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).

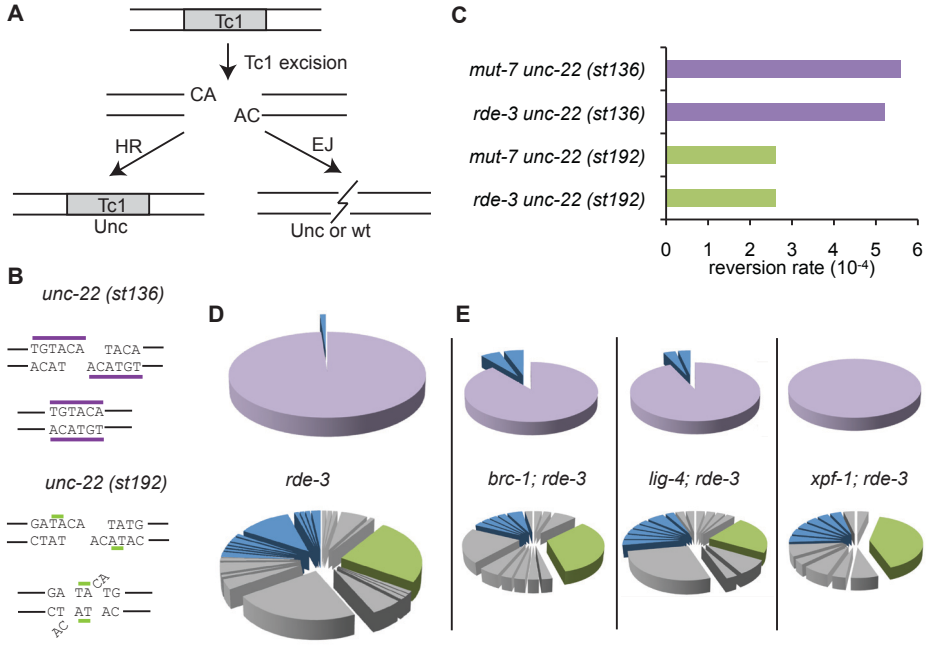

FIGURE 1. Germline repair of transposon-mediated DSBs.
A. Mechanism of Tc1-mediated DSB repair. The transposase Tc1A excises Tc1 transposons resulting in staggered cuts with 3' overhangs of two nucleotide. Such breaks can be repaired in an error-free way via HR or in an error-prone way via EJ. Only in 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

