DNA Repair Mechanisms in *C. elegans*

DNA Herstelmechanismen in *C. elegans*

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. J.C. Stoof, ingevolge van het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 6 maart 2009 des middags te 2.30 uur

door

Karin Brouwer

geboren op 3 augustus 1980 te Leiden

Promotor: Prof. dr. ir. E. P. J. G. Cuppen

Co-promotor: Dr. M. Tijsterman

Het drukken van dit proefschrift is mede mogelijk
gemaakt met een financiële bijdrage van de J.E. Jurriaanse stichting.

“It is though to beat a person who never gives up”
Babe Ruth

“Nothing shocks me, I am a scientist”
Indiana Jones

“Science is a way of trying not to fool yourself. The first principle is that you must not fool yourself, and you are the easiest person to fool”
Richard Feynman

“My mind is not perfect, but it is sincere”
The View

“Science is a wonderful thing if one does not have to earn one’s living at it”
Albert Einstein

“They say that time changes things, but you actually have to change them yourself”
Andy Warhol

“There is science, logic, reason; there is thought verified by experience. And then there is California”
Edward Abbey

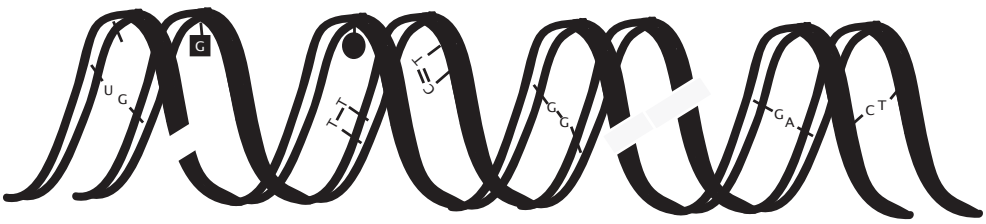
“No science is immune to the infection of politics and the corruption of power”
Jacob Bronowski

ISBN: 978-90-393-5017-1

Table of contents

Chapter 1

General introduction



DNA is the carrier of genetic information and maintenance of its structure and the conservation of the information it carries are essential for survival of cells. DNA is chemically reactive and its structure is very easily disturbed by damaging agents. These agents can originate from endogenous as well as exogenous sources. Also, prior to every cell division, the DNA has to be replicated. This replication process is extremely accurate, nonetheless it is not error-free.

Throughout evolution, cells have developed DNA repair mechanisms that enable them to protect their DNA and keep its information intact. Here, I will give a brief overview of the various mechanisms, their relevance in disease and aging, and in greater detail I will describe the repair mechanisms that are of importance for the studies performed in this thesis.

DNA Damage

DNA damage can derive from three sources. Firstly, environmental agents such as UV light, ionizing radiation and a wide variety of genotoxic chemicals target the DNA. Secondly, (by)products of normal cellular processes, like reactive oxygen species (ROS) can damage the DNA from within the cell. Lastly, chemical bonds in DNA can spontaneously disintegrate under physiological conditions. The inhibition of transcription, replication, and chromosome segregation by lesions leads to cell-cycle arrest and apoptosis. Also, when not properly repaired, DNA lesions will lead to mutations and chromosome aberrations that cause cancer, aging, and inborn disease (reviewed in [1]).

DNA Repair Mechanisms, an Overview

The plethora of DNA damages is repaired by specialized processes (as depicted in Figure 1). Direct reversal (DR) is mostly used in the repair of alkylated bases. The major advantage of DR is that it does not need sequence information to repair. Two DR mechanisms exist: 1) suicidal methyltransferases that transfer the methyl group from the DNA to a cysteine residue in the transferase itself, and 2) a newly identified AlkB family of dioxygenases that directly reverse the damage by oxidative demethylation [2].

Base excision repair (BER) is responsible for repair of the most common lesions. Short-patch BER removes the damaged base and replaces it, using the complementary strand as a template [3]. Long-patch BER works via a similar mechanism and differs in the fact that it involves the removal of a few bases instead of just one. This mechanism is used in the repair of single-strand breaks [1].

Bulky lesions that affect one of the two DNA strands and disturb the structure of the DNA helix are repaired by nucleotide excision repair (NER). NER consists of two pathways. Global genome repair deals with damages that occur throughout the genome, while transcription-coupled repair works specifically on damages in the transcribed strand of active

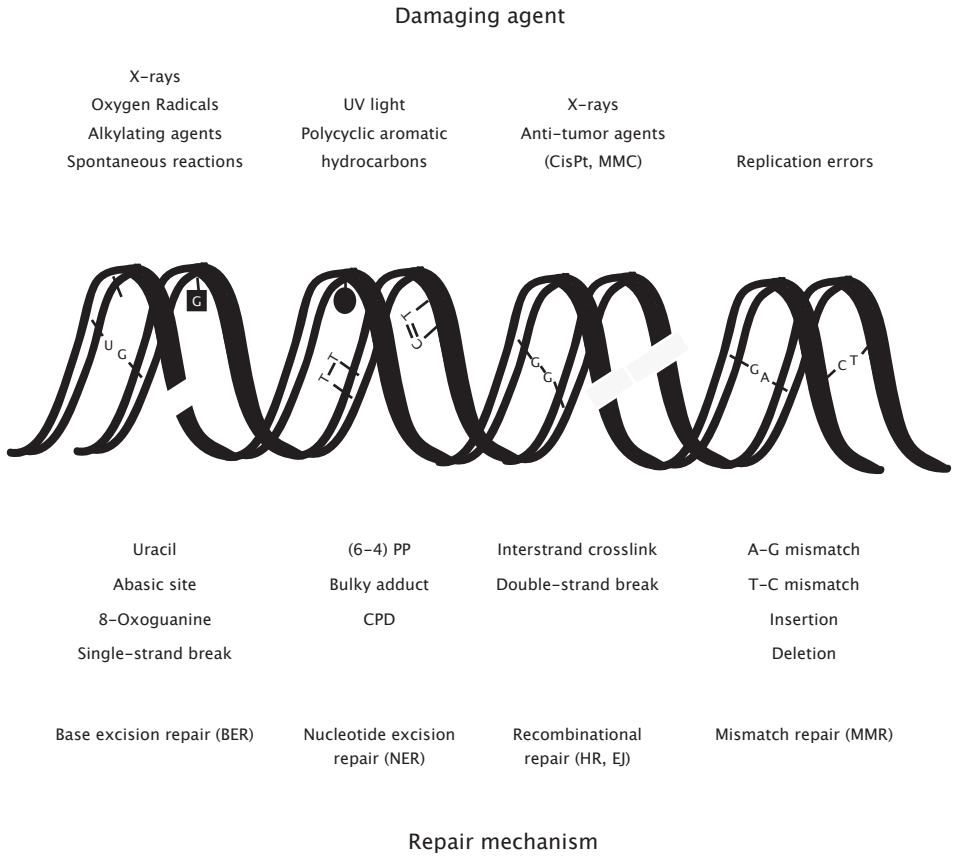


Figure 1 DNA damaging agents, DNA lesions, and DNA repair. A general overview of the types of damaging agents, the correlating DNA lesions they cause, and the DNA repair mechanisms responsible for the proper repair of these lesions (adapted from [1]).

genes. Both pathways use different mechanisms of damage recognition and subsequently use the same core mechanism for the repair of the lesion [4].

Double-strand breaks (DSBs) can be repaired via several pathways. Homologous recombination (HR) repair allows using the sister chromatid as a template for error-free repair. When no template is available, error-prone end-joining occurs: by joining the ends together using sequence homology on both sites of the break (single-strand annealing [SSA]) or, in a less specified way, by ligating the ends together (non-homologous end-joining [NHEJ]). Both pathways frequently involve the loss or gain of sequence (reviewed in [5]).

DNA adducts that form interstrand crosslinks (ICL) between the two DNA strands cause problems during replication. Because they physically link the two DNA strands, separation of the two strands by the replication machinery is no longer possible. The ICL repair response relies on a range of genes including the Fanconi Anemia genes. These genes function specifi-

cally in ICL repair. Additionally, genes that also function in other repair pathways, like NER and HR, are implicated (reviewed in [6]).

Mismatch repair (MMR) is involved in the repair of replication errors. These errors are either mismatches that result from incorporation of a wrong base by the replicating polymerase, or insertions and deletions that occur due to slippage of the polymerase during replication of, in most cases, repetitive sequences. The polymerase itself has proofreading activity that allows high-fidelity replication. The MMR system scans newly replicated DNA and removes overseen errors, further reducing the replication error rate (reviewed in [7]).

Some DNA replication initiates before DNA damage repair has been completed. The remaining damages can cause replication blocks and threaten the ability of a cell to finish the cell cycle. DNA damage tolerance mechanisms exist that relief replication arrest, leaving the damage for repair at some later time. One option is the use of mechanisms whereby the replicative machinery uses undamaged segments of the genome to copy DNA, thereby avoiding the need to replicate directly across the template-strand base damage. As a second option, the damage can be bypassed by specialized non-replicative polymerases in a process called translesion synthesis (TLS) (reviewed in [8]).

DNA Repair, Disease, and Aging

The importance of preventing genome instability is illustrated by the plethora of human (cancer susceptibility) syndromes that arise from defects in DNA repair genes.

Patients that suffer from Lynch (or HNPCC) syndrome show early onset of colon cancer and an 80% lifetime risk of colorectal cancer. The syndrome is marked by an autosomal dominant mode of inheritance. HNPCC patients carry germline mutations in MMR genes [9].

Fanconi Anemia is an autosomal recessive disorder that is characterized by a variety of symptoms. These symptoms include congenital abnormalities, defective haemopoeisis and a high risk of developing acute myeloid leukaemia and certain solid tumors. Proteins expressed by genes mutated in Fanconi patients are implicated in the ICL repair pathway [10].

NER is related to multiple syndromes, which are correlated to mutations in specific genes. Cockayne syndrome is caused by mutations in two genes encoding for proteins that function in NER, namely CSA and CSB. Patients have many developmental defects including severe physical and mental retardation, microcephaly, long limbs, bird-like face, pigmented retinophagy, gait defects and sunsensitivity [11]. Xeroderma pigmentosum (XP) syndrome symptoms are sunlight-induced pigmentation changes in the skin and dry parchment-like skin. Often there is extreme sensitivity to sunlight and photophobia and multiple skin cancers. Of seven complementation groups the mutated genes are NER factors [11]. There is one XP variant that is caused by mutation of the TLS polymerase Polη [12]. The third NER-

related syndrome is Trichothiodystrophy (TTD). TTD symptoms are sulphur-deficient brittle hair, small stature, mental retardation, ichtyotic skin, β-thalassaemia trait, unusual facial features and in many cases photosensitivity [11]. Interestingly, TTD can be caused by mutations in genes also associated with XP, i.e. certain mutations in XPD can lead to XP, while other XPD mutations lead to TTD [13], demonstrating that different mutations in the same gene can lead to very different diseases.

Multiple genes that are involved in the repair of DSBs are also related to human syndromes. Nijmegen breakage syndrome is caused by a mutation in the NBS1 gene. Symptoms include immunodeficiency, increased cancer risk, and growth retardation [14].

Bloom syndrome is caused by mutations in a single gene, namely the *BLM* gene that encodes a RecQ family helicase. Bloom syndrome is characterized by a wide variety of abnormalities, including stunted growth, immunodeficiency, fertility defects, sun sensitivity, increased frequency of diabetes, and greatly increased frequency of various types of cancer [15].

Mutations in a second gene encoding a RecQ helicase, the *WRN* gene, causes Werner syndrome. Werner syndrome is an adult progeria syndrome that includes a wide range of age-related traits having greatly accelerated onset although generally occurring after puberty. Patients die at an average age of 47, usually either from cancer or cardiovascular disease [15].

Carriers of mutations in the two tumor suppressor genes *BRCA1* and *BRCA2* (for breast cancer-related) have an increased risk to develop breast and ovarian cancers. For example, in the case of *BRCA1*, mutations lead to a breast cancer risk of about 70% at the age of 70 [16]. Both genes were shown to function in the HR pathway [15].

In a number of the above mentioned disorders, progeroid phenotypes are part of the symptoms [17]. The fact that these syndromes are caused by mutations in DNA repair genes led to the suggestion that improper repair of DNA damage causes premature aging. Indeed, mouse models for TTD exhibit many premature aging symptoms [18]. However, it was shown that the accelerated aging is not the consequence of increased genomic instability per se [19]. It has been hypothesized that repair mechanisms function on the one hand to prevent cancer by avoiding increased mutation rates caused by DNA damage and on the other hand to avoid premature aging caused by DNA damage-induced cell death/senescence/malfunction [20].

In this thesis, work will be described that involves a variety of DNA repair mechanisms. I will discuss mismatch repair, double strand break repair, nucleotide excision repair, translesion synthesis, and interstrand crosslink repair in detail. I will also discuss research done on these particular mechanisms in *C. elegans*, the model organism used in the experimental studies described in this thesis.

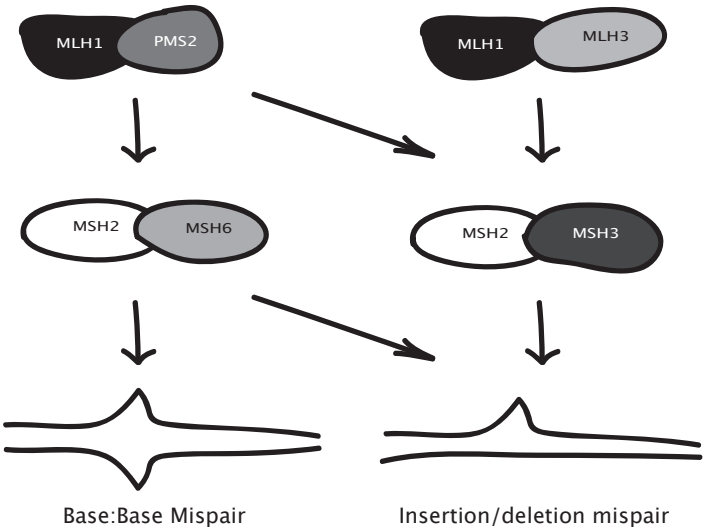


Figure 2: Schematic representation of the first steps in the eukaryotic mismatch repair mechanism. Two MutS-related complexes scan the genome for lesions that are then processed by MutL-related complexes. Overlap in specificity has been reported (adapted from [7]).

Mismatch Repair

Mismatch repair (MMR) deals with the correction of bases that are wrongfully inserted during replication and the loops that occur through slippage of the replication machinery on repetitive sequences. When not properly repaired, the first type of pre-mutagenic lesion leads to base changes, while the second leads to frameshifts.

In *E. coli* (reviewed in [21]), the MutS protein forms a homodimer that binds specifically to base-base mispairs and short loops (up to four nucleotides in length) and this complex initiates the MMR process [22]. A homodimer of MutL proteins is involved in coupling mismatch recognition by MutS to downstream steps [23-25]. These downstream steps involve MutH, a methylation-sensitive endonuclease that targets repair to the newly synthesized strand. The newly synthesized strand is not methylated yet and can therefore be distinguished from the template strand [26]. This way, it is possible to determine which of the two bases in the mispair actually is the wrongfully incorporated one. MutH induces a nick in the DNA on either side of the mismatch and subsequently exonucleases degrade the mismatch containing nicked strand [27, 28]. Because the exonucleases are single-strand specific, the activity of UvrD is needed to unwind the duplex molecule [29].

In eukaryotes, the system is more complex and less well understood (reviewed in [7, 21]). First, the variety of proteins involved is larger than in *E. coli*. Some of these proteins are homologous to MutS and MutL and therefore named MSH (MutS homologue) and MLH (MutL homologue). In addition, two parallel pathways exist that each deal with specific substrates (Figure 2). Base-base mispairs and small loops are recognized by an MSH2-MSH6

heterodimer and further processed by the MLH1-PMS1 heterodimer [30, 31]. Larger loops are mostly processed by MSH2-MSH3 and MLH1-MLH3 heterodimers [32, 33]. The MutS- and MutL-related complexes appear to have similar functions as their related complexes in *E. coli*, suggesting that recognition and initial processing of the mismatches occur via similar mechanisms. However, the downstream steps are more different and it remains unclear which proteins are involved. One of the crucial differences between *E. coli* and eukaryotes is the strand discrimination step. In eukaryotes, no hemi-methylation exists and the distinction between the template and the newly synthesized strand is not understood. This far, a variety of proteins have been implicated in the later stages of MMR, i.e. DNA polymerase δ, RPA, PCNA, RFC, Exonuclease 1, FEN1, and the exonucleases of polymerases δ and ε [34-41].

Mismatch repair in C. elegans

In *C. elegans* only four genes of the first steps in the MMR pathway appear to be conserved. Homologues for human MSH2, MSH6, MLH1, and PMS2 are present. The absence of homologues for the MSH3 pathway raises the question how this influences the repair capacity of the MMR pathway. It is possible that the mismatch repair system in *C. elegans* has evolved such that the substrate specificity of the four present homologues is less stringent than in other organisms, and that they can repair all types of lesions that are substrates for subpathways of MMR in other organisms. Alternatively, certain lesions are not repaired in *C. elegans*, possibly leading to a more dynamic genome. Because MSH3 mostly works on repeats and hardly on mismatches, this could mean that repetitive sequences are more frequently changing lengths than they do in other organisms.

In agreement with a lack of MSH3, the mutation spectrums and rates of *msh-2* and *msh-6* mutant animals are identical [42]. In both cases, mutation rates are about a hundred times higher than in wild type. This is in contrast to what is described in other organisms.

In our laboratory, the function of the *msh-6* gene was studied in detail [43]. It was shown that *msh-6* defective animals show increased somatic and germ line mutations. These mutations consist of point mutations and frameshifts. In addition, there is increased microsatellite instability similar as seen in HNPCC patients.

Independently, a study on *msh-2* was published [44]. Here, a similar increase in mutation rates and microsatellite instability was detected. Additionally, a decrease in DNA damage-induced germ line apoptosis was observed in *msh-2* mutants. Both studies report a decrease in viability in passaged cultures of mutant animals. This indicates that a functional MMR pathway is essential for both the regulation of short- and long-term genome stability.

A genome-wide RNAi screen to identify genes involved in avoiding repeat instability identified the four *C. elegans* MMR genes [45]. In addition, nine genes that have not been previously suggested in this process have been detected. Further characterization of one of these genes, *rev-1*, is described in Chapter 2 of this thesis.

Double Strand Break Repair

At least three pathways to repair DSBs are known (Figure 3): 1) Homologous recombination (HR), 2) Non-homologous end-joining (NHEJ), and 3) Single-strand annealing (SSA). Although these pathways are conserved throughout evolution, their relative contributions differ between lower and higher eukaryotes (reviewed in [5]). In yeast, DSBs are primarily repaired by HR, while in mammals HR and NHEJ are both important.

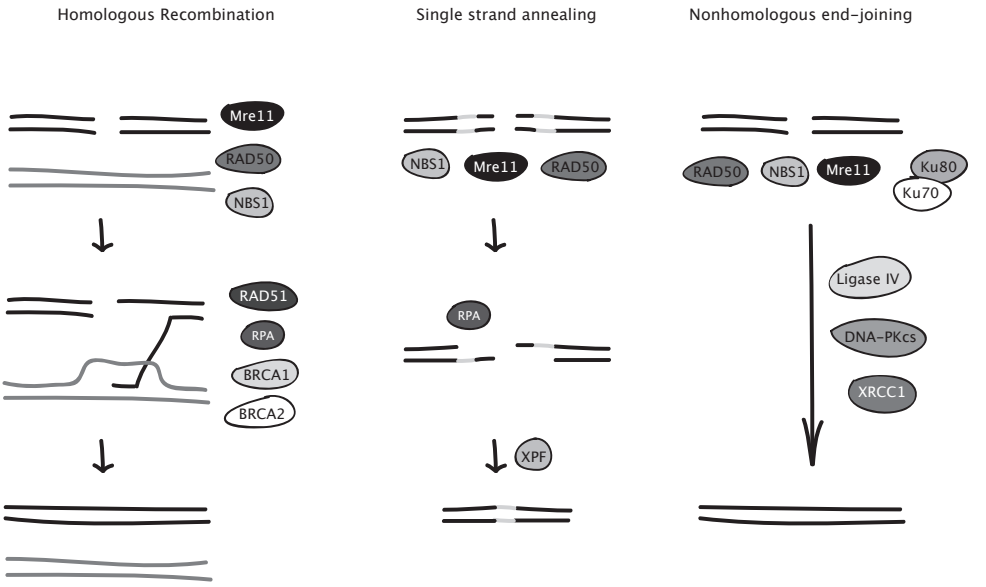


Figure 3: Overview of the three pathways involved in double-strand break repair. When a sister chromatid or homologous chromosome is present, the break can be repaired via homologous recombination. When no homologue is present, breaks can be repaired by two mechanisms that cause loss of sequence. In case of non-homologous end-joining the two ends are ligated together without much processing. Single-strand annealing requires strand resection and homologous sequences present at both sites of the break.

In HR, repair of a DSB occurs by the use of homologous sequence either on the sister chromatid or on the homologous chromosome. This allows for repair without loss of sequence and this makes HR the only conservative DSB repair pathway. In *S*. *cerevisiae*, RPA, all three replicative polymerases and the genes from the RAD52 genes group are implicated in HR [46]. In mammals, RAD52 complementation group genes were identified and due to redundancy, the number of genes is larger in comparison to yeast [5]. Additionally, in humans the breast cancer susceptibility genes BRCA1 and BRCA2 were shown to play a role in HR [47]; these proteins are not present in *S. cerevisae*.

Repair of DSBs by NHEJ is achieved by limited processing of DNA ends, followed by joining the ends, and re-ligation. NHEJ requires little or no sequence homology at the ends. Thus far, five factors involved in NHEJ have been identified: ligase IV and its associated protein

XRCC1, and the three components of the DNA-dependent protein kinase (DNA-PK) complex, namely Ku70, Ku80, and DNA-PKcs (reviewed in [5]).

SSA is the second type of end-joining and can be considered a special form of homology-driven repair. In this case, homology in both ends resulting from the break is used. After resection of the 5' ends and exposure of the regions of homology, the formation of joint molecules is possible. After removal of the non-homologous ends and DNA synthesis, repair of the break is completed by DNA ligase. Since the region in between the repeats is lost, SSA is a non-conservative mechanism (reviewed in [5]). Thus far, proteins required for SSA have only been identified in *S. cerevisiae*. The endonucleolytic activity of the Rad1p/Rad10p complex and the presence of the Msh2p and Msh3p proteins are necessary for the removal of non-homologous 3'-single-stranded ends [48-50]. Additionally, Srs2p, which has 5'-3' helicase activity, is required for SSA [51]. Rad52p and Rad59p play an important role in SSA, however in absence of these factors SSA still occurs, although at a dramatically reduced level [52]. In the absence of the Rad50p/Xrs2p/Mre11p complex, SSA is still possible, but at significantly reduced rates [53].

Double strand break repair in C. elegans

Thus far, study of HR in *C. elegans* has focused on two genes. The *C. elegans* BRCA1 homologue was found using the yeast two-hybrid system [54]. The putative ortholog of BRCA1-interacting protein BARD1, BRD-1, was used as bait in order to find a partner that could function similarly as BRCA1. This screen identified *brc-1* and depletion of both this gene and *brd-1* was shown to lead to high-incidence of males, elevated levels of p53-dependent germ cell death before and after IR, and impaired progeny survival and chromosome fragmentation after IR. This strongly implicates *brc-1* in the DNA damage response.

For the second breast cancer susceptibility gene, BRCA2, its *C. elegans* ortholog was also identified via a yeast two-hybrid interaction, in this case with RAD-51 [55]. The *brc-2* gene was shown to be much smaller than its human counterpart. Furthermore, the number of functional domains in BRC-2 is greatly reduced, compared to BRCA2. Yet, BRC-2 is functionally related to BRCA2. It binds to *rad-51* via its BRC domain and binds preferentially to single-stranded DNA via its OB-fold motif. In addition, *brc-2* mutants fail to repair meiotic or radiation-induced DSBs by HR due to inefficient RAD-51 nuclear localization and a failure in the targeting of RAD-51 to sites of DSBs. It has been suggested that BRC-2 not only regulates RAD-51 during HR, but that it can also function independently of RAD-51 in other DSB repair pathways.

Using purified proteins, the function of BRC-2 was further elucidated [56]. BRC-2 stimulates RAD-51-mediated D-loop formation and reduces the rate of ATP hydrolysis catalyzed by RAD-51. These functions depend on the direct binding to RAD-51 and the DNA binding activity of BRC-2. The RAD-51-independent role of BRC-2 might be in replacing the role of

vertebrate Rad52 that is involved in single-stranded annealing. Indeed, it was shown that BRC-2 can mediate the SSA of RPA-nucleotide complexes *in vitro*, similar to Rad52.

Controversy on the interaction of human BRCA2 with RAD51 has existed. The question was whether the BRC domain primarily interacts with RAD51 monomers or with the RAD51-DNA polymer, and whether there was a single interaction or multiple ones. The BRC motif of *C. elegans* BRC-2 was shown to contain two different RAD-51-binding regions [57]. One of these regions binds only weakly to RAD-51-DNA filaments, but strongly to RAD-51 alone. The second region appears to bind strongly to RAD-51-DNA-filaments. These results suggest a model where an interaction with RAD-51 alone is likely involved in filament nucleation, whereas a second interaction is involved in stabilization of the RAD-51-DNA filaments by BRC-2.

Currently, only one study on NHEJ in *C. elegans* is published. It describes the relative importance of HR and NHEJ in the *C. elegans* germ and somatic cells at multiple developmental stages [58]. Canonical NHEJ appears to be used exclusively by non-dividing somatic cells whereas HR is used to repair radiation-induced DNA damage in proliferating somatic cells and in germ cells. Error-prone NHEJ plays at most a small role in the repair of DSBs in the *C. elegans* germline. A study on the end-joining mechanisms that occur in the germ line is described in Chapter 6 of this thesis.

Two high-throughput in-liquid genome-wide RNAi screens performed in our laboratory aimed to identify genes involved in the response to DSBs. In the first approach, a strain that suffered elevated levels of DSBs, due to loss of transposon silencing in a mutator background, was used to screen for genes that are synthetic lethal with these elevated levels of DSBs [59]. This screen led to the identification of 32 genes that show synthetic lethality with the mutator phenotype. Of these genes, knockdown of ten led to sensitivity to DSB-inducing IR and six showed sensitivity to camptothecin, an agent that induces single-strand breaks that are transformed into DSBs during S phase. Of the six genes that are implied in the response to camptothecin, knockdown of five also showed response to IR. Of the ten genes sensitive to IR, nine appear to be involved in the prevention of IR-induced chromosomal aberrations.

The second screen directly identified genes involved in the response to IR-induced DSBs [60]. This approach led to the identification of 45 genes, including genes that were previously known to be involved in the DSB response. Knockdown of eleven genes led to impaired IR-induced cell cycle arrest and seven genes were essential for apoptosis upon exposure to IR. To further validate the identified genes, cell lines expressing siRNAs against human homologues were established. Of the eleven available lines, seven showed reduced survival after radiation.

Nucleotide Excision Repair

NER provides an important cellular defense against a variety of DNA alterations, mostly against bulky lesions that cause distortions of the helix structure. The proteins that recruit the core NER proteins depend on the genomic location of the damage. Apparently, there is a difference between the transcribed and non-transcribed regions of the genome [61, 62]. These differences led to the proposal of two subpathways: fast transcription-coupled repair (TCR), specific for the transcribed regions of the DNA, and relatively slow global genome repair (GGR) that can act throughout the genome (Figure 4). For TCR, it was suggested that it is triggered by a stalled RNA polymerase at the site of a damage in the transcribed strand [63]. The two Cockayne syndrome-related proteins CSA and CSB are involved in the displacement of the RNA polymerase and recruitment of the downstream NER factors [64, 65]. In GGR, the XPC-hHR23 is responsible for the damage-sensing step [66], while it is dispensable for TCR [4].

The core NER pathway is widely conserved through eukaryotes. The repair reaction can be subdivided into multiple steps. First, a preincision complex forms at the site of the damage and local DNA unwinding occurs; XPB and XPD, DNA-dependent ATPases with helicase function, are involved in unwinding the DNA [67, 68]. XPA and RPA are required for the formation of the pre-incision complex [69, 70]. It has been proposed that these proteins are able to double check DNA bending and unwinding and, consequently, they could serve as subunits that verify the damage-specific recruitment of NER factors. Next, on both sides of the damage the DNA is incised by XPG (3' site) and the ERCC1/XPF complex (5' site) [71]. The damage-containing oligo is removed. Finally, the resulting gap is filled by a polymerase and strand ligation completes the reaction [4].

Nucleotide excision repair in C. elegans

The first indication that the NER pathway is conserved in *C. elegans* came with the identification of nine mutants (*rad-1* to *rad-9*) that showed increased sensitivity to UV irradiation [72]. One of these mutants, *rad-3*, showed a defect in excision repair. This defect was more severe in larvae compared to embryos, suggesting a developmental regulation of DNA repair in *C. elegans* [73]. It was shown that the UV sensitivity in *rad-3* mutants is caused by a nonsense mutation in *xpa-1* [74]. Two *C. elegans* NER genes were cloned based on homology with their human counterparts, namely *xpa-1* and *xpf-1* [75, 76]. Currently, more *C. elegans* NER genes have been annotated [77], suggesting that most NER genes are conserved.

Study of the *xpa-1*/*rad-3* mutant indicated similarities to the phenotypes described in humans, but also differences came into view [74]. In absence of UV, the animals did not show noticeable developmental defects, hypersensitivity to oxidative damage and decrease in lifespan. Upon UV irradiation, the *xpa-1* mutants showed immediate growth arrest and displayed impaired survival and hypermutability. The UV-induced growth arrest and subse-

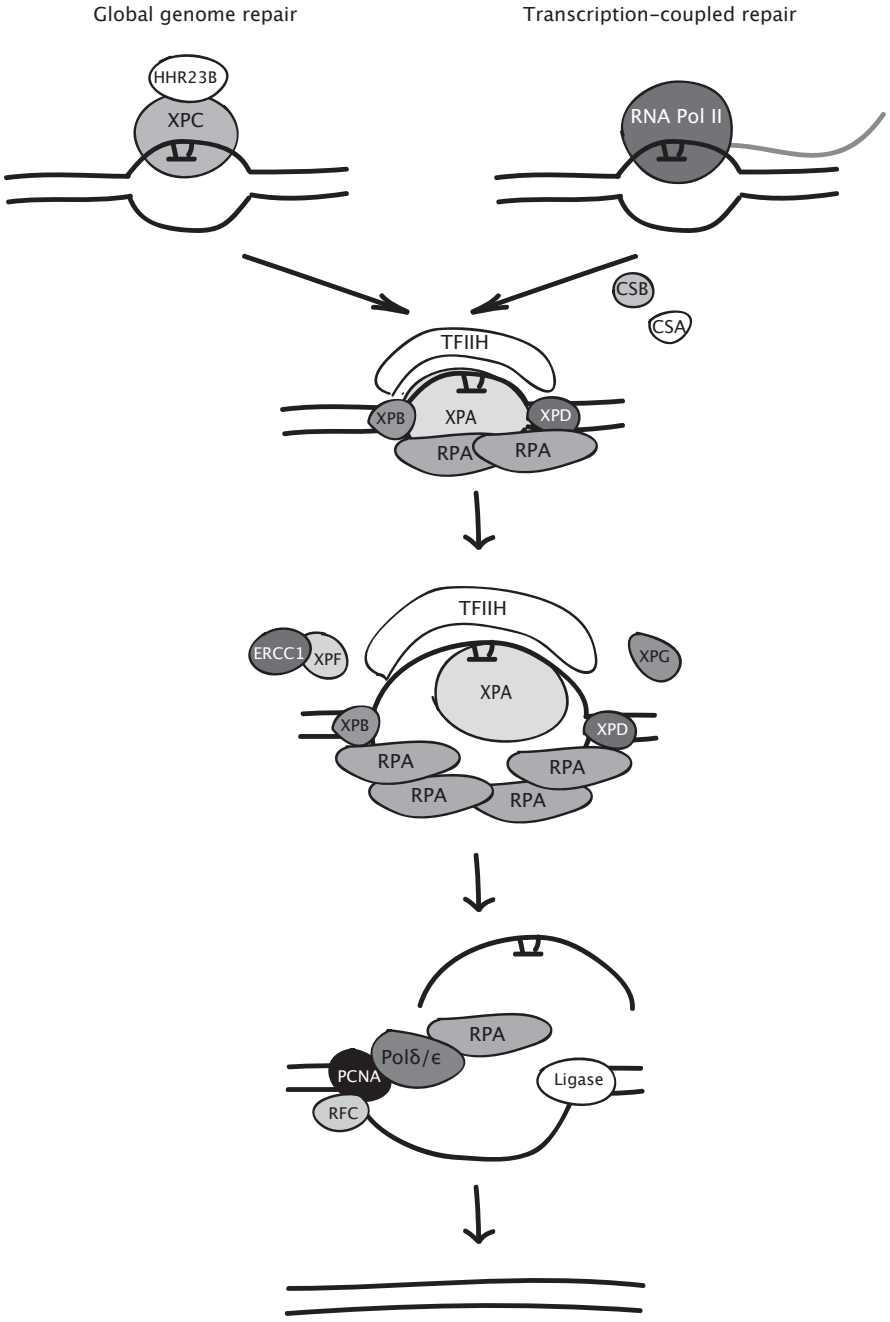


Figure 4: Mechanism of nucleotide excision repair. Damage recognition occurs in two ways, depending on the location of the damage: Adducts in transcribed regions are recognized by the stalled RNA polymerase while in the remainder of the genome they are recognized by HHR23B/XPC. The downstream steps are performed by the core mechanism.

quent death of *xpa-1* mutants was correlated with transcriptional inhibition that is associated with a decline in RNA pol II levels [74]. This study also identified an E3 ubiquitin ligase involved in the degradation of RNA pol II.

xpa-1 as well as *xpf-1* mutants were also shown to be sensitive to Cisplatin [78]. Cisplatin causes two kinds of lesions: intra- and interstrand crosslinks. Intrastrand crosslinks are repaired by NER. This partly explains the sensitivity seen in these mutants. As discussed later, XPF is also implicated in the repair of the second type of Cisplatin adducts, interstrand crosslinks [79].

Analysis of the mutation rate in *xpa-1* mutants showed that the average rate is about 28 fold higher that in wildtype [80]. Unlike MMR mutants that mostly could not be propagated beyond twenty generations, *xpa-1* mutants survived for 40 generations. In total, 24 mutations were identified in mutation accumulation lines; 17 base substitutions and 7 insdel mutations. The majority (4 out of 7) of insdel mutations in *xpa-1* animals occured at short mononucleotide runs (3-6 bp). In humans deficient for NER, 20% of insdel mutations occured at this type of runs [81].

Three studies revealed other cellular or organismal consequences of NER defects. The response of the adult germ line to UV-C radiation was studied [82], identifying a signaling pathway that induces cell cycle arrest and apoptosis after UV-C exposure. This signaling pathway contains proteins also involved in the response to IR. Importantly, the study showed that NER genes *xpa-1* and *xpc-1* are involved in the induction of cell cycle arrest and apoptosis, and are necessary for the recruitment and activation of the 9-1-1 complex.

Mutants that demonstrate a long-lived phenotype showed a higher repair rate of UV-induced pyrimidine dimers when compared to wildtype animals [83]. RNAi knockdown of *xpa-1* in these long-lived mutants decreased their resistance to UV irradiation and oxidative stress and reduced their lifespan. These data support a relation between the capacity to repair DNA damage and longevity in *C. elegans*.

Finally, a decline of nucleotide excision repair capacity in aging *C. elegans* was reported [77]. A PCR-based assay to monitor the repair of UV-C lesions demonstrated that there was a decrease in repair capacity in old adults in comparison to young adults.

Translesion Synthesis

Combined, the DNA repair systems are very accurate in the removal of DNA damage. However, remaining unrepaired damages present in the DNA when replication initiates are inevitable. Cells use DNA damage tolerance mechanisms that relief replication arrest, leaving the damage for future repair. First, the cell can use mechanisms whereby the replication machinery uses undamaged segments of the genome to copy DNA, thereby avoiding the need to replicate the damaged template-strand. Two known mechanisms are post-replication recombinational repair (or post-replicative gap filling) and replication-fork regression

(or copy-choice mechanism). In addition, the damage can be bypassed by specialized non-replicative polymerases in a process called translesion synthesis (TLS) [8].

For the studies performed in this thesis, the TLS mechanism is of interest. In TLS, bypass of the DNA lesion occurs through replication of the damaged region by specialized non-replicative polymerases that temporarily take over from the normal replication machinery (Figure 5). These polymerases can facilitate replication of damaged template lesions that cannot be performed by high-fidelity polymerases, mostly because they have a more flexible active site that allows incorporation of a base across from a damaged site. However, this property makes these polymerases less accurate. The bypass of damages by these TLS polymerases can be accurate (error free) or mutagenic (error prone) depending on the polymerase and/or the damage that needs bypass [84].

In eukaryotes, ten TLS polymerases have been identified [85]. These polymerases lack the exonuclease-activity that high-fidelity polymerases possess. They show very low fidelity when replicating undamaged DNA (some make one error in every twenty replicated bases) and show unusual error specificity. These properties are sometimes used by organisms, best illustrated in the generation of mutations in the variable regions of immunoglobin genes in which these polymerases are implicated [86].

In all models for TLS, polymerase switching is a crucial step. Since TLS polymerases have low fidelity rates, it is desirable that their activity remains restricted to damaged DNA, leaving undamaged regions to be replicated by the normal replication machinery [84]. First, high-fidelity polymerases of the replication machinery arrest when a lesion is encountered at the primer terminus and need to be replaced by specialized polymerases that can catalyze incorporation of nucleotides directly opposite the lesion. The lesion causes helix distortion that influences the base pairing in the immediate region and this precludes productive rearrangement of the replication machinery. Therefore, other specialized polymerases come in to extend the distorted primer terminus. Finally, the primer is extended to a position at which newly incorporated nucleotides are no longer susceptible to removal by exonucleolytic proofreading, and a final polymerase switch takes place (as depicted in Figure 5). The replication machinery reengages in high-fidelity replication.

A question that remains is how these polymerase switches are regulated. Exposure of replicating cells to DNA damage leads to the monoubiquitination of PCNA [87]. PCNA was identified as the polymerase sliding-clamp and interacts with many factors involved in DNA metabolism [88]. This suggests that stalling of the replication machinery leads to the monoubiquitination of PCNA, which directs the replication machinery into the TLS pathway. The monoubiquitination of PCNA led to an increased affinity of Polη for chromatin in UV-exposed cells [89, 90]. These results suggest that the first polymerase switch, from replicative to TLS polymerase, is arranged through the monoubiquitination of PCNA. The increased sensitivity of Polη for the modified PCNA would then result in a “switching out” of the replicative polymerases Polδ or Polε followed by “switching in” of Polη. The TLS polymerases Polι

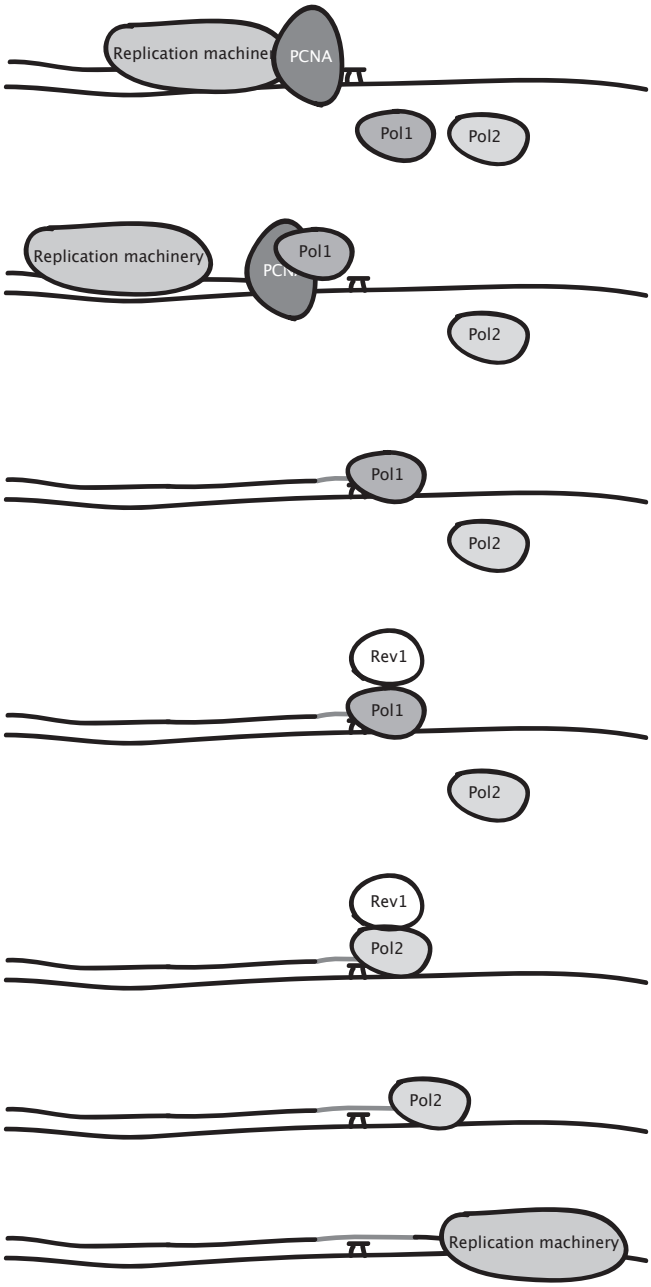


Figure 5: Translesion synthesis. Translesion synthesis functions in the bypass of DNA damages that block the repair machinery. Current models suggest that two TLS polymerases function in this bypass. First, a polymerase bypasses the damage and then a second polymerase elongates further past the damage until the helix distortion caused by the damage is not disturbing the normal replication polymerase (adapted from [8]).

and Polκ also contain PCNA binding motifs, however, it remains to be elucidated whether they physically interact with PCNA [84].

