity to this carcinogen. In spite of their error-proneness, POLH-1 and POLK-1 appear to be highly important in protection against DNA damage during embryonic development, while their role in later somatic development is limited. Finally, we used genome-wide RNAi to screen for factors that have a similar sensitivity profile leading to the identification of new factors that may play a role in the regulation of TLS.

RESULTS

Isolation of *C. elegans* mutants for *polh-1* and *polk-1*.

To study the function of Y-family TLS polymerases in the DNA damage response at different stages of animal development, we set out to isolate mutants for the *C. elegans* homologs of the Polη and Polκ genes. Figure 1A illustrates a phylogenetic tree of the Y-family polymerase members from several species including *C. elegans*. The *C. elegans* genome encodes Polη, Polκ and Rev-1, but not Polι.

Full alignment of the *C. elegans polh-1* and *polk-1* gene products with their mammalian and yeast homologs reveals their well-conserved catalytic core (figure S1). In addition, POLH-1 contains a C-terminal PIP box motif, which is essential for interaction with PCNA, and more recently, has also been shown to target the protein for degradation (Haracska *et al.* 2001; Kim and Michael 2008b). The remaining part of the C-terminus is evolutionary less conserved. Human Polη / yeast Rad30 and human Polκ contain ubiquitin binding zinc finger (UBZ) domains, mediating their interaction with PCNA

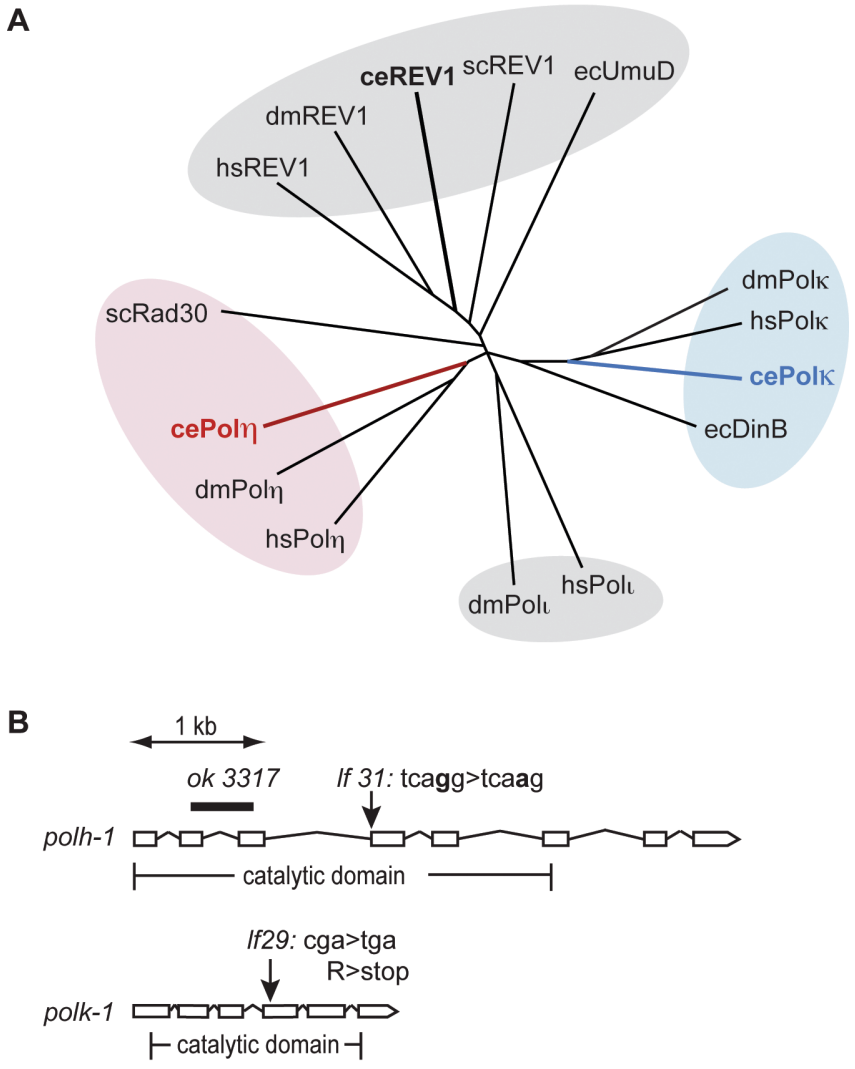


FIGURE 1. Y-family polymerases POLH-1 and POLK-1 of *C. elegans*.
A. Phylogenetic tree displaying Y-family polymerases from *C. elegans, S. cerevisiae, D. melanogaster, E. coli* and *H. sapiens*. Respectively red and blue branches show *C. elegans* POLH-1 (Polη) and POLK-1 (Polκ). **B.** Gene structure of the *C. elegans polh-1* and *polk-1* genes and the molecular nature of the alleles used in this study.

(Bienko *et al.* 2005). A UBZ domain was found in *C. elegans* POLK-1 but not in *C. elegans* POLH-1 (figure S1). Furthermore, *C. elegans* and yeast Polη and Polκ lack previously identified mammalian motifs that are mediating an interaction with the deoxycytidyl transferase Rev1 (Kosarek *et al.* 2008; Ohashi *et al.* 2009).

Using a targeted mutagenesis approach (Cuppen *et al.* 2007) we isolated mutants for *polh-1* and *polk-1* (figure 1B). *polh-1 (lf31)* has a single nucleotide substitution in the splice acceptor site of the fourth exon of the *polh-1* gene; *polk-1 (lf29)* contains a premature stop in the fourth exon and encodes a severely truncated version of POLK-1, missing at least part of the catalytic domain. In the course of this study we obtained another *polh-1* mutant from the Gene Knockout Consortium. This allele *(ok3317)* carries a deletion in *polh-1* that results in a fusion of upstream sequences of exon 2 to downstream sequences of exon 3, removing 549 coding nucleotides (figure 1B). These mutant strains were backcrossed to remove background mutations that resulted from the mutagenic treatment. No obvious abnormal phenotypes were observed for the mutant strains: the numbers of progeny of *polh-1* and *polk-1* single mutants were comparable to wildtype (data not shown); embryonic survival rates and post-embryonic development was also unaffected by the absence of POLH-1 or POLK-1. However, double mutants of *polh-1 (ok3317); polk-1 (lf29)* and of *polh-1 (lf31); polk-1 (lf29)* show a minor but significant reduction in both brood size and embryonic survival (up to five percent of the progeny died, data not shown), suggesting some level of functional redundancy in promoting fecundity.

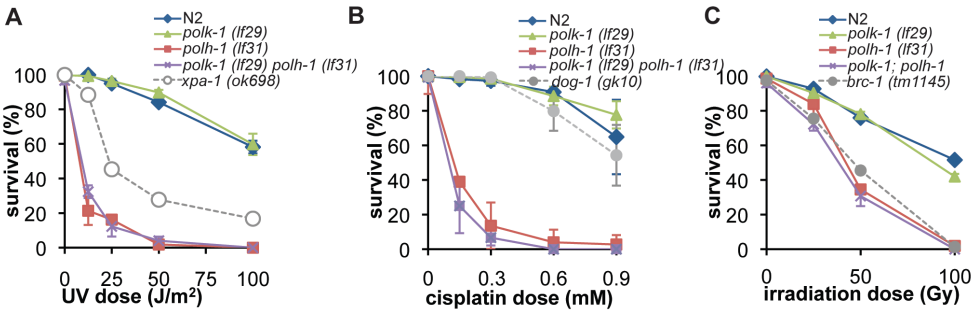


FIGURE 2. Germline sensitivity of *polh-1* and *polk-1* mutants to different sources of DNA damage.
A. Sensitivity to UV irradiation. **B.** Sensitivity to γ-irradiation. **C.** Sensitivity to cisplatin. Adults were exposed to DNA damaging treatments and survival was quantified by counting dead embryos versus living progeny in the next generation. Each line represents the mean of minimal three independent experiments. Error bars denote the s.e.m.*npp-22* knockdowns. (L) Sensitivity to MMS is not further reduced in *polh-1; polk-1* mutants by any of the tested RNAi clones. (M) Sensitivity to UV is further reduced by indicated RNAi food against *gei-17*, *ulp-1*, *npp-2* and *npp-22* in a *xpa-1* mutant background. (N) Sensitivity to a low dose of UV is not further reduced in a *polh-1* mutant background after knockdown by indicated RNAi foods.

C. elegans POLH-1 in protection against UV and cisplatin

Because UV-induced CPDs are excellent substrates for Polη-mediated TLS in yeast and mammals (Johnson *et al.* 1999b; Masutani *et al.* 1999), we tested the sensitivity of *polh-1* mutant animals to UV light by irradiating young adults and scoring progeny survival (figure 2A and figure S2A). In contrast to Polη-defective yeast and mammalian cells (Astin *et al.* 2008; McDonald *et al.* 1997), POLH-1 deficiency leads to extreme sensitivity to UV irradiation. Both *polh-1(lf31)* and *polh-1(ok3317)* mutants are more sensitive to UV than animals carrying mutations in *xpa-1,* the worm homolog of NER gene XPA, which is essential for repair of UV damage (Astin *et al.* 2008; Lans *et al.* 2010a). NER contributes to UV survival also in *polh-1* compromised conditions as animals defective in both *xpa-1* and *polh-1* are more sensitive than either of the singles. (figure S2B). In line with mammalian data, we observed that the protective role of POLH-1 is not restricted to UV-induced damage. *polh-1* worms are severely sensitized to cisplatin treatment (figure 2B and S2C). This sensitivity was even more pronounced than for the *dog-1* mutants which are defective in the homolog of the Fanconi Anemia gene FANCJ, involved in crosslink repair (Youds *et al.* 2008).

POLH-1 and XPA-1 protect against γ-irradiation induced damage in the *C. elegans* germline

Vertebrate Polη has been implicated as the polymerase responsible for extension of HR intermediates (McIlwraith *et al.* 2005; Rattray and Strathern 2005). Since HR is the predominant repair pathway in *C. elegans* for γ-irradiation-induced breaks in germ cells (Boulton *et al.* 2004; Clejan *et al.* 2006), we exposed L4 animals to γ-irradiation and scored survival of the progeny (figure 2C and figure S2D). We found that the sensitivity of *polh-1 (lf31)* and *polh-1 (ok3317)* mutants to irradiation was comparable, but not epistatic, to the sensitivity of animals that carry a mutation in *brc-1*, the worm homolog of the HR gene BRCA1 (figure 2C). Worms defective for both *polh-1* and *brc-1* are more sensitive to γ-irradiation than either of the single mutants (figure S2E-F), suggesting a *brc-1*-independent role for POLH-1 in protection against γ-irradiation. This notion is strengthened later where we will show that POLH-1 and BRC-1 protect cells against radiation at very different developmental stages.

To further test whether the sensitivity of the *polh-1* mutants to γ-irradiation is due to a possible defect in HR of DSBs, we determined the role of Polη in response to endogenously produced DSBs upon transposition. Transposition is desilenced in the germline of *rde-3* mutants (Chen *et al.* 2005) and the ensuing DSBs predominantly rely on HR for their repair (Plasterk 1991). However, embryonic lethality was not increased in *polh-1; rde-3* double mutants, in contrast to increased lethality in *brc-1; rde-3* doubles (table S2). As an independent and a direct method to address a possible *in vivo* role of *C. elegans* Polη in HR, we measured repair of a site-specific DSB using a somatic HR

reporter assay (figure 3). In this assay, that will be described in more detail elsewhere, heat shock-induced expression of the yeast endonuclease I-SceI leads to a DSB in the coding sequence of a GFP transgene that is driven by the intestinal *elt-2* promoter. This transgenic setup generates a readout for intrastrand HR, specifically in E-lineage cells, which are still proficient to enter S-phase post embryonically (in contrast to many other post embryonic cells that arrest in G1 and rely on non-homologous end-joining to repair DSBs). A functional GFP transgene is generated following DSB induction only when repair uses a downstream GFP fragment as donor sequence (figure 3A). This outcome will manifest as GFP expressing intestinal cells. While *brc-1(tm1145)* mutation resulted in a profound reduction in the number of cells that expressed GFP, *polh-1(ok3317)* mutant

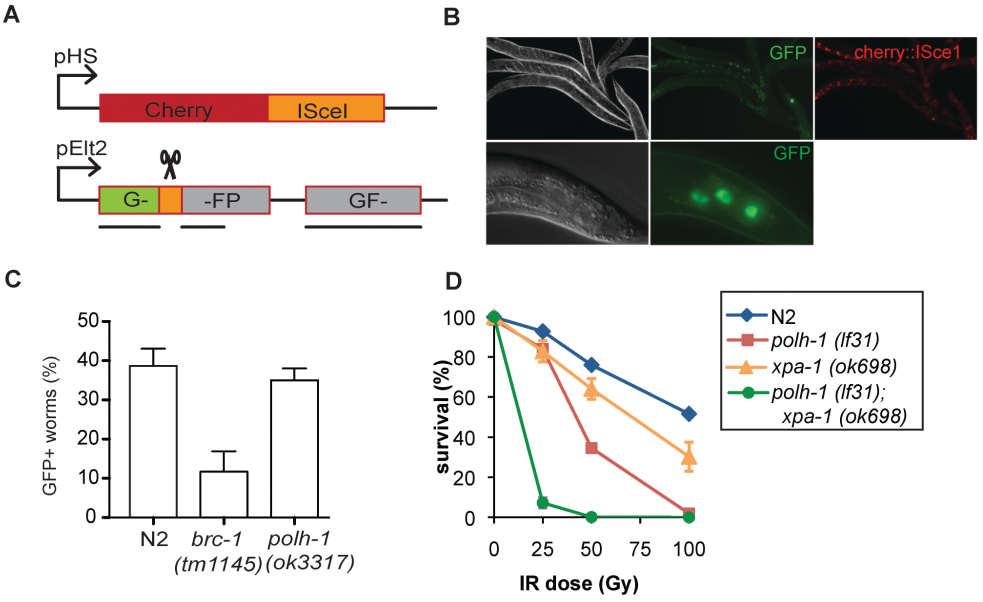


FIGURE 3. Reporter system for homologous recombination in *C. elegans*.
A. Schematic representation of the reporter transgenes. Expression of the yeast endonuclease ISceI fused to Cherry is controlled by the heat shock promotor (pHS). The reporter transgene is placed behind the intestinal *elt-2* promotor (pElt-2). Upon activation of ISceI following heat shock, the ISceI endonuclease cuts into the GFP coding sequence. Repair by gene conversion from an aborted copy of GFP results again in full length GFP. **B.** Expression of GFP and Cherry in worms containing the reporter transgenes. Upon heat shock induction all intestinal cells express cherry::ISceI. Repair of the break site by HR from an aborted GFP template results in GFP expression in some intestinal cells. **C.** Quantification of the fraction of GFP positive worms in different genetic backgrounds. Each bar represents the mean of three independent experiments. Error bars denote the s.e.m. **D.** Germline sensitivity of *polh-1 (lf31)*, *xpa-1(ok698)* and double mutants to γ-irradiation. Each line represents the mean of minimal three independent experiments. The percentage of surviving progeny is relative to the fraction of surviving progeny without any irradiation, since *polh-1(lf31)* and *xpa-1(ok698)* show about 30% synthetic lethality. Error bars denote s.e.m.

animals displayed similar numbers of cells expressing GFP with similar intensities as compared to wild type worms (figure 3B and 3C). These data further support the notion that the observed sensitivity of *polh-1* mutants to γ-irradiation is not caused by a defect in HR.

We thus explored an alternative explanation, in which the increased cytotoxicity of *polh-1* mutant animals towards γ-irradiation is the result of failed bypass of other (non-DSB) DNA lesions: apart from DSBs, γ-irradiation induces single strand breaks (SSBs) as well as 8-Oxo-dG sites and thymine glycols (Roos and Kaina 2012). We reasoned that base adducts in the DNA caused by γ-irradiation may resemble helix-distorting lesions that are substrates for NER and TLS. To address this hypothesis, we tested *xpa-1* animals as well as animals defective for both *xpa-1 xpa-1* and *polh-1* for sensitivity to γ-irradiation (figure 3D). Strikingly, a similar redundant effect of both factors is observed after exposure to γ-irradiation as seen after UV-irradiation (figure S2B). These results suggest that γ-irradiation of the germline causes replication-blocking lesions that are substrates for NER and can be bypassed by Polη. It also implies that genes previously found to be involved in γ-irradiation protection are not necessarily involved in the repair of DNA breaks (van Haaften *et al.* 2006).

Damage bypass by POLH-1 during early embryogenesis.

C. elegans polh-1 mutants are far more sensitive to various DNA damaging agents as compared to vertebrate cells. We hypothesized that the dependence on POLH-1 for damage tolerance might be specific for early embryonic development, when TLS by POLH-1 is the predominant mechanism to avoid checkpoint activation by replication fork blocks on damaged DNA (Holway *et al.* 2006). In differentiated cells, the damage response may be dominated by NER or other repair pathways. We therefore tested at which stage during development of *C. elegans* either POLH-1 mediated damage bypass or NER dominate the response to UV-irradiation. First, we exposed synchronized larvae of the L1 stage to UV light and quantified survival and growth (figure 4A and figure S3A). L1 larvae already contain 558 of the total 959 somatic cells that make up the adult animal, and thus mainly grow by cellular volume expansion as opposed to mitotic proliferation (Altun and Hall 2008). Although *xpa-1* mutants completely arrest in L1 after a low dose of UV (figure 4A, (Lans *et al.* 2010a)), in *polh-1* mutants L1 development is only slightly delayed (figure S3A). Ultimately *polh-1* mutants displayed similar survival as found for wildtype L1s following UV exposure (figure 4A), indicating that in contrast to XPA, POLH-1 plays hardly any role in the UV damage response in L1. Second, we found that germ cell maturation in *polh-1* mutants was comparable to wildtype following UV exposure (figure 4B-E), in contrast to *xpa-1* mutants that (i) display an UV-induced expansion of the pachytene region and (ii) fail to generate normal-sized oocytes (figure 4D, (Lans *et al.* 2010a)). In addition we determined the apoptotic response in the germline after UV irradiation using a *ced-1::GFP* transgene that marks germ cells in the

process of apoptosis (Schumacher *et al.* 2005b). In contrast to *xpa-1* deficient animals (Stergiou *et al.* 2007), we found no reduction in the UV-dependent apoptotic response in *polh-1* mutants as compared to wildtype animals (figure S4). Together, these data indicate that NER is essential for normal gametogenesis and L1 development following UV exposure. Apparently, in *polh-1* mutants there is sufficient time for repair of UV lesions in these developmental stages to prevent replication stress.

However, limited time for DNA repair is available immediately upon fertilization, when a *C. elegans* embryo goes through a 3 hrs period of rapid divisions, according to a fixed and time-constrained lineage program (Clejan *et al.* 2006). Thus, in this developmental stage incomplete removal of DNA damage could account for the severe embryonic lethality of UV-exposed *polh-1* mutants. To test this hypothesis, we studied the persistence of CPDs - the most abundant lesion type caused by UV – in pronuclei of oocytes, just after fertilization. We irradiated adults with 200 J/m^2 and stained after 24 hours developing embryos for CPDs. Remarkably, in wildtype embryos CPDs were still present in the paternal pronucleus, while no CPDs were observed in the maternal pronucleus (figure 4F-G). We next assayed *xpa-1* and *polh-1* mutants after a dose of 50 J/m^2 (leading to comparable levels of embryonic lethality). Mutants defective in *xpa-1* displayed CPD staining in both pronuclei (figure 4H), suggesting that in wildtype animals NER-dependent removal of CPDs has occurred during meiotic maturation of the germ cells. In contrast to *xpa-1* mutants, but similar to wildtype animals, *polh-1* mutants were proficient in removal of CPDs from the maternal pronucleus, whereas CPDs were clearly detectable in the paternal pronucleus (figure 4I). Before migration and fusion with the maternal pronucleus, the paternal genome decondenses and is replicated in less than 12 minutes (Edgar and McGhee 1988). This time span is insufficient for NER to remove DNA damage. We hypothesize that the presence of unrepaired damage from the paternal DNA poses a problem on the first mitotic divisions in *polh-1* early embryos. To address this hypothesis, we mated UV-irradiated wildtype or mutant hermaphrodites with untreated males, providing a source of undamaged sperm DNA (figure 4J). To mark the progeny we used a transgenic line expressing Pmyo-2::GFP. Indeed, lethality in the progeny of irradiated *polh-1* hermaphrodites is almost fully rescued by providing a source of undamaged sperm DNA. In contrast, mating of *xpa-1* hermaphrodites with untreated males does not affect survival of the progeny. Together, these data indicate that correct progression of early embryonic cell divisions strongly relies on POLH-1 when the genome contains DNA damage. This dependency is not restricted to UV-induced damage but also extends to DNA damage induced by γ-irradiation: the increased sensitivity of *polh-1* mutants to γ-irradiation can be completely rescued by crossing irradiated hermaphrodites with untreated males, thus providing a non-damaged paternal genome (figure S5). Importantly this is in stark contrast to the sensitivity of *brc-1* mutants, that cannot be rescued by providing non-damaged sperm. This developmental separation of the modes of action of these proteins further substantiates our findings that *polh-1* and *brc-1* act independently in protecting cells against γ-irradiation induced-DNA damage.

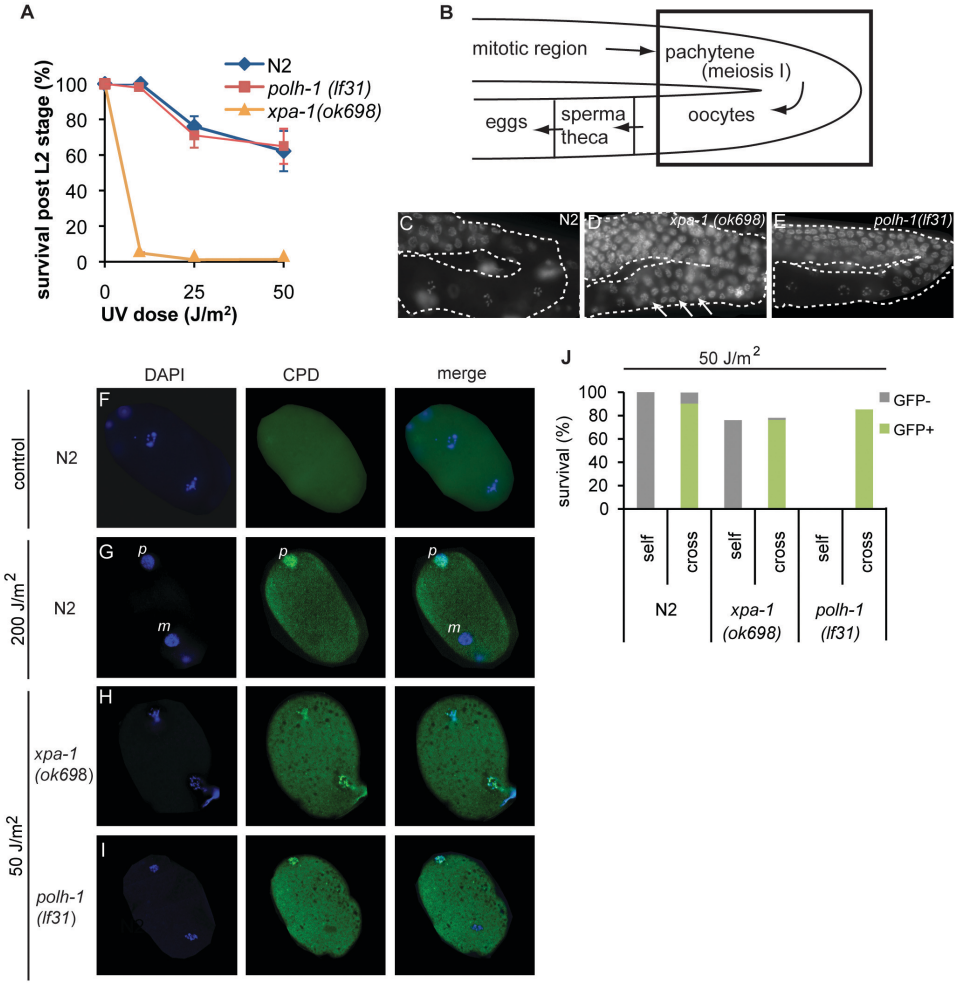


FIGURE 4. ***C. elegans polh-1*** **and** ***xpa-1*** **adults and embryos exposed to UV at different stages during development.**

A. Survival of larvae irradiated at L1 stage. Each line is the mean of three independent experiments; error bars denote s.e.m. **B.** Schematic overview of the *C. elegans* germline. Boxed area shows the transition in the germline bend from the pachytene to maturating oocytes displayed in pictures C-D. **C-E.** DAPI stainings of germlines of indicated genotype 16 hrs after exposure to 120 J/m^2 UV. Morphology of the germline is completely disrupted in *xpa-1* mutants (D) but not in wildtype (C) or *polh-1* worms (E). Oocyte maturation in irradiated *polh-1* mutants is normal (E), while most meiotic cells fail to progress into oocytes in *xpa-1* worms after UV-irradiation (D), causing expansion of the pachytene region through the germline bend (arrows). **F-I.** Presence of CPDs during the first embryonic divisions. Immunofluorescence on just fertilized N2 embryos 24 hrs after treatment with 200 J/m^2 shows that only one of the two pronuclei carries CPDs (F and G). In UV irradiated *xpa-1* embryos both pronuclei carry CPDs (H) while in *polh-1* embryos, similar to wildtype (G), only one pronucleus contains CPDs (I). N.B. a lower UV dose was used in H and I to compare doses that induced similar levels of lethality. **J.** UV-irradiated hermaphrodites were crossed with untreated males carrying a Pmyo-2::GFP transgene. UV-induced lethality is partly rescued in the cross progeny of *polh-1*, but not *xpa-1* hermaphrodites.

POLH-1 and POLK-1 act in a redundant fashion in protection against the methylating agent MMS

We next wondered whether a similar developmentally restrained function could be attributed to TLS polymerase POLK-1 To address this question, we exposed *polk-1(lf29)* mutant worms to different doses of UV, cisplatin or γ-irradiation (figure 2A-C), but found no difference in sensitivity as compared to wildtype animals, indicating that POLK-1 is not involved in protection against these sources of DNA damage in *C. elegans*. However, akin to the outcome of published RNAi experiments (Holway *et al.* 2006), *polk-1* and *polh-1* mutants both appeared to be sensitive to chronic exposure to the alkylating agent MMS, albeit that sensitivity in *polh-1* mutants was much more pronounced (figure 5A, figure S6) indicating that both POLH-1 and POLK-1 play a role in bypass of MMS-induced damage. We assayed *polh-1 (lf31); polk-1 (lf29)* double mutants and *polh-1 (ok3317) polk-1* RNAi fed animals for MMS sensitivity (figure 5A, figure S6A). Interestingly, double mutants are extremely sensitive to MMS, and complete lethality was observed at a dose that was 100 times lower than the effective dose for any of the single mutants (figure 5A, figure S6A). We did not observe any synergistic effect for any of the other types of lesions we tested (figure 2)

POLH-1 has previously been shown to be involved in avoiding checkpoint activation by replicating damaged DNA (Holway *et al.* 2006). In *C. elegans* embryogenesis, checkpoints - mediated by the *C. elegans* homologs of the checkpoint genes ATR and CHK-1 - are used to time the first asynchronous cell divisions that are essential for embryonic patterning and embryonic viability (Brauchle *et al.* 2003b). Checkpoint activation due to DNA damage interferes with the developmental role of the checkpoint, causing patterning defects and embryonic lethality. Our results with null mutants for *polh-1* and *polk-1* suggest that both POLH-1 and POLK-1 can act to avoid checkpoint activation. To test the involvement of POLK-1 in checkpoint avoidance directly, we timed the first embryonic division of *polk-1* embryos after exposure to MMS. Figure 5B illustrates a delayed first embryonic division in *polk-1* mutants when compared to wildtype embryos. Importantly, we also observed examples of *polk-1* embryos that after MMS treatment fully arrested at the 1-cell stage (movies S1 and S2), while we never observed such cases for MMS-treated wildtype embryos. Two other phenotypes are also indicative of replication stress during early embryonic divisions of MMS treated *polh-1* and *polk-1* mutant animals: first, *polh-1; polk-1* double mutant embryos displayed foci of the DSB repair marker RAD51 (figure 5C-J), indicative of DSBs resulting from trying to replicate damaged genomes (Holway *et al.* 2005). Second, DAPI stainings revealed chromatin bridges and a disrupted nuclear morphology in the early embryo (figure 5J), suggesting division of incompletely replicated or disentangled genomes. These phenotypes were less profound, but noticeable, in both single mutants, while never observed in wild type embryos at these MMS concentrations (figure 5C-H).

To investigate whether the dependency on POLH-1 and POLK-1 for tolerance to MMS

was restricted to embryogenesis - similar to the requirement of POLH-1 in UV tolerance - we followed the outgrowth of L1 animals exposed to different concentrations of MMS (figure S6B). *polh-1* animals were mildly affected in development, while we did not observe any effect for *polk-1*. As for UV, NER deficient *xpa-1* larvae are profoundly more sensitive to MMS than either *polk-1* or *polh-1* deficient larvae (figure S6B), while the opposite is true for embryonic stages: *xpa-1* embryos are less sensitive to MMS than *polh-1* embryos (Holway *et al.* 2006). This again argues that TLS outranks NER at developmental stages that require fast replication cycles.

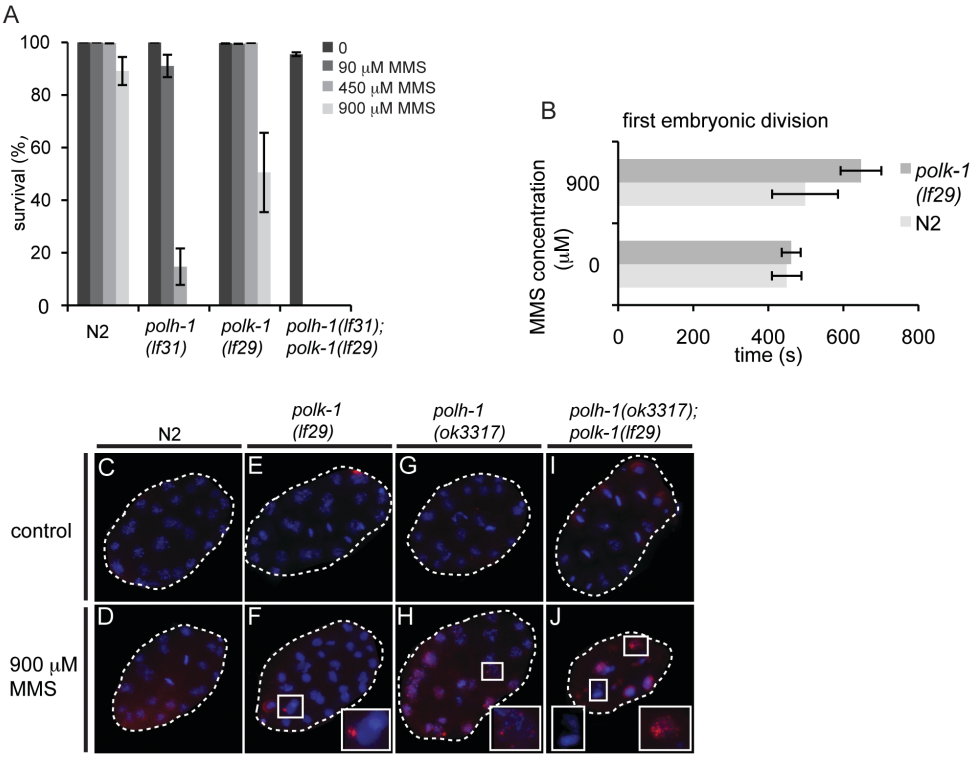


FIGURE 5. *polh-1* and *polk-1* protect in a redundant fashion against the methylating agent MMS
A. Double mutants of *polh-1* and *polk-1* are severely sensitized to MMS exposure. Results of a representative experiment are shown. Error bars denote SD. **B.** Delayed progression through the first embryonic division after MMS exposure in *polk-1* mutants. The interval between passing of the paternal pronucleus over the midline till the start of cytokinesis is timed for at least 5 embryos per datapoint. Statistical significance for the difference in delay between N2 and *polk-1* embryos treated with MMS was calculated with a student's t-test (p=0.012) **C-J.** RAD51 immunostainings of early embryos treated with MMS. Morphology of *polh-1; polk-1* double mutant embryos is abnormal after MMS exposure, displaying chromatin bridges and abundant RAD51 staining (J). *polh-1(ok3317)* and *polk-1(lf29)* single mutants show incidental RAD51 foci in embryos (F and H), while such foci were never observed in untreated controls (C, E, G and I).

A genetic approach for identifying new factors in TLS regulation in the early embryo

Since POLH-1 and POLK-1 together appear to be extremely important in protecting the developing embryo against MMS, we wondered whether there might be a general pathway underlying the regulation of the two TLS enzymes. To identify new factors regulating TLS in the early embryo, we performed a genome wide RNAi screen for genes sensitizing for MMS. Out of 16,757 genes tested (covering ~86% of all predicted *C. elegans* genes), we found 87 genes resulted in sensitivity to MMS upon knockdown including *polk-1*. *polh-1* was not picked up, probably due to insufficient knockdown of the RNAi clone targeting this gene. We next inspected these RNAi knockdowns for phenotypes reminiscent to *polh-1;polk-1* double mutants. All 87 hits were analysed by DAPI for altered nuclear morphology after exposure to MMS (figure 5A and S7). Four clones were selected for follow-up analysis based on perturbed embryonic divisions as indicated by chromatin bridges and malformed nuclei. These clones targeted the genes: *gei-17, ulp-1, npp-2* and *npp-22* (figure 6). *gei-17* encodes a SUMO-protease that was previously shown to interact with POLH-1 after DNA damage (Kim and Michael 2008b). *ulp-1* encodes a ubiquitin-like protease (ULP) which deconjugates SUMO moieties from their target proteins (Zhang *et al.* 2004). *npp-2* and *npp-22* encode two components of the *C. elegans* nuclear pore complex (NPC)(Galy *et al.* 2003). Null alleles of *gei-17*, *npp-2* and *npp-22* result in embryonic lethality. Knockdowns of the four genes reduced tolerance to MMS to a similar extent as mutations in *polh-1* and *polk-1* (figure 6K). In line with published data (Holway *et al.* 2006), *gei-17* knockdown lead to abundant RAD51 foci in embryos treated on MMS, indicative of replication stress. Also knockdown of *ulp-1*, *npp-2* and *npp-22* lead to MMS-induced RAD51 foci, although to a lesser extend than *gei-17* knockdown, which is expected from the fact that these knockdowns also display less dramatic effects on progeny survival. Foci formation was never observed in mock-treated knockdowns or wild type controls (figure 6A-J).

We hypothesized that if these genes would be in a common pathway with POLH-1 and POLK-1, then knockdown of these factors would not further increase sensitivity of *polh-1;polk-1* worms to a low dose of MMS. Indeed, MMS sensitivity is not further increased when *ulp-1*,*gei-17, npp-2* and *npp-22* are knocked down in *polh-1;polk-1* double mutant animals (figure 6L), placing all four factors in an epistatic relation to the TLS genes in the response to alkylating damage. To substantiate this epistatic relationship we also studied another source of DNA damage infliction by exposing young adults to UV light. We previously showed that *polh-1* is important for embryonic development in the presence of UV damage, and that an additional mutation in the NER factor *xpa-1* renders the animals even more sensitive to low doses of UV (figures 2A and S2B). We argued that if these factors act in a common pathway with Polη in the response to UV, we would expect their knockdowns to be epistatic with a *polh-1* mutation, but increase the sensitivity of *xpa-1* defective animals. Indeed, knockdowns of *gei-17, ulp-1, npp-2*

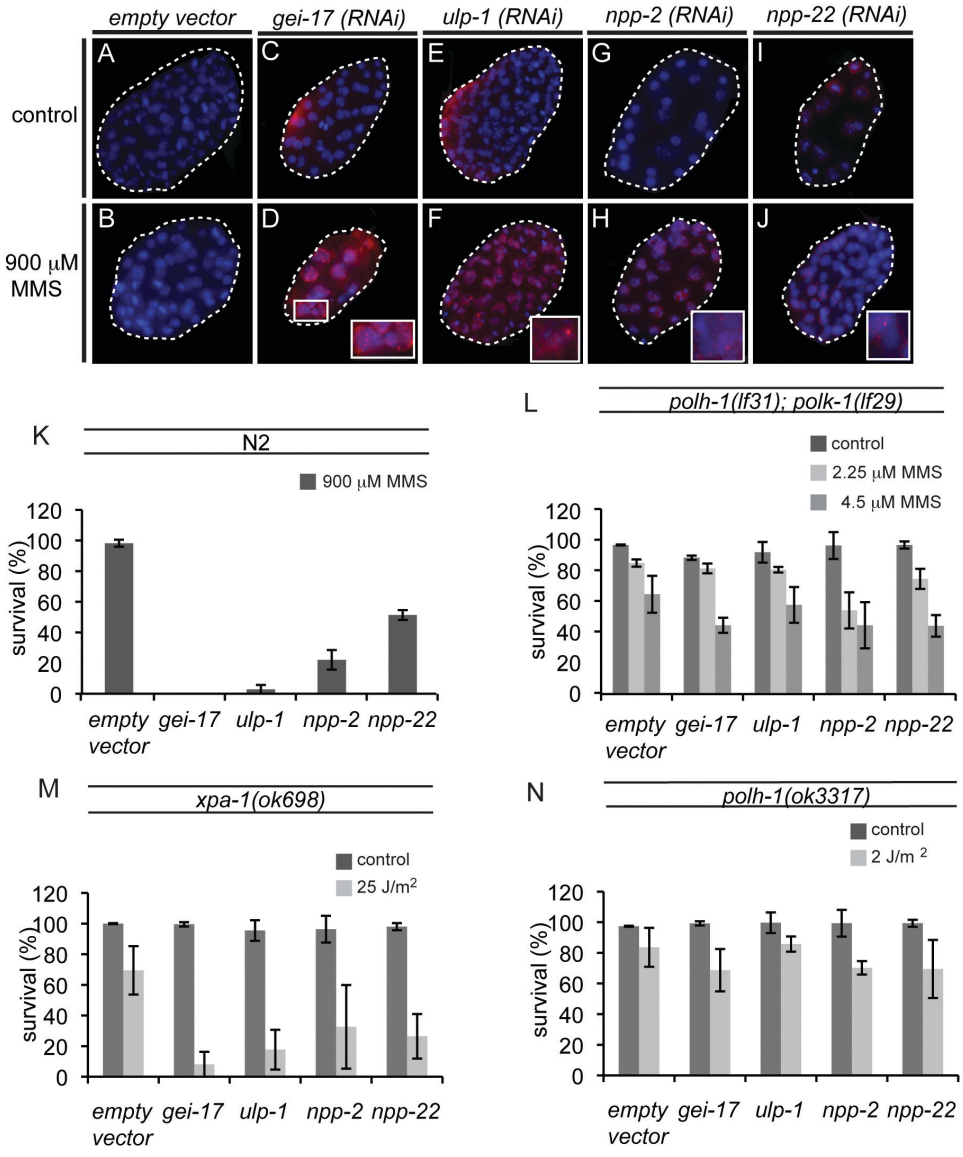


FIGURE 6. New factors in damage tolerance in the early embryo.
A-J. RAD51 stainings of early embryos treated with MMS. *gei-17* RNAi knockdown embryos show abundant RAD51 staining after treatment with MMS (D). Incidental RAD51 foci are observed in *ulp-1*, *npp-2 and npp-22* knockdown embryos (F, H, J) but not in wildtype controls (B). **K.** MMS sensitivity of *gei-17*, *ulp-1*, *npp-2* and *npp-22* knockdowns. **L.** Sensitivity to MMS is not further reduced in *polh-1; polk-1* mutants by any of the tested RNAi clones. **M.** Sensitivity to UV is further reduced by indicated RNAi food against *gei-17*, *ulp-1*, *npp-2* and *npp-22* in a *xpa-1* mutant background. **N.** Sensitivity to a low dose of UV is not further reduced in a *polh-1* mutant background after knockdown by indicated RNAi foods.

and *npp-22* all further increase the sensitivity of *xpa-1* mutant animals to a low dose of UV (figure 6M), but did not change sensitivity for *polh-1* mutants (figure 6N). Together these data indicate that *ulp-1, npp-2* and *npp-22* are novel factors in TLS mediated by POLH-1 and POLK-1 in response to DNA damage during early embryogenesis in *C. elegans*. The SUMO protease gene *gei-17* has previously been shown to promote damage tolerance by sumoylating POLH-1 (Holway *et al.* 2006; Kim and Michael 2008a); we here show that GEI-17 may be implicated in TLS mediated by both POLH-1 and POLK-1.

