Roadblocks & bypasses : protection of genome stability by translesion DNA synthesis in C. elegans

Bostelen, I. van

Ivo van Bostelen

Roadblocks & Bypasses

Protection of genome stability by translesion DNA synthesis in *C. elegans*

Roadblocks & Bypasses

Protection of genome stability by translesion DNA synthesis in *C. elegans*

Ivo van Bostelen

Author: Ivo van Bostelen
Design & layout: Eric van der Zwet & Ivo van Bostelen
Printing: Puntgaaf Drukwerk
ISBN: 978-90-9032-426-5
Email author: ivo.vanbostelen@gmail.com

Financial support for the printing of this thesis was provided by MRC Holland BV.

To my parents, who have let me be free.

Roadblocks & Bypasses

Protection of genome stability by translesion DNA synthesis in *C. elegans*

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden, op gezag van Rector Magnificus Prof. mr. C.J.J.M. Stolker, volgens besluit van het college voor Promoties, te verdedigen op dinsdag 3 december 2019 klokke 13.45 uur.

door

Ivo van Bostelen

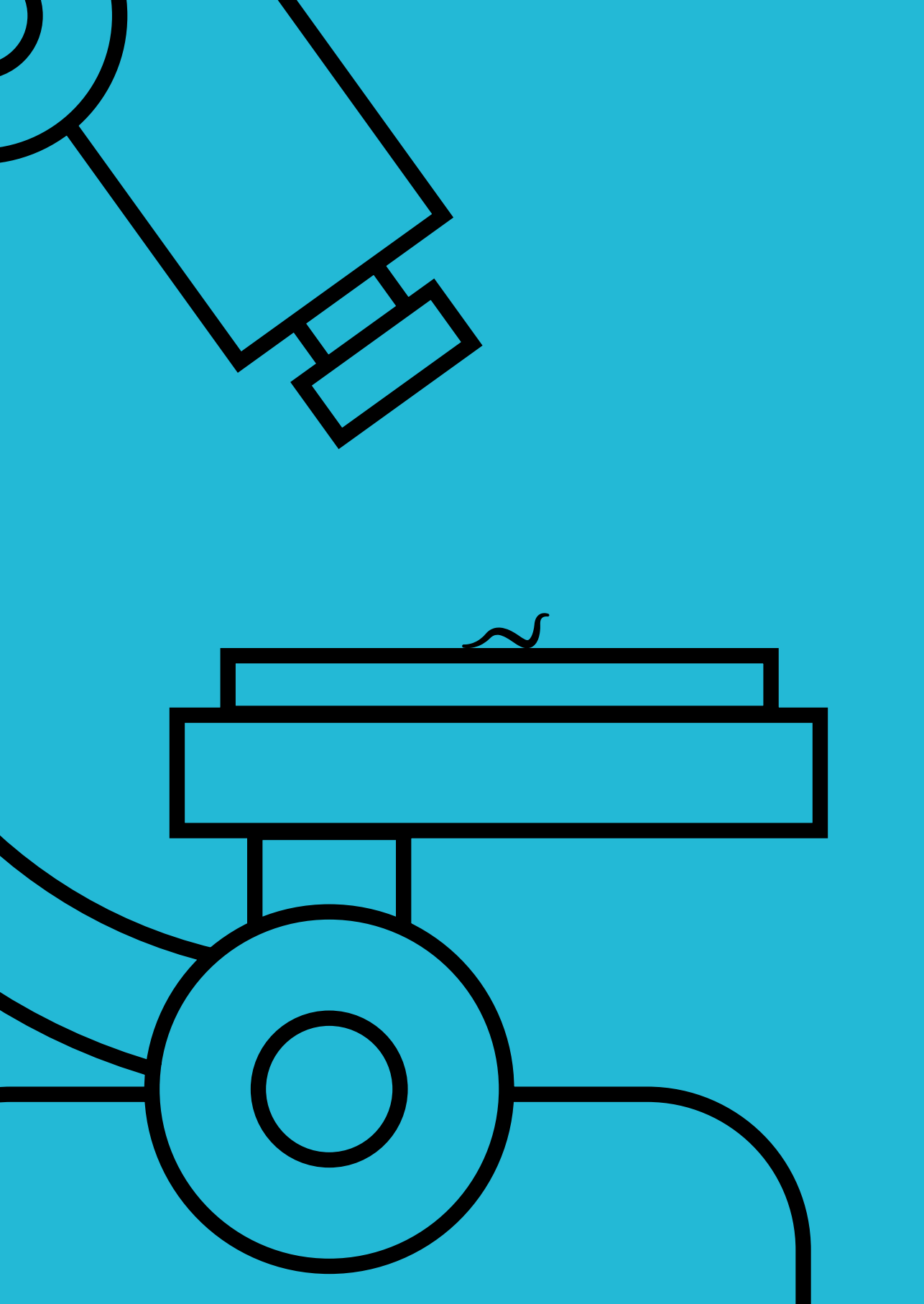
geboren te Leiderdorp, Nederland
in 1983

Promotor
Prof. dr. M. Tijsterman

Co-promotor
Prof. dr. J. Brouwer

Promotiecommissie
Prof. dr. J. T. den Dunnen
Prof. dr. S. J. L. van den Heuvel (Universiteit Utrecht)
Dr. P. Knipscheer (Hubrecht Insitute)

Contents



1

General introduction

A fragile molecule central to life on earth

Deoxyribonucleic acid - DNA - is central to life on earth. It is literally at the core of each living cell and it is fundamental to all living organisms because it encodes the genetic instructions for life. Therefore, its fragility seems counterintuitive; DNA is chemically instable and this leads to decay over time. Spontaneous reactions such as hydrolysis and oxidation damage the DNA, and threats to the stability of DNA do not end there: reactive byproducts of the cells metabolism chemically alter the DNA, and environmental sources of DNA damage are numerous [1,2]. So, don't organisms need stable genomes to guarantee their survival and proliferation? Of course they do, because lesions in DNA interfere with the most essential biological processes: transcription and replication. Additionally, DNA damage is a major driver of mutation formation - a process that causes disease like cancer, ageing, and congenital disorders. While an instable genome might be a bad thing for the health of an individual organism, a certain level of mutagenesis is advantageous for species and for life on earth in general. Without genome instability there would not be any genomic diversity, and without genomic diversity there cannot be evolution. Intriguingly, it is also evolution that has spawned a multitude of bio-molecular pathways that combat genome instability caused by DNA damage, a network of pathways that is termed the DNA damage response.

Before I go into more detail about the pathways that govern the repair of DNA damage, I will briefly introduce the sources of DNA damage. The well-known double helix of the DNA molecule with bases adenine (A), cytosine (C), guanine (G) and thymine (T) paired between sugar (deoxyribose) and phosphate backbones has the potential to undergo many different chemical alterations. Spontaneous hydrolysis of the *N*-glycosyl bond between the base and sugar group, results in an abasic site. These lesions are by far the most abundant and it is estimated that a human cell has a steady-state level of 30.000 abasic sites. Additionally, hydrolysis can lead to deamination of bases, thereby altering the chemical structure of the base. Reactive molecules - often the result of the cells own metabolism - also contribute to endogenous DNA damage; oxidation, methylation and alkylation of nucleotides result in alterations to nucleotides and pose threats to genome stability [2]. Although DNA replication is not a source of DNA damage per se, it is an endogenous source of genome instability because occasional replication errors introduce mutations. The wrong base can be inserted, but also sequence specifics can lead to replications errors: small insertions and deletions can occur at repetitive stretches of DNA. On top of that, some sequences have the potential to form secondary structures, exemplified by G-quadruplexes [3]. The micro-environment in the nucleus of a living cell already seems a harsh place for DNA to be, but when we step outside, into the sunlight, other dangers present themselves. Sunlight contains ultra violet (UV) light that can induce dimerization of adjacent pyrimidines and oxidative damage as well. Another well-known example of an exogenous source of DNA damage is ionizing

radiation (IR), which induces a wide variety of damages, including, oxidized bases and breaks of the DNA backbone. Moreover, there are many natural and synthetic genotoxic chemicals. Alkylating agents like MMS and EMS for example, cause base adducts. Another striking example is cisplatin, a chemical that can form a covalent bond between bases of opposing DNA strands. This kind of lesion is called an interstrand crosslink (ICL), and is extremely toxic because it prevents strand separation of DNA, thereby forming an absolute block of transcription and replication. These examples of chemotoxins are familiar, and there are many more chemicals that damage DNA, too many to discuss them all here. It is clear that there is a plethora of threats to the integrity of DNA, with which cells have to deal continuously. This calls for efficient and faithful mechanisms that maintain the healthy state of the genetic material and thereby safeguard transcriptional and replicative potential.

The DNA damage response

The network of molecular pathways that governs the cells response to DNA damage is called the DNA damage response (DDR); it entails damage detection, downstream signaling and DNA repair actions. While DDR includes the promotion and execution of repair, it also initiates senescence and apoptosis when the damage load is too high. All these processes together suppress the accumulation of too much DNA damage and mutagenesis and prevents cells from becoming cancerous [4]. As I have described earlier, there are many types of DNA lesions and evolution has provided cells with recognition and repair pathways that are tailored to different classes of lesion. To keep within the scope of this thesis I will limit the discussion of the DNA damage response to the most important DNA repair pathways and damage tolerance mechanisms, with a focus on the central theme of this thesis: Translesion DNA synthesis.

Base excision repair (BER)

This pathway is highly conserved from *E. coli* until humans and deleterious mutations in BER genes cause increased mutagenesis and higher chance of developing cancer [5,6]. Simply put, BER removes damaged bases and replaces them with new nucleotides, specifically for lesions that do not cause much structural change to the normal helical organization of the DNA. These lesions have the potential to induce base substitutions and cause polymerases to stall. By repairing them, BER safeguards replicative and transcriptional potential and genomic stability. Substrates include the common abasic sites, but also oxidized and alkylated bases. These last two types are recognized and removed by a lesion specific glycosylase, which thus generates abasic sites. The DNA backbone is subsequently cut by an AP endonuclease, generating a DNA single-strand break (SSB) and a single nucleotide gap. Polymerase β then fills the gap. When only one new nucleotide is incorporated this is called short-patch BER and the downstream

processing and ligation is done by XRCC1 and LIG3, respectively. Long-patch BER incorporates up to 11 new nucleotides and processing and ligation is performed by FEN1 and LIG1 [7,8]. The enzymes involved in ligation during BER are also responsible for the repair of SSBs from other sources [9].

Nucleotide excision repair (NER)

Lesions that distort the helical structure of the DNA require a different repair mechanism. Most notably, and widely studied in the DNA repair field, are pyrimidine dimers that are induced by UV light. The covalent bonds between neighboring cytosines and thymines - in the forms of cyclobutane–pyrimidine dimers (CPDs) and 6–4-pyrimidine–pyrimidone photoproducts (6–4PPs) - induce tension on the DNA molecule and cause bending of the backbone. Additional lesions that are subject to NER include a wide variety of structures: cisplatin induced intrastrand crosslinks, bulky chemical adducts, products of reactions with reactive oxygen species (ROS) like cyclopuriones, and more. Transcription and replication are performed by high fidelity polymerases that cannot use such distorted templates; therefore, all these lesions block these essential processes. Detection of NER-specific lesions happens through two separate sensing mechanisms. Transcription-coupled NER (TC-NER) safeguards efficient transcription; it is activated when the RNA polymerase stalls on a lesion. This recruits the CSA and CSB proteins, among others, that then promote repair. The other sensing mechanism of NER is referred to as global genome NER (GG-NER). GG-NER continually scans the whole genome for distortions in the DNA helix and mainly employs DDB1 (also known as XPE) for lesion recognition and promotion of repair. After damage recognition the downstream steps of the repair pathway are identical; the TFIIH complex opens the DNA, XPF and XPG excise the damaged DNA, and then a polymerase and DNA ligase restore the DNA to the original state [10], (Fig. 1).

Many NER factors are named for their association to the disease *Xeroderma pigmentosum* (XP). Deleterious mutations in the XP genes result in a deficiency to deal with UV-induced DNA damage. Patients are extremely sensitive to sunlight and have ~1000-fold elevated risk for developing skin cancer. Cockayne syndrome (CS) and trichothiodystrophy (TTD) are other genetic disorders associated with NER and, among other symptoms, also display hypersensitivity of the skin to sunlight [10,11].

Mismatch repair (MMR)

Although it is debatable if mismatches fall in the category of DNA damage, the MMR pathway is essential for the maintenance of genome stability. The polymerases that are responsible for DNA replication have incredibly high fidelity; the main replicative polymerases in mammals Polδ and Polε generate just one error for up to 10^7 correct incorporation events. The rare errors are either mismatched bases or microsatellite

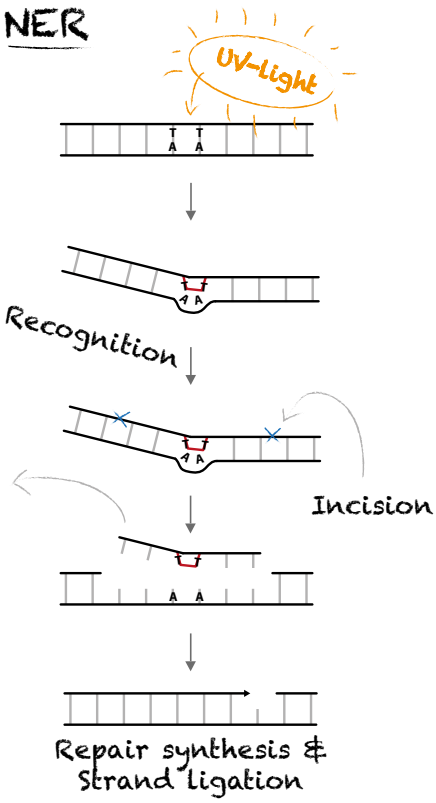


Figure 1. Nucleotide excision repair.
A schematic representation; for details see text.

mutations: small insertions or deletions (indels) in repetitive DNA sequences due to polymerase slippage. These polymerases achieve this high fidelity because their tight catalytic centers do not accommodate mismatched bases and they have proofreading activity that detects misaligned DNA after replication [12]. Nevertheless, this level of mutagenesis would still be substantial, especially if you consider the number of rounds of replication necessary to build and maintain a human being. The MMR pathway is able to recognize mismatches and small loops caused by slippage at microsatellite sequences, while also distinguishing the old and new DNA strands. In eukaryotes the MutS-homologs MSH2, MSH6 and MSH3 enable recognition, which is followed by nicking of the nascent strand by MutL-homologs MLH1 and PMS2, Then, EXO1 removes the nicked DNA, before a replicative polymerase and LIG1 take care of resynthesis and ligation, respectively. Thereby MMR reduces the error rate approximately 100-fold to an estimated 1 for every 10^9 correct base incorporations [13,14]. The importance of this pathway in the maintenance of genome stability is illustrated by the symptoms of Lynch syndrome patients. Lynch syndrome is caused by deleterious mutations in MMR genes and patients have an increased risk of developing colon cancer early in life.

A hallmark of Lynch syndrome is microsatellite instability: patients have increased genomic instability at mono- di- and tri-nucleotide repeats [15].

DNA damage tolerance mechanisms

Although BER and NER provide efficient DNA repair of lesions that stall replicative polymerases, the abundance of base damages has led to a strong evolutionary selection pressure on alternate systems that help to sustain replicative potential. These alternative pathways, together termed DNA Damage Tolerance (DDT), do not remove lesions but provide the potential to complete DNA replication in the presence of damage and thereby suppress the formation of replication associated DSBs. There are two well-known pathways described in literature: Template Switching (TS) & Translesion Synthesis (TLS). Template switching uses the undamaged strand as a template for replication past the lesion and in that way bypasses the damage without induction of mutations [16,17]. This mechanism shares intermediates with homologous recombination (HR, discussed later in more detail). It is employed in bacteria and yeast, but it is unclear whether such a pathway is also used extensively in higher eukaryotes [18-20]. Although TS is error free, it is dependent on complex HR-like reactions, which may be a time-consuming process for the cell especially when, during stages of embryogenesis, there is a demand for a quick fix to resolve replication blocks. Also, it is thought that TS can cause genomic rearrangements when the HR-like intermediates cannot be resolved properly [21,22]. The second DDR pathway - TLS - is conserved among all kingdoms of life and in multicellular eukaryotes it is arguably the most essential DDT pathway [23].

Translesion DNA synthesis (TLS)

Specialized TLS polymerases are employed to synthesize DNA opposite damaged templates. Their wider catalytic center and lack of proofreading capacity allow them to accommodate damaged bases and bulky adducts as templates, thereby permitting the continuation of DNA replication without actual repair [24]. This efficient process protects against the many deleterious effects of replication form collapse, but due to the intrinsic properties of TLS polymerases and their inability to read certain damaged templates correctly it is also mutagenic [25-27], (Fig. 2). In eukaryotes TLS is predominantly mediated by polymerases from the Y-family: Polη, Polκ, Polι, and REV1 and by B-family polymerase Polζ [28,29]. Structural differences of these enzymes define the functional specificities of TLS polymerases and bypass of more demanding lesions may require the combined efforts of more than one TLS polymerase [30]. TLS can happen directly at the replication fork by switching the replicative polymerase temporarily for a TLS polymerase or it can happen post-replication when the single strand DNA regions that are left opposite the DNA lesions are filled in by TLS polymerases [31]. Before I go into more detail about the regulation of TLS and the current models, I will first introduce the main players.

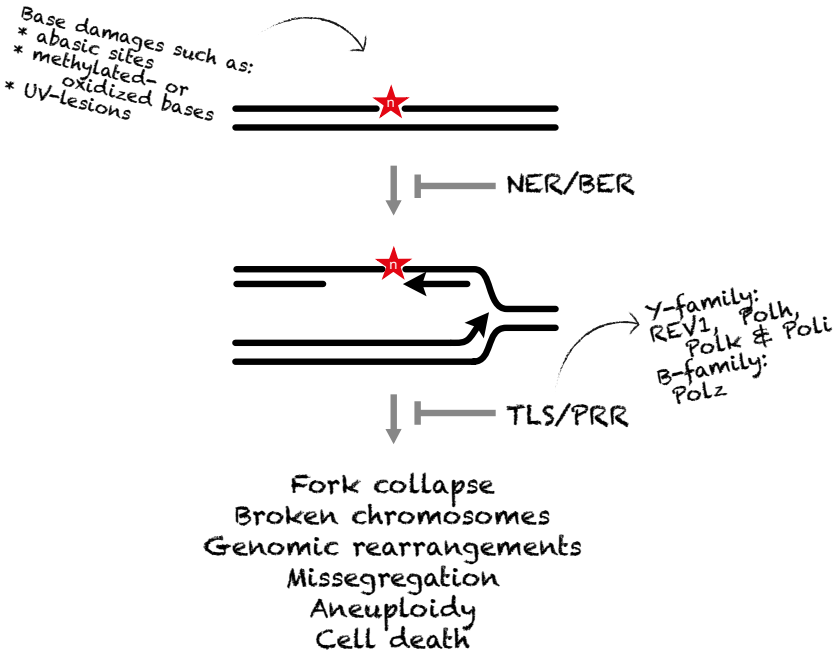


Figure 2. How translesion synthesis protect against deleterious consequences. *For details see text.*

Polymerase eta (Polη)

Polη is involved in the bypass of UV photolesions, especially CPDs and likely also 6-4PPs [32-34]. This is highlighted by the increased cancer risk and sensitivity to sunlight of *Xeroderma pigmentosum* variant (XP-V) patients that have a mutation in the Polη encoding gene [32,35,36]. In addition, Polη is important for the bypass of the most abundant oxidized base, 8oxoG, that is also indirectly induced by sunlight [37,38] and many others lesions: (+)-*trans*-*anti*-benzo[a]pyrene-N2-dG [39], acetylaminofluorene-adducted guanine [40], *O6*-methylguanine [41], thymine glycol [42], and adducts derived from cisplatin and oxaliplatin [43]. Bypass of CPDs and 8oxoG by Polη is remarkably error free, but on undamaged DNA Polη's fidelity is very low as compared to replicative polymerases. Despite this high mutagenic potential, knockdown in human cells and knockout in MEFs actually increases mutagenicity [44] and Polη loss in yeast does not change mutagenesis [32,45]. Due to redundancies between the TLS polymerases it is likely that other TLS polymerases, possibly Polι or Polκ are able to take over [46]. Also, the observations that overexpression of Polη in human cells and yeast has limited effect on the mutation rate point to very stringent regulation that restricts Polη activity only to

damaged DNA [47,48]. Different protein interactions contribute to this strict regulation: the sliding clamp PCNA can interact directly with Polη and this interaction is boosted by the Rad6-Rad18 dependent mono-ubiquitination of PCNA [49]. Mammalian Rev1 can also physically interact with Polη [50]. Work in DT40 chicken cells has shown that these two interactions regulate Polη in temporarily different ways: REV1 is needed for Polη dependent bypass at the replication fork, while PCNA mono-ubiquitination recruits Polη to post-replicative ssDNA gaps opposite lesions [51]. In *C. elegans* Polη is especially important during embryogenesis, likely because embryos are extremely sensitive to cell cycle delays and thus need a rapid solution to stalled replication [52-54].

Polymerase kappa (Polκ)

Polκ is the most conserved TLS polymerase and was identified via its homology with the *E. coli dinB* gene [55]. While the function in TLS of the *dinB* gene product, PolIV, has been extensively studied in *E. coli*, the role of Polκ in eukaryotes is less clear. This could be due to the lack of a Polκ homolog in the most abundantly studied eukaryote, *S. cerevisiae*, or because loss of Polκ does not cause profound changes in spontaneous or damage induces mutagenesis [56]. Also, Polκ has a relatively high fidelity on undamaged DNA when compared to other TLS polymerases. The catalytic activity of Polκ seems to be restricted to N^2-adducted dG lesions [57-60]. In addition, loss of Polκ sensitizes vertebrate cells and *C. elegans* to alkylating agents, suggesting Polκ also acts on these lesions, together or redundantly with Polη [52,61]. Similarly to Polη, Polκ also interacts with PCNA and REV1 [50,62]. When a lesion requires the concerted efforts of two TLS polymerase the insertion step of 1-2 nucleotides directly opposite the damaged nucleotide is performed by one polymerases and this is followed by extension of the aberrant primer terminus by another. As Polκ is able to extend mispaired primer termini and is relatively accurate on undamaged DNA it is a good candidate to fulfill the role of extender in addition to its specialized role in direct lesion bypass [24].

Polymerase iota (Polι)

Polι is the most recent addition to the Y-family and is not conserved in yeast and *C. elegans*. It is highly accurate when it uses dT templates, but very mutagenic when using dA templates. This is due to its specialized enzymatic core, which prefers *Hoogsteen* pairing [29]. This characteristic may also explain why Polι is able to bypass of 8oxoG in an error-free manner [63,64]. In addition, Polι may have a role in BER, possibly as a gap-filling polymerase [64].

REV1

The Rev1 gene was named for its reversion-less phenotype in yeast in knockout strains; Rev1 deficient mutants have reduced UV-induces mutagenesis [65-67].

REV1 is an a-typical polymerase because its catalytic activity is limited to incorporation of deoxycytidines (dC). REV1 can perform this opposite undamaged or damaged guanines but also across adenines, uracil and abasic sites *in vitro*, which can explain why it promotes mutagenesis [24,68-72]. During TLS, REV1 is important for bypass of lesions that are caused by lipid peroxidation and UV induced damage [67,73]. Although REV1 is structurally alike to the other Y-family members, its limited catalytic activity may suggest it has other roles. There are various studies showing that REV1 performs non-catalytic functions via interactions with other proteins: the BRCT domain of REV1 interacts with PCNA and is involved in the bypass of UV-C induced lesions in mouse ES cells [74-76]. Additionally, the C-terminal part of mammalian REV1 contains ubiquitin binding motifs (UBMs) that interact with ubiquitinated PCNA, a region able to interact with other Y-family polymerases, and a motif that interacts with the REV7 subunit of B-family TLS polymerase Polζ [24,50,77-79]. These interactions support a proposed role of REV1 as a master regulator of TLS. While the polymerase and BRCT-domain are conserved between mammalian REV1 and *C. elegans* REV-1, this is not obviously the case for other protein interaction domains. However, the functions may be conserved. In addition to its role in TLS, REV1 may play a role in maintenance of epigenetic stability at G4 sequences, possibly by facilitating replication across these hard-to-replicate sequences [80,81], and (together with Polζ) in homologous recombination (HR) break repair [82,83].

Polymerase zeta (Polζ)

The catalytic subunit and member of the B-family polymerases, REV3, and the accessory subunit REV7 make up the core of Polζ [28,84]. Like Rev1, these two genes were also discovered in screens for mutants that have a reduced UV-C induced mutation frequency in yeast [65,66,85]. More recently it was shown that accessory subunits Pol31 and Pol32 of Polδ are also part of the Polζ complex [86-89]. While Polη efficiently bypasses UV-C induced CPDs, Rev1 and Polζ are important for the bypass of 6-4PPs [68,84,90]. As Polζ is able to extend aberrant primer termini at a lesion but also at mismatched bases, and is relatively accurate on undamaged templates, it may also function as the extender polymerase after one of the Y-family polymerase has performed the insertion step of lesion bypass [24]. The mammalian homolog *REV3L* produces a protein twice the size of the yeast protein, which may indicate more complex or diverse functions compared to the yeast protein [28]. Rev3-/- mice are embryonic lethal due to high levels of apoptosis that is likely caused by the observed chromosomal instability. In humans, heterozygosity of deleterious mutations in *REV3L* cause a genetic developmental disorder called Möbius syndrome [91,92]. Together with REV1, Polζ may also play a yet ill-defined role in homologous recombination (HR) break repair [82,83].

Other polymerases involved in lesion bypass

The proteins I have just discussed are thought to perform the bulk of TLS. However, there are a few other polymerases that have shown bypass capacity, mainly in biochemical assays. Polβ has, next to its function in BER, the capacity to bypass certain cisplatin induced lesions [93], Polv and Polθ can bypass thymine glycols, and Polθ can efficiently synthesize DNA across abasic sites [94]. While Polκ and Polζ are able to extend from aberrant primer termini, making them suitable to perform the extension step of TLS, Polλ, Polμ and Polθ can also perform extension from mismatched primer termini [95-97].

Although Polθ was initially proposed to be involved in TLS, a genetic study in Drosophila showed that its function in tolerance to several kinds of DNA damage is due to a role in microhomology driven double strand break (DSB) repair [98]. Recently, the interest in Polθ has spiked and studies from our lab [99-103] and others [Reviewed in 104,105] have further characterized its biological roles. Because Polθ is of special interest for this thesis, I will discuss Polθ-mediated end joining (TMEJ) in more detail later, in the paragraph *Alternative end joining*.

Regulation of TLS

The high mutation rate and low processivity of TLS polymerases on undamaged DNA calls for a strict regulation that only allows access to the DNA when a lesion blocks replicative polymerases. Generally, this is not achieved by regulation of overall cellular concentration, but it appears that the localization of TLS polymerases is controlled. TLS polymerases are concentrated in replication factories dependent on PCNA ubiquitination [106,107]. Once the replication fork arrests at the site of a blocking lesion the eukaryotic sliding clamp PCNA is mono-ubiquitinated at lysine 164 by the E3 ubiquitin ligase RAD18 and the E2 conjugating enzyme RAD6 [108]. This reaction is promoted by the interaction of RAD18 with RPA coated ssDNA in close proximity to the lesion. Mono-ubiquitination of PCNA recruits Y-family polymerases via the ubiquitin binding motifs at the C-terminus of these polymerases [62]. Yeast cells that carry a K164R substitution have defects in UV-induced mutagenesis and mammalian cells with the same substitution are sensitive to genotoxic stress [109]. Additionally, in yeast PCNA can be poly-ubiquitinated at the same lysine residue, which depends on Rad5, and can also be SUMOylated. Both these signals are linked to recombinational DNA damage bypass but the crosstalk between regulation of TS and TLS is complex [110]. For example, Rad5 dependent poly-ubiquitination in yeast also seems to be important for efficient TLS [20]. Furthermore, the mammalian homologs of Rad5 - SHPRH and HLTF - have complex roles in DDT and seem to have functions in both TLS and other DDT mechanisms [18]. It is proposed that these proteins suppress mutagenesis after DNA damage by directing the most appropriate TLS polymerase or other bypass mechanisms to replication fork impediments [19].

In mammalian cells the PCNA-K164R substitution mutation strongly decreases TLS efficiency, but it is not completely abrogated [111]. In DT40 chicken cells the remaining TLS activity is largely dependent on the non-catalytic function of REV1, but whether this is also the case in other organisms is unclear [51]. Nevertheless, the many interactions of REV1 with different TLS factors support that REV1 also regulates TLS [75,76]. Additionally, post-translational modifications like phosphorylation and ubiquitination of TLS polymerases can provide further layers of regulation, but how this exactly works remains unclear; emerging evidence indicates that some Fanconi anemia (FA) factors also regulate TLS independent from their well-known role in interstrand crosslink (ICL) repair (this pathway will be discussed in detail later) [112]. Finally, TLS polymerases were shown to interact with the PCNA-like 9-1-1 clamp, which may be especially important for highly helix-distorting lesions [113], and lesion bypass also proved to be dependent on chromatin dynamics [114].

The lesion itself has a strong effect on the bypass mechanism and polymerase choice. For some lesions bypass can be quick and easy: for example UV-light-induced thymine-thymine cyclobutane pyrimidine dimers are efficiently bypassed by Polη [33]. Many other lesions may require more complex mechanisms and the collaborative efforts of more than one TLS polymerase [115]. TLS is also affected by the cell cycle, and specifically the progression through S-phase. This is illustrated by the identification of temporal distinct sub-pathways of TLS [51,116]. The development of novel techniques - especially high-resolution genetic assays - should help to determine the complex regulatory network of DDT and shed more light on possible TLS models, providing an interesting subject for future research [117]. The value of such future work is emphasized by the many and pleiotropic ways in which TLS is involved in health.

TLS in health and disease

As I have discussed earlier, TLS polymerases have low fidelity, especially when compared to the orders of magnitude higher fidelity of replicative polymerases. Interestingly, the mutagenicity of TLS polymerases is used in a way that is beneficial to human health. Somatic hypermutation is an essential part of our adaptive immune response; it requires induction of base substitutions in the variable immunoglobulin genes to produce a wide variety of antibodies. There is evidence that at least a subset of these base substitutions are generated through mutagenic TLS performed by Polζ, Polη, Polι and REV1 [118-120].

Although TLS may be mutagenic on occasion there is much evidence that TLS actually protects genome integrity by preventing replication fork collapse. Persistent replication fork stalling causes a very serious domino effect of problems linked to cancer and inborn disease: broken chromosomes, genomic rearrangements or missegregation and eventually aneuploidy and cell death (Fig. 2), [26]. Another positive effect of TLS

on human health is illustrated by the genetic disorder *Xeroderma pigmentosum* variant (XPV), caused by deleterious mutations in the gene that encodes Polη. Patients display hypersensitivity toward UV light and develop malignant skin neoplasia at early age due to increased mutagenesis [35,121-123]. Furthermore, mice with a hypomorphic allele of REV1 in a sensitized NER deficient background show increased UV induced skin carcinogenesis. It is proposed that this is caused by increased levels of replication stress that trigger DNA damage signaling, which in turn leads to induction of inflammation and senescence [124]. The recent development of increasingly sensitive and robust affordable sequencing techniques has also contributed to the identification of somatic mutations of TLS polymerase genes in cancers. Deleterious mutations in the genes coding for Polη and Polκ have been identified in melanomas, breast cancers, prostate tumors, ovarian- and kidney cancer [125]. Moreover, there is abundant evidence of misregulation of TLS in many different kinds of cancer. Given the fact that cancer cell rely on efficient DNA replication in the presence of high levels of replication stress it is perhaps not surprising that many cancers have upregulated TLS activity [126,127]. Additionally, TLS polymerases support chemoresistance of tumors, because many of cancer therapeutics function though the induction of replication blocking lesions that can be bypassed by TLS polymerases. The inhibition of TLS has thus been suggested as a promising approach in the treatment of cancer [127-129]. A TLS inhibitor may be used to sensitize cancer cells to existing chemotherapeutics or synthetic lethal interactions can be exploited, for example a tumor deficient in one of the DNA repair pathways may rely more on efficient TLS [130-132].