A second protein suggested to be involved in polymerase switching is Rev1. Rev1 is a member of the Y polymerase family and shows *in vitro* dCMP transferase activity in the bypass of abasic sites [91]. However, abolishment of the transferase activity does not abrogate its requirement for UV radiation-induced mutagenesis in yeast [91, 92]. These observations suggest a role for Rev1 in TLS independent of its dCMP activity. Rev1 interacts with a plethora of TLS polymerases, namely Polη, Polκ, Polι, Polλ, and the Rev7 subunit of Polζ [93-95]. Additionally, it was shown that binding of Rev1 to Polκ can be competed out by increasing amounts of Rev7 protein *in vitro* [93]. This indicates that Rev1 can switch between different (TLS) polymerases. Since there are multiple polymerase switches during the TLS process, it needs to be clarified in which switch Rev1 functions.

Translesion synthesis in C. elegans

Thus far, the study of translesion synthesis proteins in *C. elegans* is limited. The polymerases *polh-1*, *polk-1* and *rev-1* are coded by the genome, while a Polι homologue has not been found. ηRNAi knockdown of *polh-1* led to increased UV-C sensitivity when embryos were UV irradiated. Already laid embryos hardly showed sensitivity [96].

Secondly, *polh-1(RNAi)* and *polk-1(RNAi)* were shown to lead to MMS sensitivity in embryos [97]. Furthermore, RNAi knockdown of *polh-1* led to a delay in progression through P0 S phase dependent on MMS. This delay could be reversed by codepeletion of *chk-1*. *chk-1* is involved in the DNA damage checkpoint. Also, RAD-51 foci were detected in *polh-1(RNAi)* embryos treated with MMS.

As mentioned before, our laboratory has performed a screen to identify genes that are involved in the prevention of frameshift mutations. Surprisingly, a gene encoding a protein involved in TLS was identified [45]. RNAi knockdown of the *C. elegans* homologue of REV1, *rev-1*, was shown to increase frameshift mutations. In Chapter 2 of this thesis, we describe a functional characterization of the *rev-1* gene.

Interstrand Crosslink Repair

Interstrand crosslinks (ICLs) physically connect the two DNA strands. This leads to severe problems during replication, because the two strands cannot be separated to allow the replication machinery to copy DNA, and therefore, DNA damage tolerance mechanisms as described above cannot be used.

Repair of these lesions includes processing of the DNA substrate, translesion synthesis, and homologous recombination and is the result of the interplay of various proteins that also function in other repair pathways combined with several proteins that appear to be

restricted to ICL repair [6]. These are the Fanconi Anemia proteins: Cells derived from FA anemia patients are hypersensitive to ICL inducing agents, suggesting that the genes that are mutated in these patients are involved in the response to ICLs [98].

In total, there are 13 known FA complementation groups and for all of these the mutated genes have been identified [6, 99]. Based on protein domains, 4 out of 13 cloned FA genes provide clues about the function of the protein they encode. FANCD1 is identical to BRCA2, FANCJ and FANCM are helicases, and FANCL encodes an E3 ligase [100-103]. Most of the FA proteins (FANCA, B, C, E, F, G and L) interact to form a FA core complex [100]. FANCL can ubiquitinate FANCD2, a process that also needs FANCI. Monoubiquitination of FANCD2 triggers its assembly in nuclear foci and association with chromatin. FANCD2 then interacts with homologous recombination repair factors such as BRCA1, BRCA2/FANCD1 and RAD51 [104]. FANCJ and BRCA2/FANCD1 are dispensable for FANCD2 monoubiquitination and are therefore thought to act downstream of FANCD2 [6].

It has been suggested that the FA proteins are involved in scanning of the genome for ICLs. After recognition of the ICLs, and activation of FANCD2, the repair proteins would be involved in the actual removal of the damage. For example, the endonuclease MUS81-EME1 [105] and ERCC1-XPF [79] are involved in the conversion of the ICL to a double strand break. Still, many details of the repair of ICLs remain to be elucidated. Figure 6 depicts a current work model [6].

Interstrand crosslink repair in C. elegans

FA genes are conserved in *C. elegans*. Thus far, homologues for FANCD2 [106, 107], FANCD1/BRCA2 [55] and FANCJ [108] have been described. For the FANCD2 and FANCJ homologues, ICL repair related phenotypes have been described [106-108]. Also, homologues of ICL repair genes other than FA genes are conserved in the *C. elegans* genome, i.e. *xpf-1* and *mus-81 [109]*.

Animals mutated in the *C. elegans* FANCD2 homologue, *fcd-2*, showed increased sensitivity to interstand crosslinks and insensitivity to ionizing radiation [107]. FCD-2 is monoubiquitylated in response to DNA damage and is recruited to nuclear repair foci. In addition, this response was shown to be specific for replication stress and not the presence of a DSB per se, since FCD-2 foci only form in the presence of interstrand crosslinking damage and replication stress, but not after IR. All together, this study nicely showed that there is functional conservation of a central protein in the regulation of the response to ICLs by Fanconi Anemia proteins.

The gene *dog-1* is the homologue of FANCJ in *C. elegans*. Mutants showed sensitivity to ICL forming agents and this sensitivity was epistatic with *fcd-2*, but not with *brc-1* [108]. Recently, animals carrying mutations in *polq-1* and *hel-308* were shown to be sensitive to ICL treatment. Epistatic analysis revealed that they act in different pathways: *hel-308* acts in

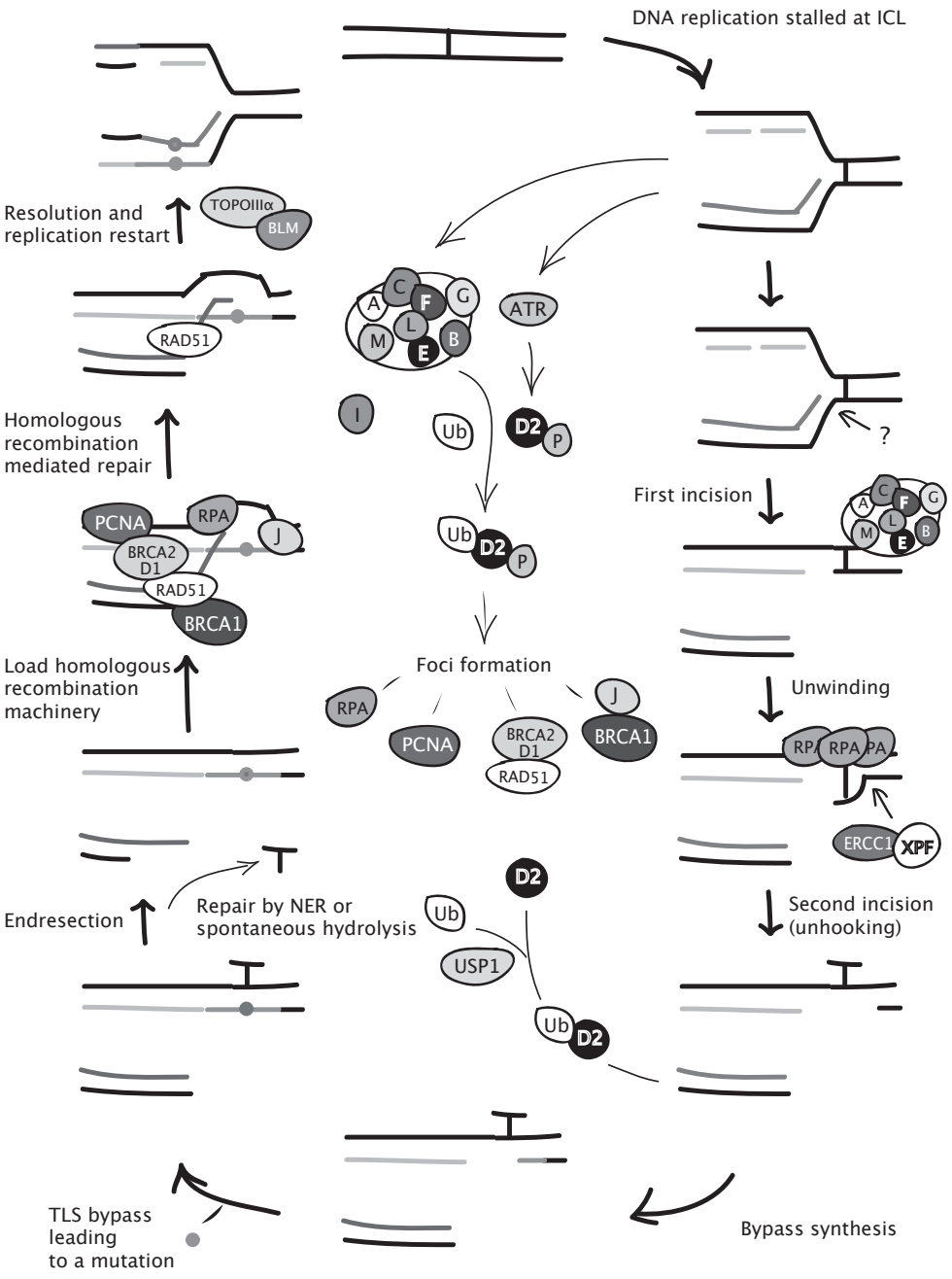


Figure 6: Proposed model of interstrand crosslink repair. An interplay of Fanconi Anemia, nucleotide excision repair, translesion synthesis and double-strand break repair factors are involved in the response to an interstrand crosslink lesion in the DNA that stalls the replication machinery (adapted from [6]).

a Fanconi Anemia-dependent pathway, and *polq-1* acts in a novel distinct *brc-1* dependent pathway [110].

Guanine tract instability

The *dog-1* gene, the C*. elegans* homologue of FANCJ*,* was previously described to be involved in the avoidance of a very specific kind of genome instability [111]. The name *dog-1* stands for deletions of guanine-rich DNA. Mutant animals showed a very specific mutation spectrum. All mutations were deletions that started around the 3' end of a polyguanine tract and terminated at variable positions at the 5' end of the tract. These deletions specifically occured at tracts that were at least 22 guanines long and about half of the guanine tracts that are present in the genome were fragile.

In vitro, guanine tracts have been shown to be capable of the formation of so-called quadruplex structures in which guanines in one strand pair to each other, leaving a very stable structure. Since the *dog-1* gene encodes a DEAH helicase, the authors presented a model where DOG-1 resolves the quadruplex structures. In absence of DOG-1 the quaduplexes are not resolved thereby causing a permanent replication block that can lead to deletion formation.

A candidate-gene approach was performed to implicate more genes in poly-G-tract instability [112] and this study showed that while some genes were capable of inducing a higher deletion rate in a *dog-1* background, none of them showed an increase in deletion rate on their own, when compared to wildtype. The genes that were identified are the BLM ortholog *him-6*, HR genes *rad-51*, *xpf-1* and *brd-1* and TLS polymerases *polh-1* and *polk-1*. NHEJ factors did not show an increase in deletion induction in both wildtype and *dog-1* background.

In addition, the RAD51 paralog *rfs-1* was implied in guanine tract maintenance. *rfs-1* is involved in HR specifically after the formation of replication blocks [78]. As in all other cases, RFS-1 on its own was not involved in maintaining guanine tract stability, while in a *dog-1* background an increase in deletion rate was detected.

In Chapter 4 of this thesis, we describe further study of the *dog-1* phenotype.

Scope of this thesis

The work described in this thesis originates from the observation that other genes than the already known DNA mismatch repair genes could prevent DNA frameshifts in *C. elegans.* A genome-wide RNAi screen performed in our laboratory [45] identified 13 genes that, when knocked down, elevated the rate of LacZ restoration in worms that had many copies of a LacZ reporter gene placed out of frame by the insertion of an A17 tract between the start codon and the ORF.

One of these genes is the *C. elegans* homolog of REV1, a member of the Y-family of translesion synthesis (TLS) proteins that is required to bypass DNA damage during replication. A role for TLS polymerases in counteracting micro-satellite instability has not been described before and we set out to investigate this phenomenon by isolating mutant alleles of *rev-1* from an EMS-induced mutant library, and subsequently analyzing these in several phenotypic assays. This work is described in **Chapter 2**.

Remarkably, we found that functional REV-1 is essential for embryonic development in *C. elegans*, in contrast to what is reported in other organisms. The essential nature of this gene makes epistatic analysis impossible and complicates other functional studies. Together with the notion that also another positive hit from the RNAi screen [45] resulted in lethality when knocked out genetically (as opposed to knocked down by RNAi) triggered us to improve our reporter assays to allow forward mutagenesis screens aimed to isolate hypomorphic alleles of potential MSI genes.

Chapter 3 describes the development of a more sensitive MSI reporter that was subsequently used in a forward genetic screen to identify new alleles of genes that are involved in MMR repair. Five of these had mutations in the already identified MMR gene, validating the approach. Thus far, other mutants have not been identified because the rate at which frameshifts occur in the cognate mutant strains is so high that mapping by conventional means is impossible (or too labor intensive).

In a parallel study, we modified these LacZ reporter transgenes to investigate the deletion forming potential of polyguanine sequences. Rose and colleagues cloned a gene that prevented the loss of poly G stretches from the *C. elegans* genome [111]. They named this gene *dog-1* for deletion of guanine-rich DNA. We show in **Chapter 4** that these sequences are mutagenic because they can adopt a G4-DNA structure. We also found that DOG-1, like its human ortholog FANCJ, is involved in protecting the organism against DNA crosslinking agents.

Next, we performed a genome-wide RNAi screen aimed to identify the set of genes that are involved in the cellular response to the chemotherapeutic drug Cisplatin, a known and clinically relevant DNA crosslinking agent. The outcome of this screen as well the description of Cisplatin-induced DNA damage responses is described in **Chapter 5**.

Chapter 6 describes the work performed to investigate end-joining repair of DSBs in the *C. elegans* germline. Some of the deletion products that stem from G4 DNA instability in *dog-1* mutated animals (Chapter 4) have characteristics that closely resemble the footprints that are observed upon repair of transposon-induced DSBs. This could mean that G4 DNAs are mutagenic because they form DSBs during DNA replication. By studying the genetic requirements and repair products of germ line DSBs that were induced upon transposon excision we propose that end-joining in this tissue is predominantly dictated by primer-template extension that is strongly driven by homology of the sequences that surrounds the DSB.

References

1. Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366-374.
2. Nordstrand, L.M., Ringvoll, J., Larsen, E., and Klungland, A. (2007). Genome instability and DNA damage accumulation in gene-targeted mice. Neuroscience *145*, 1309-1317.
3. Wilson, D.M., 3rd, and Bohr, V.A. (2007). The mechanics of base excision repair, and its relationship to aging and disease. DNA Repair (Amst) *6*, 544-559.
4. Costa, R.M., Chigancas, V., Galhardo Rda, S., Carvalho, H., and Menck, C.F. (2003). The eukaryotic nucleotide excision repair pathway. Biochimie *85*, 1083-1099.
5. Pastink, A., Eeken, J.C., and Lohman, P.H. (2001). Genomic integrity and the repair of double-strand DNA breaks. Mutat Res *480-481*, 37-50.
6. Niedernhofer, L.J., Lalai, A.S., and Hoeijmakers, J.H. (2005). Fanconi anemia (cross)linked to DNA repair. Cell *123*, 1191-1198.
7. Kolodner, R.D., and Marsischky, G.T. (1999). Eukaryotic DNA mismatch repair. Curr Opin Genet Dev *9*, 89-96.
8. Friedberg, E.C. (2005). Suffering in silence: the tolerance of DNA damage. Nat Rev Mol Cell Biol *6*, 943-953.
9. Rustgi, A.K. (2007). The genetics of hereditary colon cancer. Genes Dev *21*, 2525-2538.
10. Tischkowitz, M.D., and Hodgson, S.V. (2003). Fanconi anaemia. J Med Genet *40*, 1-10.
11. Lehmann, A.R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie *85*, 1101-1111.
12. Kannouche, P., and Stary, A. (2003). Xeroderma pigmentosum variant and error-prone DNA polymerases. Biochimie *85*, 1123-1132.
13. Nishiwaki, T., Kobayashi, N., Iwamoto, T., Yamamoto, A., Sugiura, S., Liu, Y.C., Sarasin, A., Okahashi, Y., Hirano, M., Ueno, S., and Mori, T. (2008). Comparative study of nucleotide excision repair defects between XPD-mutated fibroblasts derived from trichothiodystrophy and xeroderma pigmentosum patients. DNA Repair (Amst).
14. Demuth, I., and Digweed, M. (2007). The clinical manifestation of a defective response to DNA double-strand breaks as exemplified by Nijmegen breakage syndrome. Oncogene *26*, 7792-7798.
15. Thompson, L.H., and Schild, D. (2002). Recombinational DNA repair and human disease. Mutat Res *509*, 49-78.
16. Levy-Lahad, E., and Friedman, E. (2007). Cancer risks among BRCA1 and BRCA2 mutation carriers. Br J Cancer *96*, 11-15.
17. Schumacher, B., Garinis, G.A., and Hoeijmakers, J.H. (2008). Age to survive: DNA damage and aging. Trends Genet *24*, 77-85.
18. de Boer, J., Andressoo, J.O., de Wit, J., Huijmans, J., Beems, R.B., van Steeg, H., Weeda, G., van der Horst, G.T., van Leeuwen, W., Themmen, A.P., Meradji, M., and Hoeijmakers, J.H. (2002). Premature aging in mice deficient in DNA repair and transcription. Science *296*, 1276-1279.
19. Dolle, M.E., Busuttil, R.A., Garcia, A.M., Wijnhoven, S., van Drunen, E., Niedernhofer, L.J., van der Horst, G., Hoeijmakers, J.H., van Steeg, H., and Vijg, J. (2006). Increased genomic instability is not a prerequisite for shortened lifespan in DNA repair deficient mice. Mutat Res *596*, 22-35.
20. Hoeijmakers, J.H. (2007). Genome maintenance mechanisms are critical for preventing cancer as well as other aging-associated diseases. Mech Ageing Dev *128*, 460-462.
21. Harfe, B.D., and Jinks-Robertson, S. (2000). DNA mismatch repair and genetic instability. Annu Rev Genet *34*, 359-399.
22. Parker, B.O., and Marinus, M.G. (1992). Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. Proc Natl Acad Sci U S A *89*, 1730-1734.
23. Hall, M.C., Jordan, J.R., and Matson, S.W. (1998). Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD. Embo J *17*, 1535-1541.
24. Hall, M.C., and Matson, S.W. (1999). The Escherichia coli MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. J Biol Chem *274*, 1306-1312.
25. Wu, T.H., and Marinus, M.G. (1999). Deletion mutation analysis of the mutS gene in Escherichia coli. J Biol Chem *274*, 5948-5952.
26. Modrich, P., and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem *65*, 101-133.
27. Cooper, D.L., Lahue, R.S., and Modrich, P. (1993). Methyl-directed mismatch repair is bidirectional. J Biol Chem *268*, 11823-11829.
28. Au, K.G., Welsh, K., and Modrich, P. (1992). Initiation of methyl-directed mismatch repair. J Biol Chem *267*, 12142-12148.
29. Dao, V., and Modrich, P. (1998). Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. J Biol Chem *273*, 9202-9207.
30. Marsischky, G.T., Filosi, N., Kane, M.F., and Kolodner, R. (1996). Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev *10*, 407-420.

31. Harfe, B.D., and Jinks-Robertson, S. (2000). Mismatch repair proteins and mitotic genome stability. Mutat Res *451*, 151-167.
32. Sia, E.A., Kokoska, R.J., Dominska, M., Greenwell, P., and Petes, T.D. (1997). Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol *17*, 2851-2858.
33. Flores-Rozas, H., and Kolodner, R.D. (1998). The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. Proc Natl Acad Sci U S A *95*, 12404-12409.
34. Gu, L., Hong, Y., McCulloch, S., Watanabe, H., and Li, G.M. (1998). ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. Nucleic Acids Res *26*, 1173-1178.
35. Johnson, R.E., Kovvali, G.K., Guzder, S.N., Amin, N.S., Holm, C., Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1996). Evidence for involvement of yeast proliferating cell nuclear antigen in DNA mismatch repair. J Biol Chem *271*, 27987-27990.
36. Kirchner, J.M., Tran, H., and Resnick, M.A. (2000). A DNA polymerase epsilon mutant that specifically causes +1 frameshift mutations within homonucleotide runs in yeast. Genetics *155*, 1623-1632.
37. Lin, Y.L., Shivji, M.K., Chen, C., Kolodner, R., Wood, R.D., and Dutta, A. (1998). The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair. J Biol Chem *273*, 1453-1461.
38. Longley, M.J., Pierce, A.J., and Modrich, P. (1997). DNA polymerase delta is required for human mismatch repair in vitro. J Biol Chem *272*, 10917-10921.
39. Tran, H.T., Gordenin, D.A., and Resnick, M.A. (1999). The 3'-->5' exonucleases of DNA polymerases delta and epsilon and the 5'-->3' exonuclease Exo1 have major roles in postreplication mutation avoidance in Saccharomyces cerevisiae. Mol Cell Biol *19*, 2000-2007.
40. Umar, A., Buermeyer, A.B., Simon, J.A., Thomas, D.C., Clark, A.B., Liskay, R.M., and Kunkel, T.A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell *87*, 65-73.
41. Xie, Y., Counter, C., and Alani, E. (1999). Characterization of the repeat-tract instability and mutator phenotypes conferred by a Tn3 insertion in RFC1, the large subunit of the yeast clamp loader. Genetics *151*, 499-509.
42. Denver, D.R., Feinberg, S., Estes, S., Thomas, W.K., and Lynch, M. (2005). Mutation rates, spectra and hotspots in mismatch repair-deficient Caenorhabditis elegans. Genetics *170*, 107-113.
43. Tijsterman, M., Pothof, J., and Plasterk, R.H. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161*, 651-660.
44. Degtyareva, N.P., Greenwell, P., Hofmann, E.R., Hengartner, M.O., Zhang, L., Culotti, J.G., and Petes, T.D. (2002). Caenorhabditis elegans DNA mismatch repair gene msh-2 is required for microsatellite stability and maintenance of genome integrity. Proc Natl Acad Sci U S A *99*, 2158-2163.
45. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H., and Tijsterman, M. (2003). Identification of genes that protect the C. elegans genome against mutations by genome-wide RNAi. Genes Dev *17*, 443-448.
46. Paques, F., and Haber, J.E. (1999). Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. Microbiol Mol Biol Rev *63*, 349-404.
47. Gudmundsdottir, K., and Ashworth, A. (2006). The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. Oncogene *25*, 5864-5874.
48. Ivanov, E.L., and Haber, J.E. (1995). RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in Saccharomyces cerevisiae. Mol Cell Biol *15*, 2245-2251.
49. Saparbaev, M., Prakash, L., and Prakash, S. (1996). Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in Saccharomyces cerevisiae. Genetics *142*, 727-736.
50. Sugawara, N., Paques, F., Colaiacovo, M., and Haber, J.E. (1997). Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination. Proc Natl Acad Sci U S A *94*, 9214-9219.
51. Paques, F., and Haber, J.E. (1997). Two pathways for removal of nonhomologous DNA ends during double-strand break repair in Saccharomyces cerevisiae. Mol Cell Biol *17*, 6765-6771.
52. Sugawara, N., Ira, G., and 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.
53. Ivanov, E.L., Sugawara, N., Fishman-Lobell, J., and Haber, J.E. (1996). Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. Genetics *142*, 693-704.
54. Boulton, S.J., Martin, J.S., Polanowska, J., Hill, D.E., Gartner, A., and Vidal, M. (2004). BRCA1/BARD1 orthologs required for DNA repair in Caenorhabditis elegans. Curr Biol *14*, 33-39.
55. Martin, J.S., Winkelmann, N., Petalcorin, M.I., McIlwraith, M.J., and Boulton, S.J. (2005). RAD-51-dependent and -independent roles of a Caenorhabditis elegans BRCA2-related protein during DNA double-strand break repair. Mol Cell Biol *25*, 3127-3139.

56. Petalcorin, M.I., Sandall, J., Wigley, D.B., and Boulton, S.J. (2006). CeBRC-2 stimulates D-loop formation by RAD-51 and promotes DNA single-strand annealing. J Mol Biol *361*, 231-242.
57. Petalcorin, M.I., Galkin, V.E., Yu, X., Egelman, E.H., and Boulton, S.J. (2007). Stabilization of RAD-51-DNA filaments via an interaction domain in Caenorhabditis elegans BRCA2. Proc Natl Acad Sci U S A *104*, 8299-8304.
58. Clejan, I., Boerckel, J., and Ahmed, S. (2006). Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics *173*, 1301-1317.
59. van Haaften, G., Vastenhouw, N.L., Nollen, E.A., Plasterk, R.H., and Tijsterman, M. (2004). Gene interactions in the DNA damage-response pathway identified by genome-wide RNA-interference analysis of synthetic lethality. Proc Natl Acad Sci U S A *101*, 12992-12996.
60. van Haaften, G., Romeijn, R., Pothof, J., Koole, W., Mullenders, L.H., Pastink, A., Plasterk, R.H., and Tijsterman, M. (2006). Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. Curr Biol *16*, 1344-1350.
61. Bohr, V.A., Smith, C.A., Okumoto, D.S., and Hanawalt, P.C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. Cell *40*, 359-369.
62. Hanawalt, P.C. (1989). Preferential repair of damage in actively transcribed DNA sequences in vivo. Genome *31*, 605-611.
63. Svejstrup, J.Q. (2002). Mechanisms of transcription-coupled DNA repair. Nat Rev Mol Cell Biol *3*, 21-29.
64. van Gool, A.J., Citterio, E., Rademakers, S., van Os, R., Vermeulen, W., Constantinou, A., Egly, J.M., Bootsma, D., and Hoeijmakers, J.H. (1997). The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. Embo J *16*, 5955-5965.
65. Tu, Y., Bates, S., and Pfeifer, G.P. (1998). The transcription-repair coupling factor CSA is required for efficient repair only during the elongation stages of RNA polymerase II transcription. Mutat Res *400*, 143-151.
66. Batty, D., Rapic'-Otrin, V., Levine, A.S., and Wood, R.D. (2000). Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. J Mol Biol *300*, 275-290.
67. Weber, C.A., Salazar, E.P., Stewart, S.A., and Thompson, L.H. (1990). ERCC2: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast RAD3. Embo J *9*, 1437-1447.
68. Weeda, G., van Ham, R.C., Vermeulen, W., Bootsma, D., van der Eb, A.J., and Hoeijmakers, J.H. (1990). A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. Cell *62*, 777-791.
69. Yang, Z.G., Liu, Y., Mao, L.Y., Zhang, J.T., and Zou, Y. (2002). Dimerization of human XPA and formation of XPA2-RPA protein complex. Biochemistry *41*, 13012-13020.
70. Vasquez, K.M., Christensen, J., Li, L., Finch, R.A., and Glazer, P.M. (2002). Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. Proc Natl Acad Sci U S A *99*, 5848-5853.
71. Evans, E., Moggs, J.G., Hwang, J.R., Egly, J.M., and Wood, R.D. (1997). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. Embo J *16*, 6559-6573.
72. Hartman, P.S., and Herman, R.K. (1982). Radiation-sensitive mutants of Caenorhabditis elegans. Genetics *102*, 159-178.
73. Hartman, P.S., Hevelone, J., Dwarakanath, V., and Mitchell, D.L. (1989). Excision repair of UV radiation-induced DNA damage in Caenorhabditis elegans. Genetics *122*, 379-385.
74. Astin, J.W., O'Neil, N.J., and Kuwabara, P.E. (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 (Amst) *7*, 267-280.
75. Park, H.K., Suh, D., Hyun, M., Koo, H.S., and Ahn, B. (2004). A DNA repair gene of Caenorhabditis elegans: a homolog of human XPF. DNA Repair (Amst) *3*, 1375-1383.
76. Park, H.K., Yook, J.S., Koo, H.S., Choi, I.S., and Ahn, B. (2002). The Caenorhabditis elegans XPA homolog of human XPA. Mol Cells *14*, 50-55.
77. Meyer, J.N., Boyd, W.A., Azzam, G.A., Haugen, A.C., Freedman, J.H., and Van Houten, B. (2007). Decline of nucleotide excision repair capacity in aging Caenorhabditis elegans. Genome Biol *8*, R70.
78. Ward, J.D., Barber, L.J., Petalcorin, M.I., Yanowitz, J., and Boulton, S.J. (2007). Replication blocking lesions present a unique substrate for homologous recombination. Embo J *26*, 3384-3396.
79. Niedernhofer, L.J., Odijk, H., Budzowska, M., van Drunen, E., Maas, A., Theil, A.F., de Wit, J., Jaspers, N.G., Beverloo, H.B., Hoeijmakers, J.H., and Kanaar, R. (2004). The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. Mol Cell Biol *24*, 5776-5787.
80. Denver, D.R., Feinberg, S., Steding, C., Durbin, M., and Lynch, M. (2006). The relative roles of three DNA repair pathways in preventing Caenorhabditis elegans mutation accumulation. Genetics *174*, 57-65.
81. King, N.M., Oakley, G.G., Medvedovic, M., and Dixon, K. (2001). XPA protein alters the specificity of ultraviolet light-induced mutagenesis in vitro. Environ Mol Mutagen *37*, 329-339.

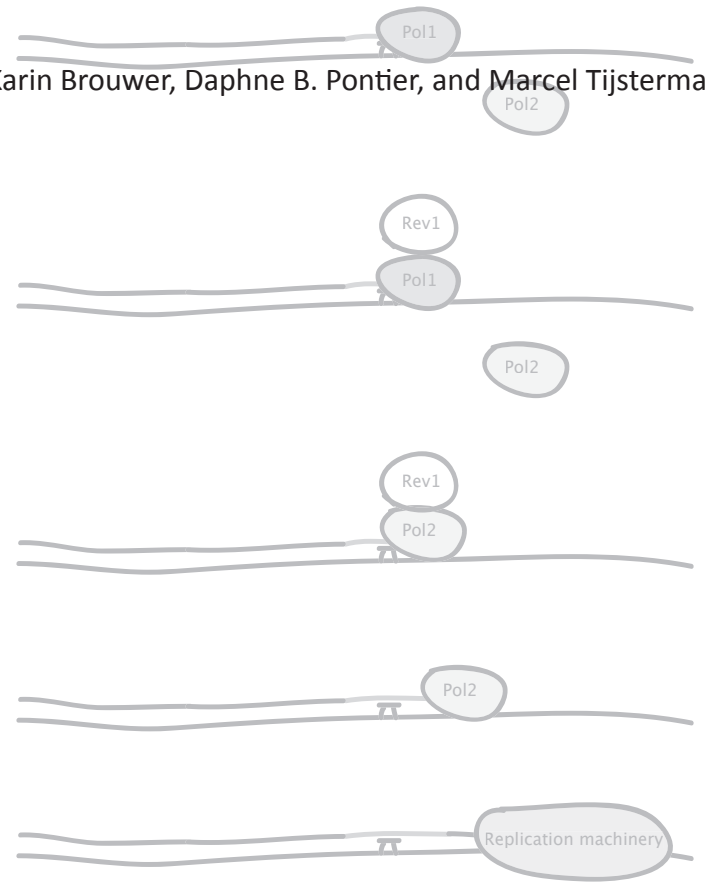
82. Stergiou, L., Doukoumetzidis, K., Sendoel, A., and Hengartner, M.O. (2007). The nucleotide excision repair pathway is required for UV-C-induced apoptosis in Caenorhabditis elegans. Cell Death Differ *14*, 1129-1138.
83. Hyun, M., Lee, J., Lee, K., May, A., Bohr, V.A., and Ahn, B. (2008). Longevity and resistance to stress correlate with DNA repair capacity in Caenorhabditis elegans. Nucleic Acids Res.
84. Friedberg, E.C., Lehmann, A.R., and Fuchs, R.P. (2005). Trading places: how do DNA polymerases switch during translesion DNA synthesis? Mol Cell *18*, 499-505.
85. Bebenek, K., and Kunkel, T.A. (2004). Functions of DNA polymerases. Adv Protein Chem *69*, 137-165.
86. Kunkel, T.A., Pavlov, Y.I., and Bebenek, K. (2003). Functions of human DNA polymerases eta, kappa and iota suggested by their properties, including fidelity with undamaged DNA templates. DNA Repair (Amst) *2*, 135-149.
87. Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141.
88. Maga, G., and Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. J Cell Sci *116*, 3051-3060.
89. Kannouche, P.L., Wing, J., and 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.
90. Watanabe, K., Tateishi, S., Kawasuji, M., Tsurimoto, T., Inoue, H., and Yamaizumi, M. (2004). Rad18 guides poleta to replication stalling sites through physical interaction and PCNA monoubiquitination. Embo J *23*, 3886-3896.
91. Lawrence, C.W. (2004). Cellular functions of DNA polymerase zeta and Rev1 protein. Adv Protein Chem *69*, 167-203.
92. Otsuka, C., Loakes, D., and Negishi, K. (2002). The role of deoxycytidyl transferase activity of yeast Rev1 protein in the bypass of abasic sites. Nucleic Acids Res Suppl, 87-88.
93. Guo, C., Fischhaber, P.L., Luk-Paszyc, M.J., Masuda, Y., Zhou, J., Kamiya, K., Kisker, C., and Friedberg, E.C. (2003). Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. Embo J *22*, 6621-6630.
94. Ohashi, E., Murakumo, Y., Kanjo, N., Akagi, J., Masutani, C., Hanaoka, F., and Ohmori, H. (2004). Interaction of hREV1 with three human Y-family DNA polymerases. Genes Cells *9*, 523-531.
95. Tissier, A., Kannouche, P., Reck, M.P., Lehmann, A.R., Fuchs, R.P., and Cordonnier, A. (2004). Co-localization in replication foci and interaction of human Y-family members, DNA polymerase pol eta and REVl protein. DNA Repair (Amst) *3*, 1503-1514.
96. Ohkumo, T., Masutani, C., Eki, T., and Hanaoka, F. (2006). Deficiency of the Caenorhabditis elegans DNA polymerase eta homologue increases sensitivity to UV radiation during germ-line development. Cell Struct Funct *31*, 29-37.
97. Holway, A.H., Kim, S.H., La Volpe, A., and Michael, W.M. (2006). Checkpoint silencing during the DNA damage response in Caenorhabditis elegans embryos. J Cell Biol *172*, 999-1008.
98. D'Andrea, A.D. (2003). The Fanconi road to cancer. Genes Dev *17*, 1933-1936.
99. Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D., and Elledge, S.J. (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289-301.
100. Meetei, A.R., Medhurst, A.L., Ling, C., Xue, Y., Singh, T.R., Bier, P., Steltenpool, J., Stone, S., Dokal, I., Mathew, C.G., Hoatlin, M., Joenje, H., de Winter, J.P., and Wang, W. (2005). A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. Nat Genet *37*, 958-963.
101. Cantor, S.B., Bell, D.W., Ganesan, S., Kass, E.M., Drapkin, R., Grossman, S., Wahrer, D.C., Sgroi, D.C., Lane, W.S., Haber, D.A., and Livingston, D.M. (2001). BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell *105*, 149-160.
102. Meetei, A.R., de Winter, J.P., Medhurst, A.L., Wallisch, M., Waisfisz, Q., van de Vrugt, H.J., Oostra, A.B., Yan, Z., Ling, C., Bishop, C.E., Hoatlin, M.E., Joenje, H., and Wang, W. (2003). A novel ubiquitin ligase is deficient in Fanconi anemia. Nat Genet *35*, 165-170.
103. Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G., Ikeda, H., Fox, E.A., and D'Andrea, A.D. (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. Science *297*, 606-609.
104. Taniguchi, T., Garcia-Higuera, I., Andreassen, P.R., Gregory, R.C., Grompe, M., and D'Andrea, A.D. (2002). S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. Blood *100*, 2414-2420.
105. Hanada, K., Budzowska, M., Modesti, M., Maas, A., Wyman, C., Essers, J., and Kanaar, R. (2006). The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. Embo J *25*, 4921-4932.
106. Dequen, F., St-Laurent, J.F., Gagnon, S.N., Carreau, M., and Desnoyers, S. (2005). The Caenorhabditis elegans FancD2 ortholog is required for survival following DNA damage. Comp Biochem Physiol B Biochem Mol Biol *141*, 453-460.
107. Collis, S.J., Barber, L.J., Ward, J.D., Martin, J.S., and Boulton, S.J. (2006). C. elegans FANCD2 responds to replication stress and functions in interstrand cross-link repair. DNA Repair (Amst) *5*, 1398-1406.

108. Youds, J.L., Barber, L.J., Ward, J.D., Collis, S.J., O'Neil, N.J., Boulton, S.J., and Rose, A.M. (2008). DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Mol Cell Biol *28*, 1470-1479.
109. O'Neil, N., and Rose, A. (2006). DNA repair. WormBook, 1-12.
110. Muzzini, D.M., Plevani, P., Boulton, S.J., Cassata, G., and Marini, F. (2008). Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. DNA Repair (Amst).
111. Cheung, I., Schertzer, M., Rose, A., and Lansdorp, P.M. (2002). Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. Nat Genet *31*, 405-409.
112. Youds, J.L., O'Neil, N.J., and Rose, A.M. (2006). Homologous recombination is required for genome stability in the absence of DOG-1 in Caenorhabditis elegans. Genetics *173*, 697-708.

Chapter 2

C. elegans translesion synthesis polymerase REV-1 counteracts microsatellite instability and is essential for embryonic development

Karin Brouwer, Daphne B. Pontier, and Marcel Tijsterman



Abstract

To safeguard the transmission of genetic information, numerous repair pathways have evolved to repair DNA, thus preventing the incorporation of incorrect bases during DNA replication. However, damages can escape repair –because they are unnoticed or introduced during S-phase- and are thus encountered by the replication machinery during DNA replication. The fidelity of replicative DNA polymerases prohibits fork progression and to prevent fork collapse, cells can opt to recruit specialized low-fidelity polymerases to bypass these damages, which frequently occurs in an error-prone manner. Here, we describe a role for *C. elegans* TLS polymerase REV-1 in preventing mutations at mononucleotide tracts under non-challenged conditions. Surprisingly, we found that functional REV-1 is essential for *C. elegans* development: null alleles arrest at late stages of embryogenesis accompanied by the appearance of RAD-51 foci, suggesting the persistence of DNA double strand breaks. We find that this essential role in embryogenesis does not require REV-1's BRCT domain: a mutation in a conserved residue in this domain does not affect worm development or growth but confers hypersensitivity to DNA damaging conditions such as UV irradiation and Cisplatin treatment.

Introduction

Cellular DNA is under continuous threat by a variety of internal and external DNA damaging agents. Specialized DNA repair pathways remove most of the damages that occur. However, persistent damages can pose a severe problem to replicative polymerases during DNA replication. In eukaryotes, a range of translesion synthesis (TLS) polymerases have been identified that can bypass damages by incorporating nucleotides across damaged bases or a-basic sites (reviewed in [1]). This bypass action is frequently error-prone, not only because DNA lesions are non- or mis-instructional, but also because these polymerases lack proofreading activity and are by themselves error-prone. To prevent mutation induction at sites of TLS, the replicative, and more reliable, DNA polymerases delta and epsilon take over soon after the damage has been bypassed.

One of the proteins involved in TLS is REV1, originally identified in the budding yeast *Saccharomyces cerevisiae*. The Rev1p protein has been shown to posses deoxycytidyl transferase activity that *in vitro* mostly incorporates dCMPs across a-basic sites [2] and non-damaged guanines [2, 3]. However, two observations indicate that this activity is not central to REV1's role in damaged-induced TLS: i) *in vivo* experiments in yeast show that Rev1p is required to bypass a-basic sites but that this predominantly leads to the incorporation of an A [3, 4]. ii) a mutation in REV1 that disrupt the transferase activity still supports translesion synthesis [5]. It has thus been proposed that REV1 may also play a non-catalytic role in controlling the action of other TLS polymerases during damage bypass (reviewed in [6]). In support for such a coordinative role, the C-terminal region of mammalian REV1 has been shown to interact

with other TLS polymerases [4, 7]. However, this region is not conserved in nematodes and yeast [8, 9]. Another motif that is thought to mediate protein-protein interaction in TLS is a BRCT motif that is conserved in all eukaryotic REV1 homologues. The BRCT domain was first identified in the breast cancer-related genes BRCA1, hence its name: BRCA1 carboxyl-terminal domain. BRCT domains are present in many other proteins involved in the cellular response to DNA damage [10], and have been shown to act as phosphopeptide-binding modules involved in protein targeting [11]. In mouse, REV1's BRCT domain interacts with PCNA, and is responsible for the nuclear localization of REV1 under non-challenged conditions [12].

Little is known about TLS polymerases in replication of non-damaged DNA. The high intrinsic error rate of TLS polymerases must mean that their activity is tightly regulated. Current models implicate a replication fork stall to trigger a molecular cascade that involves PCNA mono-ubiquitination, which alters the affinity of this replication cofactor in favor of TLS polymerases, but this could also happen to forks that stall at structural impediments or at sequences that are intrinsically difficult to replicate, such as repetitive DNA.

We previously picked up *C. elegans rev-1* in a screen for genes that protect the genome against mutations [13]. We made use of transgenic *C. elegans* strains that allowed us to monitor DNA frameshifts occurring at repetitive sequences as the result of replication slippage [14]. This genomic instability phenotype, known as micro-satellite instability (MSI), is linked to carcinogenesis, exemplified by the discovery of elevated levels of MSI in tumor cells of HNPCC patients. Surprisingly, a genome-wide RNAi screen for genes that elevated the rate of MSI-dependent LacZ restoration identified a clone that targets *C. elegans rev-1* [13], together with the four known *C. elegans* homologues of mismatch repair (MMR) genes. In the present study, we use RNAi and newly derived mutant alleles of *rev-1* to further investigate the role of the encoded protein in preventing genomic alterations under non-damaged conditions. By using a number of different frameshift reporters, we strengthen the notion that REV-1 acts to prevent frameshifts at microsatellites in the nematode's genome. Additionally, we show that *rev-1* has an essential function during *C. elegans* embryogenesis: animals homozygous for a knock out allele arrest during animal development. This essential function does not rely on a functional BRCT domain because mutation of a conserved amino acid in this domain resulted in damage-sensitivity phenotypes but not in compromised growth or loss of viability. Non-viable embryos resulting from *rev-1(RNAi)* display increased levels of RAD-51 foci, suggesting that DNA double strand breaks accumulate during animal embryogenesis in the absence of REV-1.