DISCUSSION

Here we showed that there is modulation of the choice between repair and bypass of damaged template DNA in a developing organism. *A priori* one would expect error-free repair by NER to be the favoured option in germ cells to prevent the accumulation of mutations in subsequent generations. Indeed, we and others found that both for germ cell maturation and post-embryonic somatic development, NER is indispensable in response to specific DNA damages (Astin *et al.* 2008; Lans *et al.* 2010b; Stergiou *et al.* 2007) . However, and in line with previously published data (Holway *et al.* 2006), we found that immediately after fertilization of the oocyte, during stages of rapid cell divisions in the early embryo, survival is determined by TLS factors and not by NER. The need for efficient TLS must be viewed in light of the strict timing of the developmental program, which likely does not allow to "wait" for repair processes to complete. Our observation that wild type animals can easily survive UV doses up to 50J/m^2 without substantially repairing CPDs from their sperm or decondensed paternal pronucleus implicates that TLS-proficient zygotes can replicate a damaged genome containing 10-50.000 UV lesions in less than 12 minutes - the time it takes for the paternal nucleus to double its genome content - without delaying cell division (Edgar and McGhee 1988).
We found that *C. elegans* POLH-1 has a broader substrate specificity than POLK-1; POLH-1 is involved in bypass of damaged induced upon exposure to UV light, γ-irradiation, cisplatin and MMS. We considered the option that all treatments may lead to one similar substrate that may cause the sensitivity: i.e. DSBs brought about by replication fork obstruction and collapse. This notion was fuelled by studies in vertebrates, in which Polη was suggested to act in HR repair of DSBs by extending the D-loop intermediate structure (Kawamoto *et al.* 2005; McIlwraith *et al.* 2005). However, we observed a wild type response to either transposon-mediated or ISceI-induced DSBs, thus arguing against a role for POLH-1 in DSB repair, in both germline or somatic tissue of *C. elegans*. We ascribe the sensitivity of *polh-1* mutants towards γ-irradiation to the induction of other non-DSB lesions, which may be NER substrates as we observed a synergistic relationship between *xpa-1* and *polh-1* on IR sensitivity. The induction of free radicals by ionizing radiation causes a plethora of lesions in the DNA, such as 8-oxo-dG sites,

which may require Polη-mediated bypass to prevent the induction of toxic DSBs (Lee and Pfeifer 2008).

An explanation for the broad substrate specificity of POLH-1 may reside in the flexible active site of POLH-1 which may allow for bypass of lesions that are structurally very different. Indeed, studies in chicken DT40 cells indicate that Polη is a much more versatile polymerase than the phenotype of XPV cells would suggest (Hirota *et al.* 2010). Alternatively, Polη could have an indirect role by recruiting other TLS proteins to the damage site. In human cells Rev1 is recruited to UV damages *via* an interaction with Polη (Akagi *et al.* 2009). Interestingly, Polκ has been shown to serve as a 'backup' polymerase in XPV cells in the bypass of both UV-induced CPDs as well as cisplatin adducts (Shachar *et al.* 2009; Ziv *et al.* 2009). We here show that in nematodes this genetic interaction is restricted to damage induced by the S_N1 methylating agent MMS. Its molecular effects include the formation of N-7 methylguanine (which by spontaneous apurination can lead to an abasic site), N3- methyladenine, N3-cytosine and O^6-methylguanine (Fu *et al.* 2012). Although we cannot deduce from our in vivo analysis which of these damages underlies the cytotoxicity observed in nematodes, all of these base damages have residual coding capacity, and are less disturbing than some of the DNA lesions induced by cisplatin or IR treatment. This notion may explain the redundant role of the functionally more restricted POLK-1 on MMS, while no contribution was seen on UV, IR or cisplatin.

In order to find novel factors that are directly or indirectly involved in TLS, we screened for RNAi knockdowns that rendered cells sensitive to MMS and UV, only in the context of TLS functionality. Out of ~ 17,000 clones we identified four genes whose knockdown sensitized wildtype but not TLS-deficient animals to MMS treatment. One of these genes, *gei-17*, was previously identified to regulate Polη: GEI-17 has recently been shown to sumoylate POLH-1 through its PIP-box motif resulting in protection of the protein from degradation (Kim and Michael 2008b). The profound effects of *gei-17* RNAi on cellular tolerance to MMS suggest that this SUMO-ligase most likely acts on both POLH-1 and POLK-1 (figure 5A and I); possibly *C. elegans* POLK-1 also contains a PIP-box motif (figure S1). In addition to GEI-17 we also identified the SUMO protease ULP-1 as a factor in TLS-mediated MMS- and UV-sensitivity. This result suggests that, apart from sumoylation, also desumoylation may play a role in the regulation of TLS proteins. Future work aimed to identify direct targets of ULP-1 needs to establish whether its role is direct, by desumoylation of the TLS polymerases, or indirect. Ubiquitin-like proteases (ULPs) deconjugate SUMO from their target proteins and therefore the damage sensitivity of ULP-1 knockdown may also be explained by disturbed regulation because of a shortage of SUMO. SUMO proteases and ligases may anchor to the NPCs in order to sumoylated or desumoylate their targets (Palancade and Doye 2008). Here we show that, similar to *gei-17* and *ulp-1*, RNAi against nuclear pore components *npp-2* and *npp-22* is compatible with viability and results in sensitivity to UV lesions and MMS. However, sensitivity is not further increased in the absence of POLH-1 and POLK-1. This finding implicates a

role for the NPC (or NPC subunits) in TLS mediated damage tolerance, possibly in the localization of SUMO-regulation. *npp-2* encodes the *C. elegans* homolog of yeast Nup85, which is one of the proteins of the scNup84 scaffolding complex. In yeast, mutants in the Nup84 and Nup60/Mlp1-2 complexes have similar phenotypes in the response to DNA damage as Ulp1 mutants (Palancade *et al.* 2007). A direct link of NPCs to the DNA damage response was also suggested by Nagai et al., who showed relocation of damaged DNA to nuclear pores (Nagai *et al.* 2008) and recently by Bermejo et al. who showed involvement of inner basket proteins in fork stability (Bermejo *et al.* 2011). *npp-22* encodes the *C. elegans* homolog of yeast or mammalian NDC1, which is crucial for nuclear pore assembly (Stavru *et al.* 2006). Future work on *gei-17*, *ulp-1* and the nuclear pore components *npp-2* and *npp-22* is needed to substantiate the role of sumoylation and desumoylation processes and a possible link to the NPC (subunits) in regulating TLS.

MATERIALS AND METHODS

C. elegans genetics

All strains were cultured according to standard methods (Brenner 1974). Wildtype N2 (Bristol) worms were used in all control experiments. The *polh-1 (lf31)* and *polk-1 (lf29)* mutants were isolated in our own laboratory. *polh-1 (ok3317)* worms, that were kindly provided by Joel Meyer (Duke University, Durham NC, USA), have been generated by the *C. elegans* knock-out consortium. *BCN2081*, carrying a single copy integrated Pmyo::GFP transgene, was provided by Ben Lehner (EMBL Centre for Genomic Regulation, Barcelona, Spain)(Semple *et al.* 2010). All other alleles *(xpa-1 (ok698)*; *rde-3 (ne298); brc-1 (tm1145*); *dog-1 (gk10))* and the transgenic line MD701 *(bcIs39[P(lim-7)ced-1::GFP + lin-15(+)])* were obtained from the *C. elegans* Genetics Center (St Paul, MN, USA). All mutant strains were backcrossed six times before performing experiments. Newly generated strains are listed in table S1 in the supplementary information.

Survival assays

Staged animals were exposed to different doses of various DNA damaging agents. To assess germline sensitivity three plates with three worms were allowed to lay eggs for 24 - 48 hrs per experimental condition. 24 hrs later, the number of unhatched eggs and the number of surviving progeny was determined. All experiments were performed in triplicate. To measure germline sensitivity to UV, staged young adults one day post L4 were transferred to empty NGM plates and exposed to different doses of UV-C (predominantly 254 nm, Philips). Animals were placed on fresh OP50 plates and allowed to lay eggs for 32 hrs.

To determine whether lethality could be rescued by the supply of undamaged sperm, UV irradiated hermaphrodites were mated with untreated BCN2081 worms, which have Pmyo-2::GFP transgenes integrated in their genomes. After 24 hrs of male contact, the hermaphrodites were transferred to individual plates and allowed to lay eggs for 24 hrs. The mother was subsequently removed and 24 hrs later the number of unhatched eggs and the number of GFP+ and GFP- progeny was determined.
The sensitivity of L1 larval stage animals to UV was measured as described previously (Lans *et al.* 2010a). L1s were synchronized by bleaching, and exposed to UV-C on empty NGM plates. Per plate, at least 100 L1 animals were counted. For three subsequent days the development of L1-treated animals was monitored.
To measure germline sensitivity to γ-irradiation, different doses were delivered by an X-ray generator (dose rate 7 Gy/min; YXLON International) to L4 animals. Animals were allowed to lay eggs for 48 hrs, and scored 24 hrs later for hatching.
Sensitivity to cisplatin was determined by incubating staged L4 animals for 3 hrs in M9 containing different concentrations of cisplatin (Sigma-Aldrich). After 1 hr recovery on OP50 plates, animals were placed on fresh OP50 plates and allowed to lay eggs for 48 hrs. The mother was removed and the survival of the progeny was scored 24 hrs later.
To measure sensitivity to chronic exposure to MMS, staged L4 animals were placed for 24 hrs on NGM plates containing different concentrations of MMS (Sigma-Aldrich). After 24 hrs, the number of non-hatched eggs and surviving progeny was determined.

Homologous recombination (HR) assay

A HR reporter plasmid was constructed consisting of a GFP/LacZ fusion under the control of the intestinal specific *elt-2* promotor (Fukushige *et al.* 1999). An ISceI recognition sequence was inserted that disrupted the GFP ORF. To provide a template for homologous recombination, part of the GFP coding region was PCR amplified and inserted downstream of the disrupted GFP/LacZ locus. The ISceI expressing plasmid pRP3001 (hsp-16.41::ISceI ORF) (Pontier and Tijsterman 2009), was modified to include the mCherry ORF leading to a functional ISceI/mCherry protein to visualize and monitor the expression of the ISceI endonuclease. A detailed description of the reporter system and its validation will be published elsewhere.
For reading out HR, synchronized L4 animals were transferred and incubated for 1.5 hrs at 34°C. After 24 hrs, GFP expression in the intestine was analyzed on a Leica DM6000 microscope.

Microscopy

Nuclear stainings on germlines and embryos were performed by incubation of staged young adults for 10 minutes in ethanol containing 10 μg/mL 4',6-diamidino-2-phenylindole (DAPI). After two washes with PBS, worms were mounted on object slides

in 30% glycerol.
To detect CPDs, eggs were liberated from UV-irradiated worms and fixed with 3% paraformaldehyde. Fixed eggs were permeabilized by freeze cracking and subsequently washed with 1% Triton and methanol (-20°C). CPDs were visualized by subsequent staining with an anti-CPD mouse monoclonal antibody and an Alexa488-labelled goat-anti-mouse secondary antibody (Molecular Probes Inc) combined with 10 μg/mL DAPI. Dissected worms and eggs were mounted on object slides in Vectashield®.
To study RAD51 foci formation, a similar procedure as described for CPD staining was followed. Fixed eggs were permeabilized by freeze cracking and subsequently washed with 1% Triton and methanol (-20°C). RAD51 was visualized by subsequent staining with an anti-RAD51 rabbit monoclonal antibody and an Alexa488-labelled goat-anti-rabbit secondary antibody (Molecular Probes Inc) combined with 10 μg/mL DAPI. Dissected worms and eggs were mounted on object slides in Vectashield®.
For the analysis of apoptosis, transgenic MD701 animals, expressing a *ced1*::GFP fusion behind a *lim-7* promotor, were used to visualize sheath cells surrounding apoptotic germ cells (Schumacher *et al.* 2005a; Schumacher *et al.* 2005b). All microscopy was performed with a Leica DM6000 microscope.

RNAi screen and RNAi experiments

Using the Ahringer Lab RNAi feeding library a genome-wide screen was performed for clones sensitizing animals to MMS. The procedure is an adaptation from a genome-wide RNAi screen for radiation sensitivity by Van Haaften et al, described in detail in their supplementary data (van Haaften *et al.* 2006). Briefly, L1 worms were grown to L4s in liquid on RNAi food. At the L4 stage MMS was added to a concentration of 0.01%. After three days survival of the progeny was scored by visual inspection. For knockdown of *polk-1, gei-17, ulp-1, npp-2* and *npp-22* genes, individual Ahringer clones were grown on IPTG containing NGM plates. Staged L4s were transferred to RNAi plates; analysis was performed on the progeny of these animals.

ACKNOWLEDGEMENTS

Several deletion alleles were provided by the Caenorhabditis Genetics Center (CGC). The *polh-1(ok3317)* allele has been kindly provided by Joel Meyer. We thank members of the Tijsterman-lab for discussions and Jane van Heteren for helpful comments on this manuscript. M.T. is supported by a European Research Council Starting Grant (203379, 'DSBrepair').

REFERENCES

Akagi, J.-I., C. Masutani, Y. Kataoka, T. Kan, E. Ohashi *et al.*, 2009 Interaction with DNA polymerase eta is required for nuclear accumulation of REV1 and suppression of spontaneous mutations in human cells. DNA Repair **8:** 585-599.

Altun, Z. F., and D. H. Hall, 2008 Introduction in *WormAtlas*.

Andersen, P. L., F. Xu and W. Xiao, 2008 Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA, pp. 162-173 in *Cell Res*.

Astin, J. W., N. J. O'Neil and P. E. Kuwabara, 2008 Nucleotide excision repair and the degradation of RNA pol II by the Caenorhabditis elegans XPA and Rsp5 orthologues, RAD-3 and WWP-1. DNA Repair **7:** 267-280.

Bermejo, R., T. Capra, R. Jossen, A. Colosio, C. Frattini *et al.*, 2011 The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. Cell **146:** 233-246.

Bienko, M., C. M. Green, N. Crosetto, F. Rudolf, G. Zapart *et al.*, 2005 Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis, pp. 1821-1824 in *Science*.

Boulton, S. J., J. S. Martin, J. Polanowska, D. E. Hill, A. Gartner *et al.*, 2004 BRCA1/BARD1 orthologs required for DNA repair in Caenorhabditis elegans, pp. 33-39 in *Curr Biol*.

Brauchle, M., K. Baumer and P. Gönczy, 2003a Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in C. elegans embryos, pp. 819-827 in *Curr Biol*.

Brauchle, M., K. Baumer and P. Gönczy, 2003b Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in C. elegans embryos. Current biology : CB **13:** 819-827.

Brenner, S., 1974 The genetics of Caenorhabditis elegans, pp. 71-94 in *Genetics*.

Carlson, K. D., R. E. Johnson, L. Prakash, S. Prakash and M. T. Washington, 2006 Human DNA polymerase kappa forms nonproductive complexes with matched primer termini but not with mismatched primer termini, pp. 15776-15781 in *Proc Natl Acad Sci USA*.

Chen, C.-C. G., M. J. Simard, H. Tabara, D. R. Brownell, J. A. McCollough *et al.*, 2005 A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Current biology : CB **15:** 378-383.

Ciccia, A., and S. J. Elledge, 2010 The DNA damage response: making it safe to play with knives, pp. 179-204 in *Mol Cell*.

Clejan, I., J. Boerckel and S. Ahmed, 2006 Developmental modulation of nonhomologous end joining in Caenorhabditis elegans, pp. 1301-1317 in *Genetics*.

Cuppen, E., E. Gort, E. Hazendonk, J. Mudde, J. van de Belt *et al.*, 2007 Efficient target-selected mutagenesis in Caenorhabditis elegans: toward a knockout for every gene, pp. 649-658 in *Genome Res*.

Edgar, L. G., and J. D. McGhee, 1988 DNA synthesis and the control of embryonic gene expression in C. elegans. Cell **53:** 589-599.

Encalada, S. E., P. R. Martin, J. B. Phillips, R. Lyczak, D. R. Hamill *et al.*, 2000 DNA replication defects delay cell division and disrupt cell polarity in early Caenorhabditis elegans embryos, pp. 225-238 in *Developmental Biology*.

Fu, D., J. A. Calvo and L. D. Samson, 2012 Balancing repair and tolerance of DNA damage caused by alkylating agents. Nature Reviews Cancer **12:** 104-120.

Fukushige, T., M. J. Hendzel, D. P. Bazett-Jones and J. D. McGhee, 1999 Direct visualization

of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo, pp. 11883-11888 in *Proc Natl Acad Sci USA*.

Galy, V., I. W. Mattaj and P. Askjaer, 2003 Caenorhabditis elegans nucleoporins Nup93 and Nup205 determine the limit of nuclear pore complex size exclusion in vivo. Molecular biology of the cell **14:** 5104-5115.

Haracska, L., C. M. Kondratick, I. Unk, S. Prakash and L. Prakash, 2001 Interaction with PCNA is essential for yeast DNA polymerase eta function, pp. 407-415 in *Mol Cell*.

Haracska, L., L. Prakash and S. Prakash, 2002 Role of human DNA polymerase kappa as an extender in translesion synthesis, pp. 16000-16005 in *Proc Natl Acad Sci USA*.

Hirota, K., E. Sonoda, T. Kawamoto, A. Motegi, C. Masutani *et al.*, 2010 Simultaneous disruption of two DNA polymerases, Polη and Polζ, in Avian DT40 cells unmasks the role of Polη in cellular response to various DNA lesions. PLoS Genetics **6**.

Holway, A. H., C. Hung and W. M. Michael, 2005 Systematic, RNA-interference-mediated identification of mus-101 modifier genes in Caenorhabditis elegans, pp. 1451-1460 in *Genetics*.

Holway, A. H., S.-H. Kim, A. La Volpe and W. M. Michael, 2006 Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos. The Journal of Cell Biology **172:** 999-1008.

Johnson, R. E., C. M. Kondratick, S. Prakash and L. Prakash, 1999a hRAD30 mutations in the variant form of xeroderma pigmentosum, pp. 263-265 in *Science*.

Johnson, R. E., S. Prakash and L. Prakash, 1999b Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta, pp. 1001-1004 in *Science*.

Kawamoto, T., K. Araki, E. Sonoda, Y. M. Yamashita, K. Harada *et al.*, 2005 Dual roles for DNA polymerase eta in homologous DNA recombination and translesion DNA synthesis, pp. 793-799 in *Mol Cell*.

Kim, S.-H., and W. M. Michael, 2008a Regulated proteolysis of DNA polymerase eta during the DNA-damage response in C. elegans. Molecular cell **32:** 757-766.

Kim, S.-H., and W. M. Michael, 2008b Regulated proteolysis of DNA polymerase eta during the DNA-damage response in C. elegans, pp. 757-766 in *Mol Cell*.

Kosarek, J. N., R. V. Woodruff, A. Rivera-Begeman, C. Guo, S. D'Souza *et al.*, 2008 Comparative analysis of in vivo interactions between Rev1 protein and other Y-family DNA polymerases in animals and yeasts. DNA Repair **7:** 439-451.

Lans, H., J. A. Marteijn, B. Schumacher, J. H. J. Hoeijmakers, G. Jansen *et al.*, 2010a Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development, pp. e1000941 in *PLoS Genet*.

Lans, H., J. A. Marteijn, B. Schumacher, J. H. J. Hoeijmakers, G. Jansen *et al.*, 2010b Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development. PLoS Genetics **6:** e1000941.

Lee, D.-H., and G. P. Pfeifer, 2008 Translesion synthesis of 7,8-dihydro-8-oxo-2'-deoxyguanosine by DNA polymerase eta in vivo. Mutation research **641:** 19-26.

Masutani, C., M. Araki, A. Yamada, R. Kusumoto, T. Nogimori *et al.*, 1999 Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity, pp. 3491-3501 in *EMBO J*.

McDonald, J. P., A. S. Levine and R. Woodgate, 1997 The Saccharomyces cerevisiae RAD30 gene, a homologue of Escherichia coli dinB and umuC, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism, pp. 1557-1568 in *Genetics*.

McIlwraith, M. J., M. J. McIlwraith, A. Vaisman, Y. Liu, E. Fanning *et al.*, 2005 Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination, pp. 783-792 in *Mol Cell*.

Nagai, S., K. Dubrana, M. Tsai-Pflugfelder, M. B. Davidson, T. M. Roberts *et al.*, 2008 Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase, pp. 597-602 in *Science*.

Nouspikel, T., 2009 DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility, pp. 994-1009 in *Cell Mol Life Sci*.

O'Farrell, P. H., J. Stumpff and T. T. Su, 2004 Embryonic cleavage cycles: how is a mouse like a fly? Current biology : CB **14:** R35-45.

Ogi, T., and A. R. Lehmann, 2006 The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair, pp. 640-642 in *Nat Cell Biol*.

Ogi, T., S. Limsirichaikul, R. M. Overmeer, M. Volker, K. Takenaka *et al.*, 2010 Three DNA Polymerases, Recruited by Different Mechanisms, Carry Out NER Repair Synthesis in Human Cells, pp. 714-727 in *Mol Cell*.

Ohashi, E., T. Hanafusa, K. Kamei, I. Song, J. Tomida *et al.*, 2009 Identification of a novel REV1-interacting motif necessary for DNA polymerase kappa function. Genes to cells : devoted to molecular & cellular mechanisms **14:** 101-111.

Palancade, B., and V. Doye, 2008 Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties?, pp. 174-183 in *Trends Cell Biol*.

Palancade, B., X. Liu, M. Garcia-Rubio, A. Aguilera, X. Zhao *et al.*, 2007 Nucleoporins prevent DNA damage accumulation by modulating Ulp1-dependent sumoylation processes, pp. 2912-2923 in *Mol Biol Cell*.

Plasterk, R. H., 1991 The origin of footprints of the Tc1 transposon of Caenorhabditis elegans. The EMBO journal **10:** 1919-1925.

Pontier, D. B., and M. Tijsterman, 2009 A robust network of double-strand break repair pathways governs genome integrity during C. elegans development, pp. 1384-1388 in *Curr Biol*.

Prakash, S., R. E. Johnson and L. Prakash, 2005 Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function, pp. 317-353 in *Annu Rev Biochem*.

Rattray, A. J., and J. N. Strathern, 2005 Homologous recombination is promoted by translesion polymerase poleta, pp. 658-659 in *Mol Cell*.

Roos, W. P., and B. Kaina, 2012 DNA damage-induced apoptosis: From specific DNA lesions to the DNA damage response and apoptosis. Cancer letters.

Schumacher, B., M. Hanazawa, M.-H. Lee, S. Nayak, K. Volkmann *et al.*, 2005a Translational repression of C. elegans p53 by GLD-1 regulates DNA damage-induced apoptosis, pp. 357-368 in *Cell*.

Schumacher, B., M. Hanazawa, M.-H. Lee, S. Nayak, K. Volkmann *et al.*, 2005b Translational repression of C. elegans p53 by GLD-1 regulates DNA damage-induced apoptosis. Cell **120:** 357-368.

Semple, J. I., R. Garcia-Verdugo and B. Lehner, 2010 Rapid selection of transgenic C. elegans using antibiotic resistance, pp. 725-727 in *Nat Methods*.

Shachar, S., O. Ziv, S. Avkin, S. Adar, J. Wittschieben *et al.*, 2009 Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals, pp. 383-393 in *EMBO J*.

Stavru, F., B. B. Hülsmann, A. Spang, E. Hartmann, V. C. Cordes *et al.*, 2006 NDC1: a crucial membrane-integral nucleoporin of metazoan nuclear pore complexes. The Journal of Cell Biology **173:** 509-519.

Stergiou, L., K. Doukoumetzidis, A. Sendoel and M. O. Hengartner, 2007 The nucleotide excision repair pathway is required for

UV-C-induced apoptosis in Caenorhabditis elegans. Cell death and differentiation **14:** 1129-1138.

van Haaften, G., R. Romeijn, J. Pothof, W. Koole, L. H. F. Mullenders *et al.*, 2006 Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. Current biology : CB **16:** 1344-1350.

Waters, L. S., B. K. Minesinger, M. E. Wiltrout, S. D'Souza, R. V. Woodruff *et al.*, 2009a Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance, pp. 134-154 in *Microbiol Mol Biol Rev*.

Waters, L. S., B. K. Minesinger, M. E. Wiltrout, S. D'Souza, R. V. Woodruff *et al.*, 2009b Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiology and molecular biology reviews : MMBR **73:** 134-154.

Youds, J. L., L. J. Barber, J. D. Ward, S. J. Collis, N. J. O'Neil *et al.*, 2008 DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Molecular and Cellular Biology **28:** 1470-1479.

Zhang, H., G. A. Smolen, R. Palmer, A. Christoforou, S. van den Heuvel *et al.*, 2004 SUMO modification is required for in vivo Hox gene regulation by the Caenorhabditis elegans Polycomb group protein SOP-2. Nature genetics **36:** 507-511.

Ziv, O., N. Geacintov, S. Nakajima, A. Yasui and Z. Livneh, 2009 DNA polymerase zeta cooperates with polymerases kappa and iota in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients, pp. 11552-11557 in *Proc Natl Acad Sci USA*.

SUPPLEMENTARY INFORMATION

FIGURE S1. Full alignment of *C. elegans* POLH-1 and POLK-1 with human Polη and Polκ and yeast Rad30.

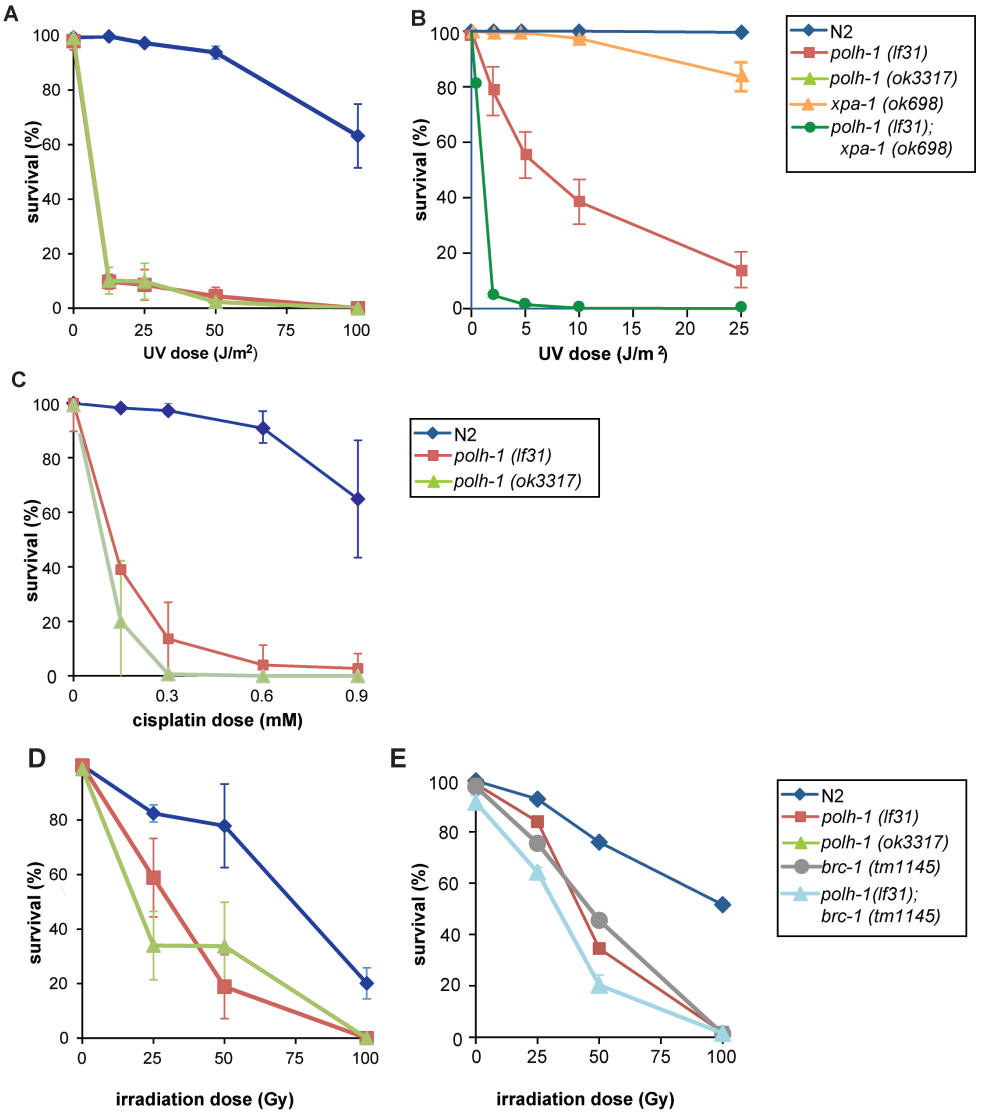


FIGURE S2. Germline sensitivity of *polh-1 (lf31)* and *polh-1 (ok3317)* mutants combined with repair defective backgrounds to different sources of DNA damage.
A-B. Sensitivity to UV irradiation. **B.** Epistasis analysis for *xpa-1* and *polh-1*. **C.** Sensitivity to cisplatin. **D-F.** Sensitivity to γ-irradiation. (E-F) Epistasis analysis of *polh-1 (lf31)* and *brc-1(tm1145)* mutants for γ-irradiation. Data have been normalized for reduced survival (about 95%) in *polh-1;brc-1* double mutants without treatment. Results of representative experiments are shown for A, C, D and F. Error bars denote SD. (A)(C)(D) Both alleles of *polh-1* lead to equal sensitivity to various damaging agents. Each line in B and E represents the mean of minimal three independent experiments. Error bars denote the s.e.m.

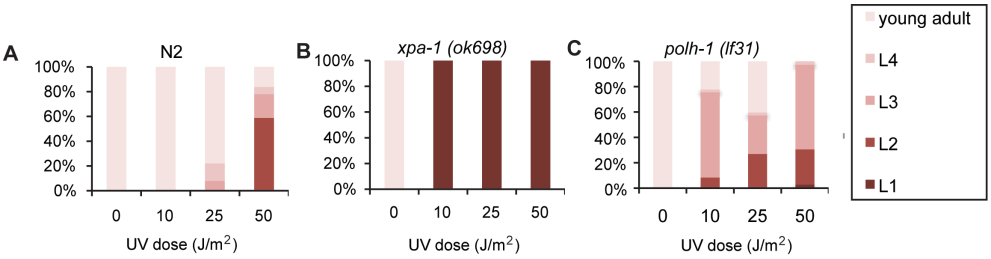


FIGURE S3. Development of L1 larvae three days after treatment with indicated UV doses.
Percentage of animals in the different larval stages (L1-L2-L3-L4-young adult) was quantified 72 hrs after exposure to UV-irradiation.

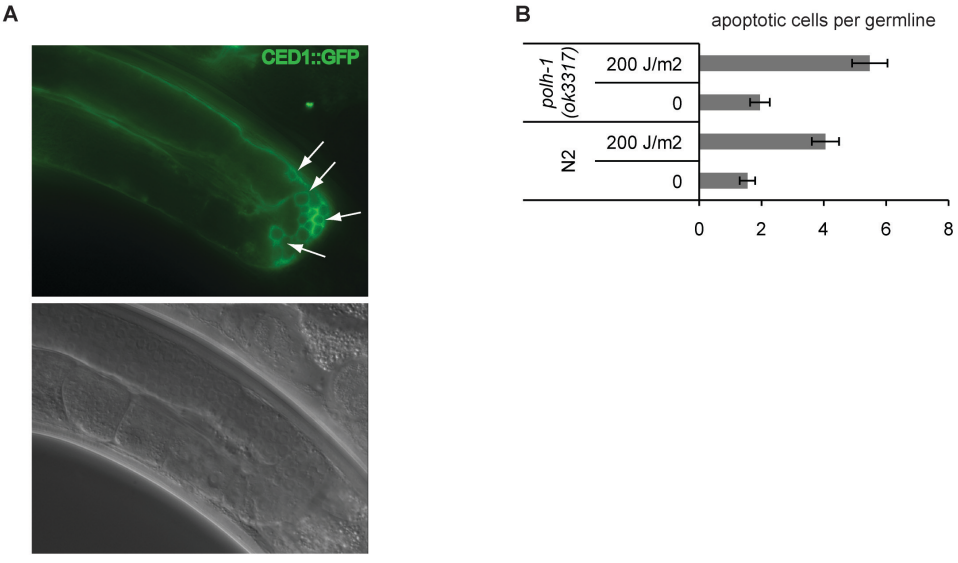


FIGURE S4. Apoptosis in the germline after UV-irradiation.
A. A transgenic line expressing pLim7::ced1::GFP is used. Examples of CED1-GFP engulfed cells in the germline bend are indicated with arrows. **B.** Quantification of CED1-GFP engulfed cells. In the *polh-1(ok3317)* mutant apoptosis is slightly increased after exposure to UV irradiation. About 40 germlines have been analysed per sample. Error bars denote s.e.m.

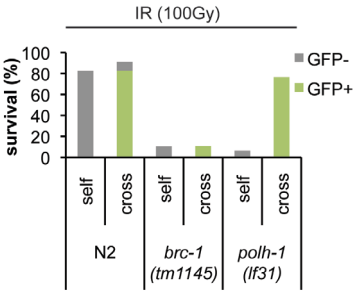


FIGURE S5. γ-irradiated hermaphrodites were crossed with non-irradiated males carrying a Pmyo-2::GFP transgene.
Lethality induced by γ-irradiation is almost fully rescued in the cross progeny of *polh-1*, but not *brc-1* hermaphrodites.

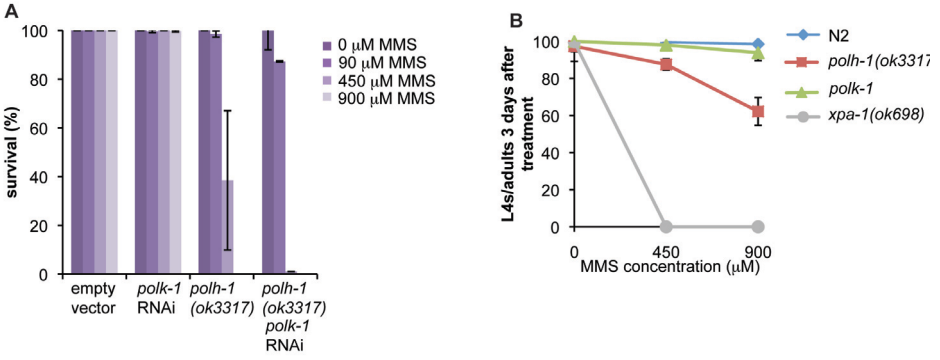


FIGURE S6. POLH-1 and POLK-1 in cellular tolerance to MMS during embryogenesis and L1 outgrowth.
A. MMS sensitivity of N2 and *polh-1(ok3317)* mutants with or without depletion of POLK-1 by RNAi. **B.** Development of larvae exposed to MMS from L1 stage. Each line is the mean of three independent experiments; error bars denote s.e.m.