Interstrand crosslink repair

Interstrand crosslinks (ICLs) are one of the most toxic types of DNA damage and require a complex network of repair pathways and TLS. The crosslinking of paired DNA strands through a covalent bond between bases in opposing strands forms an absolute block for transcription and replication because it prevents strand separation. ICLs also have a high potential to induce base substitutions, indels and chromosomal rearrangements when mutagenic repair occurs. The high toxicity of ICLs is illustrated by the effective treatment of cancer with agents that induce ICLs; especially replicating cells have great difficulty with ICLs [133]. Synthetic chemicals such as nitrogen mustards and cisplatin can induce these lesions, but there are also natural sources of ICLs: aldehydes and natural psoralens for example [134].

Studies of the genetic disorder Fanconi anemia (FA) have provided great insight in the molecular mechanism of ICL repair. Patients with this syndrome have developmental aberrations, progressive bone marrow failure and greatly increased cancer risk, and cultured FA cells are hypersensitive to ICL-inducing drugs [135]. Currently, nineteen FA genes have been identified, and all gene products act in the same ICL repair pathway

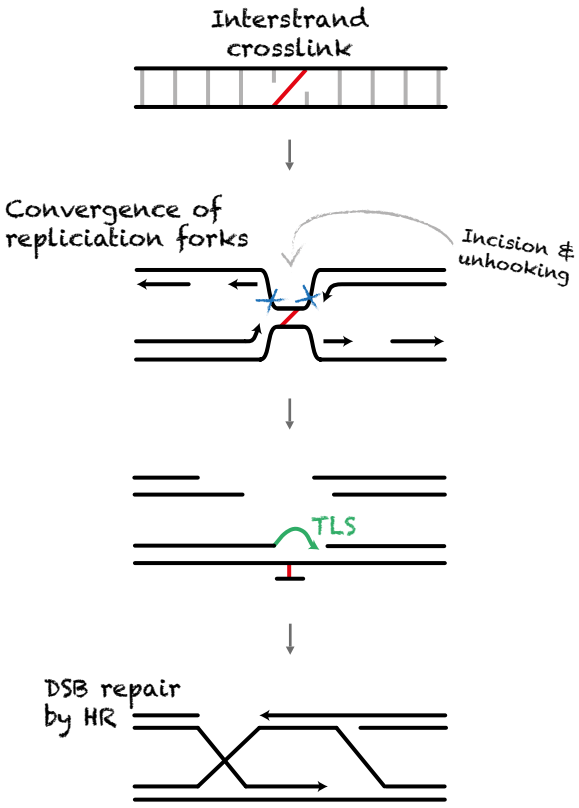


Figure 3. A schematic of interstrand crosslink repair. *For details see text.*

[136]. In the fifth chapter of this thesis I will discuss the FA pathway in more detail. Briefly, the best-known mechanism of ICL repair is as follows: upon two replication forks converging on an ICL the first step of the repair is unhooking of the crosslink by stimulating incisions in the backbone up- and downstream of the lesion through the action of by two different endonucleases. This process creates a flipped-out ICL-containing oligo still covalently attached to one of the strands. Translesion synthesis across the unhooked lesion restores the sister chromatid, before HR repair can use this as a template to repair the DSB on the other chromatid that resulted from the dual incisions (Fig. 3). TLS polymerases Polζ and REV1 are especially important for this TLS step but likely other TLS polymerases are involved too. The unhooked ICL that remains may be removed via NER or by hydrolysis, but not much is known about this step yet [137,138]. Much of what we currently know about this complex repair mechanism has come from studying the replication of plasmids that contain a single cisplatin crosslink in *Xenopus laevis* egg extracts [139-142]. While it was first thought that this was the predominant pathway for any ICL, recently it was shown that psoralen crosslinks are

repaired differently. For these lesions the DNA backbone is not incised but one of the two N-glycosyl bonds forming the crosslink is simply cleaved by a glycosylase. TLS is involved here too, but because no DSB is generated HR is not necessary for repair. This alternative pathway shows that the mechanisms of ICL repair are flexible and dependent on the chemical nature of the crosslink [143]. Both these mechanisms function in S-phase when replication forks run into ICLs. Outside the context of DNA replication the recognition and repair of ICLs is not fully understood but NER and mismatch repair factors seem to play important roles in recognition of ICLs. The efficiency of ICL repair outside S-phase is dependent on their level of distortion of the normal DNA helix structure; lesions that strongly disrupt the normal structure of the DNA helix, like cisplatin crosslinks, are more readily recognized and repaired [144]. TLS is important in replication-independent ICL repair also: Polκ and PCNA-K164 ubiquitination function in replication-independent repair in *Xenopus* egg extracts and similar a role for Polκ was confirmed in mammalian cells [145].

The sources and repair of double stranded breaks

Persistent stalling of replication - for instance when TLS is not functional - causes collapse of the replication fork eventually resulting in DNA double strand breaks (DSBs). These highly toxic and mutagenic lesions are also induced by exposure to ionizing radiation or by physical stress on the DNA molecule. Because DSBs represent physical breaks of the chromosome they can lead to loss of genomic information, but also when not properly repaired to genomic rearrangements such as translocations events that are causal to cancer development. Intriguingly, DSBs are also generated endogenously to promote meiotic crossover formation, and during V(D)J recombination. Here, DSBs are introduced to generate diversity in the antigen receptor repertoires. In all cases these breaks need to be repaired efficiently and – with exception of the breaks generated during V(D)J recombination – without aberrant loss of genetic information. The two most widely studied DSB repair pathways are homologous recombination (HR) and non-homologous end joining (NHEJ), and there are several alternative pathways [146].

Homologous recombination (HR)

This pathway uses homologous sequences from the sister chromatid or the homologous chromosome as a template to repair the chromosomal break and therefore it is *grosso modo* error free. Upon recognition of DSBs the cell cycle is halted and initiation of repair is induced via ATM signaling. Regulated processing and short-range resection of the DNA is then performed by the MRN complex (MRE11, RAD50 and NBS1), which creates small 3' single stranded overhangs. EXO1 and DNA2 can then further resect the DNA to generate long stretches of ssDNA, which become coated by the heterotrimeric RPA complex. Subsequently, RAD51 - the recombinase protein that facilitates strand invasion

- replaces RPA dependent on BRCA2. The RAD51-ssDNA filaments can then invade the sister chromatid or homologous chromosome to pair with complementary sequences, which is then used as a template for extension by a replicative polymerase. Next, the newly synthesized DNA can be displaced and anneal to the resected other end of the DSB. This reaction can use just one end of the DSB for invasion and extension, which is referred to as synthesis dependent strand annealing (SDSA). Simultaneous reactions from both sides of the DSB are also possible, leading to a complex intermediate: the double Holliday junction (dHJ). The dHJ can be resolved to form a crossover, which is essential during meiotic recombination, but non-crossovers are also a possible outcome [147]. Although HR is error free in most cases, it can lead to loss of heterozygosity when not the sister chromatid but the other homologous chromosome is used as a template in somatic cells. Also, gene conversion events are possible when another homologous sequence is used, for example a nonfunctional pseudogene [146].

HR is an essential pathway for human health; mutations in HR factors such as BRCA1, BRCA2, and RAD51 are lethal in mice and sexually reproducing organisms without functional HR cannot form meiotic crossovers and are infertile. In humans the vital function of error free repair by HR is also reflected in the various cancer predisposition syndromes associated with HR genes. Examples are: hereditary breast and ovarian cancer caused by *BRCA1* & *BRCA2* mutations, Bloom's syndrome cause by defective *BLM* (dHJ resolvase), and ataxia telangiectasia resulting from *ATM* mutations [4].

Non-homologous end joining (NHEJ)

A second major DSB repair pathway - non-homologous end joining (NHEJ) - has a much simpler mode of action: it does not use a homologous template but directly ligates the broken ends together. Arguably it is the go-to DSB repair mechanism in most somatic cells, also because the sister chromatid or homologous chromosome is not available. The major drawback of NHEJ is the potential loss of sequence at the break site, and therefore it is considered to be error-prone. NHEJ of a DSB is as follows: first, the ends of the break are bound by Ku70/Ku80 heterodimers, then DNA-PKcs tethers both ends together, and subsequently the ends are ligated by Ligase IV (Lig4) in complex with XRCC4. When the ends of the break require processing before repair this can be performed by a variety of biochemical activities, including those performed by Artemis, Polλ or Polμ [148]. Although the NHEJ pathway is conserved throughout all known eukaryotes, there are also notable differences, for instance, DNA-PKcs and Artemis, which are important NHEJ factors in mammals are not conserved in the lower eukaryotes yeast and *C. elegans* [149].

Because NHEJ is also employed to repair the programmed DSBs generated during V(D)J recombination, mutations in KU70, KU80 and DNA-PKcs can lead to severe combined immunodeficiency (SCID) [150]. In addition, in mice inactivation of NHEJ

factors XRCC4 and Lig4 is lethal, which underwrites the vital importance of this DNA repair pathway [148].

Alternative end joining

Keeping the importance of HR and NHEJ in mind and considering the tremendous toxicity of DSBs, it may not come as a surprise that there are additional alternative mechanisms a cell can employ to repair DSBs. Alternative end joining can be achieved when there is sequence homology in the flanks of the DSB, which can be used to promote end joining by annealing of the complementary sequences that are exposed after the break ends have been resected. This mode of repair results in sequence loss of the DNA between the homologous sequences. The homology in the flanks of the break can be minimal (*e.g.* up to 25 bases) in which case the repair mode is termed micro-homology mediated end joining (MMEJ). Alternatively, the term single strand annealing (SSA) is used when longer stretches of homology are used. After the ssDNA ends hybridize a polymerase can act to fill the remaining ssDNA and a ligase to ligate the DNA backbones [151].

In *C. elegans* it was found that even when HR, NHEJ and SSA were knocked out still DSB repair could be found [152]. It was subsequently found that this DSB repair route resulted a specific mutational footprint: small deletions with minimal micro-homology of 1 nt, and a subset of the deletion had inserts, which were argued to be templated from the flanks of the break. The templated inserts suggested the involvement of a polymerase in this DSB repair pathway. Following the finding that Polymerase theta (Polθ) was involved in break repair mediated by microhomology in Drosophila [98], this seemed a likely candidate. Indeed, multiple studies in recent years have shown that Polθ mediated end joining (TMEJ) plays an essential role in DSB repair in *C. elegans* [99-103]. TMEJ is especially key to repair DSBs that occur due to varying persistent replication blocks. Worms that lack DOG-1 have increased loss of sequences that can form quadruplex structures. These structures cause replication associated breaks that are repaired via TMEJ, which leaves behind a very characteristic mutagenic footprint [99,101]. The same genomic scar is found in worms that lack TLS polymerases Polη and Polκ [100]. Subsequently, it became clear that the Polθ is the key player of a very versatile repair mechanisms that is employed widely in *C. elegans*, also for DSBs from other sources [102,103]. Polθ has the ability to capture a 3' ssDNA and stabilizes minimal (1 nt) base pairing, before it can extend to stabilize joining of the break ends. When this directly leads to repair a deletion results that is characterized by single nucleotide homology at the repair junction. On occasions that the newly synthesized sequence tract detaches prior to another round of annealing and extension a templated inserts results (Fig. 4).

Parallel to the research in worms, the interest in Polθ in mammals has greatly increased over the last years [extensively reviewed in 104,105]. Importantly, POLQ is

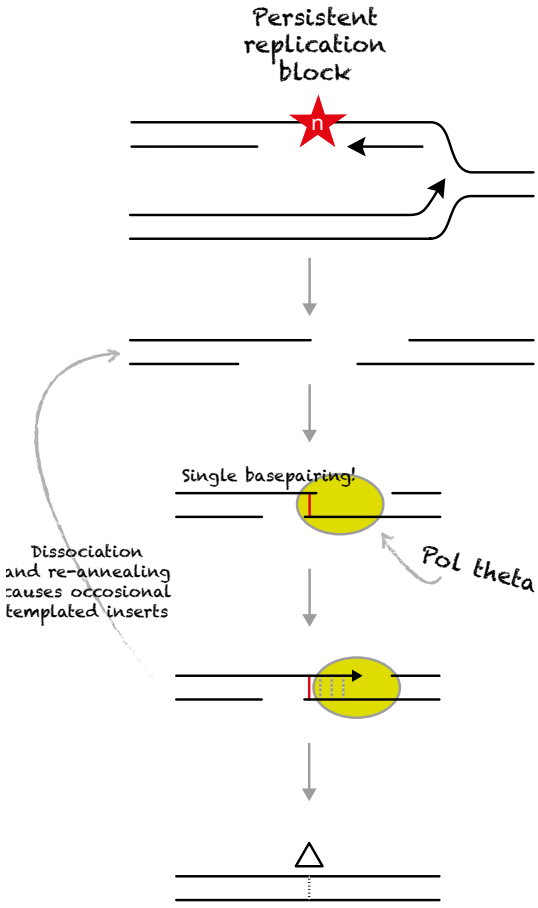


Figure 4. Simplified representation of polymerase theta mediated end joining. *For details see text.*

part of a group of upregulated genes correlated to reduced clinical outcome for a set of different cancers. More importantly HR defective tumor cells require Polθ for their growth [105]. Therefore, inhibition of Polθ is considered a promising approach to the treatment of (at least certain types of) cancer.

Methods that matter

Biological research has developed rapidly during recent history. There are several technological advances especially important for the work described in the experimental chapters that follow this introduction. I will emphasize a few that have been essential for my work, and will discuss their specific application. The discovery of DNA, its structure and function, starting almost 150 years ago, is at the base of these innovations. In the late 19th century, long before the link between chromosomes and heritability was proposed, Freidrich Miescher and Richard Altman isolated DNA and found that

it was part of chromosomes. In the early 20th century Phoebus Levene discovered the building blocks of DNA: A, T, G, and C [153]. Over the next decades the field of molecular biology leaped forward when the link between DNA and heritability became clear, when it was understood that DNA carried genes, the genetic code for the synthesis of proteins. The discovery of the double-helix structure of the DNA molecule by Watson and Crick is another iconic breakthrough of that era [154]. In the years that followed numerous advances have paved the way for modern genetic research and molecular biology. Among these are the development of techniques to determine DNA sequences and methods for genetic engineering. In addition, gaining a thorough understanding of the function and interactions of genes and their corresponding proteins would not have been possible without the use of model organisms, such as *C. elegans*.

Model organisms in genetics – Caenorhabditis elegans

In his book *Thus Spoke Zarathustra: A Book for All and None* from 1883, Friedrich Nietzsche wrote: "You have made your way from worm to man, but much within you is still worm." Although he was not discussing genetics or molecular biology, we now know that it also applies to these fields. Model organisms are widely used in biological and biomedical research, ranging from the humble *E. coli*, to plants like Arabidopsis, unicellular organisms such as yeast, multicellular animals such as *C. elegans*, fruit flies, zebra fish, up to complex mammalian models like rodents, human cell cultures and organoids. It is humbling to realize we are similar to these simple systems in so many ways, and these models have been, and will remain, incredibly valuable for genetic research.

In the sixties Sydney Brenner chose the nematode *C. elegans*, primarily to study neuronal development. Somewhat later in 1974, he established it as a genetic model with a much wider application [155]. There several characteristics that make it such a valuable model: an adult *C. elegans* has just 959 cells and an approximate length of only 1 mm; it grows and reproduces quickly and is inexpensive to maintain, within 3 days a fertilized egg grows into a fertile adult that can produce 200 to 300 offspring; and *C. elegans* can be grown easily on agar plates seeded with *E. coli* as food, at room temperature. When food becomes limited the worms go into a special survival stage (referred to as dauer stage) that can last for months. Also important: populations of worms can be frozen at -80°C for decades, if not longer. Typically, *C. elegans* exist as self-fertilizing hermaphrodites, but one in every thousand animals is born as a male. This provides an essential opportunity: males from one genetic background can be crossed to hermaphrodites from another, which allows researchers to combine mutations in different genes for example. In addition, the ability to self-fertilize allows the growth of large clonal populations. For many years, forward genetic screening and the generation of mutants have relied on random mutagenesis with DNA damaging agents. Most of

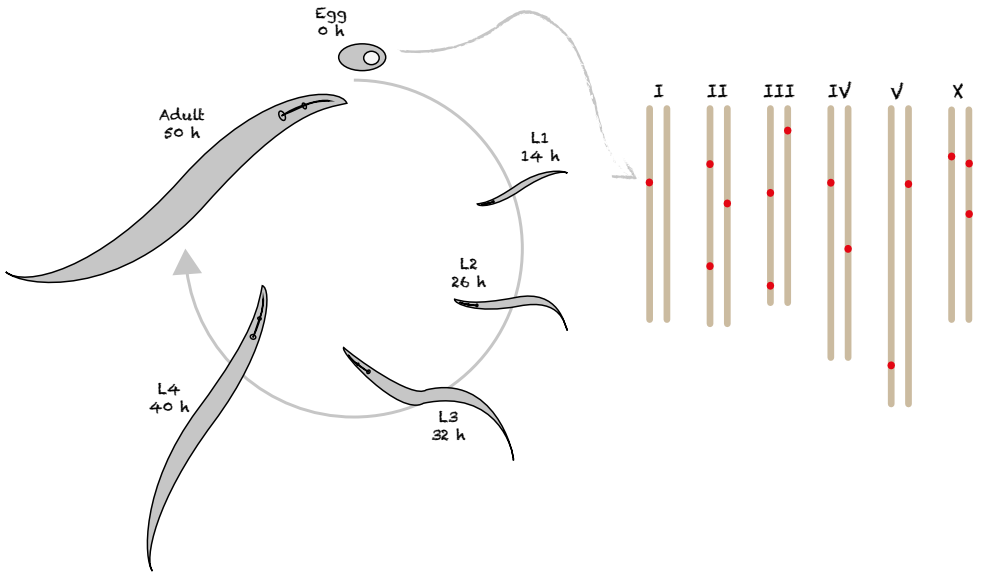


Figure 5. Mutation accumulation over multiple generations in C. elegans.

the generated strains are documented and available through the WormBase initiative and the number of available mutants exploded with the Million mutation project [156]. Alternatively, for reverse genetic screening genome wide RNAi libraries were widely used, and transgenic DNA can easily be introduced through microinjection. For many years targeted genome editing offered great difficulty, but recently this was overcome by the development of the CRISPR/Cas9 system and the application of it in *C. elegans* [157,158]. When in 1998 *C. elegans* became the first multi-cellular organism with a fully sequenced genome, it showed strong conservation of many genes and pathways. This includes the DNA repair pathways I discussed earlier and provides one of the many arguments why *C. elegans* is so well suited as a model to study DNA repair and the consequences of DNA damage [159-161].

All of the research in this thesis is performed in *C. elegans* and I exploited many of its benefits as a genetic model. For example, the high conservation of TLS, NER and POLQ and ease of crossing different genotypes served us well for chapter 2. In chapter 3 & 4 we made use of the quick life cycle and the possibility to grow clonal populations to investigate the accumulation of mutation over many generations (Fig. 5), while in chapter 5 we use microinjection of foreign DNA to study the repair of crosslinked DNA.

DNA sequencing: from 77 to 3.200.000.000 nucleotides and beyond

After it became clear that DNA holds the genetic code, the race was on to develop methods to determine the order of nucleotides in DNA and RNA molecules. Robert Holley was the first to do this and published the sequence of 77 ribonucleotides of alanine tRNA in 1965 [162]. Sequencing DNA proved to be more difficult, but in 1977 two scientists independently developed a chain termination technique to sequence DNA [163,164]. This first-generation sequencing technique, referred to as Sanger sequencing, is still widely used. For the research discussed in the following chapters Sanger sequencing has provided innumerable DNA repair footprints. Over the years that followed further advances in sequencing technique and computational power led to an iconic achievement; sequencing of the entire human genome of 3,200,000,000 bases was initiated in 1990, a first draft published in 2001, and completed in 2003 [165]. This project drove the development of next generation sequencing (NGS) techniques that are faster, cheaper and have increasingly higher throughput. The current NGS platforms can sequence up to hundreds of gigabases per run, and soon it will cost <$1000 to fully sequence an entire human genome. The dropping costs and increased sequencing speed and computational capacity have provided many new possibilities in genetic research and diagnostics and will continue to do so [166]. From the point of view of a researcher the identification of somatic mutations in tumor cells may lead to a deeper understanding of cancer, while for a patient it can be applied to assign personalized treatments that target the now identified weaknesses of the tumor. In chapter 3 & 4 I used NGS to identify novel mutations that have accumulated in the genome of *C. elegans* over a known number of generations, allowing me to identify mutational footprints and mutation frequency in a very unbiased way.

Genetic engineering

The development of our ability to read the genetic code goes hand in hand with advances to edit the genetic code. The first genetically modified viruses, bacteria and even the first transgenic mouse were already made in the seventies. Nowadays, transgenic plants are widely used as a food source, and applications of genome editing in medicine and research are countless. Therefore, I will limit myself to mention a few examples that are applied in *C. elegans* and relevant for the work described in this thesis.

Transgenic DNA can be introduced by microinjection into the gonadal syncytium of a young adult hermaphrodite. Multiple copies of the exogenous DNA then recombine to form an extrachromosomal array in the embryos. The genes on this array can, for example, rescue a genetic defect or to introduce a phenotype [167,168]. One of the most widely used phenotypic markers is green fluorescent protein (GFP). The gene that encodes GFP is from the bioluminescent jellyfish *Aequorea victoria*. Martin Chalfie was

the first to express the GFP coding sequence in another eukaryotic system: *C. elegans* [169], for which he received the Nobel Prize.

Precise and targeted editing of genomes has always been challenging, especially in *C. elegans*, but the discovery of the CRISPR/Cas9 system has taken biomedical research by storm [157,158]. CRISPR systems are adaptable immune mechanisms used by many bacteria. When a bacterium gets infected for the first time it incorporates a sequence from the invader into its own genomic CRISPR locus. Upon a next infection with the same pathogen, the invader sequence is transcribed and processed into a mature CRIPS RNA, which forms a complex with Cas9. This complex then targets the invading DNA and induces DSBs resulting in the destruction of the pathogen [170]. This system has been adapted by researchers and can now be used to induce targeted DSBs at almost any location of choice in a wide variety of organisms, ranging from *C. elegans* to mammals [Reviewed in 171]. The CRISPR/Cas9 technique can be applied to generate targeted gene knockouts or specific mutations, and to introduce foreign sequences at precise genomic locations. I applied this novel technology in the work described in this thesis; in chapters 3 and 4 I have used it to generate new knockout alleles of the REV-1 gene and to introduce a specific amino acid substitution mutation in the gene encoding PCNA (PCN-1 in *C. elegans*). Generating this PCNA mutation would not have been possible without CRISPR/Cas9.

The isolation and subsequent expression of GFP in different species and the discovery, development, and application of the CRISP/Cas9 system are both iconic victories for fundamental biological research. They show that studies in simple model organisms can have profound impact on applied research and medicine. Nowadays, gene therapy is already used to combat genetic disease and it can be expected that further development of sophisticated genome editing techniques will improve such treatments greatly. In the future we may come to rely more and more on genome editing techniques for our survival; further genetic adaptation of food crops may be necessary to sustain the ever-growing human population, for example. In any case, the field of genetic research will have a sensational and challenging future.

Aim and outline of this thesis

In this thesis I will argue that TLS is a mechanism that not only maintains replication potential and thereby promotes survival in the presence of DNA damage, but also preserves genome stability. When I started the research described in this thesis it became apparent that TMEJ is particularly important for the repair of replication associated breaks, and this thesis provides strong support of this notion.

Chapter 2 was initially titled *Nightcrawlers – When daylight kills*. In this chapter we show that a worm, when stripped of three DNA damage response mechanisms, completely loses the ability to live in the presence of regular daylight. It directly

illustrates that our natural environment on earth is inducing a harsh burden for DNA and it shows how essential the combined efforts of different DNA damage response mechanisms are, thus allowing organisms to live long and prosper in the face of DNA damaging environments.

In **chapter 3** we show that the TLS polymerase REV-1 protects genomic stability under challenged and unchallenged conditions and we confirm that Polθ mediated end-joining repairs replication associated DSBs in REV-1 deficient animals.

Chapter 4 provides evidence how TLS factors Polζ and PCNA also provide genome stability. In addition, we show that mutations accumulate in animals that lack all Y-family polymerases as well as Polζ, yet they are able to grow and reproduce.

In **chapter 5** I present a novel assay that I developed together with Robin van Schendel to specifically look at mutagenesis at single ICLs. The approach is based on the assays that have been used to study ICL repair in Xenopus egg extracts and provides a new tool to study the complex ICL repair mechanisms.

In the general discussion in **chapter 6** I will reflect on the conclusions of the experimental chapters and discuss how these impact our current understanding of evolution, and diseases like cancer and ageing. I will also consider the future perspectives of this research.

References

[1] Lindahl T. Instability and decay of the primary structure of DNA. Nature 1993;362:709–15. doi:10.1038/362709a0.

[2] Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. Toxicol Sci 2011;120 Suppl 1:S130–45. doi:10.1093/toxsci/kfq371.

[3] van Kregten M, Tijsterman M. The repair of G-quadruplex-induced DNA damage. Exp Cell Res 2014;329:178–83. doi:10.1016/j.yexcr.2014.08.038.

[4] Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell 2010;40:179–204. doi:10.1016/j.molcel.2010.09.019.

[5] Starcevic D, Dalal S, Sweasy JB. Is there a link between DNA polymerase beta and cancer? Cell Cycle 2004;3:998–1001. doi:10.4161/cc.3.8.1062.

[6] Farrington SM, Tenesa A, Barnetson R, Wiltshire A, Prendergast J, Porteous M, et al. Germline Susceptibility to Colorectal Cancer Due to Base-Excision Repair Gene Defects. The American Journal of Human Genetics 2005;77:112–9. doi:10.1086/431213.

[7] Baute J, Depicker A. Base excision repair and its role in maintaining genome stability. Crit Rev Biochem Mol Biol 2008;43:239–76. doi:10.1080/10409230802309905.

[8] Krokan HE, Bjoras M. Base Excision Repair. Cold Spring Harb Perspect Biol 2013;5:a012583–3. doi:10.1101/cshperspect.a012583.

[9] Caldecott KW. Single-strand break repair and genetic disease. Nat Rev Genet 2008;9:619–31. doi:10.1038/nrg2380.

[10] Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JHJ. Understanding nucleotide excision repair and its roles in cancer and ageing. Nature Reviews Molecular Cell Biology 2014;15:465–81. doi:10.1038/nrm3822.

[11] Friedberg EC. How nucleotide excision repair protects against cancer. Nat Rev Cancer 2001;1:22–33. doi:10.1038/35094000.

[12] McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res 2008;18:148–61. doi:10.1038/cr.2008.4.

[13] Peña-Diaz J, Jiricny J. Mammalian mismatch repair: error-free or error-prone? Trends in Biochemical Sciences 2012;37:206–14. doi:10.1016/j.tibs.2012.03.001.

[14] Fishel R. Mismatch repair. Journal of Biological Chemistry 2015;290:26395–403. doi:10.1074/jbc.R115.660142.

[15] Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RMW, et al. Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. Hum Mutat 2012;33:1617–25. doi:10.1002/humu.22168.

[16] Blastyák A, Pintér L, Unk I, Prakash L, Prakash S, Haracska L. Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell 2007;28:167–75. doi:10.1016/j.molcel.2007.07.030.

[17] Atkinson J, McGlynn P. Replication fork reversal and the maintenance of genome stability. Nucleic Acids Res 2009;37:3475–92. doi:10.1093/nar/gkp244.

[18] Unk I, Hajdu I, Blastyák A, Haracska L. Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair 2010;9:257–67. doi:10.1016/j.dnarep.2009.12.013.

[19] Lin J-R, Zeman MK, Chen J-Y, Yee M-C, Cimprich KA. SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. Mol Cell 2011;42:237–49. doi:10.1016/j.molcel.2011.02.026.

[20] Fan Q, Xu X, Zhao X, Wang Q, Xiao W, Guo Y, et al. Rad5 coordinates translesion DNA synthesis pathway by recognizing specific DNA structures in saccharomyces cerevisiae. Curr Genet 2018;62:405–11. doi:10.1007/s00294-018-0807-y.

[21] Lee JA, Carvalho C, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell 2007.

[22] Branzei D, Foiani M. Template switching: from replication fork repair to genome rearrangements. Cell 2007;131:1228–30. doi:10.1016/j.cell.2007.12.007.

[23] Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, et al. The Y-family of DNA polymerases. Mol Cell 2001;8:7–8. doi:10.1016/S1097-2765(01)00278-7.

[24] Prakash S, Johnson RE, Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem 2005;74:317–53. doi:10.1146/annurev.biochem.74.082803.133250.

[25] Knobel PA, Marti TM. Translesion DNA synthesis in the context of cancer research. Cancer Cell Int 2011;11:39. doi:10.1186/1475-2867-11-39.

[26] Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. Transl Cancer Res 2013;2:107–29. doi:10.3978/j.issn.2218-676X.2013.04.01.

[27] Jansen JG, Tsaalbi-Shtylik A, de Wind N. Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. DNA Repair 2015;29:56–64. doi:10.1016/j.dnarep.2015.01.001.

[28] Gan GN, Wittschieben JP, Wittschieben BØ, Wood RD. DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res 2008;18:174–83. doi:10.1038/cr.2007.117.

[29] Sale JE, Lehmann AR, Woodgate R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Reviews Molecular Cell Biology 2012;13:141–52. doi:10.1038/nrm3289.

[30] Prakash S, Prakash L. Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. Genes & Development 2002;16:1872–83. doi:10.1101/gad.1009802.

[31] Lehmann AR, Fuchs RP. Gaps and forks in DNA replication: Rediscovering old models. DNA Repair 2006;5:1495–8. doi:10.1016/j.dnarep.2006.07.002.

[32] McDonald JP, Levine AS, Woodgate R. The Saccharomyces cerevisiae RAD30 Gene, a Homologue of Escherichia coli dinB and umuC, Is DNA Damage Inducible and Functions in a Novel Error-Free Postreplication Repair Mechanism. Genetics 1997;147:1557–68.

[33] Johnson RE, Prakash S, Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 1999;283:1001–4.

[34] Johnson RE, Haracska L, Prakash S, Prakash L. Role of DNA polymerase eta in the bypass of a (6-4) TT photoproduct. Mol Cell Biol 2001;21:3558–63. doi:10.1128/MCB.21.10.3558-3563.2001.

[35] Johnson RE, Kondratick CM, Prakash S, Prakash L. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 1999;285:263–5. doi:10.1126/science.285.5425.263.

[36] De Palma A, Morren M-A, Ged C, Pouvelle C, Taïeb A, Aoufouchi S, et al. Diagnosis of Xeroderma pigmentosum variant in a young patient with two novel mutations in the POLH gene. Am J Med Genet A 2017. doi:10.1002/ajmg.a.38340.

[37] Haracska L, Yu SL, Johnson RE, Prakash L, Prakash S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat Genet 2000;25:458–61. doi:10.1038/78169.

[38] Lee D-H, Pfeifer GP. Translesion synthesis of 7,8-dihydro-8-oxo-2'-deoxyguanosine by DNA polymerase eta in vivo. Mutat Res 2008;641:19–26. doi:10.1016/j.mrfmmm.2008.02.006.

[39] Zhang Y, Yuan F, Wu X, Rechkoblit O, Taylor JS, Geacintov NE, et al. Error-prone lesion bypass by human DNA polymerase eta. Nucleic Acids Res 2000;28:4717–24. doi:10.1093/nar/28.23.4717.

[40] Yuan F, Zhang Y, Rajpal DK, Wu X, Guo D, Wang M, et al. Specificity of DNA lesion bypass by the yeast DNA polymerase eta. J Biol Chem 2000;275:8233–9. doi:10.1074/jbc.275.11.8233.

[41] Haracska L, Prakash S, Prakash L. Replication pastO 6-Methylguanine by Yeast and Human DNA Polymerase η. Molecular and Cellular ... 2000. doi:10.1128/MCB.20.21.8001-8007.2000.