Materials and Methods

Strains and maintenance

General methods for culturing *C. elegans* were used [15]. Strains used in this study were N2, RB864 [*xpa-1(ok698)*], NL4893 [*dog-1(pk2247)*], AZ212 [*unc-119(ed3) ruIs32[unc-119(+) pie-1::GFP::H2B]*], NL4812 [*unc-119(ed3) pkIs2175 [pRP1890 unc-119(+)]*]. The *rev-1* null (*lf34)* and missense mutation (*lf35*) were isolated from an EMS mutant library, as previously described [16].

RNAi feeding

RNAi clones for *rev-1*, *msh-2*, and *msh-6* were taken from the RNAi feeding library described by Ahringer and coworkers [17]. After growing overnight in LB containing 50 μg ml^{-1} ampicilin medium at 37 °C, expression of dsRNA was induced by 200 μg ml^{-1} isopropylthiogalactoside at 37 °C for four hours. Subsequently, agar plates containing 50 μg ml^{-1} ampicilin and 200 μg ml^{-1} isopropylthiogalactoside were seeded with these cultures and left overnight at room temperature, protected from light.

Detection of microsatellite instability

To detect MSI, staged NL4812 L1 animals were grown on L4440 control vector, *rev-1(RNAi)*, *msh-2(RNAi)*, and *msh-6(RNAi)*. Young adult populations were stained with X-gal to detect the expression of functional β-galactosidase.

Identification of embryonic stage

To determine the stage of development in which the *rev-1(RNAi)* arrest, we used eight recognizable embryonic stages and L1 stage as described in worm atlas [18] to classify the embryos (Figure 3D). L4 animals were put on feeding plates and 24 hours later young adults were singled on feeding plates. These young adults were left to lay eggs for 16 hours, after which the mothers were removed. Three hours after removal, all embryos on the plate were checked and classified.

Immunostainings

Young adult worms were dissected in egg salts (1.18M NaCl; 480 mM KCl; 20 mM $CaCl_2$; 20 mM $MgCl_2$; 250 mM Hepes pH 7.4) to release embryos. Embryos were fixed for 5 minutes in 1.85 % paraformaldehyde and then snapfrozen onto slides in liquid nitrogen. Slides were incubated in 100% methanol at -20 °C for 15 minutes, washed in PBS-1% Triton and then

blocked in PBS with 0.1% Tween and 1% BSA for 30 minutes at room temperature. RAD-51 antibody (1:200) was added and incubated for 12 hours at 4 °C. Slides were washed three times in PBS with 0.1% Tween and were incubated in a Cy-3-labelled secondary goat-anti-rabbit antibody (1:1000) and DAPI for 3 hours at room temperature. Slides were washed three times in PBS with 0.1% Tween, mounted in Vectashield and observed using an SP2-confocal microscope (Leica).

UV sensitivity

Staged young adults were irradiated with UV in a UVP CL-1000 ultraviolet crosslinker with 0, 50, 100, 150 J/m^2. After irradiation, per dose four OP50 plates with three animals were prepared and grown for 42 hours at 20 °C. Next, the animals were removed and 24 hours after, their progeny was analyzed.

Cisplatin sensitivity

L4 animals were treated with 0, 0.3 and 0.6 mM Cisplatin (Cisplatine Mayne, 50 mg/50 ml, ONCO-TAIN®) for three hours in M9. After treatment, four plates with three L4 animals per dose were incubated for 42 hours. Next, parents were removed and death eggs and living offspring were scored 24 hours thereafter.

Results

C. elegans REV-1 acts to prevent MSI in somatic cells

We previously assayed MSI with animals that contained many copies of a LacZ-based reporter system that were located on an integrated transgenic array. This reporter system identified genes that prevented deletion and insertion mutations in the ~ 1kb ORF upstream of LacZ as well as frameshifts in an A17 tract placed immediately 3' of the ORF's starting codon. For many of the RNAi clones identified in a genome-wide screen we found that their LacZ restoration potential did not result from MSI, and for some that did, we observed greatly reduced ratios of mutagenic events when we assayed transgenic strains with few copies of the reporter transgene ([13], data not shown). For validation purposes, we set out to develop a more sensitive MSI reporter that could be assayed in low copy environments. We replaced the A17 tract for a more mutagenic C23 tract, and subsequently created transgenic lines via biolistic transformation. One of these lines, which displayed limited levels of LacZ restoration in a wild type genetic background and carried few copies of the reporter (~5), was used for further analysis. Animals, synchronized at the larval L1 stage, were transferred to control RNAi or *rev-1(RNAi)* plates and the population was assayed for LacZ restoration

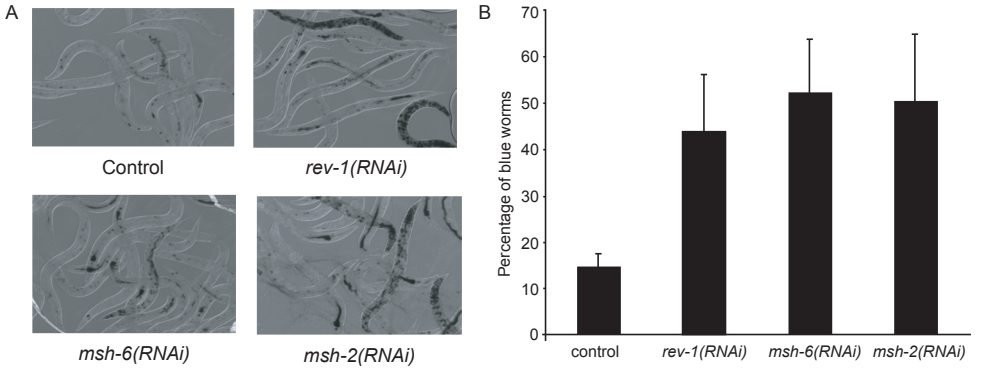


Figure 1: *rev-1(RNAi)* enhances the level of microsatellite instability. A) Transgenic reporter animals that were mock-treated control or subjected to *rev-1(RNAi)*, *msh-6(RNAi)* and *msh-2(RNAi)*. MSI is visualized by β-galactosidase expressing cells. Importantly, animals are subjected to RNAi at the larval L1 stage when they already contain ~500 of the 959 somatic cells that make up the adult animal. This means that the phenotypic consequences in the knockdowns are limited to late-stage cell divisions, after RNAi kicks in; a wild type level of MSI at the (C)23 tract takes place during embryogenesis. B) Quantification of the MSI instability. Animals were counted positive if they had at least one cell that express B-galactosidase. For each condition, we tested 10 independent cultures. Error bars indicate standard deviations.

when worms developed to young adults. We found that *rev-1(RNAi)* leads to 43.8 ± 12.3% of animals that show at least one cell that expresses B-galactosidase (Figure 1B), whereas 14.8 ± 3.0% of the animals stained positive on control RNAi clones. For comparison, we also knocked down the canonical MMR genes *msh-2* and *msh-6* by RNAi, which resulted in 50.4 ± 14.4% and 52.1 ± 11.6% positively stained animals, respectively. We also knocked down *rev-1* in MMR defective reporter animals by combining *rev-1(RNAi)* with a genetic deletion of *msh-6*. Here, however, the degree of frameshifting by genetically inactivating the MMR pathway was so high (worms accumulated many germline hits in the reporter) that it was impossible to determine a possible epistatic relationship between *rev-1* and MMR genes (data not shown).

C. elegans REV-1 shows homology to REV1 proteins from other organisms

REV-1 is well conserved throughout eukaryotic organisms, and we next compared the protein structure of *C. elegans* REV-1 to orthologous proteins in other species (Figure 2). *C. elegans rev-1* encodes for a 1027 amino acid protein, with two recognizable structural motifs: a BRCT domain and a Pol_zeta domain. Despite the high degree of similarity within these domains, which points at functional conservation, there are also noticeable differences: there is a clear lack of conservation of the ~ 100 amino acids located at the C-terminus, a region that in other systems has been shown to interact with Y-polymerases polη and polκ. Yeast two-hybrid analysis also failed to identify an interaction of *C. elegans* REV-1 with the

worm's homologs of polη (POLH-1) and polκ (POLK-1) [9]. This lack in conservation at the proteins C-terminus may also bear significance for a possible physical interaction with DNA polymerase ζ subunits: recently identified motifs in the C-terminus of yeast Rev1p have been shown to be sufficient for interaction with Rev7p [8]. We tested MSI on RNAi clones targeting *C. elegans* POLH-1, POLK-1 and the worm homolog of Rev3: all clones had wild type levels of frameshift induction.

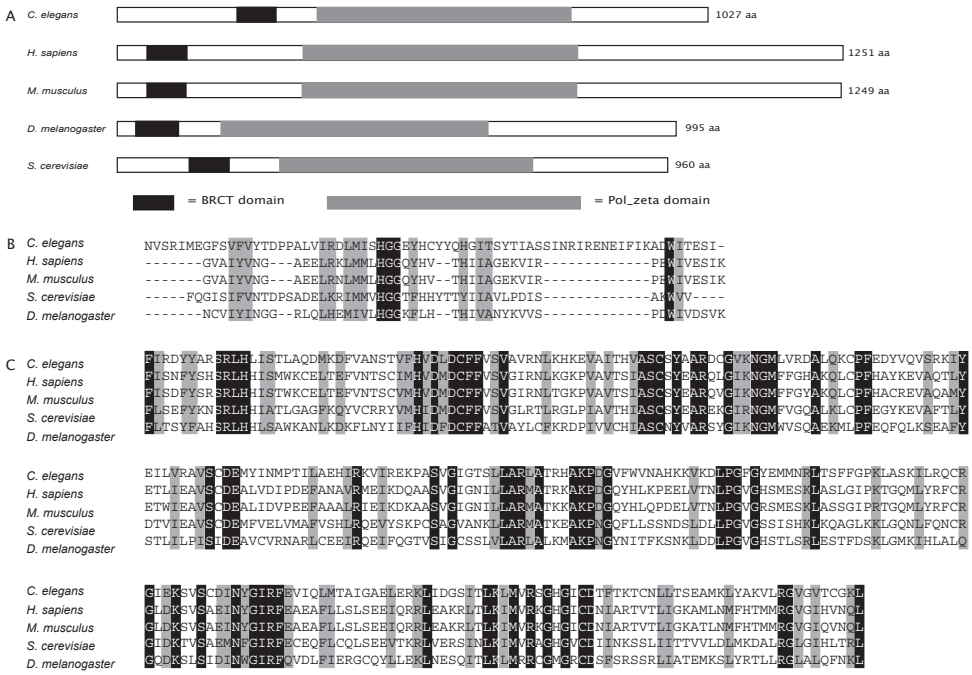


Figure 2: Evolutionarily conservation of REV-1 orthologs. A) Schematic illustration of the modular organization of REV1 protein from various organisms. The BRCT and Pol_zeta domains are shown in black and gray, respectively. Sequence conservation of B) the BRCT domain and C) the Pol_zeta domain. Conserved amino acids are boxed in black; gray boxes indicate similar residues.

C. elegans REV-1 is essential for embryonic development

To further study the role of *rev-1* in preventing MSI, we aimed to isolate genetic alleles by re-sequencing an EMS-mutagenized *C. elegans* library [16]. We identified animals that carry a premature stop codon that is located in the fourth exon of *rev-1*. The mutation is a A>T transversion, that leads to a K236* mutation (Figure 3A). Culturing animals carrying this allele (*rev-1(lf34)*) suggested accompanying embryonic lethality: the allele could only be maintained in a heterozygous state (also after 6 fold out-crossing or extensive sib-selection procedures), and dead embryos that were found on the plate were homozygous for lf34. To confirm that this embryonic lethal phenotype results from REV-1 loss and not from a

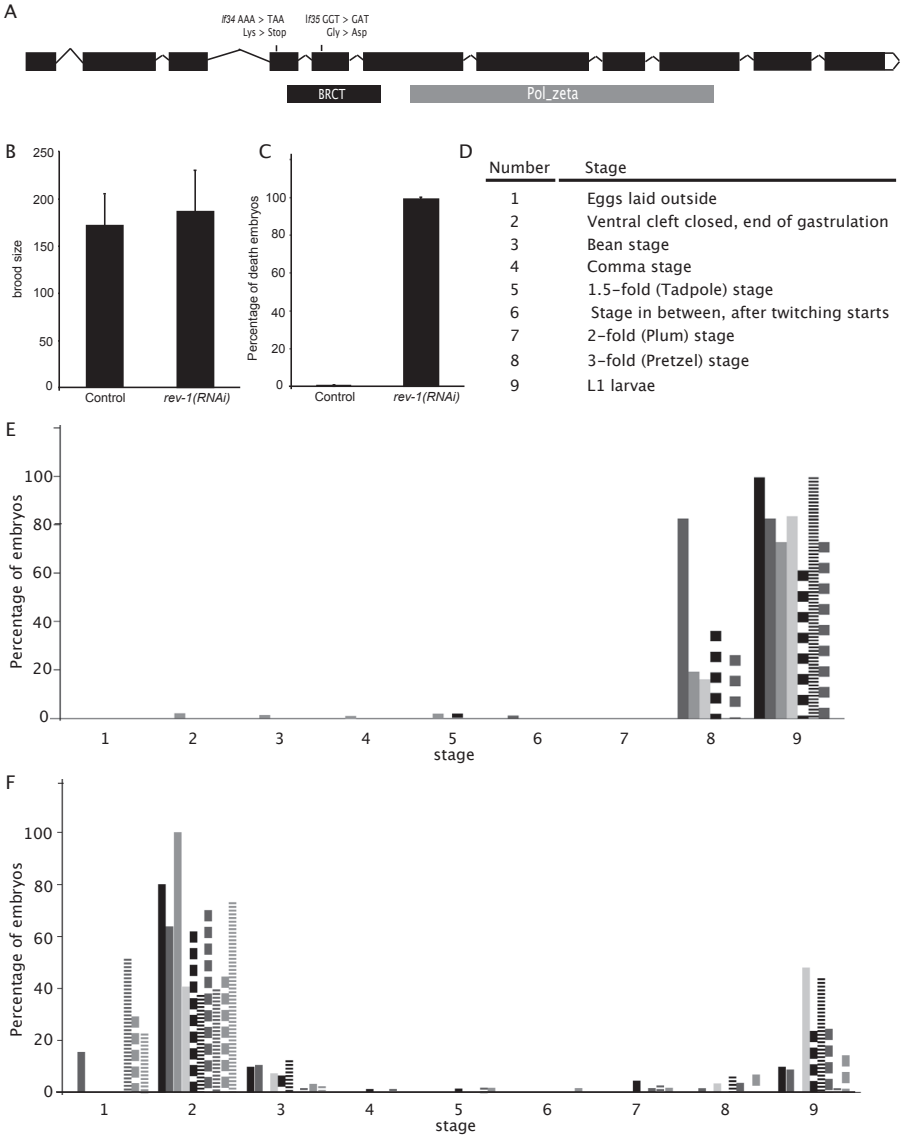


Figure 3: *C. elegans rev-1* is essential for embryonic development. A) Genomic organization of the *C. elegans rev-1* locus. The alleles described in this study are depicted above the gene structure; the regions that encode for the BRCT and Pol_zeta domains are depicted underneath. B) Brood size and C) Survival of animals (n=10) that were fed on control RNAi clones or *rev-1* dsRNA producing bacteria. Error bars indicate standard deviations. D) Distinct approximate stages of *C. elegans* embryogenesis can be distinguished by low resolution bright-field microscopy [18]. Eight of those categories, supplemented with hatched larvae as a separate class, were used to classify the developmental stages of embryos laid by E) mock treated or F) *rev-1*(RNAi) treated mothers. The progeny of seven (control) or nine (*rev-1(RNAi)*) animals were analyzed. Each type of bar represents one brood of one animal. The different developmental stages are plotted on the *x*-axis, while the *y*-axis shows the percentage of embryos in that specific stage.

nearby linked mutation, we placed wild type L4 animals on *rev-1(RNAi)* and empty feeding vector control plates and determined brood size, viability, and growth of their progeny. While the brood size (number of deposited eggs) of *rev-1(RNAi)* animals was unaffected by REV-1 knockdown, almost all embryos failed to hatch (99.1 ± 0.7% compared to 0.2 ± 0.3% for control animals; Figure 3B and C). Together these data show that REV-1 is essential for embryonic development in *C. elegans*.

We next classified the stages of development at which the embryos arrest/die. We placed wild type animals of the larval L4 stage on *rev-1(RNAi)* plates and allowed them to mature and produce eggs for 16 hours, after which the mothers were sacrificed. Three hours after removal we checked all embryos on the plate and classified them in 8 categories of recognizable embryonic sub-stages as described in worm atlas [18] (Figure 3D). On control RNAi plates, the vast majority of embryos were either in the "pretzel stage" or had already hatched (Figure 3E). *rev-1(RNAi)* embryos, in contrast, failed to reach the "bean stage" and started to degenerate at earlier developmental stages (Figure 3F). We also made time-lapse video images of the earliest zygotic cell divisions. It has been described that interfering with DNA replication (by RNAi) results in a delay of the cell cycle at the very first embryonic divisions [19]. However, we found that in contrast to RNAi of core DNA replication components, *rev-1(RNAi)*, under conditions that were sufficient to cause near 100% embryonic arrest, did not affect the cell cycle duration at the two-cell stage, when the larger anterior blastomere AB divides before the smaller posterior blastomere P(1).

Why do REV-1 depleted *C. elegans* embryos die? REV-1 is presumed to function in preventing replication fork collapse because it aids to bypass DNA damage and perhaps also facilitates replication of non-damaged DNA (e.g. micro-satellite sequences). Early embryogenesis in *C. elegans* goes through a phase of rapid cell divisions, without proper checkpoint control that would prevent cell cycle progression in the presence of damage [20]. It is thus conceivable that the early arrest that we observe in *rev-1(RNAi)* embryos is caused by excessive genomic deterioration if the cell cycle would progress in the presence of stalled or collapsed forks. The logical outcome of such a crisis would be the presence of chromosomal breaks. We thus stained embryos with an antibody against RAD-51, a protein that accumulates at resected ends of a double strand break (DSB). Indeed we found that even in the absence of applied genotoxic stress multiple RAD-51 foci formed in *rev-1(RNAi)* embryos (Figure 4). These were never observed on control RNAi. This observation is consistent with the hypothesis of increased replication stress in the absence of functional REV-1. At present we do not know whether embryos arrest because somatic cells acquire the capacity to activate a DNA damage checkpoint or for other reasons. It should be noted that embryonic development halts at similar stages when embryos are completely devoid of any DNA [21]. We used FISH to determine the ploidy in *rev-1(RNAi)* animals, but observed no difference from wild type (not shown), suggesting that *rev-1* embryos likely do not die because of massive chromosomal loss.

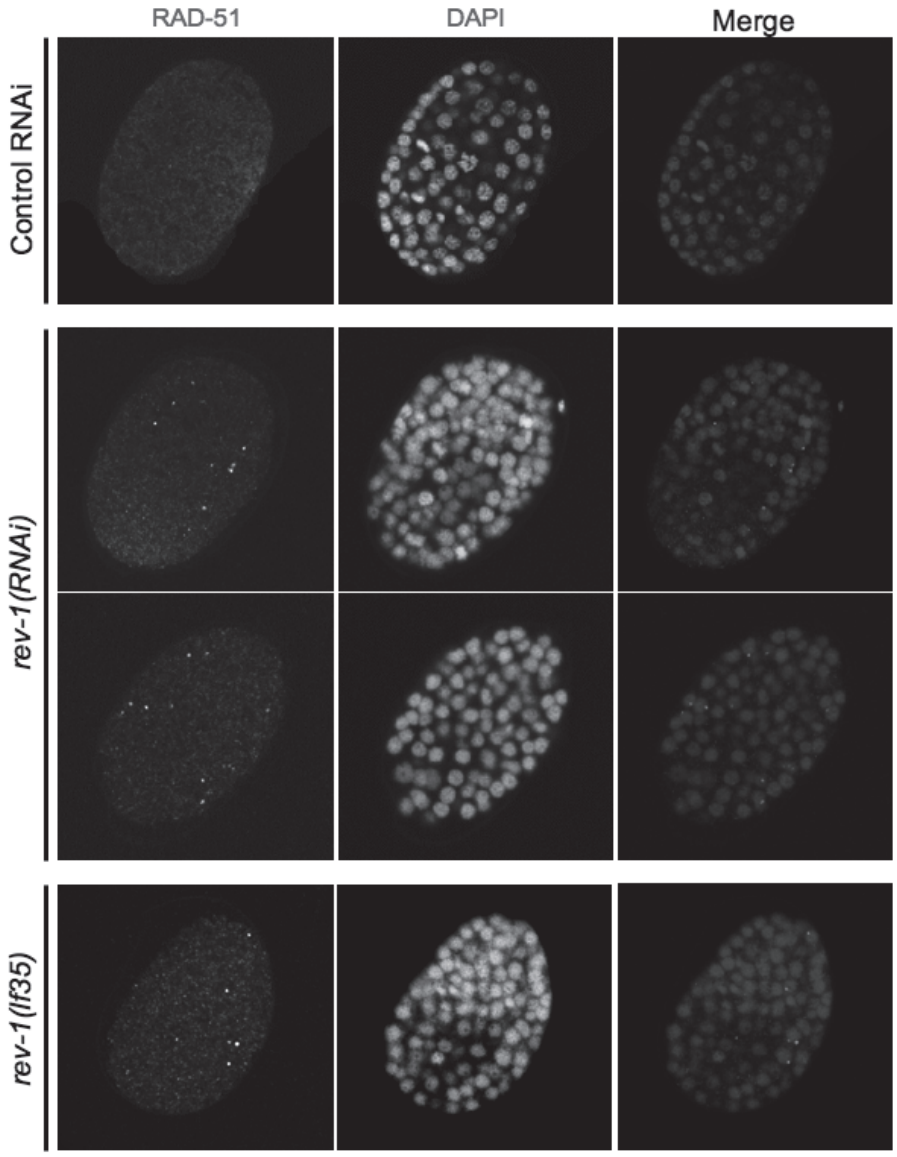


Figure 4: RAD-51 foci accumulate in *rev-1* deficient animals. Embryos of mock- or *rev-1*(RNAi)-treated mothers were analyzed by RAD-51 immuno- and DAPI-staining (left and middle panel respectively); the right panel displays the merged image. All displayed images show representative partial projections of Z-stacks.

The BRCT domain of REV-1 is required for TLS but not for viability or to counteract MSI

Apart from isolating a *rev-1* allele containing a premature stop, we also found a strain that carries a mutation located in one of REV-1's functional domains. The *rev-1(lf35)* allele carries a mutation changing the glycine at position 283 into an aspartic acid (G283D) (Figure 3A). G283 is located in the BRCT domain of the protein and is 100% conserved in all animal and yeast REV-1 proteins. In yeast, a mutant carrying a mutation of the corresponding glycine in the BRCT domain (Rev1p G193R) was isolated (*rev1-1*) [22]. This mutation renders the cells sensitive to UV irradiation. The mutant protein is defective in the bypass of T-T (6-4) UV photoproducts and abasic sites, but yet retains substantial deoxycytidyl transferase activity [23]. Mouse cells carrying a mutated REV1 BRCT domain display an UV-C-induced delay in progression through late S and G2 phases. Additionally, UV-C-mutagenesis is reduced and mutations at thymidine-thymidime dimers are absent [24].

Importantly, animals homozygous for *rev-1*(*lf35)* are viable and grow well without any effect on viability or embryonic development (data not shown), indicating that the essential role for *C. elegans* REV-1 is not associated with it's BRCT domain. We next assayed *rev-1(lf35)* animals for sensitivity to DNA damaging agents. Figure 5A illustrates that *rev-1(lf35)* animals are hypersensitive to UV irradiation, which is in excellent agreement with the yeast and mouse data [12, 22]. The UV sensitivity of *rev-1(lf35)* is further increased by inactivating nucleotide excision repair, which would remove the majority of UV-induced photoproducts. *rev-1(lf35) xpa-1(ok698)* double mutant animals are significantly more sensitive to UV irradiation than either single mutants. *rev-1(lf35)* animals are also hypersensitive to the DNA crosslinker Cisplatin. Cisplatin induces intra- and interstrand crosslinks [25, 26]. Intrastrand crosslinks are repaired by NER [27], while interstrand crosslinks are repaired by interstrand crosslink repair [28]. We show that *rev-1(lf35)* mutant animals are more sensitive to Cispla-

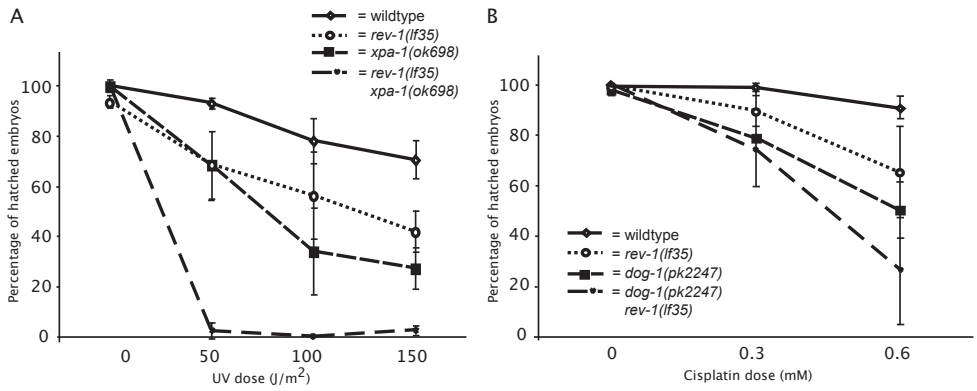


Figure 5: Sensitivity of *rev-1(lf35)* animals to DNA damaging agents. Animal survival curves for the indicated genotypes when exposed to UV irradiation (A) or to the crosslinking agent Cisplatin (B). Animals of the larval L4 stage were treated with the indicated dose, and the ratio live progeny/total laid eggs was determined for a 24 hours time window 24 hours post treatment. Error bars indicate standard deviations.

tin than wildtype animals, and additionally, that this mutation increases the sensitivity of *dog-1* mutant animals, the homologue of the Fanconi Anemia protein FANCJ [29], to this drug (Figure 5B). Together, this shows that the BRCT domain of *C. elegans* REV-1 is important for its role in DNA damage bypass. However, it is not essential for the essential role REV-1 plays during embryonic development.

Discussion

Our data broadens the significance of REV-1 in maintaining genome stability in animals. We found that functional REV-1 is essential for embryonic development in nematodes. For this essential function, the protein does not depend on an intact BRCT domain, while this domain is required for REV-1's protective role against DNA damage. A prominent role for REV1 in (allowing) animal development has been hinted at by the finding that in mice REV1 null alleles could not be created in a C57BL/6 background, while viability was virtually not compromised by REV1 inactivation in an 129/OLA genetic background [30]. It remains an open question which likely redundant pathway acts on similar endogenous substrates that have to be dealt with to allow proper embryogenesis in mice. High-throughput genetic analysis (e.g. epistatic miniarray profiling) in yeast may provide insight into the genetic interaction map of REV1. Two independent recent studies [31, 32] identified members of the RAD52 epistasis group to cause synthetic growth defect in rev1 deleted *S. cerevisiae* strains in the absence of exogenous DNA damage. Some of the members of this pathway, including RAD52 itself, are not encoded by the *C. elegans* genome, which may explain the more dramatic phenotype of REV1 inactivation in worms.

What substrate is so critically dependent or REV1 action? In mice, 129/OLA derived REV1 knockouts display pleiotropic phenotypes, even in the absence of exogenous challenges [30], suggesting the presence of an endogenous lesion that requires REV1 to be (efficiently) bypassed. We found that mutating the BRCT domain of *C. elegans* REV-1 confers UV and Cisplatin sensitivity but not loss of viability. If we interpret these phenotypes to result from the loss of TLS activity on specific forms of DNA damage, it means that those forms (or related types of DNA damage) are not the underlying cause of the genomic instability phenotype we observe in *rev-1*(RNAi) embryos, ultimately leading to embryonic arrest. Apart from damaged DNA, non-damaged but alternatively folded DNA may also constitute replication-blocking obstacles that require TLS polymerase action. Recently, yeast REV1 has been implicated in maintaining the stability of trinucleotide repeats, but only of those that have hairpin forming potential, in a manner dependent on BRCT domain function [33]. We identified a more general role for REV1 in protection of mononucleotide repeats against frameshifting errors: we found MSI at monoA and monoG tracts in *rev-1* compromised somatic cells. This leads to the question whether fork stalling at these sequences poses such a severe problem that it would ultimately lead to a permanent arrest and embryonic death.

Can the essential role of REV-1 be linked to catalytic activity of its polymerase domain? Rev1 is a translesion synthesis polymerase that *in vitro* incorporates uniquely deoxycytidine opposite damaged DNA and abasic residues.

This question is technically very difficult to address in the worm system. The essential nature of *rev-1* also precludes the generation of mutation accumulation lines, which would allow us to determine the mutation profile in the absence of REV-1, as has been done in mice [30]. However, the following rational reveals an intriguing conundrum: if REV-1 is so important that in its absence *C. elegans* embryos fail to develop, why is the spontaneous mutation frequency in this animal so low? The mutation rate was determined to be one nucleotide change per animal per generation [34]. This notion suggests that the essential function of REV-1 is in an error free pathway. Therefore, it is difficult to attribute a role of the enzyme's deoxycytidyl transferase activity, which is predominantly mutagenic, in maintaining the integrity of the replication fork at the expense of base substitutions: the margin of error is simply too small for a prominent role of a mutagenic TLS polymerase. A very recent finding may shed some light on such a function: Sale and Coworkers revealed a role for chicken REV1 in ensuring frame fidelity during TLS of UV photoproducts [35]. These authors show that in the absence of REV1, three instead of two bases are incorporated across a two "6-4" covalently linked thymidines. Bypass of this UV photoproduct was found to completely depend on Polζ's TLS action leading to the hypothesis that REV1 restrains the activity of this polymerase to ensure that nucleotides are incorporated in-frame with the template strand. Our data fits a model in which such a frame-tracking role is extended to other DNA polymerases at replication fork stalls.

Note added in press

After submission of this thesis manuscript, we discovered out that our stock rev-1 RNAi clone contained bacteria that produce dsRNA against the C. elegans rba-1 gene. Considering that the RNAi phenotypes described for rba-1 phenotypically mimics the phenotypes we observed through genetic inactivation of rev-1, we cannot exclude the possibility that the RNAi phenotypes we described in this chapter are the result of rba-1 instead of rev-1 knockdown. Future work will be aimed to resolve this issue.

References

1. Prakash, S., Johnson, R.E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem *74*, 317-353.
2. Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996). Deoxycytidyl transferase activity of yeast REV1 protein. Nature *382*, 729-731.
3. Haracska, L., Prakash, S., and Prakash, L. (2002). Yeast Rev1 protein is a G template-specific DNA polymerase. J Biol Chem *277*, 15546-15551.
4. Pages, V., Johnson, R.E., Prakash, L., and Prakash, S. (2008). Mutational specificity and genetic control of replicative bypass of an abasic site in yeast. Proc Natl Acad Sci U S A *105*, 1170-1175.

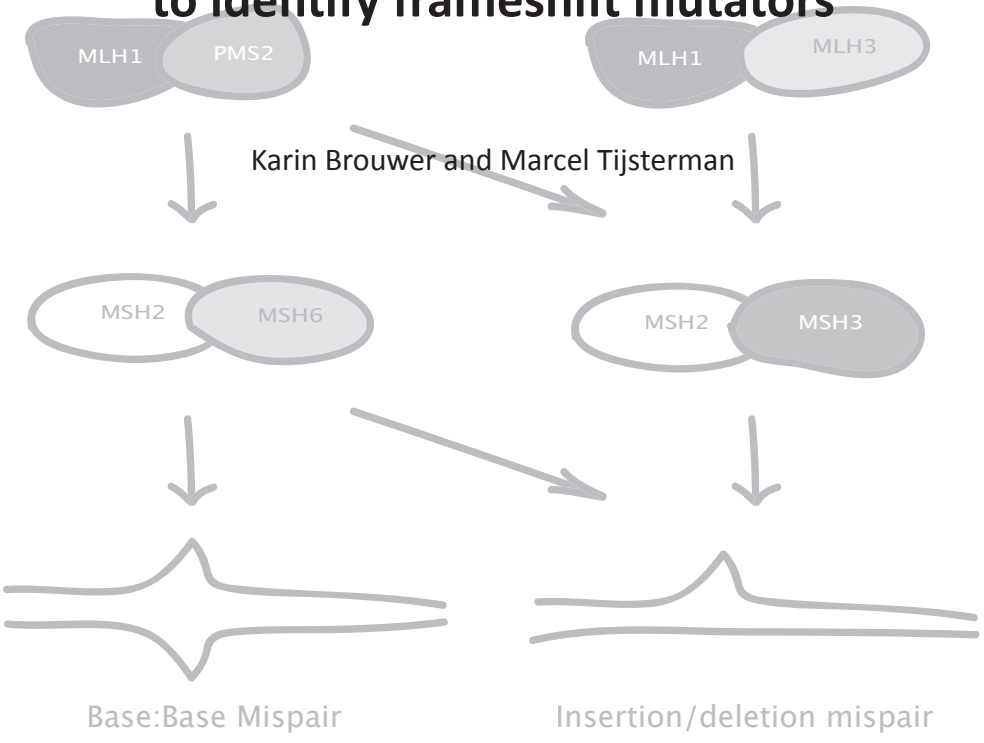
5. Otsuka, C., Loakes, D., and Negishi, K. (2002). The role of deoxycytidyl transferase activity of yeast Rev1 protein in the bypass of abasic sites. Nucleic Acids Res Suppl, 87-88.
6. Friedberg, E.C., Lehmann, A.R., and Fuchs, R.P. (2005). Trading places: how do DNA polymerases switch during translesion DNA synthesis? Mol Cell *18*, 499-505.
7. Haracska, L., Unk, I., Johnson, R.E., Johansson, E., Burgers, P.M., Prakash, S., and 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.
8. D'Souza, S., Waters, L.S., and Walker, G.C. (2008). Novel conserved motifs in Rev1 C-terminus are required for mutagenic DNA damage tolerance. DNA Repair (Amst).
9. Kosarek, J.N., Woodruff, R.V., Rivera-Begeman, A., Guo, C., D'Souza, S., Koonin, E.V., Walker, G.C., and 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 (Amst) *7*, 439-451.
10. Caldecott, K.W. (2003). Cell signaling. The BRCT domain: signaling with friends? Science *302*, 579-580.
11. Manke, I.A., Lowery, D.M., Nguyen, A., and Yaffe, M.B. (2003). BRCT repeats as phosphopeptide-binding modules involved in protein targeting. Science *302*, 636-639.
12. Guo, C., Sonoda, E., Tang, T.S., Parker, J.L., Bielen, A.B., Takeda, S., Ulrich, H.D., and 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.
13. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H., and Tijsterman, M. (2003). Identification of genes that protect the C. elegans genome against mutations by genome-wide RNAi. Genes Dev *17*, 443-448.
14. Tijsterman, M., Pothof, J., and Plasterk, R.H. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161*, 651-660.
15. Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.
16. Cuppen, E., Gort, E., Hazendonk, E., Mudde, J., van de Belt, J., Nijman, I.J., Guryev, V., and Plasterk, R.H. (2007). Efficient target-selected mutagenesis in Caenorhabditis elegans: toward a knockout for every gene. Genome Res *17*, 649-658.
17. Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G., and Ahringer, J. (2001). Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. Genome Biol *2*, RESEARCH0002.
18. Hall, D.H., and Altun, Z.F. (2008). C. elegans atlas, 1 Edition (Cold Spring Harbor Press).
19. Brauchle, M., Baumer, K., and Gonczy, P. (2003). Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in C. elegans embryos. Curr Biol *13*, 819-827.
20. Clejan, I., Boerckel, J., and Ahmed, S. (2006). Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics *173*, 1301-1317.
21. Zhong, W., Feng, H., Santiago, F.E., and Kipreos, E.T. (2003). CUL-4 ubiquitin ligase maintains genome stability by restraining DNA-replication licensing. Nature *423*, 885-889.
22. Larimer, F.W., Perry, J.R., and Hardigree, A.A. (1989). The Rev1 gene of Saccharomyces cerevisiae: Isolation, Sequence, and Functional Analysis. Journal of Bacteriology *171*, 230-237.
23. Nelson, J.R., Gibbs, P.E., Nowicka, A.M., Hinkle, D.C., and Lawrence, C.W. (2000). Evidence for a second function for Saccharomyces cerevisiae Rev1p. Mol Microbiol *37*, 549-554.
24. Jansen, J.G., Tsaalbi-Shtylik, A., Langerak, P., Calleja, F., Meijers, C.M., Jacobs, H., and de Wind, N. (2005). The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Res *33*, 356-365.
25. Fichtinger-Schepman, A.M., van der Veer, J.L., den Hartog, J.H., Lohman, P.H., and Reedijk, J. (1985). Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. Biochemistry *24*, 707-713.
26. Vrana, O., Boudny, V., and Brabec, V. (1996). Superhelical torsion controls DNA interstrand cross-linking by antitumor cis-diamminedichloroplatinum(II). Nucleic Acids Res *24*, 3918-3925.
27. De Silva, I.U., McHugh, P.J., Clingen, P.H., and Hartley, J.A. (2002). Defects in interstrand cross-link uncoupling do not account for the extreme sensitivity of ERCC1 and XPF cells to Cisplatin. Nucleic Acids Res *30*, 3848-3856.
28. Niedernhofer, L.J., Lalai, A.S., and Hoeijmakers, J.H. (2005). Fanconi anemia (cross)linked to DNA repair. Cell *123*, 1191-1198.
29. Youds, J.L., Barber, L.J., Ward, J.D., Collis, S.J., O'Neil, N.J., Boulton, S.J., and Rose, A.M. (2008). DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Mol Cell Biol *28*, 1470-1479.
30. Jansen, J.G., Langerak, P., Tsaalbi-Shtylik, A., van den Berk, P., Jacobs, H., and 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.
31. Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., Ding, H., Xu, H., Han, J., Ingvarsdottir, K., Cheng, B., Andrews, B., Boone, C., Berger, S.L., Hieter, P., Zhang, Z., Brown, G.W.,

Ingles, C.J., Emili, A., Allis, C.D., Toczyski, D.P., Weissman, J.S., Greenblatt, J.F., and Krogan, N.J. (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

32. Pan, X., Ye, P., Yuan, D.S., Wang, X., Bader, J.S., and Boeke, J.D. (2006). A DNA integrity network in the yeast Saccharomyces cerevisiae. Cell *124*, 1069-1081.
33. Collins, N.S., Bhattacharyya, S., and Lahue, R.S. (2007). Rev1 enhances CAG.CTG repeat stability in Saccharomyces cerevisiae. DNA Repair (Amst) *6*, 38-44.
34. Denver, D.R., Morris, K., Lynch, M., and Thomas, W.K. (2004). High mutation rate and predominance of insertions in the Caenorhabditis elegans nuclear genome. Nature *430*, 679-682.
35. Szuts, D., Marcus, A.P., Himoto, M., Iwai, S., and Sale, J.E. (2008). REV1 restrains DNA polymerase {zeta} to ensure frame fidelity during translesion synthesis of UV photoproducts in vivo. Nucleic Acids Res.

Chapter 3

A forward mutagenesis screen in *C. elegans* to identify frameshift mutators



Abstract

DNA replication has to occur with high fidelity to avoid accumulation of mutations. Nevertheless, during replication, errors such as base:base mispairs leading to point mutations and frameshifts resulting in insertions or deletions occur. Supplementing the proofreading capacity of replicating polymerases, DNA mismatch repair (MMR) proteins scan newly replicated DNA to detect and correct replication errors. To identify new factors that are involved in MMR, we developed a sensitive transgenic reporter system in *C. elegans* that colorimetrically detects DNA frameshifts. With this LacZ-based reporter, we performed a forward mutagenesis screen and isolated 21 mutants. Of these mutants, five contained new alleles of known MMR genes, validating the screen. Strategies to identify the other genetic defects that result in elevated levels of frameshift mutations are discussed.

Introduction

To faithfully transmit genetic material to new generations, cells use high-fidelity DNA polymerases to avoid incorrect copying of DNA during DNA replication. The mismatch repair (MMR) proteins provide an extra level of protection by correcting the errors that escape the proofreading activity of these polymerases [1, 2]. The importance of having a functional MMR system is illustrated by the existence of an MMR-related hereditary cancer syndrome. Patients that suffer from Hereditary Non-Polyposis Colorectal Cancer (HNPCC) or Lynch Syndrome carry mutations in MMR genes [3]. One of the hallmarks of HNPCC tumors is a decreased stability of microsatellites, which are repetitive sequences that exist throughout the genome.

MMR predominantly corrects two types of replication errors: 1) single basepair mismatches, which can lead to point mutations and 2) DNA loops, which result from replication slippage on repetitive sequences/microsatellites and can cause insertions or deletions [1, 2]. While the recognition steps of replication errors by MMR are fairly well understood, the downstream processing steps are less clear. In most eukaryotes, the two aforementioned substrates are initially processed by different, partially overlapping sub-pathways: base:base mispairs are mostly detected by MSH2/MSH6 heterodimers and further processed by MLH1/PMS2, while DNA loops are recognized by MSH2/MSH3 and further processed by MLH1/MLH3 [2].

In *C. elegans* and also in *Drosophila*, only one of the two pathways appears to be conserved [4]. Four *C. elegans* MMR genes have been identified; *msh-2, msh-6*, *mlh-1* and *pms-2*, but no homologue of *msh-3* or *mlh-3* could be identified in the fully sequenced genomes of these species. It is thus unclear whether the *C. elegans* MMR has a broader substrate range and can also repair DNA loops or whether such replication errors are left untouched. The observation that *C. elegans* genomes sustain higher levels of insertion/deletion muta-

genesis as compared to other species [5] could support the latter scenario. Similar to other organisms, defective MMR leads to highly elevated mutation rates [6, 7].