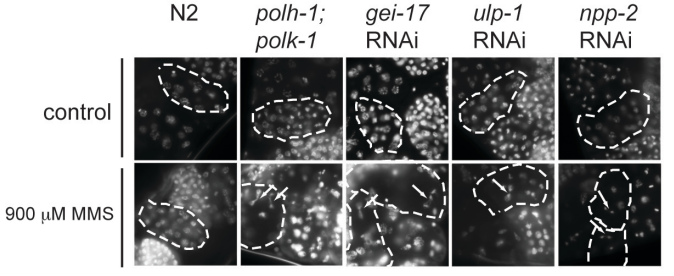


FIGURE S7. Morphological defects in MMS exposed embryos.
DAPI-stainings of whole animals exposed for 24hrs to MMS reveal a delay in development on indicated RNAi foods. Incidentally, chromatin bridges are visible (arrows) indicative of incomplete DNA replication.

TABLE S1. List of newly generated strains used in this study.

strain	genotype
XF110	*polk-1(lf29)*
XF132	*polh-1(lf31)*
XF656	*polh-1(ok3317)*
XF242	*polh-1(lf31);polk-1(lf29)*
XF507	*polh-1(lf31);polk-1(lf29)*
XF729	*polh-1(lf31); xpa-1(ok698)*
XF726	*polh-1(ok3317);xpa-1(ok698)*
XF155	*polh-1(lf31);brc-1(tm1145)*
XF153	*polh-1(lf31);dog-1(gk10)*
XF504	*polh-1(lf31);rde-3(ne298)*
XF496	*brc-1(tm1145);rde-3(ne298)*
XF503	*lfs129[P(elt-2)::GFP-HRreporter; rol-6(su1004);P(HS)::ISceI::Cherry]*
XF659	*polh-1(ok3317); lfs129[P(elt-2)::GFP-HRreporter; rol-6(su1004);P(HS)::ISceI::Cherry]*
XF550	*brc-1(tm1145); lfs129[P(elt-2)::GFP-HRreporter; rol-6(su1004);P(HS)::ISceI::Cherry]*
XF657	*polh-1(lf31); bcIs39[P(lim-7)ced-1::GFP + lin-15(+)]*

TABLE S2. *polh-1* does not influence lethality in a mutant background where transposition is desilenced. Double mutants for the HR gene *brc-1* and the mutator gene *rde-3* display synthetic lethality while *polh-1; rde-3* double mutants are comparable to *rde-3* single mutants.

Genotype	Survival*
rde-3 (ne298)	88 (±10) %
rde-3(ne298); brc-1 (tm1145)	25 (±20) %
rde-3 (ne298); polh-1 (lf31)	96 (±15) %

* progeny of at least 500 individuals from minimal 12 worms has been counted

MOVIES S1 and S2 are available at:
http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1002800#s5

A MUTAGENIC REPAIR PATHWAY DEPENDENT ON POLYMERASE θ PROTECTS THE *C. ELEGANS* GENOME AGAINST LARGE CHROMOSOMAL DELETIONS

3

Sophie F. Roerink*, Robin van Schendel*, Marcel Tijsterman
*equally contributing authors

Leiden University Medical Center,
Department of Toxicogenetics

ABSTRACT

DNA lesions that block replication fork progression are drivers of cancer-associated genome alterations, but the error prone DNA repair mechanisms acting on collapsed replication are incompletely understood, and there contribution to genome evolution largely unexplored. Here, by whole genome sequencing of mutation accumulation line, we identify a distinct class of deletions that spontaneously accumulate in *C. elegans* strains lacking translesion synthesis (TLS) polymerases. Emerging DNA double-strand breaks are repaired via an error-prone mechanism in which the outermost nucleotide of one end serves to prime DNA synthesis on the other end. This pathway critically depends on the A-family polymerase θ, which protects the genome against gross chromosomal rearrangements. By comparing the genomes of isolates of *C. elegans* from different geographical regions, we found that in fact most spontaneously evolving structural variations match the signature of polymerase Theta-mediated end joining (TMEJ), arguing that this pathway is an important source of genetic diversification.

INTRODUCTION

Identifying the mechanisms that fuel genome change is crucial for understanding evolution and carcinogenesis. Spontaneous mutagenesis is caused predominantly by misinsertions or slippage events of replicative polymerases that are missed by their proofreading domains, and not corrected by mismatch repair (Lynch *et al.* 2008). Less frequently, but with a potentially much more detrimental effect, mutations can arise when DNA damages obstruct progression of DNA replication; and stalled replication forks eventually collapse, resulting in highly mutagenic double stranded breaks (DSBs). While error-free homologous repair, where the sister chromatid is used as a template, restores the original sequence, infrequent error-prone end joining processes may give rise to spontaneous deletions and tumor-promoting translocations (Mitelman *et al.* 2007).

To circumvent fork collapse at DNA damage, cells employ various alternative polymerases that are capable of incorporating nucleotides across DNA lesions, and are hence called translesion synthesis (TLS) polymerases. TLS acts on a wide variety of DNA lesions that can result from endogenous as well as exogenous genotoxic sources: DNA lesions that result from UV-light exposure, for instance, are efficiently bypassed by the well-conserved TLS polymerase η, inactivation of which in humans leads to the variant form of the skin cancer predisposition syndrome Xeroderma Pigmentosum (Johnson *et al.* 1999; Masutani *et al.* 1999). Abundant *in vitro* studies demonstrate the involvement of TLS proteins Pol η and Pol κ in bypass of lesions that are produced by endogenous reactive compounds, arguing that these polymerases are also essential for protection of the genome under unchallenged conditions (Haracska *et al.* 2000; Kusumoto *et al.* 2002; Fischhaber *et al.* 2002).

Although error-prone while replicating, and thus potentially causing misinsertions, TLS polymerases are thought to protect cells against the more mutagenic effects of replication fork collapse (Lynch *et al.* 2008; Knobel and Marti 2011). Here, we investigate the contribution of TLS polymerases on the maintentance of genome stability and the mechanisms acting on stalled DNA replication, by characterizing *C. elegans* strains that are defective for the Y-family polymerases η and κ.

RESULTS AND DISCUSSION

TLS polymerases protect genomes against spontaneous deletions

In previous work, we established the role of the *C. elegans* homologs of TLS polymerases η and κ in protection against various a wide range of exogenous DNA damaging agents (Mitelman *et al.* 2007; Roerink *et al.* 2012). Anecdotal evidence in these studies, where we observed readily recognizable mutant phenotypes during normal culturing of *polh-1*

and *polk-1* double mutant animals, prompted us to suspect a prominent role for these Y-family of TLS polymerases in the prevention of spontaneous mutations (Figure S1). To assess mutagenesis in an unbiased way, we generated MA lines for single and double mutants in TLS for 60 generations, thus allowing spontaneous mutations to accumulate, and then sequenced their genomes (Figure 1A, Table S1 and Supplemental data file). Mutation accumulation lines of a wildtype strain (Bristol N2) and of the mismatch repair deficient strain *msh-6* - for which an ~100-fold higher mutation frequency has been reported (Johnson *et al.* 1999; Masutani *et al.* 1999; Tijsterman *et al.* 2002; Denver *et al.* 2005) - were sequenced as references.

Although Pol η and κ have reduced accuracy while replicating from undamaged as well as damaged DNA templates (Matsuda *et al.* 2000; Haracska *et al.* 2000; Ohashi *et al.* 2000; Kusumoto *et al.* 2002; Fischhaber *et al.* 2002), we found that they hardly, if at all,

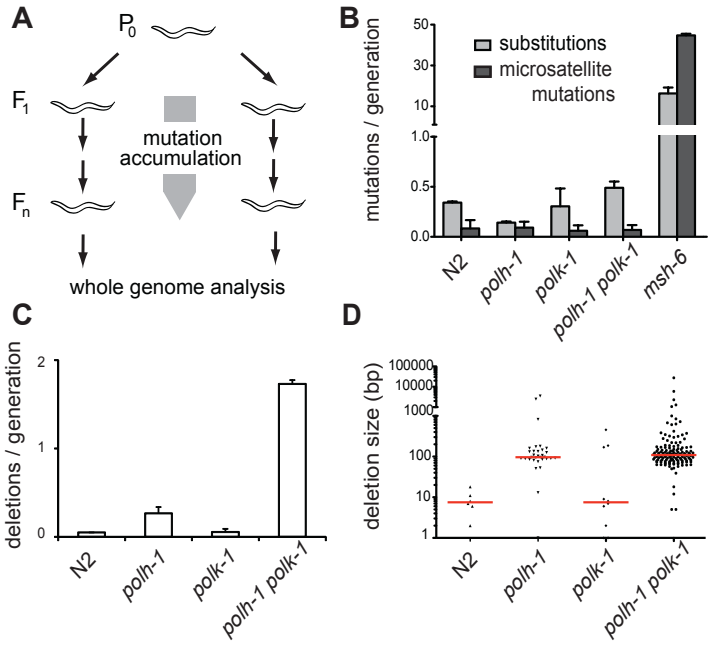


FIGURE 1. Spontaneous mutagenesis in TLS deficient strains. A. Generation of mutation accumulation (MA) lines. For each genotype multiple populations were started by cloning out single worms from a single hermaphrodite P0. Cultures were propagated by transferring animals to new plates each generation. At generation Fn, a single animal was grown to a full population of which genomic DNA was isolated and subjected to whole genome sequencing on an Illumina HiSeq. **B.** Substitution and microsatellite mutation rates for the indicated genotypes. Mutation rates are expressed as the number of mutations per generation divided by the number of nucleotides analysed. **C.** Rates of structural variations for the indicated genotypes. (D) Size distribution of deletions in the different mutant backgrounds. The median sizes are indicated in red.

contribute to base substitution processes under normal growth conditions (Figure 1B). Erroneous replication of microsatellites is another dominant basis of mutagenesis under wildtype growth conditions(Denver *et al.* 2009) for which activity of TLS polymerases has been implicated (Burr *et al.* 2006; Hile *et al.* 2011). However, our whole genome sequencing data show no elevation in the microsatellite mutation rate in *polh-1polk-1* worms compared to wildtype controls (Figure 1B), arguing that other types of genome alteration, or epigenetic changes (Sarkies *et al.* 2010) must underlie the phenotypic variations observed in TLS deficient cultures.

To detect any other structural variations, we employed Pindel software, developed to identify deletions and/or insertions in whole-genome sequencing data (Ye *et al.* 2009). Strikingly, a unique class of deletions emerged in *polh-1* and *polh-1 polk-1* mutants, which were not associated with repetitive loci or any other obvious genomic trait, and occurred at seemingly random locations throughout the genome (Figures 1C and S2). The vast majority of deletions ranged between 50 and 200 bp in size, with just a few exceptions being larger or smaller (Figure 1D). The median size, of 107 bp, was similar for *polh-1* and *polh-1polk-1*-derived deletions (Figure 1D). Control wildtype and *msh-6* samples did not display any mutations from this class. Deletions occurred in *polh-1* single mutants with a rate of ~0.4 per animal generation, which translates to an average of ~0.03 deletion per genome per cell division. *Polk-1* single mutants hardly suffered from deletions; however, *polh-1polk-1* double mutants had 5-fold increased rates of deletion induction as compared to *polh-1* single mutant animals, implying that *C. elegans* Pol η and Pol κ function redundantly on a subset of endogenous lesions.

These data also argue that Pol κ and Pol η prevent the induction of spontaneous deletions in a largely error-free manner: their joined action prevents ~2 deletions per animal generation without significantly affecting the overall substitution rate. In fact, the number of substitutions (~0.3 per generation) in TLS-proficient strains compares to only 15 percent of the number of deletions that are being induced in the absence of Pol η and Pol κ indicating that the vast majority of endogenous replication blocking lesions that use TLS to avoid deletion formation, are not leading to mutations (Figure 1B).

Footprints of error-prone DSB repair

To further investigate the origin of the high number of deletions in *polh-1polk-1* deficient strains, we looked for manifestations of genomic instability in germ cells of these animals.

We observed a mild but statistically significant increase in the number of foci of the DSB marker RAD-51 in proliferating germ cells of *polh-1polk-1* mutant animals (Figure S3A-B). Elevated levels of DSBs are also suggested by the spontaneous emergence of dominant *him* mutants in *polh-1polk-1* mutant populations (Figure S1). This phenotype, which is defined by dominant inheritance of an increased number of males (XO) in

predominantly hermaphroditic (XX) populations, has previously been found upon exposure to γ-irradiation and in mutants with enhanced telomere shortening, where it proved to result from X/autosome translocations (Herman *et al.* 1982; Meier *et al.* 2009). Despite these manifestations of enhanced replication stress in *polh-1polk-1* mutants, the levels of DSBs were insufficient to activate the two DNA damage checkpoints that operate in the *C. elegans* germline: cell cycle arrest and apoptosis (Gartner *et al.* 2000). We found neither a reduction in germ cell proliferation nor an increase of apoptotic bodies in *polh-1polk-1* mutant germlines, suggesting that TLS compromised germ cells proliferate in the presence of elevated levels of DSBs, with genomic deletions as a consequence (Figure S3C-E).

To obtain mechanistic insight on the biology of deletion formation, we performed a detailed analysis on the sequence context of 141 *polh-1polk-1*-derived deletions (Supplemental data file). While the majority had simple deletion junctions (without inserts), about 25 percent of the footprints showed insertions of short sequence stretches (Figure 2A). Cases with inserts sufficiently long to faithfully trace their origin revealed that the inserted stretch, or part of it, is identical to sequences flanking the deletion (Figure 2B-C). This finding strongly suggests that DNA close to the break site was used as a template for *de novo* synthesis before both DNA ends were joined. The close proximity of insertions to their template also suggests that the extendable end of the DSB is not subjected to extensive trimming.

A DSB repair mechanism involving DNA synthesis is also suggested by the notion of a 'priming' nucleotide in more than 80 percent of all deletions: 83 of the 102 deletions without insert contain at the junction at least one nucleotide that could have originated from either flank; in 51 cases this is restricted to a single nucleotide. To systematically assess the significance of this observation, we constructed deletion junction heat maps, which reflect the level of (micro)homology between 5' and 3' junctions (Figure 2D-F). We scored the degree of sequence identity in an 8 nt window, encompassing the 4 outermost nucleotides of the flanking sequence and the 4 nucleotides of the adjacent, but deleted, sequence. Indeed, compared to a randomly distributed simulated set, we found a very high similarity score for the nucleotide at the -1 position of the deletions and the +1 position of the opposing flanks ($p=7.3x10^{-17}$). Importantly, this profound degree of microhomology is restricted to only a single nucleotide, arguing that the underlying mechanism is distinct from previously described repair pathways involving more extended microhomology (Payen *et al.* 2008; McVey and Lee 2008; Hastings *et al.* 2009).

We investigated whether the deletion specifics would reveal the nature of the spontaneous damage underlying fork stalling and break formation in TLS compromised animals. We hypothesized that if the nascent strand, blocked at the site of base damage, defines one end of a DSBs, then the -1 position of the corresponding junction will represent the nucleotide complimentary to the damaged base: it is the first base not to be incorporated. To test this hypothesis, we plotted the base distribution for each

position of the junction and indeed found it not to be random at the -1 position, but rather being dominated by cytosines (Figure 2G). This result strongly suggests that spontaneous base damage primarily at guanines requires Pol κ and Pol η to avoid DSB induction, which may point towards N2-dG and/or 8-oxo-dG adducted sites as a primary source of spontaneous mutagenesis, as bypass activities of Pol κ and Pol η have been reported for these lesions (Haracska *et al.* 2000; Avkin *et al.* 2004; Choi *et al.* 2006). This conclusion is further supported by the notion that the degree of spontaneous deletion induction in these mutant backgrounds, follows their sensitivity towards the guanine alkylator MMS, where Pol η mutant animals are much more sensitive than Pol κ, but not as sensitive as double knockout worms (Roerink *et al.* 2012).

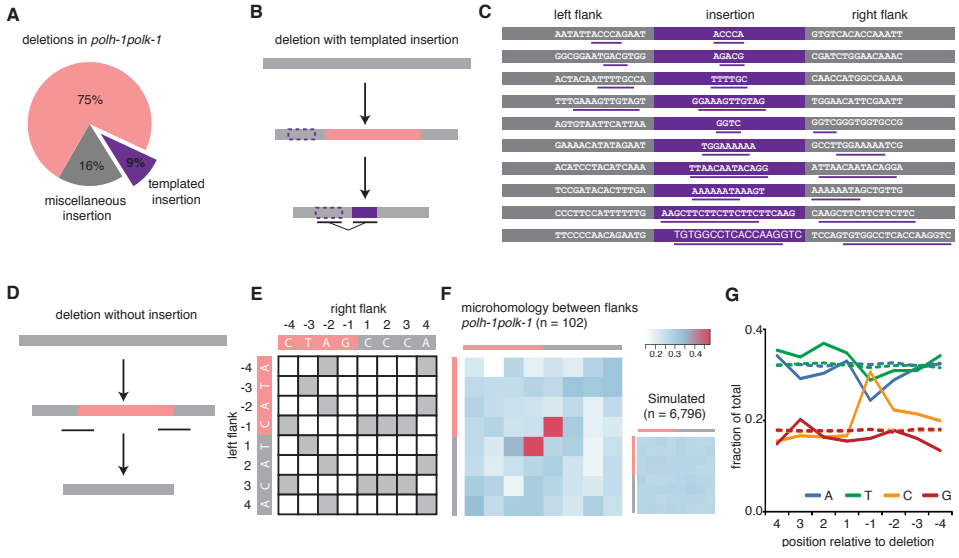


FIGURE 2. Deletion footprints in TLS mutants indicate a priming-based end joining mechanism. A. Distribution of deletion footprints in *polh-1polk-1* mutants. (B) Schematic illustration of a deletion associated with a templated insertion. Deleted sequence in pink; newly inserted sequence in purple and its template boxed; non-altered DNA in grey. **C.** Sequence context of deletions with templated insertions derived from *polh-1polk-1* animals. Matching sequences are underlined. **D.** Schematic illustration of a deletion not accompanied by insertions. Deleted sequence in pink; non-altered DNA in grey. The eight nucleotide window -capturing neighboring flanking and deleted sequences- that is used for the generation of the heat maps is underlined. **E.** The strategy to score junction homology: for each simple deletion, matching bases between the 5′ and 3′ junction were scored 1, non-matching bases were scored 0, thus creating one map per deletion. **F.** A heat map representing the sum of all individual deletion maps derived from *polh-1polk-1* animals. (n=102). A heat map for a simulated set of deletions (n=6796) with random distribution is displayed on the right. **G.** Base composition at the 5′ and 3′ junctions. The flanking sequences have positive numbers, the deleted sequence have negative; -1 being the first nucleotide within the deletion. Dotted lines indicate the relative abundance of a particular base for a simulated set of deletions (n=6796).

Polymerase θ mediates deletion formation

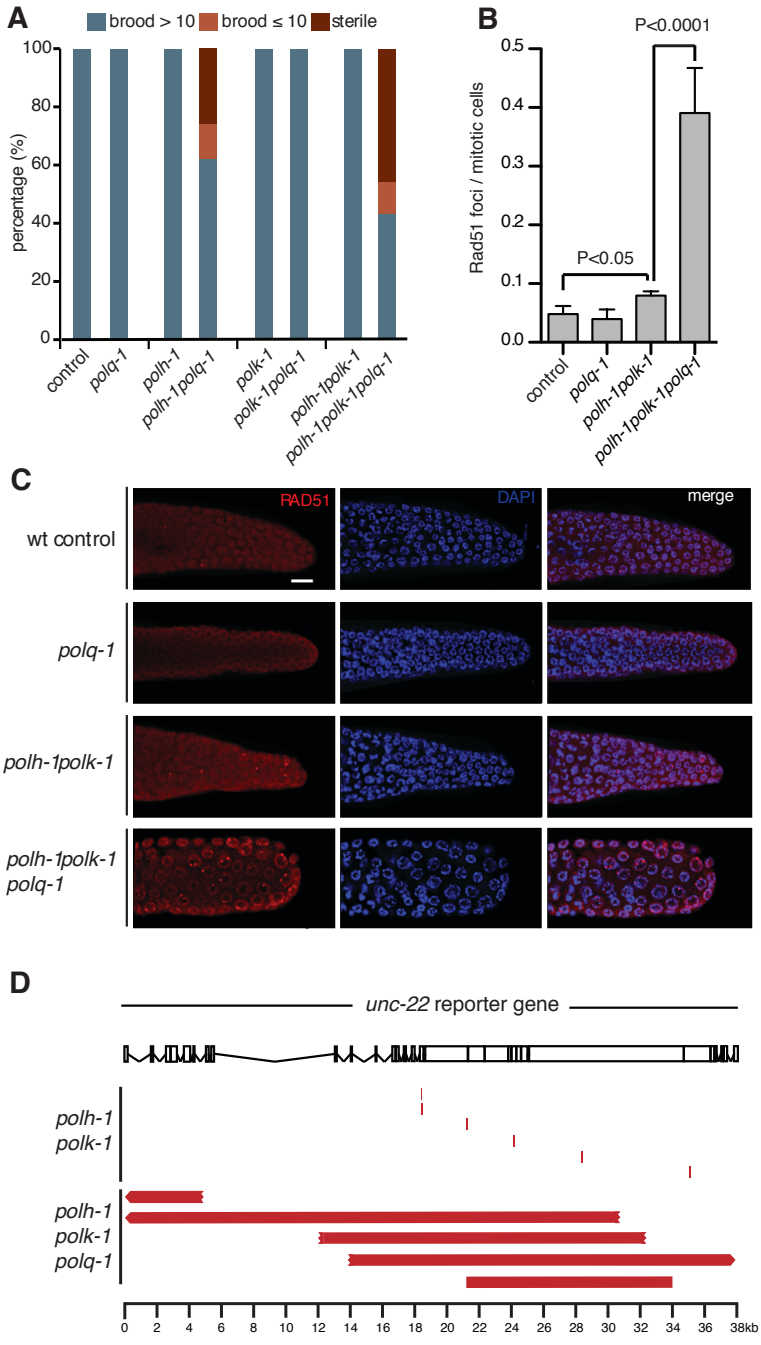
The footprints of the deletions that are suppressed by TLS polymerases fit best with a model in which one end of a DSB, induced at replication-blocking dG bases, is extended using the other end as a template, with just a single base-paired nucleotide as a primer (explaining both single nucleotide homology and templated insertions). In this model, templated inserts can be explained as the result of iterative rounds of annealing and extension (Figure S4).

We next wished to identify the polymerase responsible for inserting new nucleotides at breaks that result from TLS deficiency. One candidate is *C. elegans* Polymerase θ. Pol θ is a polymerase from the A-family, previously implicated in repair of interstrand crosslinks in various models (Shima *et al.* 2004; Muzzini *et al.* 2008; Yousefzadeh and Wood 2012), and repair of transposition-induced DSBs in Drosophila. Pol θ mutant mice display spontaneous genomic instability and increased radiosensitivity. The molecular function of this protein in these phenotypes is, however, largely unknown (Shima *et al.* 2004). In an independent study, we recently identified Pol θ in prevention of genomic instability at endogenous sequences that are able to fold into potentially replication blocking G-quadruplex structures (Koole et al, article submitted).

To test a possible role for Pol θ in deletion formation at spontaneous damage, we generated animals defective for *polh-1polk-1* and the *C. elegans* Pol θ homolog *polq-1*. Strikingly, these animals are severely compromised in normal growth: while *polq-1* and *polh-1polk-1* animals had nearly wildtype growth characteristics, *polh-1polk-1polq-1* triple mutant animals had very much reduced fertility, albeit in a stochastic fashion, ranging from complete sterility to brood sizes of 30 percent of wildtype levels (Figure 3A). Associated with these fertility defects, we observed a profound increase in the number of RAD-51 foci in the proliferative zone of the germline as well as activation of the DNA damage checkpoint (Figure 3B,C), suggesting increased DNA end-resection and DSB signaling, (Figure S3E). From this we conclude that when damage cannot be bypassed Pol θ action safeguards animal fertility by preventing undesired HR-related processing of replication-associated breaks, which would trigger checkpoint activation and prohibit proliferation.

To study the role of Pol θ in deletion formation on a molecular level, we assessed mutagenesis using an endogenous *unc-22* reporter gene (Figure 3D). We isolated spontaneous *unc-22* mutants from *polh-1polk-1* and *polh-1polk-1polq-1* populations

FIGURE 3. Pol θ mediates end joining of breaks in Pol η and Pol κ deficient animals. A. Fecundity of single, double and triple knockout mutants of Polymerase θ and TLS Polymerases η and κ. **B.** Quantification and (**C**) representative pictures of RAD-51 immunostainings on germlines of the indicated genotype. Scale bar, 10 μm **D.** Schematic representation of the *unc-22* reporter gene and spontaneous deletions (in red) isolated from either *polh-1polk-1* or *polh-1polk-1polq-1* mutant animals. Three out of five deletions extended beyond the borders of the *unc-22* locus.


A
brood > 10
brood ≤ 10
sterile
percentage (%)
100
80
60
40
20
0
control
polq-1
polh-1
polh-1polq-1
polk-1
polk-1polq-1
polh-1polk-1
polh-1polk-1polq-1
B
P<0.0001
P<0.05
Rad51 foci / mitotic cells
0.5
0.4
0.3
0.2
0.1
0.0
control
polq-1
polh-1polk-1
polh-1polk-1polq-1
C
RAD51
DAPI
merge
wt control
polq-1
polh-1polk-1
polh-1polk-1
polq-1
D
unc-22 reporter gene
polh-1
polk-1
polh-1
polk-1
polq-1
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38kb

and determined their molecular nature using PCR and Sanger sequencing. In perfect agreement with our whole-genome sequencing data, all *unc-22* mutations derived from *polh-1polk-1* animals were 50-200 bp deletions characterized by single nucleotide homology and templated insertions (Figure 3D, Table S2). In sharp contrast, *unc-22* mutants derived from *polh-1polk-1polq-1* triple mutants, while being induced at comparable rates, were of a completely different size category. Here, deletions were typically larger than 5 kb, with some spanning over 30 kb of genomic sequence, thus amplifying the deleterious impact of replication stalling lesions more than 100-fold (Figure 3D, Tables S2 and S3).

We conclude that a Pol θ-mediated end joining mechanism (TMEJ) is responsible for the small-sized deletions induced at replication fork stalling endogenous lesions. In its absence, large stretches of DNA surrounding DSBs are resected, resulting in abundant RAD-51 filament formation, mitotic checkpoint activation and a Pol θ-independent repair process accompanied by excessive loss of DNA. Reduced viability of the triple mutant may be the consequence of both processes: unscheduled checkpoint activation and loss of large genomic regions encoding essential genes.

TMEJ in wildtype *C. elegans* strains

The notion that we have uncovered TMEJ acting at spontaneous damage but only in TLS mutants, raises the question: do cells rely on TMEJ under TLS proficient conditions, or in other words, how relevant is TMEJ for animal fitness? We hypothesized that the action of an error-prone repair mechanism with such a clear and distinct signature, *i.e.* a distinct size class, single nucleotide homology and templated insertions, may leave its fingerprint in evolving genomes. For this reason, we compared the genomes of different natural isolates of *C. elegans*, to identify structural variations and defined their characteristics (Figure 4). The majority of deletions is of small size - 60 percent being smaller than 10 bp - while the number of deletions decreases with increasing size in an exponential manner. However, we found deletions in the size range 50-200 bp much more abundantly present than expected from this exponentially declining trend (Figure 4B), suggesting that the typical TMEJ-mediated deletions seen in TLS-deficient circumstances also contribute to mutation induction in wildtype natural strains (Figure 4B). Indeed, deletions in this size range bear the TMEJ signature including templated insertions and a strong overrepresentation (over 80 percent) of having at least one nucleotide homology (Figure 4C).Unexpectedly, we observed also templated insertions (2%) in the small size range of deletions, and found this class to also be dominated by ≥1 nucleotide homology at the junction (Figure 4C-D), which may argue that TMEJ is a very prominent source of mutagenesis in *C. elegans* evolution.

To further test the involvement of TMEJ in spontaneous mutation induction under non-challenged growth we used a forward mutagenesis assay that is based on the uncoordinated movement of animals carrying a dominant mutation (e1500) in the

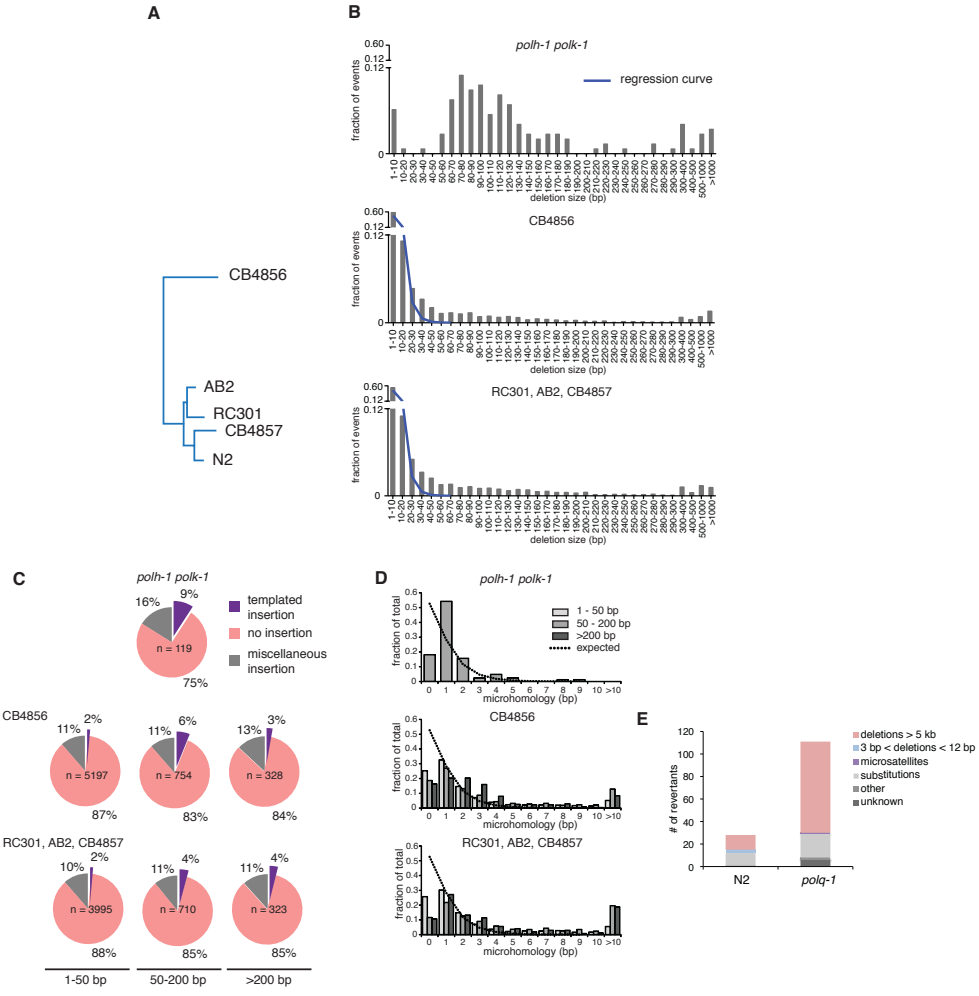


FIGURE 4. Signature of Pol θ - mediated end joining in natural isolates of *C. elegans*.
A. Phylogenetic tree diagram of the different isolates of *C. elegans* used in this study. **B.** Size distribution of deletions of evolutionary distinct *C. elegans* species compared to size distribution of *polh-1polk-1* derived deletions. An exponential regression curves describes the size distribution of deletions in both natural isolates up to 20 bp; deletions up to 300 bp are overrepresented. **C.** Deletions in natural isolates, especially in size class 50-300 bp show templated insertions analogously to deletion footprints in *polh-1polk-1* animals. **D.** Microhomology for deletions in natural isolates as compared to deletions in *polh-1polk-1* animals. **E.** *unc-93* mutagenesis in *polq-1* worms and wildtype controls.

UNC-93 protein that affects muscle contraction (De Stasio *et al.* 1997; Greenwald and Horvitz 2003). This easily recognizable phenotype is suppressed by complete loss of function of *unc-93,* or by loss of one of several extragenic suppressor genes (*e.g. sup-9*, *sup-10*). We propagated 400 populations of wildtype and *polq-1* mutant animals out of which we isolated and moleculary characterized revertants animals that had normal movement. Strikingly, the total number of revertants was increased fourfold in *polq-1* mutants, suggesting that in the absence of Pol θ, TMEJ substrates are shuttled into more mutagenic pathways (Figure 4E, Tables S4 and S5). Indeed, about 75 percent of all *polq-1*-deficient revertants were attributed to large chromosomal deletions in *unc-93* or the suppressor genes *sup-9* and *sup-10* (Figure S5); while *sup-9* and *sup-10* occupy less genomic space than *unc-93*, deletions at these sites dominate the overall spectrum, most likely because they, in contrast to *unc-93*, are located in a genomic region that is devoid of essential genes, and are therefore selective targets for large chromosomal deletions (Figure S5).

The selective increase of large deletions in *polq-1* mutants may replace mutations that are otherwise the products of TMEJ. Indeed, one class of mutations, *i.e.* small deletions of a size ≥3 bp, was exclusively found in wildtype animals (3/28 versus 0/111 in *polq-1*), suggesting that this class is the result of TMEJ action. Similar events in intronic and moreover in intragenic regions flanking these genes will give rise to large chromosomal deletions in the absence of Pol θ, explaining the increased mutagenesis in these strains.

TMEJ footprints in evolving genomes

In this study, we have identified a role for polymerase θ in preventing mitotic crisis and loss of DNA sequence at sites of stalled replication. Our data provide a mechanistic model in which Pol θ acts in an error-prone DSB repair pathway to extend one end of a DSB using the other end as a template for new DNA synthesis. This results in a stable association of both ends in the form of a canonical DNA helix, which can be further processed. In this respect, TMEJ provides an alternative to homologous recombination repair in which the sister chromatid is used as a template to generate new DNA that can subsequently be used to anneal to the other broken end. We hypothesize that TMEJ provides the cells with the ability to repair replication-associated breaks in cases where the sister chromatid cannot be used as a template because that still contains the original replication-blocking lesion (Figure S6).

The observation that Pol θ suppresses mutagenesis in wildtype animals, together with the notion that the signature of TMEJ is apparent in mutation profiles under laboratory conditions and in the genomes of natural isolates of *C. elegans* argues for a prominent role of this error-prone pathway to protect genomes against large chromosomal rearrangements.

While mutagenic processes are drivers of evolution, they also fuel malignant

transformation of cells. It is a current challenge to recognize specific classes of mutations in cancer genomes and attribute these either to underlying sources of DNA damage or to error-prone repair mechanisms. Identifying mutational signatures typifying specific repair processes is pivotal to this ambition. Templated insertions and the use of minimal homology - two characteristics of TMEJ - have frequently been observed in higher order eukaryotes and in cancer tissues (Chen *et al.* 2010; Nik-Zainal *et al.* 2012; Carvalho *et al.* 2013), and have been ascribed to either classical nonhomologous end-joining or the molecularly ill-defined process of microhomology-mediated end-joining (Honma *et al.* 2007; Kloosterman *et al.* 2012). Here, we describe a mechanistic alternative for repair of DSBs induced at stalled forks, which leaves a distinct and well-defined footprint in evolving genomes.

EXPERIMENTAL PROCEDURES

C. elegans genetics

All strains were cultured according to standard methods(Brenner 1974). Wildtype N2 (Bristol) worms were used in all control experiments. Alleles used in this study are: *polh-1 (lf31)*; *polh-1 (ok3317)*; *polk-1 (lf29)*; *polq-1(tm2026); msh-6(pk2504)*; *bcIs39[P(lim-7) ced-1::GFP + lin-15(+)]); unc-93(e1500)*. All mutant strains were backcrossed six times before performing experiments.

Whole genome sequencing of MA lines

Mutation accumulation (MA) lines were generated by cloning out F1 animals from one hermaphrodite. Each generation about five worms were transferred to new plates. MA lines were maintained for 60 generations or until severe growth defects developed. Single animals were then cloned out and propagated to obtain full plates for DNA isolation. Worms were washed off with M9 and incubated for one hour at room temperature while shaking, to remove bacteria from the animal's intestine. After two washes, worm pellets were lysed for two hours at 65°C with SDS containing lysis buffer. Genomic DNA was purified by using a DNeasy kit (Qiagen). Paired end (PE) libraries for whole genome sequencing (HiSeq2000 Illumina) were constructed from genomic DNA according to manufacturers' protocols with some adaptations. Shortly, 5 μg DNA was sheared using a Covaris S220 ultrasonicator, followed by DNA end-repair, formation of 3'A overhangs using Klenow and ligation to Illumina PE adapters. Adapter-ligated products were purified on Qiaquick spin columns (Qiagen) and PCR-amplified using Phusion DNA polymerase and barcoded Illumina PE primers for 10 cycles. PCR products of the 300 - 400 bp size range were selected on a 2% ultrapure agarose gel and purified on Qiaquick spin columns. DNA quality was assessed and quantified using an Agilent

DNA 1000 assay. Four to five barcoded libraries were pooled in one lane for sequencing on a HiSeq.

Bio-informatic analysis

Image analysis, basecalling and error calibration was performed using standard Illumina software. For the analysis of the natural isolates paired-end whole genome sequence data was downloaded from the NCBI Short Read Archive (SRP011413) (Grishkevich *et al.* 2012), and sequence reads were mapped to the *C. elegans* reference genome (Wormbase release 225) by BWA. Samtools was used for SNP and indel calling, with BAQ calculation turned off. All non-unique SNPs and indels are considered to be pre-existing and were filtered out using custom Perl scripts. To identify microsatellite mutations and deletions we used Pindel, developed by Ye et al (Ye *et al.* 2009).
A more detailed description of the bio-informatic procedures is enclosed in the supplemental information.