[42] Kusumoto R, Masutani C, Iwai S, Hanaoka F. Translesion synthesis by human DNA polymerase η across thymine glycol lesions. Biochemistry 2002. doi:10.1021/bi025549k.

[43] Vaisman A, Masutani C, Hanaoka F, Chaney SG. Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase η. Biochemistry 2000.

[44] Choi J-H, Pfeifer GP. The role of DNA polymerase eta in UV mutational spectra. DNA Repair 2005;4:211–20. doi:10.1016/j.dnarep.2004.09.006.

[45] Roush AA, Suarez M, Friedberg EC, Radman M, Siede W. Deletion of the Saccharomyces cerevisiae gene RAD30 encoding an Escherichia coli DinB homolog confers UV radiation sensitivity and altered mutability. Mol Gen Genet 1998;257:686–92. doi:10.1007/s004380050698.

[46] Jansen JG, Temviriyanukul P, Wit N, Delbos F, Reynaud C-A, Jacobs H, et al. Redundancy of mammalian Y family DNA polymerases in cellular responses to genomic DNA lesions induced by ultraviolet light. Nucleic Acids Res 2014;42:11071–82. doi:10.1093/nar/gku779.

[47] King NM, Nikolaishvili-Feinberg N, Bryant MF, Luche DD, Heffernan TP, Simpson DA, et al. Overproduction of DNA polymerase eta does not raise the spontaneous mutation rate in diploid human fibroblasts. DNA Repair 2005;4:714–24. doi:10.1016/j.dnarep.2005.04.002.

[48] Pavlov YI, Nguyen D, Kunkel TA. Mutator effects of overproducing DNA polymerase eta (Rad30) and its catalytically inactive variant in yeast. Mutat Res 2001;478:129–39.

[49] Haracska L, Kondratick CM, Unk I, Prakash S, Prakash L. Interaction with PCNA is essential for yeast DNA polymerase eta function. Mol Cell 2001;8:407–15.

[50] Guo C, Fischhaber PL, Luk-Paszyc MJ, Masuda Y, Zhou J, Kamiya K, et al. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. Embo J 2003;22:6621–30. doi:10.1093/emboj/cdg626.

[51] Edmunds CE, Simpson LJ, Sale JE. PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell 2008;30:519–29. doi:10.1016/j.molcel.2008.03.024.

[52] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and ϰ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.

[53] Brauchle M, Baumer K, Gönczy P. Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in *C. elegans* embryos. Curr Biol 2003;13:819–27. doi:10.1016/S0960-9822(03)00295-1.

[54] Holway AH, Kim S-H, La Volpe A, Michael WM. Checkpoint silencing during the DNA damage response in Caenorhabditis elegansembryos. J Cell Biol 2006;172:999–1008. doi:10.1093/nar/gkh710.

[55] Gerlach VL, Aravind L, Gotway G, Schultz RA, Koonin EV, Friedberg EC. Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. Proc Natl Acad Sci U S A 1999;96:11922–7.

[56] Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff RV, Walker GC. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiol Mol Biol Rev 2009;73:134–54. doi:10.1128/MMBR.00034-08.

[57] Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC. A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. Nature 2006;439:225–8. doi:10.1038/nature04318.

[58] Yuan B, Cao H, Jiang Y, Hong H, Wang Y. Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase in vitro and in vivo. Proc Natl Acad Sci U S A 2008;105:8679–84. doi:10.1073/pnas.0711546105.

[59] Avkin S, Goldsmith M, Velasco-Miguel S, Geacintov N, Friedberg EC, Livneh Z. Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase kappa. J Biol Chem 2004;279:53298–305. doi:10.1074/jbc.M409155200.

[60] Ogi T, Shinkai Y, Tanaka K, Ohmori H. Polkappa protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. Proc Natl Acad Sci U S A 2002;99:15548–53. doi:10.1073/pnas.222377899.

[61] Takenaka K, Ogi T, Okada T, Sonoda E, Guo C, Friedberg EC, et al. Involvement of vertebrate Polkappa in translesion DNA synthesis across DNA monoalkylation damage. J Biol Chem 2006;281:2000–4. doi:10.1074/jbc.M506153200.

[62] Andersen PL, Xu F, Xiao W. Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. Cell Res 2008;18:162–73. doi:10.1038/cr.2007.114.

[63] Kirouac KN, Ling H. Unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase iota. Proc Natl Acad Sci USA 2011;108:3210–5. doi:10.1073/pnas.1013909108.

[64] Petta TB, Nakajima S, Zlatanou A, Despras E, Couve-Privat S, Ishchenko A, et al. Human DNA polymerase iota protects cells against oxidative stress. Embo J 2008;27:2883–95. doi:10.1038/emboj.2008.210.

[65] Lemontt JF. Mutants of yeast defective in mutation induced by ultraviolet light. Genetics 1971;68:21–33.

[66] Lawrence CW, Krauss BR, Christensen RB. New mutations affecting induced mutagenesis in yeast. Mutat Res 1985;150:211–6.

[67] Larimer FW, Perry JR, Hardigree AA. The REV1 gene of Saccharomyces cerevisiae: isolation, sequence, and functional analysis. J Bacteriol 1989;171:230–7.

[68] Nelson JR, Lawrence CW, Hinkle DC. Deoxycytidyl transferase activity of yeast REV1 protein. Nature 1996;382:729–31. doi:10.1038/382729a0.

[69] Haracska L, Unk I, Johnson RE, Johansson E, Burgers PM, Prakash S, et al. Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. Genes & Development 2001;15:945–54. doi:10.1101/gad.882301.

[70] Zhang Y, Wu X, Rechkoblit O, Geacintov NE, Taylor J-S, Wang Z. Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Res 2002;30:1630–8. doi:10.1093/nar/30.7.1630.

[71] Zhou Y, Wang J, Zhang Y, Wang Z. The catalytic function of the Rev1 dCMP transferase is required in a lesion-specific manner for translesion synthesis and base damage-induced mutagenesis. Nucleic Acids Res 2010;38:5036–46. doi:10.1093/nar/gkq225.

[72] Wiltrout ME, Walker GC. The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. Genetics 2011;187:21–35. doi:10.1534/genetics.110.124172.

[73] Gibbs PE, Wang XD, Li Z, McManus TP, McGregor WG, Lawrence CW, et al. The function of the human homolog of Saccharomyces cerevisiae REV1 is required for mutagenesis induced by UV light. Proc Natl Acad Sci U S A 2000;97:4186–91. doi:10.1073/pnas.97.8.4186.

[74] Guo C, Sonoda E, Tang T-S, Parker JL, Bielen AB, Takeda S, et al. REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. Mol Cell 2006;23:265–71. doi:10.1016/j.molcel.2006.05.038.

[75] Jansen JG, Tsaalbi-Shtylik A, Langerak P, Calléja F, Meijers CM, Jacobs H, et al. The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Res 2005;33:356–65. doi:10.1093/nar/gki189.

[76] Jansen JG, Tsaalbi-Shtylik A, Hendriks G, Gali H, Hendel A, Johansson F, et al. Separate domains of Rev1 mediate two modes of DNA damage bypass in mammalian cells. Mol Cell Biol 2009;29:3113–23. doi:10.1128/MCB.00071-09.

[77] Kikuchi S, Hara K, Shimizu T, Sato M, Hashimoto H. Structural Basis of Recruitment of DNA Polymerase by Interaction between REV1 and REV7 Proteins. Journal of Biological Chemistry 2012;287:33847–52. doi:10.1074/jbc.M112.396838.

[78] Wojtaszek J, Liu J, D'Souza S, Wang S, Xue Y, Walker GC, et al. Multifaceted recognition of vertebrate Rev1 by translesion polymerases ζ and κ. Journal of Biological Chemistry 2012;287:26400–8. doi:10.1074/jbc.M112.380998.

[79] Wojtaszek J, Lee CJ, D'Souza S, Minesinger B, Kim H, D'Andrea AD, et al. Structural Basis of Rev1-mediated Assembly of a Quaternary Vertebrate Translesion Polymerase Complex Consisting of Rev1, Heterodimeric Polymerase (Pol) , and Pol. Journal of Biological Chemistry 2012;287:33836–46. doi:10.1074/jbc.M112.394841.

[80] Sarkies P, Reams C, Simpson LJ, Sale JE. Epigenetic instability due to defective replication of structured DNA. Mol Cell 2010;40:703–13. doi:10.1016/j.molcel.2010.11.009.

[81] Schiavone D, Guilbaud G, Murat P, Papadopoulou C, Sarkies P, Prioleau M-N, et al. Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. Embo J 2014;33:2507–20. doi:10.15252/embj.201488398.

[82] Okada T, Sonoda E, Yoshimura M, Kawano Y, Saya H,

Kohzaki M, et al. Multiple Roles of Vertebrate REV Genes in DNA Repair and Recombination. Mol Cell Biol 2005;25:6103–11. doi:10.1128/MCB.25.14.6103-6111.2005.

[83] Sharma S, Hicks JK, Chute CL, Brennan JR, Ahn J-Y, Glover TW, et al. REV1 and polymerase ζ facilitate homologous recombination repair. Nucleic Acids Res 2012;40:682–91. doi:10.1093/nar/gkr769.

[84] Nelson JR, Lawrence CW, Hinkle DC. Thymine-thymine dimer bypass by yeast DNA polymerase zeta. Science 1996;272:1646–9. doi:10.1126/science.272.5268.1646.

[85] Lawrence CW, Das G, Christensen RB. REV7, a new gene concerned with UV mutagenesis in yeast. Mol Gen Genet 1985;200:80–5.

[86] Baranovskiy AG, Lada AG, Siebler HM, Zhang Y, Pavlov YI, Tahirov TH. DNA Polymerase and Switch by Sharing Accessory Subunits of DNA Polymerase. Journal of Biological Chemistry 2012;287:17281–7. doi:10.1074/jbc.M112.351122.

[87] Johnson RE, Prakash L, Prakash S. Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. Proc Natl Acad Sci U S A 2012;109:12455–60. doi:10.1073/pnas.1206052109.

[88] Makarova AV, Stodola JL, Burgers PM. A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. Nucleic Acids Res 2012;40:11618–26. doi:10.1093/nar/gks948.

[89] Lee Y-S, Gregory MT, Yang W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc Natl Acad Sci U S A 2014;111:2954–9. doi:10.1073/pnas.1324001111.

[90] Nelson JR, Gibbs PE, Nowicka AM, Hinkle DC, Lawrence CW. Evidence for a second function for Saccharomyces cerevisiae Rev1p. Mol Microbiol 2000;37:549–54.

[91] Bemark M, Khamlichi AA, Davies SL, Neuberger MS. Disruption of mouse polymerase zeta (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. Curr Biol 2000;10:1213–6. doi:10.1016/S0960-9822(00)00724-7.

[92] Tomas-Roca L, Tsaalbi-Shtylik A, Jansen JG, Singh MK, Epstein JA, Altunoglu U, et al. De novo mutations in PLXND1 and REV3L cause Möbius syndrome. Nat Comms 2015;6:7199. doi:10.1038/ncomms8199.

[93] Hoffmann JS, Pillaire MJ, Maga G, Podust V, Hübscher U, Villani G. DNA polymerase beta bypasses in vitro a single d(GpG)-cisplatin adduct placed on codon 13 of the HRAS gene. Proc Natl Acad Sci U S A 1995;92:5356–60.

[94] Seki M, Masutani C, Yang LW, Schuffert A, Iwai S, Bahar I, et al. High-efficiency bypass of DNA damage by human DNA polymerase Q. Embo J 2004;23:4484–94. doi:10.1038/sj.emboj.7600424.

[95] Zhang Y, Wu X, Yuan F, Xie Z, Wang Z. Highly frequent frameshift DNA synthesis by human DNA polymerase mu. Mol Cell Biol 2001;21:7995–8006. doi:10.1128/MCB.21.23.7995-8006.2001.

[96] Picher AJ, García-Díaz M, Bebenek K, Pedersen LC, Kunkel TA, Blanco L. Promiscuous mismatch extension by human DNA polymerase lambda. Nucleic Acids Res 2006;34:3259–66. doi:10.1093/nar/gkl377.

[97] Seki M, Wood RD. DNA polymerase theta (POLQ) can extend from mismatches and from bases opposite a (6-4) photoproduct. DNA Repair 2008;7:119–27. doi:10.1016/j.dnarep.2007.08.005.

[98] Chan SH, Yu AM, McVey M. Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. PLoS Genet 2010;6:e1001005. doi:10.1371/journal.pgen.1001005.

[99] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.

[100] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.

[101] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.

[102] van Schendel R, Roerink SF, Portegijs V, van den Heuvel S, Tijsterman M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat Comms 2015;6:7394. doi:10.1038/ncomms8394.

[103] van Schendel R, van Heteren J, Welten R, Tijsterman M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.

[104] Yousefzadeh MJ, Wood RD. DNA polymerase POLQ and cellular defense against DNA damage. DNA Repair 2013;12:1–9. doi:10.1016/j.dnarep.2012.10.004.

[105] Wood RD, Doublié S. DNA polymerase θ (POLQ), double-strand break repair, and cancer. DNA Repair 2016. doi:10.1016/j.dnarep.2016.05.003.

[106] Cipolla L, Maffia A, Bertoletti F, Sabbioneda S. The Regulation of DNA Damage Tolerance by Ubiquitin and Ubiquitin-Like Modifiers. Front Gene 2016;7:105. doi:10.3389/fgene.2016.00105.

[107] Mailand N, Gibbs-Seymour I, Bekker-Jensen S. Regulation of PCNA-protein interactions for genome stability. Nature Reviews Molecular Cell Biology 2013;14:269–82. doi:10.1038/nrm3562.

[108] Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 2002;419:135–41. doi:10.1038/nature00991.

[109] Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, et al. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science 2005;310:1821–4. doi:10.1126/science.1120615.

[110] Branzei D. Ubiquitin family modifications and template switching. FEBS Lett 2011;585:2810–7. doi:10.1016/j.febslet.2011.04.053.

[111] Hendel A, Krijger PHL, Diamant N, Goren Z, Langerak P, Kim J, et al. PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLoS Genet 2011;7:e1002262. doi:10.1371/journal.pgen.1002262.

[112] Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harb Perspect Biol 2013;5:a012708. doi:10.1101/cshperspect.a012708.

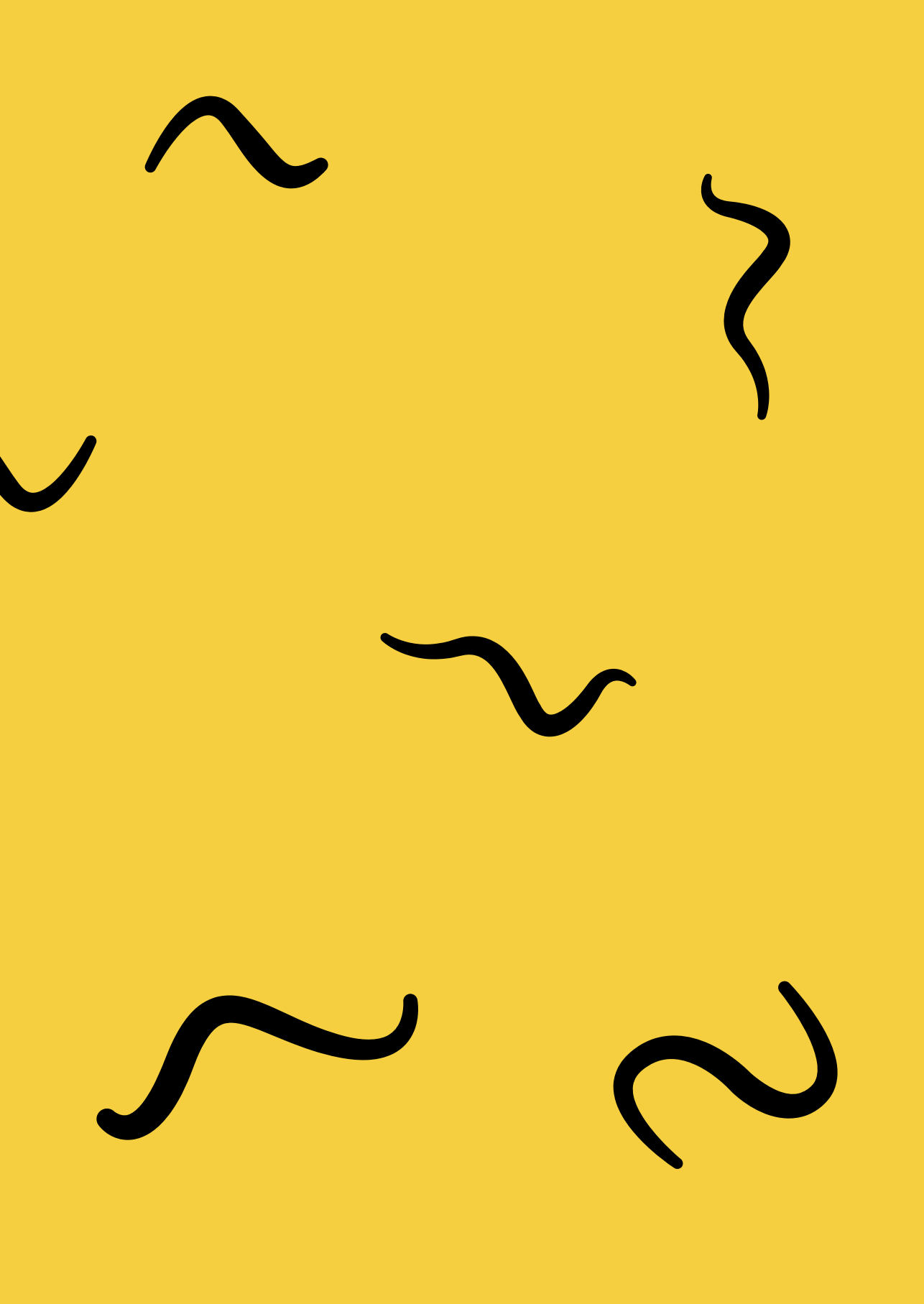
[113] Jansen JG, Fousteri MI, de Wind N. Send in the clamps: control of DNA translesion synthesis in eukaryotes. Mol Cell 2007;28:522–9. doi:10.1016/j.molcel.2007.11.005.

[114] Hung S-H, Wong RP, Ulrich HD, Kao C-F. Monoubiquitylation of histone H2B contributes to the bypass of DNA damage during and after DNA replication. Proc Natl Acad Sci U S A 2017. doi:10.1073/pnas.1612633114.

[115] Shachar S, Ziv O, Avkin S, Adar S, Wittschieben J,

Reißner T, et al. Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. Embo J 2009;28:383–93. doi:10.1038/emboj.2008.281.
[116] Temviriyanukul P, van Hees-Stuivenberg S, Delbos F, Jacobs H, de Wind N, Jansen JG. Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair 2012;11:550–8. doi:10.1016/j.dnarep.2012.03.007.
[117] Livneh Z, Cohen IS, Paz-Elizur T, Davidovsky D, Carmi D, Swain U, et al. High-resolution genomic assays provide insight into the division of labor between TLS and HDR in mammalian replication of damaged DNA. DNA Repair 2016. doi:10.1016/j.dnarep.2016.05.007.
[118] Zan H, Komori A, Li Z, Cerutti A, Schaffer A, Flajnik MF, et al. The translesion DNA polymerase zeta plays a major role in Ig and bcl-6 somatic hypermutation. Immunity 2001;14:643–53. doi:10.1016/S1074-7613(01)00142-X.
[119] Faili A, Aoufouchi S, Flatter E, Guéranger Q, Reynaud C-A, Weill J-C. Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase iota. Nature 2002;419:944–7. doi:10.1038/nature01117.
[120] Jansen JG, Langerak P, Tsaalbi-Shtylik A, van den Berk P, Jacobs H, de Wind N. Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. J Exp Med 2006;203:319–23. doi:10.1084/jem.20052227.
[121] Lehmann AR, Kirk-Bell S, Arlett CF, Paterson MC, Lohman PH, de Weerd-Kastelein EA, et al. Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. Proc Natl Acad Sci U S A 1975;72:219–23.
[122] Waters HL, Seetharam S, Seidman MM, Kraemer KH. Ultraviolet hypermutability of a shuttle vector propagated in xeroderma pigmentosum variant cells. J Invest Dermatol 1993;101:744–8.
[123] Masutani C, Araki M, Yamada A, Kusumoto R, Nogimori T, Maekawa T, et al. Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. Embo J 1999;18:3491–501. doi:10.1093/emboj/18.12.3491.
[124] Tsaalbi-Shtylik A, Verspuy JWA, Jansen JG, Rebel H, Carlée LM, van der Valk MA, et al. Error-prone translesion replication of damaged DNA suppresses skin carcinogenesis by controlling inflammatory hyperplasia. Proc Natl Acad Sci U S A 2009;106:21836–41. doi:10.1073/pnas.0909507106.
[125] Makridakis NM, Reichardt JKV. Translesion DNA Polymerases and Cancer. Front Gene 2012;3. doi:10.3389/fgene.2012.00174.
[126] Song L, McNeil EM, Ritchie A-M, Astell KR, Gourley C, Melton DW. Melanoma cells replicate through chemotherapy by reducing levels of key homologous recombination protein RAD51 and increasing expression of translesion synthesis DNA polymerase ζ. BMC Cancer 2017;17:864. doi:10.1186/s12885-017-3864-6.
[127] Zafar MK, Eoff RL. Translesion DNA Synthesis in Cancer: Molecular Mechanisms and Therapeutic Opportunities. Chem Res Toxicol 2017;30:1942–55. doi:10.1021/acs.chemrestox.7b00157.
[128] Vanarotti M, Evison BJ, Actis ML, Inoue A, McDonald ET, Shao Y, et al. Small-molecules that bind to the ubiquitin-binding motif of REV1 inhibit REV1 interaction with K164-monoubiquitinated PCNA and suppress DNA damage tolerance. Bioorg Med Chem 2018. doi:10.1016/j.bmc.2018.03.028.
[129] Choi J-S, Kim CS, Berdis A. Inhibition of Translesion DNA Synthesis as a Novel Therapeutic Strategy to Treat Brain Cancer. Cancer Res 2018;78:1083–96. doi:10.1158/0008-5472.CAN-17-2464.
[130] Choi J-S, Kim S, Motea E, Berdis A. Inhibiting translesion DNA synthesis as an approach to combat drug resistance to DNA damaging agents. Oncotarget 2017;8:40804–16. doi:10.18632/oncotarget.17254.
[131] Yamanaka K, Chatterjee N, Hemann MT, Walker GC. Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? PLoS Genet 2017;13:e1006842. doi:10.1371/journal.pgen.1006842.
[132] Zafar MK, Maddukuri L, Ketkar A, Penthala NR, Reed MR, Eddy S, et al. A Small-Molecule Inhibitor of Human DNA Polymerase η Potentiates the Effects of Cisplatin in Tumor Cells. Biochemistry 2018;57:1262–73. doi:10.1021/acs.biochem.7b01176.
[133] Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer 2011;11:467–80. doi:10.1038/nrc3088.
[134] Scharer OD. DNA Interstrand Crosslinks: Natural and Drug-Induced DNA Adducts that Induce Unique Cellular Responses. ChemBioChem 2005.
[135] Niedernhofer LJ, Lalai AS, Hoeijmakers JHJ. Fanconi anemia (cross)linked to DNA repair. Cell 2005;123:1191–8. doi:10.1016/j.cell.2005.12.009.
[136] Michl J, Zimmer J, Tarsounas M. Interplay between Fanconi anemia and homologous recombination pathways in genome integrity. Embo J 2016;35:909–23. doi:10.15252/embj.201693860.
[137] Sharma S, Canman CE. REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. Environ Mol Mutagen 2012;53:725–40. doi:10.1002/em.21736.
[138] Roy U, Schärer OD. Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. DNA Repair 2016;44:33–41. doi:10.1016/j.dnarep.2016.05.004.
[139] Räschle M, Knipscheer P, Knipsheer P, Enoiu M, Angelov T, Sun J, et al. Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 2008;134:969–80. doi:10.1016/j.cell.2008.08.030.
[140] Knipscheer P, Räschle M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 2009;326:1698–701. doi:10.1126/science.1182372.
[141] Long DT, Räschle M, Joukov V, Walter JC. Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science 2011;333:84–7. doi:10.1126/science.1204258.
[142] Budzowska M, Graham TGW, Sobeck A, Waga S, Walter JC. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. Embo J 2015;34:1971–85. doi:10.15252/embj.201490878.
[143] Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. Cell 2016;167:498–511.e14. doi:10.1016/j.cell.2016.09.008.
[144] Williams HL, Gottesman ME, Gautier J. The differences between ICL repair during and outside of S phase. Trends in Biochemical Sciences 2013;38:386–93. doi:10.1016/j.tibs.2013.05.004.
[145] Williams HL, Gottesman ME, Gautier J. Replication-independent repair of DNA interstrand crosslinks. Mol Cell 2012;47:140–7. doi:10.1016/j.molcel.2012.05.001.
[146] Aparicio T, Baer R, Gautier J. DNA double-strand break repair pathway choice and cancer. DNA Repair 2014;19:169–75. doi:10.1016/j.dnarep.2014.03.014.

[147] Liu Y, West SC. Happy Hollidays: 40th anniversary of the Holliday junction. Nature Reviews Molecular Cell Biology 2004;5:937–44. doi:10.1038/nrm1502.
[148] Lieber MR, Lu H, Gu J, Schwarz K. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. Cell Res 2008;18:125–33. doi:10.1038/cr.2007.108.
[149] Gu J, Lieber MR. Mechanistic flexibility as a conserved theme across 3 billion years of nonhomologous DNA end-joining. Genes & Development 2008;22:411–5. doi:10.1101/gad.1646608.
[150] O'Driscoll M. Diseases associated with defective responses to DNA damage. Cold Spring Harb Perspect Biol 2012;4. doi:10.1101/cshperspect.a012773.
[151] McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet 2008;24:529–38. doi:10.1016/j.tig.2008.08.007.
[152] Pontier DB, Tijsterman M. A robust network of double-strand break repair pathways governs genome integrity during *C. elegans* development. Curr Biol 2009;19:1384–8. doi:10.1016/j.cub.2009.06.045.
[153] Dahm R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. Hum Genet 2007;122:565–81. doi:10.1007/s00439-007-0433-0.
[154] WATSON JD, CRICK FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953;171:737–8.
[155] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.
[156] Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, et al. The million mutation project: a new approach to genetics in Caenorhabditis elegans. Genome Research 2013;23:1749–62. doi:10.1101/gr.157651.113.
[157] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012;337:816–21. doi:10.1126/science.1225829.
[158] Waaijers S, Portegijs V, Kerver J, Lemmens BBLG, Tijsterman M, van den Heuvel S, et al. CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics 2013;195:1187–91. doi:10.1534/genetics.113.156299.
[159] Youds JL, Barber LJ, Boulton SJ. *C. elegans*: a model of Fanconi anemia and ICL repair. Mutat Res 2009;668:103–16. doi:10.1016/j.mrfmmm.2008.11.007.
[160] Lans H, Vermeulen W. Nucleotide Excision Repair in Caenorhabditis elegans. Molecular Biology International 2011;2011:1–12. doi:10.1111/j.1574-6976.2008.00126.x.
[161] van Haaften G, Plasterk RHA, Tijsterman M. Genomic instability and cancer: scanning the Caenorhabditis elegans genome for tumor suppressors. Oncogene 2004;23:8366–75. doi:10.1038/sj.onc.1208011.
[162] HOLLEY RW, EVERETT GA, MADISON JT, ZAMIR A. NUCLEOTIDE SEQUENCES IN THE YEAST ALANINE TRANSFER RIBONUCLEIC ACID. J Biol Chem 1965;240:2122–8.
[163] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977;74:5463–7.
[164] Maxam AM, Gilbert W. A new method for sequencing DNA. Proc Natl Acad Sci U S A 1977;74:560–4.
[165] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921. doi:10.1038/35057062.
[166] Metzker ML. Sequencing technologies — the next generation. Nat Rev Genet 2010;11:31. doi:10.1038/nrg2626.
[167] Way JC, Chalfie M. mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. Cell 1988.
[168] Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. Embo J 1991;10:3959–70.
[169] Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science 1994.
[170] Fineran PC, Charpentier E. Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. Virology 2012;434:202–9. doi:10.1016/j.virol.2012.10.003.
[171] Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 2014;32:347–55. doi:10.1038/nbt.2842.



2

Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light

Ivo van Bostelen & Marcel Tijsterman

Published as:
Van Bostelen I, Tijsterman M. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light. *DNA Repair* 2017

Abstract

Infliction of DNA damage initiates a complex cellular reaction - the DNA damage response - that involves both signaling and DNA repair networks with many redundancies and parallel pathways. Here, we reveal a surprisingly simple, straightforward, linear order of events for how a simple multicellular eukaryote, *C. elegans*, deals with DNA damage induced by light. Separately inactivating repair or replicative bypass of photo-lesions results in cellular hypersensitivity towards UV-light, but impeding repair of replication associated DNA breaks does not. Yet, we observe an unprecedented synergistic relationship when these pathways are inactivated in combination. *C. elegans* mutants that lack nucleotide excision repair (NER), translesion synthesis (TLS) and alternative end joining (alt-EJ) grow undisturbed in the dark, but become sterile when grown in light. Even exposure to very low levels of normal daylight impedes animal growth. We show that NER and TLS operate to suppress the formation of lethal DNA breaks that require polymerase theta-mediated end joining (TMEJ) for their repair. Our data testifies to the enormous genotoxicity of light and to the demand of multiple layers of protection against an environmental threat that is so common.

Introduction

Already 140 years ago Downes & Blunt showed that light can be toxic to organisms [1]. Plane daylight is, in fact, one of the most common and serious environmental sources of DNA damage [2,3]. The ultra violet (UV) component in sunlight induces chemical alterations of the bases in a DNA strand [4]. UV light is categorized in UV-C (λ= 100-280 nm), UV-B (λ= 280-315 nm) and UV-A (λ= 315-400 nm). UV-C and UV-B efficiently induce dimerization of adjacent pyrimidines in a DNA strand in the form of cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) [4]. While UV-C is widely used in laboratories to induce DNA damage and many studies have shown that UV-C photolesions efficiently block gene transcription and DNA replication, normal sunlight that reaches the earth's surface contains no UV-C because the atmosphere serves as an efficient filter [5]. At the earth's surface the UV spectrum of sunlight consists of 5-10% UV-B and 90-95% UV-A. While UV-A can induce CPDs and low levels of 6-4PPs and Dewar valence isomers [5-7], it does this at a much lower efficiency than UV-B or UV-C [8-10]. UV-A, UV-B and short wavelength visible light can also induce DNA damage indirectly via the formation of reactive oxygen species that cause oxidation of bases - most abundantly in the form of 7,8-dihydro-8-oxo-2′-deoxyguanosine (8oxoG) - which can also result in genomic instability and cytotoxicity. [11-17].

Nucleotide excision repair (NER) is considered the cell's main defense strategy against the toxic effects of photolesions. There are two sub-pathways of NER that are different in the way they recognize UV lesions [18]. Transcription-coupled NER (TC-NER) acts on lesions that block ongoing transcription [19], whereas global genome NER (GG-NER) operates on damage in non-transcribed DNA [18]. Both sub-pathways recruit the same downstream NER factors, which excise lesions from a DNA strand together with approximately 30 base pairs of flanking sequence [20-26]. Subsequently, this gap is filled by DNA polymerase and DNA ligase activities [18,27]. Genetic defects in factors that are specific for TC-NER cause Cockayne syndrome (CS), which is hallmarked by a range of pleiotropic effects including accelerated and early onset ageing [28]. Genetic defects in GG-NER and essential NER enzymes cause Xeroderma Pigmentosum (XP), which manifest as an over 1000-fold increased chance in developing skin cancer [29]. While base excision repair (BER) is key for the removal of oxidized bases there is abundant evidence that, in addition, NER factors are involved in repair of oxidative DNA damage either directly or via interaction with BER proteins [5 and references therein].