The model organism *C. elegans* is an excellent system for forward genetics. Because of its small size and the short time it takes to grow large numbers, screens can be performed rapidly without demanding too much space. In addition, reverse genetics approaches are powerful in *C. elegans*. Genome-wide RNAi screens have been feasible since the creation of a library that contains about 80% of the *C. elegans* genes [8, 9]. A marked disadvantage of RNAi is that it only allows for the analysis of knockdown phenotypes and not for phenotypes caused by mutations that, by their nature, can give insight into overall gene function.

Previously, our laboratory described the development of a *C. elegans* transgenic reporter system that is able to visually monitor frameshifts in a microsatellite [7]. In this system, a LacZ gene is disrupted by the insertion of an A_{17} repeat placed immediately downstream of the start codon. This insertion puts the downstream ORF out-of-frame, however when a -1 frameshift occurs within the repeat (by the deletion of one nucleotide), the 3' end of the LacZ gene will come in-frame with the upstream start codon resulting in functional LacZ expression. Such events can thus easily be detected by staining animals with β-galactosidase. It was shown that virtually no staining is detected in transgenic animals with a wild type genetic background. However, in a *msh-6* mutant background, many cells express LacZ, as a result of frame-correcting frameshifts. This reporter was subsequently used to perform an RNAi screen to identify genes involved in avoiding MSI [10]. Several of the newly identified genes resulted in lethality when knocked down by genetic means (for example *rev-1*, described in Chapter 2 of this thesis). Because this obviously hinders genetic analysis we set out to isolate alleles of frameshift mutators via forward mutagenesis. Here, we describe the construction of a new MSI reporter, which contains several improvements compared to the one used previously. We show that this reporter is highly sensitive and specific to frameshift mutators. Transgenic animals carrying the reporter in low copy numbers were subsequently used in a clonal F2 forward mutagenesis screen. Out of 1800 mutagenized genomes, we isolated 21 mutants, five of which carried a mutation in a known MMR gene. The genetic defect in the other 16 strains is yet to be determined.

Materials and Methods

Strains and maintenance

General methods for culturing *C. elegans* were used [11]. The following strains were used: N2, CB4856, NL2511 [*msh-6(pk2504)*] and NL4812 [*unc-119(ed3) pkIs2175 [pRP1890 unc-119(+)]*]. The NLS::ATG-(C)23-GFP/LacZ reporter (pRP1890) was constructed by first oligocloning an (C)23 tract into pRP1821 [7] to generate pRP1851. An NLS sequence was inserted N-terminal to the (C)23 tract (pRP1858), and the reporter was subsequently shuttled into

the bombardment vector pRP2512 to generate pRP1890. Transgenic animals were generated via biolistic transformation, several independent lines were tested for MMR dependent LacZ correction, by crossing in an msh-6 allele. One strain was selected: NL4812, which contained an integrated array (pkIS2165) with 5-10 copies of the reporter. This allele was used in all experiments.

Forward screen

L4 animals were incubated in 0.05 M EMS in M9 solution for four hours [12]. After treatment, animals were placed on OP50 plates. Subsequently, 1500 F1 animals were singled and from each F1 plate, an F2 animal was picked to a fresh plate. When the F3/F4 generation was present on the plates, the animals were rinsed off and stained with β-galactosidase.

Mapping

For the initial mapping, the SNPs described in [13] were used. For finemapping, SNPs identified by [14] were used.

Results

The development of an improved MSI reporter

Despite the notion that the previously mentioned A_{17} MSI LacZ reporter identified all MMR genes in an RNAi based screen [10], it was not highly specific. The screen identified 61 knockdowns, of which only 13 increased the mutation rates in the micro-satellite repeat. The other 48 knockdowns also stained positive when an out-of-frame-reporter was used that lacked the A_{17} repeat. This lack of specificity may partially be explained by the fact that the reporter is present in a multi-copy complex array, thus containing repetitive sequences; rearrangements between different copies may lead to ORF restoration.

For this reason, we first aimed to make key improvements to the existing assay before initiating a time- and labor consuming genetic screen. We created new transgenic lines that differed in three aspects (Figure 1A): 1) the sensitivity of the reporter was increased by changing the length and the sequence of the microsatellite: we now used a C_{23} repeat as it was shown that DNA polymerase slippage occurs more often at cytosine tracts than on adenosine tracts [15]. 2) The specificity was addressed by using low copy integrants. 3) We added a nuclear localization signal (NLS) peptide sequence N-terminal to the LacZ protein. This allowed us to trace the events back to specific cell types and lineages.

We generated transgenic animals via biolistic transformation [16] and stained stable lines for B-galactosidase activity. Figure 1B depicts a representative population of animals. A clear

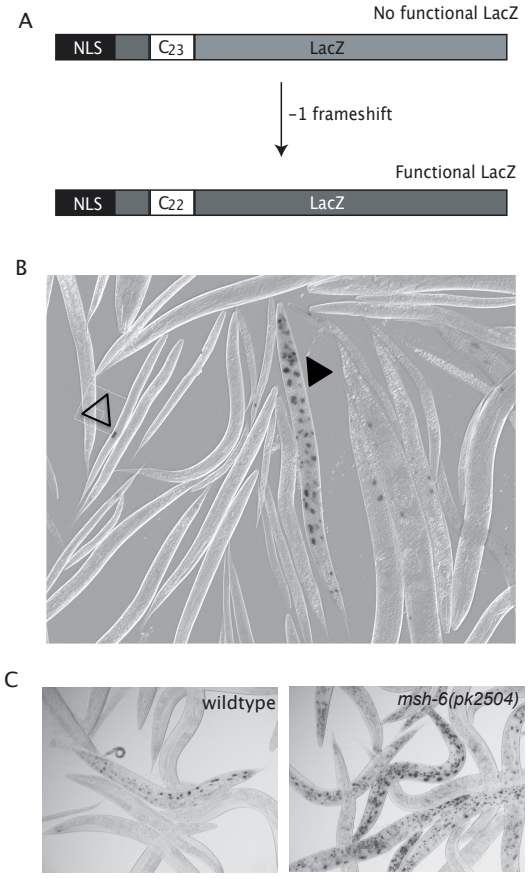


Figure 1 The newly developed C_{23} reporter. A) Schematic representation of the C_{23} reporter. The downstream LacZ ORF is placed out-of-frame with an upstream sequence that contains the start ATG and an NLS, by the insertion of a C_{23} tract. A -1 frameshift leads to an in-frame product that can be identified by staining with β-galactosidase. B) Staining of animals carrying the C_{23} MSI reporter transgene pkIs2165. Staining patterns differ between animals. The nuclear localization allows lineage tracing. Examples of early (closed triangles) and late (open triangles) events are indicated. C) MSI visualized in MMR proficient (wild type) and deficient (*msh-6(pk2504)*) genetic backgrounds.

stochastic pattern of expression can be observed indicative of spontaneous frameshifts that have occurred during DNA replication at some stage during development. Early events will lead to many positive cells (corrected copies are transmitted to progeny cells); late events will lead to animals with few spots. As can be observed many animals failed to show any cell expressing LacZ indicating that the mutation rate is below one per worm per generation.

We next assayed MSI in MMR defective animals by crossing an *msh-6(pk2504)* allele into the reporter lines. Figure 1C and 1D illustrate a strong increase in ORF correction in *msh-6(pk2504)* cultures, which demonstrates that this reporter can also be used to visualize defective MMR. To test whether this reporter was more specific than the one used previously, we tested various non-MSI knockdowns that were identified in the previous RNAi screen [10]. No increased LacZ restoration was observed for transgenic C_{23} reporter strains for all RNAi knockdowns tested, indicating that this new reporter is more specific for the MSI phenotype and will likely lead to less “false positive” hits in genetic or RNAi screens.

Forward mutagenesis screen identifies new alleles of C. elegans MMR genes

To identify mutations that cause MSI, a forward EMS mutagenesis screen [12] was performed. Mixed-stage populations were treated with EMS and 1200 F2 animals were cloned out. Because the LacZ staining protocol kills the animals, we assayed the progeny: half of the animal population were rinsed off and stained for β-galactosidase activity. This screen isolated 21 populations that showed increased LacZ staining.

Because MMR genes are the most likely hits, we first sequenced the 21 strains for the four known *C. elegans* MMR genes: *msh-2*, *msh-6*, *pms-2* and *mlh-1*. This led to the identification of five mutations in these genes (Table 1). All mutations were G/C to A/T transitions as expected, since it is the predominant mutation caused by EMS [12]. Of these mutations, one, PMS-2(Q21*), led to a premature stop codon, while the other mutations all resulted in amino acid changes. Two alleles of *mlh-1* were found and in both cases a conserved residue was mutated: MLH-1(A40V) and MLH-1(G18D). One allele of *msh-2* was identified: MSH-2(G273E). Here, an evolutionarily conserved amino acid was altered; the glycine resides in the highly conserved ATPase domain of MSH-2. The mutation in *msh-6*, which changed the proline at position 470 into a serine (MSH-6(P470S)) was not in a conserved region. Complementation analysis with e.g. *msh-6(pk2504)* will have to determine whether this mutation is causal.

We also sequenced the F45G2.3 gene, which encodes the probable *C. elegans* homologue of EXO1. EXO1 is a 5'→3' exonuclease that has been implied in MMR [17]. However, we did not identify a mutation in F45G2.3 in the 16 remaining mutant lines.

Table 1: New alleles of *C. elegans* MMR genes identified in the forward screen.

Gene	Allele	Mutation	Residue Change
mlh-1	*pk2272*	CCTCA{C>T}CGGCC	G18D
	pk2271	TGGGT{GG>AA}CGCCG	A40V
msh-2	*pk2258*	GAAGA{G>A}AACTT	G273E
msh-6	*pk2269*	TACTC{C>T}CGAAA	P470S
pms-2	*pk2275*	CCGCC{C>T}TAGGT	Q21*

Two apparent staining patterns emerge

We further analyzed the staining patterns that were detected in the identified mutants. The nuclear localization signal does allow identification of individual cells and therefore MSI events can be traced to the invariant cell lineage. In attempt to quantify the difference in various mutants, we counted blue spots in the intestinal tract. We focused on one lineage because there were too many events per individual animal to be able to count them separately. The intestinal tract was chosen for two reasons: 1) all ~ 40 cells are derived from one precursor cell E and 2) these cells are large and can easily be distinguished from neighboring

cells. We established four classes: I Worms with no staining at all, II Worms that have blue cells, but no intestinal ones, III Partial staining of the intestinal lineage, and IV Staining of all intestinal cells. As observed in figure 1B and 2, animal population with wild type genetic background (thus with functional MMR) mostly showed white worms and the second largest class was class II, with a few blue patches per worm. However, in most mutant strains the most predominant staining class is the one with all intestinal cells expressing B-gal, as is also the case for the *msh-6(pk2504)* control.

Interestingly, a few mutant strains showed a completely different staining pattern (pk2257, pk2274 and pk2277; Figure 2). They showed many LacZ positive cells, but almost no staining in the intestine was observed. This suggests that the MSI phenotype in these mutants is not caused by defective MMR, since all mutants carrying a mutation in a MMR gene showed the intestinal staining pattern. At present, we do not have an explanation for this observed phenotype.

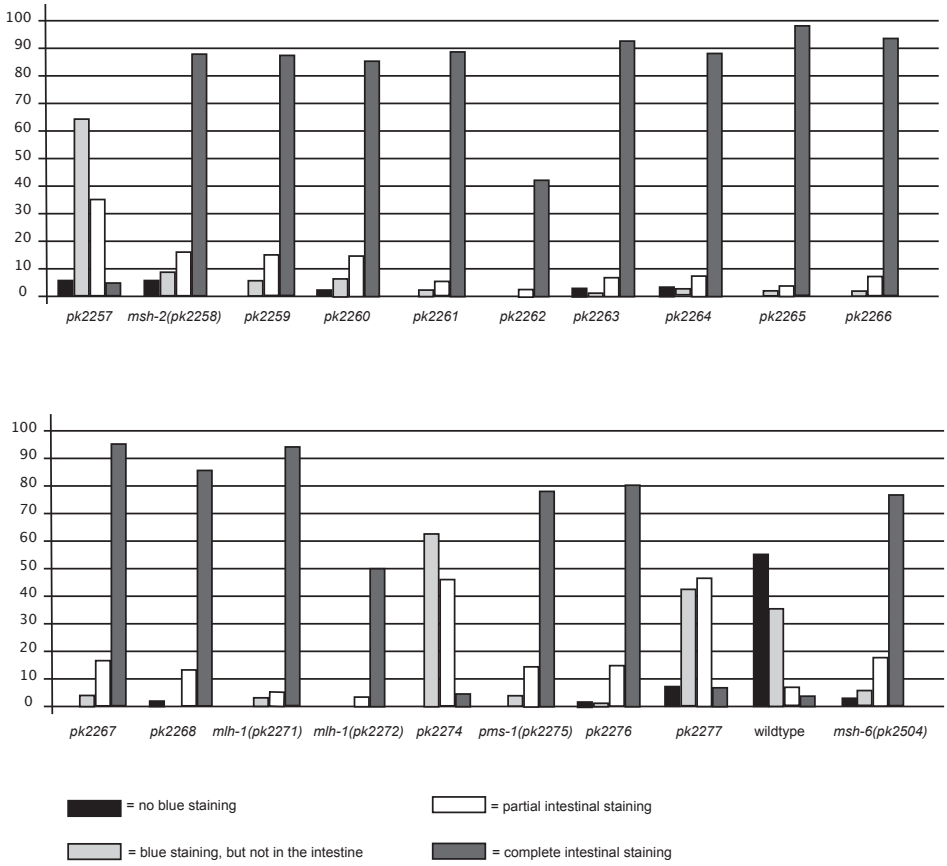


Figure 2: Quantification of staining patterns. Four classes of staining patterns were defined (see text) and the number of worms falling into every class was determined per allele.

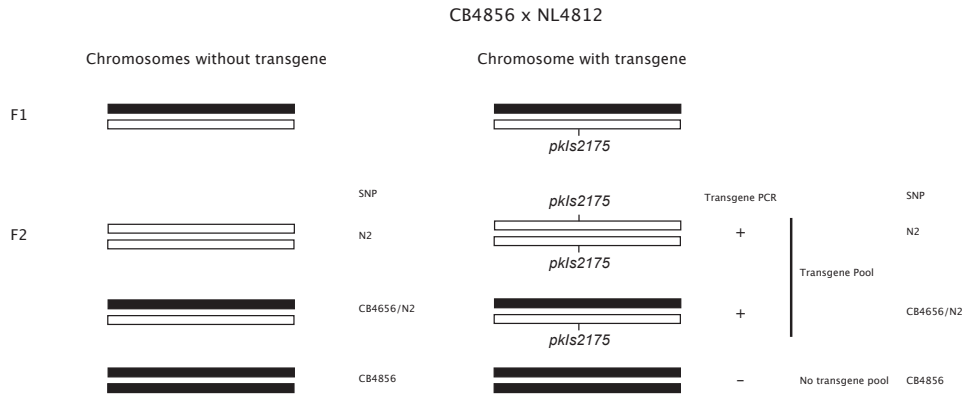


Figure 3: Stategy to map the reporter transgene. Bristol N2 worms that carry the transgene were crossed with Hawaiian isolate CB4856. F2 animals that lack the transgene should have CB4856 DNA at the location of the transgene. Therefore, an overrepresentation of CB4856 SNP in this pool will point to the genomic position of the transgene. All other unlinked genomic locations will be equally represented by N2 and CB4856 DNA.

Mapping of the MSI reporter

Five mutants have a mutation in a MMR gene. We started positional cloning efforts to identify the genetic defects in the 16 remaining mutant lines. We first determined the chromosomal location of the integrated transgene using single nucleotide polymorphisms (SNPs) that are identified in a Hawaiian isolate of *C. elegans*, CB4856. A large proportion of these SNPs lead to a modification of a restriction side, which allows rapid PCR-based mapping approaches [13, 14].

We crossed the Bristol strain that contains the MSI reporter transgene with the CB4856 strain and analyzed the F2 generation for the presence of the transgene by PCR (whether the transgene is homozygous or heterozygous cannot be determined). We made two pools: DNA of animals with or without the transgene. On both pools we performed SNP analysis with four evenly spaced SNPs per chromosome [13].

An underrepresentation of CB4856 in the DNA pool with the transgene will indicate its position. Vice versa, animals that lack the transgene should predominantly be CB4856 at that location (Figure 3). We positioned the MSI reporters on the left arm of chromosome III (Figure 4A). Next, we used additional SNPs to further finemap the location to a 6.3 cM interval (Figure 4B).

Attempts to map the remaining mutants fail

After mapping of the transgene we attempted to map the remaining mutations. This was not successful. In most cases, already in the first step we could not find linkage to any chromosome. For two alleles, *pk2266* and *pk2274*, initial linkage was found, namely to the right

arm of chromosome I and the left arm of chromosome IV, respectively. Here, the finemapping proved to be problematic. Fluctuations in the blue staining of heterozygous mutants led to improper qualifying of the phenotype. A pilot experiment in which a known allele (*msh-6(pk2504)*) was used in mapping experiments revealed that mapping strategies can principally be carried out but those strategies proved to be very labor intensive. We concluded that new approaches/strategies are required.

A

	No transgene pool	Transgene pool
I 1	N/C	N/C
I 2	N/C	N/C
I 3	N/C	N/C
I 4	N/C	N/C
II 1	N/C	N/C
II 2	N/C	N/C
II 3	N/C	N/C
II 4	N/C	N/C
III 1	C	N/C
III 2	C	N/C
III 3	N/C	N/C
III 4	N/C	N/C
IV 1	N/C	N/C
IV 2	N/C	N/C
IV 3	N/C	N/C
IV 4	N/C	N/C
V 1	N/C	N/C
V 2	N/C	N/C
V 3	N/C	N/C
V 4	N/C	N/C
X 1	N/C	N/C
X 2	N/C	N/C
X 3	N/C	N/C
X 4	N/C	N/C

B

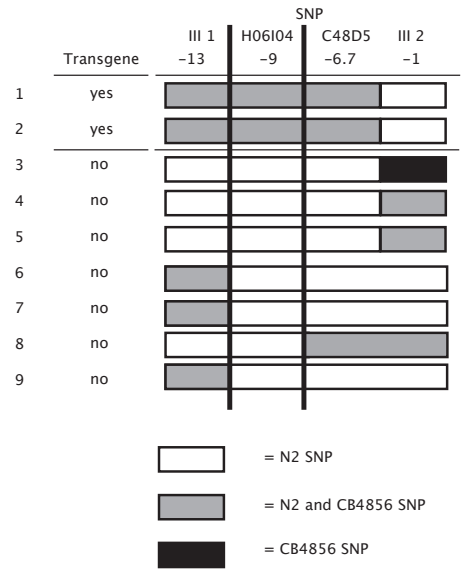


Figure 4: Mapping of the transgene. A) Initial mapping of the location of the transgene shows linkage to the left arm of chromosome III. For each of the six chromosomes (I through X), four SNP markers were analyzed. The genotype determined for each SNP marker is depicted: N stands for N2, N/C for both the N2 and CB4856 SNPs are present, and C for only the CB4856 SNP. In the DNA sample from worms that did not contain the transgene, an overrepresentation of CB4856 SNP for the III.1 and III.2 markers was present. This maps the transgene to the left arm of chromosome III. B) Finemapping of the location of the transgene. We analayzed nine individual lines (1 through 9) that had a variation in SNP pattern for the III.1 and III.2 markers. Genotypic depiction is in colors (see legend). At the -9 location all worms that do not contain the transgene, show the CB4856 SNP. At other locations, N2 SNPs are still detected. Therefore, the transgene is located in the 6.3 cM region surrounding the -9 SNP on chromosome III.

Discussion

Here, we describe a sensitive and specific MSI reporter that can be used to efficiently screen for genes involved in MMR. We performed a forward screen to identify mutants that show increased levels of MSI. In total, we found 21 mutant populations; five of those carry mutations in the four known *C. elegans* MMR genes. MMR genes are involved in the avoidance of MSI [2] and these results therefore validate the use of this reporter to screen for genes involved in MMR. The ability to screen for (new) MMR genes is of great importance. While the damage recognition steps in the MMR pathway are well understood, the mechanisms

involved in the further processing and repair of the detected mismatches are poorly understood. Identification of new MMR genes therefore can help to get a better understanding of these processes. The goal of this setup was to identify alleles of novel genes. These can be among the 16 lines we isolated in the screen, but how to identify the underlying defect? Classical positional cloning efforts proved to be unsuccessful for a number of reasons. Perhaps most importantly, the tremendously high rate of frameshifting (and LacZ correction) makes it very difficult to distinguish the mutant phenotype from germ line events. These events will transmit functional LacZ ORFs to progeny (these animals obscure the readout in genetic crosses). Very recent experiments with single copy C_{23} tracts at known genomic locations indicate the frequency of MSI in MMR defective worms is tremendously high: 20% per animal per generation. This suggests that each germ cell will have a functional LacZ ORF (~5-10 copies), making mapping and quantification impossible. Although one could principally omit animals that have LacZ expression in all cells (as these are likely caused by germ cells with a reverted copy), in our experience this does not solve the problem because there are some fluctuations in the penetrance of expression upon heat-shock (the promoter that drives LacZ expression).

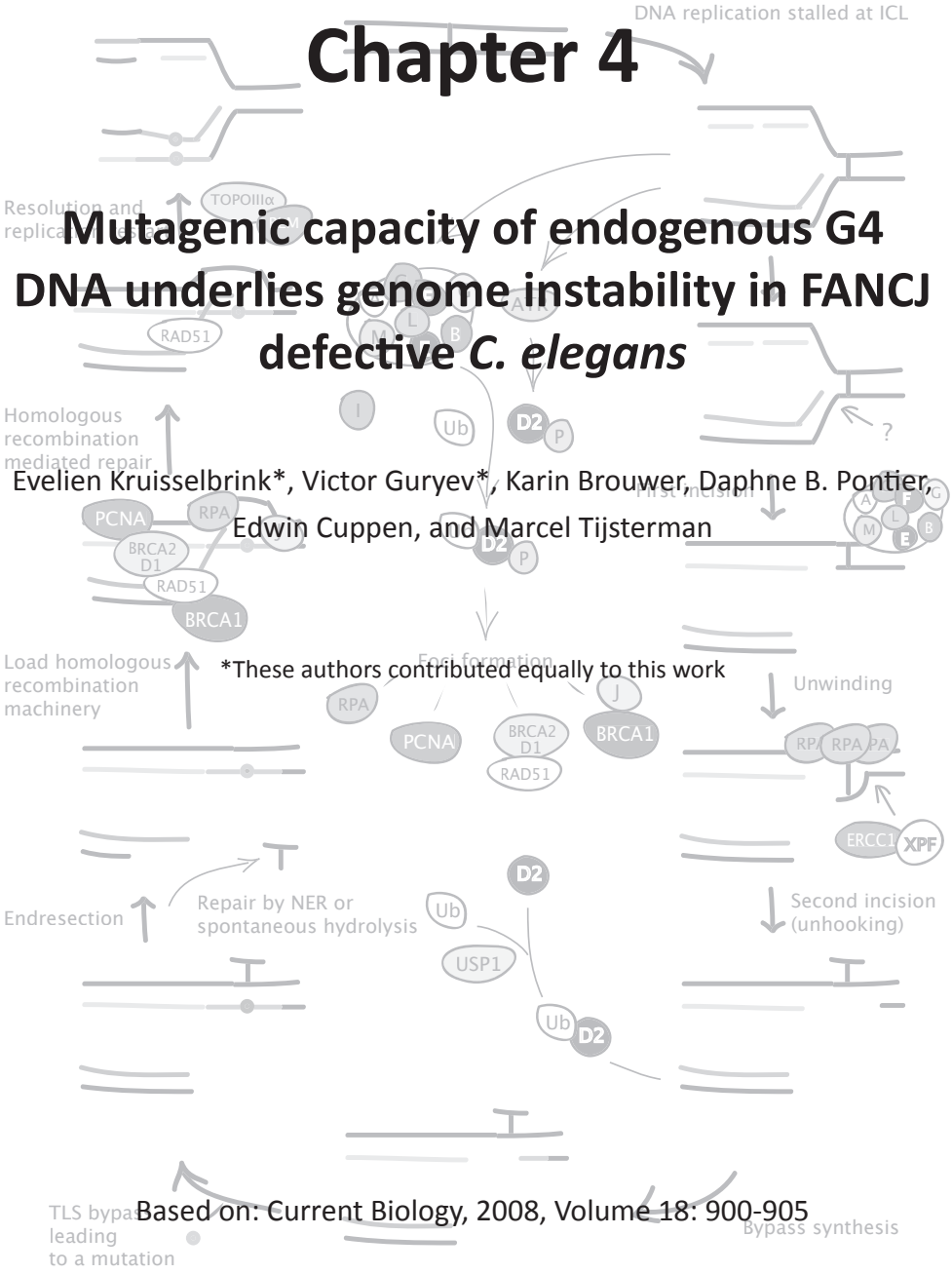
Based on these results various other schemes can be explored (single copy, multiple independent reporters, less sensitive transgenes, etc.), although it will require new screens. Alternatively, newly developed technology, e.g. whole genome sequencing, to determine the underlying defect in the 16 strains is currently available.

References

1. Harfe, B.D., and Jinks-Robertson, S. (2000). DNA mismatch repair and genetic instability. Annu Rev Genet *34*, 359-399.
2. Kolodner, R.D., and Marsischky, G.T. (1999). Eukaryotic DNA mismatch repair. Curr Opin Genet Dev *9*, 89-96.
3. Rustgi, A.K. (2007). The genetics of hereditary colon cancer. Genes Dev *21*, 2525-2538.
4. O'Neil, N., and Rose, A. (2006). DNA repair. WormBook, 1-12.
5. Denver, D.R., Morris, K., Lynch, M., and Thomas, W.K. (2004). High mutation rate and predominance of insertions in the Caenorhabditis elegans nuclear genome. Nature *430*, 679-682.
6. Degtyareva, N.P., Greenwell, P., Hofmann, E.R., Hengartner, M.O., Zhang, L., Culotti, J.G., and Petes, T.D. (2002). Caenorhabditis elegans DNA mismatch repair gene msh-2 is required for microsatellite stability and maintenance of genome integrity. Proc Natl Acad Sci U S A *99*, 2158-2163.
7. Tijsterman, M., Pothof, J., and Plasterk, R.H. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161*, 651-660.
8. Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325-330.
9. 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., and Ahringer, J. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231-237.
10. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H., and Tijsterman, M. (2003). Identification of genes that protect the C. elegans genome against mutations by genome-wide RNAi. Genes Dev *17*, 443-448.
11. Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.
12. Anderson, P. (1995). Mutagenesis. Methods Cell Biol *48*, 31-58.
13. Davis, M.W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S., and Jorgensen, E.M. (2005). Rapid single nucleotide polymorphism mapping in C. elegans. BMC Genomics *6*, 118.

14. Wicks, S.R., Yeh, R.T., Gish, W.R., Waterston, R.H., and Plasterk, R.H. (2001). Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map. Nat Genet *28*, 160-164.
15. Gragg, H., Harfe, B.D., and Jinks-Robertson, S. (2002). Base composition of mononucleotide runs affects DNA polymerase slippage and removal of frameshift intermediates by mismatch repair in Saccharomyces cerevisiae. Mol Cell Biol *22*, 8756-8762.
16. Berezikov, E., Bargmann, C.I., and Plasterk, R.H. (2004). Homologous gene targeting in Caenorhabditis elegans by biolistic transformation. Nucleic Acids Res *32*, e40.
17. Tishkoff, D.X., Boerger, A.L., Bertrand, P., Filosi, N., Gaida, G.M., Kane, M.F., and Kolodner, R.D. (1997). Identification and characterization of Saccharomyces cerevisiae EXO1, a gene encoding an exonuclease that interacts with MSH2. Proc Natl Acad Sci U S A *94*, 7487-7492.

Chapter 4



Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ defective *C. elegans*

Evelien Kruisselbrink*, Victor Guryev*, Karin Brouwer, Daphne B. Pontier, Edwin Cuppen, and Marcel Tijsterman

*These authors contributed equally to this work

Based on: Current Biology, 2008, Volume 18: 900-905

Summary

To safeguard genetic integrity, cells have evolved an accurate but not failsafe mechanism of DNA replication. Not all DNA sequences tolerate DNA replication equally well [1]. Also genomic regions that impose structural barriers to the DNA replication fork are a potential source of genetic instability [2,3]. Here, we demonstrate that G4 DNA – a sequence motif that folds into quadruplex structures in vitro [4,5]- is highly mutagenic in vivo and is removed from genomes that lack dog-1, the C. elegans ortholog of mammalian FANCJ [6,7], which is mutated in Fanconi anemia patients [8-11]. We show that sequences that match the G4 DNA signature G3-5N1-3G3-5N1-3G3-5N1-3G3-5 are deleted in germ and somatic tissues of dog-1 animals. Unbiased aCGH analyses of dog-1 genomes that were allowed to accumulate mutations in >100 replication cycles indicate that deletions are found exclusively at G4 DNA; deletion frequencies can reach 4% per site per animal generation. We found that deletion sizes fall short of Okazaki fragment lengths [12], and no significant microhomology was observed at deletion junctions. The existence of 376,000 potentially mutagenic G4 DNA sites in the human genome could have major implications for the etiology of hereditary FancJ and nonhereditary cancers.

Results and Discussion

Previously, we have built transgenic *C. elegans* strains to monitor frame-shifting errors that occur at DNA repeats [13] and observed that monoC/G tracts were much more error prone than monoA/T tracts of identical lengths. Although this observation is in agreement with worm and yeast data on endogenous repeats [14,15], we reasoned that genome rearrangements other than micro-satellite instability could also contribute. This notion was fueled by the identification of *dog-1*, a DNA helicase that protects monoC/G but not monoA/T tracts from being deleted [6]. Crossing a deletion allele of *dog-1* into reporter lines that monitor frameshifting at mono(C/G)23 tracts, however, failed to result in significantly increased reporter expression, likely because the tract is a very strong inducer of DNA frameshifts, resulting in many ORF restoring events per animal even in a mismatch repair proficient background.

To specifically investigate deletion induction at monoC/G tracts, we developed a phenotypic assay making use of the nonsymmetrical way deletions are induced at such sequences in *dog-1* deficient animals: deletions were previously found to start immediately 5' of a monoC/G tract to end at seemingly random locations a few hundred nucleotides downstream [6]. We therefore placed a monoC tract followed by mutiple stop codons downstream of a reporter gene's start codon, but in a non-essential region (Figure 1A). Only deletions that take out the mononucleotide repeat and all stop codons can bring the LacZ start codon in-frame with the downstream ORF. For clarity, we will adopt the term monoG- or G-tract induced deletions for these types of DNA rearrangements. Figure 1B shows lacZ express-

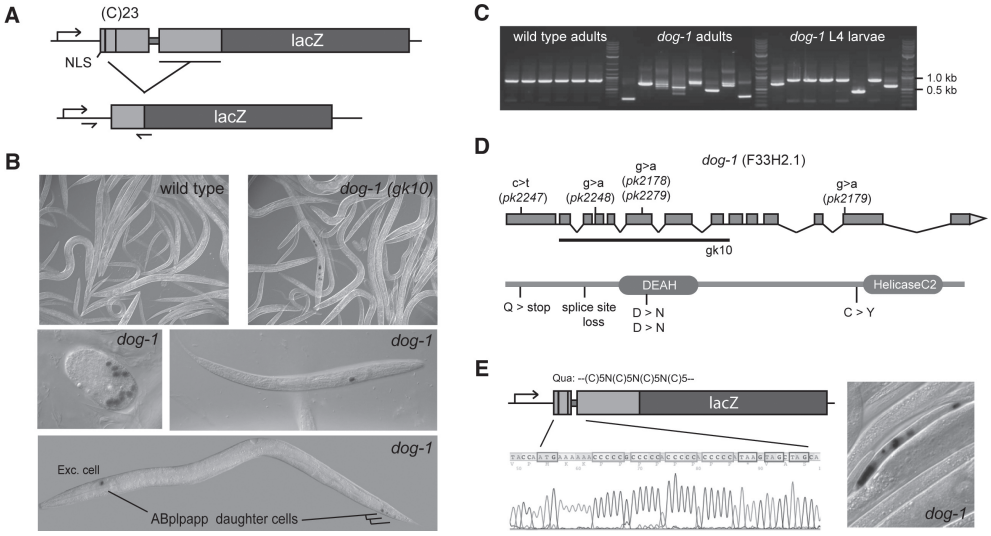


Figure 1: A transgenic reporter assay to monitor G tract induced genomic rearrangements. A) Schematic-drawing of a reporter LacZ gene that is made dysfunctional by interrupting the LacZ ORF with a $(C)_{23}$ repeat in the non-template strand and an in-frame ORF that contains stop codons. Only mutagenic events that delete the C-stretch (G tract) together with 3' flanking sequences can bring LacZ in-frame with the upstream start codon resulting in detectable B-galactosidase expression. B) *dog-1* deficient transgenic animals express B-galactosidase as a result of stochastic DNA rearrangements in somatic cells during development. The upper right panel depicts expression in a subset of intestinal cells. Unspecified early and late events are depicted in the middle panels. The bottom panel shows expression in cells located at the posterior and anterior ends of the *C. elegans* bodyplan as a result of a genomic rearrangement in the joint founder cell ABplpapp. C) Molecular analysis of G tract-induced deletion formation in single wildtype or *dog-1*-deficient adults or larvae of the L4 stage by PCR amplification of sequences flanking transgenic G tracts (primers indicated in panel A). D) Schematic illustration of independently derived *dog-1* alleles that were identified in a clonal F2 forward mutagenesis screen, as well as the *gk10* reference allele. The motif structure of the encoded protein is depicted underneath, as well as how the mutations affect the protein's structure or function. E) Schematic and sequence representation of a functional reporter system that colorimetrically visualizes G4 DNA fragility: Here, the LacZ reporter contains early stop codons immediately downstream of a $C_5NC_5NC_5NC_5$ sequence that predicts a G4 DNA quadruplex structure in the template strand. B-galactosidase-expressing cells identify G4 DNA induced genomic deletion events.

ing cells that appear in ~0.3% of *dog-1* defective animals but these are never observed in wildtype ($n > 10^6$). The observed patterns of expression were typical for stochastic events happening during DNA replication in animal development: both early and late events were observed, as reflected by many or few cells that express B-galactosidase (Figure 1B). All somatic tissues were susceptible. Note also that a single reversion event leads to expression in all progeny cells, which in combination with the nematode's invariant cell lineage allows us to trace back genomic deletions to a single somatic event during embryogenesis. Molecular PCR-based analysis of individual worms showed that these transgenes faithfully recapitulate the previously observed mutation spectrum at endogenous monoG tracts [15]: only in *dog-1*, but not in wildtype animals, we observed deletions that take away almost the

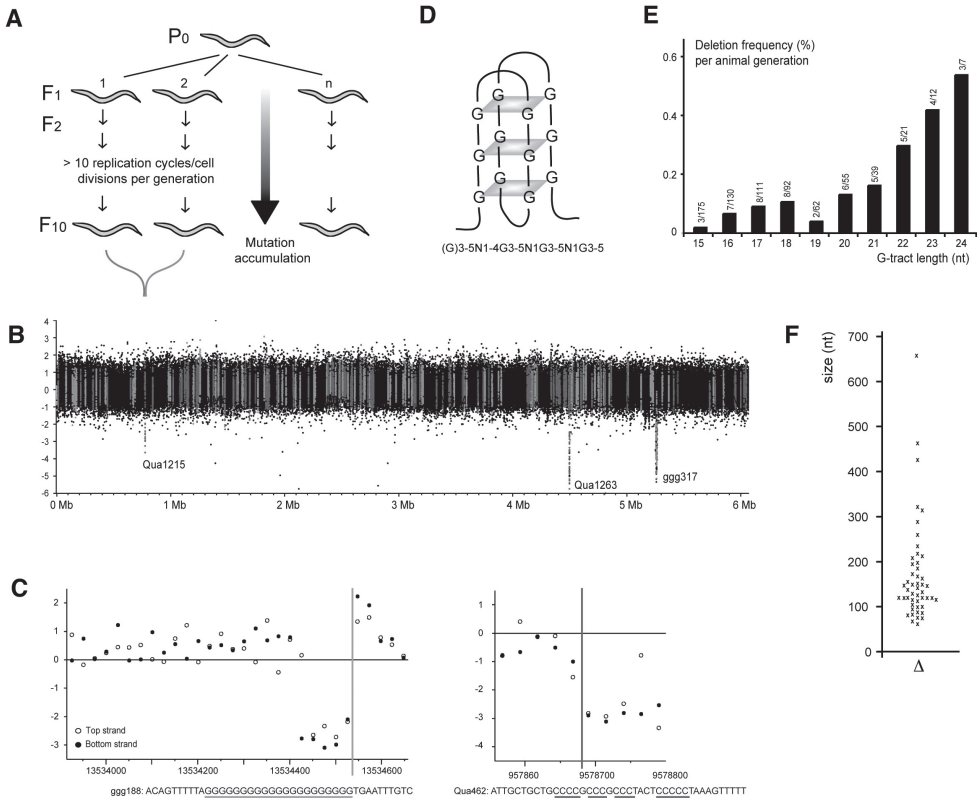


Figure 2: Array Comparative Genomic Hybridization (aCGH) identifies G4 DNA fragile sites. A) Culturing scheme leading to *dog-1* mutation accumulation (MA) lines. Here, we took advantage of the fact that *C. elegans* can be maintained as self-fertilizing hermaphrodites: we singled out the progeny of one parental animal (P0), allowed these sublines to grow for 10 generations, and then picked one (F10) animal to establish a new culture of which the DNA represents the genome of that F10 animal. These MA lines have thus independently gone through > 100 rounds of DNA replication. B) A typical log plot of ~300.000 tiled DNA probes covering the left arm of chromosome 5 is displayed for MA-14 over wildtype (N2). Probes that flank (<1000 bp) candidate fragile sites are coded in gray. This strain suffered three deletions in this interval, which are indicated. C) Graphical illustrations of a 144 bp deletion at ggg188 in MA-20 and a 157 bp deletion at Qua462 in MA-13. Closed and open circles represent probes against the chromosomal top and bottom strand. The x-axis indicates the chromosomal position on the *C. elegans* physical map; gray lines indicate the positions of the G4 DNA sequences of which the sequence is given underneath the log plots. D) Schematic representation of antiparallel quadruplex DNA in which planar rings of four guanines (G-quartets) can stack on top of each other, thereby forming an unusual DNA structure that is thermodynamically stable under physiological conditions. E) Graphical representation of deletion frequencies as a function of G tract length. The absolute number of G(n) induced deletions (for given n) are divided by the number of G(n)s in the *C. elegans* genome. These numbers are plotted above the bars. Note, that we only used deletions induced by monoG tracts that are not part of larger G4 DNA sequences: $(G)_{15\text{-}24}N_{1\text{-}7}(G)_{>2}$ and $(G)_{>2}N_{1\text{-}7}(G)_{15\text{-}24}$ were excluded. (F) Size distribution of G4 DNA induced germline deletions (n=49).

entire G tract together with 5' flanking sequence (Figure 1C + S1A). MonoA tracts are not susceptible to *dog-1* dependent deletion formation (Figure S1B).

We next questioned, what makes monoG tracts unstable? Is it because they are extremely prone to replication slippage (monoG>>monoA), and could deletions be the result of error-prone repair of DNA intermediates created by DNA mismatch repair proteins? We tested whether monoG tract instability was dependent on functional mismatch repair, but we found no effect of genetically inactivating *msh-6* on deletion induction in a *dog-1*-proficient or -deficient background (Figure S1). Rose and colleagues recently published a survey of *C. elegans* strains with mutations in various types of DNA-repair pathways, all of which had wildtype behavior for their ability to maintain monoG tracts in their genomes [7,16]. In agreement, we found that none of the major genome-maintenance pathways were involved in preventing monoG tract instability, by using our sensitive transgenic assay or population-based PCR techniques (Figures S1C to F; we estimate that we would be able to detect a deletion frequency that is < 1% of the deletion frequency observed in *dog-1* mutant animals). These also included Werner's and Bloom's Syndrome helicases –these proteins have been shown *in vitro* to unwind secondary structures that are formed in G-rich ssDNA [17,18].

In an unbiased approach to identify genetic determinants, we mutagenized transgenic animals that carry monoG tract instability reporters and assayed progeny animals that express B-galactosidase in sublineages, indicative of monoG tract induced rearrangements. We identified 5 mutants in ~ 4800 genomes, all of which were loss-of-function mutations in *dog-1* (Figure 1D). Although the forward and reverse genetic approaches presented here have not reached saturation levels, they fuel the hypothesis that monoG tracts may represent a special type of premutagenic lesion that is not acted upon by any of the known major DNA-repair pathways.

What features of monoG tracts make these sequences uniquely depend on functional DOG-1? To address this question, we took a genomics approach to identify additional fragile sites and look for common denominators. DNA of clonally grown *dog-1* cultures, to establish so-called "mutation accumulation" (MA) lines (Figure 2A), was assayed by comparative genome hybridization (CGH). We built custom-made *C. elegans* arrays, onto which we spotted a tiling path of ~300.000 overlapping probes covering the complete left arm of chromosome V, as well as ~ 81.000 probes aimed to investigate sequences surrounding candidate fragile loci: i) nucleotide repeats of various types, ii) palindrome resembling sequences potentially able to form hairpins in ssDNA, iii) G4 DNA sequences matching the signature $G_{3-5}N_{1-7}G_{3-5}N_{1-7}G_{3-5}N_{1-7}G_{3-5}$ (ssDNA containing such motifs have *in vitro* been shown to fold into four-stranded secondary structures, called quadruplexes [19,20]), and iv) G-rich telomeric DNA. Figure 2B displays a log-plot of the entire left arm of LGV for one *dog-1* MA strain compared to wildtype Bristol N2; figure 2C exemplifies two typical deletion profiles that were found at candidate fragile loci. In total, we identified 69 germ line deletions in 16 *dog-1* MA lines, versus zero in N2, and sequenced 51 of those (see supplemental table S1 for a complete

list plus characteristics), which allows us to draw conclusions with respect to frequency, distribution, size, sequence requirements, and flanking sequences. We will focus on three key findings.