Microscopy

To study RAD-51 foci formation, germlines were dissected, freeze cracked and subsequently washed with 1% Triton and methanol (-20°C). RAD-51 was visualized by using an anti-RAD-51 rabbit monoclonal antibody and an Alexa488-labelled goat-anti-rabbit secondary antibody (Molecular Probes Inc), combined with 10 μg/mL DAPI. Dissected worms and eggs were mounted using Vectashield. Apoptosis was monitored using a *lim-7* driven *CED-1*::GFP fusion, which visualises sheath cells surrounding apoptotic germ cells. All microscopy was performed with a Leica DM6000 microscope.

unc-22 mutagenesis assay

To identify spontaneous mutations in the *unc-22* muscle gene we started 50 populations by transferring a single animal to 9 cm plates seeded with OP50. In the case of the synthetically sick *polh-1polk-1polq-1* mutant, we started 200 populations with 5 worms per plate. Animals were washed off with 2 mM levamisole and transferred to 6-well plates to facilitate scoring of *unc-22* mutants, which are insensitive to the hypercontracting effects of the drug levamisole. Independent *unc-22* mutant animals were isolated. Genomic DNA was isolated from homozygous animals for subsequent PCR and sequence analysis.

unc-93 (e1500) mutagenesis assay

To generate a complete spectrum of spontaneous mutations we used a mutagenesis assay based on reversion of the socalled 'rubber band' phenotype, caused by a dominant

mutation in the muscle gene *unc-93* (De Stasio *et al.* 1997; Greenwald and Horvitz 2003). Reversion of the *unc-93(e1500)* phenotype is caused by homozygous loss of *unc-93* or one of the suppressor genes *sup-9*, *sup-10*, *sup-11* and *sup-18*. *polq-1(tm2026) unc-93(e1500)* and *unc-93(e1500)* animals were singled to 2 x 400 6 cm plates. These plates were grown till starvation and equal fractions (chunks of 2 x 2 cm) were then transferred to 9 cm plates. Before these plates were fully grown, they were inspected for wildtype moving animals. From each starting culture only one revertant animal was isolated to ensure independent events.

Large chromosomal deletions in *unc-93*, *sup-9* and *sup-10* were identified by PCR amplification of exonic regions and two regions 5 kb upstream and downstream of the respective genes. Smaller genetic changes and substitutions were first classified into events in either the *unc-93* gene or in one of the suppressor genes by their ability to complement a known *unc-93* deletion allele. All *unc-93* exons were sequenced in revertant animals that failed to complement *unc-93*, whereas all exons of *sup-9* and *sup-10* were sequenced in revertants that complemented *unc-93*. *sup-11* or *sup-18* could not be subjected to molecular analysis due to lack of sequence data. Revertants that complemented *unc-93* but had not detectable mutation of *sup-9* or *sup-10*, were classified as 'unkown'.

ACKNOWLEDGEMENTS

We thank the *C. elegans* Knockout Consortium, Shohei Mitani and the *C. elegans* Genetics Center for providing strains. We thank the Leiden Genome Technology Center for support in whole genome sequencing. We thank Wouter Koole and Jane van Heteren for critical reading of the manuscript and Bennie Lemmens and Harry Vrieling for discussions. M. T. is supported by grants from the European Research Council (203379, DSBrepair) and ZonMW/NGI-Horizon.

REFERENCES

Avkin S., Goldsmith M., Velasco-Miguel S., Geacintov N., Friedberg E. C., Livneh Z., 2004 Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa. J Biol Chem **279**: 53298–53305.

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics **77**: 71–94.

Burr K. L.-A., Velasco-Miguel S., Duvvuri V. S., McDaniel L. D., Friedberg E. C., Dubrova Y. E., 2006 Elevated mutation rates in the germline of Polκ mutant male mice. DNA Repair **5**: 860–862.

Carvalho C. M. B., Pehlivan D., Ramocki M. B., Fang P., Alleva B., Franco L. M., Belmont J. W., Hastings P. J., Lupski J. R., 2013 Replicative mechanisms for CNV formation are error prone. Nat Genet.

Chen J.-M., Cooper D. N., Férec C., Kehrer-Sawatzki H., Patrinos G. P., 2010 Genomic rearrangements in inherited disease and cancer. Semin. Cancer Biol. **20**: 222–233.

Choi J.-Y., Angel K. C., Guengerich F. P., 2006 Translesion synthesis across bulky N2-alkyl guanine DNA adducts by human DNA polymerase kappa. J Biol Chem **281**: 21062–21072.

De Stasio E., Lephoto C., Azuma L., Holst C., Stanislaus D., Uttam J., 1997 Characterization of revertants of unc-93(e1500) in Caenorhabditis elegans induced by N-ethyl-N-nitrosourea. Genetics **147**: 597–608.

Denver D. R., Dolan P. C., Wilhelm L. J., Sung W., Lucas-Lledó J. I., Howe D. K., Lewis S. C., Okamoto K., Thomas W. K., Lynch M., Baer C. F., 2009 A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci USA **106**: 16310–16314.

Denver D. R., Feinberg S., Estes S., Thomas W. K., Lynch M., 2005 Mutation rates, spectra and hotspots in mismatch repair-deficient Caenorhabditis elegans. Genetics **170**: 107–113.

Fischhaber P. L., Gerlach V. L., Feaver W. J., Hatahet Z., Wallace S. S., Friedberg E. C., 2002 Human DNA polymerase kappa bypasses and extends beyond thymine glycols during translesion synthesis in vitro, preferentially incorporating correct nucleotides. J Biol Chem **277**: 37604–37611.

Gartner A., Milstein S., Ahmed S., Hodgkin J., Hengartner M. O., 2000 A conserved checkpoint pathway mediates DNA damage--induced apoptosis and cell cycle arrest in C. elegans. Mol Cell **5**: 435–443.

Greenwald I. S., Horvitz H. R., 2003 unc-93(e1500): A behavioral mutant of Caenorhabditis elegans that defines a gene with a wild-type null phenotype. Genetics **96**: 147–164.

Grishkevich V., Ben-Elazar S., Hashimshony T., Schott D. H., Hunter C. P., Yanai I., 2012 A genomic bias for genotype-environment interactions in C. elegans. Mol. Syst. Biol. **8**: 587.

Haracska L., Yu S. L., Johnson R. E., Prakash L., Prakash S., 2000 Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat Genet **25**: 458–461.

Hastings P. J., Ira G., Lupski J. R., 2009 A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet **5**: e1000327.

Herman R. K., Kari C. K., Hartman P. S., 1982 Dominant X-chromosome nondisjunction mutants of Caenorhabditis elegans. Genetics **102**: 379–400.

Hile S. E., Wang X., Lee M. Y. W. T., Eckert K. A., 2011 Beyond translesion synthesis: polymerase {kappa} fidelity as a potential determinant of microsatellite stability. Nucleic Acids Research.

Honma M., Sakuraba M., Koizumi T., Takashima Y., Sakamoto H., Hayashi M., 2007 Non-homologous end-joining for repairing I-SceI-induced DNA double strand breaks in human cells. DNA Repair **6**: 781–788.

Johnson R. E., Prakash S., Prakash L., 1999 Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science **283**: 1001–1004.

Kloosterman W. P., Tavakoli-Yaraki M., van Roosmalen M. J., van Binsbergen E., Renkens I., Duran K., Ballarati L., Vergult S., Giardino D., Hansson K., Ruivenkamp C. A. L., Jager M., van Haeringen A., Ippel E. F., Haaf T., Passarge E., Hochstenbach R., Menten B., Larizza L., Guryev V., Poot M., Cuppen E., 2012 Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. CellReports **1**: 648–655.

Knobel P. A., Marti T. M., 2011 Translesion DNA synthesis in the context of cancer research. Cancer Cell Int. **11**: 39.

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

Lynch M., Sung W., Morris K., Coffey N., Landry C. R., Dopman E. B., Dickinson W. J., Okamoto K., Kulkarni S., Hartl D. L., Thomas W. K., 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. Proceedings of the National Academy of Sciences **105**: 9272–9277.

Masutani C., Araki M., Yamada A., Kusumoto R., Nogimori T., Maekawa T., Iwai S., Hanaoka F., 1999 Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. EMBO J **18**: 3491–3501.

Matsuda T., Bebenek K., Masutani C., Hanaoka F., Kunkel T. A., 2000 Low fidelity DNA synthesis by human DNA polymerase-eta. Nature **404**: 1011–1013.

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

Meier B., Barber L. J., Liu Y., Shtessel L., Boulton S. J., Gartner A., Ahmed S., 2009 The MRT-1 nuclease is required for DNA crosslink repair and telomerase activity in vivo in Caenorhabditis elegans. EMBO J **28**: 3549–3563.

Mitelman F., Johansson B., Mertens F., 2007 The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer **7**: 233–245.

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

Nik-Zainal S., Alexandrov L. B., Wedge D. C., Van Loo P., Greenman C. D., Raine K., Jones D., Hinton J., Marshall J., Stebbings L. A., Menzies A., Martin S., Leung K., Chen L., Leroy C., Ramakrishna M., Rance R., Lau K. W., Mudie L. J., Varela I., McBride D. J., Bignell G. R., Cooke S. L., Shlien A., Gamble J., Whitmore I., Maddison M., Tarpey P. S., Davies H. R., Papaemmanuil E., Stephens P. J., McLaren S., Butler A. P., Teague J. W., Jönsson G., Garber J. E., Silver D., Miron P., Fatima A., Boyault S., Langerød A., Tutt A., Martens J. W. M., Aparicio S. A. J. R., Borg Å., Salomon A. V., Thomas G., Børresen-Dale A.-L., Richardson A. L., Neuberger M. S., Futreal P. A., Campbell P. J., Stratton M. R., the Breast Cancer Working Group of the International Cancer Genome Consortium, 2012 Mutational Processes Molding the Genomes of 21 Breast Cancers. Cell **149**: 979–993.

Ohashi E., Bebenek K., Matsuda T., Feaver W. J., Gerlach V. L., Friedberg E. C., Ohmori H., Kunkel T. A., 2000 Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. J Biol Chem **275**: 39678–39684.

Payen C., Koszul R., Dujon B., Fischer G., 2008 Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. PLoS Genet **4**: e1000175.

Roerink S. F., Koole W., Stapel L. C., Romeijn R. J., Tijsterman M., 2012 A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during C. elegans Embryogenesis. PLoS Genet **8**: e1002800.

Sarkies P., Reams C., Simpson L. J., Sale J. E., 2010 Epigenetic instability due to defective replication of structured DNA. Mol Cell **40**: 703–713.

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

Tijsterman M., Pothof J., Plasterk R. H. A., 2002 Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics **161**: 651–660.

Ye K., Schulz M. H., Long Q., Apweiler R., Ning Z., 2009 Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics **25**: 2865–2871.

Yousefzadeh M. J., Wood R. D., 2012 DNA polymerase POLQ and cellular defense against DNA damage. DNA Repair.

SUPPLEMENTARY INFORMATION

Bio-informatic analysis

Image analysis, basecalling and error calibration was performed using standard Illumina software. Alignments to the annotated sequence of *C. elegans* available at WormBase WS225 were performed by BWA. Samtools was used for SNP and indel calling, with BAQ calculation turned off. All non-unique SNPs and indels are considered to be pre-existing and were filtered out using custom Perl scripts. We considered a SNP to be real if at least 80% of the called bases were non-wildtype for SNPs that are covered ≥4 times. As a second filter at least one of MA lines of the same genotype should be called as wildtype: having a coverage ≥4 of which ≥80% was of wildtype nature according to pileup generated with mpileup. Sanger sequencing of predicted SNPs validated these criteria.

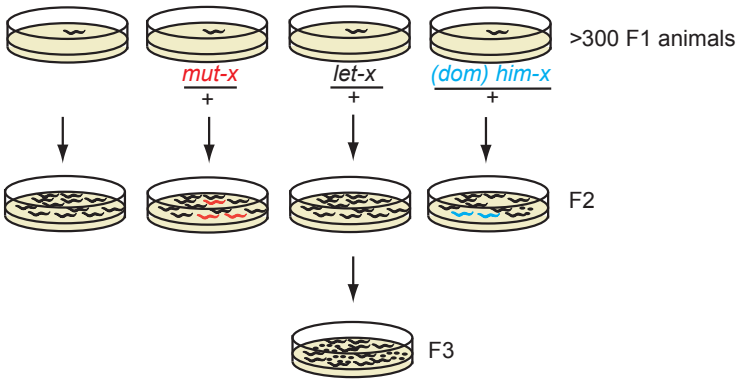
To identify microsatellite mutations and deletions we used Pindel, developed by Ye et al (Ye *et al.* 2009). Standard settings were used, except for changing the maximum allowed mismatch rate (-u) from 0.02 to 0.01. In addition, the sequencing error rate (-e) was adjusted to 0.0001. Pindel results were filtered using custom Java software that selects structural variations that are unique for one sample or the variation was supported by more than eight reads in one sample and eight times less often seen in any other sample. The latter criterium allowed us to also detect variations in homopolymers, which frequently contain sequencing errors. A deletion or deletion with template insertion of ≥20 nt was selected if the local average coverage was below five, while the directly surrounding area was covered >20 fold. Deletions and deletions with template insertions of <20 nt were included when both Pindel and samtools reported the variation and samtools reported it to be homozygous (≥80% non-wildtype). A representative set of deletions was validated by Sanger sequencing. The complete sequencing data have been submitted to the NCBI Short Read Archive (SRA) with accession ID SRP020555.

The simulated set was created in two steps. First, the N2 reference genome was edited by creating random deletions with a size distribution similar to *polh-1;polk-1* throughout the genome using a custom Perl script. Second, the software tool wgsim was used to generate pair-end reads with a 100x coverage based on the modified genome. Subsequently, the reads were analyzed using the same pipeline as used for the MA lines.

For the analysis of the natural isolates paired-end whole genome sequence data was downloaded from the NCBI Short Read Archive (SRP011413), and sequence reads were mapped to the *C. elegans* reference genome (Wormbase release 225). The average base coverage was 176x, 164x, 166x and 75x for AB2, CB4857, RC301 and CB4856, respectively. Pindel was used to detect structural variations (SVs) in the natural isolates as compared to the N2 reference genome. We included only SVs that had at least 10 reads supporting the SV and no reads supporting the reference genome. Next, we grouped

all SVs from AB2, CB4857, RC301 together as the phylogenetic tree indicate that they are closely related (as compared to CB4856. As a second criterion, we collected only those SVs that were N2-like in CB4856, but were flagged as a SV in the AB2, CB4857 or RC301 group (and vice versa). We found 6,279 and 5,028 deletions of at least 1bp for CB4856 and the combined group that consists of CB4857, AB2 and RC301, respectively. The regression curves displayed in figure 4 were created by iterating over the natural isolate datasets.

A



B

Genotype	# analyzed plates	Mutants found
N2	340	0
polh-1(ok3317)	340	0
polh-1(lf31)	340	0
polk-1(lf29)	340	0
polh-1(ok3317);polk-1(lf29)	740	*dpy(3); ste(1); let(15); him(6)*
polh-1(lf31);polk-1(lf29)	340	*dpy(3); let(5); him(3)*
msh-6	300	20 visible mutants

Figure S1. Occurrence of spontaneous visible mutants in TLS defective strains.
A. Experimental set-up to determine spontaneous mutagenesis: the F1 brood of non-mutant segregating hermaphrodites (P0) were singled to establish individual populations. These were inspected for mendelian segregation of abnormal phenotypes indicating the occurrence of a recessive mutations in the gametes of the P0. Mutants affecting body morphology (e.g. dumpy/dpy) or movement (i.e. uncoordinated/unc) can be scored in the F2 progeny. Mutations in essential genes (i.e. lethal/let) give rise to islands of dead eggs when populations are allowed to clear the food supply. Elevated numbers of males in the F2 progeny indicate a high incidence of males (him) phenotype, arguing for a dominant him mutation in the F1. **B.** Quantification of visible mutant phenotypes. The data for msh-6 mutants have been published previously (Tijsterman *et al.* 2002).

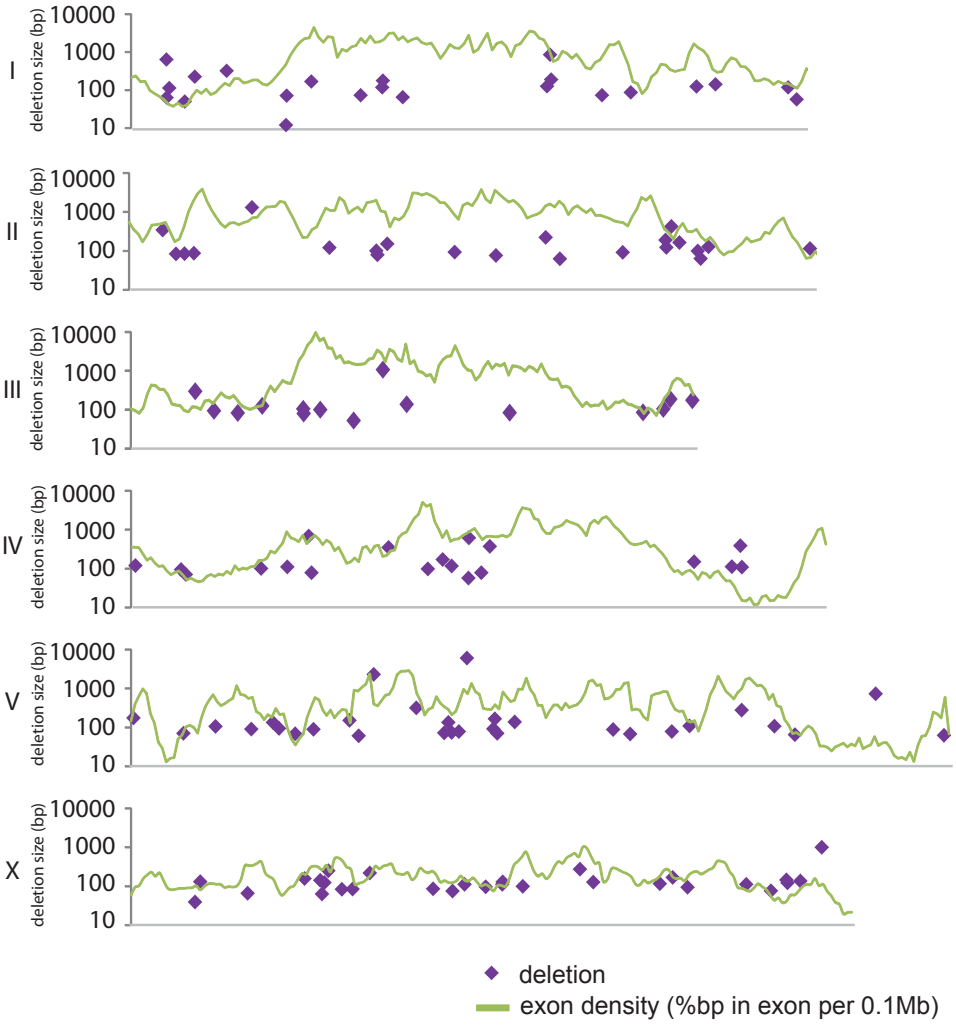


Figure S2. Genomic distribution of deletions in *polh-1polk-1* mutant animals.
Individual deletions (purple) were plotted onto a physical map of the *C. elegans* genome. The y-axisshows the size of the deletion on a logarithmic scale. The exon density is displayed in green (y-axis not shown). The length of the graph shows the size of the indicated chromosomes relative to each other.

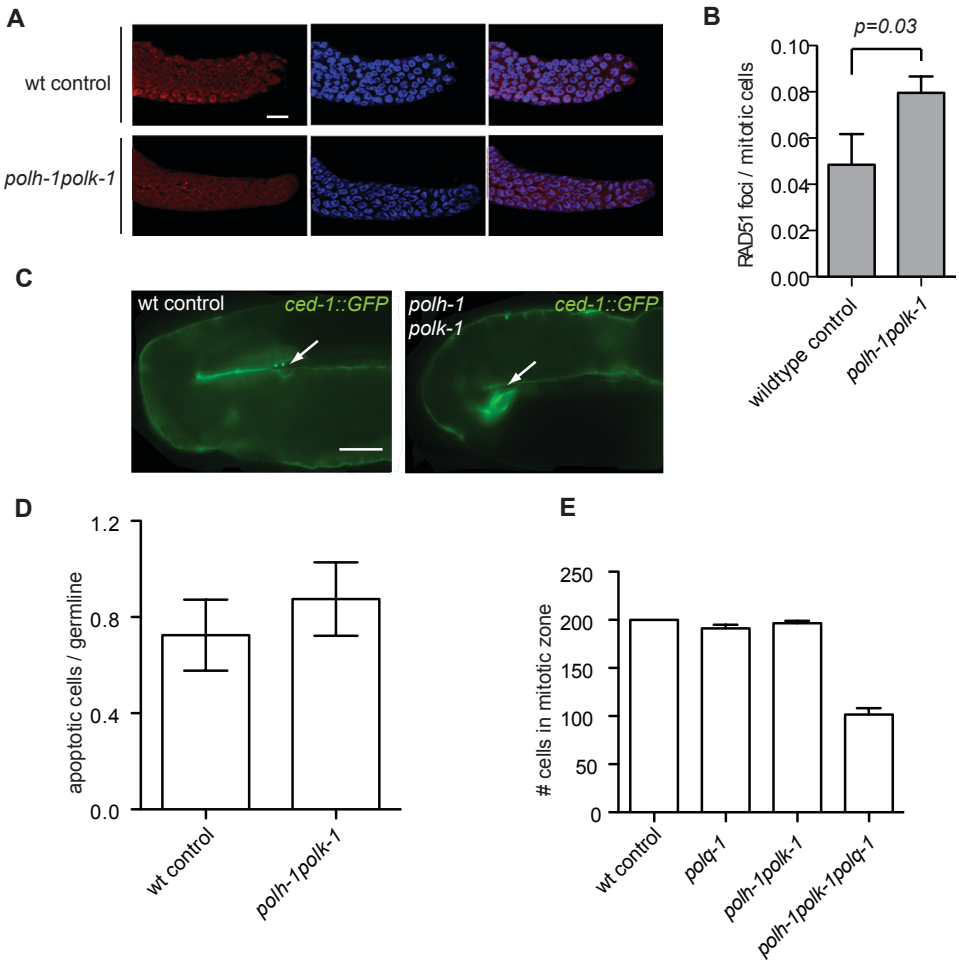


Figure S3. Analysis of checkpoint activation in the *C. elegans* germline.
A. Representative images and **B.** quantification of RAD-51 foci for the indicated genotype in nuclei present in the proliferative compartment of the *C. elegans* reproductive system. DAPI stainings in blue, RAD-51 in red. Scale bar, 10 μm **C.** Representative images of the bend of the gonad arm of animals transgenic for the apoptotic marker ced1::GFP; cells in the process of apoptotic engulfment are indicated with arrows. Scale bar, 10 μm **D.** Quantification of apoptotic cells in *polh-1polk-1* mutant animals and wildtype controls. **E.** Quantification of the number of nuclei in the mitotic region of the germline. A reduction in the number of cells in this region is an established outcome of checkpoint activation.

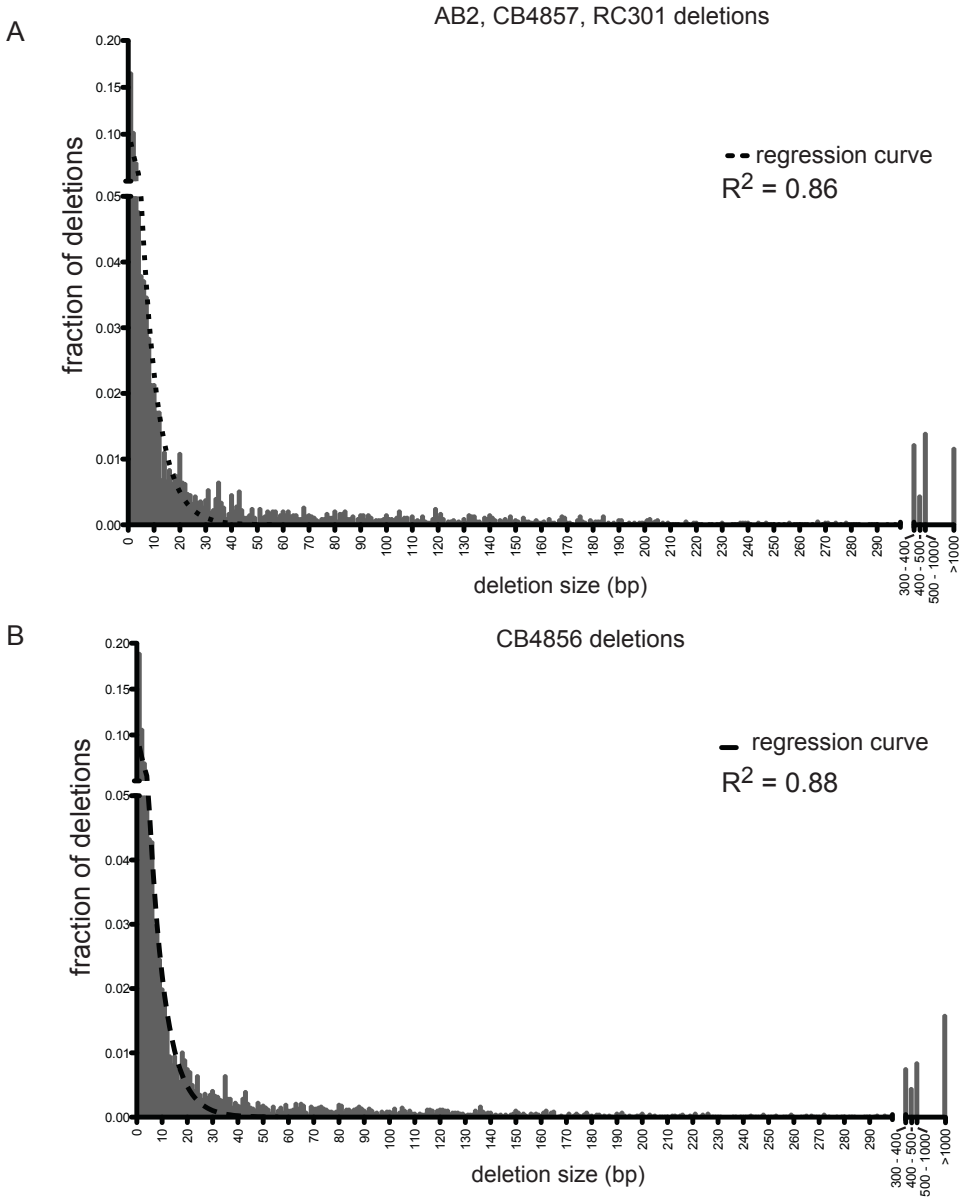
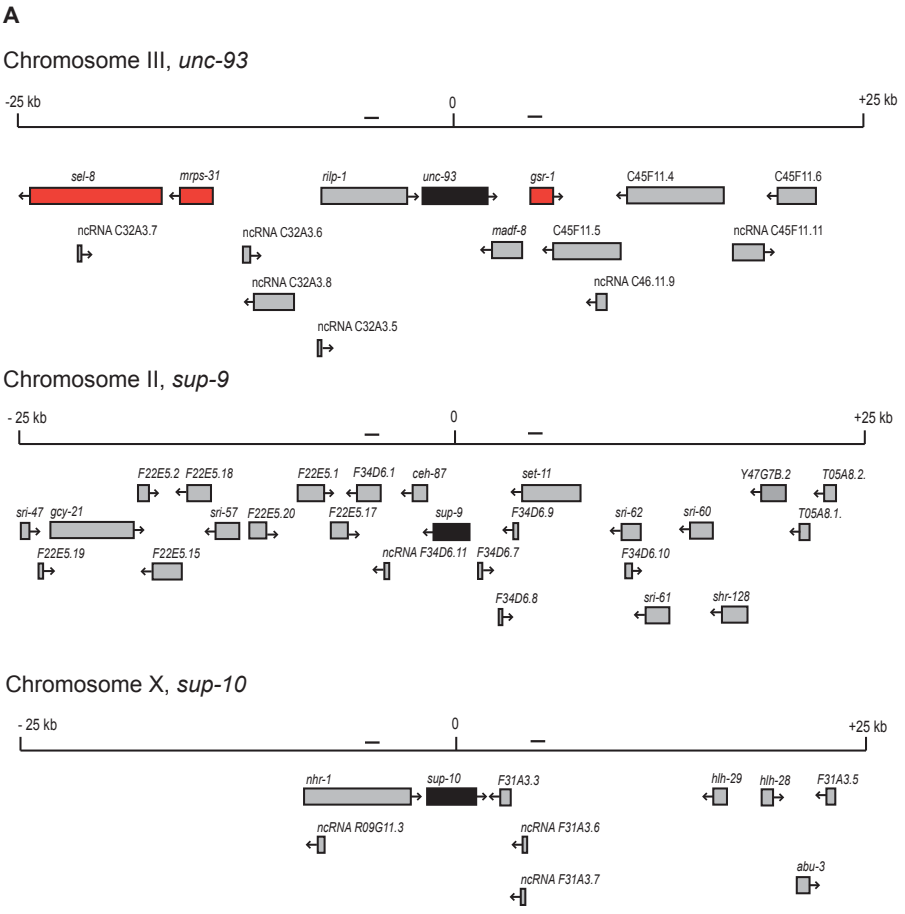


Figure S4. Histogram of size distribution of deletions in various *C. elegans* natural isolates that were analyzed.

Regression analysis showed that an exponential fit for deletion sizes up to 20bp approaches the actual distribution best. **A.** the grouped distribution for AB2, CB4857 and RC301. **B.** as in **A.** but now for CB4856.

A



B

	unc-93	*sup-9*	*sup-10*
wildtype	0	2	11
polq-1	6	20	55

Figure S5. Selective occurence of large chromosomal deletions in regions that are devoid of essential genes in the *unc-93* mutagenesis assay.
A. Schematic representation of 50 kb regions surrounding the *unc-93*, *sup-9* and *sup-10* genes. Known essential genes are depicted in red. While *unc-93* is flanked by two essential genes, no essential genes are known in the 50 kb intervals around *sup-9* and *sup-10*. To estimate deletion sizes, amplification of PCR products at -5kb and +5kb positions has been tested. **B.** Number of deletions larger than 5kb in *unc-93*, *sup-9* and *sup-10*.

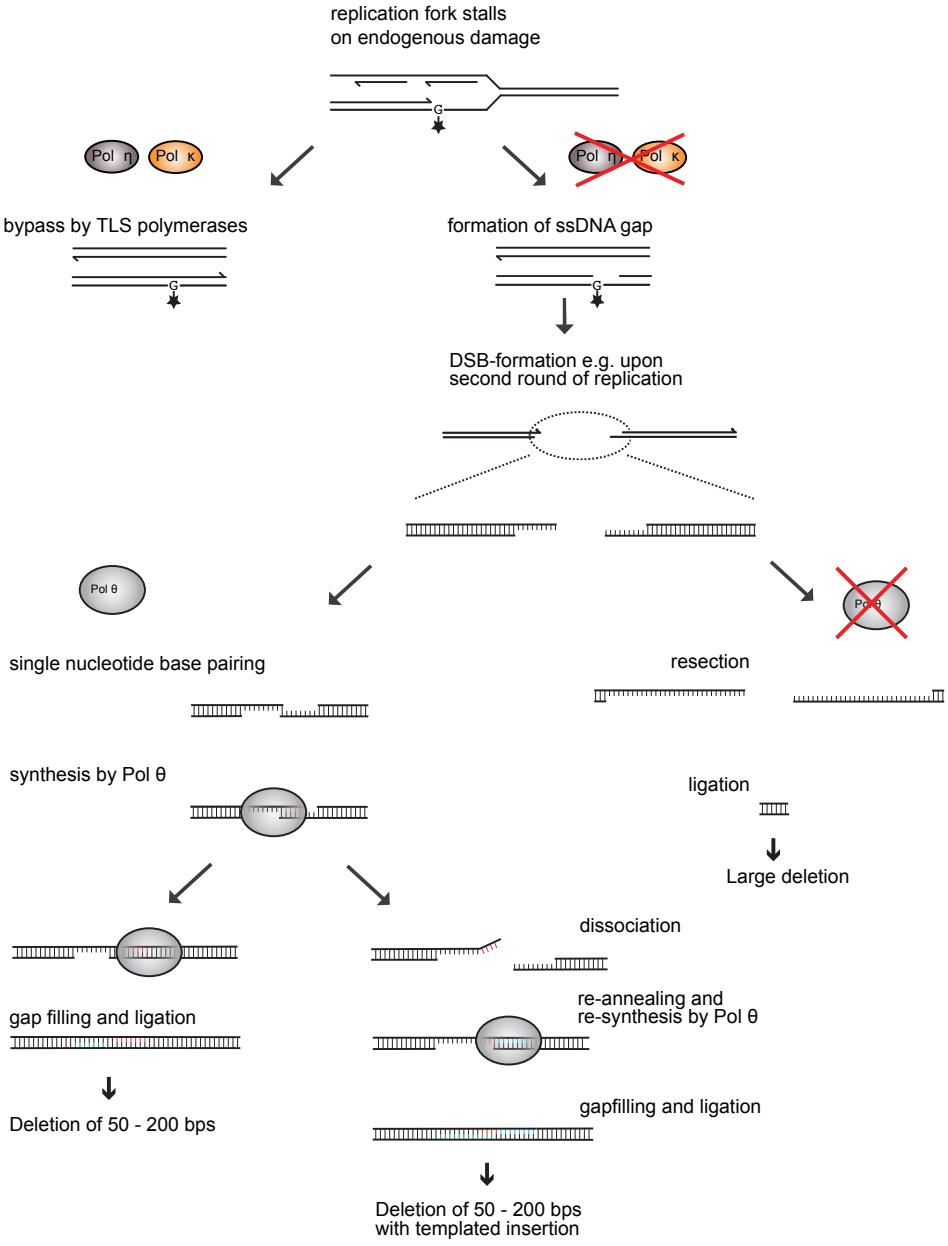


Figure S6. Proposed model for TMEJ of breaks induced at spontaneous replication fork barriers. Endogenous damage - spontaneously, or with increased frequency in the absence of functional TLS - causes replication fork blocks, leading to double stranded breaks. The broken ends can be joined by Pol θ-mediated end-joining (TMEJ) either in a continuous process, giving rise to deletions with a single nucleotide homology (left) or in subsequent steps, giving rise to deletions with templated insertions (right).

Table S1. Whole genome sequencing statistics.

genotype	sample	# generations	# reads	average coverage	# bp >= 4x covered
N2	N2	60	45,258,326	28x	100,140,732
	N4	60	23,693,826	16x	99,675,920
polh-1(lf31)	H7	60	46,203,688	39x	100,229,062
	H8	60	44,982,616	37x	100,238,324
polk-1(lf29)	K1	60	41,517,548	21x	99,970,233
	K4	60	39,275,458	30x	100,235,635
	K9	60	40,037,564	24x	100,120,773
polh-1(lf31);polk-1(lf29)	D4	32	46,284,780	21x	99,911,564
	D13	25	38,712,292	29x	100,224,845
	D14	25	59,163,976	27x	100,202,641
msh-6 (pk2504)	M13	10	48,338,722	19x	99,236,278
	M15	10	44,129,942	12x	99,799,729

Table S2. *unc-22* deletions in *polh-1polk-1* and *polh-1polk-1polq-1*.

	size	left flank	deletion left	deletion right	right flank	insertion
polh-1polk-1						
A	83 bp	GTACCTACTCA	CGTCCAAATG	TTATCGAAAA	GAACGTGTGC	-
B	74 bp	AATCCAGAAGT	CGATGACACC	CTTGGTTAGT	TATTTTTTGG	-
C	153 bp	ACAAGGCTGGG	CCTGGACAAC	TAAAGGCTGG	AGCCACTGTT	-
						AATCTGACTATCAAAGGAAATC
D	119 bp	GACTATCAAGG	CTGGTCAATC	TGATAACCCA	GAATACCAAT	TCAAGAATCTGACTATCAAAG
E	93 bp	CTTGCAAAGGA	TCCATTTGGA	CACGTGACAA	CGGTGGATCA	-
F	71 bp	TGTGAAGCCTT	ACGGAACTGA	ACCACCAGTT	GTTACTTGGC	-
G		not identified				
polh-1polk-1 polq-1						
A	>4.7 kb					
B	>30.5 kb					
C	19 - 20.6 kb					
D	12660 bp	AAATGAGCACA	CTATTCTGTG	GAACAGGAGC	ATTTGGAGTT	
E	> 23.7 kb					
F		not identified				

Table S3. Frequency of *unc-22* mutations in *polq-1*, *polh-1polk-1* and *polh-1polk-1polq-1*.

Strain		total # plates scored	# plates containing one or more twitchers	estimated mutation rate
N2	wildtype	40	0	0.00E+00
XF152	*polq-1*	40	0	0.00E+00
XF507	*polh-1polk-1*	46	7	8.00E-06
XF840	*polh-1polq-1polk-1*	39	6	8.00E-06

Table S4. Sequence analysis of reversion mutants for *unc-93(e1500)*.

wildtype			
unc-93			
deletions > 5kb	0		
substitutions	6	cagttt(g>a)tctggc; C>Y gacacg(t>a)cacagt; V>D tgtctg(g>c)aatact; G>A aaatat(c>t)gatttt; R>L ggaatc(g>a)cggctt; T>A tgttag(g>t)taatgg; splice	
other	1	gaatat(tcga>deleted)aaaactt	3bp > deletion > 12 bp
sup-9			
deletions > 5kb	2		
substitutions	1	ccattg(g>a)gactta; G>stop	
other	2	ccaata(gtga>deleted)cgtcat tctgta(ccgggtgggga>deleted)ggtctg	3bp > deletion > 12 bp 3bp > deletion > 12 bp
sup-10			
deletions > 5kb	11		
substitutions	3	cagttc(t>a)cttgta; L>H tggaat(a>g)tggtcgg; M>V* agccag(g>t)tttgta;; splice site mutation	
unknown	2		

*also tctttt(t>c)caacca in intron 150 bp upstream

Table S5. Sequence analysis of reversion mutants for *polq-1; unc-93(e1500)*.

polq-1			
unc-93			
deletions > 5kb	6		
substitutions	12	tgcgga(c>a)aagtcg; Q>K cgttga(c>a)gattttc; T>K gatctc(g>a)gatctg; G>R ttccat(c>t)atttat; S>L tttcta(c>a)ctcatg; T>N tttcat(g>t)attgta; M>I ggggag(c>a)caaatg; A>D aagtcg(t>a)cggaaa; V>D tccttt(c>t)gagaca; R>stop tctata(c>a)attgtc; Y>stop aatata(t>a)ttgctg; Y>stop tgttag(g>a)taatgg; splice site mutation	
other	2	acgtca(ca>deleted)gttgaa ttttac(t>deleted)ttttag	other microsatellite
sup-9			
deletions > 5kb	20		
substitutions	7	tcttcg(g>a)gctcac; G>E gggtac(c>a)agtgga; Q>K gtggag(c>a)atttta; A>E ccattg(g>a)gactta; G>stop aggcta(c>a)ggtcat; Y>stop tccctg(c>t)aaactc; Q>stop caagta(c>a)aacatg; Y>stop	
sup-10			
deletions > 5kb	55		
substitutions	1	atgtta(a>t)tataag; N>I	
other	1	gtgatg(a>deleted)catcaa	hairpin
unknown	7		

POL θ - MEDIATED END JOINING (TMEJ) OF TRANSPOSITION-INDUCED DOUBLE STRAND BREAKS IN *C. ELEGANS* GERM CELLS

4

Sophie F. Roerink, Karin Brouwer, Marcel Tijsterman

Leiden University Medical Center,
Department of Toxicogenetics

ABSTRACT

Double strand breaks (DSBs) are potentially harmful lesions in the DNA that can be repaired, either by error-free homologous repair (HR) or error-prone end joining (EJ) mechanisms. While canonical non-homologous end joining (NHEJ) is responsible for most EJ events in organisms, also alternative EJ pathways have been described to contribute to mutagenic and possibly tumorigenic repair events.