Another layer of protection against the cytotoxic effect of light-induced DNA damage is translesion synthesis (TLS). Contrary to replicative polymerases δ and ε, TLS polymerases are able to synthesize DNA across damaged templates, thereby avoiding persistent replication blocks and thus permitting cells to complete replication and cell cycle progression [30]. In eukaryotes TLS is performed by Y-family polymerases

Polη, Polκ, Polι and REV1 [31] and B-family Polζ, consisting of core subunits REV3 and REV7 [32] and two accessory subunits - Pol31 & Pol32 in yeast, PolD2 & PolD3 in human cells - that are shared with polymerase δ [33-36]. A wide variety of damages (*e.g.* UV-induced pyrimidine dimers and oxidized bases, but also abasic sites) are bypassed by TLS [37]. Patients with a mutation in the Polη encoding gene display hypersensitivity toward UV light and have increased cancer risk, highlighting the important role of Polη in the bypass of UV lesions [38-40]. TLS polymerases are rather promiscuous enzymes: numerous studies have described roles for REV1 and Polζ in the bypass of UV lesions, especially 6-4PPs [41-46], and redundant roles for Polι and Polκ [47,48], whereas Polη is also important for the bypass of 8oxoG [49,50].

In the absence of repair and TLS, photolesions are thought to generate insurmountable blocks to replication. Persistent stalled replication can cause a domino effect of problems linked to cancer and inborn disease: broken chromosomes, genomic rearrangements or missegregation and eventually aneuploidy and cell death [51]. We previously provided genetic evidence in *C. elegans* that persistent replication blocks can survive mitosis and result in double strand breaks (DSBs) each subsequent round of replication. These replication-associated DSBs are repaired by polymerase θ (Polθ) mediated end-joining leaving behind a genomic scar with a characteristic footprint [52-55]. Here, we report how NER, TLS and Polθ-mediated end-joining form three mechanistically different layers of protection against DNA damage induced by regular light. Without these three pathways nematodes become intolerant to light.

Materials & Methods

General culturing and strains used

All strains were cultured according to standard methods as described in [56]. The N2 Bristol strain was used as WT control. The alleles *polq-1(tm2026), rev-1(gk455794), rev-3(gk919715), polh-1(ok3317), xpa-1(ok0698)*, were obtained from the Caenorhabditis Genetics Center, Minnesota, USA. The *polh-1(lf0031)* and *polk-1(0029)* alleles were isolated in our own laboratory [57]. Crossing single and/or double mutants made combinations of the different alleles.

UV-C and UV-A Survival assays

To measure germline sensitivity to UV, staged young adults (one day post L4) were transferred to empty NGM plates and exposed to different doses of UV light. Per dose and genotype 3 plates with 3 adults were transferred to fresh NGM plates with OP50 and allowed to lay eggs for 20 hours. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later the numbers of non-hatched eggs

and surviving progeny were determined. UV-C and UV-A experiments were performed in a similar fashion. UV-C source: predominantly 254 nm, Philips. UV-A source: predominantly 366 nm, GE lighting F8T5 BLB U.S.A. Before every UV-C exposure the irradiance of our light source was determined using an International Light photometer (model: IL1400BL, ser. nr.: 7819). This varied slightly between experiments with an average of 26,0 (±1,0) µW*cm-2 (equals 0,260 J*m-2*s-1). Before every UV-A exposure the irradiance of our source was determined using a Blak-ray® long wave ultraviolet meter (model: J221, ser. #12994). Measurements varied slightly between experiments with an average of 140 (±10) µW*cm-2 (equals 1,40 J*m-2*s-1). For both UV-C and UV-A the exposure times determined the dose.

UV mutagenesis & unc-22 assay

The *unc-22* locus is used as a phenotypic readout to isolate UV-induced mutations that randomly occur throughout the genome of *C. elegans*. Loss of one or two of the two copies of the gene results in a phenotype referred to as "twitcher"; these worms twitch when they are soaked in buffer that contains 2 mM levamisole, while WT animals are paralyzed by this drug. An extra advantage of using this gene as a mutational target is that it has a very large open reading frame (>38 kb) [58]. Populations of young adult P0 animals were exposed to different UV-C doses. These animals were allowed to produce progeny that were then scored for heterozygous *unc-22* mutations. Single (F1) animals were isolated and allowed to produce offspring, and from each F1 a homozygous F2 animal was picked and propagated to be used for DNA analysis. Novel *unc-22* mutants from all three doses (2,5; 7,5 and 15 J/m2) were isolated. PCR of the entire gene (61 amplicons) was performed and samples were subjected to gel electrophoresis to identify deletions. The PCR fragments were Sanger sequenced to determine the molecular nature of the deletions. Bioinformatic analysis for the construction of the heat map was performed as described in [55].

Light vs. dark

To examine the sensitivity of animals towards plain daylight and artificial light in our laboratory, identical transparent plastic containers were used, one of which was coated with aluminum isolation tape to prevent any light from passing through, one was exposed to the artificial light in the lab, and one was exposed to only daylight (behind glass, in a room without artificial light). 20 L4-stage hermaphrodite animals were transferred to 9 cm plates and allowed to proliferate. Normally 20 WT animals will grow out to a full plate with 50000-100000 animals in approximately 7 days. On day 7 the number of living worms on each plate was determined. For each condition 10 plates were assayed. Since all plates that were not exposed to light were completely and equally full, only 3 of 10 were quantified.

Radiospectroscopy of artificial light

The spectral irradiance of the artificial light in our lab and the light that penetrates through a plastic container was determined with a ruggedized wideband spectroradiometer (International Light, model RPS900-R). SpectrILight III Spectral Analysis/Spectrometer Control Software was used to make the spectral readout and SpectrILight Tools Analysis Pak v.1.0 Irradiance software was used to plot and analyze the radiospectrogram.

N-Acetyl-L-cysteine exposure

Regular NGM plates were poured with medium supplemented with N-Acetyl-L-cysteine (NAC) (Sigma-Aldrich). These plates were seeded with OP50 (at least) one day before use. Animals were picked onto NGM plates (20 L4-stage animals per 9 cm plate) and allowed to reproduce for 7 days, either exposed to artificial light, to daylight (behind glass, in a room without artificial light), or not exposed to light. The total number of living worms on each plate was counted on day 7. UV-A dose-response curves were generated using animals that were transferred at the larval L4-stage to NAC containing plates 24 hours prior to UV-A exposure. All other parameters of the experiment were as described in section UV-C and UV-A Survival assays.

RAD51 staining of exposed young embryos

The protocol described in [59] was adapted for the staining of early embryos. Young adult animals (1-day post L4 stage) were exposed to 2,5 J/m2 UV-C and allowed to recover for 4h at 20C. These were subsequently picked into a 10 µl drop of dissection buffer (EGG buffer, 0.1% TWEEN 20, 0.2 mM Levamisole) on a 20x20 mm coverglass and cut in two at the site of their vulva to extract the embryos. 10-20 animals were dissected per slide. To extract as many embryos as possible a pipet (set to 5 µl) was used to pipet up and down a few times. Then, the sample was sandwiched between the coverglass and a Superfrost Plus slide from Thermo Scientific, placed on a frozen block (metal block immersed in dry ice) and incubated until the sample was completely crystallized. Following, the sample was freeze-cracked by taking off the coverglass and immediately transferred to cold (-20°C) methanol for 1 min. Subsequently, the slides were post-fixed with 500 µl of 4% PFA mix (PBS, 80 mM HEPES (pH 7.4), 0.8 mM EDTA, 4% formaldehyde, 1.6 mM MgSO4) and incubated 30 minutes @RT. After washing, the slides were blocked with 0.5% BSA in PBST. For the immuno-staining the slides were incubated with RAD51 antibody (rabbit polyclonal from SDIX/Novus Biologicals, cat# 29480002, used at 1:10000 in PBST+0,5%BSA) overnight at room temperature. Alexa anti-rabbit 488, 1:500, was used as secondary antibody (2h incubation at room temperature). DNA was visualized by DAPI staining and the slides were finished with Vectashield. Imaging and processing were done on a Leica DM6000B microscope.

Results & Discussion

Loss of NER and TLS confers sensitivity to UV-C exposure

UV-C light with a wavelength of 254 nm induces both mildly helix distorting CPDs and strongly helix distorting 6-4PPs, at a ratio of approximately 4:1 [60]. Efficient bypass of CPDs is performed by Polη [38-40,61] with possible redundant roles for Polι and Polκ [47,48]; bypass of 6-4PPs seems to be more complex due to the helix distorting nature of this lesion. Purified Polη and Polι were shown to polymerize DNA across the 3' T of a (6-4)T-T photoproduct but then abrogate further extension [62-64]. Polζ is then able to extend this substrate [48,65] perhaps facilitated by REV1 [46,48]. In order to study which TLS polymerases are involved in UV-C lesion bypass in *C. elegans* we tested strains with mutations in Polη, REV-1 and the catalytic subunit of Polζ: REV-3, or combinations thereof, for sensitivity towards UV-light. *C. elegans* does not encode Polι and mutation in Polκ confers hypersensitivity towards alkylating agents but not UV [57]. Figure 1a shows that loss of Polη confers the greatest hypersensitivity to UV-C exposure, whilst loss of REV-1 or REV-3 results in moderately increased sensitivity. Importantly, loss of REV-1 or REV-3 did not further increase the hypersensitivity measured in *polh-1* mutants (Suppl. fig. S1). In fact, the hypersensitivity towards UV in animals that lack all TLS polymerases is comparable to that observed in *polh-1* single mutant animals (Fig. 1a). We conclude that TLS past UV-C induced damage is fully lost in Polη deficient nematodes. The observation that loss of REV-1 and REV-3 enhances the UV-sensitivity of otherwise wild type animals but not of *polh-1* mutants may be explained by a different genetic requirement for different photoproducts: Polη being sufficient to bypass CPDs, but additionally depending on Polζ/REV-1 to bypass the more helix-distorting 6-4PPs. A direct interaction between POLH-1 and REV-1 has, however, not been established [66]. Such a model where REV-1 and Polζ act together on a subset of UV lesions is further substantiated by our finding that *rev-1* and *rev-3* single mutant animals are equally sensitive towards UV-C as *rev-1;rev-3* double mutant animals (Suppl. fig. S1b).

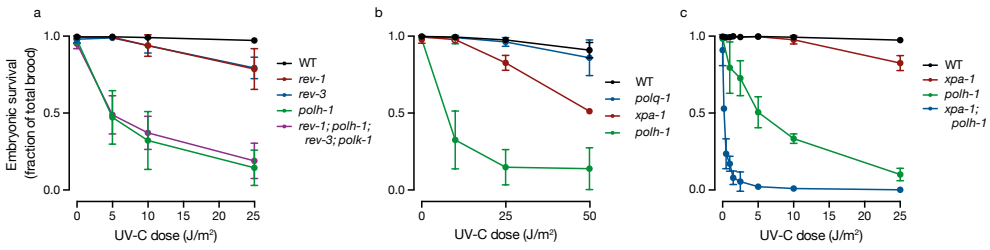


Figure 1. UV hypersensitivity of C. elegans NER and TLS mutants. *a-c) Hermaphrodite adult animals of indicated genotypes were exposed to different doses of UV-C and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20h time period post irradiation.*

Previous studies have shown that loss of a core NER component, such as the protein XPA, completely abrogates NER resulting in cellular and organismal hypersensitivity to UV exposure [18,67-70]. The fact that polh-1 mutant animals are more hypersensitive to UV than NER deficient animals (Fig. 1b) can be understood by realizing that *C. elegans* toxicity assays that encompass early embryonic cell divisions are extremely sensitive to perturbations of DNA replication progression [57,71,72]. As UV photolesions are repaired by NER or bypassed by TLS we expected to find a synergistic relationship between NER and TLS. Indeed, *xpa-1;polh-1* double mutants are profoundly more sensitive to UV than either single mutant (Fig. 1c), consistent with the observed synthetic lethality of NER and loss of TLS polymerase activity in other species [73,74]. Here, however, because *C. elegans* TLS at UV lesions completely depend on Polη (Fig. 1a), the *xpa-1;polh-1* double mutant allows us to study the consequences of a complete loss of repair and of bypass.

UV-C induced mutagenesis in TLS deficient animals

A failure to bypass UV-photoproducts has profound cytotoxic effects on animal development (Fig. 1). However, because Polη deficient persons develop UV-induced skin cancer we also suspect genotoxic consequences of Polη loss [38]. To test this hypothesis, we exposed populations of *polh-1* animals to UV and monitored the appearance of loss of function mutations in the *unc-22* locus, which results in an easy observable phenotypic change in movement. The use of this gene as an endogenous reporter allows us to determine the influence of UV on the rate of mutation induction as well as the isolation of mutants for sequence analysis. Figure 2a shows that a dose dependent relationship between the dose of UV-C exposure and the frequency of mutations in the population. An inverse correlation is observed between UV dose and the potential of animals to produce viable offspring. At different UV doses we isolated *unc-22* loss of function alleles, in total 28, which were analyzed by PCR. Of these, 17 contain a 50-500 bp deletion, 1 has a deletion of 2-4 kb and 10 have mutations that are smaller than ~20 bp, and are likely SNVs (Suppl. table 1). All deletions in the size range of 50-500 were sequenced to determine the genomic location and the deletion junctions at nucleotide resolution (Fig 2b,e; Suppl. table 2). We found a mean size of approximately 100 bp (Fig. 2b). One deletion allele contains an insert that can be mapped to the sequences immediate flanking the deletion (Fig. 2c; Suppl. table 2). For the deletions without insert we found micro-homology at their junctions (Fig. 2d). This specific mutational signature has recently been found to result from Polθ-mediated end joining (TMEJ) repair of replication associated DSBs [52,54,55]. From this we postulate that UV-light can induce substrates for TMEJ when TLS is deficient.

Strong synergistic relation between NER, TLS & TMEJ

The data above projects a simple and straightforward route for the fate of UV-

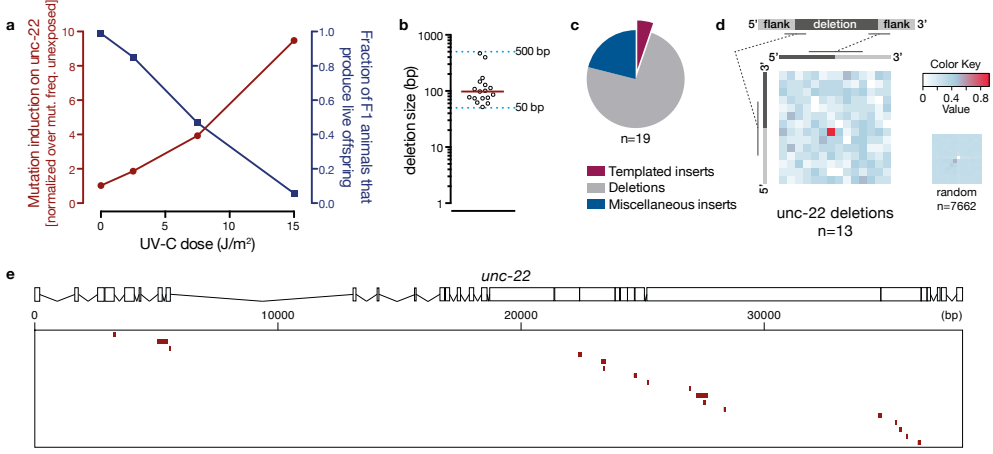


Figure 2. UV-C induced mutagenesis in TLS deficient animals show characteristics of Polθ mediated repair.
a) Dose-dependent increased frequency of unc-22 loss of function mutations upon UV-C exposure (in red). Dose dependent decrease in animal fertility (in blue). b) Size representation of UV-induced unc-22 deletion alleles (n=17). c) Distribution of deletion alleles: the majority are simple deletions (in grey), without containing insertions in between both breakpoints. Three alleles have small miscellaneous inserts of 1 to 6 bp (in blue), whereas one allele (in magenta) has an insert that is templated from the flank of the break (see also Suppl. table 2). d) A heat map representation of micro-homology at deletion breakpoints of simple deletion alleles. The bases that flank the right and left deletion breakpoints and are either retained or lost in the deletion alleles are plotted against each other. The heat map, representing 15 simple deletions, reveals overrepresentation of 1 nucleotide of micro-homology. A heat map for a simulated set of random deletions (n=7662) with random distribution is displayed on the right. e) A schematic representation of the unc-22 gene with underneath the position and size of the deletion alleles represented as red bars.

photoproducts: NER repairs photolesions, but those that escape detection are not interfering with proliferation as Polη prevents under-replication by performing TLS. In cases where photoproducts persist (*e.g.* because of lesion overload or genetic defects) stalled replication forks produce DNA breaks that are substrate to Polθ–mediated DSB repair. To further test this, we combined the defects in NER and TLS with a knockout allele of the polq-1 gene, which encodes *C. elegans* Polθ. However, upon generating these NER/TLS/TMEJ mutant animals we encountered great difficulty maintaining this triple mutant strain under normal laboratory conditions. Surprisingly we found that a failure to produce viable progeny was only observed when plates containing animals were maintained on the bench, but not when placed in a closed incubator. This led to the suggestion that NER/TLS/TMEJ deficient animals cannot cope with plain daylight and/or artificial light. To examine this, we transferred NER/TLS/TMEJ deficient animals onto agar plates and cultured these in transparent boxes in the presence of normal laboratory lighting (artificial and daylight), in the presence of only daylight in a room without artificial light, and in a third box that was covered by light-impermeable coating (Fig. 3a). Indeed, we found that animals proliferate normally in the dark, but

became sterile in the first or second generation when grown in light (Fig. 3b). Despite a markedly different wave-length spectrum (Suppl. fig. S2a), both light sources were almost equally potent in inducing sterility, although populations exposed to daylight were slightly (but statistically significant) less affected than those exposed to artificial light (Fig. 3b). None of the other mutant combinations manifested any difference between light and dark conditions (data not shown). We conclude that in the absence of repair and bypass, the inability to repair DSBs resulting from stalled replication at persistent damage renders nematodes incapable to tolerating even extremely small amounts of light.

UV-A can account for cytotoxicity in NER/TLS/TMEJ mutant animals

As UV-C and most UV-B are efficiently absorbed by the ozone layer, daylight contains 90-95% UV-A and just 5-10% UV-B. The artificial lighting in our laboratory emits low levels of UV-A but no light of smaller wavelengths (Suppl. fig. 2a,b). We thus tested whether exposure to just UV-A can account for the observed sensitivity of NER/TLS/TMEJ mutant animals to light. Indeed, we found that these animals displayed complete sterility at a dose where single mutant animals showed no, and double mutant animals only very mild sensitivity (Fig. 3c). Such dose equates to just ~4 minutes daylight in North-Western Europe [75].

UV-A can induce photolesions directly but also cause base damages indirectly, via triggering the formation of radical oxidation species (ROS). To address whether light-induced toxicity in NER/TLS/TMEJ mutants is (at least in part) caused by ROS, we performed similar toxicity assays with animals that are chronically exposed to N-Acetyl-L-cysteine (NAC). NAC has been shown to be an effective anti-oxidant in *C. elegans* to reduce cellular ROS levels [76-78]. Strikingly, we found that almost all UV-A-induced toxicity disappeared when animals were grown on NAC containing medium (Suppl. fig. S4), pointing towards ROS species as the causal agent. Such strong effect of NAC was, however, not observed when populations were grown under artificial or normal daylight conditions (Suppl. fig. S3a-c), perhaps because of a very high toxicity dose overriding any suppressing effect of a scavenger, or because of a higher ratio of direct versus indirect DNA damages for these light sources.

DSB accumulation increases when alt-EJ is inactivated

To investigate whether the observed inability to sustain proliferative growth was indeed caused by DSB formation, we monitored the accumulation of RAD51, a marker for DSBs, in TLS and TMEJ deficient genetic backgrounds upon exposure to UV light. Here, we used NER proficient backgrounds because experimental variation increased to impracticable levels at the extreme low UV-C dose that was required to separate *xpa-1; polh-1* from *xpa-1; polh-1; polq-1* (Suppl. Fig.1b). To quantitatively study how TLS and/or TMEJ

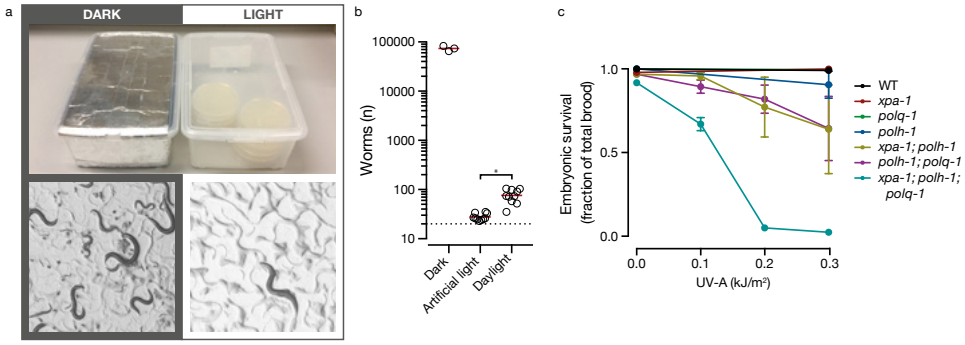


Figure 3. Strong synergistic relation between NER, TLS and TMEJ. *Images of the experimental conditions that were used to test animal sensitivity to normal light. Coating boxes with alumina foil created dark conditions (top left). In the transparent box (top right) animals are exposed to low intensity light. Nematode plates incubated in boxes without alumina foil are exposed to low intensity light. Underneath are representative images of NER/TLS/TMEJ triple deficient animals upon prolonged growth in either light or dark conditions. Under light conditions (bottom right) animals fail to produce progeny and plates only contain adult animals. When protected from light, animals produce healthy broods, the population proliferates normally, and different larval stages and eggs can be observed on the plate (bottom left). b) Proliferation of NER/TLS/TMEJ triple deficient animals that are either grown in the dark or exposed to different light sources. Each data point represents the total number of living animals on a plate 7 days after a population was started with 20 L4-stage animals (dotted lines). Red lines indicate the means. * indicates p<0,0001 as determined by a Mann-Whitney test. c) Hermaphrodite adult animals of indicated genotypes were exposed to different doses of UV-A and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20h time period post irradiation.*

suppresses the formation or accumulation of DSB during embryogenesis we counted RAD51 foci in the first embryonic divisions 4 hours after exposing gravid animals to 2,5 J/m2 UV-C, a relatively low dose to prevent irreversible cell-cycle arrest. We found that RAD51 foci in animals that are proficient in TLS (WT and *polq-1*) are extremely rare; in most cases not a single spot is observed (Fig. 4a). However, in the absence of TLS, DSBs manifest, but only after the damaged genome has replicated at least once: RAD51 foci were first observed at the 2-cell embryo stage (Fig. 4a). Knocking out *polq-1* in this TLS deficient background led to increased numbers of foci, especially in later stages of early embryonic divisions, providing molecular evidence that TMEJ acts to repair DSB that result from stalled replication at persistent DNA damage (Fig. 4a,b).

Interestingly, although UV-damage is inflicted at pre-fertilization stages (sperm, oocyte or meiotic prophase of germ nuclei) the total number of RAD51 foci increase during embryogenesis (Fig. 4c). This observation fits well with recent data arguing that persistent replication blocks can survive mitotic division and result in a new DSB at each cycle of DNA replication [54]. This observed increase in RAD51 foci, as opposed to their disappearance, also supports the idea that DSBs that occur in early stages of embryogenesis cannot be repaired by other means than TMEJ. This conclusion is in

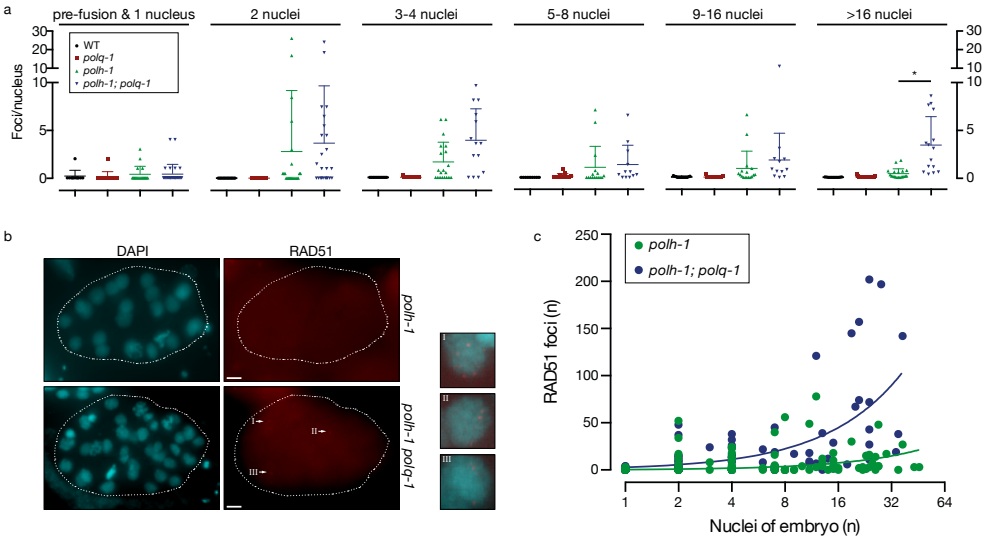


Figure 4. Polθ suppresses DSB accumulation when persistent replication blocks are induced.
*a) Quantification of RAD51 foci in young embryo's 4h after exposure to 2,5 J/m2 UV-C. Each data point represents the average number of foci per nucleus of one embryo. The number of nuclei defines the stage of the embryo. A Mann-Whitney test was used to define statistical significance, * indicates p<0,0001. b) Representative images polh-1 and polh-1;polq-1 embryos at the embryonic stage of >16 nuclei. The arrows and roman numerals correspond to the enlarged images at the right. Scale bars are 5 µm. c) A plot representing the total number of RAD51 foci in embryos at different stage. Each dot represents a single embryo. The data displayed in a) and c) are the sum of two independent experiments.*

line with the recent observation that repair of CRISPR/Cas9-induced breaks in *C. elegans* totally depend on functional TMEJ and not on NHEJ factors [79]. Also, Zebrafish embryos mutant for Polq cannot repair DSBs that are induced by CRISPR/Cas9 or ionizing radiation [80].

Conclusions

The notion that light, apart from being a necessity for biological systems, also has detrimental effects has been realized already more than a century ago when Downes & Blunt found that light exposure could reduce bacterial growth [1]. Over the last decades, a lot of work, mostly pioneered in bacterial systems, have elucidated how light, both UV and visible, can damage and mutate DNA [11-17], and a diverse set of molecular pathways were uncovered that protect the genome of organisms against light-induced DNA damage [Reviewed in 81]. We here reveal how profoundly toxic normal light is to a multicellular biological system when it has lost three layers of defense. Previously, it has been demonstrated that genetic inactivation of DNA repair or of translesion synthesis results in hypersensitivity to UV-C, but we here show that this level of

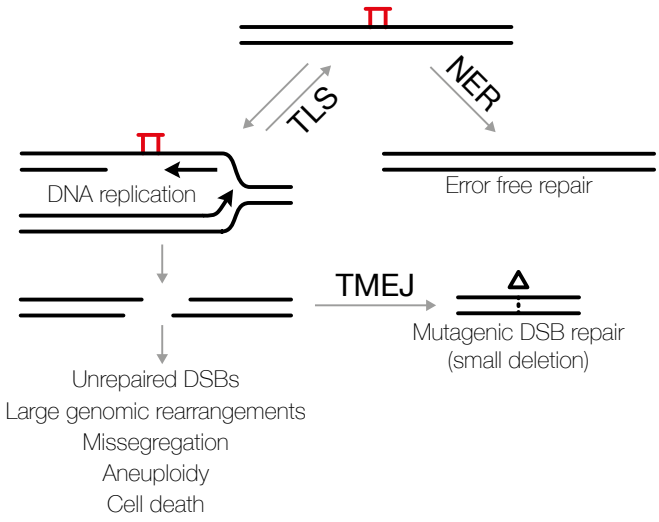


Figure 5. The interplay of mechanisms that protect cells against DNA damage induced by light.
Depending on whether DNA is being replicated or not, lesions will lead to translesion synthesis (TLS) or be the substrate of nucleotide excision repair (NER). NER restores the nucleotide composition of damaged DNA in a grosso modo error free fashion, whereas TLS is considered more mutagenic. TLS does not repair the damage but generates a new substrate for NER meanwhile suppressing the formation of replication-associated DSBs. These DSBs are repaired by TMEJ, which protects the cell against genomic havoc, at the price of increased mutagenesis.

hypersensitivity does not at all reflect the total damaging power of light toward living systems. Organisms have evolved mechanistically diverse pathways that together act to accommodate proliferation in the presence of such a ubiquitous threat to genetic integrity (Fig. 5).

Our data argue that NER, TLS and TMEJ act to protect cells from photoproducts that are induced directly by UV but also from DNA damages that result from ROS species. In fact, we find that oxidative damage is the main cytotoxic product of UV-A light, but to a lesser extend dictates the toxicity of visible light. The observation that UV-A toxicity in NER/TLS/TMEJ animals is higher than in all other single or double mutant genetic background, argues for a cytotoxic NER substrate that provides an obstacle to the replicative polymerases but not to Polη, which could be 8oxoG. It is interesting in this respect that the *C. elegans* genome does not encode OGG1, the primary glycosylase responsible for removing 8oxoG during base excision repair in other systems [Reviewed in 82].

The conservation of NER and TLS mechanisms from bacteria to mammals demonstrates the strong evolutionary pressure on organisms to be able to deal with

damaged bases: either through repairing them in an error free manner (by NER) or to prevent them from forming an insurmountable obstacle to replication (by TLS) [30,83-85]. Bacteria employ very efficient DNA damage tolerance (DDT) and damage avoidance (DA) pathways that allow for DNA replication of damaged DNA, also making use of the newly synthesized non-damaged sister chromatid. It has been proposed that this repair mode is preferred over TLS in bacteria because it is error-free, while TLS is mutagenic, although the extend of this dominance is depending on the type of lesion [86]. Template switching (TS)-based mechanisms have also been described for eukaryotic cells such as yeast and may also function in higher eukaryotes [87]. While it is unambiguous that TLS plays a key role in mammals, the role and importance of TS remains unclear. We have not found any evidence of such a pathway functioning in *C. elegans* to avoid persistent replication blocks during embryogenesis [54]. Perhaps the rapid divisions that take place during embryogenesis do not provide enough time to execute and complete complex reactions such as the formation and resolution of homologous recombination (HR) intermediates; early embryonic cell divisions are also extremely sensitive to perturbations of DNA replication progression [57,71,72]. In light of this rationale it is worth noting that, Wolters *et al*. recently found that UV sensitivity can result from the formation of toxic homologous HR intermediates in *C. elegans* [88]. In agreement, we here find increased levels of RAD51 foci in TLS deficient animals upon UV exposure, arguing for the formation of such intermediates. TMEJ may act to suppress RAD51 foci formation and thus prevent setting in motion HR related processing, by providing a quick fix to DSBs that result at persistent replication blocks. For a discussion on why replication-associated DSBs at persistent lesions are exclusive substrates for TMEJ and not for classical NHEJ or HR we wish to refer the reader to recently published work that addresses this issue using G-quadruplexes as a model substrate to block DNA replication [53,54].

In summary, in the absence of the NER, cells use TLS to prevent DSB generation at replication fork barriers. We here show that in the absence of TLS, in which case NER removes most photo-lesions, TMEJ deals with the DSBs that result at sites of stalled replication and as such prevents genomic havoc. Interestingly, TMEJ deficiency does not itself results in hypersensitivity towards UV light (Fig. 1b) likely because the capacity of NER and TLS combined prevents UV-lesions in becoming persistent replication blocks. However, inactivation of these pathways in combination, in which a single UV lesion may be an insurmountable block to replication and thus cause cell death, reveals that simple daylight is already so toxic that without these cellular protection mechanisms nematodes simply cannot survive.

Acknowledgements

Several strains were provided by the CGC, which is funded by the NIH Office of Research

Infrastructure Programs . This work is supported by the European Research Council (ERC) and the Netherlands Organization for Scientific Research (NWO). We would like to thank Ron Romeijn for experimental assistance, Robin van Schendel for bioinformatical support, and Threes Smijs for radiospectroscopic analysis.