Most importantly, we found that G4 DNA sites, but only G4 DNA sites, are fragile in *dog-1* deficient genomes. The chromosome V tiling path, covering 7% of the entire genome, identified eight deletions, all of which map to monoG tracts. In support of such a narrow spectrum of fragile sequences, we found no DNA rearrangements at telomeric sites at 93 $A_{17\text{-}24}$, at 599 $(A/G)_{17\text{-}24}$, and at ~1000 sequences that could potentially form DNA hairpins. In contrast, the candidate approach revealed 62 deletions at large monoG tracts ($G_{n>14)}$), 4 deletions at monoG-like sequences (having 1 or 2 nucleotides that interrupt a monoG tract), and 3 deletions at sequences that deviate from the monoG type (e.g. GGGGGagtaGGGcGGGcGGGG), but have the potential to fold into a quadruplex structure: having four stretches of at least three guanines interrupted by nucleotides of any type (Table S1; Figures 2C and 2D). In all deletions, the fragile site is taken out almost completely together with 5' (but not 3') flanking DNA. We argue that the unique feature that makes monoG tracts unstable is that they match the G4 signature and thus have the ability to adopt a quadruplex structure. Although the relative abundance of deletions at monoG tracts versus non-monoG G4 DNA (62 versus 3) seems to contradict the relative abundance of the two classes of DNA sequence in the *C. elegans* genome (525 versus 1742), this can be fully explained by assuming that longer monoG sequences have a greater degree of liberty to fold into a quadruplex. For example, non-monoG Qua462 (Figure 2C) can fold into six different quadruplex structures, while a monoG sequence of similar size (22nt) offers 362 possibilities. In support of this notion, we observed that the deletion-inducing capacity of G4 DNA increased with tract length and G-ratio (Figure 2E; data not shown).

Another notable observation from the aCGH data is the total lack of sequence similarity around the deletion junctions. We found no evidence for microhomology driven repair (or bypass) by scanning 51 germline deletions at different genomic loci or 18 deletions at one transgenic G4 DNA sequence in somatic cells (Table S1 and Figure S1A). In particular, the lack of apparent homology between nucleotides positioned upstream but within the deleted segment with unaffected sequences immediately downstream of the fragile site argues against mechanisms involving microhomology-dependent DSB repair at G4 DNA-blocked replication forks.

Finally, we observed that deletions are predominantly of small size. The array design combined with the algorithms we developed allow us to detect deletions larger than ~ 50-70 basepairs, but there is no upper constraint, apart from *in vivo* limitations where large deletions may cause lethality. Indeed, we identified deletions in the range of 63 – 7347 nucleotides. However, in 88% of the cases, less than 300 bp have been deleted, the average size of which is 141 bp. Such size distribution hints at a model in which the deletion size is determined by the distance between a lagging strand replication fork that is stalled at a G4

DNA sequence to the nearest upstream Okazaki fragment (the average length of eukaryotic Okazaki fragments being ~ 300 bp [12]).

We estimate that our CGH data is derived from ~ 1600 replicated genomes. Although most deleted sites were represented only once we found two exceptions: ggg317 (3x) and ggg463 (2x) This indicates that the mutagenic potential of these sequences can be enormous: ~ 4% for ggg317 per animal generation. The majority of G4 DNA sites, however, have not been deleted, likely because the rate of deletion induction per site is far below the rate required to see it here. Alternatively, many sites are not intrinsically mutagenic; previous work suggested that only half of the monoG tracts are fragile, leading to the speculation that monoG tract fragility could predict whether DNA sequences are replicated via the lead-

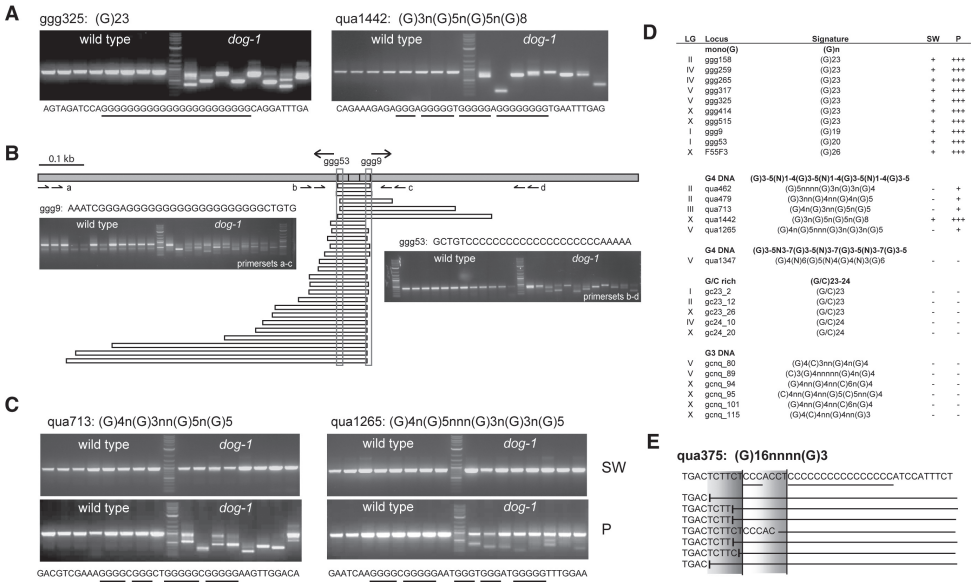


Figure 3: Mutagenic determinants of endogenous G4 DNA sequences. (A) PCR analysis of endogenous G4 DNA sequences ggg325 and qua1442 on single wildtype and *dog-1(pk2247)* gravid animals (n=8), separated by a DNA size marker containing lane. Both wildtype products are ~0.9 kb. (B) Schematic diagram of an endogenous locus that contains two G4 DNA tracts having opposite polarity. The sequences are given above gel images that display PCR analysis on these loci in wildtype and *dog-1* mutant animals (n=12). For each genetic background, five gravid animals were pooled in one sample. A representative set of smaller than wildtype bands, obtained with primers a and d, were purified and sequenced. The resulting deletions are graphically displayed beneath the locus diagram. The vertical boxing is to highlight that all deletions are on one site flanked by G4 DNA flanking sequence; that border defines the fragile G4 sequence. (C) Single worm (SW) and population (P) based (10^4-10^5 animals) PCR analysis of G4 DNA sequences qua713 and qua1265 for the indicated genotype (n=8). Wildtype products are ~ 1 kb. (D) Schematic sum up of endogenous G4 DNA instability in *dog-1(pk2247)* ranked per type and chromosome number (LG). A + marking indicates a deletion frequency of > 6% (2 smaller bands in 32 samples); +++ means that individual bands were difficult to discriminate (instead DNA smears are observed resulting from amplification of many differently sized (deletion-containing) fragments per DNA sample). (E) A graphic illustration of the deletion start sites that occur at G4 DNA sequence qua375. Sequences that are not deleted are depicted. For qua375, there are two potential deletion initiation sites: deletions in the left gray zone likely result from a DNA replication blocking quadruplex that includes the 3' $(G)_3$ sequence (or 5' $(C)_3$), whereas a quadruplex made up of only the upstream (G)16 sequence is predicted to trigger deletion induction in the right gray zone.

ing or lagging strand [6]. We here show that this is not the case: all tested monoG$_{>20}$ tracts are mutagenic (Figures 3A and 3D). In addition, we identified two DNA tracts that were only 39 bp apart but located in opposite DNA strands, and both are fragile (Figure 3B). While we concur with the idea that quadruplexes are preferentially formed in the lagging strand, our observation is consistent with data from other systems suggesting that there are no fixed origins of replication during development [21].

Next, we addressed the question whether all sequences that match the used G4 DNA signature are fragile. First, we tested a custom-made G4 DNA sequence G$_5$NG$_5$NG$_5$NG$_5$ in transgenic reporter animals and found it to induce deletions in a *dog-1*-dependent manner (Figure 1E). Second, we developed more sensitive PCR assays on endogenous loci -we titrated PCR conditions to optimally amplify smaller than wildtype products- and found that all sequences that match the G4 DNA signature but have a limited number of nucleotides in between the G4 DNA legs are fragile (Figures 3C and 3D). In contrast, G3 DNA sites (sequences that resemble G4 but miss one "leg" of a possible quadruplex) were never fragile. We also investigated one case in which a monoG tract had a (G)$_3$ tract 4 nt away at its 3' flank. We found that six out of seven deletions started close to the (G)$_3$ sequence (Figure 3E), >7 nucleotides away from the monoG tract, indicating that the start of the fragile site was determined by the extra (G)$_3$ leg and not by the monoG tract.

Our molecular analyses demonstrate that endogenous sequences that have the ability to fold into quadruplex structures depend on DOG-1 to persist in *C. elegans* genomes. How do cells deal with these replication-blocking sequences? One previously suggested explanation [6] involves the unwinding of the quadruplex structure by DOG-1/FancJ's helicase activity to allow replication to proceed. Such a scenario would, however, not explain why *dog-1* animals are also sensitive to DNA crosslinks ([7] + Figure S2), because these cannot be unwound. This could point to a dual function of DOG-1: operating together with other Fanc proteins in crosslink repair but acting independently of them in a genome-maintenance pathway that prevents loss of G4 DNA. Alternatively, quadruplexed DNA constitutes strong replication impediments also in the presence of DOG-1, and replicative bypass is established not via unwinding of the quadruplex but via invasion and subsequent replication of the nearby already-replicated leading strand. The helicase activity of DOG-1 could help to unwind the replicated leading strands dsDNA to allow DNA synthesis. Such activity could also be envisaged for repair of crosslink damage [22].

The question of which enzymatic activities are involved in converting the premutagenic lesion (the quadruplex) to loss of sequence information (a deletion) is unanswered. Rose and colleagues have identified factors (e.*g.* homologous recombination proteins) whose loss stimulated deletion induction at monoG tracts in *dog-1* animals [16], suggesting that these factors play a role in an error-free way of dealing with G4 DNA. Loss of G4 DNA is, however, a consequence of an error-prone pathway. Candidate gene approaches have thus far suggested that none of the repair pathways that operate on DNA double-strand breaks

(DSBs) are involved: inactivation of components of nonhomologous end joining (NHEJ), HR, or single-strand annealing (SSA) did not suppress deletion formation in a *dog-1* background ([16] and Figure S1). The notion that NHEJ components are not required is all the more surprising in the light of the observed lack of obvious homology at the deletion junctions.

Our transgenic setup, which perfectly mimics endogenous G4 DNA fragility, now allows a further investigation of the genetic and molecular determinants that influence replication progression *in vivo* at known locations that are amenable to genetic manipulation.

In summary, we have provided the first evidence that G4 DNA sequences that have the potential to fold into replication blocking quadruplex structures are intrinsically mutagenic in live animals. To prevent massive genome rearrangements at G4 DNA sites, cells require a specialized genome protection mechanism that involves *C. elegans* FancJ, but not any of the other genes causally linked to Fanconi anemia [7,23,24]. Future work will have to uncover whether these fragile sites -there are more than 376,000 predicted G4 sites in the human genome [25] - are causally linked to at least part of the genomic rearrangements seen in tumors of human FancJ patients or in nonhereditary cancers.

Acknowledgements

We thank Ronald Plasterk for generous support; Marit Kosters and Astrid Eijkelenboom for technical assistance; the CGC, the knockout consortia, Simon Boulton, and Shohei Mitani for strains; and all Wormbase curators for building the platform that allows *C. elegans* genomic research. This work was supported by a ZonMW VIDI grant to M.T.

Experimental Procedures

Strains and culturing conditions

See supplemental data for the *C. elegans* strains that were used in this study Animals were grown at 20 °C [26]. Transgenic strains were created via biolistic transformation or via gonadal injections followed by integration of extrachromosomal arrays by X-ray irradiation. B-galactosidase expression was assayed as described previously [13].

Array design and bio-informatical analyses

We used WS140 *C. elegans* genome built to design 40 to 60-mer probes with fixed 72 °C Tm for Nimblegen 388.5k microarray chips according to manufacturer's instructions. On average, adjacent probes overlap 50%. The array (precise design available upon request) contained a 7.7 Mb tiling path of LGV as well as probes directed at sequences that flank candidate fragile sites: all candidate fragile sites were flanked on each site with 10 probes

(top and bottom strand) except monoG tracts; these were flanked by 60 upstream and 10 downstream probes. Nimblegen performed hybridizations for 20 DNA samples: 2 Bistol N2, 16 *dog-1* MA lines, and 2 DNA samples derived from the *dog-1* P0 animal. We predicted deletions by comparing hybridization intensity ratios of every MA line to N2. We created a deletion-candidate probe set that included probes with the most extreme ratios (0.1% from both tails of ratio distribution). Deletions were called if a sequence segment was represented by at least two of three consecutive probes in the candidate probeset. 96 DNA segments were chosen for PCR and sequence verification. aCGH data were confirmed by analyzing 96 genomic loci by PCR amplification and sequencing. Primers were designed to target strong candidate deletions -those were handpicked upon visual inspection of log plots of all regions that scored positive using our algorithms- as well as a number of negative controls and cases that were ambiguous. This led to a 100% verification rate for the highest scoring subset of candidate loci.

DNA analysis and reporter cloning

Analysis of fragile sites on endogenous loci was performed with nested sets of primers (sequences available upon request) and PCR conditions were optimized per primerset to favor the amplification of smaller than wildtype bands. Reporter transgene cloning: for all variants we started with pRP1821 [13] that contains a heat-shock driven GPF::lacZ ORF lacking a start codon. For pRP1878 (pkIs2165), we oligo-cloned an ATG-NLS-(C)23 sequence upstream of the GFP::lacZ ORF and then placed a stop codons at the XhoI site of GFP. For pRP3020 (lfIs17), we placed an ATG-(G4 DNA)-stop codon sequence in front of the GFP::LacZ fusion. pRP1889 contained a monoA tract at that position.

References

1. Harfe, B. D. & Jinks-Robertson, S. DNA mismatch repair and genetic instability. (2000). Annu Rev Genet *34,* 359-399.
2. Brewer, B. J., Lockshon, D. & Fangman, W. L. (1992). The arrest of replication forks in the rDNA of yeast occurs independently of transcription. Cell *71,* 267-276.
3. Lambert, S., Watson, A., Sheedy, D. M., Martin, B. & Carr, A. M. (2005). Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. Cell *121,* 689-702.
4. Gellert, M., Lipsett, M. N. & Davies, D. R. (1962). Helix formation by guanylic acid. Proc Natl Acad Sci U S A *48,* 2013-2018.
5. Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature *334,* 364-366.
6. Cheung, I., Schertzer, M., Rose, A. & Lansdorp, P. M. (2002). Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. Nat Genet *31,* 405-409.
7. Youds, J. L., et al. (2008). DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Mol Cell Biol *28,* 1470-1479.
8. Levran, O. et al. (2005). The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. Nat Genet *37,* 931-933.
9. Levitus, M. et al. (2005). The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. Nat Genet *37,* 934-935.
10. Bridge, W. L., Vandenberg, C. J., Franklin, R. J. & Hiom, K. (2005). The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. Nat Genet *37,* 953-957.

11. Litman, R. et al. (2005). BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ. Cancer Cell *8,* 255-265.
12. Abdurashidova, G. et al. (2000). Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science *287,* 2023-2026.
13. Tijsterman, M., Pothof, J. & Plasterk, R. H. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161,* 651-660.
14. 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.
15. Sia, E. A., Kokoska, R. J., Dominska, M., Greenwell, P. & Petes, T. D. (1997). Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol *17,* 2851-2858.
16. Youds, J. L., O'Neil, N. J. & Rose, A. M. (2006). Homologous recombination is required for genome stability in the absence of DOG-1 in Caenorhabditis elegans. Genetics *173,* 697-708.
17. Fry, M. & Loeb, L. A. (1999). Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)n. J Biol Chem *274,* 12797-12802.
18. Sun, H., Karow, J. K., Hickson, I. D. & Maizels, N. (1998). The Bloom's syndrome helicase unwinds G4 DNA. J Biol Chem *273,* 27587-27592.
19. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res *34,* 5402-5415.
20. Maizels, N. (2006). Dynamic roles for G4 DNA in the biology of eukaryotic cells. Nat Struct Mol Biol *13,* 1055-1059.
21. Gilbert, D. M. (2005). Origins go plastic. Mol Cell *20,* 657-658.
22. Niedernhofer, L. J., Lalai, A. S. & Hoeijmakers, J. H. (2005). Fanconi anemia (cross)linked to DNA repair. Cell *123,* 1191-1198.
23. Mirchandani, K. D. & D'Andrea, A. D. (2006). The Fanconi anemia/BRCA pathway: a coordinator of cross-link repair. Exp Cell Res *312,* 2647-2653.
24. Levitus, M., Joenje, H. & de Winter, J. P. (2006). The Fanconi anemia pathway of genomic maintenance. Cell Oncol *28,* 3-29.
25. Huppert, J. L. & Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. Nucleic Acids Res *33,* 2908-16.
26. Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77,* 71-94.

Supplemental Experimental Procedures

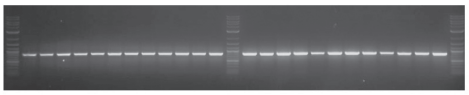
C. elegans strains that were used in this study are N2; *dog-1(gk10); dog-1(pk2247); dog-1(pk2179); dog-1(pk2178); dog-1(pk2248; msh-6(pk2504); div-1(or148); lig-4(ok716); cku-80(ok861); brc-1(tm1145); fcd-2(tm1298); brc-2(tm1086); him-9/xpf-1 (e1487); mre-11(ok179) IV/+; dpy-13(e184) rad-51(lg8701) IV/nT1; brd-1(dw-1); rcq-5(ok660); him-6(ok412); wrn-1(tm764); mus-81(tm1937); F10G8.7(tm2073); mrt-2(e2663); atm-1(gk186); xpa-1(ok698); pkIs2165[pRP1878: hsp-16.41::ATG-(C)23-stops-LacZ; unc-119]; pkIs2165 dog-1(gk10); pkIs2165 dog-1(pk2247); msh-6(pk2504) pkIs2165; msh-6(pk2504) dog-1(pk2247); lig-4(ok716) dog-1(gk10); cku-80(ok861) dog-1(gk10); brc-1(tm1145) dog-1(gk10); fcd-2(tm1298) dog-1(pk2247); brc-2(tm1086)/+ dog-1(pk2247); him-9/xpf-1 (e1487) dog-1(pk2247); mre-11(ok179)/+ dog-1(gk10); dpy-13(e148) rad-51 (lg8701) IV/+ dog-1(pk2247); wrn-1(tm764) him-6(ok412); msh-6(pk2504) dog-1(gk10) pkIs2165. pkIs2172* [*pRP1889: hsp-16.41::ATG-(monoA)-stops-LacZ; unc-119]; dog-1(pk2247) pkIs2172; lfIs17 [pRP3020: hsp-16.41::ATG-Quadruplex-stops-LacZ; pRF4: rol-6(su10060)]; dog-1(pk2247) lfIs17.*

A

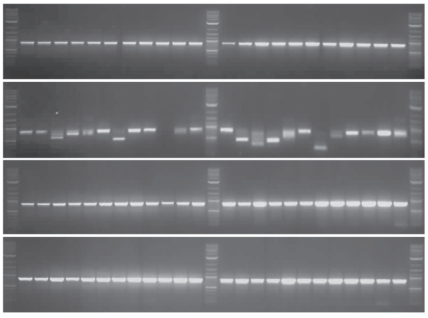
Upstream sequence	3'-end of deleted seq.	downstream sequence	del size	Insertion
CCCGGGTACGATGACCCCCCCCC	TCCCAATTCTTGTTG	AATTAGATGGTGATG	81	-
CCCGGGTACGATGACCCCCCCCC	TGAATTAGATGGTGA	TGTTAATGGGCACAA	95	-
CCCGCGTACGATGACCCCC	GTTAATGGGCACAAA	TTTTCTGTCAGTGGA	115	-
CCCGGGTACGATGA	GATGTTAATGGGCAC	AAATTTTCTGTCAGT	117	AGA
CCCGGGTACGATGACCC	CAACATACGGAAAAC	TTACCCTTAAATTTA	163	-
CCCGGGTACGATGACCCCCCCC	TAAATTTATTTGCAC	TACTGGAAAACTACC	180	-
CCCGGGTACGATGAC	TTGCACTACTGGAAA	ACTACCTGTTCCATG	196	AC
CCCGGGTACGATGA	CCATGGGTAAGTTTA	AACATATATATACTA	222	-
CCCGGGTACGAT	AGTTTAAACATATAT	ATACTAACTAACCCT	233	-
CCCGGGTACGATGACCCCCC	AACCCTGATTATTTA	AATTTTCAGCCAACA	249	AC
CCCGGGTACGATGAC	ATTTAAATTTTCAGC	CAACACTTGTCACTA	264	AGT
CCCGGGTACGATGACCCC	CGACCATGGATAGGG	ATAACAGGGTAATAG	318	-
CCCGGGTACGATGACCCCCCC	ATAGGGATAACAGGG	TAATAGTCGAGATAC	324	ACATACGT
CCCGGGTACGATGACCCC	CGGCATGACTTTTTC	AAGAGTGCCATGCCC	372	-
CCCGGGTACGATGACCCCCCC	TGTACAGGAAAGAAC	TATATTTTTCAAAGA	407	-
CCCGGGTACGATGACCCCCCCC	ACCCTTGTTAATAGA	ATCGAGTTAAAAGGT	533	-
CCCGGGTACGATGACCCCCCCC	GTATTGATTTTAAAG	AAGATGGAAACATTC	561	-
CCCGGGTACGATGACCCCCCC	GTTGTAAGTTTAAAC	ATGATTTTACTAACT	669	-

CCCGGGTACGATGACCCCCCCCCCCCCCCCCCCCCCCCATACGTACCGCTAGCAAAAAAAAG TAA

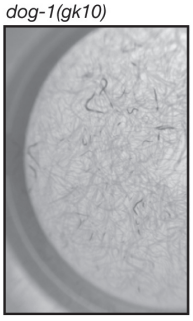
B



C

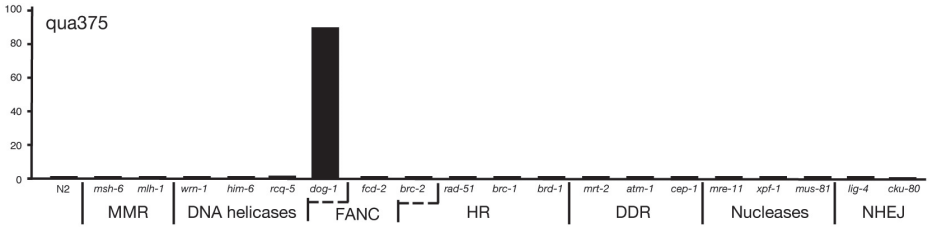


D

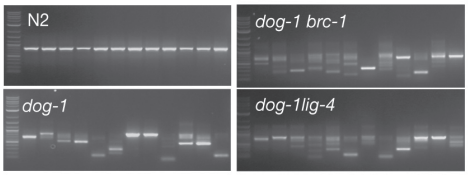


N2	-
dog-1	++
msh-6	-
dog-1 msh-6	++
lig-4	-
lig-4 dog-1	++
cku-80	-
dog-1 cku-80	++
brc-1	-
dog-1 brc-1	++
brd-1	-
rcq-5	-
him-6	-
wrn-1	-
mus-81	-

E



F



G

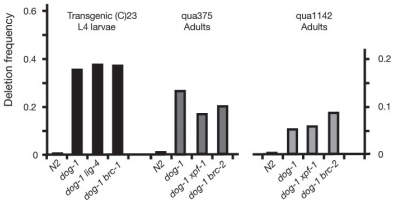


Figure S1 (left page): Genetic requirements for G4 DNA stability. A) The molecular signature of G tract induced genomic rearrangements was determined by PCR amplification of sequences flanking transgenic G tracts (primers indicated in Figure 1A) followed by sequencing of smaller than wildtype bands. The sequences that remain present (upstream/downstream sequence) are listed, as well as the 3' end of the deleted segment, the size of the deletion, and sequences that were found inserted. B) Molecular analysis of monoA-tract induced deletion formation in single *dog-1* deficient transgenic (pkIs2172) adults by PCR amplification of sequences flanking transgenic A-tracts (identical primers as used in (A)). C) Examples of PCR analysis of single gravid animals (n=24) for endogenous G4 DNA sequence qua375. Gel images are portrait for wildtype N2, *dog-1(pk2247)*, *msh-6* (*pk2504*) and *rad-51*(*lg8701*). The *lg8701* allele is linked to a *dpy-13* visible marker, and Dpy animals were picked and analysed for G tract instability and for homozygosity of the *rad-51* allele. D) Populations of animals with the indicated genotype were assayed for the G tract instability phenotype using reporter transgenes. Several thousand animals stained for B-galactosidase expression are displayed. Parallel cultures (n>5) were assayed and scoring was performed without the researcher knowing the identity of the respective genotypes. (++) indicates a similar to identical frequency of X-gal stained animals as compared to a reference *dog-1*(*gk10*) containing strain. (–) indicates that cultures were completely devoid of B-galactosidase expressing animals. E) Graphical representation of deletion frequencies at the endogenous qua375 sequence for animals with mutations in various DNA damage repair and/or signaling genes (MMR: Mismatch repair, FANC: Fanconi anemia genes, HR: Homologous recombination, DDR: DNA damage responses, NHEJ: Nonhomologous end-joining. For all genotypes at least 24 populations of 5 animals were assayed by PCR. F) PCR analysis of endogenous G4 DNA sequence qua375 on wildtype N2, *dog-1(gk10)* single- and *dog-1(gk10) brc-1*(*tm1145*) and *dog-1(gk10) lig-4*(*ok716*) double mutant animals (n=12). Here, every sample contains the DNA of 5 gravid animals. G) Quantification of deletion frequencies at transgenic (C)23 sites (n=32), and endogenous sequences qua375 and qua1142 (n=96) in single animals of the indicated genotype and developmental stage.

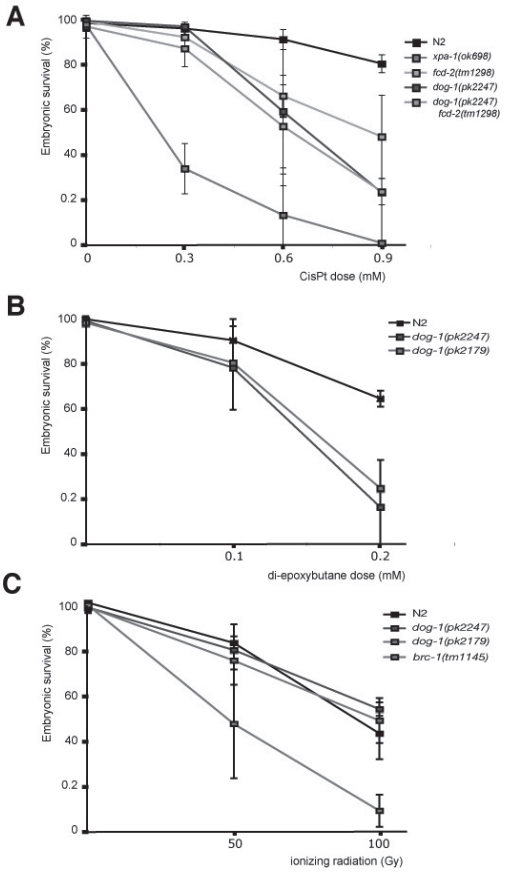


Figure S2: Sensitivity to DNA damaging agents. Animal survival curves for the indicated genotypes when exposed to the crosslinking agents Cisplatin (A) and di-epoxybutane (B) or ionizing radiation (C). Animals of the larval L4 stage were treated with the indicated dose, and the ratio live progeny/total laid eggs was determined for a 24 hours time window 24 hours post treatment.

Table S1. Deletions identified by aCGH

LG	locus	(G)n	upsteam	G4 DNA	downsteam	strain	size	Left flank	Del: 5'----3'	right flank	Insert
I	ggg039	20	ACTTCGAGTA	CCCCCCCCCCCCCCCCCCCC	TACACTGTAC	dog1_13	140	CACTTCGAGTAC	(C)19---GTGTGTATGTAG	TGCATAACTTGT	AT
I	ggg054	20	TTTTAGAGAG	CCCCCCCCCCCCCCCCCCCC	GGCCACTTGT	dog1_3	116	TCATGTTTTTAG	AGAG(C)20---TTAAGGCCC	GGATGCCCAGAT	
I	ggg063	18	CCAGTTTTTG	CCCCCCCCCCCCCCCCCC	TCCTCTTAGG	dog1_8	121	CCAGTTTTTG	(C)18---CATAAACTAGTC	CCCCCCCCTTGTG	
I	ggg064	23	CGTCTCTATT	CCCCCCCCCCCCCCCCCCCCCC	TTCTCCCCGT	dog1_4	169	AACGTCTCTATT	(C)23---ATGTGGAGCTCT	GTTTGCTTTTTTT	
I	ggg067	17	ATTGCGGCGG	CCCCCCCCCCCCCCCC	ACATTCCCTA	dog1_17	663	ATTGCGGCGGCC	(C)15---GTTAACGGCTTT	TTAATACCTCCA	
I	ggg075	19	ACAAGAAGCA	CCCCCCCCCCCCCCCCCCC	TCCCGCCCAT	dog1_4	84	CAAGAAGCACCC	(C)16---TATTCGGATTTT	TATGTGCAATTT	
I	qua0015	14	CAACGTAAGA	CCCTCCCCCCCCCCCCC	GCCAAACCTT	dog1_16	151	ACAACGTAAGAC	CCT(C)14---ATCGCGAGAC	GTCGTTGAATC	
I	qua0066	15	TTTAAAAGTG	CCCCACCCCCCCCCCCCCC	TATTTCCTAT	dog1_16	117	CTTTAAAAGTGC	CCCA(C)15---GCTATTGATATC	TGTCGGCATCTA	AAGAAAAG
I	qua0124	12	CGCAATAAT	CCCCCGGGGTACCCCCCCCCCCCGCCC	TAAAACCTTT	dog1_17	149	ATCCCCCGGGGT	A(C)12---TCAATCGATCGT	TTTTCAACGGGT	
II	ggg104	17	GTATTTACCT	CCCCCCCCCCCCCCCC	GATCTACCCG	dog1_4	148	TGTATTTACCTC	(C)16---CATTATATTATT	AGGTTGTTTTTT	
II	ggg106	21	CATAAACTTA	CCCCCCCCCCCCCCCCCCCC	AATTTTTTGC	dog1_17	464	TAAACTTACCCC	(C)17---CAGAGCTCTG	ACTGAAATTTC	TGAAATT
II	ggg107	18	CACAGTCTCA	CCCCCCCCCCCCCCCCC	ACATTTTAAC	dog1_1	427	ACACAGACACAG	TCTCA(C)18---CCTTCAATG	CTATAGGAACCC	
II	ggg110	22	TCAAGGTGTG	CCCCCCCCCCCCCCCCCCCCC	GTCCGTCTAA	dog1_2	151	TCAAGGTGTGCCC	(C)18---GGCTTTACC	TCTCACACATCT	
II	ggg121	17	ATTATTATTG	CCCCCCCCCCCCCCCC	TCCAGTCTGT	dog1_20	187	TATTATTCGCCC	(C)17---GAAATCCAAAA	AATTTTAATTTTG	
II	qua0262	16	TCTTTGAAAT	CCCCTCCCCTCCCCCCCCCCCCCCC	TTCCAGAAAC	dog1_2	122	TTGAAATCCCCT	(C)16---CTTTGCAATCGA	TTTTTAACAGAAAAA	15 nt
II	qua0375	16	TGACTCTTCT	CCCACCTCCCCCCCCCCCCCC	ATCCATTTCT	dog1_20	196	TTTATTGACT	TTCTCCCACCT(C)16--CGTGCGT	GGATATGGAGAA	
II	qua0429	8	TCAAAACATT	CCCCCCCCTCCCTCCCTTCACTCCC	GGTTAAAATT	dog1_8	96	TCAAAACATTCC	(QUA)---TTTATTTGCCAC	TTTTTTATTGAC	AAAAAACA
II	qua0462	5	ATTGCTGCTG	CCCCGCCCGCCCTACTCCCCC	TAAAGTTTTT	dog1_13	157	GATTGCTGCTG	(QUA)---CACTTTTTGTGCT	TTGTTAAAATTC	
III	ggg172	17	GAGCGACGAT	CCCCCCCCCCCCCCCC	ATCAACGTTC	dog1_17	214	GAGAGCGACGATC	(C)16---TAACGCACTTTC	TCTATGTTTCTC	
III	ggg188	21	GACAAATTCA	CCCCCCCCCCCCCCCCCCCC	TAAAAACTGT	dog1_20	144	TTGACAAATTCA	(C)21---CCGTACTTTTTT	TCCTTTTTTTTCTCT	
III	ggg217	19	TTTTTTCTCT	CCCCCCCCCCCCCCCCCC	ACAAATACAT	dog1_20	164	CGCTAGTTTTTT	CTCT(C)19---GAGAGAGTCG	AACTAGGTTCAC	(T)16
IV	ggg233	18	AATTTGTTTG	CCCCCCCCCCCCCCCCC	ACCTCATCGT	dog1_5	121	AAATTTGTTTGC	(C)17---AAATTCGGA	AATTTGAATGTA	
IV	ggg237	24	TTCATGCCTT	CCCCCCCCCCCCCCCCCCCCCCCC	GATAAGGTCG	dog1_5	290	TATTCATCCCTT	(C)24---GTGAGAGTGAAA	AGCGCCAAGAGG	G
IV	ggg257	17	AAATAATTTG	CCCCCCCCCCCCCCCCC	AGTGCCTTCA	dog1_17	210	GAAATAATTTGCCC	(C)14---GTTTTTTCTGTG	ACTATTTTCAAG	
IV	ggg268	18	GGTTTTTTGTG	CCCCCCCCCCCCCCCCCC	AACTGTTTTC	dog1_8	236	CGGGTTTTTGTG	(C)18---CTGGTTCTAT	TGATTAATTCTT	TTATTTTTGTG
IV	qua0788	16	AACATGACTA	CCCCCCCCCCCCCCCCACCCC	GCACCGCAGC	dog1_4	82	AAAACATGACTA	(C)16---AAAGTTGCAAAT	TATAAAAATGTA	
IV	qua0945	16	ATGTAGAATA	CCCCCCCCCCCCCCCCGCCCCC	GCCTCAAAAT	dog1_2	134	TATAATGTAGAA	TA(C)16---AAATTTCTTTT	TTTTTGTATTTT	AAAAA
V	ggg317	23	GCATTTTACG	CCCCCCCCCCCCCCCCCCCCCCC	GCCTCTTGTC	dog1_14	7347	AAATATTTCGAG	CATATCAAAC---AGTTTGAATA	AACCTTGACACG	
V	ggg317	23	GCATTTTACG	CCCCCCCCCCCCCCCCCCCCCCC	GCCTCTTGTC	dog1_2	219	CGGCATTTTACG	(C)23---GGATTTCTAACA	ATTGCTGATGGC	
V	ggg317	23	GCATTTTACG	CCCCCCCCCCCCCCCCCCCCCCC	GCCTCTTGTC	dog1_5	113	TTTACGCCCCC	(C)18---AAATTGGATTCA	CTTCTTCTTATA	
V	ggg326	18	AGACCCACTG	CCCCCCCCCCCCCCCCC	TCCCCCCCCG	dog1_16	106	AGACCCACTGCCCC	(C)14---CACGCAAGACCC	TGCGAGACCTTG	
V	ggg333	21	ACTAACCACT	CCCCCCCCCCCCCCCCCCCC	AATCTTTAAG	dog1_19	78	ACTAACCACT(C)8	(C)13---GTTTTTTAATGC	TTTCATAGAGCTTT	25 nt
V	ggg344	21	AATAAAGCCT	CCCCCCCCCCCCCCCCCCCC	ACCCAAAAAA	dog1_17	101	AACAATAAAGCC	T(C)21---GCATCACAGACC	GAAATACTTTTA	
V	ggg349	24	CATCATCGCT	CCCCCCCCCCCCCCCCCCCCCCCC	ACCGACCGAT	dog1_19	261	CATCATCGC	T(C)24---TTTTCGTCGT	CGTCGTCGTCGT	
V	ggg372	17	TGTTGTAGGATA	CCCCCCCCCCCCCCCC	ACTCTGAAAT	all strains	121	TGTTGTAGGATA	(C)17---CATGCTTTTTCA	ATTTTCGAACTC	
V	qua1162	14	ATAGAGGTCT	CCCCCCCACCCCCCCCCCCCC	GCTGCCGCCG	dog1_8	76	TTTAATAGAGGT	CTCCCCCCCA(C)14--TCTGTTC	GTTCTTCAATAG	

Table S1_continued. Deletions identified by aCGH

LG	locus	(G)n	upsteam	G4 DNA	downsteam	strain	size	Left flank	Del: 5'----3'	right flank	Insert
V	qua1215	16	GAATTTTTCT	CCCCCCCCCCCCCCC	ATGGTCTGTG	dog1_14	316	TTTGAATTTTTC	T(C)16---TTCTCAAATATT	TATACGAAAGAA	ATCA
V	qua1263	16	TGTTTCGTGT	CCCCCCCCCCCCCCC	ACCAACACCG	dog1_14	1328	CAAATTTTGTTT	CGTGT(C)16---GTAGGCTTGC	ACCCGTCTTGC	
V	qua1324	16	TCAAATCAAG	CCCCCCCCCCCCCCCCTTCAACCC	TTAAGCCCGG	dog1_5	153	TTCAAATCAAGC	(C)15---TATTTAGTAA	AAAAATGCGTGA	T
X	ggg385	17	CTCCAGTACT	CCCCCCCCCCCCCCCC	AGCAACTCCC	dog1_15	63	TCTCCAGTACTC	(C)16---GTGTACGTGTGC	TCATATATTTGT	AATATTTT
X	ggg389	20	TCACGTCGCG	CCCCCCCCCCCCCCCCCCC	ATTCTCGAAG	dog1_4	121	TTCACGTCGCGC	(C)19---TTGATTTGGAGA	AAGTGAAGAAG	GAAG
X	ggg410	20	TTCCCTAAGA	CCCCCCCCCCCCCCCCCCC	ATCTTACCAA	dog1_13	70	CCCTAAGACCCC	(C)16---CATTTTGGAATG	TTTTGTTATGAT	
X	ggg428	20	CCAAATGAAA	CCCCCCCCCCCCCCCCCCC	AAAATTTCAT	dog1_15	102	AATGAAACCCC	(C)16---TTTTTGTTTTTC	TGCACTTTTTTC	
X	ggg461	20	CTTTTTCCTA	CCCCCCCCCCCCCCCCCCC	TCCTACTAGT	dog1_15	131	CTTTTTCCTACC	(C)18---TTTTCCGCAAT	TTCCTCAATTCT	
X	ggg463	22	TTGTTTCAAA	CCCCCCCCCCCCCCCCCCCCC	GCCTACACCT	dog1_16	199	ATGTATTGTTTC	AA(C)22---GTGAAACTGGG	TCATTTATCCGG	TTCTCATGTC
X	ggg463	22	TTGTTTCAAA	CCCCCCCCCCCCCCCCCCCCC	GCCTACACCT	dog1_17	87	TTGTTTCAAA(C)10	(C)12---CACTCAAGTTTC	AAAAGGCGTCTA	(A)8
X	ggg510	22	TCCTTTTTCA	CCCCCCCCCCCCCCCCCCCCC	TCCCTATTCC	dog1_15	89	ACTCCTTTTTCA	(C)22---AGTAAACCTAGA	AGTATTATTTCC	
X	ggg510	22	TCCTTTTTCA	CCCCCCCCCCCCCCCCCCCCC	TCCCTATTCC	dog1_20	127	CCTTTTTCA(C)10	(C)12---TGTAGTTTCCTC	TTTTTTTTTTTGAA	AAATTTTT
X	ggg525	19	CCGCGCCGA	CCCCCCCCCCCCCCCCCC	GCTTACCTGC	all strains	323	ACCGCGCCGACC	(C)19---AATTAATTATGT	AAAATCATTTTC	
X	qua1442	8	CTCAAATTCA	CCCCCCCCTCCCCCACCCCCTCCC	TCTCTTTCTG	dog1_13	174	ACCTCTCAAATT	CA(QUA)---CTTGGAGAAATT	TTCCATTCTTG	
X	qua1590	15	GGCGCCTGAG	CCCCCCCCCCCCCCC	AAAAAAAGTG	dog1_2	121	TGGCGCCTGAGC	(C)14---TTTCGTTTGTTT	AAAATAACATTT	AAA

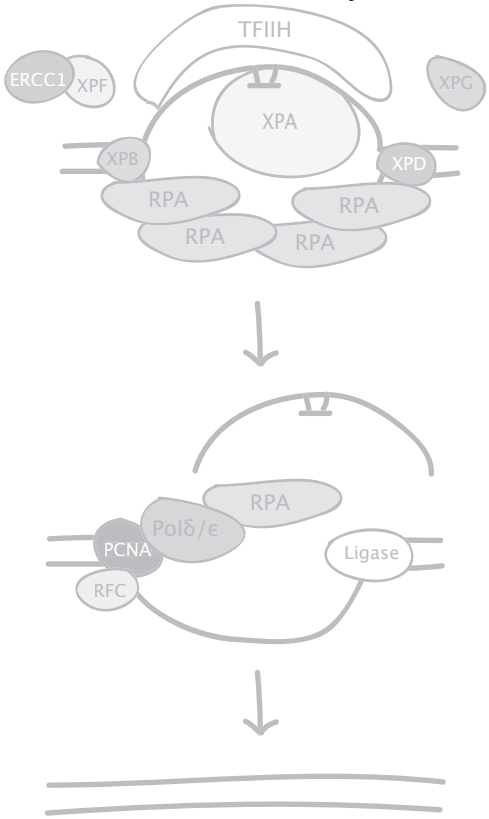
LG	locus	(G)n	upsteam	G4 DNA	downsteam	strain	deletion size	verification status
I	ggg012	18*	TCTCACACTG	CCCCCCCCCCCCCCCCC	ACCTAAAAGC	dog1_3	10 probes	PCR comfirmed
I	ggg064	23	CGTCTCTATT	CCCCCCCCCCCCCCCCCCCCCC	TTCTCCCCGT	dog1_14	3 probes	Non overlapping Smith-Waterman prediction
I	qua0195	16	GTGACTGATT	CCCCCCCCCCCCCCCCGATCCCCC	ACTCACTTGC	dog1_6	2 probes	Non overlapping Smith-Waterman prediction
II	ggg137	22	CTGTTCTATT	CCCCCCCCCCCCCCCCCCCCCC	GCCTACCTTG	all strains	5 probes	PCR confirmed, no sequence
III	ggg181	18*	TTTACCCATT	CCCCCCCCCCCCCCCCC	AAAAAAAAGT	dog1_2	>700 nt	Absent in PCR
III	ggg200	21*	TTTTTGAAAA	CCCCCCCCCCCCCCCCCCCC	GATTTTTTTT	dog1_3	6 probes	nd
III	ggg202	21*	CTAAACCTTT	CCCCCCCCCCCCCCCCCCCC	TCCCGTTGTC	dog1_10	6 probes	PCR confirmed, no sequence
III	qua0606	15	TGGTCATCGT	CCCCCCCCCCCCCC	ACCTCCTTCG	dog1_2	168	168 bp deletion= 52 nt downstream of C-tract
IV	ggg283	18*	ATTTTTCTTA	CCCCCCCCCCCCCCCCC	AACCACTCAA	dog1_8	>25 probes	nd: endpoint not predicted
IV	ggg290	18	GTTTCCCCAA	CCCCCCCCCCCCCCCCC	AATCTCGCCA	dog1_1	>25 probes	Refined script prediction
V	ggg310	24*	CAAAAAACTG	CCCCCCCCCCCCCCCCCCCCCCCC	GATTGTAGCA	dog1_10	12-13 probes	PCR confirmed: Complex rearrangement/insert
V	ggg324	21	CCCAGGACCA	CCCCCCCCCCCCCCCCCCCC	AATAAATTTC	dog1_16	3 probes	Non overlapping Smith-waterman prediction
V	qua1178	15	TTAGTCATGA	CCCCCCCCCCCCCCTCCC	TGAGCCACCT	dog1_8	>5 probes	Refined script prediction
X	ggg408	19	TGCAAGAGTA	CCCCCCCCCCCCCCCCCC	ACCTTCTCCA	dog1_14	5 probes	Refined script prediction
X	ggg413	17*	CCCCACGCCA	CCCCCCCCCCCCCCCC	TTCCCTCCCA	dog1_3	6 probes	(C)14---complex rearrangement
X	ggg414	23	ATGCACTTTG	CCCCCCCCCCCCCCCCCCCCCC	GACTATTACT	dog1_20	5 probes	Refined script prediction
X	ggg464	22	GCAATTATAG	CCCCCCCCCCCCCCCCCCCCC	ACACTGATAT	dog1_2	3 probes	Refined script prediction
X	qua1548	14	GGAGTTTCTG	CCCCCCCCCCCCCTTCCCCCC	AAAAGAATGA	dog1_10	>5 probes	Refined script prediction

* Also included in Figure 2e because these deletions are either PCR confirmed or part of a larger data set that scored a 100% verification rate.