In the previous chapter, EJ dependent on the A-family Polymerase θ and therefore denoted as Pol θ - mediated end joining (TMEJ) has been identified as a prominent route for repair of stalled replication forks in the *C. elegans* germline, most likely via a DSB intermediate. Here, we use a transposition-based assay to study EJ in the *C. elegans* germline. We analyzed EJ repair for two Tc1-mediated DNA breaks which have categorically different flanking sequences: i) six basepairs of perfect homology, ii) no overt micro-homology. Strikingly, efficient repair of both types of breaks was not dependent on canonical NHEJ factors, but highly dependent on Polymerase θ, identifying TMEJ as the major route for error-prone repair of transposon-induced breaks in the *C. elegans* germline. Using a large collection of footprints generated by TMEJ action, we present a detailed mechanistic model for this DSB repair pathway, in which opposing ends of a DNA break are joined together via one or more cycles of strand extension that is initiated by base pairing of a single nucleotide.

INTRODUCTION

DNA double strand breaks (DSBs) are extremely toxic to cells, because they can lead to sequence loss, genomic rearrangements, aneuploidy and cell death (Hoeijmakers 2001). To counteract these detrimental effects, multiple repair pathways have evolved (Wyman and Kanaar 2006). Breaks can be repaired in an error-free way by homologous repair (HR), which predominantly makes use of the undamaged sister chromatid as a repair template. Alternatively, non-homologous end joining (NHEJ) ligates the ends of the break site without the use of a repair template. Because this process frequently goes together with loss of insertion of a few nucleotides, NHEJ repair is grosso modo error-prone. The choice between these two repair modes largely depends on cell type and stage of the cell cycle: in non-dividing mammalian cells, NHEJ is the predominant repair route, while in cycling yeast cells HR is the principal route to repair DSBs (Shrivastav *et al.* 2008). In *C. elegans*, LigaseIV-dependent NHEJ dominates repair of ionizing irradiation-induced DSBs in non-dividing somatic cells, while proliferating somatic cells and germline cells use HR to repair these breaks (Clejan *et al.* 2006).

In addition to canonical NHEJ, several lines of evidence state the existence of EJ mechanisms independent of the classical factors LigaseIV and Ku, commonly classified as alternative end-joining (Davies *et al.* 2008; McVey and Lee 2008). In our lab, we have observed efficient repair of an endonuclease-induced break in a transgenic *C. elegans* strain that was defective for all canonical DSB repair activities, suggestive of efficient repair activities other than NHEJ or HR (Pontier and Tijsterman 2009).

Factors implicated in alternative end-joining processes include poly(ADP-ribose) polymerase-1 (PARP-1) and Ligase III (Audebert *et al.* 2004; Wang *et al.* 2006). Since microhomology of a few nucleotides at both ends of the break has been shown as a determinant for the repair outcome, alternative end-joining processes are also often referred to as microhomology-mediated end joining (MMEJ)(McVey and Lee 2008).

Recently, we identified Pol θ - mediated end-joining (TMEJ) as an important contributor to error-prone repair processes in *C. elegans*. We found that TMEJ acts in two different backgrounds that increase replication fork stalling in the germline: i) knocking out the *C. elegans* FANCJ homolog which is necessary for unwinding stable G4 structures in the DNA (Koole et al, submitted for publication) ii) knocking out translesion synthesis polymerases which can bypass damaged DNA during replication (Chapter 3, this thesis). In both cases, deletions from a distinct size class (50-300 bp) appeared, either at the location of G4 sequences (i) or randomly distributed over the genome (ii). This mutational signature is further characterized by a predominance of homology of a single nucleotide at the deletion junction or by small insertions templated from flanking sequences, and was completely dependent on the presence of the A-family member Polymerase θ. Importantly, this mutational signature is also observed in comparisons of evolutionary distinct natural isolates of *C. elegans*, suggestive of a central role for TMEJ in genome change during evolution (Chapter 3, this thesis). We propose that prolonged

replication fork stalling in the *C. elegans* germline results in single strand gaps, that will cause DSBs in the next replication round, that are subject to TMEJ.

To address the question whether TMEJ can act on DSBs directly, we study repair of a natural source of DSBs in *C. elegans*: the excision of Tc1 transposons. The Tc1/mariner family of transposable elements is probably the most widespread family of transposons in nature: members have been found in fungi, plants and animals (Plasterk *et al.* 1999). They contain a single gene, encoding a transposase, flanked by terminal inverted repeats (TIRs). The Tc1 element is liberated from its host genome via a staggered cut at the terminal ends of the TIR made by the cognate transposase and can be reinserted into target DNA at a TA dinucleotide target site, resulting in spreading over the genome. The excision of the Tc1 element by the transposase results in a DSB with 3′ overhangs of two nucleotides at either side.
In wildtype *C. elegans* strains, excision of Tc1 is restricted to somatic cells (Emmons and Yesner 1984), where break repair is performed by Ligase IV (*lig-4*) dependent NHEJ (Clejan *et al.* 2006; Davies *et al.* 2008). Tc1 transposons are kept silent in the germline of many strains of *C. elegans* by an RNAi-related genome protection mechanism (Bessereau 2006). Germline transposition can, however, be found in some wild strains, and can be activated in silenced strains by mutating genes involved in silencing mechanisms. These genes were originally termed mutator genes, because mutants of these genes have increased levels of (transposon-related) mutagenesis (Ketting *et al.* 1999; Tabara *et al.* 1999).
Tc1-induced breaks in germ cells are predominantly repaired by error-free HR, resulting in a net increase in copy number of the Tc1 element (Plasterk 1991). However, loss of the element has also been observed, suggesting activity of an error-prone EJ mechanism (Plasterk 1991). Similar results were obtained with Mos1 excision - a Tc1 member derived from flies (Robert and Bessereau 2007). Surprisingly, non-homologous repair of Mos1 excision sites in germ cells is not dependent on the homologues of classical NHEJ genes Ku80 and Ligase IV (Robert and Bessereau 2007). A strong contribution of a non-canonical EJ mechanisms is also found in other metazoans: in the fruitfly *D. melanogaster*, EJ after excision of a transposable P element is not dependent on Ligase IV (McVey *et al.* 2004).

Here, we study genetic requirements for repair of two Tc1-induced break sites in the *unc-22* muscle gene, either having six nucleotides of microhomology around the break site, or having no overt microhomologous sequences. We found that DSB repair by EJ is strongly directed by the sequence context of the break, and almost completely dependent on the presence of functional Pol θ, demonstrating that TMEJ acts as a major pathway in the *C. elegans* germline for error-prone repair of DSBs.

RESULTS

Microhomology directs end-joining in the *C. elegans* germline

To study DSB repair in the *C. elegans* germline, we combined Tc1 alleles in the muscle gene *unc-22*, which disrupt the open reading frame leading to a phenotypically easily recognizable "twitching" phenotype, with a mutator mutation that desilences germline transposition. Error-prone break repair of DSBs resulting from Tc1 excision can restore the *unc-22* ORF, and thus to wildtype movement of animals derived from these germ cells (Figure 1).

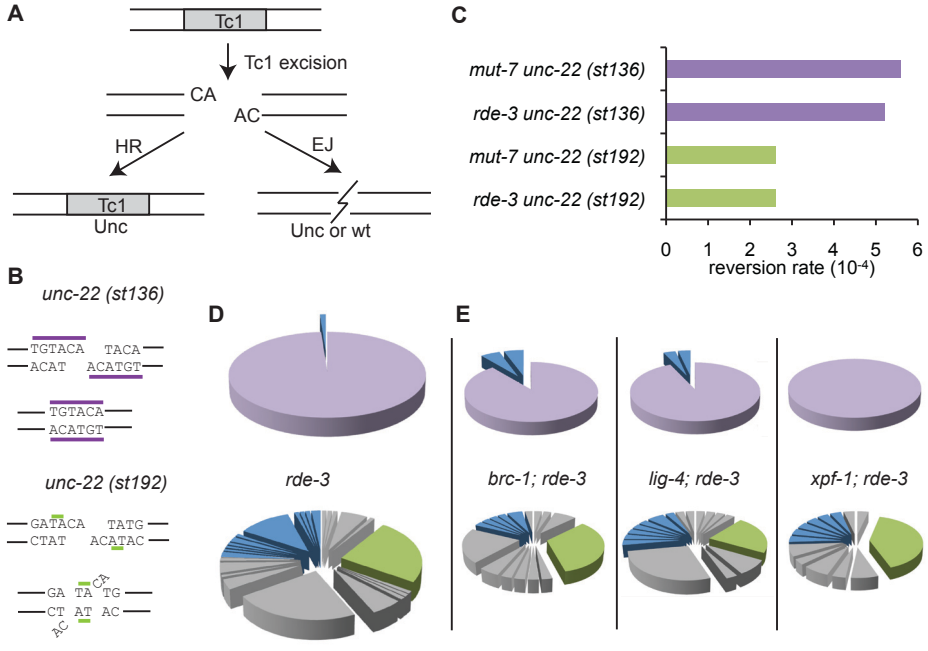


FIGURE 1. Germline repair of transposon-mediated DSBs.
A. Mechanism of Tc1-mediated DSB repair. The transposase Tc1A excises Tc1 transposons resulting in staggered cuts with 3' overhangs of two nucleotide. Such breaks can be repaired in an error-free way via HR or in an error-prone way via EJ. Only in case of ORF correction of the *unc-22* gene by error-prone repair, animals will revert to wildtype movement. **B.** The sequence context of the breaks that result from transposon excision at the *unc-22(st136)* allele displaying six basepairs microhomology flanking the break (upper panel), and the *unc-22(st192)* allele displaying maximally two nucleotides of possible microhomology (lower panel). **C.** Reversion rates of *unc-22(st136)* and *unc-22(st192)* alleles in two different mutator strains (*rde-3* and *mut-7*) that lost transposon silencing. **D.** Distribution of reversion events in *rde-3; unc-22(st136)* (upper panel) and *rde-3; unc-22(st192)* (lower panel). Almost all reversion events of the *unc-22(st136)* allele show footprints of repair mediated by six basepair microhomology, depicted in purple. A diverse spectrum of footprints is displayed for the *unc-22(st192)* allele; the footprint that results from the constitution as depicted in B is colored in green. Flank duplication events that are predominantly found in the *st192* allele are indicated in blue. **E.** Distribution of reversion footprints in animals defective for *brc-1*, *lig-4* and *xpf-1* respectively is not significantly altered as compared to wildtype.

To investigate a potential role for microhomology in the repair of germline DSBs, we made use of two transposon insertions: *unc-22 (st136::Tc1)* and *unc-22 (st192::Tc1)* which have categorically different flanking sequences (Figure 1). Whereas excision of Tc1 in *st192* generates a break with no overt microhomology flanking the non-complementary staggered cut, excision of Tc1 in *st136* creates a break in which the outer six nucleotides at the 3′ end on both sides of the break are perfectly complementary. Use of these six nucleotides in microhomology-directed repair will lead to *unc-22* ORF correction, and thus to reversion of the Unc phenotype (Figure 1B). A mutation in the mutator gene *mut-7* was crossed into both *unc-22:Tc1* alleles, to allow Tc1 excision in germline tissue (Ketting *et al.* 1999). The reversion frequency for the *st136:Tc1* strain was about twofold higher than for *st192:Tc1* (Figure 1C). Similar frequencies were obtained when another mutator locus *rde-3* was mutated to confer Tc1 jumping.

We next determined the molecular nature of ~100 independently derived revertants of either allele and found that the sequence context is a critical determinant in DSB repair: 94 out of 95 reversions of *st136::Tc1* (six nt homology) show loss of Tc1 as well as one copy of the six nt homologous stretch that was present in both flanks of the break. This outcome strongly suggests that microhomology-mediated repair dominates (table 1).

In sharp contrast, the reversion spectrum for *st192::Tc1* (no overt microhomology) is highly variant, showing 26 different footprints in 103 analyzed animals (Table 1).

In 25 out of 103 footprints we observed that loss of Tc1 was associated with DNA insertions of up to 38 nts. The sequence inserted at the break site is similar, if not identical, to the DNA sequence immediately flanking the break and the majority of the insertions were below ten basepairs. Three larger inserts contained two or more copies of the adjacent DNA sequence. Another prominent class of footprints (36 out of 103) comprised of 'simple' deletions where Tc1 was lost as well as a limited number of flanking basepairs. Remarkably, 34 out of these 36 cases contained at the junction a single nucleotide that could have originated from either flank, a feature we have termed single nucleotide homology (Chapter 3). A final dominant class of footprints (24 out of 103) had lost the Tc1 element and had four nucleotide inserts that comprised the nucleotide overhangs present at both sides of the break. Seven different footprints made up the residual 16 cases, which did not easily classify into the above-mentioned categories. Here, Tc1 loss was associated with both loss and gain of a few nucleotides at the break site. The molecular mechanisms operating to generate these diverse repair products will be discussed later.

End-joining in the *C. elegans* germline does not depend on canonical DSB repair factors or PARP-1.

To analyze the pathways involved in EJ of Tc1-induced breaks, we studied functional knockouts in different canonical DNA repair proteins. In line with Plasterk and colleagues, who showed that HR is able to repair Tc1-induced DSBs in germline tissue,

TABLE 1. Repair after excision of a Tc1 element at two different positions in *unc-22*.

rde-3 (ne298): unc-22 (st136)			
	AGATTGACGAGATCCATAAGGAAGGATGTACA		TACATTGAACTGGAAGCCTCCAACTG
	TCTAACTGCTCTAGGTATTCCTTCCTACAT		ACATGTAACTTGACCTTCGGAGGTTGAC
Class I: deletions			
94	AGATTGACGAGATCCATAAGGAAGGATGTA		CATTGAACTGGAAGCCTCCAACTG
Class III: flank duplications			
1	AGATTGACGAGATCCATAAGGAAGGATGT	TTGAACTGGAAGCCTCCA	ACATTGAACTGGAAGCCTCCAACTG
rde-3(ne298): unc-22 (st192)			
	AGACGACGGTGGTTCTCCAATTTTGGGATACA		TATGTCGTTGAACGTTTTGAGAAGAG
	TCTGCTGCCACCAAGAGGTTAAAACCCTAT		ACATACAGCAACTTGCAAAACTCTTCTC
Class I: deletions			
1	AGACG		*(54 bp deletion)*
1	AGACGACGGTGGTTCTCCAATT		GAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGG		TCGTTGAACGTTTTGAGAAGAG
6	AGACGACGGTGGTTCTCCAATTTTG		TATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTTGGG		TATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGA		GTTGAACGTTTTGAGAAGAG
Class IA: 2bp deletions			
25	AGACGACGGTGGTTCTCCAATTTTGGGATA		TGTCGTTGAACGTTTTGAGAAGAG
Class II: indels (≤ 4bp)			
1	AGACGACGGTGGTTCTCCAATTTTGGG		GTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTTGGGA		TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATAC		TATGTCGTTGAACGTTTTGAGAAGAG
6	AGACGACGGTGGTTCTCCAATTTTGGGATACA		ATGTCGTTGAACGTTTTGAGAAGAG
4	AGACGACGGTGGTTCTCCAATTTTGGGATACA		TATATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATACG		TGTATGTCGTTGAACGTTTTGAGAAGAG
4	AGACGACGGTGGTTCTCCAATTTTGGGATATA		TGTATGTCGTTGAACGTTTTGAGAAGAG
Class IIA: 2bp insertions			
24	AGACGACGGTGGTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAGAAGAG
Class III: flank duplications			
1	AGACGACGGTGGTTCTCCAATTTTGGGATACA	GGA	TGTATGTCGTTGAACGTTTTGAGAAGAG
2	AGACGACGGTGGTTCTCCAATTTTGGGATAC	TTTGGGA	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATACA	ATTTTGG	TGTATGTCGTTGAACGTTTTGAGAAGAG
12	AGACGACGGTGGTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGAT	TTTGGGATTTTGGAT	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATAC	GATATTTGGGATA	TGTATGTCGTTGAACGTTTTGAGAAGAG
2	AGACGACGGTGGTTCTCCAATTTTGGGATA	TTTTGG	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATACA	ATTTTGGGATTTTTTTTG GGATTGTTGATTTTGGGA	TGTATGTCGTTGAACGTTTTGAGAAGAG
2	AGACGACGGTGGTTCTCCAATTTTGGGATACA	TGTCGT	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATACA	TGTCGTTGT	TGTATGTCGTTGAACGTTTTGAGAAGAG
1	AGACGACGGTGGTTCTCCAATTTTGGGATACG	TATGTCGTT	TGTATGTCGTTGAACGTTTTGAGAAGAG

we found that in mutants for the HR gene *brc-1* - the worm homolog of mammalian breast cancer gene BRCA1 - germline transposition compromised the growth of animals (table S1).
However, error-prone repair leading to *unc-22* ORF correction is unlikely the result of this HR pathway as the distribution of repair footprints is unaffected by the *brc-1* status (Figure 2E, Supplementary information).
Classical NHEJ would be the obvious candidate to explain most of the junctions observed for *st192::Tc1*. However, the junction profile of *st192::Tc1* nor that of *st136::Tc1* was affected by knocking out classical NHEJ factor *lig-4* (Figure 2E and supplementary information).
We next addressed a possible involvement of SSA components. Although SSA tends to use larger stretches of homology (Sugawara *et al.* 2000), the genetic requirements for certain molecular steps may be similar for SSA and the micro-homology-mediated repair of *st136::Tc1*. The endonuclease XPF/ERCC1 has been shown to be involved in the removal of the 3′ flap during SSA. We asked whether microhomology-mediated repair of *st136::Tc1* was dependent on the presence of the *C. elegans* homologs of the XPF and ERCC1 endonucleases. Yet, also in *xpf-1* and *ercc-1* mutants the *st136::Tc1* strain shows almost exclusively microhomology-mediated repair junctions, while also for *st192:Tc1* we found the distribution of the repair events over the different classes allele largely unchanged (Figure 2E and supplementary information).
Thus, knockout of canonical DSB repair pathways (i.e. HR, NHEJ and SSA) did not influence the frequency or the distribution of the EJ footprints.
Several studies have pointed towards a role for PARP proteins in alternative end joining processes (Audebert *et al.* 2004; Wang *et al.* 2006; Mansour *et al.* 2008). Numeral PARP like proteins are present in eukaryotes, although PARP-1 is responsible for the majority of poly(ADP)ribose production (Krishnakumar and Kraus 2010). In the worm six PARP homologs are known, of which *pme-1* shares most homology to PARP-1 (Gagnon *et al.* 2002). To test whether PARP-1-mediated end joining could be responsible for the spectrum of junctions observed, we crossed a deletion allele of *pme-1(ok988)* in our mutator strains. However, also here, we found the distribution of repair events to be unchanged as compared to wildtype (supplementary information). Although we formally cannot exclude the possibility that the other *pme* genes have redundant roles in alternative EJ pathways in *C. elegans*, unpublished data from our lab using an IsceI-based reporter assay do not support this hypothesis for the two most conserved PARP members *pme-1* and *pme-2*.

Efficient break repair is dependent on Pol θ

Characteristics of the repair footprint of Tc1 induced breaks - flank insertions and use of a single nucleotide homology - showed similarity to Pol θ-mediated end-joining (TMEJ) - a recently identified mechanism for repair of DSBs arising at replication blocks in the

C. elegans germline. We therefore tested repair efficiency and outcomes of Tc1-induced breaks in *polq-1* mutant worms - lacking the worm homolog of Pol θ.
Strikingly, in the absence of a functional *polq-1* gene both *unc-22(st136)* and *unc-22(st192)* animals hardly reverted to wildtype movement (Figure 2). Reversion frequencies dropped about 23-fold for *unc-22(st192)* and 15-fold for *unc-22(st136)* (Figure 2A), indicating that Pol θ is required for DSB repair of Tc1-induced breaks in *C. elegans* germ cells. We cloned out *polq-1 unc-22* mutants on a large scale to determine the

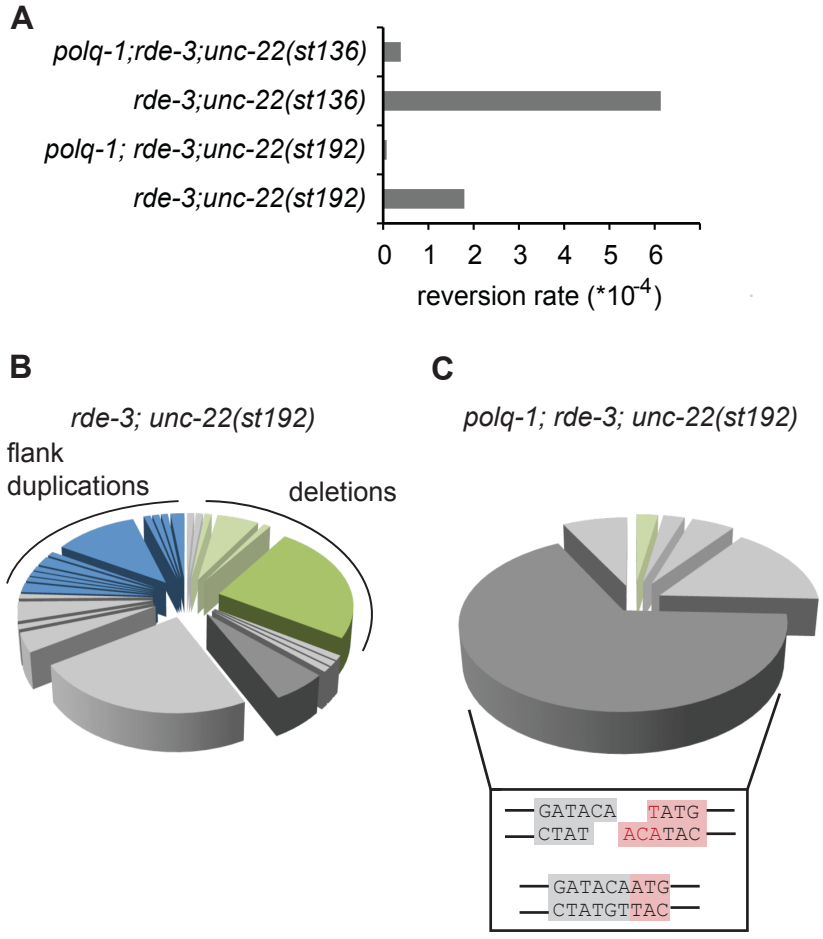


FIGURE 2. *polq-1* deficiency suppresses *unc-22* reversion.
A. Reversion frequency of *unc-22(st136)* and *unc-22(st192)* alleles in *rde-3* and *polq-1;rde-3* background. **B.** Distribution of footprints in *rde-3;unc-22(st192)* mutants. Footprints that display flank duplications are indicated in blue; deletion footprints displaying ≥1 nt microhomology around the break are in indicated in green. **C.** Distribution of footprints in *polq-1; rde-3;unc-22(st192)* mutants. The major footprint (in grey) is drawn out below the chart.

footprints of repair events taking place in the absence of *polq-1*. For the *unc-22(st192)* allele we were able to determine the molecular nature of 39 junction footprints; none of these displayed nucleotide insertions. Also the other footprints that dominated the spectrum in *polq-1* proficient conditions were absent or greatly reduced, and only 1 out of 39 footprints displayed single nucleotide homology, a feature that dominated the spectrum in wildtype animals. Instead, 32 out of 39 *unc-22* revertants isolated in *polq-1* mutants showed a footprint that was only occasionally seen in *polq-1* proficient conditions (Figure 2B): while the CA overhang is retained on one of the flanking sides, a single base pair has been deleted from the other side.
Although the reversion frequency of *unc-22(st136)* was also greatly affected in *polq-1* deficient animals, the change in the profile of repair junctions was far less dramatic: all 34 footprints analyzed showed loss of the six nucleotides homology around the break (supplementary information), arguing that both a Pol θ-dependent as well as a Pol θ-independent mechanism of DSB repair will use the presence of such homology if present.

Other factors tested for EJ repair in *C. elegans* germ cells

In yeast, homologs of translesion synthesis (TLS) polymerases pol η and pol ζ have been implicated in microhomology-mediated end repair (Lee and Lee 2007). We crossed in null mutants for *polh-1* as well as a hypomorphic allele of *rev-1*, containing a mutation in the BRCT domain of the REV1 protein, which operates as a key factor in TLS processes (chapter 5 of this thesis). We found that reversion frequencies and junction profiles were not affected in mutants for the worm homologs of Y-family polymerases Polη and Rev-1.
In addition to polymerase activity, some deletion footprints suggest that 3′-overhanging ends are subject to limited trimming, which may require separate exonuclease activity. Exonuclease 1 (Exo1) is a versatile nuclease that acts both during end resection in HR and the excision step in mismatch repair (Genschel and Modrich 2003; Eid *et al.* 2010). Plasmid repair assays in *S. pombe* and *S. cerevisiae* suggested that Exo1 might also affect end joining processes (Decottignies 2007; Bahmed *et al.* 2011). However, junction profiles of *unc-22(st192)* were not affected by an *exo-1* mutation (supplementary information).
In addition, we wondered whether DNA mismatch repair might influence DSB repair since repair footprints incidentally show misinsertions in duplicated flanks (Table 3). However, we did not observe any effect on reversion frequency or junction profiles of *unc-22(st192)* in animals with a mutation in the mismatch repair gene *mlh-1* (although more subtle effects may be missed due to the limited number of events) (supplementary information).
So far, *polq-1* is the only genetic factor found, which affected EJ of Tc1-induced breaks in the *C. elegans* germline.

DISCUSSION

A molecular mechanism for Pol θ-mediated End Joining (TMEJ) of DSBs in *C. elegans* germ cells

In the present study we used tranposon-induced breaks as model substrates to identify a prominent role for Pol θ in error-prone DSB repair in the germline of *C. elegans*. While well-known DSB pathways, such as HR, SSA and NHEJ had no effect on error-prone repair of Tc1-induced breaks, knockout of Pol θ lead to a 15 to 33-fold reduction in DSB repair-dependent ORF correction of the endogenous reporter gene *unc-22*. We analyzed two break sites with categorically different sequence contexts: *st136::Tc1* which is characterized by an identical six nt DNA sequence present at either side of the break, and which use in repair would restore the *unc-22* ORF, and another, *st192::Tc1*, that was devoid of any overt sequence homology surrounding the break. We found that the six nt homologous sequence guided repair in both Pol θ - dependent and independent DSB repair routes. In contrast, the various Pol θ -dependent footprints isolated at *st192::Tc1* allowed us to propose a detailed molecular mechanism for TMEJ. To increase the resolution we included the footprints derived from all Pol θ - proficient genetic backgrounds and compared them to *polq-1* mutant lines.

TABLE 2. Repair after excision of a Tc1 element at two different positions in *unc-22* in *polq-1* mutants.

polq-1; rde-3 (ne298); unc-22 (st136)		
	AGATTGACGAGATCCATAAGGAAGGATGTACA	TACATTGAACTGGAAGCCTCCAACTG
	TCTAACTGCTCTAGGTATTCCTTCCTACAT	ACATGTAACTTGACCTTCGGAGGTTGAC
Class I: deletions		
34	AGATTGACGAGATCCATAAGGAAGGATGTA	CATTGAACTGGAAGCCTCCAACTG
polq-1; rde-3(ne298); unc-22 (st192)		
	AGACGACGGTGGTTCTCCAATTTTGGGATACA	TATGTCGTTGAACGTTTTGAGAAGAG
	TCTGCTGCCACCAAGAGGTTAAAACCCTAT	ACATACAGCAACTTGCAAAACTCTTCTC
Class I: deletions		
1	AGACGACGGTGGTT	*(57 bp deletion)*
1	AGACGACGGTGGTTCTCCAATTTTGGGAT	ATGTCGTTGAACGTTTTGAGAAGAG
Class II: indels ≤ 4 bp		
2	AGACGACGGTGGTTCTCCAATTTTGG	TGTATGTCGTTGAACGTTTTGAGAAGAG
6	AGACGACGGTGGTTCTCCAATTTTGGGAT	TGTATGTCGTTGAACGTTTTGAGAAGAG
26	AGACGACGGTGGTTCTCCAATTTTGGGATACA	ATGTCGTTGAACGTTTTGAGAAGAG
class IIA: 2bp insertions		
3	AGACGACGGTGGTTCTCCAATTTTGGGATACA	TGTATGTCGTTGAACGTTTTGAGAAGAG

Two features of *Tc1::st192* footpints are almost completely dependent on the presence of functional Pol θ: I) single nucleotide homology and II) the presence of templated insertions (Figure 2, Tables 1 and 2). We characterized 152 deletions without associated insertions, of which all but one had at least a single nucleotide that could have been derived from either side of the break (Table 3). Over 99 percent of the deletions were thus characterized by single nucleotide homology, greatly exceeding the ~45 percent expected based on a random distribution (chapter 3 of this thesis). While we found deletions up to 100 basepairs, the predominant fraction was below five basepairs (Table 3), which suggests that the ends are not subject to extensive trimming by exonucleases before an EJ mechanism repairs the break.
The same rules, i.e. use of one nt homology and stable 3′ ends, applied to the 70 templated insertions that have been isolated from various Pol θ - proficient backgrounds (Table 4). Most inserts seemed to be derived from the left flank of the break site: 55 cases over 13 cases in which the right flank was used as a template. Nevertheless, insertions from the right flank may be underrepresented, as in total 80 cases of class IIA, in which four nucleotide inserts comprised the nucleotide overhangs present at both sides of the break, may also be explained by an insertion of three nt templated from the right flank (table 4). In two cases sequences derived from both flanking sides were found.

We envisage the following scenario for TMEJ of Tc1-induced DSBs: upon physical contact of the 3′overhanging tails or upon invasion of one 3′end in the other side's double stranded end, a single basepairing nucleotide is sufficient to trigger the synthesis of a nucleotide tract by Pol θ**.** The continuous tract of duplexed DNA, generated by Pol θ, is sufficiently stable to trigger gap closure and subsequent sealing by ligases - other than the canonical EJ factor LigIV. A recent paper by Simsek and coworkers demonstrates the involvement of both LigIII and LigI in alternative pathways of break repair leading to chromosomal translocations in mammalian cells (Simsek *et al.* 2011). Interestingly, LigIII-deficient cells lose their bias towards use of microhomology at break points, while LigI-deficient cells display less insertions. Possibly both alternative ligases play a role in TMEJ processes, generating different repair junctions.

A single round of annealing, extension and ligation explains class I deletions, which are characterized by single nucleotide homology. Base pairing of a single nucleotide is sufficient but also necessary for Pol θ to trigger repair (Table 3 and Chapter 3).
However, the molecular nature of the templated inserts suggests that the extended 3′ end of one flank (by templated synthesis) may be considered as an intermediate structure with two fates: i) as described above, both ends are extended and newly generated DNA is ligated to parental DNA or ii) the extended strand dissociates, resulting in a new 3′end that again searches for a single nucleotide homology.
Careful inspection of the templated inserts points out that almost all templated inserts can be explained by iterative cycles of strand annealing, single nucleotide - driven elongation,

dissociation, and reannealing (Table 4, Figure 3). We observed footprints suggestive of one to even four of such cycles. What exactly determines whether one round is sufficient to repair the break or why retries are required is unknown, but perhaps repair can only be finished if the 3′ overhanging ends of both strands simultaneously pair and thus can be extended. Dissociation of the intermediate may follow if the opposite ends cannot be extended by Pol θ (*e.g.* if it has a mismatched end). Figure 3C exemplifies how cycles of Pol θ - mediated strand extension explain a complex combinatorial footprint isolated in an *xpf-1* defective background.

TABLE 3. Class I deletions in *unc-22(st192)* mutants.

	***unc-22 (st192)*: break left after Tc1 excision**		
	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG		
	GGTGGTTCTCTGCTGCCACCAAGAGGTTAAAACCCTAT ACATACAGCAACTTGCAAAACTCTTCTCTCCACCGC		
			genotype
1	CC	*(99 bp deletion)*	*lig-4*
1	CCACCAAGAGACG	GTGGCG	wt
1	CCACCAAGAGACGACGGT	TTTGAGAAGAGAGGTGGCG	*polh-1*
1	CCACCAAGAGACGACGGTG	AACGTTTTGAGAAGAGAGGTGGCG	*exo-1*
1	CCACCAAGAGACGACGGTGG	CG	*rev-1BRCT*
1	CCACCAAGAGACGACGGTCGTT	TTGAGAAGAGAGGTGGCG	*polh-1*
1	CCACCAAGAGACGACGGTCGTT	TTGAGAAGAGAGGTGGCG	*exo-1*
1	CCACCAAGAGACGACGGTGGTTCT	GAACGTTTTGAGAAGAGAGGTGGCG	*lig-4*
1	CCACCAAGAGACGACGGTGGTTCTC	GTTTTGAGAAGAGAGGTGGCG	*brc-1*
1	CCACCAAGAGACGACGGTGGTTCTCCA	CGTTTTGAGAAGAGAGGTGGCG	*lig-4*
1	CCACCAAGAGACGACGGTGGTTCTCCAATT	GAACGTTTTGAGAAGAGAGGTGGCG	wt
1	CCACCAAGAGACGACGGTGGTTCTCCAATTT	ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*brc-1*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTT	ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*rev-1BRCT*
6	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	wt
3	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*brc-1*
6	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*ercc-1*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*pme-1*
2	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*exo-1*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*polh-1*
3	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGG	TCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*rev-1BRCT*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGA	GTTGAACGTTTTGAGAAGAGAGGTGGCG	wt
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTGAGAAGAGAGGTGGCG	*lig-4*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTGAGAAGAGAGGTGGCG	*xpf-1*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTGAGAAGAGAGGTGGCG	*pme-1*
1	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTGAGAAGAGAGGTGGCG	*mlh-1*
25	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	wt
13	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*brc-1*
8	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*lig-4*
11	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*xpf-1*
11	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*ercc-1*
14	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*pme-1*
9	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*exo-1*
3	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*mlh-1*
9	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*polh-1*
9	CCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCG	*rev-1BRCT*

TABLE 4. Templated insertions in *unc-22(st192)* mutants.