References

[1] Downes A, Blunt TP. Researches on the effect of light upon bacteria and other organisms, 1877.

[2] Kubitschek HE. Mutagenesis by near-visible light. Science 1967. doi:10.1126/science.155.3769.1545.

[3] Webb RB, Malina MM. Mutagenesis in Escherichia coli by visible light. Science 1967;156:1104–5. doi:10.1126/science.156.3778.1104.

[4] Beukers R, Eker APM, Lohman PHM. 50 years thymine dimer. DNA Repair 2008;7:530–43. doi:10.1016/j.dnarep.2007.11.010.

[5] Schuch AP, Moreno NC, Schuch NJ, Menck CFM, Garcia CCM. Sunlight damage to cellular DNA: Focus on oxidatively generated lesions. Free Radic Biol Med 2017. doi:10.1016/j.freeradbiomed.2017.01.029.

[6] Douki T, Reynaud-Angelin A, Cadet J, Sage E. Bipyrimidine photoproducts rather than oxidative lesions are the main type of DNA damage involved in the genotoxic effect of solar UVA radiation. Biochemistry 2003;42:9221–6. doi:10.1021/bi034593c.

[7] Rochette PJ, Therrien J-P, Drouin R, Perdiz D, Bastien N, Drobetsky EA, et al. UVA-induced cyclobutane pyrimidine dimers form predominantly at thymine-thymine dipyrimidines and correlate with the mutation spectrum in rodent cells. Nucleic Acids Res 2003;31:2786–94. doi:10.1093/nar/gkg402.

[8] Besaratinia A, Synold TW, Chen H-H, Chang C, Xi B, Riggs AD, et al. DNA lesions induced by UV A1 and B radiation in human cells: comparative analyses in the overall genome and in the p53 tumor suppressor gene. Proc Natl Acad Sci U S A 2005;102:10058–63. doi:10.1073/pnas.0502311102.

[9] Gates KS. An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. Chem Res Toxicol 2009;22:1747–60. doi:10.1021/tx900242k.

[10] Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. Toxicol Sci 2011;120 Suppl 1:S130–45. doi:10.1093/toxsci/kfq371.

[11] Kielbassa C, Roza L, Epe B. Wavelength dependence of oxidative DNA damage induced by UV and visible light. Carcinogenesis 1997;18:811–6.

[12] Kuluncsics Z, Perdiz D, Brulay E, Muel B, Sage E. Wavelength dependence of ultraviolet-induced DNA damage distribution: involvement of direct or indirect mechanisms and possible artefacts. J Photochem Photobiol B, Biol 1999;49:71–80. doi:10.1016/S1011-1344(99)00034-2.

[13] Zhang X, Rosenstein BS, Wang Y, Lebwohl M, Mitchell DM, Wei H. Induction of 8-oxo-7,8-dihydro-2'-deoxyguanosine by ultraviolet radiation in calf thymus DNA and HeLa cells. Photochem Photobiol 1997;65:119–24.

[14] Jostes RF, Dewey WC, Hopwood LE. Mutagenesis by fluorescent light in mammalian cell cultures. Mutat Res 1977;42:139–44.

[15] Burki HJ, Lam CK. Comparison of the lethal and mutagenic effects of gold and white fluorescent lights on cultured mammalian cells. Mutat Res 1978;54:373–7.

[16] Jones CA, Huberman E, Cunningham ML, Peak MJ. Mutagenesis and cytotoxicity in human epithelial cells by far- and near-ultraviolet radiations: action spectra. Radiat Res 1987;110:244–54.

[17] Godley BF, Shamsi FA, Liang F-Q, Jarrett SG, Davies S, Boulton M. Blue light induces mitochondrial DNA damage and free radical production in epithelial cells. J Biol Chem 2005;280:21061–6. doi:10.1074/jbc.M502194200.

[18] Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JHJ. Understanding nucleotide excision repair and its roles in cancer and ageing. Nature Reviews Molecular Cell Biology 2014;15:465–81. doi:10.1038/nrm3822.

[19] Vermeulen W, Fousteri M. Mammalian transcription-coupled excision repair. Cold Spring Harb Perspect Biol 2013;5:a012625. doi:10.1101/cshperspect.a012625.

[20] Fagbemi AF, Orelli B, Schärer OD. Regulation of endonuclease activity in human nucleotide excision repair. DNA Repair 2011;10:722–9. doi:10.1016/j.dnarep.2011.04.022.

[21] de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes & Development 1998;12:2598–609. doi:10.1101/gad.12.16.2598.

[22] Dunand-Sauthier I, Hohl M, Thorel F, Jaquier-Gubler P, Clarkson SG, Schärer OD. The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. J Biol Chem 2005;280:7030–7. doi:10.1074/jbc.M412228200.

[23] Zotter A, Luijsterburg MS, Warmerdam DO, Ibrahim S, Nigg A, van Cappellen WA, et al. Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. Mol Cell Biol 2006;26:8868–79. doi:10.1128/MCB.00695-06.

[24] Staresincic L, Fagbemi AF, Enzlin JH, Gourdin AM, Wijgers N, Dunand-Sauthier I, et al. Coordination of dual incision and repair synthesis in human nucleotide excision repair. Embo J 2009;28:1111–20. doi:10.1038/emboj.2009.49.

[25] Tsodikov OV, Ivanov D, Orelli B, Staresincic L, Shoshani I, Oberman R, et al. Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. Embo J 2007;26:4768–76. doi:10.1038/sj.emboj.7601894.

[26] Mocquet V, Lainé JP, Riedl T, Yajin Z, Lee MY, Egly JM. Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. Embo J 2008;27:155–67. doi:10.1038/sj.emboj.7601948.

[27] Ogi T, Limsirichaikul S, Overmeer RM, Volker M, Takenaka K, Cloney R, et al. Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. Mol Cell 2010;37:714–27. doi:10.1016/j.molcel.2010.02.009.
[28] Laugel V. Cockayne syndrome: the expanding clinical and mutational spectrum. Mech Ageing Dev 2013;134:161–70. doi:10.1016/j.mad.2013.02.006.
[29] DiGiovanna JJ, Kraemer KH. Shining a light on xeroderma pigmentosum. J Invest Dermatol 2012;132:785–96. doi:10.1038/jid.2011.426.
[30] Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harb Perspect Biol 2013;5:a012708. doi:10.1101/cshperspect.a012708.
[31] Sale JE, Lehmann AR, Woodgate R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Reviews Molecular Cell Biology 2012;13:141–52. doi:10.1038/nrm3289.
[32] Gan GN, Wittschieben JP, Wittschieben BØ, Wood RD. DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res 2008;18:174–83. doi:10.1038/cr.2007.117.
[33] Baranovskiy AG, Lada AG, Siebler HM, Zhang Y, Pavlov YI, Tahirov TH. DNA Polymerase and Switch by Sharing Accessory Subunits of DNA Polymerase. Journal of Biological Chemistry 2012;287:17281–7. doi:10.1074/jbc.M112.351122.
[34] Johnson RE, Prakash L, Prakash S. Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. Proc Natl Acad Sci U S A 2012;109:12455–60. doi:10.1073/pnas.1206052109.
[35] Makarova AV, Stodola JL, Burgers PM. A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. Nucleic Acids Res 2012;40:11618–26. doi:10.1093/nar/gks948.
[36] Lee Y-S, Gregory MT, Yang W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc Natl Acad Sci U S A 2014;111:2954–9. doi:10.1073/pnas.1324001111.
[37] Prakash S, Johnson RE, Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem 2005;74:317–53. doi:10.1146/annurev.biochem.74.082803.133250.
[38] Johnson RE, Kondratick CM, Prakash S, Prakash L. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 1999;285:263–5. doi:10.1126/science.285.5425.263.
[39] Johnson RE, Prakash S, Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 1999;283:1001–4.
[40] McDonald JP, Levine AS, Woodgate R. The Saccharomyces cerevisiae RAD30 Gene, a Homologue of Escherichia coli dinB and umuC, Is DNA Damage Inducible and Functions in a Novel Error-Free Postreplication Repair Mechanism. Genetics 1997;147:1557–68.
[41] Nelson JR, Lawrence CW, Hinkle DC. Deoxycytidyl transferase activity of yeast REV1 protein. Nature 1996;382:729–31. doi:10.1038/382729a0.
[42] Nelson JR, Lawrence CW, Hinkle DC. Thymine-thymine dimer bypass by yeast DNA polymerase zeta. Science 1996;272:1646–9. doi:10.1126/science.272.5268.1646.
[43] Nelson JR, Gibbs PE, Nowicka AM, Hinkle DC, Lawrence CW. Evidence for a second function for Saccharomyces cerevisiae Rev1p. Mol Microbiol 2000;37:549–54.
[44] Jansen JG, Tsaalbi-Shtylik A, de Wind N. Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. DNA Repair 2015;29:56–64. doi:10.1016/j.dnarep.2015.01.001.
[45] Yoon J-H, Park J, Conde J, Wakamiya M, Prakash L, Prakash S. Rev1 promotes replication through UV lesions in conjunction with DNA polymerases η, ι, and κ but not DNA polymerase ζ. Genes & Development 2015;29:2588–602. doi:10.1101/gad.272229.115.
[46] Jansen JG, Tsaalbi-Shtylik A, Hendriks G, Gali H, Hendel A, Johansson F, et al. Separate domains of Rev1 mediate two modes of DNA damage bypass in mammalian cells. Mol Cell Biol 2009;29:3113–23. doi:10.1128/MCB.00071-09.
[47] Jansen JG, Temviriyanukul P, Wit N, Delbos F, Reynaud C-A, Jacobs H, et al. Redundancy of mammalian Y family DNA polymerases in cellular responses to genomic DNA lesions induced by ultraviolet light. Nucleic Acids Res 2014;42:11071–82. doi:10.1093/nar/gku779.
[48] Temviriyanukul P, van Hees-Stuivenberg S, Delbos F, Jacobs H, de Wind N, Jansen JG. Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair 2012;11:550–8. doi:10.1016/j.dnarep.2012.03.007.
[49] Prakash S, Haracska L, Yu S-L, Johnson RE, Prakash L. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase - Nature Genetics. Nat Genet 2000;25:458–61. doi:10.1038/78169.
[50] Lee D-H, Pfeifer GP. Translesion synthesis of 7,8-dihydro-8-oxo-2'-deoxyguanosine by DNA polymerase eta in vivo. Mutat Res 2008;641:19–26. doi:10.1016/j.mrfmmm.2008.02.006.
[51] Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. Transl Cancer Res 2013;2:107–29. doi:10.3978/j.issn.2218-676X.2013.04.01.
[52] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.
[53] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.
[54] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.
[55] van Schendel R, van Heteren J, Welten R, Tijsterman M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.
[56] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.
[57] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.
[58] Moerman DG, Benian GM, Waterston RH. Molecular cloning of the muscle gene unc-22 in Caenorhabditis elegans by Tc1 transposon tagging. Proc Natl Acad Sci U S A 1986;83:2579–83.
[59] Colaiácovo MP, MacQueen AJ, Martinez-Perez E,

McDonald K, Adamo A, La Volpe A, et al. Synaptonemal complex assembly in *C. elegans* is dispensable for loading strand-exchange proteins but critical for proper completion of recombination. Dev Cell 2003;5:463–74.
[60] Perdiz D, Grof P, Mezzina M, Nikaido O, Moustacchi E, Sage E. Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells. Possible role of Dewar photoproducts in solar mutagenesis. J Biol Chem 2000;275:26732–42. doi:10.1074/jbc.M001450200.
[61] Biertümpfel C, Zhao Y, Kondo Y, Ramón-Maiques S, Gregory M, Lee JY, et al. Structure and mechanism of human DNA polymerase eta. Nature 2010;465:1044–8. doi:10.1038/nature09196.
[62] Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. Nature 2000;406:1015–9. doi:10.1038/35023030.
[63] Johnson RE, Haracska L, Prakash S, Prakash L. Role of DNA polymerase eta in the bypass of a (6-4) TT photoproduct. Mol Cell Biol 2001;21:3558–63. doi:10.1128/MCB.21.10.3558-3563.2001.
[64] Haracska L, Johnson RE, Unk I, Phillips BB, Hurwitz J, Prakash L, et al. Targeting of human DNA polymerase iota to the replication machinery via interaction with PCNA. Proc Natl Acad Sci U S A 2001;98:14256–61. doi:10.1073/pnas.261560798.
[65] Yoon JH, Prakash L, Prakash S. Error-free replicative bypass of (6-4) photoproducts by DNA polymerase in mouse and human cells. Genes & Development 2010;24:123–8. doi:10.1101/gad.1872810.
[66] Kosarek JN, Woodruff RV, Rivera-Begeman A, Guo C, D'Souza S, Koonin EV, et al. Comparative analysis of in vivo interactions between Rev1 protein and other Y-family DNA polymerases in animals and yeasts. DNA Repair 2008;7:439–51. doi:10.1016/j.dnarep.2007.11.016.
[67] Hartman PS, Herman RK. Radiation-sensitive mutants of Caenorhabditis elegans. Genetics 1982;102:159–78.
[68] Astin JW, O'Neil NJ, Kuwabara PE. 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 2008;7:267–80. doi:10.1016/j.dnarep.2007.10.004.
[69] Stergiou L, Eberhard R, Doukoumetzidis K, Hengartner MO. NER and HR pathways act sequentially to promote UV-C-induced germ cell apoptosis in Caenorhabditis elegans. Cell Death Differ 2011;18:897–906. doi:10.1038/cdd.2010.158.
[70] Lans H, Marteijn JA, Schumacher B, Hoeijmakers JHJ, Jansen G, Vermeulen W. Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development. PLoS Genet 2010;6:e1000941. doi:10.1371/journal.pgen.1000941.
[71] Brauchle M, Baumer K, Gönczy P. Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in *C. elegans* embryos. Curr Biol 2003;13:819–27. doi:10.1016/S0960-9822(03)00295-1.
[72] Gönczy P, Schnabel H, Kaletta T, Amores AD, Hyman T, Schnabel R. Dissection of cell division processes in the one cell stage Caenorhabditis elegans embryo by mutational analysis. J Cell Biol 1999;144:927–46.
[73] Nakajima S, Lan L, Kanno S-I, Takao M, Yamamoto K, Eker APM, et al. UV light-induced DNA damage and tolerance for the survival of nucleotide excision repair-deficient human cells. J Biol Chem 2004;279:46674–7. doi:10.1074/jbc.M406070200.
[74] Tsaalbi-Shtylik A, Verspuy JWA, Jansen JG, Rebel H, Carlée LM, van der Valk MA, et al. Error-prone translesion replication of damaged DNA suppresses skin carcinogenesis by controlling inflammatory hyperplasia. Proc Natl Acad Sci U S A 2009;106:21836–41. doi:10.1073/pnas.0909507106.
[75] Moseley H, Davison M, Mackie RM. Measurement of daylight UVA in Glasgow. Phys Med Biol 1983;28:589–97. doi:10.1088/0031-9155/28/5/011.
[76] Schulz TJ, Zarse K, Voigt A, Urban N, Birringer M, Ristow M. Glucose restriction extends Caenorhabditis elegans life span by inducing mitochondrial respiration and increasing oxidative stress. Cell Metab 2007;6:280–93. doi:10.1016/j.cmet.2007.08.011.
[77] Yang W, Hekimi S. A mitochondrial superoxide signal triggers increased longevity in Caenorhabditis elegans. PLoS Biol 2010;8:e1000556. doi:10.1371/journal.pbio.1000556.
[78] Desjardins D, Cacho-Valadez B, Liu J-L, Wang Y, Yee C, Bernard K, et al. Antioxidants reveal an inverted U-shaped dose-response relationship between reactive oxygen species levels and the rate of aging in Caenorhabditis elegans. Aging Cell 2017;16:104–12. doi:10.1111/acel.12528.
[79] van Schendel R, Roerink SF, Portegijs V, van den Heuvel S, Tijsterman M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat Comms 2015;6:7394. doi:10.1038/ncomms8394.
[80] Thyme SB, Schier AF. Polq-Mediated End Joining Is Essential for Surviving DNA Double-Strand Breaks during Early Zebrafish Development. Cell Rep 2016. doi:10.1016/j.celrep.2016.03.072.
[81] Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell 2010;40:179–204. doi:10.1016/j.molcel.2010.09.019.
[82] Boiteux S, Coste F, Castaing B. Repair of 8-oxo-7,8-dihydroguanine in prokaryotic and eukaryotic cells: Properties and biological roles of the Fpg and OGG1 DNA N-glycosylases. Free Radic Biol Med 2016. doi:10.1016/j.freeradbiomed.2016.11.042.
[83] Jarosz DF, Beuning PJ, Cohen SE, Walker GC. Y-family DNA polymerases in Escherichia coli. Trends Microbiol 2007;15:70–7. doi:10.1016/j.tim.2006.12.004.
[84] Simmons LA, Foti JJ, Cohen SE, Walker GC. The SOS Regulatory Network. EcoSal Plus 2008;3. doi:10.1128/ecosalplus.5.4.3.
[85] Schärer OD. Nucleotide excision repair in eukaryotes. Cold Spring Harb Perspect Biol 2013;5:a012609. doi:10.1101/cshperspect.a012609.
[86] Fuchs RP. Tolerance of lesions in E. coli: Chronological competition between Translesion Synthesis and Damage Avoidance. DNA Repair 2016;44:51–8. doi:10.1016/j.dnarep.2016.05.006.
[87] Unk I, Hajdu I, Blastyák A, Haracska L. Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair 2010;9:257–67. doi:10.1016/j.dnarep.2009.12.013.
[88] Wolters S, Ermolaeva MA, Bickel JS, Fingerhut JM, Khanikar J, Chan RC, et al. Loss of Caenorhabditis elegans BRCA1 promotes genome stability during replication in smc-5 mutants. Genetics 2014;196:985–99. doi:10.1534/genetics.113.158295.

Supplemental information

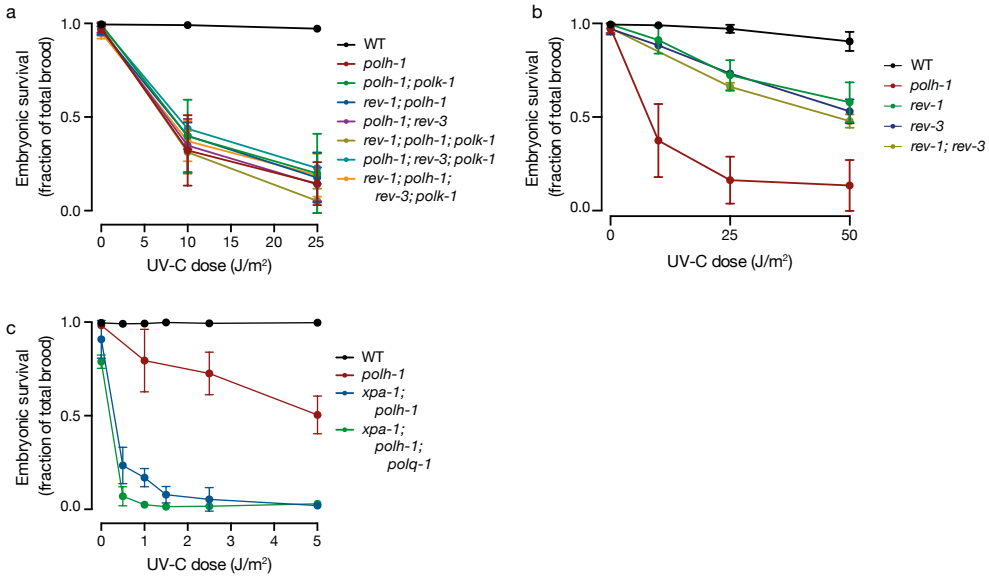


Figure S1. *a+b) Hermaphrodite adult animals of indicated genotypes were exposed to different doses of UV-C and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20 hours time period post irradiation.*

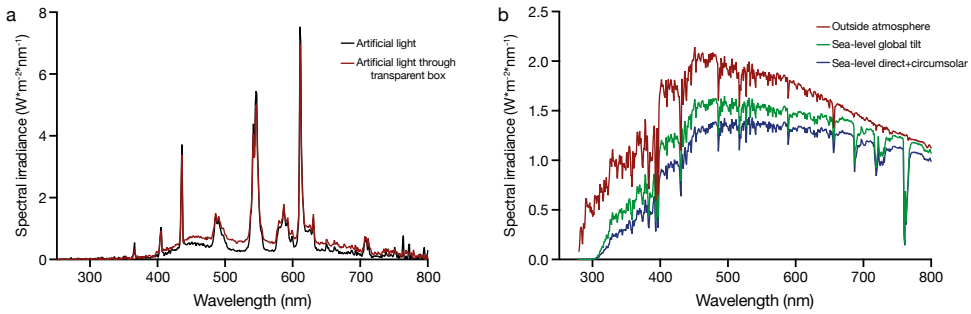


Figure S2*. a) Radiospectrogram of the spectral irradience of the artificial light in our lab. It contains low levels of UV-A and -B with a clear peak at 366 nm and also substantial ammounts of high energy violet end blue light (400-500 nm). b) Radiospectrogram of regular daylight measured in spectral irradiance. Source: ASTM Terrestrial Reference Spectra for Photovoltaic Performance Evaluation (http://rredc.nrel.gov/solar/spectra/am1.5). No data was available for wavelengths lower than 280 nm. However, it is known that UV-C light with wavelengths between 100 and 280 nm is efficiently absorbed by the ozone layer and does not reach the Earth's surface.*

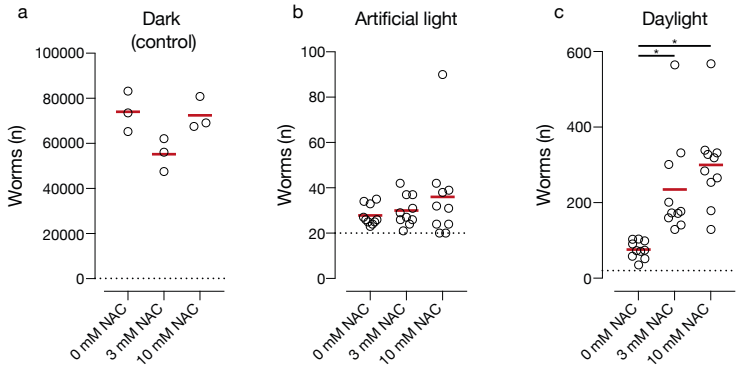


***Figure S3**. Proliferation of NAC treated populations of NER/TLS/TMEJ triple deficient animals when exposed to different light sources. Each data point represents the total number of living animals on a plate after 7 days from a starting population of 20 L4-stage animals (dotted lines). Red lines indicate the mean of each NAC dose and the thick black lines with a * indicate statistically significant differences determined by Mann-Whitney tests. a) When animals are not exposed to any light they proliferate at a normal rate and grow out to a full population of 50000-100000 animals. No significant differences were observed between the control and different doses of NAC. b) Exposure to the artificial light in the lab is extremely toxic to the NER/TLS/TMEJ triple deficient animals. No significant differences were observed between the control and different doses of NAC. c) When animals were exposed to just daylight in our lab (behind glass, in a room without artificial light) addition of NAC significantly increased proliferation of the populations. No significant difference was observed between the different NAC doses.*

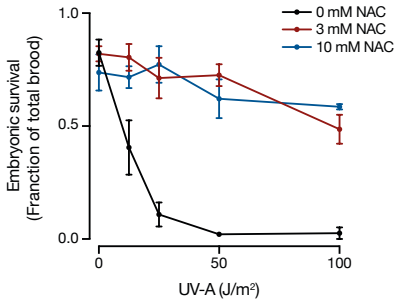


***Figure S4**. Hermaphrodite adult animals, triple deficient for NER/TLS/TMEJ were exposed to different doses of UV-A and different doses of NAC and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20 hours time period post irradiation. Exposure to the antioxidant NAC strongly reduces the embryonic lethality of the triple mutant when exposed to UV-A.*

Suppemental table 1: *Quantitative data of UV induced mutagenesis at unc-22*

Dose (J/m²)	Viable F1 animals	50-500 bp deletion found	Larger deletion found (2-4 kb)	Unkown mutation
2,5	16	13	0	3
7,5	10	3	1	6
15	2	1	0	1
Total	28	17	1	10

Suppemental table 2: *Nucleotide sequences of UV induced deletions (50-500 bp) in unc-22 locus in polh-1 mutant animals.*

	Position on unc-22						
Dose (J/m2)	Start	End	Size (nt)	Left flank	Deletion	Right flank*	Insert*
2,5	36324	36424	101	ACTCTGCAGGA-CAAGAACAAAC-CAGAGCAA-CATTGACAGT	GAAAGGAGATCAACCACT-TCTCAATGGACACGCTGGA-CAGGCTGTTGAAAGTGAACT-TCGTGTAACAAAGCACTTGG-GAGGAGAAATTGTGAATAAT-GGAG	AGTCAGTTACATTTGAAGCTA-GAGTGCAAGGAACACCAGAA-GAAGTGT	T
2,5	35566	35651	86	GCCGATGAA-GAAGGCACCAA-GTCCACCAA-GAGTTGAAGAA	TTCAAGGAGAGAAGATCTG-CACCCTTCTTCACATTCCATCT-CAGAAATCGTTTGATTCAAAA-GAACCATCAGTGCAAATT-GACATG	TTCTTTGCAAG-GAAACCCTAATCCAACAATT-GAATGGATGAAGGACGG	GAAAAT
2,5	35398	35471	74	ATGAAGAGGTA-CAATGGAAAT-GATTATGGACT-TACCATT	AACCGAGTAAAGGGAGATGA-TAAGGGAGAATACACAGTC-CGTGCAAAGAACTCATACG-GAACCAAGGAAGAAAT	TGTATTCTTGAATGTTACCCGT-CACTCTGAACCACTCAAATTC-GAGCCAT	-
2,5	34689	34816	128	AATCCGTCAA-GGTTACAA-CAGGAACTGC-CGAATTTGCCG	CTCCAGAAGTTGCCGAAGG-CAAGCCAGTCGGTTATTACAC-CGATATGTGGAGCGTTGGAGT-TCTCTCTTACATTCTTCTTTC-CGGACTTTCACCATTTGGAG-GAGAGAACGATGATGAAA-CATTGAGG	AATGTTAAGGTATGTTTTAT-TTTAATTATTAACTTGATAT-TAAACGTTAA	-
2,5	27510	27607	98	ATTCCGTGT-CAAGGCAGT-CAATCTTCAAG-GAGAAT	CAAAACCATTGGAAGCTGAA-GAACCAATTATTGCAAAGAAT-CAATTTGATGTTCCTGATC-CAGTTGACAAACCAGAGGT-TACTGACTGGGATAAGGAT	AGAATTGATATTAAGTG-GAACCCAACTGCAAACAATG-GAGGAGCTCCAG	-

	Position on unc-22						
2,5	27205	27674	470	CTGCTGTTCGT-GGAGATACT-GGAGTTTA-CAAAATCAT	TGTTGAAAATGAGCATG-GAAAAGATACTGCTCAGTG-CAATGTTACTGTTCTTGATG-TACCAGGAACTCCAGAAGGAC-CACTCAAGATTGACGAGATC-CATAAGGAAGGATGTACATT-GAACTGGAAGCCTCCAACT-GATAACGGAGGAACTGATGT-TCTTCACTACATTGTTGAGAA-GATGGATACTTCTCGTGGAA-CATGGCAGGAAGTCGGAACT-TTCCCAGATTGTACAGCCAA-GGTTAATAAGCTTGTTCCTG-GAAAGGAATACGCATTCCGT-GTCAAGGCAGTCAATCTTCAA-GGAGAATCAAAACCATTG-GAAGCTGAAGAACCAAT-TATTGCAAAGAATCAATTT-GATGTTCCTGATCCAGTTGA-CAAACCAGAGGTTACTGACTG-GGATAAGGATAGAATTGATAT-TAAGTGGAACCCAACTGCAAA-CAATGGAGGAGCTCCAGT-CACTGGATATATTGTTG	AGAAGAAGGAGAAGGGAAG-CGCAATCTGGACAGAAGCCG-GAAAGACTCC	C
2,5	26933	27009	77	TTCCGTGT-GAAGGCTGT-CAATAAGGCTG-GGCCAGGAAAA	CCATCTGATCCAACAGGAAAT-GTTGTTGCCAAACCAAGAA-GAATGGCTCCAAAACT-TAACCTCGCCGGACTTTTGGA	TCTCCGTATCAAGGCTGGAA-CACCCATCAAGCTCGATATCGC	-
2,5	25212	25263	52	TCTAAATCAAT-TACTAATAAT-CATTGAAT-TAAATTATAG	TTCCACCAGTCATCGATCG-TAACTCGATTCAAGAAATCAA-GGTCAAGGCTGG	TTGGAGGCTCGTGAAGC-TATTATCGCCAAGGATCCAT-TCGATCGTGCTG	-
2,5	23378	23437	60	AGGCTCAAATT-GAAAATGAGC-CATACATCTC-GAGATTTG	CTTTGCCAAAGGCACTTCG-TAAGCAAAGTGGAAAATATAC-CATCACTGCAACCAACATTA	ATGGAACTGACAGTGTCAC-TATCAATATCAAGGTAAAAAG-CAAGCCAAC	GAACT-GACAGT-GTCAC
2,5	23324	23493	170	CAGATG-TAACATGGT-CATTCAATG-GAAAAG	GAATCGGAGAGAGCAAGGCT-CAAATTGAAAATGAGCCATA-CATCTCGAGATTTGCTTTGC-CAAAGGCACTTCGTAAGCAAA-GTGGAAAATATACCATCACTG-CAACCAACATTAATGGAACT-GACAGTGTCACTATCAATAT-CAAGGTAAAAAGCAAGC-CAACGAAACCA	AAGGGACCAATCGAGGTAACT-GATGTCTTCGAAGATCGTG-CAACTCTTG	AATT
2,5	22355	22501	147	ATCAAGGTAG-GAAACGATGTG-GAATTCGATGT-TCCAGTA	CGCGGAGAACCACCACCGAA-GAAGGAATGGATCTTCAAT-GAGAAACCAGTCGATGAT-CAAAAGATCAGGGTAAGAT-TTTATATATTTTTGAACAT-TCTAATTTACGATTATTTCA-GATTGAAAGCGAAGACTACAA-GACCCGATTTG	TGCTCCGTGGAGCAACTCG-CAAGCATGCTGGTTTGTA-CACTCTTACT	-
2,5	5525	5598	74	AAGAA-CAGAAATCG-CAGACGCT-TCAAGCT-GAAACCAAAA	AGCGAGTGGCGAGACGAAG-CAAGTCAAAGAGTAAGAGTC-CGGTAAGTTATACTATTTTAAT-TATTTTATCAAAA	GGAAATCAGTATGCATTATAAT-TATTTCAAGCTGATACCTTTC-CAGGC	-
2,5	3232	3344	113	AAAAAATCAAT-GAAATCGAGG-GAAGGAACTC-CAAAACGT	ACCCTGAAACCAAGAGAGGGT-TCCCCATCGAAAAAGTTGAG-GTGAGCATTTCATTTGAAAT-TATATGGAGAAACAAAAT-TATAAAATTTTGCAAAAT-TAAATTGTATCGTTTT	TCAAATTTCCCAAT-CATCCCGATTTTCTCGCT-TCTAATTGTAAGAATT	-

	Position on unc-22						
7,5	35857	35932	76	CTTGGGAGTT-GATGTCTC-CGAGTGTG-TACTCACAGT-TCAA	ACTAAAGGAGGTGAACCAAT-TCCACGTGTTTCTTCGTTCA-GACCCCGAAGAGCTTATGACA-CATTATCAACTGGAA	CTGATGTCGAAAGATCACAT-TCGTATGCTGATATGAGAA-GAAGATCCC	-
7,5	24644	24751	108	GTATTCCAATT-GAAAACTATT-GATCGAAAAG-TACGATACT	GCAAGTGGAAGATGGGT-TCCAGCTGCAAAGGTCGCT-GGAGATAAGACTACAGCT-GTTGTTGACGGTCTTATTCCTG-GACATGAATATAAAGTGAGAAT-TTTTTAATCT	TGAACTCAATTGCCTACAA-CATTTTGATTTTCAGTTCCGT-GTCGCTGC	-
7,5	5082	5480	399	CCAT-CAAAATAAT-TATCTCAAAAT-TTTCAGAGATG-GTACA	TTGGTCAGAAATTCTTC-CGAATATTCACAGTCGTTCAAT-GGATCAATAGCTAAACTGCAA-GTGAACAAGCTGACCGAA-GAGAAATCGGGTCTCTATAAAT-GTCATGCAAAGTGTGACTAT-GGAGAAGGTCAAAGCAGTG-CAATGGTCAAAATCGAA-CAGTCTGGTAGGTTTTCTCT-GATTCTCTTTTTGATAA-CACGATTGATAACACAAAAAG-TATAGAATTTTTCGAAATTAA-CAAGAAAATAGTATGAACCAGT-GGCCGGATGGTCTAGAGGTAT-GATTCTCGCTTCGGGTGCGAGA-GGTCCCGGGTTCGATTCCCGGT-TCGGCCCAAACATTTTATAAT-TCAGATGTGGAAGAAGAACT-CATGAAGCATAGAAAAGACG-CGGAGGATGAATATCA	AAAAGAAGAACAGAAATCG-CAGACGCTTCAAGCTGAAAC-CAAAAAGCG	-
15	28348	28409	62	TGGAACAC-CAGGAAAGC-CAGGAAGAC-CAGAAATTGTT-GAT	ACTGATAATGATCATATCGA-TATCAAATGGGATCCTCCACGT-GACAACGGTGGATCACCAGT	TGATCATTACGACATTGA-GAGGAAGGATGCAAAGACTG-GACGCTGGAT	-

* Underlined sequence is templated

3

Suppression of genome instability by Y-family TLS polymerase REV-1 in *C. elegans*

Ivo van Bostelen, Robin van Schendel, Sophie Roerink & Marcel Tijsterman

Abstract

Translesion synthesis (TLS) polymerases promote the replication of damaged DNA through their ability to catalyze nucleotide addition to growing DNA chains at sites where DNA lesions block the replicative polymerases. By this action, TLS polymerases allow cells to complete the duplication of damage-containing genomes and thus prevent checkpoint activation, genome instability and cell death. The price to pay is mutation induction because TLS polymerases have, in comparison to replicative polymerases, reduced fidelity and are thus generally considered to be pro-mutagenic. In this study we have used *C. elegans* to determine the contribution of the Y-family polymerase REV-1 on long-term stability of an animal genome. Surprisingly, we found REV-1 to both stimulate and suppress spontaneous mutagenesis during unperturbed propagation. By stimulating bypass REV-1 prevents the persistence of ssDNA gaps that are converted to small deletions by alternative end joining of ensuing double-strand breaks. Thus, opposite of what is the current dogma, the action of REV-1 during unperturbed growth is predominantly anti-mutagenic: it prevents the accumulation of deletions at the cost of less detrimental SNVs. In addition, we found the level of spontaneous lesions that depend on REV-1 action to be surprisingly low: only 1 lesion in ~100 rounds of genome replication, which equates to 1 block per ~10^{10} bases, requires REV-1 action. Our findings augment the concept that ensuring replication progression outweighs near-perfect conservation of genetic information in animal cells.