Chapter 5

A genome-wide RNAi screen in *C. elegans* identifies modulators of the cellular response to the DNA crosslinking agent Cisplatin

Karin Brouwer and Marcel Tijsterman



Abstract

Cisplatin is one of the mostwidely used anticancer drugs. Chemotherapeutic activity has been demonstrated for this compound against a variety of cancers. We employ the model system *C. elegans* to identify genes that protect cells against Cisplatin-induced toxicity. We show that exposure of whole animals to the drug results in DNA damage response phenotypes (cell cycle arrest and apoptosis) in germ cells; on the organismal level, we find embryonic arrest and brood size reduction. We find that Cisplatin treatment primarily results in DNA deletions of various lengths in the *C. elegans* genome. A genome-wide RNAi screen, targeting ~ 16,000 *C. elegans* genes, identified 51 genes that influence the sensitivity of the worm to Cisplatin. These include genes acting in chromatin remodeling, signal transduction, and posttranslational modification, such as ubiquitinylation.

Introduction

Cisplatin is one of the most widely used anticancer drugs. Platinum-based chemotherapy is remedial for the majority of testicular cancers [1, 2]. Also, activity against lung, ovary, bladder, head and neck, esophagus, and endometrium cancers has been demonstrated. Identification of genes that, when absent, cause increased sensitivity to Cisplatin, could lead to the discovery of new drug targets to intensify Cisplatin-based chemotherapies.

Cisplatin targets cells by binding to DNA: it preferentially (90%) forms intrastrand crosslinks on GpG and ApG sequences [3] and these adducts are repaired by the nucleotide excision repair (NER) pathway [4]. The toxicity is largely attributed to the formation of interstrand crosslinks (ICLs), which covalently link the two DNA strands, mostly between two guanine residues in GpC and CpG sequences [5]. Interstrand crosslinks are repaired by the interstrand crosslink repair pathway [6]: in a complex, multistep reaction, the DNA is repaired by the action of proteins encoded by the Fanconi Anemia genes, by proteins that also act in NER, and by proteins that have been linked to the repair of DNA double-stranded breaks (DSBs). For many of the proteins that have been implicated in ICL repair (based on the fact that mutants in the encoding genes confer increased cellular sensitivity to crosslinking agents), the exact function remains to be determined. Both types of Cisplatin adducts are cyto- and genotoxic as they block DNA transcription and replication. Elimination of these lesions is thus essential for cellular survival.

C. elegans provides a simple multicellular model in which the cellular responses to DNA damaging agents can be studied [7, 8]. Most DNA damage response genes are conserved throughout evolution and have homologues encoded by the *C. elegans* genome. Two RNAi libraries that contain dsRNA clones to systematically knockdown 17,000 and 11,000 (mostly overlapping) *C. elegans* genes, respectively, are available. These reagents provide a tremendous opportunity to connect genotype to phenotype in a systemic way, and genome-wide screens can be profitable, provided that the assay or readout for biological function is sensitive, specific and scalable [9, 10]. For *C. elegans*, one of the easiest phenotypes to readout is

sterility or the failure to (rapidly) produce viable progeny. This, combined with the development of protocols to perform genome-wide RNAi in liquid 96-well culture plates, makes the nematode a suitable system for pharmacological screens aimed to identify genes that are important for cellular survival. Such an approach can link a drug to cyto- or genotoxic response pathways. Reading out a negative impact on the fate of the progeny from exposed mothers means that the target tissue of the drug/compound are the cells that are located in the germ line of the animal. *C. elegans* germ cells start their lives as mitotically active cells in the distal region of the gonadal syncitium. Cells in the mitotic zone proliferate but will start a meiotic program when they are outside the reach of the distal tip cell (DTC); GLP-1 signaling in the germ line negatively regulates meiotic entry in response to a signal from the DTC throughout larval and adult stages. Once germ cells have entered the meiotic pathway, germ cell chromosomes undergo homologous recombination. As meiotic chromosomes progress from diplotene to diakinesis, they become highly condensed, forming six discrete bivalents that can be clearly visualized in oocytes, where the cell cycle is halted until the oocyte is fertilized by sperm in the spermatheca. During oogenesis in *C. elegans,* many germ cells undergo the apoptotic program in the loop region of the gonad and are engulfed by the gonadal sheath cells. Apart from their detrimental effects on embryogenesis, DNA adducts can induce a physiologically detectable cell cycle arrest in the mitotic compartment of the germ line as well as an increased level of apoptotic cells in the loop region. In this study, we have first ascertained the cyto- and genotoxic responses of the nematode to the DNA-damaging agent Cisplatin and then performed a genome-wide RNAi screen to identify genes that modulate the cellular response to this drug.

Materials and Methods

Strains

The following *C. elegans* strains were used: wild-type Bristol N2, *xpa-1(ok698)*, *fcd-2(tm1298)*, *xpf-1(e1487)*, *set-2(ok952)*, *coh-3(gk112)*, *skpt-1(ok851)*, R05G6.10*(ok1159)*, *ulp-1(ok1768)*, *unc-93(e1500)* and *bcIs39(*P_{lim}*::ced-1::GFP)*.

Sensitivity to Cisplatin and ionizing radiation

L4 animals were treated with 0, 0.3, 0.6 and 0.9 mM Cisplatin (Cisplatine Mayne, 50 mg/50 ml, ONCO-TAIN®) for three hours in M9 or irradiated with 0, 50 and 100 Gy (in a Gammacell 1000 [Cs-137]) on agar plates. After treatment, four plates with three L4's per dose were incubated for 42 hours. Then, parents were removed and dead eggs and live offspring were scored 24 hours thereafter.

Brood size determination

Worms were treated with Cisplatin as described above. After treatment, 10 L4 worms per dose were singled. Every 24 hours, parents were transferred to fresh plates. Directly after removal of the parent and 24 hours later, the progeny was counted.

unc-93(e1500) *reversion assay*

L4 *unc-93(e1500)* animals were treated with 0, 0.15 and 0.3 mM Cisplatin for three hours. After treatment, 20 animals were placed on 9 cm NGM plates with OP50. The F2 generation was scored for the occurrence of revertants. The reversion rate was calculated using the following formula: r = (number of plates segregating revertants)/(2 x number of F1 animals per plate x number of plates) [11]. The number of F1 animals per plate was determined to be 653 and 562 for treatment with 0.15 and 0.30 mM Cisplatin respectively.
The reversion-causing mutations were determined by PCR and subsequent sequencing of the coding regions of the *unc-93*, *sup-10* and *sup-11* genes.

Cell cycle arrest and apoptosis

Wild-type Bristol N2 animals or wild-type animals carrying the *bcIs39(P_{lim}::ced-1::GFP)* transgene expressing CED-1::GFP were treated with 0.9 mM Cisplatin or irradiated with 120 Gy ionizing radiation. Cell cycle arrest was determined by quantification of the number of nuclei in the mitotic zone in N2 animals 18 hours after treatment. Apoptotic response was determined 24 hours after treatment. Cells completely surrounded by CED-1::GFP were considered apoptotic.

Genome-wide RNAi screen for Cisplatin sensitivity

A bacterial culture of 500 µl of LB medium containing 50 µg/ml ampicillin, inoculated from the RNAi library [10] was grown overnight in deep-well blocks at 37 °C. Next, the cultures were induced with IPTG (250 µg/ml) for 4 hours at 37 °C. The animals were bleached and the resulting embryos were synchronized overnight in M9. ~50 synchronized L1's were grown per well of a flat-bottom 96-well tissue-culture plate in 50 µl $M9^+$ (M9 with 10 µg/ml cholesterol, 50 µg/ml ampicillin, 12 µg/ml tetracycline, 200 µg/ml IPTG and 0.1 µg/ml fungizone). 70 µl of induced bacterial suspension was added per well. The RNAi cultures were grown at 20 °C while shaking 150-200 rpm [12]. After 36-39 hours of growth L4 animals were pipetted into M9 buffer and incubated with 0.75 mM Cisplatin for three hours. After treatment animals were pipetted back into fresh RNAi food plates containing 70 µl induced bacterial cultures and 50 µl $M9^+$. After four days, lack of progeny was scored. The dose of Cisplatin used was such that in most cases, there were only a few surviving progeny. Wells that totally

lacked progeny in the treated sample but did contain a normal amount of progeny in the untreated control were scored as positive. RNAi foods scored positive were repeated in two independent experiments.

Results and Discussion

Cisplatin exposure induces DNA damage responses in the C. elegans *germline*

We first determined appropriate conditions for investigating the cellular response to Cisplatin and performed RNAi knockdown experiments on a genome-wide scale. Previous work [13] indicated that animals are most sensitive to Cisplatin when treated at the larval L4 stage. At this stage, the worm's gonad, which contains the proliferative mitotic germ cells that differentiate to become sperm/oocytes, has formed, and an approximately four-fold amplification in total germ cell numbers occurs during maturation into gravid adult stages [14]. We exposed L4 animals for 3 hours to various doses of Cisplatin and subsequently determined the effect of this treatment on next generation's embryonic development and on the size of the brood (schematically illustrated in Figure 1A). The number of eggs that were

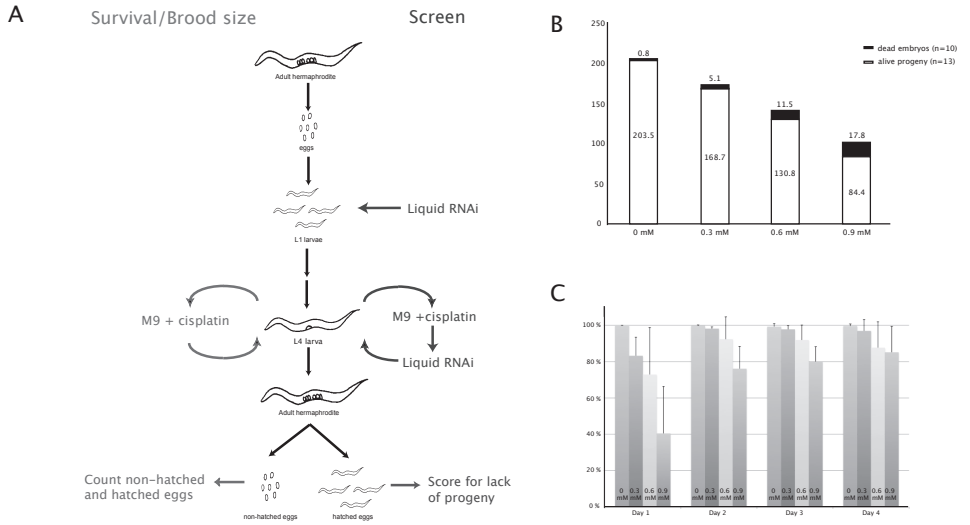


Figure 1 Cisplatin-induced toxicity effect on *C. elegans* progeny formation. A) Experimental design of the Cisplatin sensitivity assay on a small number of individual worms (right) or on small cultures in a high-throughput 96-well format (left). For small-scaled tests, animals were grown on OP50-seeded NGM plates and at L4 stage transferred to M9 buffer and treated with Cisplatin. After treatment, animals were placed back on fresh plates and the live versus dead progeny were counted. For the genome-wide RNAi analysis, we grew staged L1 larvae in liquid RNAi cultures until they reached the L4 stage. The animals were then transferred to M9 buffer-containing plates, treated with Cisplatin and again transferred to fresh liquid cultures. Four days after treatment, plates were scored for Cisplatin-dependent reduction of progeny count. B) Average broodsize (n=10) of animals that were treated with the indicated dose of Cisplatin; alive progeny in white, arrested/dead embryos in black. C) Embryonic survival (as a percentage of the total brood) in different time windows after Cisplatin exposure.

deposited as well as their fitness was monitored for five days after treatment. Untreated L4 animals reach maturity in approximately 18 hours and then lay all of their eggs in the subsequent 3 days. Nearly all (99%) of these embryos develop normally and hatch as L1 larvae (Figure 1B). We observed a dose-dependent decrease in brood size and viability of eggs that was most prominent in the first day after treatment (here, numbers are small) but remained detectable on all subsequent days of culturing: fewer eggs were laid and of those, fewer developed into healthy animals (Figures 1B and C). The reduction in brood size was thus not because of a delayed gametogenesis: after 5 days, all worms, both treated and mock-treated, ceased egg laying.

We next asked whether the Cisplatin-induced increase in lethality and reduction in brood size is accompanied by DNA-damage responses in the germ line cells of exposed hermaphrodites. In *C. elegans*, two DNA-damage responses have been described for germ line cells [15]: cell cycle arrest and apoptosis. The most distal end of the germ line consists of mitotic stem cells, which upon induction of DNA damage (by e.g. ionizing radiation) undergo cell cycle arrest; nuclei temporarily fail to proliferate but grow larger in size, and as a consequence, there are fewer but larger cells occupying a defined volume of the mitotic zone. Another outcome of DNA damage is apoptosis, which can only be seen when meiotic germ nuclei are in late pachytene and pass the bend of the gonada late stage of meiotic nuclei just before they develop into oocytes where the DNA arrests at the diakinesis stage.

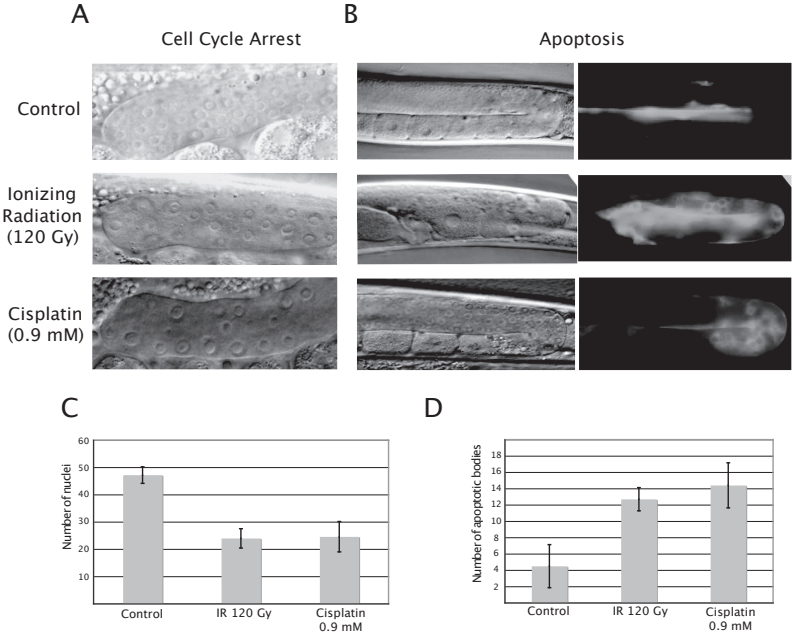


Figure 2 Cisplatin induces DNA damage responses in the *C. elegans* germ line. A) Cell cycle arrest of mitotic cells located in the distal region of the *C.elegans* germ lines. The middle panel displays enlarged nuclei in germ lines of worm treated with ionizing radiation or Cisplatin. B) Increased apoptosis in the pachythene region in germ lines of wild-type animals after treatment with Cisplatin and ionizing radiation is visualized by CED-1::GFP. C) Quantification of the Cisplatin-induced cell cycle arrest. D) Quantification of the Cisplatin-induced increase in apoptosis.

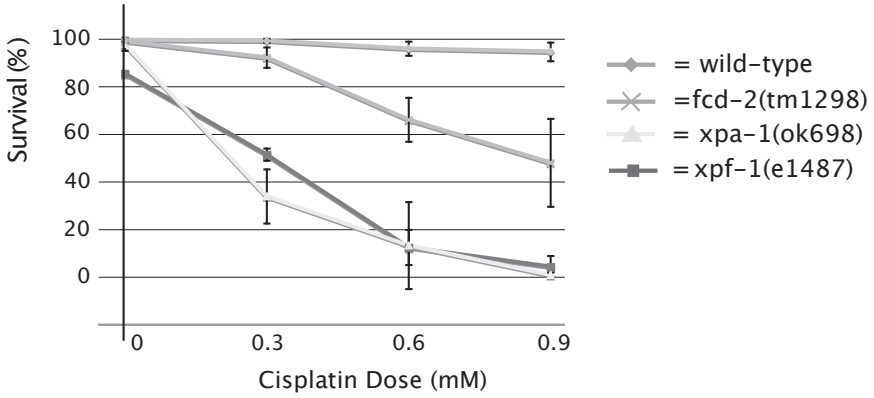
We observe a clear cell cycle arrest in the mitotic zone upon treatment with 0.9 mM Cisplatin (Figures 2A and C): the number of nuclei counted in a fixed volume is 24.6 ± 5.5, which is about 50% less than that of non-treated control animals (47.2 ± 3.0 nuclei). This reduction is comparable to the effect observed when animals are exposed to 120 Gy of ionizing radiation (24 ± 3.5). The checkpoint response is temporal: no difference between treated or mock-treated germ lines are noticeable 42 hours after treatment, which corresponds to 24 hours after observing a clear cell cycle arrest phenotype (data not shown). There are different ways to score for apoptotic cells in the *C. elegans* germ line, including cell and nuclear morphology, staining with acridine orange or performing a TUNEL assay. Here, we used cellular localization of the apoptotic marker CED-1::GFP fusion [16, 17]: apoptotic cells are engulfed by surrounding sheat cells during which CED-1 is expressed and localized to the membrane. GFP halos are thus thought to mark apoptotic cells (Figure 2B). Using this marker, we counted 14.4 ± 2.8 apoptotic corpses 24 hours after treatment (Figure 2D), while in mock-treated control animals 4.5 ± 2.6 apoptotic corpses per gonad were observed. Also, in this case, the response is comparable to 120 Gy of ionizing radiation (12.7 ± 1.4).

Together, these data show that treatment with Cisplatin induces conventional DNA-damage responses in the *C. elegans* germ line and has detrimental consequences for progeny formation, in both number and fitness.

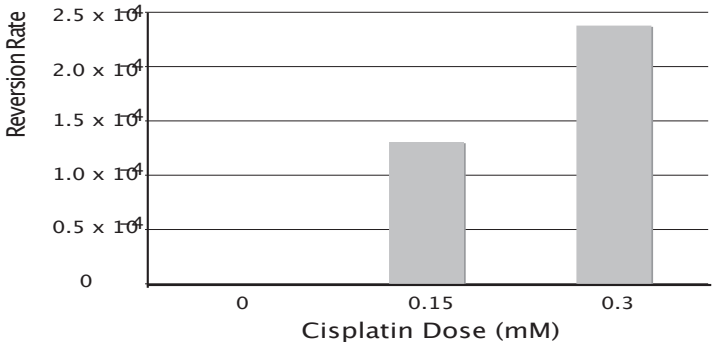
DNA damage induction and mutation profile of Cisplatin treatment

The notion that Cisplatin induces cell cycle arrest and apoptosis is in perfect agreement with the assumptions that treatment results in DNA damage and that the observed embryonic lethality is the consequence of a failure to properly deal with these replication-blocking DNA adducts. We thus tested animals carrying mutations in *C. elegans* orthologs of genes known to be involved in removing ICLs. L4 larvae were incubated for 3 hrs with Cisplatin in M9 buffer and subsequently transferred to agar plates to allow maturation and egg deposition. 42 hours after the treatment the parental animals were removed and the progeny (eggs and larvae) were counted immediately and after a 24-hour developmental window. Under non-challenged conditions, *C. elegans* eggs hatch 16-18 hours hours after they are laid. Wild-type animals are fairly resistant to Cisplatin-induced embryonic lethality up to a concentration of 0.9 mM, in contrast to animals that are defective in nucleotide excision repair or ICL repair. Embryogenesis is severely compromised in *xpa-1* and *xpf-1* mutant animals at this dose, while animals with a defect in the worm ortholog of FANCD2 (FCD-2) are intermediately sensitive (Figure 3A). Here, it should be noted that we observed substantial variations if experiments were performed on different days or with different batches of Cisplatin. Therefore, throughout this study, we only compared data points that are derived from experiments that were performed in parallel. In their trends, all experiments are entirely consistent.

A



B



C

Mutated gene	Type of mutation	Frequency	Sequence
sup-10	5 nt deletion	3	CAAAGTGTCCTCAG[CCATC]TTGGTCAATAGTAGCCAG
	4 nt deletion	1	CCATTAGATGGAT[TGAC]GCATGTATTGCTA
	7 nt deletion	1	TTTTGCAGAATAG[CCCAAGA]TCACGGGATCCAACC
	22 nt deletion, 12 nt insertion	1	AAATA[AGCGCA-//-TATTAG]{TATTTGTATTAA}TTGGT
	deletion of exons 5 and 6	2	not determined[1]
	deletion of gene	8	not determined[1]
sup-9	6 nt deletion	1	TTAACGGGCTCACC[GAGCCC]TCTTACCTGCAGAAT
	286 nt deletion	1	GATTGCAGCTTCT[TCTTA--GCCAA]AGTTATACCCACC
	120 nt deletion	1	CATCAGCAGTAGGAT[CACCA--TTCGGT]GTTCATAGTCAA
	deletion of last 6 exons	1	not determined[1]
unc-93	Single-base-pair substitution	1	GTTCCCTAGTCTTCG(G>A)GCCACTGATCAAGCTT
	3 nt deletion	1	TTCGGAATTTATTTC[AGC]TCCTACTTATTCCGTGG
	7 nt deletion	1	GGCCTAGGAAAAGAG[CAGCTGG]AAGAAGAAGCCAG
	8 nt deletion	1	TCAATATTCTGTGGA[ATCGCGGC]TTCATGTATCTGGGG
	142 bp deletion	1	GAATTTAGCGTATT[CTTAG--AATTT]TAAGCGATTCAAT
sup-18	unknown[2]	5	

[1] Scored based on failure to detect PCR product for mentioned regions
[2] Sequence of *sup-18* is not known, scored based on phenotype

Figure 3 DNA damage phenotypes that results from Cisplatin treatment. A) Cisplatin survival curves of animals with the indicated genotype. B) Cisplatin-dose dependent increase in reversion rates of the *unc-93(e1500)* rubber-band phenotype. C) Molecular nature of the mutations caused by Cisplatin.

The observation that DNA repair defects leads to increased levels of lethality indicates that this phenotype can be attributed to DNA adducts and not the result of other cytotoxic consequences of drug treatment. We do not know whether intra- or interstrand DNA crosslinks are the most toxic lesions in our experimental setup. The notion that NER components XPA-1 and XPF-1 contribute more strongly to Cisplatin resistance than the ICL-protein FCD-2 may suggest a more prominent role for intrastrand crosslinked NER substrates; however, NER proteins, in particular XPF-1, have also been implicated in the repair of interstrand crosslinks.

Next, we determined the mutagenic consequences of Cisplatin-induced DNA adducts in the worm. To determine the nature of the mutations induced by Cisplatin, we used the *unc-93(e1500)* reversion assay. Animals homozygous for the semi-dominant allele *unc-93(e1500)* show uncoordinated movement ("rubber band" phenotype) and have an egg-laying deficiency [11]. This phenotype can be suppressed by loss-of-function mutations in *unc-93* itself or by mutations in the genes *sup-9*, *sup-10*, *sup-11* and *sup-18*. These genes can serve as a mutational target because phenotypic revertants are easily scored: they move wild-type and are egg-laying and can thus be clearly distinguished from the *unc-93(e1500)* phenotype. By sequencing the aforementioned genes in these revertants, the nature of the Cisplatin-induced mutations can be determined.

We treated populations of *unc-93(e1500)* L4 larvae with 0.15 and 0.3 mM Cisplatin, transferred aliquots each containing 20 animals to fresh culture dishes and inspected the F2/F3 progeny for Unc suppression. Because *C. elegans* is cultured as hermaphrodites that produce sperm and oocytes, germline mutations that are induced in the gametes of exposed P0 animals will homozygose in the F2 generation. In the mock-treated control, we found not a single revertant on 80 culture plates. In contract, plates seeded with Cisplatin-treated animals had many: 0.15 mM Cisplatin treatment resulted in 14 out of 82 positive plates, whereas 0.3 mM Cisplatin treatment resulted in 22 out of 82 plates, corresponding to reversion rates of 1.3×10^{-4} and 2.4×10^{-4}, respectively (Figure 3B).

We found five of the revertants linked to *sup-18*. In *sup-18*-mediated suppression of unc-93(e1500), only the egg-laying phenotype is reversed, while the uncoordinated movement is hardly affected [18]. Because *sup-18* has not yet been cloned, the molecular nature of these mutations could not be identified. We determined the molecular lesion in 25 revertants, for which we performed PCR and sequence analysis of the *unc-93*, *sup-9* and *sup-10* loci. Of these, 5 carried a mutated version of *unc-93*, 4 had a mutation in *sup-9* and 16 either had small frameshifting deletions in the *sup-10* ORF or had large deletions within the genomic region that contains *sup-10*. The molecular nature of all mutations are listed in Figure 3C. We found that Cisplatin predominantly induces deletions: only 1 out 25 mutations is a single-base-pair substitution. Nine of the deletions are small (in the 3-7 nt range), 3 contain 100-300 nt sized deletions and for 11 revertants, larger deletions occurred. In 8 cases, we failed to detect the *sup-10* gene completely by PCR. The exact size of the large deletions was

not determined. The predominance of deletions in the Cisplatin-induced mutation pool is remarkable, since in both *E. coli* and *S. cerevisiae*, an overrepresentation of single-base-pair substitutions has been found [19-21]. Although the mutation detection assays used in the *E. coli* studies can only detect point mutations in one case [19] and point mutations and frameshifts in the other [20], the yeast system was capable of detecting mutations of any kind, and this led to a spectrum of predominantly point mutations [21]. This is in complete contrast to our data in *C. elegans*, but a molecular understanding of this discrepancy is currently unknown. In both systems, but in contrast to *e.g.* chicken DT40 cells, NER appears to be a major protection mechanism. *xpa-1* mutants are significantly more sensitive to Cisplatin than mutant animals that have a mutated version of the ICL gene *fcd-2*. Interestingly, while *xpf-1* has been implicated in NER and ICL repair, loss of the NER-exclusive factor *xpa-1* leads to a sensitivity similar to that caused by loss of *xpf-1*. Whether NER plays a profound role in mammalian systems is unclear because of conflicting data: Hamster cells carrying mutations in NER genes XPB, XPD, XPG and CSB display only mild sensitivity to Cisplatin, but have compromised ICL removal [4]. XPF-deficient hamster cells, however, are extremely sensitive to Cisplatin. Human fibroblasts lacking functional NER factor XPA were shown to be more defective in the removal of ICL adducts than Fanconi Anemia factor FAA deficient fibroblasts [22].

A genome-wide RNAi screen identifies genes that influence Cisplatin sensitivity

Next, we performed a genome-wide RNAi screen to identify genes that modulate the cellular response to Cisplatin treatment. Of the phenotypes we described (cell cycle arrest, apoptosis, brood size reduction, embryonic lethality and mutation induction), we considered only the detrimental effects on progeny formation amenable to large-scale analysis. This is supported by the notion that genetic alleles of *xpa-1* and *xpf-1* greatly sensitize animals to the treatment. We used the in-liquid-screening method previously developed in our laboratory [12, 23]. This method allows genome-wide screening of RNAi libraries in 96-well format. A schematic representation is outlined in Figure 1A: we added synchronized L1 larvae to liquid RNAi cultures and allowed these to grow under RNAi conditions until the vast majority of the cultures reached the L4 stage. We then transferred animals to 96-well plates containing M9 buffer. This transfer is necessary to avoid fluctuations in the effective dose caused by the fact that the bacteria in the liquid cultures take up Cisplatin in a density-dependent manner. After transfer, the worms were incubated with Cisplatin and then transferred back to fresh liquid RNAi cultures. These cultures were allowed to grow for four days, after which we scored for the presence or absence of progeny in the Cisplatin-treated cultures. We subsequently inspected the mock-treated controls that were assayed in parallel. Clones that induced a Cisplatin-dependent reduction of brood size were repeated twice. This resulted in 51 genes that modulated the worm's sensitivity to Cisplatin when knocked down (Table

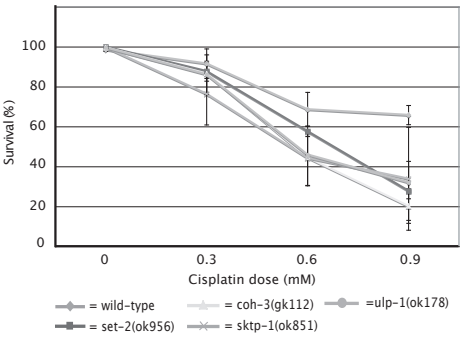
Table 1 Genes identified in a genome-wide RNAi screen for increased Cisplatin sensitivity.

seq. name	gene	HOMOLOGY Hs	Dm	Sc	Hs	Dm	Sc	Remarks
UBIQUITIN RELATED								
F10G7.10		+	+	+	NP_742067.2	CG1531-PB	UBR2	
T10F2.3	*ulp-1*	+	+	+	SENP1	CG11023-PA	ULP1	Protease, Ulp1 family [KOG0778]
F48E8.7	*skpt-1*	+	+		SKP2	CG9772-PA		SCF ubiquitin ligase, Skp2 component [KOG2120]
F58E1.8/9	fbxb-18/19							
CHROMATIN RELATED								
C47D12.1	*trr-1*	+	+	+	TRAPP	Nipped-A-PA	TRA1	Histone acetyltransferase SAGA, TRRAP/TRA1 component, PI-3 kinase superfamily [KOG0889]
C26E6.9a	*set-2*	+	+	+	Q9UPS6	trx_PE	SET1	Histone H3 (Lys4) methyltransferase complex, subunit SET1 and related methyltransferases [KOG1080]
D2096.8		+	+	+	NAP1L1	Nap1-PA	NAP1	Nucleosome assembly protein NAP-1 [KOG1507]
Y2H9A.1	*mes-4*	+	+	+	NP_060248.2	Mes-4-PA	SET2	Transcription factor NSD1 and related SET domain proteins [KOG1081]
B0035.11	*coh-3*	+	+	+	LEO1	Atu-PA	LEO1	Uncharacterized conserved protein [KOG2428]
F08H9.1								rad21/rec8-like family of cohesin proteins expressed in germline
CHECKPOINT KINASES								
R05G6.10	*kin-34*	+	+	+	RASGEF1B	CG7369-PA	CDC25	Predicted guanine nucleotide exchange factor [KOG3541]
R02C2.2		+	+	+	CHEK1	grp-PB	CHK1	
OTHER KINASES								
W03A5.1		+	+	+	FRK	Btk29A-PA	STE20	Fibroblast/platelet-derived growth factor receptor and related receptor tyrosine kinases [KOG0200]
R07H5.8		+	+	+	P55263-2	CG11255-PA	ADO1	Possible pfkB family carbohydrate kinase [KOG2854]
RECEPTORS								
F33H1.5	*srd-1*							Chemoreceptor/7TM receptor [LSE0147]
F36G9.8	*sri-9*							Predicted olfactory G-protein coupled receptor [LSE0262]
B0240.3	*daf-11*	+	+		NPR2	CG31183-PA		Natriuretic peptide receptor, guanylate cyclase [KOG1023]
D1069.4		+	+		GP139_HUMAN	DmsR-1PA		Unnamed protein [TWOG0217]
F40A3.7		+	+		BLR1	CG16752-PA		Unnamed protein [LSE0693]
T03D3.11/6	*srj-44/45*							
CYTOSKELETON COMPONENTS								
W03F11.6a	*afd-1*	+	+	+	AF-6	cno-PA	MUC1	Actin filament-binding protein Afadin [KOG1892]
F58A4.8	*tbg-1*	+	+	+	TUBG1	gammaTub37C-PA	TUB4	Gamma tubulin [KOG1374]
STRESS RESPONSE								
W03A5.7	*dnj-24*	+	+	+	DnaJ (Fragment)	mrj-PD	YDJ1	
RNA PROCESSING & TRAFFICKING								
F56D12.5	*vig-1*	+	+	+	PAI-RBP1	vig-PB	GAR1	Predicted RNA-binding protein
K07H8.10		+	+	+	NCL	CG13214-PA	MDN1	
ZK596.1		+	+	+	KRT10	GCR(ich)-PA	NOP1	Predicted alpha-helical protein [LSE0521]
F45F2.10			+	+		ear-PA	SRP40	
F33H1.3		+	+		WBP11	CG2685-PA		Uncharacterized conserved low complexity protein [KOG4672]
METABOLIC ENZYMES								
T04F3.1		+	+	+	NP_115981.1	Stm-Mck-PD	USO1	
T10D2.11	*ugt-41*	+	+	+	UGT1A9	Ugt86Dc-PA		UDP-glucuronosyl and UDP-glucosyl transferase [KOG1192]
T05E11.6		+	+	+	PIGK	CG4406-PA	GPI8	GPI-anchor transamidase [KOG1349]
F01G4.2	*ard-1*	+	+	+	HADH2	scu-PA	YDL114W	Short-chain alcohol dehydrogenase/3-hydroxyacyl-CoA dehydrogenase [KOG1199]
OTHER								
Y105E8A.8		+	+		WDR79	CG9226-PA		
Y39G10AR.7								Unnamed protein [LSE4141]
T20D3.9		+	+		FA38A	CG8486-PB		Uncharacterized conserved protein [KOG1893]
C40D2.1	*math-19*							Uncharacterized protein, contains BTB/POZ domain [LSE0150]
T16A1.1	*math-42*							Uncharacterized protein, contains BTB/POZ domain [LSE0150]
F56C9.6								
F13A2.5								
F07C3.2								
C01G5.8		+			KIAA1018			Uncharacterized conserved protein [KOG2143]
Y37H9A.1								
F30A10.2								
Y32B12A.4								Unnamed protein [LSE0747]
F59A1.8	*fbxa-129*							
R10E8.2								Unnamed protein [LSE0736]
ZK1193.4								
F56C3.4								
Y41D4B.13	*ced-2*	+	+	+	CRKL	Crk-PC	LSB3	Crk family adapters [KOG4792]
C26F1.3		+	+	+	protoglycan precursor	tnc-PA	SIR4	
T27A10.6								

1). Several gene-ontology classes were found to be overrepresented. We identified proteins that are thought to act in chromatin remodeling, signal transduction and posttranslational protein modification such as ubiquitinylation.

To validate our RNAi approach we tested five genes of which a mutant allele exists: *coh-3*, *set-2, ulp-1*, *sktp-1* and R05G6.10. For mutant strains having mutations in *set-2*, *coh-3*, *ulp-1* and *sktp-1*, we found an increased sensitivity to Cisplatin (Figure 4). For R05G6.10, the mutant strain was not different from wild type, suggesting the presence of some false positives hits in our data set. Next, we tested the specificity of these gene products in their ability to protect cells against genomic inflictions. All of the four aforementioned validated hits also conferred increased sensitivity towards ionizing radiation. Sensitivity to another DNA crosslinker, diepoxybutane [24], was found for *set-2*, *ulp-1* and *sktp-1*.

A



B

Gene	Allele	CisPt	DEB	IR
R05G6.10	ok1159			
set-2	ok952	+	+	+
ulp-1	ok1786	+	+	+
sktp-1	ok851	+	+	+
coh-3	gk112	+		+

Figure 4 Validated modulators of the cellular response to Cisplatin. A) Cisplatin survival curves of animals with the indicated genotype. B) Table depicting the sensitivity of five genetic knockouts to Cisplatin, diepoxybutane (DEB) and ionizing radiation (IR). These genes correspond to RNAi clones that were identified in the genome-wide RNAi screen. A '+ 'mark indicates an increased level of sensitivity.

Unexpectedly, we failed to pick up genes that are known to be involved in repair of Cisplatin-induced DNA adducts, such as the proteins acting in NER. We therefore specifically re-tested these RNAi clones in liquid and on agar plates, but found no enhanced sensitivity for cisplation for these clones, suggesting that these foods are not effective in knocking down their targets.

For many of these proteins, we could find links to the DNA damage response in the literature. The ubiquitin-related group consists of three genes that have all been implicated in DNA damage response or cell cycle regulation. ULP1, the yeast homologue of *ulp-1*, is involved in SUMO1 and smt3 cleavage and was shown to play an essential role in the G2/M phase of the cell cycle [25]. In mice, loss of the ubiquitin E3 ligase UBR2 (homologue of F10G7.10) leads to sensitivity to the DNA crosslinking agent mitomycin C, defective homologous recombination repair and chromosome fragility [26]. Finally, degradation of p21 in response to UV damage was shown to be dependent on the ubiquitin/skp2 pathway in human cells [27]. Skp2 is the human homologue of the *sktp-1* gene identified here. Of the

four kinases identified in this screen, two function in checkpoint regulation. These are the yeast homologues of R05G10.6 and *kin-34*, Cdc25 and Chk1 respectively, which act in one pathway to regulate Cdc2 in response to DNA damage, thereby avoiding mitotic entry [28]. Among the chromatin-related genes, we found genes that were previously linked to DNA damage response pathways. The *trr-1* homologue TRA1 is related to ATM and DNA-dependent kinase and is part of two histone acetyltransferase complexes, SAGA and NuA4 [29, 30]. The catalytic subunit of Nu4A is ESA1[29], which is essential for cell cycle progression [31]. SET1, the homologue of *set-2*, has been linked to both DNA repair and telomere maintenance processes [32, 33].

For other classes of gene products (e.g. RNA processing and trafficking; receptors), a functional link to Cisplatin action is more difficult to establish and further investigation will have to reveal whether these proteins act in a directed fashion on DNA adducts or whether the physiology of the drug-exposed cells are changed such that they become more sensitive to this crosslinker. DNA adduct-independent roles of Cisplatin may be exemplified by the identification of two cytoskeleton components (*afd-1* and *tbg-1*) and one cytoskeleton-related enzyme Uso1 (yeast homologue of T04F3.1), which is involved in intracellular protein transport [34]. A role for Cisplatin in the disruption of microtubules has been described [35] and in our knockdown conditions, such an effect may be enhanced.

Acknowledgements

We thank Ron Romeijn and Albert Pastink of the Department of Toxicogenetics at the Leiden University Medical Center in Leiden, The Netherlands for supplying the Cisplatin, Nigel O'Neil for sharing the molecular nature of the *xpf-1*(e1487) mutation and the CGC for providing worm strains.