***unc-22 (st192)*: break left after Tc1 excision**				
	GTTCTCCAATTTTGGGATACA		TATGTCGTTGAACGTTTTGAG	
	CAAGAGGTTAAAACCCTAT		ACATACAGCAACTTGCAAAACTC	
Class IIA: 2bp insertions				
24	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*wt*
8	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*brc-1*
10	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*lig-4*
2	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*xpf-1*
9	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*ercc-1*
5	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*pme-1*
3	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*mlh-1*
7	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*polh-1*
12	GTTCTCCAATTTTGGGATACA		TGTATGTCGTTGAACGTTTTGAG	*rev-1BRCT*
Class III: flank duplications				
1	GTTCTCCAATTTTGGGATACA	AAA	TGTATGTCGTTGAACGTTTTGAG	*exo-1*
1	GTTCTCCAATTTTGGGATACA	GGA	TGTATGTCGTTGAACGTTTTGAG	*wt*
2	GTTCTCCAATTTTGGGATA	TTTTGG	GTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGAT	TTGGGA	TGTATGTCGTTGAACGTTTTGAG	*ercc-1*
2	GTTCTCCAATTTTGGGATAC	TTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATACA	TTTTTGG	GTATGTCGTTGAACGTTTTGAG	*lig-4*
1	GTTCTCCAATTTTGGGATACA	ATTTTGG	GTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATACA	ATTTTGG	GTATGTCGTTGAACGTTTTGAG	*xpf-1*
1	GTTCTCCAATTTTGGGATACA	ATTTTGG	GTATGTCGTTGAACGTTTTGAG	*pme-1*
1	GTTCTCCAATTTTGGGATAC	TTTTGGG	ATGTCGTTGAACGTTTTGAG	*polh-1*
1	GTTCTCCAATTTTGGGATACA	ATTTTGG	GTATGTCGTTGAACGTTTTGAG	*polh-1*
12	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*wt*
2	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*brc-1*
2	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*lig-4*
2	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*pme-1*
1	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*exo-1*
2	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*polh-1*
3	GTTCTCCAATTTTGGGATACA	ATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*rev-1BRCT*
1	GTTCTCCAATTTTGGGA	ATTTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*lig-4*
1	GTTCTCCAATTTTGGGATAC	TTTGGGTATA	TGTATGTCGTTGAACGTTTTGAG	*lig-4*
1	GTTCTCCAATTTTGGGATACA	TATATTTTGG	GTATGTCGTTGAACGTTTTGAG	*lig-4*
1	GTTCTCCAATTTTGGGATACA	TTTTGGGATATA	TGTATGTCGTTGAACGTTTTGAG	*brc-1*
1	GTTCTCCAATTTTGGGATACA	TTTTGGGATGGA	TGTATGTCGTTGAACGTTTTGAG	*xpf-1*
1	GTTCTCCAATTTTGGGATACA	TTTTGGGATGGA	TGTATGTCGTTGAACGTTTTGAG	*rev-1BRCT*
1	GTTCTCCAATTTTGGGATACA	TTTTGGGATACA	TGTATGTCGTTGAACGTTTTGAG	*lig-4*
2	GTTCTCCAATTTTGGGATACA	TTTTGGGATACA	TGTATGTCGTTGAACGTTTTGAG	*rev-1BRCT*
1	GTTCTCCAATTTTGGGAT	GGAATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*mlh-1*
1	GTTCTCCAATTTTGGGATAC	GATATTTGGGATA	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATAC	TTCTCCAATTTTGG	GTATGTCGTTGAACGTTTTGAG	*pme-1*
1	GTTCTCCAATTTTGGGATACA	TAATTTTGGGATATA	TGTATGTCGTTGAACGTTTTGAG	*exo-1*
1	GTTCTCCAATTTTGGGAT	TTTGGGATTTTGGAT	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATACA	GGTTCTCCAATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*ercc-1*
1	GTTCTCCAATTTTGGGATACA	ATTTTGGGATTTTTTTGGGATTGTTGATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATAC	GTGATTTTGGGATTTTGGGATAATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*ercc-1*
1	GTTCTCCAATTTTGGGATACA	TTTTTGGGATTTTGGGATTTTGGGATTTTGG	GTATGTCGTTGAACGTTTTGAG	*rev-1BRCT*
1	GTTCTCCAATTTTGGGATACA	TGA	TGTATGTCGTTGAACGTTTTGAG	*brc-1*
1	GTTCTCCAATTTTGGGATA	TGTC	TGTCGTTGAACGTTTTGAG	*ercc-1*
1	GTTCTCCAATTTTGGGATAC	GTTG	ATGTCGTTGAACGTTTTGAG	*lig-4*
1	GTTCTCCAATTTTGGGATACA	TGTC	GTATGTCGTTGAACGTTTTGAG	*brc-1*
1	GTTCTCCAATTTTGGGATAC	GTTGA	GTATGTCGTTGAACGTTTTGAG	*xpf-1*
2	GTTCTCCAATTTTGGGATACA	TGTCGT	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATAC	TTTTCGT	TGTATGTCGTTGAACGTTTTGAG	*brc-1*
1	GTTCTCCAATTTTGGGATACA	TGTCGTTGT	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATACA	TGTCATTGA	ATGTCGTTGAACGTTTTGAG	*pme-1*
1	GTTCTCCAATTTTGGGATAC	GTATGTCGTT	TGTATGTCGTTGAACGTTTTGAG	*wt*
1	GTTCTCCAATTTTGGGATA	TGTCGTTGAA	TATGTCGTTGAACGTTTTGAG	*pme-1*
1	GTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTGA	GTATGTCGTTGAACGTTTTGAG	*brc-1*
1	GTTCTCCAATTTTGGGATACA	TGTCGTTGAAATTTTGGGATACATGTCGTTGGA	TGTATGTCGTTGAACGTTTTGAG	*xpf-1*
1	GTTCTCCAATTTTGGGATACA	GTTTTGGGATGTATGTCGTTGAAAATTTTGGGA	TGTATGTCGTTGAACGTTTTGAG	*polh-1*

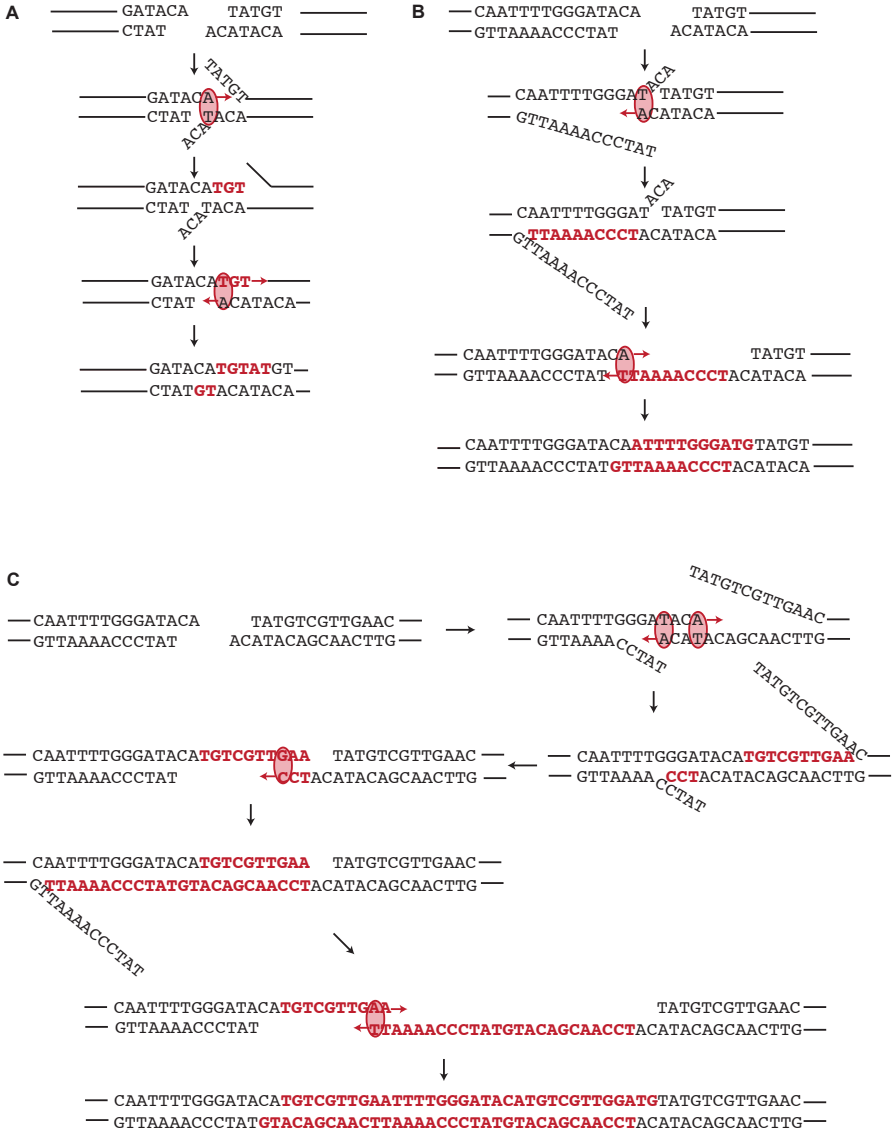


FIGURE 3. TMEJ explains complex repair footprints after transposon excision.
Extension from a single basepair is catalyzed by Pol θ, depicted by a red oval. Newly synthesized nucleotides are marked in red. **A.** Apparent blunt end joining explained via a mechanism of single nucleotide base pair priming, followed by a second cycle of annealing and extension. **B.** A similar scenario to explain a frequently (n=24) observed left flank duplication through two rounds of extensions starting with single base pair priming. **C.** A highly complex footprint, derived from an *xpf-1* negative background, shows multiple templated insertions derived from both flanks of the break. Also here, the same simple mechanistic model comprised of iterative rounds of Pol θ-mediated extension primed at basepaired 3′ ends, intersected by strand dissociation and re-annealing explains very complex insertions.

In the absence of Pol θ, the repair frequency by EJ dropped dramatically, suggesting TMEJ as the pathway of choice for EJ repair of transposition-induced breaks. In Drosophila, a subclass of DSB repair using larger microhomologies is greatly reduced in mutants for Pol θ, suggestive of a conserved role for Pol θ in repairing transposon-induced breaks (Chan *et al.* 2010; Hendel *et al.* 2011). However, in our assay, six basepairs of microhomology were guiding the repair process even in the absence of Pol θ, although the reversion frequency was drastically reduced. It is currently unknown which factors contribute to these repair outcomes; possibly canonical EJ. Possibly, these repair products result from EJ on DSBs that that are that are not normally channeled into a Pol θ-dependent pathway, for instance if these breaks are induced at another developmental stage (still contributing to the germ lineage) or induced at another cell cycle stage.

Pol θ in promoting genome stability

We demonstrate a key role for TMEJ in DSB repair in the *C. elegans* germline. While data from other systems describe microhomology-mediated repair mostly as a backup mechanism active in the absence of canonical NHEJ factors, we show that TMEJ dominates error-prone repair of Tc1 breaks in *C. elegans* germ cells. This dependency may reflect a specific developmental or cellular context for break induction by Tc1 that excludes canonical NHEJ activity. Indeed, also the repair of radiation-induced DSBs in the *C. elegans* germline is largely independent of canonical NHEJ (Clejan *et al.* 2006; Robert *et al.* 2008). Active mechanisms may even inhibit suppress NHEJ in *C. elegans* germ cells (Lemmens *et al.* 2013).
In the course of the study described in this chapter, we found evidence for at least two other substrates for TMEJ - mediated repair: (i) stretches of guanines that are capable of forming stable G4 structures (Koole, article submitted for publication) and (ii) spontaneously stalled forks due to replicational stress (Chapter 3). In those cases we envisaged DSB induction due to prolonged replication fork stalling, followed by TMEJ (Chapter 3). Here, we clearly demonstrated the involvement of TMEJ in repair of germline DSBs, which further supports our tentative model for TMEJ of replication fork stalls via a DSB intermediate.

A striking feature of TMEJ is the occasional insertion of flanking sequences at the break site. Although a mechanism of repeated rounds of extension from not fully complementary dsDNA intermediates might intuitively not probable, we show here that a limited sequence of steps suffices in explaining virtually all footprints. The mechanism proposed bears resemblance to serial replication slippage, which has been proposed to explain complex rearrangements in mammals (Chen *et al.* 2005). Insertion of templated nucleotides has also been observed at translocation breakpoints of certain tumours (Roth *et al.* 1985; Welzel *et al.* 2001; Marculescu *et al.* 2006; Edmunds *et al.* 2008; Murga Penas *et al.* 2010). Copy number variations (CNVs) in humans have also been

attributed to mechanisms of rearranging DNA based on microhomology and flexibility of primer-template intermediates, although in those cases inserted sequences were found to be much longer (Lee *et al.* 2007; Hastings *et al.* 2009). In addition to Pol θ, higher eukaryotes may have a larger set of polymerases at their disposal that may act in EJ repair. Nevertheless, the underlying principles of repetitive cycles of one nucleotide pairing and extension may be conserved over all species and contribute significantly to genome evolution.

MATERIALS AND METHODS

C. elegans genetics

General methods for culturing *C. elegans* were used (Brenner 1974).The following alleles were used in this study: *rde-3 (ne298); mut-7 (pk204); unc-22 (st136::Tc1); unc-22 (st192::Tc1); lig-4(ok716); xpf-1 (e1487); ercc-1 (tm2073); brc-1 (tm1145); pme-1(ok988); exo-1(tm1842); mlh-1(gl516); polh-1(lf31); rev-1 (lf35); polq-1 (tm2026).*

Reversion assay to identify mutations caused by Tc1 transposition

Animals carrying either *unc-22 (st136::Tc1)* or *unc-22 (st192::Tc1)* were crossed with *rde-3(ne298)* or *mut-7(pk204)* males, to allow for germline transposition in these strains. These alleles were subsequently crossed into various genetic backgrounds defective for DSB repair genes. Animals were kept in culture by selecting for worms that display the Unc phenotype. To study EJ repair at the Tc1 site, animals were singled on about fifty 6 cm agar plates seeded with OP50 and propagated till starvation. Non-unc wildtype moving revertants were isolated from these plates. The reversion frequency is calculated by assuming a Poisson distribution for reversion events (Mori *et al.* 1988): Reversion frequency = $-\ln(P_0) / 2n$, where P_0 is the fraction of plates that did not yield revertants, and n is the number of animals that were screened per plate. To identify the molecular nature of the event that restored UNC-22 function, PCR analysis and sequencing was performed on the isolated revertants.

REFERENCES

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

Bahmed K., Seth A., Nitiss K. C., Nitiss J. L., 2011 End-processing during non-homologous end-joining: a role for exonuclease 1. Nucleic Acids Research **39**: 970–978.

Bessereau J.-L., 2006 Transposons in C. elegans. WormBook : the online review of C elegans biology: 1–13.

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics **77**: 71–94.

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

Chen J.-M., Chuzhanova N., Stenson P. D., Férec C., Cooper D. N., 2005 Complex gene rearrangements caused by serial replication slippage. Hum. Mutat. **26**: 125–134.

Clejan I., Boerckel J., Ahmed S., 2006 Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics **173**: 1301–1317.

Davies A. A., Huttner D., Daigaku Y., Chen S., Ulrich H. D., 2008 Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. Mol Cell **29**: 625–636.

Decottignies A., 2007 Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination. Genetics **176**: 1403–1415.

Edmunds C. E., Simpson L. J., Sale J. E., 2008 PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell **30**: 519–529.

Eid W., Steger M., El-Shemerly M., Ferretti L. P., Peña-Diaz J., König C., Valtorta E., Sartori A. A., Ferrari S., 2010 DNA end resection by CtIP and exonuclease 1 prevents genomic instability. EMBO Rep **11**: 962–968.

Emmons S. W., Yesner L., 1984 High-frequency excision of transposable element Tc 1 in the nematode Caenorhabditis elegans is limited to somatic cells. Cell **36**: 599–605.

Gagnon S. N., Hengartner M. O., Desnoyers S., 2002 The genes pme-1 and pme-2 encode two poly(ADP-ribose) polymerases in Caenorhabditis elegans. Biochem J **368**: 263–271.

Genschel J., Modrich P., 2003 Mechanism of 5'-directed excision in human mismatch repair. Mol Cell **12**: 1077–1086.

Hastings P. J., Ira G., Lupski J. R., 2009 A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet **5**: e1000327.

Hendel A., Krijger P. H. L., Diamant N., Goren Z., Langerak P., Kim J., Reissner T., Lee K.-Y., Geacintov N. E., Carell T., Myung K., Tateishi S., D'Andrea A., Jacobs H., Livneh Z., 2011 PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLoS Genet **7**: e1002262.

Hoeijmakers J. H. J., 2001 Genome maintenance mechanisms for preventing cancer. Nature **411**: 366–374.

Ketting R. F., Haverkamp T. H., van Luenen H. G., Plasterk R. H., 1999 Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell **99**: 133–141.

Krishnakumar R., Kraus W. L., 2010 The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. Mol Cell **39**: 8–24.

Lee J. A., Carvalho C. M. B., Lupski J. R., 2007 A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell **131**: 1235–1247.

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

Lemmens B. B. L. G., Johnson N. M., Tijsterman M., 2013 COM-1 promotes homologous recombination during Caenorhabditis elegans meiosis by antagonizing Ku-mediated non-homologous end joining. PLoS Genet **9**: e1003276.

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

Marculescu R., Vanura K., Montpellier B., Roulland S., Le T., Navarro J.-M., Jäger U., McBlane F., Nadel B., 2006 Recombinase, chromosomal translocations and lymphoid neoplasia: targeting mistakes and repair failures. DNA Repair **5**: 1246–1258.

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

McVey M., Radut D., Sekelsky J. J., 2004 End-joining repair of double-strand breaks in Drosophila melanogaster is largely DNA ligase IV independent. Genetics **168**: 2067–2076.

Mori I., Moerman D. G., Waterston R. H., 1988 Analysis of a mutator activity necessary for germline transposition and excision of Tc1 transposable elements in Caenorhabditis elegans. Genetics **120**: 397–407.

Murga Penas E. M., Callet-Bauchu E., Ye H., Gazzo S., Berger F., Schilling G., Albert-Konetzny N., Vettorazzi E., Salles G., Wlodarska I., Du M.-Q., Bokemeyer C., Dierlamm J., 2010 The t(14;18)(q32;q21)/IGH-MALT1 translocation in MALT lymphomas contains templated nucleotide insertions and a major breakpoint region similar to follicular and mantle cell lymphoma. Blood **115**: 2214–2219.

Plasterk R. H., 1991 The origin of footprints of the Tc1 transposon of Caenorhabditis elegans. EMBO J **10**: 1919–1925.

Plasterk R. H., Izsvák Z., Ivics Z., 1999 Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet **15**: 326–332.

Pontier D. B., Tijsterman M., 2009 A robust network of double-strand break repair pathways governs genome integrity during C. elegans development. Curr Biol **19**: 1384–1388.

Robert V., Bessereau J.-L., 2007 Targeted engineering of the Caenorhabditis elegans genome following Mos1-triggered chromosomal breaks. EMBO J **26**: 170–183.

Robert V. J., Davis M. W., Jorgensen E. M., Bessereau J.-L., 2008 Gene conversion and end-joining-repair double-strand breaks in the Caenorhabditis elegans germline. Genetics **180**: 673–679.

Roth D. B., Porter T. N., Wilson J. H., 1985 Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol **5**: 2599–2607.

Shrivastav M., De Haro L. P., Nickoloff J. A., 2008 Regulation of DNA double-strand break repair pathway choice. Cell Res **18**: 134–147.

Simsek D., Brunet E., Wong S. Y.-W., Katyal S., Gao Y., Mckinnon P. J., Lou J., Zhang L., Li J., Rebar E. J., Gregory P. D., Holmes M. C., Jasin M., 2011 DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. PLoS Genet **7**: e1002080.

Sugawara N., Ira G., Haber J. E., 2000 DNA length dependence of the single-strand annealing pathway and the role of Saccharomyces

cerevisiae RAD59 in double-strand break repair. Mol Cell Biol **20**: 5300–5309.

Tabara H., Sarkissian M., Kelly W. G., Fleenor J., Grishok A., Timmons L., Fire A., Mello C. C., 1999 The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell **99**: 123–132.

Wang M., Wu W., Wu W., Rosidi B., Zhang L., Wang H., Iliakis G., 2006 PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic Acids Research **34**: 6170–6182.

Welzel N., Le T., Marculescu R., Mitterbauer G., Chott A., Pott C., Kneba M., Du M. Q., Kusec R., Drach J., Raderer M., Mannhalter C., Lechner K., Nadel B., Jaeger U., 2001 Templated nucleotide addition and immunoglobulin JH-gene utilization in t(11;14) junctions: implications for the mechanism of translocation and the origin of mantle cell lymphoma. Cancer Res **61**: 1629–1636.

Wyman C., Kanaar R., 2006 DNA double-strand break repair: all's well that ends well. Annu Rev Genet **40**: 363–383.

SUPPLEMENTARY INFORMATION

TABLE S1. Repair after excision of a Tc1 element at two different positions in *unc-22*.

Genotype	Survival*
rde-3 (ne298)	88 (±10) %
rde-3(ne298); brc-1 (tm1145)	25 (±20) %

TABLE S2. *unc-22(st136)* repair footprints arranged by genotype

```
break left by transposon
     AGATTGACGAGATCCATAAGGAAGGATGTACA                                    TACATTGAACTGGAAGCCTCCAACT
     TCTAACTGCTCTAGGTATTCCTTCCTACAT                                      ACATGTAACTTGACCTTCGGAGGTTGA

brc-1 (tm1145); unc-22 (st136)

class I: deletions
  17 AGATTGACGAGATCCATAAGGAAGGATGTACA                                        TTGAACTGGAAGCCTCCAACT

class III: flank duplications
   1 AGATTGACGAGATCCATAAGGAAGGATGTA   TGTACATA                                 GAACTGGAAGCCTCCAACT
   1 AGATTGACGAGATCCATAAGGAAGGATGTACA                  TCATTGAACACA            TTGAACTGGAAGCCTCCAACT

lig-4 (ok716); unc-22 (st136)

class I: deletions
  28 AGATTGACGAGATCCATAAGGAAGGATGTACA                                        TTGAACTGGAAGCCTCCAACT

other: large insertions
   1 AGATTGACGAGATCCATAAGGAAGGAT      ATCTTTTTTGGCCAGCAC              TGTACATTGAACTGGAAGCCTCCAACT
   1 AGATTGACGAGATCCATAAGGAAGGATGT    GAAAAAG-54bp-ATCTTTTTGGCCAGCAC TGTACATTGAACTGGAAGCCTCCAACT

xpf-1 (e1487); unc-22 (st136)

class I: deletions
  30 AGATTGACGAGATCCATAAGGAAGGATGTACA                                        TTGAACTGGAAGCCTCCAACT

pme-1 (ok988); unc-22 (st136)

class I: deletions
  29 AGATTGACGAGATCCATAAGGAAGGATGTACA                                        TTGAACTGGAAGCCTCCAACT
   1 AGATTGACGAGATCCATAAGGAAGGATGTACA                                      ACATTGAACTGGAAGCCTCCAACT

ercc-1(tm2073); unc-22 (st136)

class I: deletions
  13 AGATTGACGAGATCCATAAGGAAGGATGTACA                                        TTGAACTGGAAGCCTCCAACT
   1 AGATTGACGAGATCCATAAGGAAGGATGTACA                                      ACATTGAACTGGAAGCCTCCAACT

polh-1(lf31); unc-22(st136)

class I: deletions
   2 AGATTGACGAGATCCATAAGGAAGG                                          TACATTGAACTGGAAGCCTCCAACT
  37 AGATTGACGAGATCCATAAGGAAGGATGTACA                                       TTGAACTGGAAGCCTCCAACT

class II: small indels
   1 AGATTGACGAGATCCATAAGGAAGGATGTA                                    TGTACATTGAACTGGAAGCCTCCAACT

other: large insertions
   1 AGATTGACGAGATCCATAAGGAAGGATGTA   87 bp insertie                      ACATTGAACTGGAAGCCTCCAACT
```

TABLE S3. *unc-22(st192)* repair footprints arranged by genotype

break left by transposon

```
     AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                  TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
     TTGGGTGGTTCTCTGCTGCCACCAAGAGGTTAAAACCCTAT                                   ACATACAGCAACTTGCAAAACTCTTCTCTCCACCGCCACTAA
```

brc-1 (tm1145); unc-22 (st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGTGGTTCTC                                                                  GTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTT                                                       ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTTG                                                    TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  13 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                  TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGG                                                                      AGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTC                                                      TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGG                                                    TTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                 TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                                ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   8 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                            TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                  TGTC  GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                   TGA TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGAGGGTGGTTCTCCAATTTTGGGATAC                                 TTTTCGT TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGAGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                  TGTCGTTGAACGTTTTGA  GTATGTCGTTGAACGTTTTGAAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTGGGATATA                       TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

lig-4; unc-22(st192)

```
class I: deletions
   1 AACCC
   1 AACCCACCAAGAGACGACGGTGGTTCT                                                                  GAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCA                                                                CGTTTTGAGAAGAGAGGTGGCGGTGATT
   8 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                 CGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT                                                  TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                 GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  10 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                              TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                              CGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                               ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                  GTTG    ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGA     ATTTTTGGGA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC  TTTGGGTATA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTGGGATACA                       TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TATATTTTGGG                          TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTTGG                              GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

xpf-1 unc-22(st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                     CGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  11 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                  TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAA                                                           TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                 GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                  TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATATA                                               TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                               TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAACAAT                                                           CGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                         AGGA TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                  GTTGA  GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGG                             GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TGTCGTTGAAATTTTGGGATACATGTCGTTGGA    TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTGGGATGGA                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TCAGAAACTTCAAG                       TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

TABLE S3 (continued)

ercc-1; unc-22(st192)

```
class I: deletions
   6 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                                TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  11 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                             TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                            GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   9 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                  TGTC      TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT    TTGGGA                              TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC  GTGATTTTGGGATTTTGGGATAATTTTGGGA      TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA GGTTCTCCAATTTTGGGA                  TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

pme-1(ok988); unc-22(st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                                TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                             CGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  14 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                            TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT                                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTTT                                             TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                           GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                        ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   5 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                          TGTCGTTGAA   TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC  TTCTCCAATTTTGG                      GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGG                            GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TGTCATTGA                          ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

exo-1(tm1842); unc-22 (st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGTG                                                                AACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTT                                                               TTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                              TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   9 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                           TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                            GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                        ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATATA                                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA AAA                                    TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                              TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TAATTTTGGGATATA                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

mlh-1(gk516); unc-22 (st192)

```
class I: deletions
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                             TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                              GTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAAT                                                    GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                        ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                      TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGGTCTCCAATTTTGGGATATA                                      TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACG                                      TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT    GGAATTTTGGGA                          TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

TABLE S3 (continued)

polh-1(lf31); unc-22 (st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGT                                                                                  TTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTCGTT                                                                               TTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                                                 GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                                               TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   9 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                            TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT                                                           TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   4 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                            GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   7 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACG                                                              TTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                                        TATATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC  TTTTGGG                                                    ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                                                TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGG                                                   TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA GTTTTGGGATGTATGTCGTTGAAAATTTTGGGA    TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

rev-1 (lf35); unc-22 (st192)

```
class I: deletions
   1 AACCCACCAAGAGACGACGGTGG                                                                                CGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTT                                                                  ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG                                                               TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   9 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA                                                            TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class II: small indels
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                            GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAC                                                          TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                                          ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
  12 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                                        TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                          TAA TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA                                  TGTCGA TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT

class III: flank duplications
   3 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA ATTTTGGGA                                                TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTGGGATGGA                                             TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
   2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTGGGATACA                                             TGTATGTCGTTGAACGTTTTGAGAACAGAGGTGGCGGTGATT
   1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATACA TTTTTGGGATTTTGGGATTTTGGGATTTTGGG                           TATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATT
```

THE TRANSLESION SYNTHESIS POLYMERASE REV-1 PLAYS AN ESSENTIAL ROLE IN GENOME STABILITY AND MITOTIC PROLIFERATION IN THE *C. ELEGANS* GERMLINE

5

Sophie F. Roerink, Robin van Schendel, Marcel Tijsterman

Leiden University Medical Center,
Department of Toxicogenetics

ABSTRACT

REV1 is deoxycytidyl transferase that plays a role in replication of damaged DNA in a process called translesion synthesis (TLS). Although its catalytic activity is restricted to the incorporation of deoxycytosines across DNA damage, other domains including a BRCA1 C-terminal (BRCT) domain mediate its role in protection against various sources of DNA damage.

Here, we characterize a functional knockout for *rev-1* in the nematode *C. elegans* as well as a mutant form of *rev-1*, which carries a mutation of a conserved amino acid in its BRCT domain. While *rev-1BRCT* mutant animals are hypersensitive to various sources of DNA damage, *rev-1(null)* mutants are also severely affected in growth and development during non-challenged conditions, and show a defect in germ cell proliferation. The severe defects in *rev-1(null)* mutants are accompanied by a highly elevated mutation rate under normal growth conditions. A specific subtype of mutations consisting of 50 - 200 bp deletions is observed in *rev-1(null)* mutants, suggestive of replication fork stalling due to endogenous damage.

INTRODUCTION

Organisms possess specialized polymerases that can bypass base damage in template DNA in a process called translesion synthesis (TLS). The most studied bypass polymerases are the Y-family polymerases, which are efficient in replicating across various damaged templates, but lack proofreading activity and are therefore intrinsically error-prone (Guo *et al.* 2009; Sale *et al.* 2012). Although the REV1 protein shares its basic structure with the more versatile members of the family Pol η, Pol κ and Pol ι, its catalytic activity is limited to incorporation of deoxycytosines (Prakash *et al.* 2005). *In vitro* assays demonstrated the capacity of REV1 to incorporate dC residues not only opposite guanines but also adenines, uracil, abasic sites and damaged guanine bases (Nelson *et al.* 1996; Haracska *et al.* 2001; Zhang *et al.* 2002). *In vivo* experiments showed the involvement of deoxycytidyl transferase activity in bypass of N^2-dG lesions and 1-N^6-ethenoadenines, lesions that are possibly caused by lipid peroxidation products in the cell (Zhou *et al.* 2010; Wiltrout and Walker 2011).

In addition to its direct role, there is abundant evidence for a non-catalytic role of REV1 in damage bypass and mutagenesis. REV1 deficient cells and yeast strains are severely impaired in survival upon exposure to UV or carcinogens as cisplatin and MMS while catalytic dead mutants are unaffected (Ross *et al.* 2005; D'Souza *et al.* 2008). This non-catalytical role may reside in two other functional domains that have been identified in the REV1 protein: the N-terminal BRCT domain and the C-terminus. BRCT domains - named after the carboxyl terminus of Breast Cancer Associated Protein 1 - are phosphopeptide binding domains identified in various proteins involved in the DNA damage response (Callebaut and Mornon 1997; Yu *et al.* 2003; Gerloff *et al.* 2012). The BRCT domain of REV1 interacts with the DNA sliding clamp proliferating cell nuclear antigen (PCNA) (Guo *et al.* 2006), and is involved in mutagenic bypass of UV-induced damage in mammalian cells (Jansen 2005; Jansen *et al.* 2009; Guo *et al.* 2009). The outer C-terminus of REV1 contains tandem ubiquitin binding motifs (UBMs) and a region that is able to bind each of the other Y-family polymerases (Guo *et al.* 2003; Prakash *et al.* 2005).

Recent data suggest that REV1 acts in bypass of various DNA lesions in two temporally distinct phases: i) actual bypass of the lesion, which depends on the C-terminal domain and/or the BRCT domain, and ii) subsequent postreplicative gap-filling possibly via REV1 dependent PCNA ubiquitination (Edmunds *et al.* 2008; Jansen *et al.* 2009; Temviriyanukul *et al.* 2012).

In line with a prominent role for REV1 in the DNA damage response, *Rev1-/-* mice are severely affected: they are born in submendelian ratios and display growth retardation, proliferation defects and somatic hypermutation (Jansen *et al.* 2006). A yet unrecognized role for REV1 in epigenetic stability was identified in studies in DT40 chicken cells: *rev1-/-* cells loss repressive histone marks in the vicinity of DNA capable of forming G4 structures (Sarkies *et al.* 2010).

Here, we used the nematode *C. elegans* as a new model system to address the role of REV1 in genome protection and replication progression. We isolated *rev-1* knockout and *rev-1BRCT* mutant strains and analyzed viability and the response to DNA damage in these strains. The fast growing properties and compact genome of *C. elegans* enabled us to study mutagenesis in *rev-1* deficient animals in an unbiased way by whole genome sequencing of mutation accumulation lines.

RESULTS

Isolation of *rev-1(null)* and *rev-1BRCT* mutant alleles in *C. elegans*

Using a targeted mutagenesis approach (Cuppen *et al.* 2007) we isolated mutants for the *C. elegans* homolog of REV1. We identified two substitution mutations respectively in the fourth and the fifth exon of the *rev-1* gene (Figure 1A).
The mutation in the fourth exon causes a premature stop. The transcript of the corrupted gene may either encode a non-functional protein that lacks both the catalytic site and the DNA binding BRCT domain or likely be subject to nonsense-mediated decay. For this reason we denote the *rev-1(lf34)* allele as a *rev-1(null)* allele.
The second mutation causes a G>D substitution in the conserved G283 residue in the BRCT domain (Figure 1). Alignment of this part of the BRCT domain with REV1 proteins from other species reveals that the mutated amino acid in *rev-1(lf35)* worms is homologous to the exact same residue that has been mutated in yeast and mice (Lawrence 2004; Guo *et al.* 2006)(Figure 1B). G193>R and G76>R substitutions in respectively yeast and mice abolish the function of the BRCT domain of the REV1 protein. Analogous to the situation in yeast and mammals, we denote the allele encoding a G283>D substitution in worms as a *rev1BRCT* allele.

rev-1(null) mutants display a stochastic loss of fecundity

We had severe difficulties in maintaining *rev-1(null)* mutants in culture, which could be the result of REV1 loss or the result of background mutations that resulted from the mutagenic treatment. To reduce the number of such possibly deleterious background mutations we first backcrossed the *rev-1(null)* allele several times to wild type N2 and then combined it with a closely positioned *dpy-10 (e128)* marker. By subsequent uncoupling of the *rev-1(null)* allele and *dpy-10* we map the region causing the impaired fertility phenotype to the right arm of chromosome II. Whole genome sequencing data show that in addition to the *rev-1* mutation, ten other substitution mutations have been introduced in this region (Table S1). Nine mutations are intronic or synonymous substitutions; one mutation causes an Ala>Thr mutation in the *cpna-2* gene, encoding the worm homolog of mammalian Copine-4. However, deletion mutants for *cpna-*

2(gk428) were perfectly viable, adding proof for a causal role for the *rev-1* deficiency in the observed growth defect (Table S2), as it is the only other amino-acid changing mutation in the mapped interval.
About 70 percent of the *rev-1(null)* population is sterile (Figure 2A): overall, progeny numbers are low, although some animals produce a brood of about 50 progeny (Figure 2B). In contrast, fecundity of *rev-1BRCT* animals after outcrossing is comparable to wildtype (Figure 2).

REV1 is needed for proliferation in the mitotic zone of the germline

We wondered whether the loss of fecundity in *rev-1(null)* mutants was due to an obvious defect in germ cell maturation. We analyzed germ cell morphology by DAPI stainings on dissected gonads (Figure 2C-D). Maturating germ cells are spatially organized in

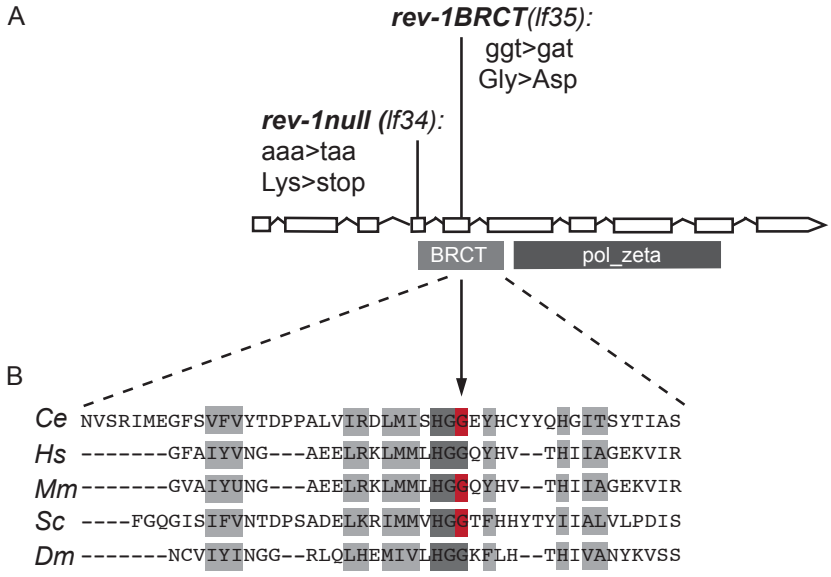


FIGURE 1. *rev-1(null)* and *rev1BRCT* alleles in *C. elegans* used in this study.
A. Gene structure of *C. elegans rev-1* and the molecular nature of two newly derived alleles. An early stop mutation *(lf34)* induces most likely nonsense mediated decay, resulting in a null phenotype. A point mutation *(lf35)* induces a substitution in a conserved residue of the BRCT domain. **B.** Alignment of the *C. elegans* REV-1 BRCT domain with that of human (Hs), mice (Mm), yeast (Sc) and flies (Dm). The arrow indicates the amino acid that has been substituted in the *rev-1BRCT* allele. Mutant alleles described in this study and elsewhere are indicated in red (Guo *et al.* 2006).

the *C. elegans* germline with extensive proliferation happening in the mitotic tip of the germline. The border of the mitotic zone is marked by the appearance of crescent shaped nuclei due to chromatin condensation as chromosomes enter meiotic prophase (Kimble and Crittenden 2005).

Strikingly, the number cells in the mitotic zone of the germline of *rev-1(null)* mutants is strongly reduced as compared to wildtype (Figure 2C-E).

A

	brood size	sterile worms (%)
N2	254 (±12)	0
rev-1BRCT	242 (±24)	0
rev-1null	5.2 (±10)	70

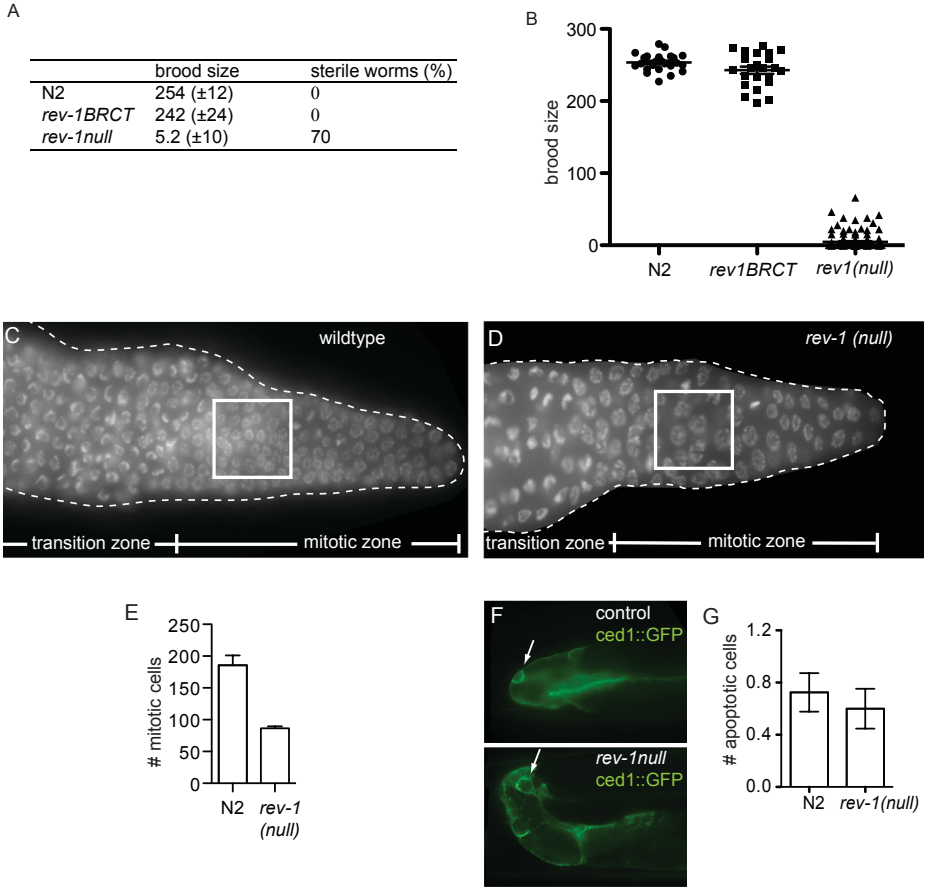


FIGURE 2. *rev-1* is essential for germ cell proliferation

A. *rev-1(null)* worms but not *rev-1BRCT* worms display sterility and **B.** an overall low broodsize. **C.** and **D.** DAPI stainings of germlines in synchronized young adults. The transition zone characterized by crescent shaped nuclei marks the border of the mitotic area. Boxed squares indicate a 20 μm x 20 μm area in the mitotic zone. In *rev-1(null)* mutants (D), there are fewer mitotic cells and cells are enlarged as compared to wildtype (C). **E.** Quantification of the number of cells in the mitotic zone in *rev-1(null)* mutants and N2 controls. **F.** Visualisation of apoptosis in *rev-1(null)* mutants and N2 controls by a transgenic CED1::GFP marker that engulfs germ cells bound for apoptosis. **G.** Quantification of CED1::GFP positive cells in wildtype and *rev-1(null)* backgrounds.

Enlarged mitotic cells can be a manifestation of checkpoint activation in the *C. elegans* germline. In addition to the mitotic checkpoint, a second checkpoint in the germline results in apoptosis induction in the germline bend (Gartner *et al.* 2000; 2008). To check for checkpoint activation, we quantified apoptosis by crossing in a transgenic ced1::GFP marker, which marks germline cells bound for apoptosis (Schumacher *et al.* 2005). Apoptosis was however not increased in *rev-1(null)* mutants (Figure 2F-G).

rev-1BRCT mutants are hypersensitive to exposure to DNA damage in the *C. elegans* germline

We wondered whether the severe effect of REV-1 depletion in worms could be attributed to a role in the DNA damage response in the germline. Since *rev-1(null)* mutants were already severely affected without any treatment, we tested the possibly hypomorphic *rev-1BRCT* mutants for germline sensitivity to different sources of DNA damage.
We exposed *rev-1BRCT* young adults and wildtype controls to UVC-irradiation, γ-irradiation and exposure to methyl methane sulphonate (MMS) and scored for survival in the next generation (Figure 3).

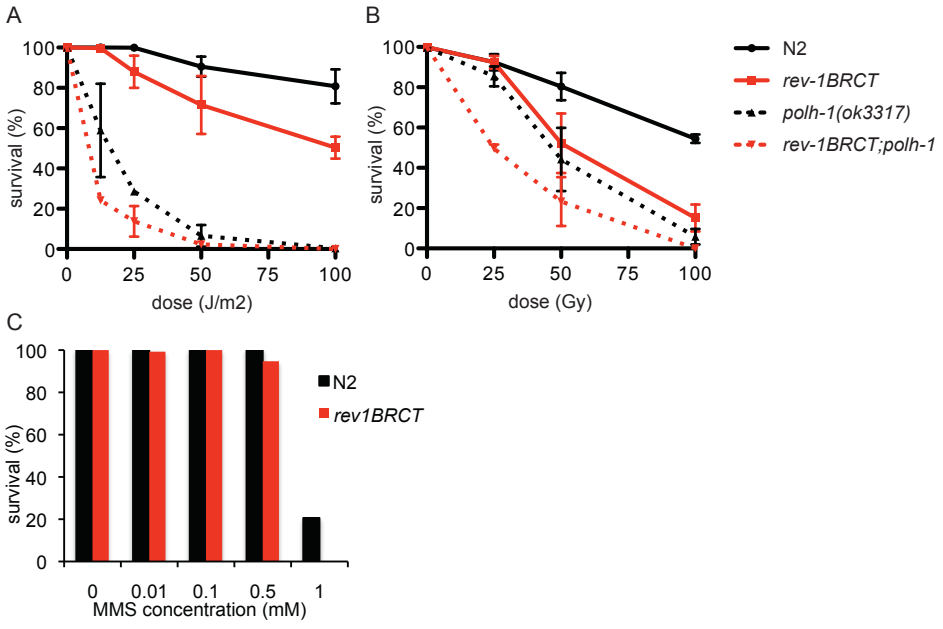


FIGURE 3. The BRCT domain of REV-1 protects against DNA damage in the *C. elegans* germline
A. Germline sensitivity to UV-irradiation. **B.** Germline sensitivity to γ-irradiation. *rev-1BRCT* mutants are hypersensitive to both sources of damage as compared to wildtype controls. The additional knockout of TLS polymerase η further increases sensitivity for both UV-irradiation (A) and γ-irradiation (B). **C.** Germline sensitivity to MMS.

UVC-irradiation causes mainly replication blocking lesions as cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs). Sensitivity to UV was elevated in *rev-1BRCT* mutants but not to the extent as we showed for mutants defective in Pol η, a TLS polymerase previously implicated in protection against UV (Roerink *et al.* 2012)Figure 3a).
γ-irradiation is mainly described as an inducer of double strand breaks (DSBs) but in addition it causes single strand breaks (SSBs) as well as 8-oxo-dG sites and thymine glycols (Roos and Kaina 2012) that can block replication. Sensitivity to γ-irradiation was also increased in *rev-1BRCT* mutants, as compared to wildtype controls (Figure 3B).
MMS alkylates DNA and proteins and can cause various bulky adducts on the DNA in both indirect and direct ways (Fu *et al.* 2012). Sensitivity to MMS was increased in *rev-1BRCT* mutants as compared to wildtype controls (Figure 3C).

rev-1BRCT acts non-epistatic to TLS polymerase η in the protection against DNA damage

In our previous work we identified an important role for Y-family polymerase η in the protection against various sources of DNA damage (Roerink *et al.* 2012). We wondered whether reduced survival in *rev-1BRCT* mutants on exposure to DNA damage could be attributed to defective recruitment or activity of Pol η. To test this hypothesis, we crossed *rev-1BRCT* mutants with *polh-1(ok3317)* mutants and studied sensitivity of *rev-1BRCT* mutants to UV-irradiation and γ-irradiation in a *polh-1* defective background (Figure 3A-B). For both phenotypes the effect of a *rev-1BRCT* mutation was additive to the effect of Pol η depletion, showing independent roles for both Y-family polymerases.

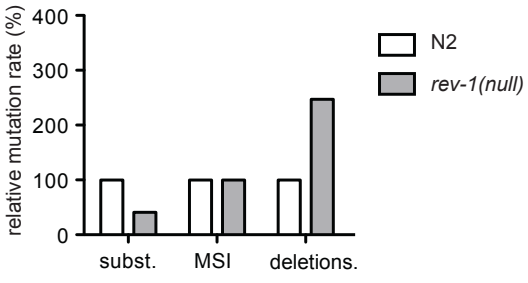
REV-1 protects the *C. elegans* genome against deletions under normal growth conditions

Genome instability due to defects in the response to endogenous DNA damage may be the underlying cause for the severe phenotype of *rev-1(null)* mutants. To study accumulation of spontaneous mutations in *rev-1(null)* mutants, we performed whole-genome sequencing of several mutation accumulation (MA) lines (Denver *et al.* 2009). Due to the severe growth defects in this strain, we were unable to maintain the lines for more than 13-24 generations. We sequenced the genomes of *rev-1(null)* animals after a closely monitored number of generations and compared these data with N2 lines that had been grown for 60 generations. In this way we were able to assay mutagenesis in a largely unbiased way.
Strikingly, a specific subtype of mutations was detected in *rev-1(null)* mutants, which was completely absent in wildtype controls: 8 deletions, in size ranging from 50 to 200 bps were found in 86 generations of REV1 deprived growth, versus 0 in 120 generations

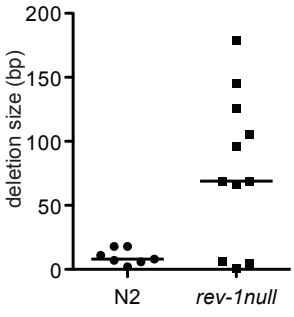
A

	generations	mutations / generation			
		substitutions	MSI mutations	deletions<50bp	deletions≥50bp
N2					
MA N2 2	60	0.35	0.20	0.03	0.0
MA N2 4	60	0.33	0.12	0.03	0.0
		0.34 ± 0.01	**0.16 ± 0.06**	**0.03**	**0.0**
rev-1 (null)					
MA *rev-1* A	13	0.23	0.31	0.00	0.23
MA *rev-1* B	14	0.07	0.07	0.00	0
MA *rev-1* C	15	0.13	0.07	0.13	0.13
MA *rev-1* D	34	0.15	0.18	0.00	0.09
		0.14 ± 0.07	**0.16 ± 0.11**	**0.03 ± 0.07**	**0.11 ± 0.1**

B



C



D

	left flank	deletion				right flank
		left	size	inserted	right	
	CATAAAACCT	ccagaacctc	69		tgtaggcgca	GTCACTGAAC
	TGCCAATTTT	ccttgcaaaa	96		tttcctttga	ATGTTTTCAT
priming	AAAGTAGCAG	agttttttt	66		gatccaagga	AGTGCCGCAA
	AGCGATAAAA	ttgtaatggc	105		acaattgtgc	TCAAAAGGAC
	ACGAGGCAGG	cgtaggtcgc	126		aacatgtttt	CCTTTTCAGT
	AAAACAATTA	aatgaagtta	179		agaaacaggc	AAAAAAGGGG
insertions	AAATTAGACC	aagggatagc	69	gacc	taaataaagt	TTTCAATATT
	CTGTGCTGCT	aacaatcttg	145	tgctgca	gatacagggt	CCACGTAGAA

FIGURE 4. *rev-1(null)* mutants accumulate spontaneous deletions
A. Mutation rates in MA-lines of *rev-1(null)* mutants and wildtype controls that have been growing for the indicated number of generations. **B.** Overall mutation rate in *rev-1(null)* lines normalized to wildtype controls. **C.** Deletion size in *rev-1(null)* mutants. Larger deletions ranging from 50 to 200 bps are exclusively found in *rev-1(null)* mutants. **D.** Sequence context of 50-200 bp deletions found in *rev-1(null)* mutants. In two cases, a deletion was accompanied by the insertion of novel DNA, possibly resulting from duplication of flanking sequences. Four out of six deletions without associated insertions had one or more basepairs at the breakpoint that could not be unambiguously attributed to only one deletion junction. Both features are hallmarks of Pol θ - mediated end joining (TMEJ) a priming-based mechanism of DSB repair (chapter 3, this thesis).

under REV1 proficient growth (Figure 4A-C).
Closer inspection of the sequence context of the deletions revealed several footprints that are in agreement with a priming based mechanism (Figure 4D, Chapter 3 and 4, this thesis). In these cases, one or more bases at the flanking site are identical to the deleted bases at the other end of the deletion. Two cases show a insertion of a short stretch at the break site which can be mapped to the sequence flanking the deletion.

DISCUSSION

Our data support a key role for REV-1 in maintaining genome stability and thereby growth and survival of a multicellular organism. *rev-1* deficient worms display high numbers of sterility, possibly due to defects in germline proliferation. These results resemble mouse data: in *Rev1-/-* mice fertility and germ cell proliferation are also severely compromised, *Rev1-/-* mice are born in submendelian ratios and could not been constructed at all in a C57BL/6 background (Jansen *et al.* 2006).
While the BRCT domain of REV-1 is not required for this essential function (fertility of *rev-1BRCT* mutant worms was unaffected) it is required for REV1's role in protecting cells against a wide range of DNA damaging agents. It is unclear by which mechanism the BRCT domain promotes translesion synthesis across damaged DNA. Possibly its interaction with monoubiquitylated PCNA is needed for recruitment of other TLS polymerases (Guo *et al.* 2006). Previous work identified Polη as a key factor in protection against UV and γ-irradiation in the *C. elegans* (Roerink *et al.* 2012). However, epistasis analyses showed that knocking out *polh-1* further increased the DNA damage sensitivity of *rev-1BRCT* mutant animals, and vice versa. The B-family member Polζ has not been tested and may contribute to survival upon exposure to DNA damage in a REV1BRCT dependent fashion.
The role of REV1 in damage protection is not restricted to exogenous DNA damage: apart from their severe reduction in fertility, *rev-1(null)* mutants grown for several generations under unchallenged conditions accumulate deletions ranging from 50 to 200 basepairs. The size and sequence context of the breaks bear resemblance to spontaneous deletions accumulating in *polh-1* and *polh-1polk-1* deficient mutant strains (Chapter 3, this thesis).
Chromosomal breaks resulting from replication fork stalling on endogenous lesions may be a consequence of impaired translesion synthesis in the absence of REV-1, either by the absence of itsthe catalytic activity or by defective recruitment of other TLS polymerases. The catalytic activity of REV1 itself on bypass of endogenous lesions appears to be quite limited, although bypass activity on N^2-dG and 1-N^6-ethenoadenine adducted sites has been reported, which are known to be generated endogenously (Zhou *et al.* 2010; Wiltrout and Walker 2011).
The interaction of REV1 with other polymerases depends mostly on its outer C-terminal

domain, which is not conserved in *C. elegans* and yeast (Kosarek *et al.* 2008). However, in yeast an interaction of the C-terminal domain with Rev7 - the catalytic subunit of Polζ - was demonstrated (D'Souza and Walker 2006). Alternatively, other yet-unknown domains might be mediating the recruitment of the TLS polymerases in *C. elegans*.

The sequence context of spontaneous deletions in *rev-1* mutants does not only show similarity to spontaneous *polh-1polk-1* - mediated mutations but also to repair footprints of Tc1-induced breaks and deletions of quadruplex-forming guanine stretches (Chapter 4, this thesis, W. Koole, article submitted). In these various forms of repair of (replication-dependent) double strand breaks in the *C. elegans* germline we identified an important role for the A-family polymerase Pol θ (Chapters 3 and 4, this thesis). The characteristics of TMEJ or Pol θ-mediated repair - priming based on a single homologous nucleotide at the break site, and flank site duplications - are also recognized in *rev-1* deletion footprints, further substantiating TMEJ as a general mechanism for break repair in the germline.

The severe defects in growth and fertility in *rev-1* mutants are possibly not only due to genome instability. *polh-1polk-1* mutants display a spontaneous mutation rate twice as high as that of *rev-1* mutants, but these animals do not display severe defects in growth and fertility (chapter 2, this thesis). REV1 may have dual roles, not only mediating damage bypass but also maintaining epigenetic stability of replicated DNA (Sarkies *et al.* 2010). Derepression of silent chromatin and aberrant gene expression may explain some of the phenotypes of REV1 deficient worms and mice.

Sale and coworkers have shown that REV1 enhances replication of G-quadruplex (G4) DNA, which can explain its requirement to maintain the repressive chromatin state at some G4 containing genes (Sarkies *et al.* 2010). However, defective replication of G4 structures does not explain the spontaneous mutations in *rev-1 C. elegans*: neither in our unbiased sequence analysis nor in specific reporter (Kruisselbrink *et al.* 2008) or PCR assays on endogenous G4 sites (unpublished results) did we observe spontaneous mutations at G4 sites in *rev-1* mutants.

In conclusion, we present evidence that REV1 is required for genome stability under genotoxically non-challenged conditions in a multicellular organism. Animals have very poor fertility in REV-1 deprived conditions which goes together with the increased occurrence of mutations that are typical for replication-associated DNA breaks. The genetically amenable *C. elegans* model offers a valuable tool for further understanding the molecular and cellular nature of the essential role of REV1 in organism growth and development.

MATERIALS AND METHODS

C. elegans genetics

All strains were cultured according to standard methods (Brenner 1974). Wildtype N2 (Bristol) worms were used in all control experiments. *rev-1(lf34)*, *rev-1(lf35)* and *polk-1 (lf29)* mutants were isolated in our own laboratory. *polh-1(ok3317)*, *cpna-2(gk428)* and *dpy-10(e128); unc* and the transgenic line MD701 (*bcIs39[P(lim-7)ced-1::GFP+lin-15(+)]*) were obtained from the Caenorhabditis Genetics Center, Minnesota, USA.

Brood size assays

To determine the brood size, we singled L4 animals on OP50 plates. Every second day, we quantified the number of embryos and larvae on the plate, and transferred the mother to a new plate. The next day we determined the hatching rate of the eggs on the plates. For each strain we quantified the brood of at least 24 parents, or until a brood of about 650 animals was counted.

Microscopy

Nuclear staining on germlines and embryos was performed by incubation of staged young adults for 10 minutes in ethanol containing 10 μg/mL 4',6-diamidino-2-phenylindole (DAPI). After two washes with PBS, worms were mounted on object slides in 30% glycerol. Microscopy was performed with a Leica DM6000 microscope. For the analysis of apoptosis transgenic MD701 animals expressing a CED1::GFP fusion by the lin-7 promotor were used to visualize sheath cells surrounding apoptotic germ cells (Schumacher *et al.* 2005).

Survival assays

Staged animals were exposed to different doses of various DNA damaging agents. To assess germline sensitivity three plates with three worms were allowed to lay eggs for 24 - 48 hrs per experimental condition. 24 hrs later, the number of non-hatched eggs and the number of surviving progeny was determined. All experiments were performed at least in duplicate. To measure germline sensitivity to UV, staged young adults (one day post L4) were transferred to empty NGM plates and exposed to different doses of UV-C (predominantly 254 nm). Animals were placed on fresh OP50 plates and allowed to lay eggs for 32 hrs.

To measure germline sensitivity to γ-irradiation, L4 animals were exposed to varying doses of radiation generated by an X-ray generator (dose rate 7 Gy/min; YXLON International). Animals were allowed to lay eggs for 48 hrs, and scored 24 hrs later for

hatching.
To measure sensitivity to chronic exposure to MMS, staged L4 animals were placed for 24 hrs on NGM plates containing different concentrations of MMS (Sigma-Aldrich). After 24 hrs, the number of non-hatched eggs and surviving progeny was determined.

Whole genome sequencing of MA lines

Mutation accumulation (MA) lines were generated by cloning out several F1 animals from one P0 plate. Each generation five worms were transferred to new plates. MA lines were maintained for 60 generations or until severe growth defects developed. Single animals were then cloned out and propagated to obtain full plates for DNA isolation. Worms were washed off with M9 and incubated for one hour at room temperature while shaking, to remove bacteria from the intestines. After two washes, worm pellets were lysed for two hours at 65°C with SDS containing lysis buffer. Genomic DNA was purified by using a DNeasy kit (Qiagen). Paired end (PE) libraries for whole genome sequencing (HiSeq2000 Illumina) were constructed from genomic DNA according to manufacturers' protocols with some adaptations. Shortly, 5 μg DNA was sheared using a Covaris S220 ultrasonicator, followed by DNA end-repair, formation of 3'A overhangs using Klenow and ligation to Illumina PE adapters. Adapter-ligated products were purified on Qiaquick spin columns (Qiagen) and PCR-amplified using Phusion DNA polymerase and barcoded Illumina PE primers for 10 cycles. PCR products of the 300 - 400 bp size range were selected on a 2% ultrapure agarose gel and purified on Qiaquick spin columns. DNA quality was assessed and quantified using an Agilent DNA 1000 assay. Four to five barcoded libraries were pooled in one lane for sequencing on a HiSeq.

Bio-informatic analysis

Image analysis, basecalling and error calibration was performed using standard Illumina software. Alignments to the annotated sequence of *C. elegans* available at WormBase WS225 were performed by BWA. Samtools was used for SNP and indel calling, with BAQ calculation turned off. All non-unique SNPs and indels are considered to be pre-existing and were filtered out using custom Perl scripts. We considered a SNP to be real if at least 80% of the called bases were non-wildtype for SNPs that are covered ≥4 times. As a second filter at least one of MA lines of the same genotype should be called as wildtype: having a coverage ≥4 of which ≥80% was of wildtype nature according to pileup generated with mpileup. Sanger sequencing of predicted SNPs validated these criteria.
To identify microsatellite mutations and deletions we used Pindel, developed by Ye et al (Ye *et al.* 2009). For deletion identification we applied similar bio-informatic settings as described in Chapter 3.

REFERENCES

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics **77**: 71–94.

Callebaut I., Mornon J. P., 1997 From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. FEBS Lett **400**: 25–30.

Cuppen E., Gort E., Hazendonk E., Mudde J., van de Belt J., Nijman I. J., Guryev V., Plasterk R. H. A., 2007 Efficient target-selected mutagenesis in Caenorhabditis elegans: toward a knockout for every gene. Genome Res **17**: 649–658.

D'Souza S., Walker G. C., 2006 Novel role for the C terminus of Saccharomyces cerevisiae Rev1 in mediating protein-protein interactions. Mol Cell Biol **26**: 8173–8182.

D'Souza S., Waters L. S., Walker G. C., 2008 Novel conserved motifs in Rev1 C-terminus are required for mutagenic DNA damage tolerance. DNA Repair **7**: 1455–1470.

Denver D. R., Dolan P. C., Wilhelm L. J., Sung W., Lucas-Lledó J. I., Howe D. K., Lewis S. C., Okamoto K., Thomas W. K., Lynch M., Baer C. F., 2009 A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci USA **106**: 16310–16314.

Edmunds C. E., Simpson L. J., Sale J. E., 2008 PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell **30**: 519–529.

Fu D., Calvo J. A., Samson L. D., 2012 Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat Rev Cancer **12**: 104–120.

Gartner A., Boag P. R., Blackwell T. K., 2008 Germline survival and apoptosis. WormBook : the online review of C elegans biology: 1–20.

Gartner A., Milstein S., Ahmed S., Hodgkin J., Hengartner M. O., 2000 A conserved checkpoint pathway mediates DNA damage--induced apoptosis and cell cycle arrest in C. elegans. Mol Cell **5**: 435–443.

Gerloff D. L., Woods N. T., Farago A. A., Monteiro A. N. A., 2012 BRCT domains: A little more than kin, and less than kind. FEBS Lett **586**: 2711–2716.

Guo C., Fischhaber P. L., Luk-Paszyc M. J., Masuda Y., Zhou J., Kamiya K., Kisker C., Friedberg E. C., 2003 Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. EMBO J **22**: 6621–6630.

Guo C., Kosarek-Stancel J. N., Tang T.-S., Friedberg E. C., 2009 Y-family DNA polymerases in mammalian cells. Cell Mol Life Sci **66**: 2363–2381.

Guo C., Sonoda E., Tang T.-S., Parker J. L., Bielen A. B., Takeda S., Ulrich H. D., Friedberg E. C., 2006 REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. Mol Cell **23**: 265–271.

Haracska L., Unk I., Johnson R. E., Johansson E., Burgers P. M., Prakash S., Prakash L., 2001 Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. Genes Dev **15**: 945–954.

Jansen J. G., 2005 The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Research **33**: 356–365.

Jansen J. G., Langerak P., Tsaalbi-Shtylik A., van den Berk P., Jacobs H., de Wind N., 2006 Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. J Exp Med **203**: 319–323.

Jansen J. G., Tsaalbi-Shtylik A., Hendriks G., Gali H., Hendel A., Johansson F., Erixon K., Livneh Z., Mullenders L. H. F., Haracska L., de Wind N., 2009 Separate domains of

Rev1 mediate two modes of DNA damage bypass in mammalian cells. Mol Cell Biol **29**: 3113–3123.

Kimble J., Crittenden S. L., 2005 Germline proliferation and its control. WormBook : the online review of C elegans biology: 1–14.

Kosarek J. N., Woodruff R. V., Rivera-Begeman A., Guo C., D'Souza S., Koonin E. V., Walker G. C., Friedberg E. C., 2008 Comparative analysis of in vivo interactions between Rev1 protein and other Y-family DNA polymerases in animals and yeasts. DNA Repair **7**: 439–451.

Kruisselbrink E., Guryev V., Brouwer K., Pontier D. B., Cuppen E., Tijsterman M., 2008 Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ-defective C. elegans. Curr Biol **18**: 900–905.

Lawrence C. W., 2004 Cellular functions of DNA polymerase zeta and Rev1 protein. Adv Protein Chem **69**: 167–203.

Nelson J. R., Lawrence C. W., Hinkle D. C., 1996 Deoxycytidyl transferase activity of yeast REV1 protein. Nature **382**: 729–731.

Prakash S., Johnson R. E., Prakash L., 2005 EUKARYOTIC TRANSLESION SYNTHESIS DNA POLYMERASES: Specificity of Structure and Function. Annu Rev Biochem **74**: 317–353.

Roerink S. F., Koole W., Stapel L. C., Romeijn R. J., Tijsterman M., 2012 A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during C. elegans Embryogenesis. PLoS Genet **8**: e1002800.

Roos W. P., Kaina B., 2012 DNA damage-induced apoptosis: From specific DNA lesions to the DNA damage response and apoptosis. Cancer Lett.

Ross A.-L., Simpson L. J., Sale J. E., 2005 Vertebrate DNA damage tolerance requires the C-terminus but not BRCT or transferase domains of REV1. Nucleic Acids Research **33**: 1280–1289.

Sale J. E., Lehmann A. R., Woodgate R., 2012 Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat Rev Mol Cell Biol **13**: 141–152.

Sarkies P., Reams C., Simpson L. J., Sale J. E., 2010 Epigenetic instability due to defective replication of structured DNA. Mol Cell **40**: 703–713.

Schumacher B., Schertel C., Wittenburg N., Tuck S., Mitani S., Gartner A., Conradt B., Shaham S., 2005 C. elegans ced-13 can promote apoptosis and is induced in response to DNA damage. Cell Death Differ **12**: 153–161.

Temviriyanukul P., van Hees-Stuivenberg S., Delbos F., Jacobs H., de Wind N., Jansen J. G., 2012 Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair **11**: 550–558.

Wiltrout M. E., Walker G. C., 2011 The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. Genetics **187**: 21–35.

Ye K., Schulz M. H., Long Q., Apweiler R., Ning Z., 2009 Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics **25**: 2865–2871.

Yu X., Chini C. C. S., He M., Mer G., Chen J., 2003 The BRCT domain is a phospho-protein binding domain. Science **302**: 639–642.

Zhang Y., Wu X., Rechkoblit O., Geacintov N. E., Taylor J.-S., Wang Z., 2002 Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Research **30**: 1630–1638.

Zhou Y., Wang J., Zhang Y., Wang Z., 2010 The catalytic function of the Rev1 dCMP transferase is required in a lesion-specific manner for translesion synthesis and base damage-induced mutagenesis. Nucleic Acids Research **38**: 5036–5046.

SUPPLEMENTARY INFORMATION

TABLE S1. Mutations identified by whole genome sequencing in mapped region of *rev-1(null)* mutants.

position	reference	mutation	gene	coding?	protein change
6779390	A	G	null	non coding	
6782822	G	A	null	non coding	
6876160	G	A	C44B7.6a.1	intron	
7457130	A	T	D1022.4	intron	
7592229	G	A	F35D2.5a	intron	
7651965	G	A	null	non coding	
7715589	G	A	B0495.10a	intron	
7720243	G	A	null	non coding	
7751841	G	A	B0228.4c *(cpna-2)*	exon	GCC(Ala) -> aCC(Thr)
7802782	G	A	null	non coding	
7901796	A	T	ZK675.2 *(rev-1)*	exon	AAA(Lys) -> tAA(stop)

TABLE S2. Fecundity in *cpna-2(gk428)* deletion mutants is comparable to wildtype;
A background substitution mutation in *cpna-2* (Table S1) is therefore highly unlikely to be causative for the major fertility problems in the *rev-1null* mutant.

	broodsize	sterile worms
N2	254(±12)	0
rev-1null	5.2(±10)	70
cpna-2(gk428)	216(±76)	0

6 GENERAL DISCUSSION

C. ELEGANS AS A MODEL TO MONITOR GENOME EVOLUTION

The research described in this thesis started with exploring the possibility of the use of the model organism *C. elegans* to monitor mutagenesis by large scale sequencing approaches. The compact genome of *C. elegans* (about 100 Mbp) and its rapid growth properties make this model excellently suitable for analysis by whole genome sequencing. The advantage of this approach is that mutagenesis can be studied in an unbiased fashion because there is no selection for a specific mutational target.
While a similar strategy has been exploited by others to study mutation accumulation in wildtype *C. elegans* strains (Denver *et al.* 2009; Weber *et al.* 2010), we applied this method for the first time to strains that have a specific defect in replicational bypass of lesions by creating knock-out alleles for the three Y-family polymerase members that are present in the worm and monitoring their genomes over time. In chapter 3 we evaluated the effect of knocking out *polh-1* and/or *polk-1*, while chapter 5 describes the effect of a *rev-1(null)* mutation.
In order to study the full spectrum of mutations that are being induced, various bio-informatic analysis tools have been evaluated and further developed in the course of this study (chapters 3 and 5). During recent years, *C. elegans*-specific software packages have been developed for fast and straightforward analysis of whole genome sequence data that were focussed towards rapid identification of mutants (Sarin *et al.* 2008; 2010). However, for our in-depth analysis of mutation spectra, we benefited most from combining standard alignment tools with Pindel - an algorithm that searches for genomic rearrangements (Ye *et al.* 2009).
As sequencing techniques are becoming more widely available and less expensive, this newly developed sequencing and analysis pipeline will be highly valuable in evaluating other genome defence mechanisms. Currently, other mutants in key DNA repair processes are being subjected to similar analysis in our laboratory.

ERROR-PRONE POLYMERASES IN GENOME PROTECTION

It remains an intriguing question how error-prone polymerases contribute to genome stability, while they do result in mutation induction at the same time. It is proposed that error-prone polymerases prevent gross chromosomal instability by preventing prolonged replication fork stalling and eventually replication fork collapse at sites where a replication fork hits unrepaired base damage (Figure 1, Knobel and Marti 2011).
Indeed, the absence of TLS polymerases results in micronuclei induction and γ-H2Ax phosphorylation in UV-irradiated mammalian cells, both indicators of induction of double strand breaks (DSBs) that may result from replication fork collapse (Temviriyanukul *et al.* 2012). Nevertheless, it is unknown to which extent this mechanism operates to

maintain genomic stability in the absence of exogenous stressors as UV. A single study reports similar phenotypes (micronuclei induction and γ-H2Ax phosphorylation) in non-exposed Pol η knockdown cells due to defective replication of common fragile sites (Rey *et al.* 2009).