Introduction

Although mutagenesis is a prerequisite for evolution, spontaneous mutations are also life threatening as they are at the basis of inborn diseases and age-related pathologies like cancer. To suppress these detrimental effects several mechanisms have evolved to prevent the occurrence of mutations. For example, during the copying of DNA, the combined action of exonuclease activity of the replicative polymerases, which removes erroneously incorporated nucleotides during synthesis, and the mismatch repair (MMR) pathway, which repairs mismatched bases after the fact, provides an estimated 10.000-fold increase in copying accuracy [1,2]. Apart from replication errors, so-called spontaneous mutations can result from replicating damaged DNA caused by endogenous processes in the cell. For example, oxidative metabolites can react with DNA, damage bases, and in that way hamper replication [3,4]. Efficient and unperturbed DNA synthesis is essential for survival since stalling of replication can lead to collapse of the replication fork followed by formation of highly toxic and mutagenic DNA double stranded breaks (DSBs) that may result in genomic rearrangements or cell death. A number of pathways have evolved to remove potential replication blocks and repair the DNA, including base excision repair (BER) and nucleotide excision repair (NER), where newly synthesized DNA replaces the damaged DNA at the site of the potential replication block [4,5]. Additionally, damage tolerance pathways have evolved to allow for continuation of the cell cycle in the presence of DNA lesions. A well-studied mechanism to tolerate DNA damage is translesion synthesis (TLS). While replicative polymerases stall at damaged bases, specialized TLS polymerases are able to synthesize DNA opposite of these blocks. Lesions can be bypassed directly when the replicative polymerase is temporarily switched with a TLS polymerase at the replication fork during S-phase, or single strand DNA gaps at the site of lesions remain and bypass and gap filling occurs after S-phase [6].

In eukaryotes TLS is mediated by Y-family polymerases Polη, Polκ, Polι and Rev1 and the B-family polymerase Polζ composed of catalytic subunit Rev3 and regulatory subunit Rev7. These TLS polymerases lack proofreading activity and have wide catalytic centers to allow for replication across damaged bases and DNA synthesis from misaligned primer termini. These characteristics cause TLS polymerases to have lower fidelity then replicative polymerases, making them inherently error prone [7,8]. Whereas some types of lesions require only a specific Y-family polymerase, other types require the sequential action of a two or more TLS polymerases [9-11]. Two modes for TLS have been proposed. The first acts on lesions that have low impact on the DNA helix structure and are still 'readable'. For these lesions the default TLS polymerase Polη is suggested, mostly resulting in error-free bypass. The second mode of action is required for lesions that are harder to read and strongly distort the DNA helix. In that case the concerted efforts of one or more Y-family polymerases in combination with the Rev1

dependent activity of Polζ are required. This second mechanism is thought to be much more mutagenic [6,12].

REV1 has a similar structure as the other members of the Y-family, Polη, Polκ and Polι, but its catalytic activity is limited to the incorporation of deoxycytidine (dC) residues [7]. *In vitro*, REV1 was shown incorporate dCs across undamaged or damaged guanines but also opposite adenines, uracil and abasic sites [13-15]. *In vivo*, REV1 plays a role in bypass of lesions that are caused by lipid peroxidation [16,17]. Various studies have demonstrated that REV1 plays non-catalytic roles via interactions with other proteins. The BRCT domain of REV1 interacts with PCNA and is involved in the bypass of UV-C induced lesions in mouse ES cells [18-20]. The C-terminal part of mammalian REV1 contains ubiquitin binding motifs (UBMs) that can interact with ubiquitinated proteins like PCNA-K164Ub, a region able to interact with other Y-family polymerases, and a motif that interacts with the REV7 subunit of B-family TLS polymerase Polζ [7,21-24] (Suppl. fig. 1). REV1 could act as a master regulator of TLS, instead of acting as a true polymerase: via its interactions with PCNA, Y-family polymerases and Polζ, REV1 may provide a 'molecular scaffold' that is central to TLS [6,12]. REV1 might also have functions outside of canonical TLS since research in avian DT40 cells has shown a role for this polymerase in maintenance of epigenetic stability at G-quadruplex structures, possibly by facilitating replication through these hard-to-replicate secondary structures [25,26]. Finally, a yet unexplained role for REV1 (together with Polζ) in homologous recombination (HR) break repair has been described [27,28].

The action of TLS polymerases results in base substitutions, which may appear a detrimental process for cells, but an occasional base substitution may outweigh the deleterious consequences of complete replication fork blocks. Unrepaired breaks can result from collapsed replication and lead to cell cycle arrest and cell death [29,30]. Although the molecular details of translesion synthesis become more and more understood it remains unclear how TLS action affects genome maintenance or influences spontaneous mutagenesis either positively or negatively, on a genome-wide scale. The model system *C. elegans* is well suited to address this question through whole genome sequencing (WGS) revealing the diverse types of mutations, such as base substitutions, insertions/deletions (indels) or large genomic rearrangements that accumulate over generations. Here, we report that the TLS polymerase REV-1 safeguards replicative potential and genomic stability that are threatened by spontaneous DNA lesions.

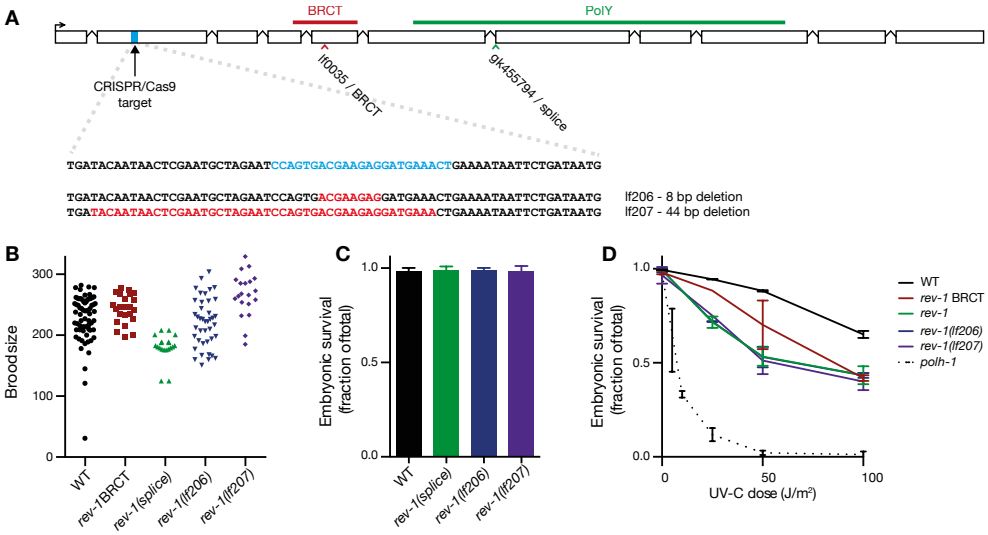


Figure 1. Generation and characterization of rev-1 alleles. *A) Schematic representation of the rev-1 ORF. Exons are to scale, introns are not. In blue the CRISPR/Cas9 sgRNA target is indicated with the two generated knock-out alleles below. The red bar indicates the BRCT domain with the location of the point mutation (lf0035/BRCT) below. The green bar indicates the polymerase domain with the splice site mutation (gk455794) below. B) Brood size quantification of the different rev-1 mutant alleles. Each data point represents the total brood of one animal. C) Quantification of embryonic lethality of the same broods as in B. D) Embryonic survival in response to UV-C exposure of the indicated genotypes.*

Results

Generation and characterization of the *rev-1* alleles

To study the role of REV-1 in maintenance of genome stability we obtained several loss-of-function alleles. The allele most used in this study is *rev-1(gk455794)*, generated in the million mutation project [31], has a point mutation in the acceptor splice site of exon 7 that generates a potential truncated product lacking the polymerase domain and C-terminal part. Although it is likely that this mutant is a null allele because it lacks the whole Y-family polymerase domain and other C-terminal parts, we also generated two mutants targeting exon 2 of the rev-1 gene via the CRISPR/Cas-9 method [32-38]. We independently isolated two small genomic deletions at the CRISPR targeted site that both lead to frame shifts running into early stops (Fig. 1A). We consider the alleles *rev-1(gk455794)*, *rev-1(lf206)* and *rev-1(lf207)* to be null alleles and will refer to *rev-1(gk455794)* simply as *rev-1* from now on. Using a different technology [39], we also isolated a mutation causing a G283>D amino acid substitution in the highly conserved G283 residue of the BRCT domain: the *C. elegans* G283 residue aligns to G193 in yeast and

G76 in mice Rev1 (Suppl. fig. 1). The widely studied G193>R, G193>D and G193>V mutants in yeast, and the G76>R mutant in mice were shown to abolish the functionality of the BRCT domain [21,40,41]. We refer to this allele as *rev-1BRCT*.

None of the *rev-1* mutants showed brood sizes that were substantially different from wild type (WT) controls (Fig. 1B). Also, *rev-1* knock-out mutants showed similar levels of very marginal embryonic lethality as WT controls (Fig. 1C). Both these observations show that REV-1 is not essential for proliferation under unchallenged conditions. The hypersensitivity to UV-C-induced DNA lesions observed in REV1 deficient yeast and MES cells suggests a conserved role for REV1 in TLS of photoproducts [18,42,43]. To test whether this is true for *C. elegans* REV-1 we exposed young adult worms to different doses of UV-C light and determined the embryonic lethality of their brood. We established that *rev 1* mutants have reduced embryonic survival as compared to WT controls in response to UV-C exposure but not to the extent as we showed for mutants defective in Polη, a TLS polymerase previously implicated to be essential in protection against UV induced damage [44,45] (Fig. 1D). A similar degree of sensitivity is observed in all three different *rev-1* knockout mutants; the *rev-1BRCT* hypomorph shows an intermediate phenotype, which is consistent with a partial loss of function.

rev-1 and *rev-1BRCT* mutants show increased levels of spontaneous DSBs

When TLS is impaired replication forks can collapse at sites of damaged bases and form DSBs. In the mitotic compartment of the *C. elegans* gonad DSBs can be visualized by staining for RAD51 foci. RAD51 foci in the mitotic compartment in the distal tip of the gonadal arms are rare events in healthy WT controls. However, after damage induction or as a result of spontaneous DNA damage in TLS mutants the number of foci increases while the brood size and embryonic lethality remain similar to WT control [46,47]. Both *rev-1* and *rev-1BRCT* mutants have significantly increased levels of spontaneous RAD51 foci in the mitotic compartment of the gonad, arguing that REV-1 suppresses the formation of DSBs ($p<0.01$; Fig 2A, B).

REV-1 protects against the formation of genomic deletions and rearrangements larger than 50 base pairs.

While *rev-1* mutants displayed elevated levels of spontaneous DSBs no effect was observed on proliferation (Fig 1B, C & Fig. 2A), arguing for repair of RAD51-coated DSBs. Because repair of DSBs can be mutagenic, we asked whether the observed increase of spontaneous breaks affects mutation induction in *rev-1* mutants. To answer this we made use of an established mutagenesis assay; the *unc-93* reversion assay [48,49]. Animals carrying the toxic allele *unc-93(e1500)* have very poor capacity to move and also grow slowly. A mutation that kills the protein via mutation of an essential amino acid or via disruption of the ORF will lead to reversion to a WT-like phenotype

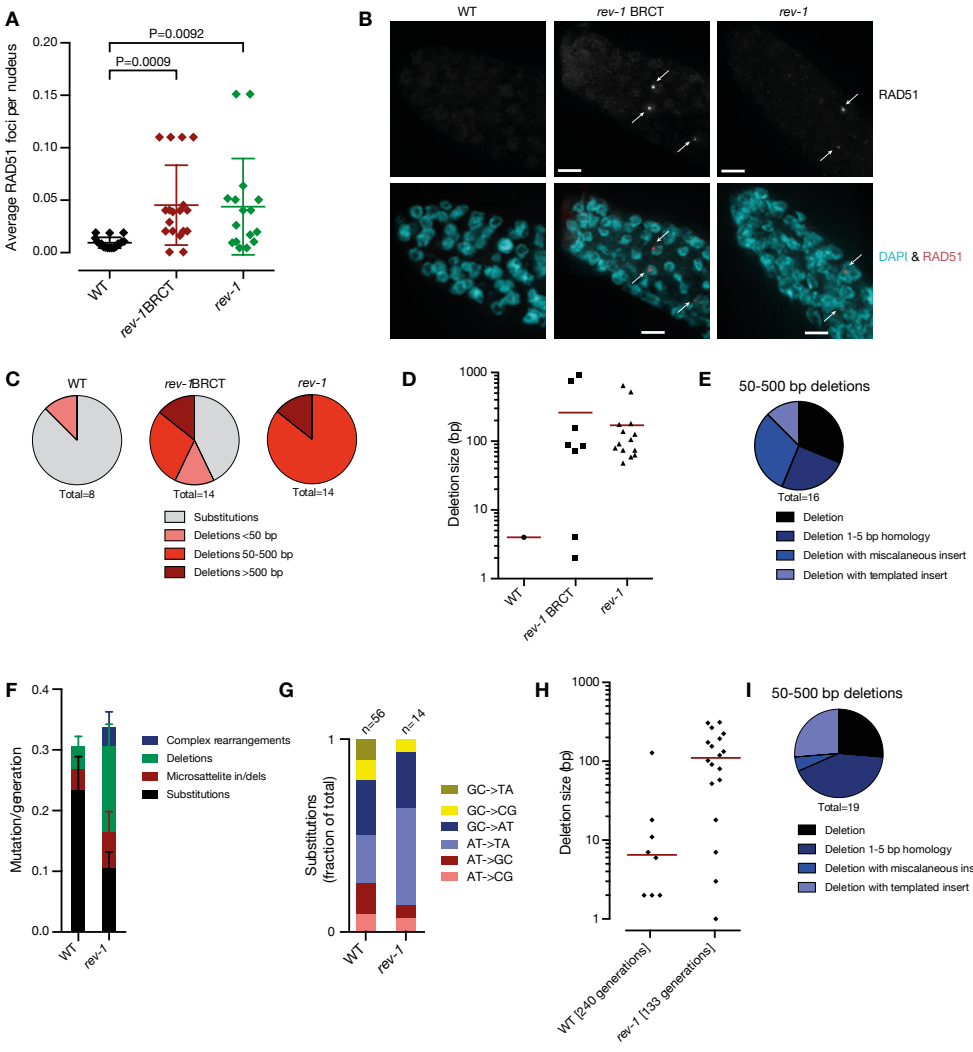


Figure 2. Spontaneous DNA damage and mutagenesis in REV-1 deficient animals. *A) Quantification of spontaneous RAD51 foci in the mitotic compartment in the distal tip of the gonadal arms. B) Representative images of RAD51- and DAPI-stained distal tips of gonads. C) Mutations found in unc-93 assay in the different genetic backgrounds. D) Size of the deletions found the unc-93 assay. E) Properties of the 50-500 bp deletions found the unc-93 assay in both rev-1 mutants. F) The frequencies of the different categories of mutations accumulated over generations in WT and rev-1 mutants. G) SNP distribution in the mutation accumulation (MA) data. H) Size distribution of the deletions from MA data, excluding microsatellite indels. I) Characteristics of the 50-500 bp deletions in the MA data. See supplemental table 1 for sequence data.*

allowing for efficient selection of spontaneous mutants. Mutations in the suppressor genes sup-9, sup-10, sup-11 or sup-18 will also revert animals to a WT-like phenotype. We isolated 30 revertants for each genotype and subsequently sequenced the entire unc-93 locus. In WT, *rev-1BRCT* and *rev-1* backgrounds we found 8, 14, and 14 causative mutations in *unc-93(e1500)*, respectively. The other revertants likely carry mutations in one of the suppressor genes, but these genes were not sequenced. In the WT background, 7 of the 8 mutations are single nucleotide variations (SNVs) that disrupt gene function via amino acid substitution, introduction of an early stop or loss of a splice site. One deletion allele of 4 bp was found. Similar base substitutions and deletions of few bases were also found in the *rev-1BRCT* mutant, but 6 of the 14 mutations are larger deletions (>50 bp). In worms lacking any REV-1 activity all causative mutations were found to be >50 bp in size (Fig. 2C, D). From this data we conclude that the BRCT domain of REV-1 is involved in bypass of endogenous lesions and that REV-1 has an important role in spontaneous mutagenesis by suppressing the formation of small genomic deletions.

REV-1 dependent TLS results in base substitutions but protects the genome from genomic deletions and rearrangements

While the *unc-93* reversion assay is an established method to study mutagenesis it has a few drawbacks: it is limited to the genomic locus of *unc-93* and the only mutations that will be found are those that disrupt gene function. Deletions, for example, are more likely to disrupt gene function then SNVs, creating a bias towards detecting deletions over SNVs and making it difficult to measure mutations frequencies. In order to study spontaneous mutagenesis in an unbiased way and to allow us to determine mutation frequencies accurately we performed WGS of animals of populations that were grown for multiple generations. In WT animals we find on average ~0,23 base substitutions per animal generation, ~0,04 microsatellite indels and ~0,04 other small deletions. While the frequency of microsatellite indels remains more or less the same in the *rev-1* mutant the frequency of substitutions is halved to ~0,11 per generation and the frequency of deletion formation increases 3,5-fold to ~0,14 per generation. Additionally, few larger complex genomic rearrangements like tandem duplications and inversions occur in the *rev-1* mutant animals; these were not observed in WT animals (Fig. 2F). The distribution of substitutions in the *rev-1* mutant compared to the distribution in WT does not show marked difference (Fig. 2G); the variations are indicative of experimental variation in the small set (n=14) of substitutions that were found in the *rev-1* mutant. In addition to the striking increase in the frequency of non-microsatellite deletions there is also a marked difference in the size of the deletions between those observed in WT and *rev-1*: deletions of 50-500 bp are now the prominent category, while these are very rare in WT animals (Fig. 2H).

Assuming that lesions that absolutely require REV-1 for bypass will result in a deletion or complex rearrangement in the absence of REV-1, and estimating the number of cell divisions needed to go from one generation to the next based in the extensive knowledge on *C. elegans* cell lineage [50], we find that the occurrence of spontaneous lesions that depend on REV-1 action is surprisingly low: only 1 lesion in ~100 rounds of genome replication, which equates to 1 block per ~10^{10} bases, requires REV-1 action.

Loss of REV-1 does not lead to G-quadruplex instability in *C. elegans*

A potential role in replication of sequences that can form G-quadruplexes was described for REV1 in DT40 cells [25,26]. We tested whether *C. elegans* REV-1 is required for the maintenance of genome stability at G-quadruplexes as has been described previously for *C. elegans* FANCJ/DOG-1 [51,52]. We first searched for the presence of G-quadruplex motifs within the deleted sequence in *rev-1* mutant animals in the data from the *unc-93* and MA assays, but found none, arguing that spontaneous mutation induction is at sites of base damage and not on structured DNA. For further conformation, we tested if loss of functional REV-1 affected the nature and frequencies of G-quadruplex deletions that are formed in animals that lack the G-quadruplex resolving helicase DOG-1. In these animals G-quadruplex are more likely to constitute a block to replication [51,52]. However, we did not observe any significant difference on frequency and deletion size between *dog-1* and *dog-1; rev-1* mutant animals studying the endogenous G-quadruplex site qua1466, a very potent inducer of deletions in *dog-1* animals (Suppl. fig. 2).

Absence of REV-1 leads to substrates for TMEJ/Polθ

The size distribution of the spontaneous deletions that form over generations when lacking REV-1 activity is the same as for those found in TLS mutants lacking POLH-1 and POLK-1 [47]. This mutational footprint results from Polθ mediated end-joining (TMEJ) of replication associated DSBs and is further characterized by microhomology in the break sites and occasional inserts homologous to sequences flanking the deletion [47,52,53]. When we study the 50-500 bp deletions obtained in MA assay in the *rev-1* mutant we find similar characteristics: 36% shows the use of microhomology at the break sites and 23% contain templated inserts that originate from the flanks of the breaks (Fig 2I), comparable results are observed in the *unc-93* assay (Fig. 2I). In order to verify that TMEJ repairs spontaneous DSBs in *rev-1* mutants we employed the *unc-93* reversion assay in animals deficient for REV-1 and Polθ *(rev-1;polq-1)* and compared the repair footprints to the mutations found in the *rev-1* mutant. For the lager deletions the exact break points could not be determined by PCR and sequencing because all amplicons covering the unc-93 gene were deleted. We estimated the size of the large deletions by the loss of the amplicons in the flanks of unc-93 and found that all *rev-1;polq-1* revertants carry large deletions spanning several kb. Two of the five deletions extend into the neighboring

lethal genes indicting a deletion size >10 kb (Fig. 3A). We conclude that in a *rev-1* deficient worm small deletions form induced by stalled replication forks at endogenous lesions via TMEJ. To substantiate our conclusion that Polθ acts on substrates that result from the inability to bypass damaged DNA we tested Polθ dependency also in response to known bona fide replication blocks *i.e.* UV-C-induced photoproducts. We found that loss of just Polθ does not sensitizes animals to UV-C but combined loss of Polθ and REV-1 *(rev-1;polq 1)* results in a significantly higher sensitivity than the *rev-1* single mutant (Fig. 3B). This finding reinforces the notion that replication blocking lesions require TLS polymerases to prevent larger genomic insult, such as deletions resulting from DSB repair.

Discussion

We here show that REV-1 safeguards survival in *C. elegans* upon exposure to UV light but is not needed under unchallenged conditions in the laboratory environment. During unperturbed growth, spontaneous DNA lesions arise that are bypassed by REV-1 dependent TLS, which can result in base substitutions. Without REV-1 these lesions cause persistent replication blocks that lead to the formation of small genomic deletions ranging from 50 to 500 bp in size and to the formation of larger complex genomic rearrangements. The characteristics of the deletions are consistent with TMEJ of replication associated DSBs. Accordingly, loss of TMEJ activity in a REV-1 deficient background further sensitizes animals to UV-C induced replication blocks and completely alters the mutational footprint of spontaneous replications block: the observed deletions become at least several kilobases in length, analogous to the mutational consequences of other persistent replication blocks in TMEJ deficient *C. elegans* [47,52].

Rev1 was named after its 'reversion-less' phenotype in yeast: in rev1 mutants UV-C induced mutagenesis is considerably lower than in WT [42,54]. Similar function has been assigned to human REV1 in response to UV induced DNA damage [43,55] and in mouse cells [19,56]. In line with the mutagenicity of REV1 dependent TLS we observe a reduction in base substitution frequency in the *rev-1* mutant. If the deoxycytidine transferase activity of REV-1 would be responsible for a substantial part of substitutions, *e.g.* the incorporation of dCs opposite abasic sites, REV-1 activity would be responsible for AT>GC, AT>CG and GC>CG substitutions. There are minor reductions in AT>GC and AT>CG substitutions, however we also see a loss of GC>TA substitutions in the *rev-1* mutant. Therefore, it is unlikely that the polymerase activity of REV-1 is responsible for a mayor part of base substitutions in WT animals. We propose that the lower levels in substitutions in the *rev-1* mutants are caused by a generally lower TLS efficiency of a mutagenic sub-pathway, due to the loss of the regulatory function of REV-1 in TLS. In previous studies it was reported that Polη and Polκ perform relatively error-free TLS

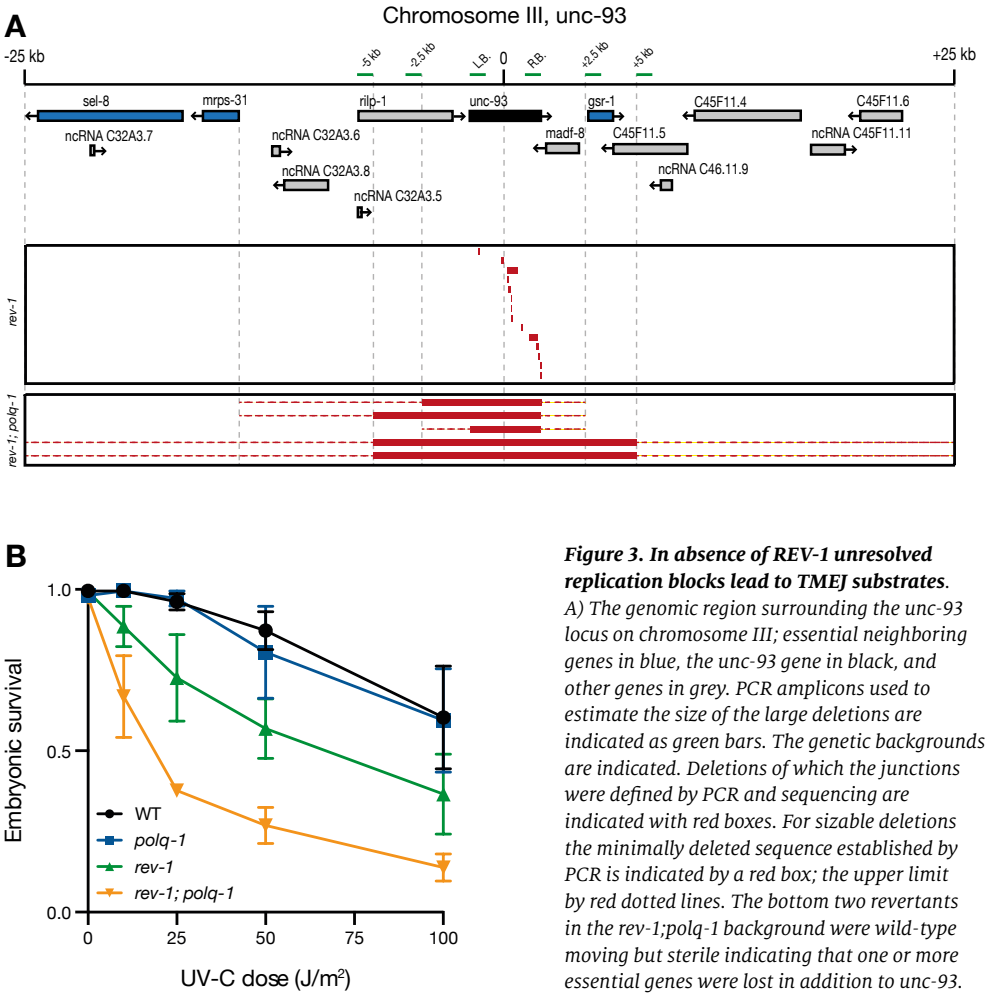


Figure 3. In absence of REV-1 unresolved replication blocks lead to TMEJ substrates. *A) The genomic region surrounding the unc-93 locus on chromosome III; essential neighboring genes in blue, the unc-93 gene in black, and other genes in grey. PCR amplicons used to estimate the size of the large deletions are indicated as green bars. The genetic backgrounds are indicated. Deletions of which the junctions were defined by PCR and sequencing are indicated with red boxes. For sizable deletions the minimally deleted sequence established by PCR is indicated by a red box; the upper limit by red dotted lines. The bottom two revertants in the rev-1;polq-1 background were wild-type moving but sterile indicating that one or more essential genes were lost in addition to unc-93. B) Embryonic survival in response to UV-C exposure of the indicated genotypes.*

on lesions that are more abundant than those that require REV-1: loss of POLH-1 and POLK-1 induced ~2 deletions per generation, which is 14-fold more than observed in the *rev-1* mutant [44]. An explanation could be that REV-1 is required for a subset of lesions that needs the concerted efforts of multiple TLS polymerases: Polη or Polκ performs insertion of one or a few bases opposite the damage followed by REV-1 dependent extension of the aberrant primer terminus by Polζ or possibly Polκ. It is thought this complex sub-pathway of TLS induces more base substitutions [6].

It is interesting that REV1 deficient mouse embryonic fibroblasts, yeast and human fibroblast cell lines have no proliferative problems, yet Rev1-/- mice develop poorly arguing that the consequences of REV1 loss more prominently manifest on an

organismal level [19,42,43,56]. Unresolved replication blocks in Rev1 mutant mice may become a more serious problem due to accumulation of genomic rearrangements or cell death in tissues harboring cycling cells leading to loss of proliferative potential and ageing. The fact that we observe no proliferative defect in worms may be explained by the relatively high tolerance of worms to DNA damage and by relatively small size of the *C. elegans* genome resulting in less replication stress. Our observations suggest that the levels of spontaneous DSBs that form in REV-1 deficient worms are low and efficiently repaired. As a result, the level of replication-associated DSBs are insufficient to induce noticeable proliferative defects.

While it is widely established that REV1 acts in TLS, some studies have described roles in other pathways. Okada *et al*. and Sharma *et al*. describe involvement of REV1 in DNA repair via homologous recombination (HR) [27,28]. We cannot formally rule this out but have found no experimental support for a role for *C. elegans* REV-1 in HR; defective HR results in reduced brood size and high embryonic lethality [57]. Additionally, the spontaneous mutations we find in *rev-1* mutant genomes are consistent with the footprints found resulting from unresolved replication blocks in TLS mutants *polh-1* and *polh 1; polk-1*, that still have functional REV-1. Also, we have not found a role for *C. elegans* REV-1 in replicating through G-quadruplex structures as was observed in DT40 chicken B lymphocyte cells [25,26]. Both these additional functions could be reserved for vertebrate REV1 and dependent on C-terminal parts of the protein that are lacking in yeast and *C. elegans* REV-1. We do show a conservation of function of the BRCT domain from mice to worms. Jansen *et al*. describe that in mammalian cells the BRCT domain of Rev1 was important for bypass of UV-C induced lesions [18]. Here we confirm that loss of a functional BRCT domain in REV-1 sensitizes *C. elegans* for UV-C and we show a novel role of this domain in the bypass of endogenous lesions.