References

1. Einhorn, L.H. (1990). Treatment of testicular cancer: a new and improved model. J Clin Oncol *8*, 1777-1781.
2. Einhorn, L.H., and Donohue, J.P. (1998). Advanced testicular cancer: update for urologists. J Urol *160*, 1964-1969.
3. Fichtinger-Schepman, A.M., van der Veer, J.L., den Hartog, J.H., Lohman, P.H., and Reedijk, J. (1985). Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. Biochemistry *24*, 707-713.
4. De Silva, I.U., McHugh, P.J., Clingen, P.H., and Hartley, J.A. (2002). Defects in interstrand cross-link uncoupling do not account for the extreme sensitivity of ERCC1 and XPF cells to Cisplatin. Nucleic Acids Res *30*, 3848-3856.
5. Vrana, O., Boudny, V., and Brabec, V. (1996). Superhelical torsion controls DNA interstrand cross-linking by antitumor cis-diamminedichloroplatinum(II). Nucleic Acids Res *24*, 3918-3925.
6. Niedernhofer, L.J., Lalai, A.S., and Hoeijmakers, J.H. (2005). Fanconi anemia (cross)linked to DNA repair. Cell *123*, 1191-1198.
7. Gartner, A., MacQueen, A.J., and Villeneuve, A.M. (2004). Methods for analyzing checkpoint responses in Caenorhabditis elegans. Methods Mol Biol *280*, 257-274.
8. O'Neil, N., and Rose, A. (2006). DNA repair. WormBook, 1-12.
9. Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. Nature *408*, 325-330.

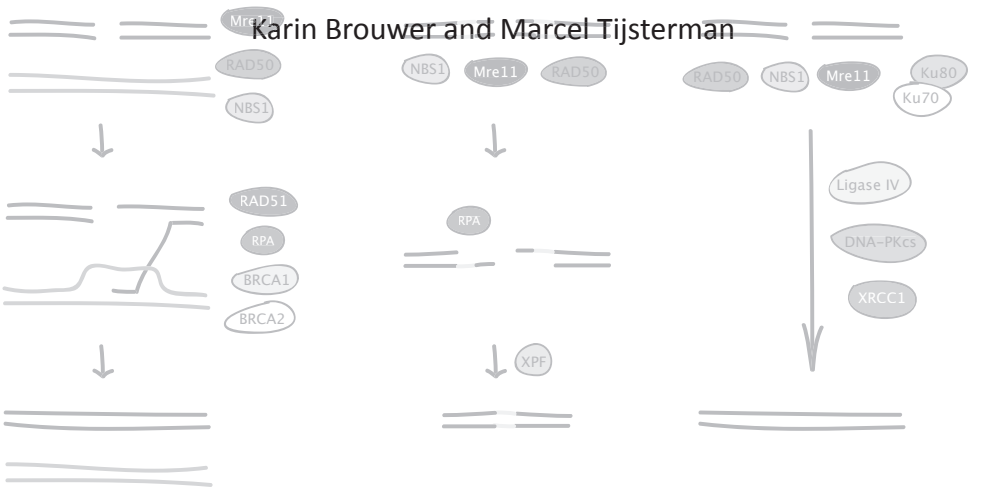
10. Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., and Ahringer, J. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231-237.
11. Greenwald, I.S., and Horvitz, H.R. (1980). unc-93(e1500): A behavioral mutant of Caenorhabditis elegans that defines a gene with a wild-type null phenotype. Genetics *96*, 147-164.
12. van Haaften, G., Vastenhouw, N.L., Nollen, E.A., Plasterk, R.H., and Tijsterman, M. (2004). Gene interactions in the DNA damage-response pathway identified by genome-wide RNA-interference analysis of synthetic lethality. Proc Natl Acad Sci U S A *101*, 12992-12996.
13. van Haaften, G., Romeijn, R., Pothof, J., Koole, W., Mullenders, L.H., Pastink, A., Plasterk, R.H., and Tijsterman, M. (2006). Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. Curr Biol *16*, 1344-1350.
14. Wood, W.B. (1988). Introduction to C. elegans Biology. In The Nematode Caenorhabditis elegans, W.B. Wood, ed. (Cold Spring Harbor: Cold Spring Harbor Press).
15. Gartner, A., Milstein, S., Ahmed, S., Hodgkin, J., and 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.
16. Bhalla, N., and Dernburg, A.F. (2005). A conserved checkpoint monitors meiotic chromosome synapsis in Caenorhabditis elegans. Science *310*, 1683-1686.
17. Schumacher, B., Schertel, C., Wittenburg, N., Tuck, S., Mitani, S., Gartner, A., Conradt, B., and Shaham, S. (2005). C. elegans ced-13 can promote apoptosis and is induced in response to DNA damage. Cell Death Differ *12*, 153-161.
18. Greenwald, I., and Horvitz, H.R. (1986). A visible allele of the muscle gene sup-10X of C. elegans. Genetics *113*, 63-72.
19. Brouwer, J., van de Putte, P., Fichtinger-Schepman, A.M., and Reedijk, J. (1981). Base-pair substitution hotspots in GAG and GCG nucleotide sequences in Escherichia coli K-12 induced by cis-diamminedichloroplatinum (II). Proc Natl Acad Sci U S A *78*, 7010-7014.
20. Burnouf, D., Duane, M., and Fuchs, R.P. (1987). Spectrum of Cisplatin-induced mutations in Escherichia coli. Proc Natl Acad Sci U S A *84*, 3758-3762.
21. Mis, J.R., and Kunz, B.A. (1990). Analysis of mutations induced in the SUP4-o gene of Saccharomyces cerevisiae by cis-diammine dichloroplatinum(II). Carcinogenesis *11*, 633-638.
22. Zhen, W., Evans, M.K., Haggerty, C.M., and Bohr, V.A. (1993). Deficient gene specific repair of Cisplatin-induced lesions in Xeroderma pigmentosum and Fanconi's anemia cell lines. Carcinogenesis *14*, 919-924.
23. Nollen, E.A., Garcia, S.M., van Haaften, G., Kim, S., Chavez, A., Morimoto, R.I., and Plasterk, R.H. (2004). Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation. Proc Natl Acad Sci U S A *101*, 6403-6408.
24. Park, S., Anderson, C., Loeber, R., Seetharaman, M., Jones, R., and Tretyakova, N. (2005). Interstrand and Intrastrand DNA-DNA Cross-linking by 1,2,3,4-Diepoxybutane: Role of Stereochemistry. J Am Chem Society *127*, 14355-14365.
25. Li, S.J., and Hochstrasser, M. (1999). A new protease required for cell-cycle progression in yeast. Nature *398*, 246-251.
26. Ouyang, Y., Kwon, Y.T., An, J.Y., Eller, D., Tsai, S.C., Diaz-Perez, S., Troke, J.J., Teitell, M.A., and Marahrens, Y. (2006). Loss of Ubr2, an E3 ubiquitin ligase, leads to chromosome fragility and impaired homologous recombinational repair. Mutat Res *596*, 64-75.
27. Bendjennat, M., Boulaire, J., Jascur, T., Brickner, H., Barbier, V., Sarasin, A., Fotedar, A., and Fotedar, R. (2003). UV irradiation triggers ubiquitin-dependent degradation of p21(WAF1) to promote DNA repair. Cell *114*, 599-610.
28. Walworth, N.C. (2001). DNA damage: Chk1 and Cdc25, more than meets the eye. Curr Opin Genet Dev *11*, 78-82.
29. Allard, S., Utley, R.T., Savard, J., Clarke, A., Grant, P., Brandl, C.J., Pillus, L., Workman, J.L., and Cote, J. (1999). NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. Embo J *18*, 5108-5119.
30. Grant, P.A., Schieltz, D., Pray-Grant, M.G., Yates, J.R., 3rd, and Workman, J.L. (1998). The ATM-related cofactor Tra1 is a component of the purified SAGA complex. Mol Cell *2*, 863-867.
31. Clarke, A.S., Lowell, J.E., Jacobson, S.J., and Pillus, L. (1999). Esa1p is an essential histone acetyltransferase required for cell cycle progression. Mol Cell Biol *19*, 2515-2526.
32. Corda, Y., Schramke, V., Longhese, M.P., Smokvina, T., Paciotti, V., Brevet, V., Gilson, E., and Geli, V. (1999). Interaction between Set1p and checkpoint protein Mec3p in DNA repair and telomere functions. Nat Genet *21*, 204-208.
33. Schramke, V., Neecke, H., Brevet, V., Corda, Y., Lucchini, G., Longhese, M.P., Gilson, E., and Geli, V. (2001). The set1Delta mutation unveils a novel signaling pathway relayed by the Rad53-dependent hyperphosphorylation of replication protein A that leads to transcriptional activation of repair genes. Genes Dev *15*, 1845-1858.
34. Nakajima, H., Hirata, A., Ogawa, Y., Yonehara, T., Yoda, K., and Yamasaki, M. (1991). A cytoskeleton-related gene, uso1, is required for intracellular protein transport in Saccharomyces cerevisiae. J Cell Biol *113*, 245-260.

35. Tulub, A.A., and Stefanov, V.E. (2001). Cisplatin stops tubulin assembly into microtubules. A new insight into the mechanism of antitumor activity of platinum complexes. Int J Biol Macromol *28*, 191-198.

Chapter 6

Non-classical end-joining repair of DNA double strand breaks in *C. elegans* germ cells

Karin Brouwer and Marcel Tijsterman



Abstract

The DNA within our cells is constantly being damaged by both environmental and endogenous agents; of the many forms of DNA damage, the DNA double strand break (DSB) is considered the most dangerous**.** In *C. elegans* germ cells, repair of DSBs occurs via homologous recombination repair in which the sister or homologous chromosome serves as a template. Simple end-joining (EJ) of broken chromosomes also takes place but this mechanism appears to be mechanistically different from canonical non-homologous end-joining (NHEJ) that dominates repair in somatic cells. Here, we studied the genetic requirements of EJ in *C. elegans* germ nuclei, and determined the influence of the break's sequence context. We show that the preferred pathway of EJ repair is independent of NHEJ component LIG-4 and that flanking microhomology is a very strong determinant in the outcome of repair. Interestingly, this homology-driven repair does not depend on the XPF/ERCC1 endonuclease complex, which has been implicated in single-strand annealing. Based on careful inspection of joined molecules, including duplicated flanking sequences, we propose that EJ repair is predominantly dictated by primer-template extension.

Introduction

DNA double-strand breaks (DSBs) are extremely genotoxic to cells and several pathways have evolved to repair or join broken molecules [1]. Three main repair mechanisms have been described (Figure 1A): homologous recombination repair (HR), single-strand annealing (SSA), and non-homologous end-joining (NHEJ) [2]. The pathway of choice largely depends on cell type and stage of the cell cycle, *e.g.* in yeast HR is the predominant form of DSB repair, whereas NHEJ is the route of choice in somatic human cells. In *C. elegans*, homologues of genes functioning in each pathway have been identified [3]. Similar to the situation in mammals, LigaseIV-dependent NHEJ dominates repair in non-dividing somatic cells, while proliferating somatic cells and germline cells use HR to repair IR irradiation induced DSBs [4]. Nevertheless, non-HR repair does occur in germline tissues: transposon-induced breaks were successfully used to identify end-joining products in the *C. elegans* germ line [5]. Tc1 and Tc3 are the most active and best-characterized transposons in *C. elegans* [6]. These transposons move via a "cut-and-paste" mechanism (Figure 1B): transposase protein binds to the TIRs of the cognate transposon and catalyzes excision and subsequent reinsertion of the element into target DNA. This always occurs at a TA dinucleotide site, and results in duplication of the TA sequence. In germline tissue, the double-strand break (DSB) at the empty site is repaired by HR, which leads to restoration of the element at the original position. The net result is an increase in copy number: one new insertion at a new site while DSB repair generates a "new" copy at the old location. However, at 1% of the rate at which HR using the homologous chromosome is observed, the element is lost at the break site suggestive of another form of DSB repair. The observed "footprints" of DNA transposition suggest a

repair mechanism similar to NHEJ: the DNA flanks are joined which frequently is accompanied by the gain or loss of a few nucleotides [5, 7]. Identical results were recently obtained with a Tc1 family member derived from flies. Also in this case, *Mos1*-induced breaks in the *C. elegans* germline are mostly repaired via HR [8]. Surprisingly, however, and conflicting with studies using a plasmid-based assay [9], end-joining of these breaks does not depend on homologues of NHEJ genes Ku80 and ligase IV [10]. In addition, repair of germline Tc1 and *Mos1* breaks sporadically is accompanied by a small DNA insertion of sequences that appear to be duplicated from immediate flaking regions [5, 10]. These types of insertions, suggesting polymerase activity at the break site, were also identified in *Drosophila* [11, 12] and mammalian cells [13].

To study the specifics of end-joining mechanisms that function in the *C. elegans* germline, we analyzed the genetic requirements and the influence of surrounding sequences on Tc1 transposon-induced DSBs. We used two Tc1 alleles of the *unc-22* muscle gene: one that has significant microhomology directly flanking the break and one that has not. We determined the end-joining products in strains that have defects in particular DNA DSB repair pathways including NHEJ and SSA. We found that the presence of flanking microhomology strongly influenced the outcome of DSB repair. Surprisingly, we found that none of the tested defective backgrounds dramatically influenced the outcomes of DSB repair, suggesting that an alternative pathway exists to join broken chromosomes in the *C. elegans* germline. Careful inspection of the footprints suggests that this pathway is dependent on a DNA polymerase action that initiates at single nucleotide basepairing.

Materials and Methods

Strains and maintenance

General methods for culturing *C. elegans* were used [14]. The following strains were used in this study: *rde-3(ne298) unc-22(st136::Tc1)*, *lig-4(ok716) rde-3(ne298) unc-22(st136::Tc1)*, *xpf-1(e1487) rde-3(ne298) unc-22(st136::Tc1)*, *brc-1(tm1145) rde-3(ne298) unc-22(st136::Tc1)*, *rde-3(ne298) unc-22(st192::Tc1)*, *lig-4(ok716) rde-3(ne298) unc-22(st192::Tc1)*, *xpf-1(e1487) rde-3(ne298) unc-22(st192::Tc1)*, *brc-1(tm1145) rde-3(ne298) unc-22(st192::Tc1)*, *mut-7(pk204) unc-22(st136::Tc1)*, *F10G8.7(tm2073) mut-7(pk204) unc-22(st136::Tc1)*, *pme-1(ok988) mut-7(pk204) unc-22(st136::Tc1)*, *mut-7(pk204) unc-22(st192::Tc1)*, *F10G8.7(tm2073) mut-7(pk204) unc-22(st192::Tc1)*, *pme-1(ok988) mut-7(pk204) unc-22(st192::Tc1)*, *spo-11(ok79)IV; mIs11(myo-2::GFP)IV, mut-7(pk204) spo-11(ok79)IV; mIs11(myo-2::GFP)IV*.

Reversion assay to identify mutations caused by Tc1 transposition

Animals carrying either *unc-22(st136::Tc1)* and *unc-22(st192::Tc1)* were crossed with *rde-3(ne298)* or *mut-7(pk204)* males; cross progeny males were crossed to various genetic mutant backgrounds to establish strains that had defective DSB repair, were permissive for germline transposition (mut), and contained the *unc-22* Tc1 alleles (twitching phenotype). The choice of mut mutator largely depended on the chromosomal location of the DSB repair genes that were tested in this study. Animals were kept in culture by picking UNCs. To study EJ repair at the Tc1 site, animals were singled on 6 cm agar plates seeded with OP50 and the plates were grown until starvation. Next, WT moving animals were picked (one per 6 cm plate) and sequenced to identify the molecular nature of the event that restored *unc-22* function.

Results and Discussion

To analyze the products of DSB repair at known genomic locations we made use of chromosomal breaks that result from DNA transposition. We chose Tc1 transposons that were located in the *unc-22* locus. Mutant alleles of *unc-22* can be easily recognized under the microscope by their abnormal movement: animals defective in UNC-22 move uncoordinatedly (they twitch), are thin and are unable to hyper-contract. In mutator strains, *unc-22::Tc1* alleles can revert: the Tc1 element can hop out, leaving a non-complementary staggered cut with 2 nucleotide 3'overhangs (CA-3'OH). If such a break is repaired in an error prone fashion, loss of the element and functional *unc-22* expression can result, but only if the left and right flanks are joined such that the DNA encodes an in frame ORF (Figure 1A). Scoring for wild type animals in populations of twitchers can identify these events rather easily. However, it should be noted that only a subset of possible repair products are identified, because the selection procedure demands ORF restoration.

Micro-homology directs end joining in the C. elegans germline

We chose to study two alleles, *unc-22(st136::Tc1)* and *unc-22(st192::Tc1)*, because these have categorically different flanking sequences: whereas excision of Tc1 in st192 generates a break with just 2 nt of micro-homology immediately flanking the non-complementary staggered cut, excision of st136 creates a break in which the outer 6 nucleotides at the 3' end on both sides of the break are perfectly complementary (Figure 1B). Single twitching hermaphrodites of genotype *unc-22(st136::Tc1) rde-3* and *unc-22(st192::Tc1) rde-3* were placed on 6 cm culture dishes seeded with *E. coli*. After the populations consumed all the food (~10^4 -10^5 animals per plate), we inspected the plates for the presence of wildtype moving animals. Based on the number of plates that contained revertants we calculated the reversion frequency to be 5.7 x 10^{-4} for *st136* and 2.6 x 10^{-4} for *st192* (similar frequencies

were obtained in strains that use another mutator locus to initiate Tc1 jumping, Figure 1C). We suspect the higher frequency observed for *st136* not to be the result of a higher rate of DSB induction, but resulting from a higher ratio of repair products that restore the *unc-22* ORF. We thus determined the repair products by sequencing the *unc-22* gene in revertant animals. Strikingly, the observed spectra were completely different: whereas the footprints in st192 revertants show a plethora of end-joining products (11 different footprints in 23 animals), just one type of footprint was observed in 30 independently derived st136 revertants (Table 1). In 30/30 st136 revertants, repair led to the exact loss of one copy of the 6 nt microhomology that flanked the break. In contrast, we only found 6/23 cases for st192 that lost a 2 nt microhomologous sequence (as well as the sequence in between). This suggests that the degree of homology in the sequence surrounding the DSB is strongly influencing the manner in which the break is repaired. For clarity, we categorized the footprints derived

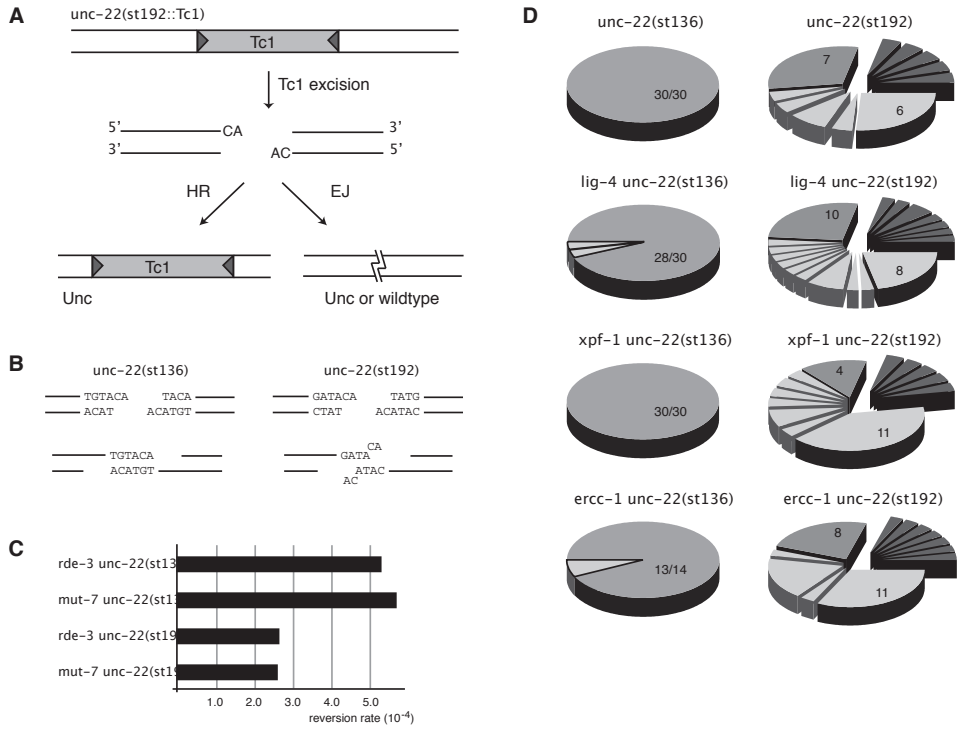


Figure 1 Germline repair of transposon mediated double-strand breaks. A) Mechanism of Tc1 mediated DSB repair. Transposases excise cognate transposons resulting in staggered cuts with 2 nucleotide 3' hydroxyl overhangs at donor sites. Breaks can be repaired in an error free way via HR or in an error prone way via EJ. Only in case the footprint restores the ORF of the *unc-22* host gene, animals will move wildtype and are picked up in our assay. B) The sequence context of the breaks that result of transposon excision at the *unc-22(st136)* or *unc-22(st192)* allele. Underneath, a possible repair intermediate is depicted that makes use of break-flanking microhomologies. C) The reversion rate of *unc-22(st136)* or *unc-22(st192)* alleles in two different mutators (*rde-3* and *mut-7*) that release transposon silencing. D) Footprint spectra for the indicated genotype, typically consisting of ~30 mutants. Color-coding for *st192* spectra: we subdivided the spectra in EJ products with (dark gray) or without duplicated flanking sequences; the second darkest gray and light gray with black border represent the two most abundant footprints.

from st192 into two subtypes of repair-joining products: in 18 out of 23 cases, small insertions or deletions have occurred, whereas in 5 out of 23 cases, repair involved duplication of sequences that flank the break site. The latter suggest a repair mechanism that is dependent on the action of a DNA polymerase.

End-joining in the C. elegans germ line does not depend on NHEJ component LIG-4

Previous work on *MosI* transposition demonstrated that repair of germ line breaks with 3-nt overhangs did not depend on canonical NHEJ components LigaseIV or Ku80: both the frequency of DSB repair and the spectrum of repair products in wild type worms were indistinguishable from animals that carried mutations in *cku-80* or *lig-4* [10]. However, a role for these components in germline DSB repair was suggested in another study that used Zinc Finger endonucleases to break the DNA *in vivo*. We determined the footprint of DSB repair at Tc1 breaks in *lig-4* deficient animals and found them to be identical to those in wildtype (Figure 1D). For st136, 28 out of 30 revertants used the micro-homologous sequence to repair the break, indicating that this route does not require functional LIG-4. Remarkably, and in agreement with the data on *MosI* transposition, we found that EJ products that are more characteristic of NHEJ also occur in a *lig-4* independent manner: 29 out of 36 revertants generated by DSBs at the st192 site contained small deletions or insertions (81% compared to 78% in wild type), whereas 7 had larger duplications of flanking sequences inserted in the break site (19% compared to 22%). Moreover, similar numbers were found for the most common footprints: 22% in *lig-4* versus 26% in wildtype had one typical 4 nt deletion; 28% in *lig-4* versus 30% in wildtype resulted from EJ in which no two nucleotides had been gained. Together with the notion that the reversion rate was identical for all these genetic backgrounds, this indeed suggests that EJ repair in the C. elegans germline is not canonical NHEJ. A candidate to regulate non-canonical EJ pathways is *pme-1*, the *C. elegans* homologue of PARP. PARP, likely together with Ligase III, has been shown to act in an end-joining pathway that serves as a backup for canonical NHEJ pathway in mammalian cell lines [15]. We have tested animals that carry a deletion in *pme-1(ok988)*, but found that the spectrum of footprints derived from 32 revertants of *st192* in a *pme-1* defective genetic background is identical to that of wildtype or *lig-4* mutant animals (Suppl. Table 1).

Micro-homology driven repair in C. elegans germ cells is independent of XPF/ERCC1

Next, we questioned which genetic components could be involved in the micro-homology driven repair pathway, which is completely dominating the repair of DSB in st136. One candidate is the XPF/ERCC1 complex as it has been implicated in DSB repair mechanisms in yeast and mammals that involve annealing of complementary resected DNA strands. However, we found no influence of knocking out *xpf-1* or it's binding partner *ERCC1/F10G8.7* on the rate

and the outcomes of DSB repair at the st136 locus (Figure 1D; supplemental Table 2). One could argue that a flap endonuclease is not required for processing SSA intermediates that are formed at the st136 breaksite since upon break induction, the outer nucleotides at both 3' are complimentary (see figure 1B). We therefore also determined the footprints at *st192* in *xpf-1* and F10G8.7 defective animals, because one of the most abundant repair products (a deletion of 4 nt which makes up 26% of the spectrum in repair proficient animals) may result from processing the break using a TA dinucleotide that flank the staggered cut (Figure 1B). We found that this particular end-product constitutes 39% and 34% in mutation spectra of *xpf-1* and F10G8.7 deficient animals, respectively. Together, this indicates that the XPF1/ERCC1 complex is not involved in error prone repair of Tc1-mediated DSBs in the *C. elegans* germline.

Table 1 Tc1 footprints of *unc-22(st192)* and *unc-22(st136)* derived alleles.

***unc-22(st192)* empty site**			
5'CTCCAATTTTGGGATACA			TATGTCGTTGAACGTTTTG
3'GAGGTTAAAACCCTAT			ACATACAGCAACTTGCAAAAC
***unc-22(st192)* footprints**			
6 CTCCAATTTTGGGATA			TGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGA			GTTGAACGTTTTG
2 CTCCAATTTTG			TATGTCGTTGAACGTTTTG
1 CTCCAATTTTTGGGAT			GTATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGG			GTATGTCGTTGAACGTTTTG
7 CTCCAATTTTGGGATACA			TGTATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGATACA	ATTTTGGGA		TGTATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGATACA	TGTCGTTGT		TGTATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGAT	CCAATTTTGGG		ATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGAT	TTTTGGGATTTTGG		TATGTCGTTGAACGTTTTG
1 CTCCAATTTTGGGATACA	ATTTTGTTTTTTTTGGGATGTATGATTTTGGGA		TGTATGTCGTTGAACGTTTTG
***unc-22(st136)* empty site**			
5' ATAAGGAAGGATGTACA			TACATTGAACTGGAAGCCTC
3' TATTCCTTCCTACAT			ACATGTAACTTGACCTTCGGAG
***unc-22(st136)* footprints**			
30 ATAAGGAAGGATGTACA			TTGAACTGGAAGCCTC

A molecular mechanism for EJ repair in C. elegans germ cells

Our data excludes NHEJ and SSA in EJ repair of DSB in the *C. elegans* germline. To get insight into a possible molecular mechanism and contributing factors we thus more carefully analyzed the repair products in an effort to deduce necessary enzymatic activities. We first concentrated on footprints that involved duplication of flanking sequences. These products require by definition two enzymatic activities: DNA polymerization (to copy in the flanking sequence) and DNA ligation (to join newly generated DNA to parental DNA). In repair proficient worms, we found 5 out of 23 revertants that had DNA insertions that could be traced back to patches of 8-11 nucleotides in the immediate vicinity of the breaks (Table 1). Three could be exclusively mapped to the left side, one to the right side and one insert combined three duplications coming from both left and right flanks. These types of inserts were seen in all tested mutant backgrounds (A list of all these type of footprints is compiled in Suppl. Table 3). We were able to construct a multistep pathway to explain all these outcomes based

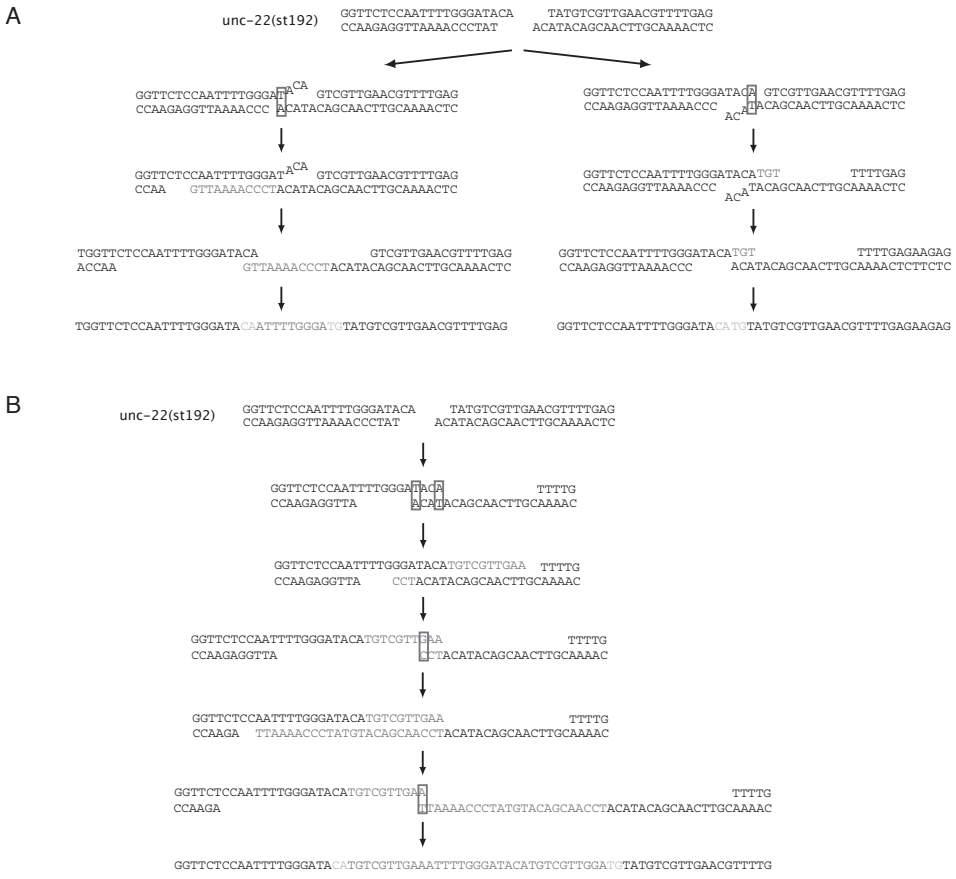


Figure 2 A model for end joining repair of DSB in germ cells. Schematic illustration of primer-template directed repair explaining the footprints observed after repair of a DSB located at *unc-22* (*st192::Tc1*). A) The duplication in the left panel was observed in 6/155 revertants; the footprint on the right side was found in 36/155 revertants. B) One example of a complex footprint that has duplicated flanking sequences from both flanks of the break. The mechanism is identical for all depicted examples: in a first step, the 3'overhanging nucleotide on one side of the break pairs with a complementary base on the other's side opposite strand. This allows *de novo* DNA synthesis of a small tract of variable length. At least for some of these intermediates structures, the invading/extended strand dissociates to re-anneal with other available bases. It is tempting to hypothesize that extension of the 3' end displaces the parental 5' strand because competition between new and old strand could explain why the new strand is a) of limited length, b) dissociates, and c) subsequently primarily pairs with bases that are at the outermost end of the template (as if other template bases or not available). Only in cases where both the 3' overhangs pair with each other at the outermost end, repair can be completed by DNA synthesis and ligation to parental DNA.

on very simple principles (drawn in detail in Figure 2). In a first step the 3'overhanging nucleotide on one side pairs with the first available complementary nucleotide on the opposite strand of the other flank. This then allows DNA polymerase action to extend the 3' end by incorporation of matching nucleotides. The footprints suggest that *de novo* DNA synthesis does not extend further than ~10 nucleotides (Suppl. Table 2). Then, this intermediate structure has two fates: a) the 3'end of the other end is also extended and newly generated DNA

is ligated to parental DNA or b) the extended strand dissociates resulting in a new 3' end that can repeat annealing and extension (See figure 2A, for one example, and supplemental Figure 1). We observed footprints that suggest 1- 4 of such cycles, and we hypothesize that repair can only be finished if the 3' overhanging ends of both strands simultaneously pair and can thus be extended. Figure 2B exemplifies how this multistep pathway perfectly explains a complex combinatorial footprint that we isolated in an *xpf-1* defective background. Several duplication-containing footprints suggest that 3-overhanging nucleotides are subject to trimming activity, which may require a separate exonuclease or which is brought about by exonuclease activity of DNA polymerases. This rather simple model that entails one or more rounds of limited primer-template extensions can also explain the footprints in which the staggered cut is "filled up". In this footprint (making up for ~30% of all revertants), the DSB-flanking sequences remain, while the 3' CA nucleotide overhangs also end up in the footprint. Figure 2A illustrates that this footprint also follows the rules that were laid out above: pairing of the left side's 3'A nucleotide to the first available T complementary strand on the right side. De novo synthesis of 3 nucleotides then generates a stretch that after dissociation can perfectly pair to the outer 3' end of the right side. Both ends can now prime DNA polymerase action and ligation is required to complete repair.

We next asked whether this model could also explain footprints that do not contain duplicated flanking sequences. This is indeed the case for most (if not all) *st192* footprints provided that the 3'ends are subject to trimming activity (Suppl. Figure 1 explains this for the second-most abundant footprint). Importantly, this model also provides the explanation for the dominating footprint (130/134) found at the *st136* site. Here, after DSB formation, pairing of both 3'A nucleotides to near Ts on opposite template strands generates a 6 bp perfect complementary region that can be extended on both ends without the need for dissociation or trimming activity.

In summary, we present evidence that EJ repair in the *C. elegans* germline is independent of NHEJ and SSA and propose a simple molecular mechanism that explains simple and complex DSB repair products. A limited number of enzymatic activities are required: i) resection or helicase action to liberate the template for de novo DNA synthesis, ii) polymerase activity and iii) ligase acitivity. Our data indicate that elongated strands dissociate to allow annealing of the 3' end to new target sites. What determines the elongated strand to dissociate? Is it influenced by the degree of 5' resection, by competing basepairing of displaced 5' ends, or by non-perfectly matching basepairs just upstream of the elongation initiating base pair? The notion that none of the tested genes act in this mechanism suggests that there are additional factors involved in DSB repair in the *C. elegans* germ line that have not been identified. We suspect that a similar mechanism can explain the identical types of repair products that have also been described in *Drosophila* [11, 12] and mammalian cells [13].

References

1. Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366-374.
2. Pastink, A., Eeken, J.C., and Lohman, P.H. (2001). Genomic integrity and the repair of double-strand DNA breaks. Mutat Res *480-481*, 37-50.
3. O'Neil, N., and Rose, A. (2006). DNA repair. WormBook, 1-12.
4. Clejan, I., Boerckel, J., and Ahmed, S. (2006). Developmental modulation of nonhomologous end joining in Caenorhabditis elegans. Genetics *173*, 1301-1317.
5. Plasterk, R.H. (1991). The origin of footprints of the Tc1 transposon of Caenorhabditis elegans. Embo J *10*, 1919-1925.
6. Plasterk, R.H., and van Luenen, H.G. (1997). Transposons. In C. elegans II, D.L. Riddle, T. Blumenthal, B.J. Meyer and J.R. Priess, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).
7. Eide, D., and Anderson, P. (1985). Transposition of Tc1 in the nematode Caenorhabditis elegans. Proc Natl Acad Sci U S A *82*, 1756-1760.
8. Robert, V.J., Davis, M.W., Jorgensen, E.M., and Bessereau, J.L. (2008). Gene Conversion and End-Joining-Repair Double-Strand Breaks in the Caenorhabditis elegans Germline. Genetics *180*, 673-679.
9. Morton, J., Davis, M.W., Jorgensen, E.M., and Carroll, D. (2006). Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. Proc Natl Acad Sci U S A *103*, 16370-16375.
10. Robert, V., and Bessereau, J.L. (2007). Targeted engineering of the Caenorhabditis elegans genome following Mos1-triggered chromosomal breaks. Embo J *26*, 170-183.
11. McVey, M., Radut, D., and Sekelsky, J.J. (2004). End-joining repair of double-strand breaks in Drosophila melanogaster is largely DNA ligase IV independent. Genetics *168*, 2067-2076.
12. Adams, M.D., McVey, M., and Sekelsky, J.J. (2003). Drosophila BLM in double-strand break repair by synthesis-dependent strand annealing. Science *299*, 265-267.
13. Roth, D.B., Porter, T.N., and Wilson, J.H. (1985). Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol *5*, 2599-2607.
14. Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.
15. Wang, M., Wu, W., Wu, W., Rosidi, B., Zhang, L., Wang, H., and Iliakis, G. (2006). PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic Acids Res *34*, 6170-6182.

Supplemental Table 1 Tc1 footprints of *unc-22(st192)* derived alleles in the indicated genetic backgrounds.

unc-22(st192)			
6	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGG	GTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
2	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG	TTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGA	GTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
7	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	ATTTTTGGGA	TGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTT	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT	CCAAATTTTGGG	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT	TTTTGGGAATTTTGGAT	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	ATTTTTGTTTTTTTTTGGGATGTATGATTTTTGGGA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
lig-4 unc-22(st192)			
8	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	ATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
3	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCT	GAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCA	CGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGAC	CTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
10	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	TGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTTG	ATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TTTTTGG	ATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
2	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	ATTTTTGGGATACA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TATATTTTTGGG	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TTTGGGTATA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGA	ATTTTTTGGGA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
xpf-1 unc-22(st192)			
11	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
2	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	CGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
2	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
2	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	GTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TTTTGGGATGGA	TGTATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TCAGGAAACTTCAAG	TATGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA	TGTCGTTGAAAATTTTGGGATACA	TGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAGAGACGACGGTGGTTCTCCAAACAAT	CGTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	
1	AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGAAGGA	93bp-Tc1	TGTCGTTGAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA
1	AACCCACCAAG	TTGAAACGTTTTTGAGAAGAGAGGTGGCGGTGATTTGGCTCCAGTCAAGATGCCAATGGTCAAGGGAACTGAATGCATTGTGCCA	

F10G8.7 unc-22(st192)

11 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGG
6 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG
8 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA

pme-1 unc-22(st192)

14 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA
2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGAT
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTTTT
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTG
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATA
5 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
2 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA
1 AACCCACCAAGAGACGACGGTGGTTCTCCAATTTTGGGATAAA

TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TAATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TAATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
GGA
TGTC
TTGGGA
ATTTTGGGA
GTCGTTGAAAA
GGTTCTCCAATTTTTGGGA
GTGATTTTGGGATTTTGGGATAATTTTGGGA

TGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
GTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
ATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TAATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
CGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
TGTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
GTATGTCGTTGAACGTTTTGAGAAGAGAGGTGGCGGTGATTGGGCTCCAGTCAAGATGCCAATGGTCAAGGAACTGAATGCATTGTGCCA
ATTTTGG
TGTCATTGA
ATTTTGGGA
TGTCGTTGAA
TTCTCCAAATTTTGG

unc-22(st136)			
30	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
xpf-1(e1487) unc-22(st136)			
30	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
F10G8.7(tm2073) unc-22(st136)			
13	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
1	AGATTGACGAGATCCATAAGGAAGGATGTACA		ACATTGAACTGGAAGCCTCCAACT
lig-4(ok716) unc-22(st136)			
28	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
1	AGATTGACGAGATCCATAAGGAAGGAT	ATCTTTTTTGGCCAGCAC	TGTACATTGAACTGGAAGCCTCCAACT
1	AGATTGACGAGATCCATAAGGAAGGATGT	GAAAAAAAGTc1(54bp)ATCTTTTTTGGCCAGCAC	TGTACATTGAACTGGAAGCCTCCAACT
pme-1(ok988) unc-22(st136)			
29	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
1	AGATTGACGAGATCCATAAGGAAGGATGTACA		ACATTGAACTGGAAGCCTCCAACT
brc-1(tm1145) unc-22(st136)			
17	AGATTGACGAGATCCATAAGGAAGGATGTACA		TTGAACTGGAAGCCTCCAACT
1	AGATTGACGAGATC*CATA*AGGAAGGATGTA	TGTACATA	GAACTGGAAGCCTCCAACT
1	AGATTGACGAGATCCATAAGGAAGGATGTACA	TCATTGAACACA	TTGAACTGGAAGCCTCCAACT

Supplemental Table 2 Tc1 footprints of *unc-22(st136)* derived alleles in the indicated genetic backgrounds.

Supplemental Table 3 *unc-22(st192)* footprints with duplicated flanks.

Chapter 7

Summarizing discussion

The identification of the TLS polymerase *rev-1* as a protector against microsatellite instability (MSI) [1] was the basis for the studies described in **Chapter 2**. This observation is unique for *rev-1*; to date, none of the other *C. elegans* TLS polymerases have been implicated in MSI prevention. Additionally, in other organisms a role for REV1 in MSI avoidance has not been described. We first confirmed this role of *rev-1* with a more sensitive MSI reporter system. The positive result led us to pursue the isolation of a *rev-1* knockout mutant to initiate further characterization of *rev-1* in MSI prevention. To our surprise, we found that a mutation that introduces a premature stopcodon in the *rev-1* ORF is (in a homozygous state) not compatible with *C. elegans* embryogenesis. Such an essential role for viability of REV1 was also described in one specific mouse strain [2]. Because of REV-1's essential role, *rev-1* knockout animals could not be used for further analysis. We assume that we failed to observe complete lethality in our earlier RNAi experiments because of an incomplete penetrance of the RNAi treatment: RNAi frequently results in knockdown-not knockout conditions. Later experiments that are described in **Chapter 2** showed that in our current RNAi setup, *rev-1(RNAi)* leads to almost 100% lethality, likely because our protocols have become more robust and efficient. Nevertheless, we are currently investigating whether we can titrate the RNAi effect to get to "hypomorphic" conditions, especially because some of our current genome instability assays require germline transmission. We recently engineered a strong MSI tract in a selectable endogenous marker gene and this would allow us to determine the molecular nature of the mutations that are induced in a *rev-1* mutant background. Thus far, it remains unclear whether these MSI mutations are similar to those caused by MMR genes [1, 3, 4] and therefore it is not known whether *rev-1* acts dependently or independently of the MMR pathway in the prevention of MSI. We used our RNAi conditions to further elucidate the essential function of *rev-1*.