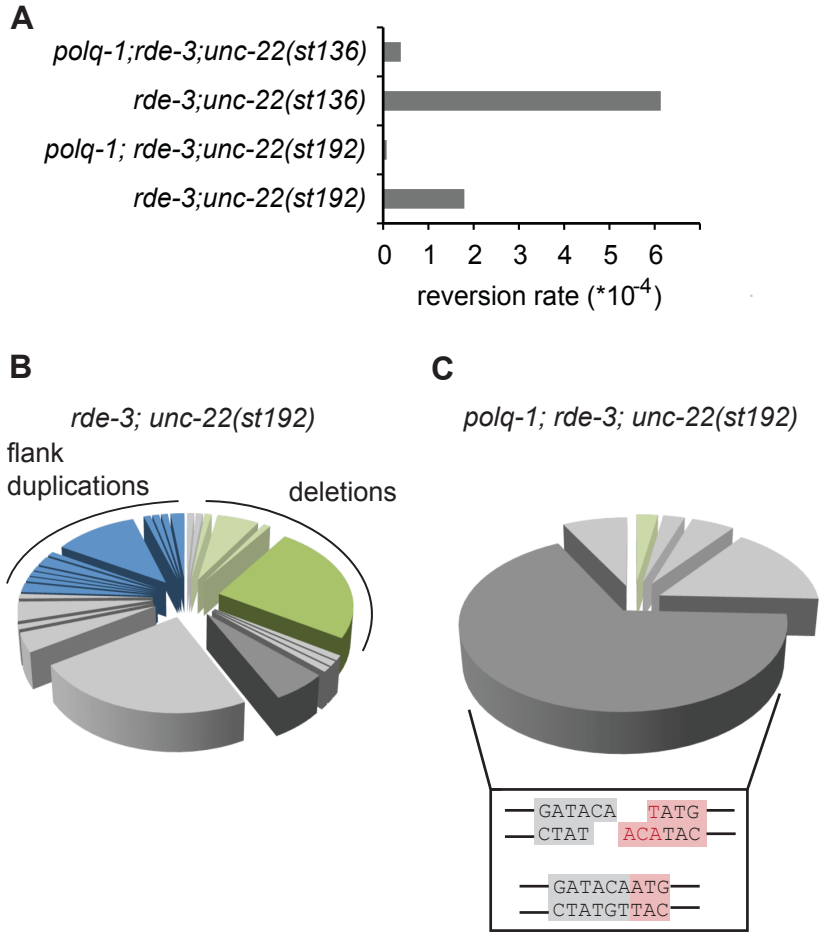

FIGURE 2. *polq-1* deficiency suppresses *unc-22* reversion.
A. Reversion frequency of *unc-22(st136)* and *unc-22(st192)* alleles in *rde-3* and *polq-1;rde-3* background. **B.** Distribution of footprints in *rde-3;unc-22(st192)* mutants. Footprints that display flank duplications are indicated in 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*

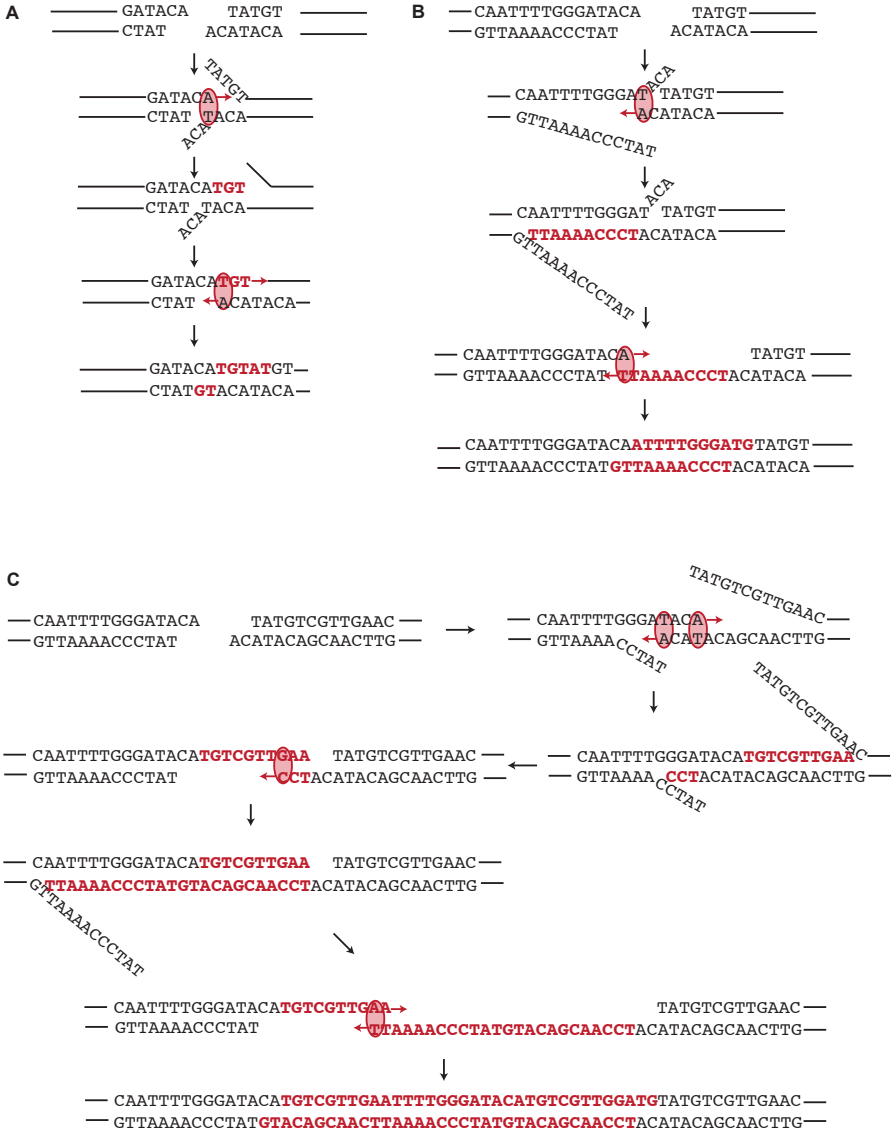

FIGURE 3. TMEJ explains complex repair footprints after transposon excision.
Extension from a single basepair is catalyzed by Pol θ, depicted by a red oval. Newly synthesized nucleotides are marked in red. **A.** Apparent blunt end joining explained via a mechanism of single nucleotide base pair priming, 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
```