Based on our NGS data we conclude that during unchallenged *C. elegans* growth the number of spontaneous lesions that for their bypass depend on REV-1 action is remarkably low: in *rev-1* mutants we found a reduction of SNVs of ~0.1 per generation (each generation has at least 10-15 cell divisions), arguing that REV-1 action produces ~1 mutation per 10^{10} bases. Strikingly, we found a matching increase in the number of deletions in the absence of REV-1. This unexpected outcome argues that REV-1 action is mostly mutagenic during TLS of spontaneous base damage: if *e.g*. 99% of REV-1 dependent TLS would be error free, then we would have expected the number of deletions to go up manifold in a *rev-1* mutant background. The matching levels of SNV reduction and deletion increase is most simply explained by assuming that only 1 lesion per 10^{10} bases absolutely depend on REV-1 to be bypassed, while in the absence of REV-1 such lesions will constitute a persistent block to replication, leading to DSBs that require TMEJ. An important inference from our data is that although REV-1-dependent TLS introduces base substitutions, it protects the genome against the formation of

small genomic deletions and larger complex rearrangements and thereby helps to maintain genome stability. Because the latter outcomes are per case more detrimental to the integrity of our genetic information than SNVs, our study leads to the notion that the action of REV-1 during unperturbed growth is predominantly anti-mutagenic, challenging the current dogma that TLS action is harmful for accurate transmission of our genetic information. While it is so, its action prevents a far greater harm *i.e.* deletions and genomic rearrangements that have a much more prominent disturbing effect. Our findings thus support the notion that ensuring replication progression outweighs near-perfect replication of genetic material.

Methods

C. elegans genetics

All strains were cultured according to standard methods [58]. The N2 Bristol strain was used as WT control. The alleles *polq-1(tm2026), rev-1(gk455794), unc-93(e1500)* were obtained from the Caenorhabditis Genetics Center, Minnesota, USA. The *rev-1(BRCT)* allele was isolated in our own lab via a random mutagenesis approach described in [39]. The *rev-1(lf206)* and *rev-1(fl207)* KO alleles were obtained via CRISPR/Cas9 mediated genome editing as described in [32-37] and below in further detail.

CRISPR/Cas9 mediated generation of rev-1 K.O. alleles

For CRISPR/Cas9 mediated targeting we used the following sequence (which includes the PAM site) in exon 2 of rev-1: AGTTTCATCCTCTTCGTCACTGG. Cloning in the appropriate vectors was done as described in [32-37]. Plasmids were injected using standard *C. elegans* microinjection procedures. Briefly, 1 day before injection, L4 animals were transferred to new plates and cultured at 15 degrees. Gonads of young adults were injected with a solution containing 20 ng/µl, pDD162 (Peft-3::Cas9, Addgene 47549; ref. [59]), 20 ng/µl pMB70 (u6::sgRNA with rev-1 target, 10 ng/µl pGH8, 2.5 ng/µl pCFJ90 and 5 ng/µl pCFJ104. Progeny (F1) animals that express mCherry were picked to new plates 3–4 days post injection and allowed to produce offspring. Of each F1 plate 10 F2 animals were pooled, lysed and genotyped. Genotyping was done by PCR amplification of a 480 bp product around the CRISPR/Cas9 target site. Subsequent restriction with MaeIII enzyme of the WT sequence would result in 2 fragments (91 bp + 389 bp). A mutation at the site of the predicted break site would likely disrupt the MaeIII recognition site resulting in an uncut PCR product. We isolated 4 alleles, 2 of which were small out of frame deletions. (Fig.1)

Brood size and embryonic lethality assay

To determine the brood size, we singled L4 animals on OP50 plates. Every day, we

transferred the mother to a new plate and one day later quantified the number of embryos and larvae on the plate. We quantified the broods and embryonic lethality of at least 20 animals per genotype.

UV-C survival assays

To measure germline sensitivity to UV, staged young adults (one day post L4) were transferred to empty NGM plates and exposed to different doses of UV-C. Per dose and genotype 3 plates with 3 adults were set up on fresh NGM plates with OP50 and allowed to lay eggs for 24 hours. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later the number of non-hatched eggs and surviving progeny was determined.

RAD51 antibody staining, imaging and quantification

Animals were synchronized by picking L4 stage worms 22h before dissection. Worms were dissected in Egg buffer with 0.1% Tween and Levamisole to expose gonads. Most of the buffer was removed and the sample with cover glass was transferred to a Superfrost Plus slide and flash frozen on a metal block in dry ice. Upon complete crystallization of the sample the cover glass was quickly removed (freeze-cracking) followed by post-fixation in 4% PFA. After washing the samples were incubated with RAD51 antibody (rabbit polyclonal from SDIX/Novus Biologicals, cat# 29480002, used at 1:1000 in PBST+0,5%BSA) overnight at room temperature. Alexa anti-rabbit 488, 1:500, was used as secondary antibody and incubated for 2h at room temperature. DNA was visualized by DAPI staining and the slides were finished with Vectashield. Imaging and processing were done on a Leica DM6000 B microscope. The data obtained (Fig. 2A, B) are from at least 3 independent experiments. Each data point represents an average value of the mitotic zone of one gonad.

Mutation accumulation lines and whole genome sequencing

Mutation accumulation (MA) lines were established by transferring single F1 animals that originated from a single parent, starting 6 clonal MA lines. Of each line three worms of the next generations were transferred to a new plate, marking every generation. MA lines were propagated for approximately 50 generations. At the end of MA timeline, of each line single animals were transferred to new plates and propagated to obtain full clonal plates for DNA isolation. Worms were washed off in M9 and to remove as much bacteria as possible from the intestines the worms wore incubated on a shaker for 2h at RT. Subsequently, genomic DNA was isolated using a Blood and Tissue Culture Kit (Qiagen). Whole genome sequencing and bio-informatics were performed as described in [53].

G-quadruplex stability on qua1466

Qua1466 is a genomic sequence [GGGAGGGCGGGCGGG] with genomic location IV:11326500..11326514 (build WBcel235), that can potentially form a G-quadruplex structure. To assay genomic instability and the formation of deletions at this site we perform a nested PCR reaction on lysed animals using the following primers: external forward CAAATAAGTATTGGGCCGAAACC; external reverse AAGGAACACCTTCAAGACTCC, internal forward CTGCGAACTTCTGACGAATTTG, internal reverse TTGACTCCTCCTCTTCTGGC. As template for the external PCR 1 µl of a 15 µl lysis with 5 worms was used. 0.5 µl of the external PCR was used as template for internal PCR. 10 µl of internal PCR product was run for 1 hour at 120V on a 1% agarose gel.

unc-93(e1500) mutagenesis assay

To pick up spontaneous mutations in the *rev-1(gk455794)* and *rev-1(gk455794);polq-1(tm2026)* backgrounds, we used a mutagenesis assay based on reversion of the so-called "rubber band" phenotype, caused by a dominant mutation in the muscle gene unc-93 [48,49]. Reversion of the *unc-93(e1500)* phenotype is caused by homozygous loss of *unc-93(e1500*) or one of the suppressor genes sup-9, sup-10, sup-11, and sup-18. For both genotypes 400 animals were singled to 9 cm plates. These plates were grown until starvation and of each plate an equal amount (chunks of 2 x 2 cm) were transferred to fresh 9 cm plates. Before these plates reached starvation, they were inspected for wild-type moving animals. From each starting culture, only one revertant animal was isolated to ensure independent events. Of each genetic background we randomly selected 30 revertants and sequenced the *unc-93* gene. When large deletions occurred that deleted the amplicons used for sequencing *unc-93*, we established the approximate size of the deletion with PCR amplicons of approximately 500 bp located at the borders of the gene and 2.5 kb and 5 kb up and downstream of the *unc-93* gene. The locations of these amplicons are indicated with green bars in figure 3A.

Acknowledgements

Several strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Programs . This work is supported by the European Research Council (ERC) and the Netherlands Organization for Scientific Research (NWO).

References

[1] McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res 2008;18:148–61. doi:10.1038/cr.2008.4.

[2] Fishel R. Mismatch repair. Journal of Biological Chemistry 2015;290:26395–403. doi:10.1074/jbc.R115.660142.

[3] Barnes DE, Lindahl T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu Rev Genet 2004;38:445–76. doi:10.1146/annurev.genet.38.072902.092448.

[4] Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell 2010;40:179–204. doi:10.1016/j.molcel.2010.09.019.

[5] Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JHJ. Understanding nucleotide excision repair and its roles in cancer and ageing. Nature Reviews Molecular Cell Biology 2014;15:465–81. doi:10.1038/nrm3822.

[6] Jansen JG, Fousteri MI, de Wind N. Send in the clamps: control of DNA translesion synthesis in eukaryotes. Mol Cell 2007;28:522–9. doi:10.1016/j.molcel.2007.11.005.

[7] Prakash S, Johnson RE, Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem 2005;74:317–53. doi:10.1146/annurev.biochem.74.082803.133250.

[8] Knobel PA, Marti TM. Translesion DNA synthesis in the context of cancer research. Cancer Cell Int 2011;11:39. doi:10.1186/1475-2867-11-39.

[9] Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. Nature 2000;406:1015–9. doi:10.1038/35023030.

[10] Prakash S, Prakash L. Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. Genes & Development 2002;16:1872–83. doi:10.1101/gad.1009802.

[11] Gan GN, Wittschieben JP, Wittschieben BØ, Wood RD. DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res 2008;18:174–83. doi:10.1038/cr.2007.117.

[12] Jansen JG, Tsaalbi-Shtylik A, de Wind N. Roles of mutagenic translesion synthesis in mammalian genome stability, health and disease. DNA Repair 2015;29:56–64. doi:10.1016/j.dnarep.2015.01.001.

[13] Nelson JR, Lawrence CW, Hinkle DC. Deoxycytidyl transferase activity of yeast REV1 protein. Nature 1996;382:729–31. doi:10.1038/382729a0.

[14] Haracska L, Unk I, Johnson RE, Johansson E, Burgers PM, Prakash S, et al. Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. Genes & Development 2001;15:945–54. doi:10.1101/gad.882301.

[15] Zhang Y, Wu X, Rechkoblit O, Geacintov NE, Taylor J-S, Wang Z. Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Res 2002;30:1630–8. doi:10.1093/nar/30.7.1630.

[16] Zhou Y, Wang J, Zhang Y, Wang Z. The catalytic function of the Rev1 dCMP transferase is required in a lesion-specific manner for translesion synthesis and base damage-induced mutagenesis. Nucleic Acids Res 2010;38:5036–46. doi:10.1093/nar/gkq225.

[17] Wiltrout ME, Walker GC. The DNA polymerase activity of Saccharomyces cerevisiae Rev1 is biologically significant. Genetics 2011;187:21–35. doi:10.1534/genetics.110.124172.

[18] Jansen JG, Tsaalbi-Shtylik A, Langerak P, Calléja F, Meijers CM, Jacobs H, et al. The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Res 2005;33:356–65. doi:10.1093/nar/gki189.

[19] Jansen JG, Tsaalbi-Shtylik A, Hendriks G, Gali H, Hendel A, Johansson F, et al. Separate domains of Rev1 mediate two modes of DNA damage bypass in mammalian cells. Mol Cell Biol 2009;29:3113–23. doi:10.1128/MCB.00071-09.

[20] Guo C, Sonoda E, Tang T-S, Parker JL, Bielen AB, Takeda S, et al. REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. Mol Cell 2006;23:265–71. doi:10.1016/j.molcel.2006.05.038.

[21] Guo C, Fischhaber PL, Luk-Paszyc MJ, Masuda Y, Zhou J, Kamiya K, et al. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. Embo J 2003;22:6621–30. doi:10.1093/emboj/cdg626.

[22] Kikuchi S, Hara K, Shimizu T, Sato M, Hashimoto H. Structural Basis of Recruitment of DNA Polymerase by Interaction between REV1 and REV7 Proteins. Journal of Biological Chemistry 2012;287:33847–52. doi:10.1074/jbc.M112.396838.

[23] Wojtaszek J, Liu J, D'Souza S, Wang S, Xue Y, Walker GC, et al. Multifaceted recognition of vertebrate Rev1 by translesion polymerases ζ and κ. Journal of Biological Chemistry 2012;287:26400–8. doi:10.1074/jbc.M112.380998.

[24] Wojtaszek J, Lee CJ, D'Souza S, Minesinger B, Kim H, D'Andrea AD, et al. Structural Basis of Rev1-mediated Assembly of a Quaternary Vertebrate Translesion Polymerase Complex Consisting of Rev1, Heterodimeric Polymerase (Pol) , and Pol. Journal of Biological Chemistry 2012;287:33836–46. doi:10.1074/jbc.M112.394841.

[25] Sarkies P, Reams C, Simpson LJ, Sale JE. Epigenetic instability due to defective replication of structured DNA. Mol Cell 2010;40:703–13. doi:10.1016/j.molcel.2010.11.009.

[26] Schiavone D, Guilbaud G, Murat P, Papadopoulou C, Sarkies P, Prioleau M-N, et al. Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. Embo J 2014;33:2507–20. doi:10.15252/embj.201488398.

[27] Okada T, Sonoda E, Yoshimura M, Kawano Y, Saya H, Kohzaki M, et al. Multiple Roles of Vertebrate REV Genes in DNA Repair and Recombination. Mol Cell Biol 2005;25:6103–11. doi:10.1128/MCB.25.14.6103-6111.2005.

[28] Sharma S, Hicks JK, Chute CL, Brennan JR, Ahn J-Y, Glover TW, et al. REV1 and polymerase ζ facilitate homologous recombination repair. Nucleic Acids Res 2012;40:682–91. doi:10.1093/nar/gkr769.

[29] Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harb Perspect Biol 2013;5:a012708. doi:10.1101/cshperspect.a012708.

[30] Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. Transl Cancer Res 2013;2:107–29. doi:10.3978/j.issn.2218-676X.2013.04.01.

[31] Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, et al. The million mutation project: a new approach to genetics in Caenorhabditis elegans. Genome Research 2013;23:1749–62. doi:10.1101/gr.157651.113.

[32] Chiu H, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. Genetics 2013;195:1167–71. doi:10.1534/genetics.113.155879.

[33] Cho SW, Lee J, Carroll D, Kim J-S, Lee J. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics 2013;195:1177–80. doi:10.1534/genetics.113.155853.
[34] Katic I, Großhans H. Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. Genetics 2013;195:1173–6. doi:10.1534/genetics.113.155754.
[35] Lo T-W, Pickle CS, Lin S, Ralston EJ, Gurling M, Schartner CM, et al. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics 2013;195:331–48. doi:10.1534/genetics.113.155382.
[36] Tzur YB, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiácovo MP. Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. Genetics 2013;195:1181–5. doi:10.1534/genetics.113.156075.
[37] Waaijers S, Portegijs V, Kerver J, Lemmens BBLG, Tijsterman M, van den Heuvel S, et al. CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics 2013;195:1187–91. doi:10.1534/genetics.113.156299.
[38] Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nature Methods 2013;10:741–3. doi:doi:10.1038/nmeth.2532.
[39] Cuppen E, Gort E, Hazendonk E, Mudde J, van de Belt J, Nijman IJ, et al. Efficient target-selected mutagenesis in Caenorhabditis elegans: toward a knockout for every gene. Genome Research 2007;17:649–58. doi:10.1101/gr.6080607.
[40] Lawrence CW. Cellular functions of DNA polymerase zeta and Rev1 protein. Adv Protein Chem 2004;69:167–203. doi:10.1016/S0065-3233(04)69006-1.
[41] Pryor JM, Gakhar L, Washington MT. Structure and Functional Analysis of the BRCT Domain of Translesion Synthesis DNA Polymerase Rev1. Biochemistry 2013;52:254–63. doi:10.1021/bi301572z.
[42] Larimer FW, Perry JR, Hardigree AA. The REV1 gene of Saccharomyces cerevisiae: isolation, sequence, and functional analysis. J Bacteriol 1989;171:230–7.
[43] Gibbs PE, Wang XD, Li Z, McManus TP, McGregor WG, Lawrence CW, et al. The function of the human homolog of Saccharomyces cerevisiae REV1 is required for mutagenesis induced by UV light. Proc Natl Acad Sci U S A 2000;97:4186–91. doi:10.1073/pnas.97.8.4186.
[44] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.
[45] van Bostelen I, Tijsterman M. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light. DNA Repair 2017. doi:10.1371/journal.pbio.1001378.
[46] Alpi A, Pasierbek P, Gartner A, Loidl J. Genetic and cytological characterization of the recombination protein RAD-51 in Caenorhabditis elegans. Chromosoma 2003;112:6–16. doi:10.1007/s00412-003-0237-5.
[47] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.
[48] De Stasio E, Lephoto C, Azuma L, Holst C, Stanislaus D, Uttam J. Characterization of revertants of unc-93(e1500) in Caenorhabditis elegans induced by N-ethyl-N-nitrosourea. Genetics 1997;147:597–608.
[49] Greenwald IS, Horvitz HR. unc-93(e1500): A behavioral mutant of Caenorhabditis elegans that defines a gene with a wild-type null phenotype. Genetics 1980;96:147–64.
[50] Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol 1983;100:64–119. doi:10.1016/0012-1606(83)90201-4.
[51] Kruisselbrink E, Guryev V, Brouwer K, Pontier DB, Cuppen E, Tijsterman M. Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ-defective *C. elegans*. Curr Biol 2008;18:900–5. doi:10.1016/j.cub.2008.05.013.
[52] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.
[53] van Schendel R, Roerink SF, Portegijs V, van den Heuvel S, Tijsterman M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat Comms 2015;6:7394. doi:10.1038/ncomms8394.
[54] Lemontt JF. Mutants of yeast defective in mutation induced by ultraviolet light. Genetics 1971;68:21–33.
[55] Clark DR, Zacharias W, Panaitescu L, McGregor WG. Ribozyme-mediated REV1 inhibition reduces the frequency of UV-induced mutations in the human HPRT gene. Nucleic Acids Res 2003;31:4981–8. doi:10.1093/nar/gkg725.
[56] Jansen JG, Langerak P, Tsaalbi-Shtylik A, van den Berk P, Jacobs H, de Wind N. Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. J Exp Med 2006;203:319–23. doi:10.1084/jem.20052227.
[57] Lemmens BBLG, Tijsterman M. DNA double-strand break repair in Caenorhabditis elegans. Chromosoma 2011;120:1–21. doi:10.1007/s00412-010-0296-3.
[58] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.
[59] Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nature Methods 2013;10:1028–34. doi:10.1038/nmeth.2641.

Supplemental information

Supplemental figure 1.

Alignment of REV1 protein sequences and phylogenetic tree.

Sc = Saccharomyces cerevisiae
Ce = Caenorhabditis elegans
Dm = Drosophila melanogaster
Dr = Danio rerio
Xl = Xenopus leavis
Gg = Gallus gallus
Mm = Mus musculus
Hs = Homo sapiens

A) CLUSTAL multiple sequence alignment of REV1 protein sequences

Known domains in human REV1 are indicated:

BRCT	With mutated residues in this and other studies in BRCT mutants indicated as: G
Pol	Y-family polymerase domain
UMB	Ubiquitin binding motifs (UBM)
TLSinteract	C-terminal domain interacting with other TLS polymerases

```
Sc      -------------------------------------MGEHGGLVDLLDSD--------       14
Ce      MHSPDPFTDHPDSDDSFEKFVPTPNVNRRRSDAIIVSDEGSFDGITGYSQVTVIENCPNT       60
Dm      ------------------------------------------------------------       0
Dr      ------------------------------------------------------------       0
Xl      ------------------------------------------------------------       0
Gg      ------------------------------------------------------------       0
Mm      ------------------------------------------------------------       0
Hs      ------------------------------------------------------------       0


Sc      --LEYSINRETPDKN--------------------NCLSQ----QSVNDS---HLTA-KT       44
Ce      TVLEDENKRANPDDTITRMLESSDEEDETENNSDNDALADLLKSFRTEDQKRKEVDIEQA       120
Dm      ------------------------------------------------------------       0
Dr      ------------------------------------------------------------       0
Xl      ------------------------------------------------------------       0
Gg      ------------------------------------------------------------       0
Mm      ------------------------------------------------------------       0
Hs      ------------------------------------------------------------       0


Sc      GGLNARSFLSTL----SDDSLIEYVNQLSQ--------------TNKNNSNPTAGTLRFT       86
Ce      ANISIDIFGDLIEEVDSEHEAAEYPDREGGNSMDKQGTMLPVEVPIQNNIKLIPRKRRLE       180
Dm      ------------------------------------------------------------       0
Dr      ------------------------------------------------------------       0
Xl      ------------------------------------------------------------       0
Gg      ------------------------------------------------------------       0
Mm      ------------------------------------------------------------       0
Hs      ------------------------------------------------------------       0


Sc      TKNISCDELHADLGGGEDSPIARSVIEIQESDSNGDDVKK---------NTVYTREAYFH       137
Ce      PKSYDCEPSTSAAHG----P-----VGVVSSDSGEYWNQQKTITIHGQNFEINNFNDYMR       231
Dm      ---------------------------------------MT-----RDEDNGFSEWGGYFE       17
Dr      ------------------------------MSRDGWRKK-----ASENDGWGGQGGYMA       24
Xl      ------------------------------MRQGGWRNR-----ASEKDGWTAWGGYIA       24
Gg      ------------------------------MRRGGWRKR-----AGEGDGWGGWGGYMS       24
Mm      ------------------------------MRRGGWRKR-----T-ENDGWEKWGGYMA       23
Hs      ------------------------------MRRGGWRKR-----A-ENDGWETWGGYMA       23
                                                                  *:
```