For various *C. elegans* replication factors that are essential for viability, RNAi knockdowns displayed disruption of timing of the first cell divisions in the embryo [5-7]. Since *rev-1* is a polymerase, we speculated that the lethality might be caused by delayed cell divisions, with developmental defects as a consequence. However, we did not detect an effect of *rev-1(RNAi)* on the timing of the cell divisions in the earliest stages of embryonic cell divisions. We next considered the hypothesis that accumulation of DNA damage may be underlying embryonic arrest/death. Although, limited FISH analysis did not reveal gross ploidy changes in *rev-1(RNAi)* embryos, we observed a profound increase in the number of RAD-51 foci in *rev-1(RNAi)* embryos (RAD-51 foci imply the presence of double-strand breaks). We favor the hypothesis that in the absence of REV-1 DNA replication blocks/stalls either collapse more frequently or need more time to be resolved. Early embryogenesis in *C. elegans* goes without a proper checkpoint control and unfinished DNA replication may thus lead to chromosomal breaks. The accumulation of breaks might interfere with embryonic development and therefore cause lethality.

Apart from using reverse genetics approaches to identify factors that prevent MSI, we also used forward genetic screen. In **Chapter 3,** we aimed to identify new alleles of MMR components and of novel genes. The genetic approach would also help us to obtain hypomorphic alleles of essential genes. The identification of alleles of four known *C. elegans* MMR genes validated the approach. The screen also resulted in strains that by sequence analysis did not have mutations in the known MMR genes and are thus candidate novel genome stability factors. However, initial attempts to map the mutations failed, likely because MSI in these backgrounds is so high that the reporter catches some germline mutations when strains are kept in culture. Current novel approaches make use of reporters as well as a genomic MSI locus at an endogenous marker gene. This setup is easier to control and preliminary screens look promising.

In our laboratory, we have developed a variety of genome instability reporters after the MSI reporters had proven to be successful (specific and scalable). One of these reporters is capable of detecting G-tract instability and is described in **Chapter 4**. Previously, it was shown that animals that lack a gene called *dog-1* (for deletion of guanine-rich DNA) display a very specific genome instability phenotype: deletions occur at sites of polyguanine stretches that have a minimum length of ~20 bases [8]. We subsequently showed that polyG sequences are mutagenic because they have the ability to form G4 DNA structures: all sequences that match the G4-DNA signature G3-5N1-3G3-5N1-3G3-5N1-3G3-5 are mutagenic in a *dog-1* mutant background. These so-called quadruplex structures can be formed in ssDNA and are likely to be sufficiently stable to block ongoing DNA replication. Thus far, *dog-1* is the only gene known to be involved in avoiding this type of lesions. We aimed to find additional genes with a *dog-1*-like phenotype by use of the G-tract instability reporter in forward genetic screens. All mutants we have identified in such screens (five) carried mutations in the *dog-1* gene, showing that the reporter is highly specific. It remains unresolved whether *dog-1* is the only gene that is involved in guanine-tract instability: the screens have by far not reached saturation. *dog-1* is a large gene, which may be the reason why it is found with high frequency (5 in 3200 genomes). It could also be that additional genes are essential and therefore cannot be isolated in a forward genetic screen. It is interesting to note that knockdown of the replication cofactor RPA was found to be synthetic lethal to *dog-1*. Because *dog-1* is the *C. elegans* ortholog of FANCJ [9], other Fanconi Anemia genes would be good candidates. However, we showed that the homologue of FANCD2, *fcd-2*, did not display G-tract instability; additional FANC homologues have yet to be identified in the worm system.

In addition, we found that *dog-1* animals are hypersensitive to the DNA crosslinking agent Cisplatin. This compound is a widely used drug in chemotherapies in order to cure a variety of cancers, being most successful against testicular cancers [10]. In **Chapter 5**, we describe a genome-wide RNAi approach to identify genes involved in the response to this chemotherapeutic drug. A better understanding of the response to Cisplatin and the genes involved might lead to the identification of additional drug targets that improve the Cisplatin

cure rate. Before we performed our RNAi screen, we first characterized the response of *C. elegans* to Cisplatin. We observed dose-dependent broodsize reduction and embryonic lethality phenotypes. In wild type animals, treatment with Cisplatin predominantly induced deletions in the genome that varied in size. We analyzed animals deficient for NER or ICL repair for the response to Cisplatin and showed, in agreement with data from other species [11, 12], that both pathways are required to counteract the genotoxic effects of this compound, also indicating that the repair pathways involved in the repair of Cisplatin are functionally conserved in *C. elegans*. Next, we performed a genome-wide RNAi screen and found 51 RNAi clones that conferred a Cisplatin hypersensitive phenotype. We were able to confirm a number of genes with genetic mutants and to relate others via literature to the Cisplatin response. Strikingly, we didn't find many genes that were known to be involved in the repair of Cisplatin-induced DNA lesions. Therefore, we performed an additional targeted gene approach with RNAi against a variety of DNA repair genes, including genes of which we used knockout alleles in our exploring studies. Remarkably, none of these scored positive under our assay conditions suggesting that the efficiency of these clones to knockdown the gene of interest is not sufficient to confer a Cisplatin hypersensitive phenotype.

In **Chapter 6**, we present data that suggest a new end-joining pathway in the *C. elegans* germline. We found end-joining products that contained inserted sequences that are homologous to sequences adjacent to the breaks. This led to the proposal that end-joining repair in the *C. elegans* germline is predominantly dictated by primer-template extension which is strongly influenced by microhomology at the breaksite. This pathway is distinct from classical non-homologous end-joining; LIG-4 was shown not to be required for this pathway. In addition, the *C. elegans* homologue of the XPF1/ERCC1 endonuclease, implicated in single-strand annealing, was not necessary for this end-joining pathway.

Template-directed insertions have previously been observed in worms [13], flies [14] and man [15, 16]. The size of our data set allowed us to analyze a large number of insertions and propose a model that requires polymerase activity. Thus far, we have not identified the polymerase responsible for the insertion of the homologous sequences (several have been tested) or the ligase that eventually connects the newly generated DNA with the parental strand. The assay used in **Chapter 6** only allows a candidate gene approach and only of genes that do not interfere with viability.

References

1. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H., and Tijsterman, M. (2003). Identification of genes that protect the C. elegans genome against mutations by genome-wide RNAi. Genes Dev *17*, 443-448.
2. Jansen, J.G., Langerak, P., Tsaalbi-Shtylik, A., van den Berk, P., Jacobs, H., and 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.
3. Kolodner, R.D., and Marsischky, G.T. (1999). Eukaryotic DNA mismatch repair. Curr Opin Genet Dev *9*, 89-96.
4. Tijsterman, M., Pothof, J., and Plasterk, R.H. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161*, 651-660.

5. Brauchle, M., Baumer, K., and Gonczy, P. (2003). Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in C. elegans embryos. Curr Biol *13*, 819-827.
6. Encalada, S.E., Martin, P.R., Phillips, J.B., Lyczak, R., Hamill, D.R., Swan, K.A., and Bowerman, B. (2000). DNA replication defects delay cell division and disrupt cell polarity in early Caenorhabditis elegans embryos. Dev Biol *228*, 225-238.
7. Sonnichsen, B., Koski, L.B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A.M., Artelt, J., Bettencourt, P., Cassin, E., Hewitson, M., Holz, C., Khan, M., Lazik, S., Martin, C., Nitzsche, B., Ruer, M., Stamford, J., Winzi, M., Heinkel, R., Roder, M., Finell, J., Hantsch, H., Jones, S.J., Jones, M., Piano, F., Gunsalus, K.C., Oegema, K., Gonczy, P., Coulson, A., Hyman, A.A., and Echeverri, C.J. (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature *434*, 462-469.
8. Cheung, I., Schertzer, M., Rose, A., and Lansdorp, P.M. (2002). Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. Nat Genet *31*, 405-409.
9. Youds, J.L., Barber, L.J., Ward, J.D., Collis, S.J., O'Neil, N.J., Boulton, S.J., and Rose, A.M. (2008). DOG-1 is the Caenorhabditis elegans BRIP1/FANCJ homologue and functions in interstrand cross-link repair. Mol Cell Biol *28*, 1470-1479.
10. Einhorn, L.H., and Donohue, J.P. (1998). Advanced testicular cancer: update for urologists. J Urol *160*, 1964-1969.
11. Reed, E. (1998). Platinum-DNA adduct, nucleotide excision repair and platinum based anti-cancer chemotherapy. Cancer Treat Rev *24*, 331-344.
12. Niedernhofer, L.J., Lalai, A.S., and Hoeijmakers, J.H. (2005). Fanconi anemia (cross)linked to DNA repair. Cell *123*, 1191-1198.
13. Robert, V., and Bessereau, J.L. (2007). Targeted engineering of the Caenorhabditis elegans genome following Mos1-triggered chromosomal breaks. Embo J *26*, 170-183.
14. McVey, M., Radut, D., and Sekelsky, J.J. (2004). End-joining repair of double-strand breaks in Drosophila melanogaster is largely DNA ligase IV independent. Genetics *168*, 2067-2076.
15. Roth, D.B., Porter, T.N., and Wilson, J.H. (1985). Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol *5*, 2599-2607.
16. Adams, M.D., McVey, M., and Sekelsky, J.J. (2003). Drosophila BLM in double-strand break repair by synthesis-dependent strand annealing. Science *299*, 265-267.

Nederlandse samenvatting voor de leek

Elk levend organisme bestaat uit één of meerdere cellen. Een gist is een eencellig organisme, terwijl het menselijk lichaam ongeveer 100 000 000 000 000 cellen bevat. Die cellen hebben allemaal verschillende functies en maken bijvoorbeeld deel uit van ogen, huid of darmen. Alle cellen bevatten verschillende onderdelen, organellen genaamd, die bijdragen aan het goed functioneren van de cel. Een van deze organellen is de kern, waarin zich de chromosomen bevinden die bestaan uit DNA. Het DNA is het erfelijk materiaal en bevat genen. Deze genen coderen voor RNA dat op zijn beurt codeert voor eiwitten. Deze eiwitten zorgen dat een cel zijn functie kan uitvoeren. Elke cel heeft onderhoudseiwitten nodig die zorgen voor bijvoorbeeld de energieverbranding en celgroei, maar ook eiwitten die specifiek zijn voor de functie van de cel. Al deze eiwitten worden dus gecodeerd door genen en alle genen van de cellen samen bepalen, in het geval van een mens, bijvoorbeeld de kleur van je ogen, de grootte van je neus en het geluid van je stem.

DNA bestaat uit twee lange strengen die om elkaar heen gewikkeld zijn. Elke streng wordt gevormd met 4 bouwstenen: A (adenine), T (thymine), C (cytosine) en G (guanine). Deze bouwstenen, basen genoemd, bepalen de code van de genen. Elke combinatie van drie basen codeert voor een bepaald aminozuur, dat zijn de bouwstenen van eiwitten. Zo leidt AAG in het DNA tot het inbouwen van een phenylalanine, terwijl AAC leidt tot de inbouw van een leucine. De volgorde van de DNA basen bepaalt zodoende de vorm en functie van het eiwit.

De twee DNA strengen kunnen om elkaar heen wikkelen omdat de twee strengen complementen van elkaar zijn. De basen in de ene streng paren met de basen uit de andere streng, waarbij een A altijd paart met een T en een C altijd met een G een paar vormt. Als een cel gaat delen, moet het DNA worden vermenigvuldigd (of gerepliceerd), zodat elke dochtercel een kopie ontvangt. Het vermenigvuldigen wordt uitgevoerd door replicatie-enzymen: DNA polymerases. De twee DNA strengen worden uit elkaar getrokken en aan elke streng wordt een nieuwe streng gemaakt. Het polymerase bouwt tegenover elke T een A, elke A een T, elke C een G en elke G en C. Hierbij wordt dus een nieuwe complement gebouwd. Als de replicatie is afgerond, zijn er twee identieke DNA moleculen ontstaan en zo blijft de erfelijke informatie behouden na de celdeling.

De replicatie van het DNA is niet altijd foutloos en soms wordt de verkeerde base in gebouwd. Ook kunnen basen worden beschadigd waardoor het replicatie-enzym de base niet meer kan herkennen en dus niet weet welke base het moet inbouwen. DNA schades kunnen afkomstig zijn van verschillende bronnen. Deze kunnen afkomstig zijn vanuit buiten de cel, zoals UV licht, Röntgen straling, uitlaatgassen en tabaksrook. Ook binnen in de cel ontstaan bij verschillende processen tussenproducten die het DNA kunnen beschadigen, zoals zuurstofradicalen. Schades leiden tot het inbouwen van verkeerde basen tijdens de replicatie en dit heeft grote gevolgen. Zoals eerder vermeld, leidt elke combinatie van drie

basen tot de inbouw van een ander aminozuur in het eiwit. In het genoemde voorbeeld, is een verandering van een AAG naar een AAC genoeg voor een verandering in de opbouw van het eiwit, omdat er nu een ander aminozuur wordt ingebouwd. Veranderingen in DNA basen worden mutaties genoemd. Deze mutaties leiden tot verandering in eiwitopbouw wat kan leiden tot verandering in eiwitfunctie. Als er veel mutaties optreden, kan het voorkomen dat een cel niet meer goed functioneert, omdat te veel eiwitten in de cel gestoord zijn in hun functie door de opgelopen mutaties. Een mogelijk gevolg is dan dat een cel in een bepaald orgaan, die normaal langzaam of niet deelt, een snel delende cel wordt. Deze snel delende cellen zijn tumorcellen en vormen samen een tumor. Het is aangetoond dat tumorcellen veel meer mutaties bevatten dan een normale cel. Tumorcellen zijn dan ook vaak gestoord in één van de mechanismen die betrokken zijn bij het behoud van genetische informatie en dus bij het voorkomen van mutaties. Deze DNA schade herstelmechanismen kunnen beschadigde basen repareren en replicatiefouten verbeteren en daarmee voorkomen ze mutaties. Dit proefschrift beschrijft de bestudering van een aantal van deze herstelmechanismen.

Voor dit onderzoek werd het modelorganisme *Caenorhabditis elegans* gebruikt. *C. elegans* is een nematode (of rondworm) die ongeveer 1 mm lang kan worden en bestaat uit 959 cellen. Alhoewel het een veel simpeler organisme is dan alle zoogdieren, heeft het wel veel genen en biologische processen gemeen met complexere organismen. Andere voordelen van het gebruik van *C. elegans* als modelorganisme zijn dat het makkelijk in grote aantallen te groeien is, makkelijke genetica heeft en dat de zorg eenvoudig is. Daarnaast bestaat er voor *C. elegans* een makkelijk methode om RNAi (RNA interference) uit te voeren. Bij RNAi wordt het RNA, dat de tussenstap is tussen DNA en eiwit, afgebroken. RNA is in tegenstelling tot DNA enkelstrengs. Als er complementaire RNA moleculen aanwezig zijn, wordt dubbelstrengs RNA gevormd dat kan worden afgebroken door de RNAi enzymen. Op deze manier is er geen RNA aanwezig om eiwit mee te maken en wordt het gen dus uitgeschakeld. Door deze methode te gebruiken, kan de rol van een eiwit in het organisme bestudeerd worden. We kunnen nu kijken wat er gebeurt als het eiwit niet meer aanwezig is en daardoor de functie van het eiwit beter begrijpen. Voor *C. elegans* is er een verzameling waarin voor elk gen een complementair stuk RNA zit dat het gen kan uitschakelen. Die RNAs zitten in de bacteriën die de wormen eten. Zo kan dus heel makkelijk RNAi tegen alle genen worden uitgevoerd.

In hoofdstuk 2 van dit proefschrift wordt de functie van het gen *rev-1* bestudeerd. Het REV1 eiwit is een speciaal soort polymerase. DNA herstelmechanismen reparen voortdurend schades in het DNA, maar er zijn altijd schades die achterblijven en aanwezig zijn als replicatie begint. Bij het begin van de replicatie kunnen er nog beschadigde basen aanwezig. Deze basen geven een probleem tijdens de replicatie, want het normale polymerase weet niet welke base het tegenover een beschadigde base moet zetten. Dat komt omdat het polymerase heel nauwkeurig is in het inbouwen. Als het polymerase een schade tegenkomt, kan het

niet verder gaan met de replicatie, de nauwkeurigheid dan in gevaar komt. Het polymerase kan alleen als er met hoge nauwkeurigheid kan worden gewerkt funtioneren. De replicatie moet echter doorgaan, want anders gaat de cel dood. Om celdood te voorkomen, wordt er gekozen om replicatie door te laten gaan en het risico op mutaties te accepteren. Allereerst moet het normale polymerase van het DNA verwijderd worden om vervolgens de replicatie te laten overnemen door polymerases die minder nauwkeurig zijn, maar wel tegenover een beschadigde base kunnen inbouwen. Deze inbouw gaat wel vaak gepaard met de inbouw van een verkeerde base en leidt dus tot mutaties. REV1 is één van die onnauwkeurige polymerases en lijkt ook een rol te spelen bij de regulatie van dit proces.

Onderzoek in ons laboratorium heeft laten zien dat *rev-1* betrokken is bij het voorkomen van een bepaalde vorm van mutaties. Deze mutaties heten frameverschuivingen. Zoals eerder vermeld, coderen steeds drie basen voor een aminozuur. Stel dat er ergens opeens een extra base wordt toegevoegd in een DNA sequentie. Nu zijn er opeens andere combinaties van drie basen, omdat alle basen doorschuiven door de extra base. Nu zal het stuk DNA dus voor compleet andere aminozuren en dus een compleet ander eiwit coderen. Dit wordt een verandering in frame of frameverschuiving genoemd.

Frameverschuivingen kunnen ontstaan door het toevoegen en of verwijderen van basen. Dit gebeurt vooral op stukken waar veel dezelfde basen op een rij zitten, bijvoorbeeld 17 A's of 23 C's. Die verschuivingen vinden daar plaats, omdat het polymerase soms gaat schuiven op zo'n serie bases en dat niet doorheeft omdat de basen hetzelfde zijn en dus een base te veel of te weinig in bouwt. In ons laboratorium is een indicator van frameverschuivingen ontwikkeld. Hierbij is een regio met 17 A's ingebouwd in het gen LacZ. Dit is zo gedaan dat de codering van het LacZ gen is verschoven en dus niet het goede eiwit wordt gemaakt. Als er een frameverschuiving optreed in de 17 A's wordt de codering voor het LacZ wel correct en wordt het goede eiwit gemaakt. LacZ codeert voor het eiwit B-galactosidase dat makkelijk te herkennen is, omdat het wormen blauw kan aankleuren. Als er in de indicator een frameverschuiving optreedt, worden de wormen blauw en zijn ze makkelijk te onderscheiden. Tot nu toe was van vier genen in *C. elegans* bekend dat ze bij het voorkomen van frameverschuivingen betrokken waren, de zo genaamde mismatch herstelgenen. Als deze genen werden uitgeschakeld met RNAi waren er meer blauwe wormen en dus meer frameverschuivingen in de populatie dan zonder uitschakeling. Dit bevestigde dat de indicator werkte. Daarna werd met RNAi één voor één elk gen uitgeschakeld en gekeken bij welke genen er een verhoging in het aantal blauwe wormen te zien was. Eén van de genen die werd gevonden was *rev-1*. Nog niet eerder was een rol voor *rev-1* in het voorkomen van frameverschuivingen aangetoond.

In hoofdstuk 2 wordt een rol voor *rev-1* in het voorkomen van frameverschuivingen bevestigd met gebruik van een tweede indicator waar een serie van 23 C's is ingebouwd in het LacZ gen. Voor verdere analyses is het beter om een genetische mutant van een gen te hebben dan alleen RNAi te gebruiken. In genetische mutanten zorgt een mutatie in het gen

dat het gen niet meer voor functioneel eiwit codeert en is het gen dus ook uitgeschakeld. In een mutantenverzameling werd gezocht naar een *rev-1* mutant en vervolgonderzoek wees uit dat een functioneel *rev-1* gen essentieel is voor overleving. Zonder een functioneel *rev-1* gen overleven de wormen niet; de embryo's ontwikkelen zich niet en de eieren komen niet uit. Wat de rol van *rev-1* is in de ontwikkeling van *C. elegans* is nog niet bekend.

In hoofdstuk 3 worden frameverschuivingen verder bestudeerd. De C23 indicator bleek nauwkeuriger te zijn dan de A17 indicator. In het geval van de A17 reporter werd een groot aantal genen gevonden die bij verdere analyse vals-positieven bleken te zijn. Ook werd de zoektocht uitgevoerd met RNAi, dat niet altijd werkt en dus tot vals-negatieven leidt. Om een beter beeld te krijgen van de betrokken genen werd een genetische screen uitgevoerd met de C23 reporter. Hier werd gebruik gemaakt van een stofje, EMS, dat overal in het genoom van de wormen mutaties aanbrengt. Als de mutatie in een gen optreedt, kan het leiden tot uitschakeling van dat gen. Als de concentratie EMS hoog genoeg is, bevat elke worm een aantal uitgeschakelde genen. Na behandeling met EMS werd een verzameling mutante dieren gemaakt. In die verzameling werd met de C23 indicator gezocht naar mutanten die een verhoging van frameverschuivingen vertoonden. Deze mutanten hebben hoogstwaarschijnlijk een mutatie die de functie van een gen betrokken bij het voorkomen van frameverschuivingen verstoord. De vier bekende mismatch herstelgenen zijn goede kandidaten om in deze mutanten mutaties te hebben opgelopen. DNA sequentie analyse bevestigde deze veronderstelling en voor alle vier de genen werd een mutant gevonden. Er bleven echter mutanten over waarvan de mutatie niet in één van de vier genen lag. Er zijn genetische technieken beschikbaar waarbij de mutatie te achterhalen is. Deze technieken werden toegepast op de overgebleven mutanten, maar dit bleek niet succesvol. Analyse van de gebruikte methoden gaf aan dat de opzet van de proef nu verkeerd was en het achterhalen van de mutaties met deze opzet niet mogelijk was of heel veel werk zou kosten. Als deze problemen kunnen worden opgelost, kan de indicator wel gebruikt worden.

Hoofdstuk 4 bespreekt een tweede type mutatie. Een lab in Vancouver beschreef het verdwijnen van enkele honderden basen in het DNA in wormen met een mutatie in het *dog-1* gen. Het verdwijnen van basen heet een deletie. Deleties komen vaker voor, maar deze deleties waren specifiek voor bepaalde plaatsen in het genoom. Ze vinden alleen plaats op plekken waar minstens 20 G's op een rij voorkomen. De deleties beginnen dan altijd een paar honderd basen voor die regio met G's, maar eindigen altijd precies na die G's. Dit type mutatie is dusverre alleen in *C. elegans* gevonden en komt alleen voor in *dog-1* mutanten.

Een mogelijk verklaring voor het feit dat de deleties alleen voorkomen bij G's is dat bij experimenten met DNA in een reageerbuis is aangetoond dat in een serie G's interacties optreden tussen die G's onderling. Als dit gebeurt, ontstaat er een hele stabiele structuur, een G4 structuur, in het DNA en verondersteld wordt dat replicatie langs deze structuur niet

kan plaatsvinden, met welk DNA polymerase dan ook. De stop in replicatie kan leiden tot een breuk in het DNA en als de twee einden van het DNA verkeerd aan elkaar geplakt worden, ontstaat er een deletie. In hoofdstuk 4 wordt aangetoond dat de G4 structuur inderdaad de oorzaak is. In een uitgebreide analyse bleek dat de deleties ook optreden in regio's waarin rijtjes van G's worden onderbroken door andere basen. Als deze onderbrekingen de vorming van de G4 structuur niet in de weg staat, kunnen deleties worden waargenomen.

Ook voor deze vorm van mutaties is een LacZ indicator ontwikkeld. Deze bestaat uit een LacZ gen onderbroken door een stuk DNA van een paar honderd basen en een regio G's. Als er nu een deletie optreedt, ontstaat er een functioneel LacZ gen en kleuren wormen met zo'n deletie blauw. De indicator is heel specifiek, want in normale wormen werd nooit een blauwe worm gezien, terwijl bij *dog-1* mutanten er regelmatig blauwe wormen te zien zijn. Omdat er pas van één gen, *dog-1*, bekend is dat het betrokken is bij de G4 deleties, werd een genetische screen uitgevoerd om meer genen te vinden. De screen werd uitgevoerd zoals eerder beschreven. Er werden vijf mutanten gevonden die de G4 deletie indicator blauw lieten kleuren. Na sequentie analyse bleek dat alle vijf de mutanten een mutatie in *dog-1* hadden gekregen door de EMS behandeling. Dit kan betekenen dat *dog-1* het enige gen is dat G4 deleties voorkomt, maar dat is niet de verwachting. Een betere verklaring is dat de andere genen die G4 deleties voorkomen een essentiële functie in *C. elegans* ontwikkeling hebben. Dan kunnen er geen mutanten gemaakt worden, want dan overleeft de worm niet meer, zoals het geval was bij *rev-1*. Het kan ook zijn dat de analyse met een verzameling die meer mutanten bevat moet worden overgedaan en dat er dan wel andere genen gevonden worden.

In hoofdstuk 5 wordt de respons van wormen op het medicijn Cisplatina bestudeerd. Cisplatina wordt veel gebruikt in chemotherapie bij de behandeling van kankerpatiënten. Cisplatina bindt aan DNA en brengt dus beschadigingen aan. Het bindt aan het DNA van alle cellen in het lichaam, maar tumorcellen hebben moeite met de schades om te gaan. Zoals eerder gezegd, werken de DNA herstelmechanismen van tumorcellen niet goed meer en het herstellen van de schades gaat dus niet goed meer. Ook delen tumorcellen veel sneller dan gewone cellen en is er dus minder tijd om de schades te repareren omdat er voortdurend replicatie moet plaatsvinden. Als er te veel schades achter blijven, ondergaat de cel apoptose. Bij apoptose laat de cel zichzelf doodgaan omdat het niet meer goed functioneert en een gevaar is voor het lichaam. Omdat de hoeveelheid achtergebleven schades in tumoren hoog is gaan deze cellen dood, terwijl de gezonde cellen weinig schades hebben en dus overleven. Snel delende cellen als haarcellen hebben ook geen tijd om de schades op tijd te herstellen en zullen dus ook doodgaan. Dit is de reden dat patiënten kaal worden van chemotherapie.

In hoofdstuk 5 wordt aangetoond dat in *C. elegans* de DNA schade die door Cisplatina wordt aangebracht door dezelfde herstelmechanismen als in andere organismen wordt

gerepareerd. Deze mechanismen zijn dus geconserveerd en is voor worm en mens vergelijkbaar. Cisplatina heeft ook invloed op de ontwikkeling van de embryo's; bij behandeling met oplopende doses Cisplatina zijn er steeds minder embryo's die overleven. Door de DNA schades wordt de ontwikkeling die gepaard gaat met snelle celdelingen te veel verstoord.

Cisplatina is een veel gebruikt chemotherapiemedicijn, maar de behandeling is niet altijd effectief. Dit komt omdat bepaalde tumoren resistent zijn tegen de Cisplatina. Om meer tumoren succesvol te kunnen behandelen, is het belangrijk om te weten of er genen zijn die we kunnen uitschakelen en daardoor een verhoging in Cisplatina gevoeligheid krijgen. Deze genen zouden dan goede kandidaten zijn om te dienen als "drug target". Door aan de behandeling een medicijn toe te voegen die dat gen uitschakelt verhoog je de Cisplatina gevoeligheid van de cellen. Ook kan geprobeerd worden het gen met RNAi uit te schakelen. Zoals eerder verteld, is er een *C. elegans* RNAi verzameling en deze verzameling is gebruikt om te kijken welke uitgeschakelde genen tot verhoogde gevoeligheid voor Cisplatina leiden. Na het analyseren van 17 000 genen was het resultaat 51 genen. Deze genen zijn voornamelijk betrokken bij het onderhoud van cellen en het transport in cellen. De volgende stap is nu om te kijken of het uitschakelen van de gevonden genen ook tumorcellen gevoeliger maakt.

In hoofdstuk 6 wordt gekeken naar het herstel van breuken in het DNA. Een van de meest gevaarlijke DNA schades voor een cel is een breuk. Eén breuk in het gehele genoom kan genoeg zijn om een cel in apoptose te laten gaan. Voor andere types schade kan een cel overleven met veel hogere aantallen.

Bij een DNA breuk is het chromosoom helemaal doorgebroken en zijn er dus twee losse einden DNA die weer goed aan elkaar gezet moeten worden. De einden zijn heel gevoelig voor afbraak en vaak worden er stukjes vanaf gegeten, waarbij dus genetische informatie verloren kan gaan. Er zijn verschillende manieren om de breuk weer te plakken. De eenvoudigste is de recht-toe-recht-aan manier en dat is het aan elkaar plakken van de einden. Omdat er vaak van de einden basen zijn afgegeten, leidt deze manier dus vaak tot verlies van sequentie, een deletie in het DNA. Een tweede manier om de einden aan elkaar te plakken maakt gebruik van op-elkaar-lijkende stukjes sequentie die aan beide kanten van de breuken voorkomen. Als er aan beide kant van de breuk een rijtje basen complementair zijn, kunnen de beide stukken DNA daar met elkaar paren. De twee strengen DNA smelten eerst uit elkaar en vervolgens paart een van de strengen aan de ene kant van de breuk met een streng aan de andere kant van de breuk. Aangezien de sequenties die paren meestal niet precies aan het uiteinde zitten, ontstaan er losse eindjes die worden afgeknipt. De basen die tussen de sequenties zaten zullen dus verdwijnen en er is bij deze manier ook weer sprake van een deletie. Er is een derde manier om de breuk te repareren en deze manier leidt niet tot verlies van basen. In elke cel zijn van elk chromosoom twee kopieën aanwezig. Een gebroken chromosoom kan ook de tweede kopie gebruiken om de breuk te herstellen. Er

vindt dan replicatie op het andere chromosoom plaats. Een eind wordt dan zo ver verlengd dat het weer op de het andere eind past en daarmee paart. Het eind van het gerepliceerde stuk is complementair aan het andere einde en als ze nu worden vastgezet zijn er geen deleties en is de hele sequentie hersteld.

In hoofdstuk 6 is gekeken naar de sequenties die ontstaan na het repareren van een breuk. Analyse van de data gaf aan dat er nog een andere manier moet zijn om de breuken te herstellen. Sequenties die het gevolg waren van een van de drie bovengenoemde manieren werden gevonden. Echter, er werden ook sequenties gevonden die niet het gevolg kunnen zijn van een van de bekende manieren. Bij deze sequenties werden juist extra basen, inserties, gevonden. Dat betekend dat er replicatie heeft plaatsgevonden voordat de breuk werd gerepareerd. Welke genen er betrokken zijn bij deze manier van breuk herstel, is nog niet bekend. Voor een aantal genen waarvan bekend is dat ze via een van de drie bekende manieren betrokken zijn, werd aangetoond dat ze niet betrokken zijn bij deze nieuwe manier.

Dankwoord

Na vijf jaar zit het er dan eindelijk op, het proefschrift is af! Nu is het tijd om mijn dank uit te spreken aan iedereen die heeft bijgedragen tot de totstandkoming van dit proefschrift. Omdat ik de laatste jaren tussen af en toe een topje ook flink wat dalen heb gezien, zijn het vooral de morele steunbijdragen die van cruciaal belang zijn geweest bij het halen van de eindstreep. Daarom is dit deel van het proefschrift niet alleen het meest gelezen deel, maar voor mij zeker ook het belangrijkste.

After five years it's finally over, the thesis is done! Now it's time to thank the people that contributed to the creation of this thesis. Through the last years, I experienced an up every now and then, however, there were quite a lot of downs on the way. Because of this, ithe moral support-contributions were crucial in making it to the finish line. Therefore, this part of the thesis is not only the most read, but for me certainly the most important.

Marcel, it has been a rocky road. Ik denk dat we er een allebei hoop van geleerd hebben. Bedankt dat je me de kans hebt gegeven om het af te maken. Veel succes in Leiden.

Evelien, wij begonnen tegelijk en ik vond het leuk om je te werken. Ik ben er trots op dat ik een (kleine) bijdrage aan jouw *dog-1* verhaal heb mogen leveren. Je bent altijd een gezellige gast, ook al gooi je meestal iets om.

Kristy, you helped me when it counted the most: at the beginning and the end of my time in the lab. Thank you for that. I liked sharing an office with you in the last years a lot. I'll miss the chats and the snacks. I'll have to start looking for somebody else to feed me or finally go for that size 8! Maybe one day we can start our own party or baking company?

Sophie, Daphne, Bennie, Wouter, mede-OIO's in de Tijstermangroep. Jullie kregen met mij te maken op het moment dat ik eigenlijk nog maar één ding wilde: weg hier. Hopelijk hebben jullie daar niet te veel last van gehad. Ik vond het leuk om met jullie samen te werken en te discussiëren over de resultaten. Bedankt voor de goede tijd op en buiten het lab.

Dan degenen die met mij de annex bevolkten in de tijd dat het eigenlijk het genoom instabiliteit/high-throughput hok was. Het was aangenaam om daar in die goede sfeer te werken. Joris, je was altijd bereid te helpen. Op elke vraag had je een antwoord. Proeven konden altijd in twee weken, jammer dat dat in de praktijk niet altijd klopte..... Ron, ook jij was altijd positief en geïnteresseerd in mij. Jouw verhalen over voornamelijk je avonturen met je buren, waren een mooie afleiding van de dagelijkse beslommeringen. Eelke, bedankt voor de praatjes en het feit dat je altijd even vroeg hoe het ervoor stond.

(High-throughput) Gijs, zonder jou was dit boekje er niet geweest. Jij steunde en hielp me op het moment dat degenen die dat hadden moeten doen dat niet deden. Jouw interesse in mijn projecten hebben me veel geholpen. Sorry voor al die keren dat ik me lekker overal mee bemoeide, het was goed bedoeld.

Florian, dude, thanks for all the good times, chats, support, and everything else. You were always there to listen to me when I needed to get my thoughts straight on something. Your jokes, our gossips, and talk about CRAZY people helped me to get through the day. Good luck in Seattle. I'll come by to explain you all you need (but probably not want) to know about the Mariners soon.

Josien, ook al spuugde je altijd in mijn epjes, je hebt me veel geholpen met alles. Zonder jou had ik het niet gered. Ik kon altijd bij jou terecht voor een praatje, hulp bij proeven of een feitje uit de enorme hoeveelheid kennis die je altijd bereid hebt. Je ongelooflijke geduld en doorzettingsvermogen zijn bijzonder, maar je moet niet overdrijven. Arthur, bedankt voor de wok-to-gos en technical support. Bedankt voor jullie vriendschap. Ik zie jullie aan de West Coast!

Marjo, Leonie, Hsin-Yi, Luc, Suzanne, Kay, Lucas, Regis, Joost, Florian and Josien from the Ketting Group. Thank you for adopting me! The fun I've had with you guys almost made me wanting to work on small RNAs just so I could officially join your group. Luckily, I could always hang out on your site of the lab.

Ira, bedankt voor al je hulp door de jaren heen. En ook voor alle keren dat je me een fijne avond kwam wensen.

In mijn tijd in de Plasterk en Tijsterman labs, heb ik met veel mensen lab- en kantoorrruimte gedeeld. Ik ga geen namen noemen, want dan ben ik bang dat ik iemand vergeet. Wel wil ik zeggen: bedankt voor de goede tijden op het lab en daarbuiten.

Ana, I became less Dutch because of you, I'm grateful for that. Thank you for being angry with my boss on behalf of me, when I couldn't because of political reasons. You created a monster and even though, now, I'm more adorable than you, you're still my friend, amazing. Good luck with finishing your PhD, I'll be there for the mental support! Keep up the feminism; you have to compensate for me!

Your office was my safe haven to relax whenever I wanted to escape my third floor troubles. To all Ana's office mates: thanks for putting up with me. Sylvia and Daniel: thank you for the friendship.

Alle andere Hubrechters: bedankt voor de goede sfeer in het Instituut. Het Hubrecht zou niet kunnen draaien zonder een aantal mensen. Jules, Romke, Richard, Thea, Anneke, de mediakeuken, Arjan en de rest van de IT, de technische dienst, de administratie, receptie en personeelszaken; bedankt voor al jullie hulp aan mij en jullie werk waardoor werken in het Hubrecht zo fijn was. Susann, bedankt dat je me altijd vroeg hoe het ging met me. Thea; bedankt voor de gezellige praatjes als ik de deur in- of uitging.

Dalila, op dezelfde dag op het Hubrecht van start en gelijk een goede vriendschap. Fijn dat ik altijd tegen je aan mag klagen en dat we samen flink kunnen lachen over alles op het lab, maar ook ver daar buiten. Keep the soul alive! Alle concerten, samen met Hajar, hebben me door de laatste jaren gerockt. Glamorous indie rock&roll is what I want, it's in my soul, is what I need. Wetenschap is niet zaligmakend!

Erwin, wij begonnen samen aan een Utrechts avontuur en vond het fijn dat we elkaar daarin konden steunen. Heerlijk, elke dag in de stoptrein. Gelukkig stapte jij in Lammenschans altijd vrolijk in! Jij en Vicky bedankt voor de hulp bij het afronden van dit proefschrift. De weekendjes weg en de avondjes met de andere scheiko's Rob, Martijn en Jasper bevielen me prima.

Lars en Sandra (+kids), mijn bezoekjes aan jullie in San Diego waren superleuk.

Brandon and Karen, it was great visiting you in SF.

Sophie en Astrid, Boston was geweldig en zo gezellig. Ik vergeet dat gezicht van dat mens bij de Cheesecake Factory nooit meer!

As, bedankt voor alle tips over gezond eten. Nu weet ik wanneer ik slecht bezig ben, dat scheelt een stuk. We moeten maar weer eens een potje honkbal gaan kijken, Fenway of Haarlem, veel verschil maakt het niet....

Chris, Elke, Sjoerd, Sharief en Bjørn, de dartavondjes in de WW waren een zeer goede reden om in Leiden langs te komen.

Bij BV Oegstgeest voel ik me altijd thuis, een unieke vereniging en elke keer weer leuk om langs te komen. Wout, een eer dat ik je assistent mocht zijn! Families van Eenbergen, Star en Creyghton; jullie zijn mijn basketballfamillie, bedankt voor alle gezelligheid op de tribune.

Marcel, jouw boekje was een inspiratie voor de mijne, fijn dat we elkaar altijd lekker direct over de stand van zaken in onze levens kunnen vertellen.

Iris, filmpje, 90's Now!, pubquiz, succesvolle mannenjacht (tja...) of drankje, altijd gezellig! Ik zie je in Tennessee!

Luc, neehee EROP! De espresso-lesbobar zal ons missen.

De stichting: jullie vroegen altijd naar mijn wormen (vooral de blauwe) en mijn leventje in Utrecht. Bedankt voor jullie interesse.

Eric, Jeanette, Immer, Liekje en Tijm; bedankt voor alle leuke vakanties en alle andere goede tijden. Van mistroostig Schots weer tot Mexicaanse zon of gewoon lekker thuis, we hebben het altijd gezellig.

Lieve Tess en Evi, jullie hebben me goed geholpen met de wormen eten geven. De afleiding die jullie me gaven, hielp goed bij het relativeren van alles in dit boekje en zo veel meer. Kom snel weer eens logeren. Karin en Leon, bedankt voor jullie steun en vriendschap.

Oop, bedankt voor de sinaasappels. Het onvoorwaardelijke geloof in al uw kleinkinderen is iets fantastisch. Ik mis oma nog elke dag en ik vind het jammer dat ze de afronding van dit boekje niet heeft kunnen meemaken.

Peet, bedankt dat ik je lievelingszus mag zijn, ook al heb je er maar een. Je bent ook de mijne! Het is fijn dat we nu goed met elkaar overweg kunnen en het veel vaker met elkaar eens zijn dan vroeger. Ik vind het altijd leuk om je verhalen over je avonturen op school te horen. Ik vind het heel knap dat je je werk kon combineren met studeren en basketball. Ik ben trots op je.

Papa en mama, bedankt voor al jullie hulp en steun de afgelopen jaren. Fijn dat ons gezin zo hecht is en dat we altijd bij elkaar terecht kunnen. Dat is heel bijzonder. Fijn dat jullie toch af en toe mijn was wilden doen, tegen alle principes in.

Jacob, meestal ben je wat bedachtzaam, maar voor mij ging je zonder lang na te denken, ik ben blij dat we de rest van ons leven samen zullen zijn. Bedankt voor je liefde, steun en hulp tijdens de laatste loodjes.

Karin

Curriculum Vitae

Karin Brouwer was born August 3 1980 in Leiden. In 1998, she received a VWO diploma from the Rijnlands Lyceum in Oegstgeest. In September of the same year, she started studying Chemistry at the University Leiden. Her dissertation project "A study on the *RAD5* and *RAD16* genes of *Saccharomyces cerevisiae*" was performed in the Molecular Genetics department headed by Prof. Dr. Jaap Brouwer. In May 2003, she received her MSc in Chemistry. Her PhD research started in September 2003 in the Functional Genomics group led by Prof. Dr. Ronald Plasterk at the Hubrecht Institute in Utrecht. From February 2007 through August 2008, her research was performed in the Genome Dynamics and Stability group led by Dr. Marcel Tijsterman, also at the Hubrecht Institute. From December 2008, she is performing a Museum Education Internship at the Oregon Museum for Science and Industry in Portland, Oregon, USA.

Karin Brouwer werd op 3 augustus 1980 geboren in Leiden. In 1998 behaalde zij haar VWO diploma aan het Rijnlands Lyceum in Oegstgeest. In september van dat jaar begon zij aan haar studie Scheikunde aan de Universiteit Leiden. Haar hoofdvakstage "A study on the *RAD5* and *RAD16* genes of *Saccharomyces cerevisiae*" werd uitgevoerd in de groep Moleculaire Genetica van Prof. Dr. Jaap Brouwer. In mei 2003 behaalde zij haar doctoraal examen. Haar promotieonderzoek werd aangevangen in september 2003 in de Functional Genomics groep onder leiding van Prof. Dr. Ronald Plasterk in het Hubrecht Instituut in Utrecht. Van februari 2007 tot en met september 2008 is het onderzoek vervolgd in de Genome Dynamics and Stability groep van Dr. Marcel Tijsterman, ook in het Hubrecht Instituut. Vanaf december voert ze een Museum Educatie stage uit bij het Oregon Museum for Science and Industry in Portland, Oregon, Verenigde Staten.