The research described in this thesis clearly demonstrates the induction of chromosomal aberrations in the absence of TLS polymerases, under normal growth conditions (chapter 3). In the absence of functional TLS, spontaneous deletions ranging from 50 - 300 basepairs are induced that were not observed in wildtype strains. In REV1 deficient strains these deletions occurred about once in every ten generations and in Pol η deficient strains about once in every three generations. Pol κ knockout on its own did not have any effect on mutagenesis, but combined knockout of Pol η and Pol κ lead to highly increased deletion induction - about two deletions per generation - that caused an easily recognizable mutator phenotype for this strain. Nevertheless, these conclusions may be difficult to derive from systems with a lower resolution analysis, stressing the strength of the fast growing and reproducing model organism *C. elegans* for these assays. From our mutation profiles, we conclude that under normal growth conditions Y-family polymerases play both redundant and non-redundant roles in bypass of endogenous lesion sites. Importantly, TLS activity on endogenous lesions appears to be mostly error-free, since wildtype substitution rates are much lower than deletion rates in specific polymerase knockouts.

It remains an intriguing question which endogenous lesions are the exact substrates for the different Y-family members. A bias towards deletion induction on dC residues in a *polh-1polk-1* defective background hints that replication fork stalling in this background

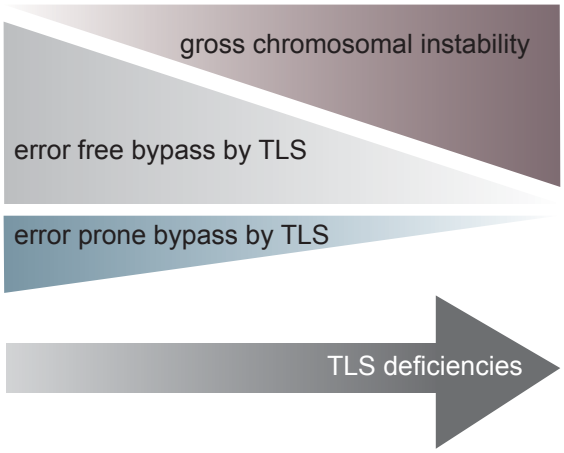


FIGURE 1. Trade off between mutation induction by TLS polymerases and chromosomal instability in the absence of TLS (adapted from (Knobel and Marti 2011).

occurs more often opposite dG residues (Figure 4, Chapter 3). The most abundant oxidative lesions are 8-oxo-dGs, responsible for the majority of transversion mutations (Barnes and Lindahl 2004). Bypass activity of yeast and mammalian Pol η on these lesions has been established in *in vitro* assays (Haracska *et al.* 2000) and may account for a large proportion of mutagenic events.

However, sensitivity assays in chapter 2 show a role of Pol η in tolerance to very different sources of DNA damage, suggesting a broad role for Pol η in damage bypass that may reflect bypass of various structurally different endogenous lesion sites. The role of Pol κ appears to be much more restricted, as *polk-1* mutants are hypersensitive to alkylating damage by methyl methane sulfonate but not to other sources of DNA damage. Nevertheless, additional knockout of Pol κ in Pol η - deficient worms increases the number of deletions under normal growth conditions about ten fold, suggesting that a large subset of lesions is bypassed by Pol κ in the absence of Pol η (Figure 1, chapter 3). Notably, this bypass is most likely error-free, as substitution rates are not increased in *polh-1* mutants. Additional knockout of Pol κ also increases cytotoxicity by MMS in Pol η - defective strains by at least a magnitude of ten, hinting at a similar redundant role for Pol κ in the absence of Pol η in bypassing alkylating damage (Chapter 2). MMS may directly methylate bases, leading to N7-methylguanines, N3-methylguanines and O^6-methylguanines but may also lead to more complex base adducts via reactions with proteins or metabolites in the cell (Fu *et al.* 2012). Endogenous byproducts of oxidation may result in a similar range of base adducts.

Chapter 2 also sheds more light on the strict regulation of error-prone polymerases during development: we show that TLS is specifically important during early embryogenesis in *C. elegans*, when embryos endure a series of fast cell divisions. Exposure to different sources of damaging agents activates a checkpoint in the first embryonic divisions in TLS mutants, resulting in embryonic lethality. Kim et al. identified the sumoylation factor GEI-17 as an important regulator in the activity of TLS protein Pol η (Kim and Michael 2008). Our genome-wide RNAi screen identified genetic interactions of both Pol η and Pol κ with GEI-17. In addition, we show involvement of three new factors: the sumo protease ULP-1, and two nuclear pore proteins (NPP-2 and NPP-22), that are essential for TLS-dependent damage bypass in the early embryo. Together these data suggest that sumoylation processes are essential for regulation of TLS polymerases after exposure to DNA damage in the early embryo. A recent paper by Mosbech and coauthors also describes a central role for the ubiquitin responsive protein p97 and its adaptor protein DVC-1 in regulation of Pol η in *C. elegans* (Mosbech *et al.* 2012).

POL THETA-MEDIATED END JOINING (TMEJ)

One of the most striking observations that we did during analysis of mutation profiles of Y-family polymerase - deficient mutants was the very distinct character of the

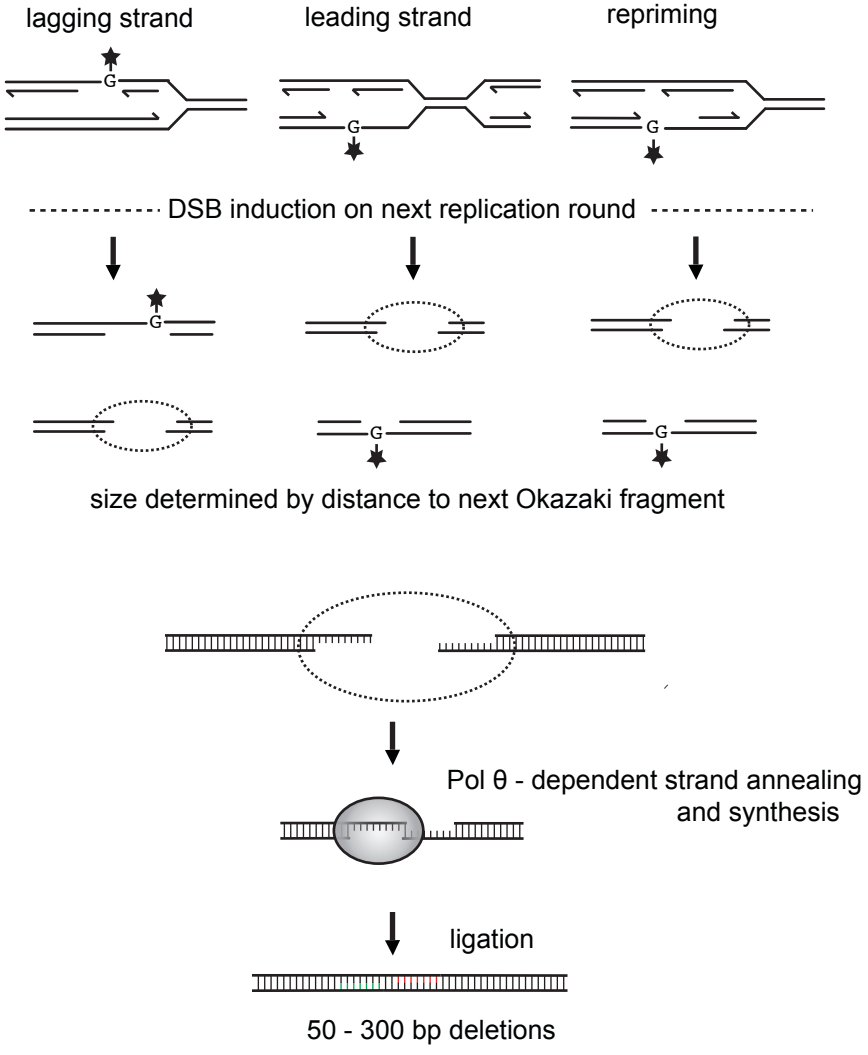


FIGURE 2. Different models explaining TMEJ - mediated deletions in absence of TLS polymerases.
After replication fork stalling at the damaged base, either a gap is formed spanning the sequence from the stall to the next Okazaki fragment, or the damaged base is looped out. Such a ssDNA gap will generate a DSB in the next round of replication, which is subsequently repaired in a Pol θ - dependent manner.

chromosomal aberrations in these strains. In *polh-1*, *polh-1polk-1* and *rev-1* mutants, we saw occurrence of a class of 50-300 bp deletions that were not seen in wildtype control strains and suggested error-prone repair of collapsed replication forks (chapters 3 and 5).

In an unrelated study - described in chapter 4 - we identified a new mechanism for germline repair of transposition-induced DSBs.

Detailed inspection of repair footprints of Tc1-induced breaks and *polh-1polk-1*-derived deletions revealed two shared characteristics in both assays: i) bias towards at least one nucleotide of microhomology and ii) insertion of fragments up to ~20 nucleotides duplicated from the flanks. A third assay, described by W. Koole et al. (submitted for publication) displays similar outcomes in FANCJ/*dog-1* mutants at genomic sequences that can fold in G-quadruplex structures. We show that all of these repair outcomes are the result of the action of the A-family Polymerase θ, hereby describing a novel repair pathway operating in the *C. elegans* germline, termed Pol θ - mediated end joining (TMEJ). This mechanism is based on one or more cycles of annealing of both flanking sites based on at least a single basepair homology and subsequent primer extension, followed by ligation.

At present, we do not know the determinants for the choice for TMEJ repair of DSBs in the *C. elegans* germline as opposed to the two canonical routes non-homologous end joining (NHEJ) and homologous repair (HR). Studies from our lab and others showed HR as the preferred pathway in the germline, while NHEJ is operating in somatic cells (Clejan *et al.* 2006; Pontier and Tijsterman 2009). However, HR depends on the availability of a sister chromatid as a repair template and may operate only in S-phase, when DNA has already replicated.

TMEJ at replication-blocking lesions in *polh-1polk-1*-defective backgrounds and at G4 stretches in *dog-1* mutants suggests that the mechanism is linked to replication. Besides, the distinct size of the deletions may give clues about the molecular mechanism: different scenarios can be envisioned that result in DSBs where the ends are approximately 50-300 bps apart and would thus result in deletions with such a distinct size (Figure 2). Notably, the average distance between Okazaki fragments in the lagging strand in eukaryotes is in the range of 100 - 200 bps (Abdurashidova *et al.* 2000; Smith and Whitehouse 2012), offering an attractive explanation for the deletion length as the gap length determined by the distance of the stalled fork to the next Okazaki fragment. Alternatively, repriming may occur downstream of the lesion site. In both scenarios a single stranded gap may be converted into a DSB intermediate in the next round of replication. The analogy to repair of Tc1-induced breaks by Pol θ strongly suggests that the repair of stalled forks involves a DSB intermediate. However, we cannot exclude an alternative scenario for replication-blocking lesions that does not require a DSB intermediate but is explained by direct priming on the same strand after loop-out of the damaged base. In this scenario, deletion size may be governed by the size of the loop.

FUTURE PERSPECTIVES

One of the most prominent questions raised by the research described in this thesis is to what extent the described error-prone TMEJ mechanism is active in higher eukaryotes under physiological or cancerous conditions. In this thesis, we show that for different sources of replicational stress or DSBs, TMEJ dominates repair in the *C. elegans* germline. Possibly, this repair mode is developmentally strictly regulated and therefore missed in many previous studies on break repair in mammals. Expression analysis showed only low levels of human and mice Pol θ in somatic tissues other than testes (Seki *et al.* 2003).

Nevertheless, Pol θ-dependent footprints bear similarity to certain deletion or translocation events identified in cancer genomes ((Roth *et al.* 1985; Welzel *et al.* 2001; Murga Penas *et al.* 2010), chapter 3), which could argue that cancer cells have acquired the capacity to deal with replication stress by employing Pol θ. Indeed, Pol θ expression was found upregulated in several cancers including breast cancer and colorectal cancer, and higher Pol θ levels are correlating with a reduced prognosis (Kawamura *et al.* 2004; Pillaire *et al.* 2009; Lemée *et al.* 2010).

The work described in this thesis, revealing an unanticipated role for Pol θ in the maintenance of genome stability in *C. elegans*, strongly encourages an in-depth analysis of the function of this protein under physiological and cancerous conditions in humans, as this may provide a new target for therapy.

REFERENCES

Abdurashidova G., Deganuto M., Klima R., Riva S., Biamonti G., Giacca M., Falaschi A., 2000 Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science **287**: 2023–2026.

Barnes D. E., Lindahl T., 2004 Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu Rev Genet **38**: 445–476.

Clejan I., Boerckel J., Ahmed S., 2006 Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics **173**: 1301–1317.

Denver D. R., Dolan P. C., Wilhelm L. J., Sung W., Lucas-Lledó J. I., Howe D. K., Lewis S. C., Okamoto K., Thomas W. K., Lynch M., Baer C. F., 2009 A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci USA **106**: 16310–16314.

Fu D., Calvo J. A., Samson L. D., 2012 Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat Rev Cancer **12**: 104–120.

Haracska L., Yu S. L., Johnson R. E., Prakash L., Prakash S., 2000 Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat Genet **25**: 458–461.

Kawamura K., Bahar R., Seimiya M., Chiyo M., Wada A., Okada S., Hatano M., Tokuhisa T., Kimura H., Watanabe S., Honda I., Sakiyama S., Tagawa M., O-Wang J., 2004 DNA polymerase theta is preferentially expressed in lymphoid

tissues and upregulated in human cancers. Int. J. Cancer **109**: 9–16.

Kim S.-H., Michael W. M., 2008 Regulated proteolysis of DNA polymerase eta during the DNA-damage response in C. elegans. Mol Cell **32**: 757–766.

Knobel P. A., Marti T. M., 2011 Translesion DNA synthesis in the context of cancer research. Cancer Cell Int. **11**: 39.

Lemée F., Bergoglio V., Fernandez-Vidal A., Machado-Silva A., Pillaire M.-J., Bieth A., Gentil C., Baker L., Martin A.-L., Leduc C., Lam E., Magdeleine E., Filleron T., Oumouhou N., Kaina B., Seki M., Grimal F., Lacroix-Triki M., Thompson A., Roché H., Bourdon J.-C., Wood R. D., Hoffmann J.-S., Cazaux C., 2010 DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proceedings of the National Academy of Sciences **107**: 13390–13395.

Mosbech A., Gibbs-Seymour I., Kagias K., Thorslund T., Beli P., Povlsen L., Nielsen S. V., Smedegaard S., Sedgwick G., Lukas C., Hartmann-Petersen R., Lukas J., Choudhary C., Pocock R., Bekker-Jensen S., Mailand N., 2012 DVC1 (C1orf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. Nat Struct Mol Biol **19**: 1084–1092.

Murga Penas E. M., Callet-Bauchu E., Ye H., Gazzo S., Berger F., Schilling G., Albert-Konetzny N., Vettorazzi E., Salles G., Wlodarska I., Du M.-Q., Bokemeyer C., Dierlamm J., 2010 The t(14;18)(q32;q21)/IGH-MALT1 translocation in MALT lymphomas contains templated nucleotide insertions and a major breakpoint region similar to follicular and mantle cell lymphoma. Blood **115**: 2214–2219.

Pillaire M. J., Selves J., Gordien K., Gouraud P.-A., Gentil C., Danjoux M., Do C., Negre V., Bieth A., Guimbaud R., Trouche D., Pasero P., Méchali M., Hoffmann J.-S., Cazaux C., 2009 A "DNA replication" signature of progression and negative outcome in colorectal cancer. Oncogene **29**: 876–887.

Pontier D. B., Tijsterman M., 2009 A robust network of double-strand break repair pathways governs genome integrity during C. elegans development. Curr Biol **19**: 1384–1388.

Rey L., Sidorova J. M., Puget N., Boudsocq F., Biard D. S. F., Monnat R. J., Cazaux C., Hoffmann J.-S., 2009 Human DNA Polymerase Is Required for Common Fragile Site Stability during Unperturbed DNA Replication. Mol Cell Biol **29**: 3344–3354.

Roth D. B., Porter T. N., Wilson J. H., 1985 Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol **5**: 2599–2607.

Sarin S., Bertrand V., Bigelow H., Boyanov A., Doitsidou M., Poole R. J., Narula S., Hobert O., 2010 Analysis of multiple ethyl methanesulfonate-mutagenized Caenorhabditis elegans strains by whole-genome sequencing. Genetics **185**: 417–430.

Sarin S., Prabhu S., O'Meara M. M., Pe'er I., Hobert O., 2008 Caenorhabditis elegans mutant allele identification by whole-genome sequencing. Nat Methods **5**: 865–867.

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

Smith D. J., Whitehouse I., 2012 Intrinsic coupling of lagging-strand synthesis to chromatin assembly. Nature **483**: 434–438.

Temviriyanukul P., van Hees-Stuivenberg S., Delbos F., Jacobs H., de Wind N., Jansen J. G., 2012 Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair **11**: 550–558.

Weber K. P., De S., Kozarewa I., Turner D. J., Babu M. M., de Bono M., 2010 Whole genome sequencing highlights genetic changes associated with laboratory domestication of C. elegans. PLoS ONE **5**: e13922.

Welzel N., Le T., Marculescu R., Mitterbauer G., Chott A., Pott C., Kneba M., Du M. Q., Kusec R., Drach J., Raderer M., Mannhalter C., Lechner K., Nadel B., Jaeger U., 2001 Templated nucleotide addition and immunoglobulin JH-gene utilization in t(11;14) junctions: implications for the mechanism of translocation and the origin of mantle cell lymphoma. Cancer Res **61**: 1629–1636.

Ye K., Schulz M. H., Long Q., Apweiler R., Ning Z., 2009 Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics **25**: 2865–2871.

ADDENDUM

SUMMARY

DNA - the carrier of genetic information in the cell - is a chemically reactive molecule. It is constantly attacked by various sources of damage, from endogenous as well as exogenous origin. Base damage interferes with cellular processes such as replication and transcription and may result in incorporation of incorrect nucleotides in the genome. Accumulation of mutations may threaten viability of the cell, or even worse, result in tumorigenesis if the 'brake' on cellular proliferation is also mutated. However, a limited source of genetic variability is desirable, because it also allows for evolutionary changes.

The key question in this thesis is how organisms ensure genome stability - how do they protect their DNA from mutations caused by base damage? At the same time, this system should not be infallible, as mutations are also needed for evolutionary adaptation to take place.
In my thesis I am studying the contribution of DNA polymerases on the maintenance of genome stability. DNA polymerases copy the DNA content of a cell during replication. In most cases, this is done by the very precise replicative polymerases, that contain proofreading domains which correct occasional misinsertions. However, in the case of base damage, 'normal' replicative polymerases cannot pass the damaged template strand. The cell employs an altenative strategy by means of specialized translesion synthesis (TLS) polymerases. These enzymes can bypass damaged bases, and prevent replication fork stalling. The downside is that TLS polymerases are error-prone; they are less precise and can occasionally result in incorporation of the wrong nucleotide.

Chapter 2 and 3 focus on the contribution on genome stability of two members of the Y-family of TLS polymerases, Pol η and Pol κ, , while chapter 5 focuses on a third member, REV1. In chapter 2 I introduce two new mutant alleles for Pol η and Pol κ in *C. elegans*, and study their function in the protection against various exogenous sources of DNA damage. I conclude that Pol η has a key role in protection of a developing *C. elegans* embryo against different damaging agents, such as the cytostatic cisplatin, UV-irradiation and X-rays. In some cases Pol κ functions redundantly to Pol η. Furthermore I compare results of a screen for sensitivity against a methylating agent, to find new interactors with Pol η and Pol κ in the cell.

In chapter 3 I ask to which extent Pol η and Pol κ contribute to genome stability in the absence of exogenous sources of base damage, thus to the effects of endogenous damage. To study this, we studied mutation accumulation in the mutants isolated in chapter 2. We kept these strains in culture for many successive generations and then determined the profile of the spontaneous mutations that arose.

Strikingly, we observed a very characteristic pattern of mutagenesis in TLS deficient animals: deletions of ~100 basepairs. Two other characteristics were shared by many deletions: microhomology of a single nucleotide on two sides of the deletion, or insertion of small stretches of duplicated DNA from the flanks. This mutational pattern suggested an error-prone mechanism to repair breaks that are the result of stalled replication in the absence of TLS. Insertion of DNA predicted the involvement of another DNA polymerase.
We identified this polymerase to be the A-family member Pol θ. In the absence of Pol θ, large stretches of DNA are resected, resulting in loss of DNA, checkpoint activation and cell death.

In chapter 4 I use a direct source of DNA breaks to study the function of Pol θ: activation of transposons in the germline. Transposons are DNA fragments that can under certain conditions be excised from the DNA, resulting in DNA breaks. Normally, this process is silenced in the germline; however if socalled mutator genes are knocked out, transposition results in breaks in DNA of germ cells. To analyze Pol θ - mediated repair at a molecular level, I study transposons in the muscle gene *unc-22* with two different sequence contexts: either surrounded by microhomology or without any microhomology. In both cases efficient repair is largely dependent of functional Pol θ, demonstrating a direct role for Pol θ in repair of double strand breaks.

REV1 - described in chapter 5 - plays also a role in genome protection both against endogenous and exogenous damage, and protects against spontaneous deletions, analogous to the other Y-family members Pol η and Pol κ. However, its cellular function may be even broader, as Rev1 knockout worms also suffer from progressive loss of maturating germ cells, resulting in sterility.

In conclusion, this study presents the first functional analysis of the Y-family polymerases in genome protection against endogenous and exogenous sources of DNA damage in *C. elegans*. Unexpectedly, this study lead to identification of a new error-prone pathway to repair double stranded breaks, mediated by the A-family polymerase Pol θ and hence termed Pol θ-mediated end joining (TMEJ). Traces of mutagenesis by TMEJ are abundant in evolutionary separated *C. elegans* strains and bear also resemblance to mutational patterns identified in cancers, highly encouraging further analysis of Pol θ function under physiological and cancerous conditions.

NEDERLANDSE SAMENVATTING (populair-wetenschappelijk)

DNA moleculen zijn de dragers van genetische informatie in de cel. De informatie ligt opgeslagen in de opeenvolging van vier verschillende bouwstenen: de basen A, C, T en G, die er samen voor zorgen dat de juiste eiwitten in de cel gemaakt worden. Beschadigingen in deze basen door straling of door reacties in de cel kunnen ertoe leiden dat onjuiste informatie wordt afgelezen, en eiwitten niet of op de verkeerde manier gemaakt worden. Voor het functioneren van een organisme is het noodzakelijk dat bij iedere celdeling beide dochtercellen correcte kopieën van al het DNA ontvangen. Een opeenstapeling van fouten ofwel mutaties in een cel kan er uiteindelijk toe leiden dat de 'rem' op celdeling gemuteerd raakt. De cel kan dan ongeremd gaan delen en een tumorcel worden.
Echter, een beperkte mate van mutaties in het DNA is essentieel voor het leven. Op die manier ontstaat er namelijk genetische variatie, die nodig is tijdens de evolutie. Selectie op kleine veranderingen kan leiden tot betere aanpassing van een soort aan de omgeving.

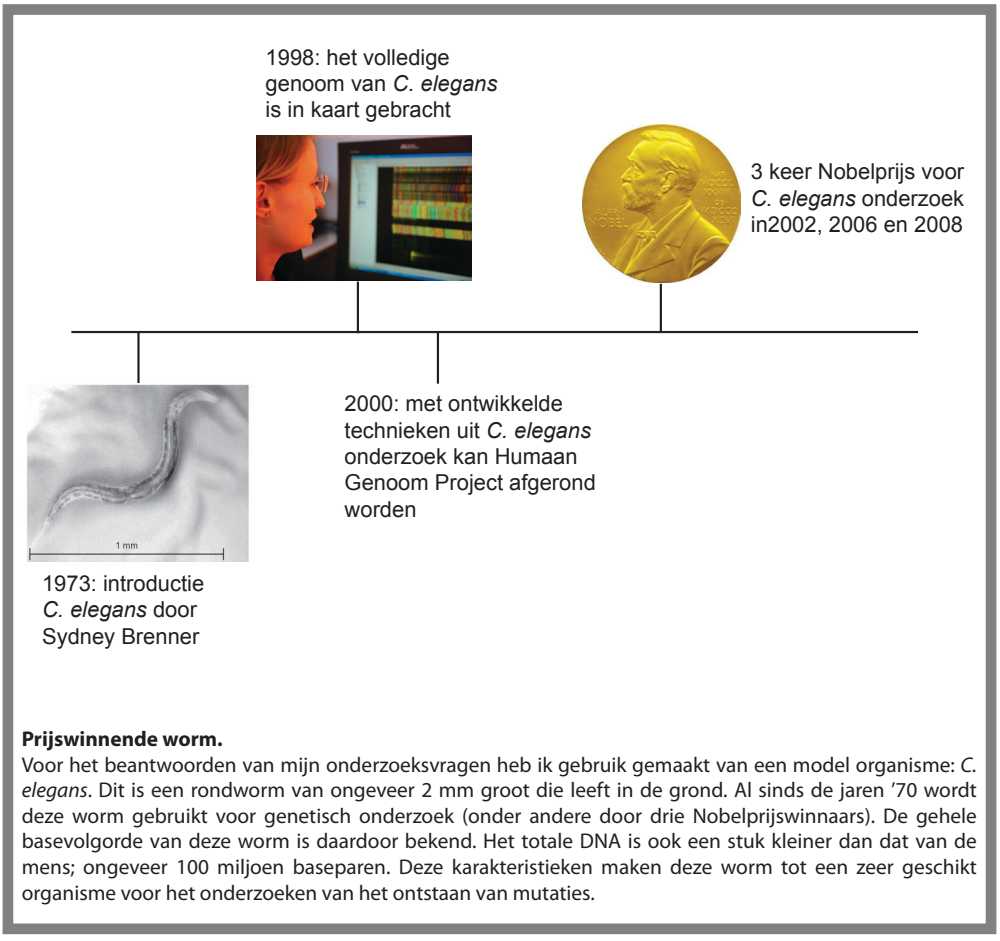
De kernvraag in dit proefschrift is hoe organismen zorgen voor genoomstabiliteit - dus het intact houden van de correcte basevolgorde in hun DNA. Hoe worden mutaties in het DNA door beschadigingen door straling of reactieve moleculen in de cel vermeden? Tegelijk moet dit systeem niet honderd procent onfeilbaar zijn, want af en toe een foutje is juist nodig om evolutionaire aanpassing mogelijk te maken.

In mijn proefschrift kijk ik naar de effecten van een groep enzymen op genoomstabiliteit: DNA polymerases. Dit zijn eiwitten die nodig zijn bij de replicatie: het kopiëren van het DNA voor de celdeling. In de meeste gevallen wordt dit gedaan door zeer precieze polymerases, die eventuele fouten zelf kunnen herstellen. Echter, in sommige gevallen, als het DNA molecuul dat gekopieerd moet worden, beschadigd is, kunnen deze polymerases hun werk niet doen. In dat geval kan de cel overschakelen op een alternatieve strategie met behulp van translesie synthese (TLS) polymerases. Deze eiwitten kunnen wel langs de beschadiging en daarmee voorkomen dat de cel doodgaat. Door 'bypass' van de schade, voorkomen ze ook dat het DNA breekt op de plaats van de beschadiging tijdens de replicatie. Zo'n breuk is erg gevaarlijk voor een cel, want als de eindjes van de breuk verkeerd aan elkaar geplakt worden, kan er genetische informatie verloren gaan. De keerzijde van bypass door TLS polymerases is wel dat deze enzymen zelf niet zo precies zijn: af en toe bouwen ze de verkeerde base in en laten ze dus ook mutaties achter.

TLS polymerases: Quick and dirty

Een veelgebruikte manier om de functie van een eiwit te bestuderen, is het bijbehorende

gen uitschakelen. Vergelijk de situatie zonder dit gen met de situatie mét dit gen, en je hebt een idee wat het eiwit (gemaakt door het gen) nu eigenlijk doet in de cel. Met deze gedachte bestudeer ik in hoofdstuk 2 van mijn proefschrift wat het effect is van het uitschakelen van twee belangrijke TLS polymerases tijdens de ontwikkeling.
Ik heb vergeleken hoe wormenstammen met of zonder de TLS polymerases Pol η en Pol κ reageren op verschillende soorten DNA schade. Uit deze experimenten bleek dat vooral Pol η van groot belang is bij de bescherming van het DNA tegen verschillende bronnen van beschadiging. Dit geldt onder andere voor UV-straling, γ-straling en de stof cisplatina, die ook in chemotherapie gebruikt wordt. Wormen blijken veel gevoeliger dan zoogdiercellen in weefselkweek. Een mogelijke reden is dat TLS polymerases vooral heel belangrijk zijn tijdens de eerste delingen van een organisme. Later in de ontwikkeling zijn er meer reparatie-enzymen beschikbaar, die de beschadiging kunnen verwijderen.



Prijswinnende worm.
Voor het beantwoorden van mijn onderzoeksvragen heb ik gebruik gemaakt van een model organisme: *C. elegans*. Dit is een rondworm van ongeveer 2 mm groot die leeft in de grond. Al sinds de jaren '70 wordt deze worm gebruikt voor genetisch onderzoek (onder andere door drie Nobelprijswinnaars). De gehele basevolgorde van deze worm is daardoor bekend. Het totale DNA is ook een stuk kleiner dan dat van de mens; ongeveer 100 miljoen baseparen. Deze karakteristieken maken deze worm tot een zeer geschikt organisme voor het onderzoeken van het ontstaan van mutaties.

Maar in de vroege embryogenese moet een cel vooral heel vaak achter elkaar delen. Er is dan geen tijd voor reparatie; dus de 'quick and dirty' methode waarmee de TLS polymerases schade kunnen passeren komt dan erg van pas.

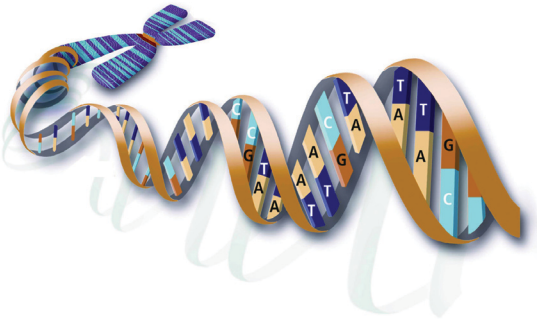
Verschillende lagen van bescherming

In hoofdstuk 3 ga ik verder in op de rol van de TLS eiwitten Pol η en Pol κ. In dit hoofdstuk bestudeer ik wat het effect is van deze eiwitten als de wormen onder 'normale' condities groeien, dus zonder blootstelling aan een hoge dosis DNA schade, zoals beschreven in hoofdstuk 2. Hiervoor laat ik parallelle lijnen groeien en breng ik vervolgens hun gehele basenvolgorde in het DNA in kaart. Op die manier kunnen we zien welke veranderingen in het DNA gedurende die tijd ontstaan zijn. Tot onze verbazing zagen we een heel specifieke soort verandering optreden: fragmenten van ongeveer 100 baseparen worden op willekeurige plaatsen uit het DNA geknipt. Deze mutaties worden deleties genoemd, en treden alleen op als Pol η en Pol κ hun werk niet kunnen doen. Ze duiden erop dat het DNA inderdaad breekt als Pol η en Pol κ niet in de buurt zijn bij een beschadiging.
Vervolgens heb ik de DNA volgorde rond deze deleties zorgvuldig bestudeerd. Daarbij sprongen twee kenmerken in het oog. Ten eerste is in de meeste gevallen de eerste base links van de deletie identiek aan de base aan de rechterkant. Daarnaast zijn er een behoorlijk aantal gevallen waar DNA van de linker- of de rechterkant van de deletie in het 'gat' geplakt is.
Deze drie kenmerken; de deletiegrootte, de identieke basen links en rechts, en de toevoeging van extra stukjes DNA, deden ons vermoeden dat een ander polymerase betrokken is bij het genereren van deleties (zie kader). Uit andere experimenten vermoedden we dat dit polymerase Pol θ is.
Vervolgens heb ik getest wat er gebeurt als ik niet alleen de TLS polymerases, maar ook Pol θ uitschakel. Inderdaad treden er geen deleties rond de 100 basenparen meer op. In plaats daarvan verdwijnen nu veel grotere stukken van het DNA, van minstens 10.000 basenparen. Je zou dus eigenlijk kunnen zeggen dat er twee 'lagen' van bescherming zijn: Pol η en Pol κ zorgen voor 'bypass' van DNA schade en zorgen ervoor dat er geen breuken in het DNA optreden. Pol θ zorgt ervoor, dat als er dan toch een breuk ontstaat, er niet al te veel DNA verloren gaat.

Springend DNA

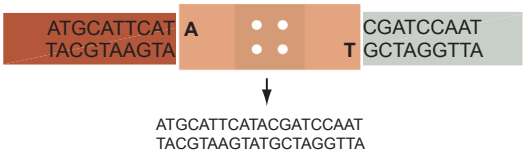
In hoofdstuk 4 gebruik ik een heel andere proefopzet om naar de functie van Pol θ te kijken. Hier ontstaan de breuken in het DNA niet op een 'indirecte' manier door de TLS polymerases uit te schakelen (zoals in hoofdstuk 3), maar maak ik breuken door transposons in het DNA te laten 'springen'. Transposons zijn fragmenten in het DNA die er spontaan uitgeknipt kunnen worden, en dan dus een breuk in het DNA achterlaten. Normaal gesproken gebeurt dit niet in de geslachtscellen van een organisme; al die

breuken in het DNA zouden tot veel te veel fouten leiden. Maar als we de factoren uitschakelen die dit proces normaal gesproken remmen, ontstaan er spontaan breuken in het DNA.
In dit hoofdstuk bestudeer ik de reparatie van zo'n breuk in een spiergen van de worm. Als dit gen niet goed werkt, gaat de worm 'twitchen'; hij beweegt alsof hij onder stroom staat. Het handige van dit gen is dat het heel groot is en er redelijk eentonig uitziet. Als er een klein stukje ontbreekt, maar de eindjes worden op de juiste manier aan elkaar geplakt, dan doet het gen het prima, en beweegt de worm weer normaal.
Dit gen kunnen we goed gebruiken om reparatie van een breuk af te lezen. We nemen wormen met een transposon in het spiergen, die dus 'twitchen' omdat het spiergen niet goed werkt. Vervolgens laten we het transposon er uitspringen, en een breuk ontstaan. Gerepareerde breuken kunnen we nu herkennen, omdat deze wormen weer normaal kunnen bewegen. Dit doen we ook in wormen waar bovendien Pol θ uitgeschakeld is, omdat we uit het vorige hoofdstuk weten dat dit gen betrokken is bij het repareren van breuken.



Chromosomen bestaan uit lange DNA moleculen, die weer bestaan uit twee strengen met de vier basen die de bouwstenen vormen. De basen uit beide strengen kunnen maar op één manier aan elkaar binden: A bindt aan T, C bindt aan G.

Een breuk in het DNA kan gerepareerd worden door knip- en plakwerk van Pol θ. Dit enzym kan bijvoorbeeld een A aan de ene kant aan een T aan de andere kant van de breuk plakken.



Soms voegt Pol θ een stukje van de DNA volgorde die naast de breuk ligt in.

Pol θ plakt de eindjes van een breuk aan elkaar.

Inderdaad, is de reparatie met Pol θ veel efficienter dan zonder Pol θ. Bovendien zien we, net als in het vorige hoofdstuk dat Pol θ bijna altijd een identieke base links en rechts van de breuk achterlaat, en weer af en toe een stukje DNA invoegt.
Deze resultaten vormen een bewijs dat Pol θ inderdaad breuken kan repareren; iets wat we hiervoor alleen op een indirecte manier konden aantonen.

Zieke wormen

Hoofdstuk 5 is een soortgelijke studie als het onderzoek naar Pol η en Pol κ in hoofdstuk 2 en 3, maar nu kijk ik naar een derde TLS polymerase: Rev1. Als Rev1 uitgeschakeld wordt zie ik vergelijkbare effecten: wormen zijn gevoeliger voor DNA schade en er ontstaan spontaan deleties.
Echter, met deze wormen is meer aan de hand. Ze zijn ook nog eens heel vaak steriel (onvruchtbaar) of krijgen maar heel weinig nageslacht.
Het lijkt er dus sterk op dat Rev1 twee zeer verschillende functies heeft in de cel: bypass van schade net als de andere TLS polymerases. Maar daarnaast zorgt het er ook voor dat er voldoende geslachtscellen kunnen rijpen. Hoe Rev1 dat precies doet, is nog onduidelijk. Mogelijk is Rev1 nodig om bepaalde stukken DNA actief te maken of juist te remmen.

Conclusies

Uit dit onderzoek blijkt dat de TLS polymerases Pol η en Pol κ en Rev1 alledrie van groot belang zijn voor het beschermen van een organisme, zowel tegen spontane schade in de cel, als tegen schade van buitenaf door bijvoorbeeld UV-licht. Vergeleken met eerder onderzoek in geïsoleerde cellen, is hun rol veel groter dan gedacht, omdat ze specifiek in een delend embryo hun werk kunnen doen.
Daarnaast stuitten we op een onverwachte vinding: er ontstaan spontaan breuken, die specifiek door een ander polymerase, Pol θ, gerepareerd worden. Dit is belangrijk omdat dit een heel andere route is voor het repareren van breuken dan tot nu toe bekend was.
Pol θ doet dat altijd op een heel specifieke wijze; door steeds dezelfde base voor en na de breuk aan elkaar te plakken, en soms een extra stukje flankerend DNA in te voegen. Het herkennen van dit patroon is erg belangrijk, want bij tumoren worden soms dergelijke mutatiepatronen gevonden. Mogelijk speelt ook hier Pol θ een rol, en leidt dit tot de fouten die ervoor gezorgd hebben dat een normale cel in een tumorcel is veranderd.

DANKWOORD

Met heel veel plezier en voldoening kijk ik vandaag terug op mijn promotietraject nu alleen 'bypass' van de laatste blokkade - de verdediging - mij nog rest. Ik prijs mij heel gelukkig met een betrokken omgeving van collega's, vrienden en familie, die ervoor gezorgd hebben dat de hindernissen op mijn pad nooit tot een voortijdige 'stall' of 'collapse' geleid hebben.

Marcel, jouw positieve blik, scherpte, nieuwsgierigheid, en bovenal plezier in het bedrijven van wetenschap zijn een grote inspiratie voor me. Er was een verkiezingsuitkomst voor nodig om me naar jouw groep te brengen, maar van die overstap heb ik nooit spijt gehad. Niet in de laatste plaats door het inspirerende gezelschap van alle Tijstermannen en -vrouwen gedurende de afgelopen jaren.
Tijstervrouwen van het eerste uur in het Hubrecht: Evelien, Daphne, Karin en Jennemiek, dank voor alle gezelligheid aan de bench en daarbuiten, en de leerzame lessen over het *C. elegans* onderzoek. Wouter, vanaf het begin van onze Leidse tijd ben jij jarenlang mijn eerste aanspreekpunt geweest in lab en kantoor, en was je gelukkig bijna altijd een vrolijk aanspreekpunt. Blij dat je me nu ook als paranimf bij wil staan. Robin, zonder jou was ik nooit ergens gekomen met de sequencing data, onze samenwerking heeft tot een prachtresultaat geleid. Bennie, bij jou ontstaan ideeën sneller dan ze ooit in woorden gevat kunnen worden, blij dat ik in ieder geval een deel mee heb kunnen pikken. Ron, de proefopzet kan jou niet groot genoeg zijn, dat zal de significantie van de beschreven data zeker ten goede gekomen zijn. Jane, ik hoop dat je de Tijstermannen nog een beetje de baas bent als ik weg ben, al heb ik daar niet al te veel twijfels over. Ivo, al pratend over eten waren er zo weer honderd wormen gesingeld. Maartje, na al die jaren zaten we ineens weer samen aan de bench, en gelukkig beviel het nog even goed als tijdens ons eerste practicum. Heel erg bedankt voor je vriendschap en steun al die jaren, en de gastvrijheid bij jou en Peter. Ik ben heel blij dat je mijn paranimf wilt zijn. Previous lab members Nick, Marijn, Evelina, Jordi, thanks for your contributions to the group both scientifically and socially. Carine, wat bofte ik met zo'n getalenteerde studente, met wie ik het ook nog eens zo goed kon vinden.
Dank ook aan alle andere collega's van Toxicogenetica voor alle hulp en gezelligheid op de afdeling. Werden die Utrechters en hun wormen eerst nog met enig argwaan beschouwd, algauw werden we toch liefdevol opgenomen op onze nieuwe stek. Thank you, Anastasia, Fabienne, Jaap, Pop, Wouter for the good atmosphere in our office. Jaap, heel veel dank voor je commentaar op hoofdstuk 1 en 2.

Heel veel dank ook aan mijn rijke omgeving buiten het lab, voor interesse in mijn werk of juist de nodige relativering. Emma, ons duo is misschien wel één van de beste overblijfselen van mijn Hubrechttijd, ik hoop dat we nog veel mooie muziek blijven maken samen. Josien, dank dat je ons samenspel zo'n vliegende start hebt gegeven,

jouw plaatsje in het Hubrecht trio bleek onvervangbaar.
Dank ook aan de vele NIVONners die altijd te porren zijn voor bergen, buiten zijn en lekker eten...ook een perfecte manier om los te komen van het lab. Laura, thanks for joining this crowd now and then.
In mijn hernieuwde Leidse episode was er gelukkig ook weer vaak tijd voor mijn Leidse vrienden, dank dat ik altijd bij jullie terecht kon na een lange werkdag! Lieve Sas, het Grieks in mijn proefschrift is weliswaar beperkt tot drie letters, desalniettemin was het delen van ervaringen gedurende deze periode onmisbaar.
Met al dat gereis was het niet altijd makkelijk om me ook in Utrecht thuis te voelen, maar gelukkig is dat uiteindelijk helemaal goedgekomen, met grote dank aan mijn omgeving daar, thanks Chinks for being such a kind neighbour, Linda, Tesse, dank voor jullie betrokkenheid en vriendschap.

Lieve Okke, door jou ben ik nog meer een nomade in de Randstad geworden, maar dit geheel zonder spijt. Met jou erbij is de eindsprint van dit traject zoveel leuker!
Ten slotte heel veel dank aan mijn familie, die me op cruciale momenten aan stelling 12 herinnert. Hugo, super dat jij er zo'n mooi boek van hebt gemaakt. Pap, mam, ik heb zoveel meer aan jullie te danken dan een setje goed gerepliceerde genen.

CURRICULUM VITAE

Sophie Roerink is geboren op 21 augustus 1982 te Almelo. Zij behaalde in 2000 haar gymnasiumdiploma aan scholengemeenschap het Noordik te Almelo, en begon in datzelfde jaar met een studie Bio-Farmaceutische Wetenschappen aan de universiteit Leiden.

Tijdens haar studie deed zij een onderzoeksstage bij de afdeling Medicinal Chemistry van het Leiden Amsterdam Center for Drug Research onder begeleiding van Prof. IJzerman, naar nieuwe liganden voor de Adenosine A1 receptor. Tevens behaalde zij in deze tijd een propedeuse (cum laude) in Griekse en Latijnse Talen en Culturen aan de Universiteit Leiden. Voor haar tweede onderzoeksstage vertrok zij naar Edinburgh (UK) om in de groep van Catherina Becker aan het Center for Neuroscience Research onderzoek te doen naar regeneratie van axonen in zebravissen.

Na het succesvol afronden van haar doctoraalopleiding, begon Sophie in augustus 2006 begon aan promotieonderzoek in het lab van prof. Ronald Plasterk aan het Hubrecht Instituut te Utrecht. In 2007 kwam zij in de groep van Marcel Tijsterman, eveneens aan het Hubrecht Instituut, waar ze haar onderzoek startte naar genoomstabiliteit in *C. elegans*. Dit onderzoek werd per augustus 2009 voortgezet in het Leids Universitair Medisch Centrum, bij de afdeling Toxicogenetica. Resultaten van dit onderzoek worden beschreven in dit proefschrift.

In januari 2014 zal zij haar onderzoekscarrière voortzetten als Postdoctoral Research Fellow in de groep van prof. Mike Stratton aan het Sanger Institute (Cambridge, UK).

LIST OF PUBLICATIONS

S.F. Roerink, W. Koole, L.C. Stapel, R.J. Romeijn and M. Tijsterman. A broad requirement for polymerases η and κ, and interacting sumoylation and nuclear pore proteins, in lesion bypass during *C. elegans* embryogenesis.
Plos Genet 2012 vol 8 (6)

S.F. Roerink*, R. van Schendel* and M. Tijsterman. A mutagenic repair pathway dependent on polymerase θ protects the *C. elegans* genome against large chromosomal deletions.
*these authors contributed equally to this work
Manuscript submitted for publication

S.F. Roerink, K. Brouwer and M. Tijsterman. Pol θ - mediated end joining of transposition-induced double strand breaks in the *C. elegans* germline
Manuscript in preparation for publication

E.C. Klaasse, R.A.F de Ligt, **S.F. Roerink**, A.Lorenzen, G. Milligan, R. Leurs, and A.P. IJzerman. Allosteric modulation and constitutive activity of fusion proteins between the adenosine A1 receptor and different 351Cys-mutated Gi alpha-subunits.
Eur. J. Pharmacol 2004, vol. 499 (pp 91-98).

E.C. Klaasse, G. van den Hout, **S.F. Roerink**, W.J. de Grip, A.P. IJzerman and M.W. Beukers. Allosteric modulators affect the internalization of human adenosine A1 receptors.
Eur. J. Pharmacol. 2005, vol. 522 (pp 1-8).

L.C.W. Chang, J.K. von Frijtag Drabbe Künzel, T. Mulder-Krieger, R.F. Spanjersberg, **S.F. Roerink**, G. van den Hout, M.W. Beukers, J. Brussee, and A.P. Ijzerman. A series of ligands displaying a remarkable agonistic-antagonistic profile at the adenosine A1 receptor.
J. Med. Chem. 2005, vol. 48 (pp 2045-2053).