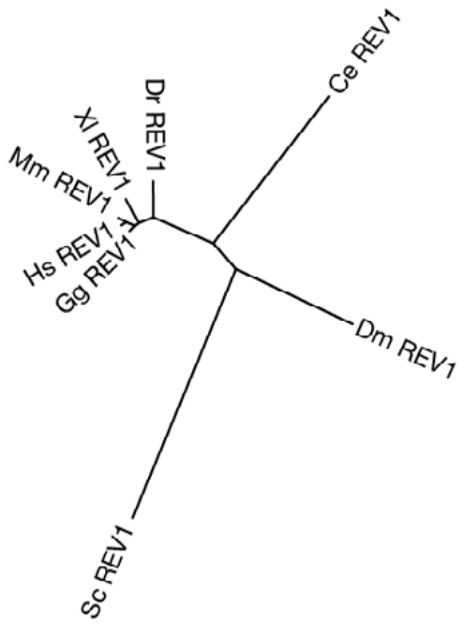
```
                                        BRCT>_____________________*_____
Sc      EKAHGQTLQDQILKDQYK-DQISSQSSKIFKNCVIYINGYTKPGRLQLHEMIVLHGGKFL      196
Ce      MK--ITKLNHQVNHGYAK---PLENVSRIMEGFSVFVNGYTDPPALVIRDLMISHGGEYH      286
Dm      AK--KSKLEEQFAAASDP-----FRKSDLFQGISIFVNGRTDPSADELKRIMMVHGGTFH      70
Dr      AK--VSKLEEQFQKDAPREQKKDGKSSCIFSGVAIYVNGYTDPSAAELRRLMMLHGGQFH      82
Xl      AK--VQKLDEQFRTDAPLQQQKDGTSSKIFNGVSIYVNGYTDPTADQLRHLMMLHGGQYH      82
Gg      AK--VKKLEDQFRSDSAIQHQRDGNSSSIFSGVAIYVNGFTDPSADELRRLMMLHGGQYH      82
Mm      AK--VQKLEEQFRTDAAN--QKDGTASAIFSGVAIYVNGYTDPSAEELRNLMMLHGGQYH      79
Hs      AK--VQKLEEQFRSDAAM--QKDGTSSTIFSGVAIYVNGYTDPSAEELRKLMMLHGGQYH      79
         *    .*:.*.              * ::.   :::** *.*    :: ::: *** :

                                ____________________________________
Sc      HYLSSKKTVTHIVASNLPLKKRIEFAN-YKVVSPDWIVDSVKEARLLPWQNYSLTSKLDE      255
Ce      CYYQHGI-TSYTIASSIATAKINRIRENEIFIKADWITESIAAGKPLDYRDFLIYEKGSV      345
Dm      HY-ERSH-TTYIIASVLPDVKVRNMNL-SKFISAKWVVDCLEKKKIVDYKPYLLYTNQKT      127
Dr      LYYSRSK-TTHIIATNLPNFKIQELKG-EKVVRPEWIIDSIKAGRQLSYIQYQLYAKQ--      138
Xl      VYYSRSK-TTHIIATNLPNNKVNELKD-EKVVRPEWITESIKAGRLLSYVPFQLYTKKSS      140
Gg      VYYSRSK-TTHIIATNLPNAKIKELKG-EKVVRPEWIVESIKAGRLLSHIPYQLYTKQSS      140
Mm      VYYSRSK-TTHIIATNLPNAKIKELKG-EKVIRPEWIVESIKAGRLLSSAPYQLYTKPSA      137
Hs      VYYSRSK-TTHIIATNLPNAKIKELKG-EKVIRPEWIVESIKAGRLLSYIPYQLYTKQSS      137
         * .    .:: :*: :   *  .:     .:  .*: :.:   : :    : :  :

Sc      QQKKLDN---CKTVNSIPLPSE----------------------------------------      274
Ce      EKGQMQ------------------------------------------------------      351
Dm      SQPMLIFGKP---KDNGANE--SKSDVEPPK----------------------DKAEVE      159
Dr      --KGLNFTRVSGIEGQDPSSN-HGPDIKPQPSHTLSKDVLK--STQANHMANYTEKTDHQ      193
Xl      VQKGLNFTSICKRDDSLPGPSNISRDFSQRENSTIKDFSARPISVLLNGVERL-EEDEHK      199
Gg      VQKGLSFNSICKPEDAMPGPSNIAKDLNR-VNH-IKQCEMES-EITPNGISSWNEEEEED      197
Mm      AQKSLNFNPVCKPEDPGPGPSNRAKQLNNRVNHIIKKIETES-EVKANGLSSWNEDGVN-      195
Hs      VQKGLSFNPVCRPEDPLPGPSNIAKQLNNRVNHIVKKIETEN-EVKVNGMNSWNEEDEN-      195
            :

Sc      ------------------------------------------------------------      274
Ce      ------------------------------------------------------------      351
Dm      VD--------------------------------------STKDETQM-------ELGG      173
Dr      LDIRLRNGLHSIISEVDAEKSRVNGIHDDDDDDDIACSILPRGSQDTLLTNGHVHPV-NG      252
Xl      NG-----DFEMLDLEQIFSDHKPNGL--------------QKHIGNTDVCSKHIPSS-NG      239
Gg      SD-----GLGFTKLDQILPERKQNGI--------------QSHKDSTAIFNGHTHNTCTS      238
Mm      -D-----DFSFEDLEHTFPGRKQNGV--------------MHPRDTAVIFNGHTHSS-NG      234
Hs      -N-----DFSFVDLEQTSPGRKQNGI--------------PHPRGSTAIFNGHTPSS-NG      234

Sc      ------------------------------------------------------------      274
Ce      ----------------------------------------------QFLT-----------      355
Dm      ILKNLQQA------------------------------------------------------      181
Dr      ALKPQDFADHHPHVVDKASNQSHLPQDKYREGEEPCCSYTESKANSHLRSLISAASPAKQ      312
Xl      ALKSLDSTVHFNSSFESGL--DLLQ--QQEKPDQCCANITDCTV-EYLQ-----------      283
Gg      ALKTQDCLVPSSNSVASRFSPGPVQ--EEGKPEKGIVDFRDCTM-QQLQ-----------      284
Mm      ALKTQDCLVPVGNSVASRLSLDSTQ--EEKRAEKSNADFRDCTV-QHLQ-----------      280
Hs      ALKTQDCLVPMVNSVASRLSPAFSQ--EEDKAEKSSTDFRDCTL-QQLQ-----------      280

Sc      -------------------------TSLHK--GSKCVGSALLPVEQQ-------SPVNL      299
Ce      ------------ISTRNT--------------------GGN------------------N      365
Dm      ---------VATSPEKEAS----AS--------ESKIT-----------------NLST-      202
Dr      PGPHNMQTDLHQKPTSNSGTLGAPQKASLHDQTTVRLNGSYHVTS----------NSSTL      362
Xl      --PNSRQTDVSCNPHRTVPL--P-SSSCLHT--NSKINGANHSLL-GPSSTNALS----P      331
Gg      --QSNKNTDFSWNPHRTMSN--SSSSSSLHS--NTKINGAHHSTVQGPSSTKSTS-VPTP      337
Mm      --HSTRSADALRSPHRTNS----LSP-SLHS--NTKINGAHHSTVQGPSSTKSTS-VLTL      330
Hs      --QSTRNTDALRNPHRTNS----FSLSPLHS--NTKINGAHHSTVQGPSSTKSTSSVSTF      332
```

```
Sc      NNLEAKRIVACDDPDFLTSYFAHSRLHHLSAWKANLKDKFLNENIHK------------      346
Ce      ETTSSNQFSDARNPNFIRDYYARSRLHLISTLAQDMKDFVANLKLEGKLT-EKCFEEKEL      424
Dm      TSNNSTTARTAADPNFLSEFYKNSRLHHIATLGAGFKQYVCRLRQKH---GTQGFPKRET      259
Dr      ANSSNQSGKSSAEAGIISEFFSHSRLHHISTWRNEFSEYVNSLQSRRRAAGGAVFSGREK      422
Xl      AKLHSPQLPKTADPNFISDFYSHSRLHHISTWKCEFTEFVRNLQTQS----NRGFPGRER      387
Gg      SKAASLSVSKPSDCSFISDFYSRSRLHHISTWKCELTEFVNSLQRKN----SGVFPGREK      393
Mm      SKVAPSVPSKPSDCNFISDFYSRSRLHHISTWKCELTEFVNTLQRQS----SGIFPGREK      386
Hs      SKAAPSVPSKPSDCNFISNFYSHSRLHHISMWKCELTEFVNTLQRQS----NGIFPGREK      388
         .          :  :: .:: .**** ::     :.: .   . .

                                          Pol>_____________________
Sc      YTK-------------------ITDKDTYIIFHIDFDCFFATVAYLCRSSSFSACDFKRD      387
Ce      ID-MK--------------SLSNEISRESTVFHVDLDCFFVSVAVRNRID------LKHK      463
Dm      LKSLANSH---------------HNCLERYVMHIDMDCFFVSVGLRTRPE------LRGL      298
Dr      LKKLKANCNSVSHFDPGSLMA-APQVRQSCVLHVDMDCFFVSVGIRHRPD------LIGK      475
Xl      LKKLKPGN----------LPPSAFPKCQNCIIHVDMDCFFVSVAIRNHAD------LKGK      431
Gg      LKKWKAGRSAL-KTDTGNVSVASSAKPQSCIMHVDMDCFFVSVAIRNRPD------LKGK      446
Mm      LKKVKTGRSSLVVTDTGTMSVLSSPRHQSCVMHVDMDCFFVSVGIRNRPD------LKGK      440
Hs      LKKMKTGRSALVVTDTGDMSVLNSPRHQSCIMHVDMDCFFVSVGIRNRPD------LKGK      442
                                    ::*:*:****.:*.   : .      :

        _____________________________________________________________
Sc      PIVVCH------------------------------------------------------      393
Ce      EVAITHSKGTIS------------------------------------------------      475
Dm      PIAVTHSKGGNAATDVPVHPQADRKAELELFAQRFEHHFHD---GD---------KAEKV      346
Dr      PVAVTSNRGPGRV---AQRPGANPQLEFQYYQNKQKHYRKEKTDGDL-EMTPSPQDGEVP      531
Xl      PVAVTSNRGAGTT---ITREGSNPQLEFQYYQNRILKGKAGQPMS--------ATSDSAQ      480
Gg      PVAVTSNRGAGKA---PLRPGANPQLEWQYYQNKLLNGKAEIRIPDKLDSWVWEHSDSAH      503
Mm      PVAVTSNRGTGTA---PLRPGANPQLEWQYYQNRALRGKAA----DIPDSSVWENQDSTQ      493
Hs      PVAVTSNRGTGRA---PLRPGANPQLEWQYYQNKILKGKAA----DIPDSSLWENPDSAQ      495
         :.:

        _____________________________________________________________
Sc      ---------GTKNSDIASCNYVARSYGIKNGMWVSQAEKMLPNGIKLISLPYTFEQFQLK      444
Ce      ----------NSMSEVASCSYAARDCGVKNGMLVRDALQKCPQ---LTLLPYQFEDYVQV      522
Dm      RSGFDK---KMSLSEIASCSYEAREKGIRNGMFVGQALKLCPE---LKTIPYDFEGYKEV      400
Dr      SNGVHEDLAALSMAEIASCSYEARQAGVRNGMFFGRAKQLCPD---LQSVPYDFHAYKEV      588
Xl      QNGLDQDISHLSMAEIASCSYEARQAGVKNGMFFGRAKQLCPE---LQAVPYDFDAYKEV      537
Gg      MNGVDCDLTVLSMAEIASCSYEARQAGIKNGMFFGQAKKLCPN---LQAVSYDFNAYKEV      560
Mm      TNGIDS---VLSKAEIASCSYEARQVGIKNGMFFGYAKQLCPN---LQAVPYDFHACREV      547
Hs      ANGIDS---VLSRAEIASCSYEARQLGIKNGMFFGHAKQLCPN---LQAVPYDFHAYKEV      549
                  . :::***.* **. *::*** .  * :  *:   *  : * *.

        _____________________________________________________________
Sc      SEAFYSTLKRLNIFNLILPISIDEAVCVRIIPDNIHNTNT-LNARLCEEIRQEIFQGTNG      503
Ce      SRKIYEILASYTL--EVRAVSCDEMYIN--MSSFCEKYEINDPTILAEHIRKVIR-EKTQ      577
Dm      AFTLYDTVAQYTL--NIEAVSCDEMFVE--LTDLAHELNV-DVMAFVSHLRQEVY-SKTG      454
Dr      ALAMYEILASYTH--NIEALSCDEALVD--ATALLVELGV-SPDELARSIREDIK-EKTG      642
Xl      AMNMYKILASYTH--DIEAVSCDEALAD--ITGILTETRL-TPDEASNAIRTEIK-EKTG      591
Gg      AQTVYEILASYTH--NIEAVSCDEALVD--ITEILTETRL-TPDELANAIRDEIK-AQTK      614
Mm      AQAMYETLASYTH--SIEAVSCDEALID--VTDILAETKL-SPEEFAAALRIEIK-DKTK      601
Hs      AQTLYETLASYTH--NIEAVSCDEALVD--ITEILAETKL-TPDEFANAVRMEIK-DQTK      603
        :  .*. :   .    :  :* **            :            :*  :    .

        _____________________________________________________
Sc      CTVSIGCSDSLVLARLALKMAKPNGYNITFKSNLSEEFWSSFKLDDLPGVGHSTLSRLES      563
Ce      CPASVGIGSTSLLARLATRHAKPDGVFWVNAH-KKNEFISEEKVKDLPGFGYEMMNRLTS      636
Dm      CPCSAGVAGNKLLARMATKEAKPNGQFLLDSSNDILAYMAPMSLDLLPGVGSSISHKLKQ      514
Dr      CCASVGMSSNILLARMATRKAKPKGQYLLRSE-EVDDFIRDQPVSSLPGVGRSMSSKLTS      701
Xl      CAASIGIGSNILLARMATRRAKPDGQYHLKPE-EVDDFIRGQLVNNLPGVGRTMDCKLSS      650
Gg      CTASVGMGSNILLARMATRKAKPDGQYHLKPE-EVDDFIRGQLVTNLPGVGRSMESKLAS      673
Mm      CAASVGIGSNILLARMATKKAKPDGQYHLQPD-EVDDFIRGQLVTNLPGVGRSMESKLAS      660
Hs      CAASVGIGSNILLARMATRKAKPDGQYHLKPE-EVDDFIRGQLVTNLPGVGHSMESKLAS      662
        *  * *  . .***.* . ***.*           .    .  ***.*    .* .
```

```
Sc      TFDSPHSLNDLRKRYTLDALKASVGSKLGMKIHLALQGQDDEESLKILYDPKEVLQRKSL      623
Ce      FFGDITKCRELQL-KTERELVPVFGPKLASKILRQCRGIEEDPDD-FW----ATHVRKSV      690
Dm      A--GLNNCGDVQN-TTLEKMEKVLGKKLGQNLFQNCRGIDDRPLA-------YEQIRKTV      564
Dr      L--GVSTCGDLQQ-LSLSQLQREFGPRTGQTLFRFCRGLDDRPVR-------SEKERKSV      751
Xl      L--GVKTCGELQN-ITMAKLQKEFGPKTGQMLYRFCRGLDDRPIR-------KEKERKSV      700
Gg      L--GIRTCGDLQC-ASMSKLQKEFGPKTGQMLYRFCRGLDDRPVR-------TEKERKSV      723
Mm      L--GIKTCGDLQC-LTMAKLQKEFGPKTGQMLYRFCRGLDDRPVR-------TEKERKSV      710
Hs      L--GIKTCGDLQY-MTMAKLQKEFGPKTGQMLYRFCRGLDDRPVR-------TEKERKSV      712
              .  :::   :   :   .* : .  :    :* ::              **::

Sc      SIDINWGIRFKNITQVDLFIERGCQYLLEKLNEINKTTSQITLKLMRRCKDAPIEPPKYM      683
Ce      SCDINYGIRFTKRGEVIQLMTAIGAELERKLIDSKLTAGSITLKLMVRSANAPIQTSKFM      750
Dm      SAEMNFGIRFTNSVECEQFLCQLSEEVTKRLVEIRRKARSINLKIMVRAAEAPVETSKYM      624
Dr      SAEMNYNIRFTQVEEAESFLNNLSMEVQKRLEGAGLRGRRVTLKVMMRKPGAPVEPAKYG      811
Xl      SAEINYGIRFTQTSEAEVFLMNLSEEIQRRLEVVGMKGKKLTLKIMVRKAGAPIESAKYG      760
Gg      SAEINYGIRFTQPKEAEAFLLSLSEEIQRRLEAAGMKGKRLTLKIMVRKAGAPVEPAKYG      783
Mm      SAEINYGIRFTQPKEAEAFLLSLSEEIQRRLEAAGMKGKRLTLKIMVRKPGAPIETAKFG      770
Hs      SAEINYGIRFTQPKEAEAFLLSLSEEIQRRLEATGMKGKRLTLKIMVRKPGAPVETAKFG      772
        * ::*: ***.:  :   ::      : .:*         :.**:* *   **::  *:

Sc      GMGRCDSFSRSSRLGIPTNEFGIIATEMKSLYRTLGCPPMELRGLALQFNKLVDVGPDNN      743
Ce      GHGICDTFTKTCNLNVPTTRGESLTSEAMKLYAKVSPKVEDLRGVGVTCGKLKSKLKKDA      810
Dm      GHGVCDIINKSSLIKYATDDVNVITTVVLDLMKDADIPPDELRGLGIHLTRLEDANEVRK      684
Dr      GHGICDNFARSVLLAQPTDSGRVIASEAIKLFHAMKLNVKDMRGVGLQVQQLDGSHAD--      869
Xl      GHGICDNIARTVTLHQATASAKVIGKEAVDMFHTMKLNISDMRGVGLQVHQLIPVGGLSI      820
Gg      GHGICDNIARTVTLDHATDSAKVIGKETLNMFHTMKLNISDMRGVGIQVQQLVPISKTT-      842
Mm      GHGICDNIARTVTLDQATDSAKIIGKATLNMFHTMKLNISDMRGVGIQVNQLVPANSNLS      830
Hs      GHGICDNIARTVTLDQATDNAKIIGKAMLNMFHTMKLNISDMRGVGIHVNQLVPTNLNPS      832
        * * ** : ::  :   *     : .   .:         ::**:.:   :*

Sc      QLKLRLPFKTIVTNRAFEALPEDVKN------------------------DINNEFEKRN      779
Ce      --------------------ATAVQEMFGKTSRVGNMARTDEQLNIIPRNEDELDK----      846
Dm      E--------------------NNIKEMFGKMSEMR---KDKPIPQG-AVGDKSIGDDKVN      720
Dr      ------------PSGQGPSRGRSIRDLLLAKQSAHSPSKESPSQDGLISASSSRAFSSNQ      917
Xl      --FGSTL----VKSGHLPGGSRSLMDMFQGQKMKS-FSEDNDGKRDTITAVDEIGFASGT      873
Gg      --SAQSA----VQSGRLPGGSHSVIDLLHVQKAKK-CSEEEHKEVFVAAMDLEISSDSRT      895
Mm      TCSSRPS----AQSSLFSGRPHSVRDLFQLQKAKK-PTEEEHKEVFLAAVDLEVSSTSRA      885
Hs      TCPSRPS----VQSSHFPSGSYSVRDVFQVQKAKK-STEEEHKEVFRAAVDLEISSASRT      887
                             : :                          .

Sc      Y-------KRKESGLTSNS---------------LSSKKKGFAIS--------------R      803
Ce      ---EPVIIPAVTEDLEVQS-----IPRQI-----NFRVANDIDIPEIVKSTLLNGRSDNN      893
Dm      KPLVFEN-KPKPRE-PRNVLSMLTAAAVSRKSVTEDRSQRGTSKPIT-------------      765
Dr      KRVTPALSPPPSTSSSSSPPPSFTIIDPMP--GTSKGDYHHPHTPNHVRA-------CLN      968
Xl      ---CYSI-PPYNKS---SQ-----APSTSK--AASTGQKNGFSSPNNVKS-------GLR      912
Gg      ---CTVL-PSRGTHLTAGL-----NSNVSK--TDSTVKLNGLHSPISVKS-------RLN      937
Mm      ---CGLL-SPLSAHLAASV-----SPDTNS--GECSRKWNGLHSPVSGQS-------RLN      927
Hs      ---CTFL-PPFPAHLPT-------SPDTNK--AESSGKWNGLHTPVSVQS-------RLN      927
                                                    .

              UBM1>
Sc      L-EVNDLPSTMEEQFMNELPTQIRAEVRHDLRIQKKIQQTKLGNLQEKIKRREE-SLQNE      861
Ce      LELEDLEDDGLVENRLA----------ELECIIQ-NEPTKESVT--NLQEM-LSTLLNYG      939
Dm      --RPLSLVPKLDEDVLAQLPEDIRLEVIANR-------------------------------      794
Dr      ISIEVPSPSQVDPSVLKALPIDIRRQVEETWRYREEQPSTSSHY--STPPR--PPSTP--      1022
Xl      FAIEVPSPSQIDPSVLEALPSDIREQVEQTFGFQQKTLLTESKN--DPAQKGSSELKQ--      968
Gg      LSIEVPSASQLDKSVLEALPPDLREQVEQIYTIQQGETYGDSKR--EPINGCNTALLS--      993
Mm      LSIEVPSPSQIDQSVLEALPLDLREQIEQVCAAQQGEPRG-KKK--EPVNGCSSGVLP--      982
Hs      LSIEVPSPSQLDQSVLEALPPDLREQVEQVCAVQQAESHGDKKK--EPVNGCNTGILP--      983
                  :  . :
```

```
                                                  UBM2>
Sc      ------------------------------------------------------------      861
Ce      KLSAFETVFRKFEEISMDSSRCDTAWFSVFHLMVPFIEEESKDLLEFPIATTRNLGAMIA      999
Dm      ----------------------------------------------------EEHLCIAEY      803
Dr      --PPGPSLVLQLPNQ--PGQPCTT---GIILELPDFSQVDPDVFAALPRELQEELCSA-Y      1074
Xl      --QPVATVLLQIPNL---SDQ-GEEQGINVIALPAFSQVDPEVFAALPADLQEELRAA-Y      1021
Gg      --QPVGTVLLQVPEL---QEP-NANMGINVIALPAFSQVDPEVFAALPAELQAELKDA-Y      1046
Mm      --HPVGTVLLQIPEP---QEPCNSDSKISVIALPAFSQVDPDVFAALPAELQKELKAA-Y      1036
Hs      --QPVGTVLLQIPEP---QES-NSDAGINLIALPAFSQVDPEVFAALPAELQRELKAA-Y      1036

        _____
Sc      ---------------KNHFMGQNSI-----------------------------------      871
Ce      S-----NTDNTSISSPEPFVPKEVFVI------DSTKHT---------------------      1027
Dm      DGY-RSPQ----------------YPTLRSPPLLN---PYVTTNV-------SPLKATDL      836
Dr      RNKGNAQA-----QASTVVEQKNSFPQLKQPAVGKLKRRYKKRNTSPAKNGSSPLKKMFP      1129
Xl      GQRNKQTNNI---NINPTFVSKNPLLQLRKPLEKSKKRSRKKNKGSPTKNIHSPLKPKLF      1078
Gg      DQRQKQPE----QQPANAFVSKNPCLQLKHATTKNKKKIRKKNPVSPVKKIQSPLKNKLL      1102
Mm      DQRQRQGEDTTHQQPTSTSVPKNPLLQLKPPAMKDKR-NKRKNLIGSPRK--SPLKNKLL      1093
Hs      DQRQRQGENSTHQQSASASVPKNPLLHLKAAVKEKKR-NKKKKTIGSPKRIQSPLNNKLL      1095

                                                        TLSinteract>_____
Sc      ------------------------------------------------------------      871
Ce      ------------------------------------------------------------      1027
Dm      KPSTSRA---AVARLQKRKERKEQEHYIRSDQI-----VADY------IDDLPDFVNPHI      882
Dr      GNSPAKTS------PSK-----TLPLPLREQDLKCPSSL---------STDVPET---LP      1166
Xl      S-SPLKNGF-SAGSPQK-----LKDGPLKQEAKPSGQSQVELGPSTSNDA---KSLAQSH      1128
Gg      G-SPAKNMPAASGSPQK-----LIDGFLKQEGAAAQLE---AVPSTSDASDPSALQTEQC      1153
Mm      S-SPAKTLPGAYGSPQK-----LMDGFLQHEGMASERPLEEVSASTPGAQDLSSLLPGQS      1147
Hs      N-SPAKTLPGACGSPQK-----LIDGFLKHEGPPAEKPLEELSASTSGVPGLSSLQSDPA      1149

        ____________________________________________________________
Sc      ----FQPIKFQNLTRFKKICQLVKQWVAETLGDGGPHEKDVKLFVKYLIKLCDSNRVHLV      927
Ce      ------------------------------------------------------------      1027
Dm      LKLISHPVEMPELLMGDNYKDLLNDWVSRE---EVPKPNDVDLILKQVSRMIKNDQLDHV      939
Dr      KPIPRPAPTLAGAHEFSEIRTLLREWVTTI---SEPMEEDILQVVKYCTELVEDKDLEKL      1223
Xl      SSFKPKPPNLAGAIEFSDVKTLLREWITTI---SDPMEEDILQVVRYCTDLIDEKDLEKL      1185
Gg      GSFRPQAPNLAGAVEFNDVKTLLKEWITTI---SDPMEEDILQVVKYCTDLIEEKDLEKL      1210
Mm      SCFRPAAPNLAGAVEFSDVKTLLKEWITTI---SDPMEEDILQVVRYCTDLIEEKDLEKL      1204
Hs      GCVRPPAPNLAGAVEFNDVKTLLREWITTI---SDPMEEDILQVVKYCTDLIEEKDLEKL      1206

        ____________________________________________________
Sc      LHLSNLISRELNLCAFLNQDHSGFQTWERIL---LNDIIPLLNRNKHTYQTVRKLDMDFE      984
Ce      ------------------------------------------------------------      1027
Dm      CDVMKYW------CRIINMKRSSSCCWHVAYKHIEESIQNQM---LTIEGYSLLFIEYIR      990
Dr      YLVIKYM------KRL--MQQSSESVWSMAFDFVLDNVQVVV---QQTYGSTLKIT----      1268
Xl      DLVIKYM------KRL--MQQSVESVWNMAFDFILDNIQVVL---QQTYGSTLKVV----      1230
Gg      DLVVKYM------KRL--MQSSVESVWNMAFDFILDNVQVVL---QQTYGSTLKVI----      1255
Mm      DLVIKYM------KRL--MQQSVESVWNMAFDFILDNVQVVL---QQTYGSTLKVT----      1249
Hs      DLVIKYM------KRL--MQQSVESVWNMAFDFILDNVQVVL---QQTYGSTLKVT----      1251

Sc      V----  985
Ce      -----  1027
Dm      CIKCS  995
Dr      -----  1268
Xl      -----  1230
Gg      -----  1255
Mm      -----  1249
Hs      -----  1251
```

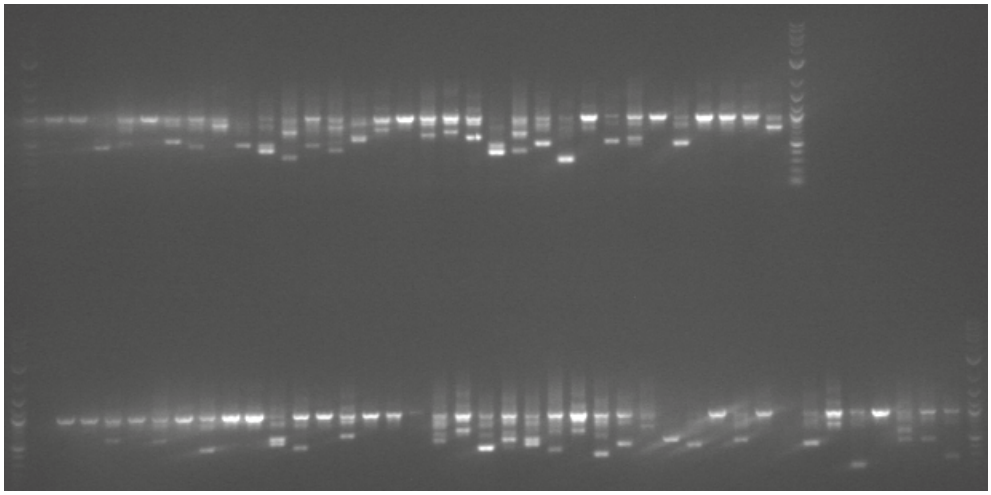
B) Phylogenetic tree of REV1 genes based on the aligned sequences above



Supplemental figure 2.

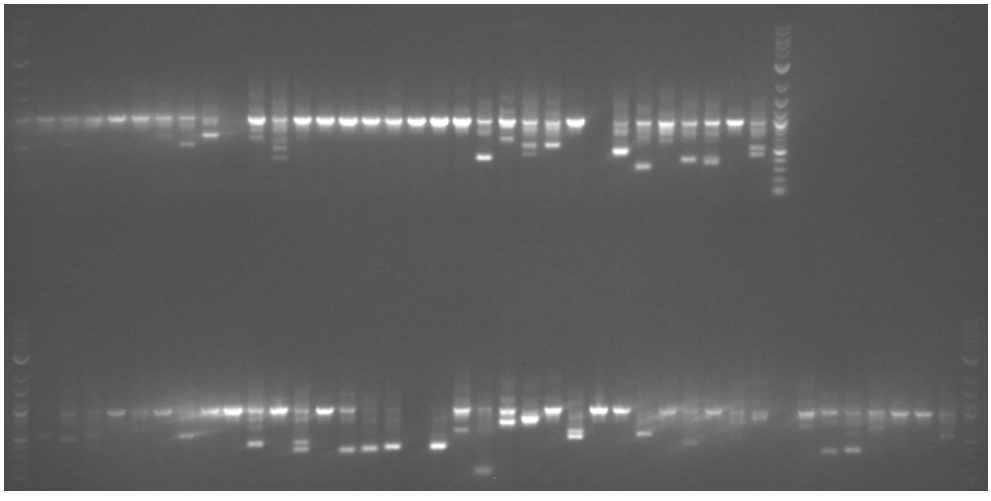
PCR assay on genomic G4 site qua1466. Lyses were made by picking 5 adult worms in 15 ul lyses buffer. Of the lyses 1 ul was used as a template for the PCR. In the dog-1 background we did not observe additiona genomic instability at qua1466 after loss of REV-1.

dog-1(gk10)



69 succesful PCRs with 50 samples positive for one or more deletions

dog-1(gk10); rev-1(splice)



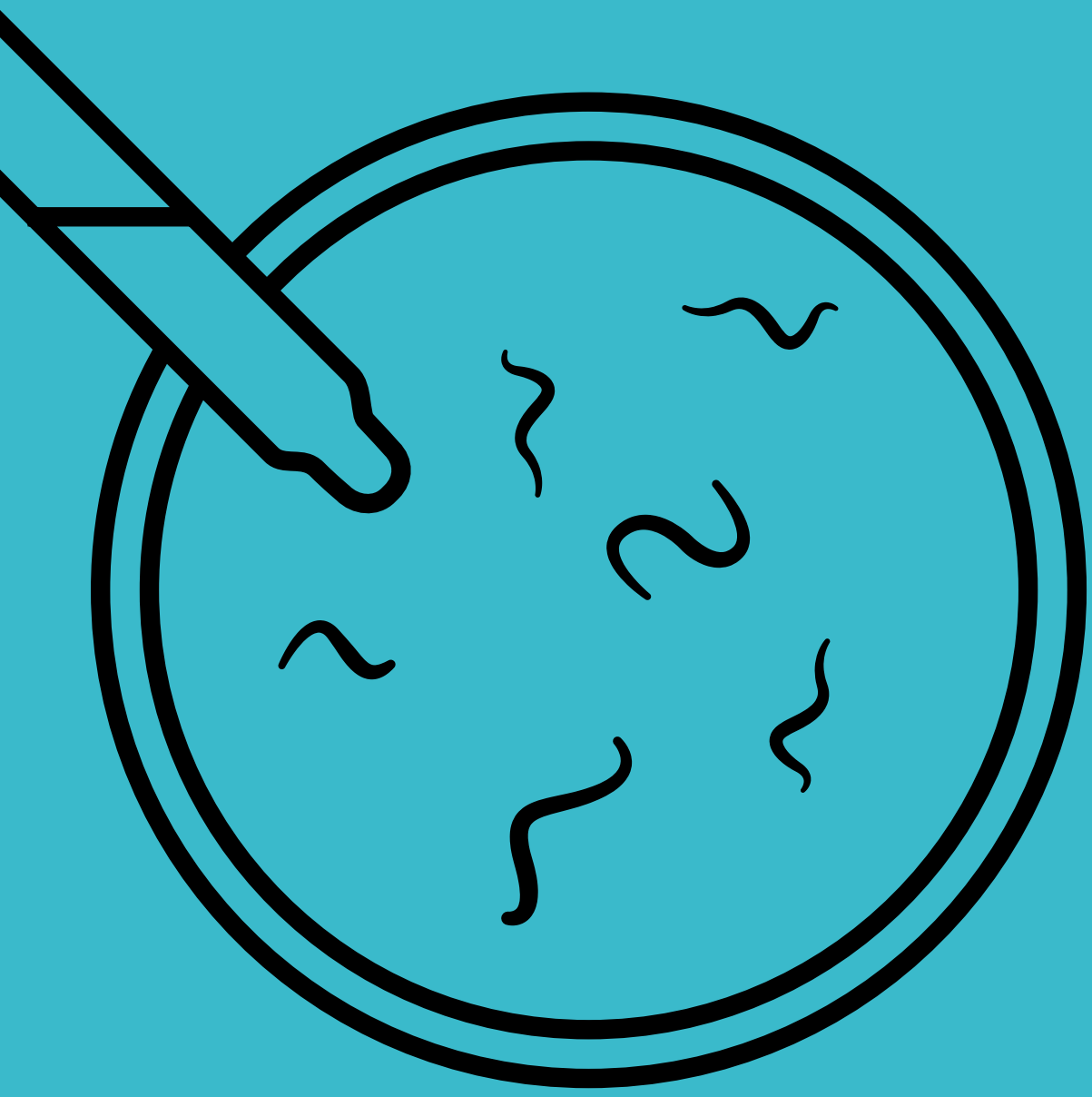
67 succesful PCRs with 36 samples positive for one or more deletions

Supplemental table 1

Genotype	Chromosome	Size of deletion	Start	End	5' flank	Deleted sequence	3' flank	Inserted sequence
N2	CHROMOSOME_I	0	7667699	7667699	CACGACAAAAGGTCA		CATACACAAGAAAAA	caaaaggt
N2	CHROMOSOME_II	2	11519275	11519277	ATATAAAATTTTTGA	tg	AAATTTTTTTGTTTT	
N2	CHROMOSOME_V	2	12400560	12400562	AATTATCTCAAAATT	tc	AAAACGTATAATGAA	caaagcg
N2	CHROMOSOME_II	2	13447128	13447130	AATCCTACTTTTTCG	ga	GAAATATCCTTTTTA	
N2	CHROMOSOME_V	6	18324061	18324067	TTGAAATAGATTATT	tcaggc	TCAACGTCGTTGAAA	
N2	CHROMOSOME_II	7	10938399	10938406	GGGCAAAGGGTAATT	atctgga	ATAGTTACCTCATTC	
N2	CHROMOSOME_X	11	16256717	16256728	TTATGCCAATTATTG	tacatatatct	TACATGGATCAATTT	
N2	CHROMOSOME_I	18	4336436	4336454	GTTTTTTGAAGTTTC	atatttcaaatttcagaa	ATATCGTAATCGTGG	
N2	CHROMOSOME_III	128	1407853	1407981	AATTTTCGATCACAG	tagatggttgagcaattttttctcagtgactgcatgcctgaaacagttagcaaaacgtgtcaggtaaatgctctgataactgtcgaaaatatttatcagtagagagctcagttatgagcccataatg	AGTTTGTTAGTTGAA	agtcac
rev-1	CHROMOSOME_X	1	4211498	4211499	AAGAAAAGAAAAGTG	a	TTATTTCAAAAAAAA	
rev-1	CHROMOSOME_X	3	6824150	6824153	AGAATATCGAGGAGT	taa	TAAGTAGTGGTGTAA	
rev-1	CHROMOSOME_IV	7	14354880	14354887	ATTTTCAGAATCAGT	ttaaaac	TATGAAACAAATGTA	
rev-1	CHROMOSOME_I	18	3281427	3281445	AGGCTTAGGTTGGGT	ttaggcttaggctgaggc	TTAGGCTTAGTTACA	
rev-1	CHROMOSOME_V	52	3945436	3945488	AAAACTCCCACCGGG	tcccatagctttcccataatcggaaatattcctcaagtggcttattatgtc	TGGAAGACGGAGGGA	
rev-1	CHROMOSOME_V	58	1066424	1066482	CGAGAAGCTTTCGTT	tttcctcgaatgtcattctttttgtcggaactttgacctccttctctttttctgaaaga	AAGTTTTTAACATTT	
rev-1	CHROMOSOME_III	80	1418981	1419061	TTCGCACAAAAATCA	ctggttttcactattttttacgctatttttttggttttttcttaattttcaggcacaaatatcgaatttaagaaggattag	CTCATAAAAATCAAG	

Genotype	Chromosome	Size of deletion	Start	End	5' flank	Deleted sequence	3' flank	Inserted sequence
rev-1	CHROMOSOME_IV	91	11887607	11887698	TTCAACTTTTACAAG	atctgggatgttcaaaagaccgttatgctattttccaaattatatttttacaagaaagcatcaaatattataaaattatgtttgtttctac	TGTATTATCTTTGTT	
rev-1	CHROMOSOME_V	102	14695448	14695550	ACAGCGAATGTCCCA	tccgaaacaatcagcttatgtagatgcgcacgcgtgcactcgaatagtcaacattaacaggggttaggatgaaaagaaaaagaagaaagagaagagcaactg	GTAGGAGATGAGAAG	
rev-1	CHROMOSOME_X	119	9279675	9279794	ATCATTATTCAAAAA	ctatgactaacatcccactatctatttttcagattgtggtcacaatctatgtaggagctgcatcaacaaactcactggcaacggaatcgtcaagtgtccatttgaccgtttggatacgc	GCGTCAGAGTGACTG	
rev-1	CHROMOSOME_V	133	625968	626101	GTAATTTGGATCCGT	acacaaatttcgtcgcattgcacaatcgctaactcctgtcgcattgtttatttttgtcatggcatggatcagaggtagtggatgttgatggggaaagtgtggtgccattgcccgatatgactggaagcaaaaat	TTTTTAGTGTGGTAG	g
rev-1	CHROMOSOME_III	155	10687677	10687832	GGGTGTATTGCAGTG	ctgtgcggctctcggcgagagccgaaaaaaattttgactcggctctcggcgcgagccgagagtcgacttttttcaagagccgcacagcactggtgtattgccttaaagaaggagtatcgtcaatggggaaattgttttaaaatgtagtatttgta	CTCAAACTTCCAATT	
rev-1	CHROMOSOME_V	174	16485754	16485928	AAGGTACTGTAGTGG	gtccgcaagggatactgtaaaattactgtaaggttcctgtagctcggggaaaatttgaattttcagcttttgaagaggttttttgttatttttgttgtcagtttggaatcttagcctgactcgagtgcactttttccaaaaaaaaaactcgctgtcaataacaagtagttttttaa	TCAACAATGTGAATT	tcaacaatg
rev-1	CHROMOSOME_I	194	10249974	10250168	AACTGGCAGTTCTTT	ttaagcttgtctccaccattatcaaggtttttctataagttgagcggtgactcagatgaaattagatattcatagtcgctattttctcacgtctaactatactgataagcccattttgcaaataagatacgcagatccctaaaaatcacatgctcacacctcttgaatgtgtccccatcatgagcccctgtcag	TTGCACCTGCCTCTT	ctttt

Genotype	Chromosome	Size of deletion	Start	End	5' flank	Deleted sequence	3' flank	Inserted sequence
rev-1	CHROMOSOME_I	224	14832621	14832845	TCCACCTTATCCTTT	ttcaattcaactag-caattagtaatatctt-gaacaacagcagct-ggcggcttgctgat-tcctctttttccttgat-ttctctttaaaaaacga-cacctagcaattaaca-gagaggagaacagg-caatatccaatctggct-tcttaagaaataagaaag-catcagttccaattatg-cgctgtgcttagtgaact-ggaaaaaacagcgg-catgcgaaacctacccag	GCGAGCCCGAA-CAAG	
rev-1	CHROMOSOME_V	266	12384944	12385210	TTTTTAGAACGAACT	cttgactatggttactct-gggaattgggcatgat-gtaggtgcagaggt-taaactgaaagagg-taattgatgaatgccacg-cataccataaaaacagt-ttctaactttcagctttatc-caaaaattgatttttttcg-gagcagatccag-cagcagtgatcaataaa-gatctttatgagaata-cattaggaggaaaatat-ttccaatatgctgtcag-cggagaaagtggaatg-caacgctcaagtgtatat-gaaagtaagccaaatggt	TATCCATAAAAT-TAA	tccatcaaa
rev-1	CHROMOSOME_V	306	15582024	15582330	CTATACGCCACCCTG	agatttttgttgaaaat-cagtgatattgccaaat-tatcaacctttttcag-gaccaacttttttaattgt-tctccaattttcgag-tacccgtgtatatttcaac-tattctcataagat-tttcacgtcacagaat-ttttataatttttttggtcct-gtaagccaactttg-ggagccgaccact-gatggcgcccttatct-cacgttccagccctc-caaaaattctgaaat-ttttttcagtgctaaa-ggagaatgcttacat-ttcatactataactgc-caaagttgtaggttctgt-ttaaataattctgt	AATTTGTTTTGT-GTT	
rev-1	CHROMOSOME_V	312	1010572	1010884	TAAAACACTTTTTTC	aatttcaatttccaatat-ttcagacttgccccac-caaagtacactgtgt-tcaaaagattaacat-atctaatttgcgaagaag-cgtacgaacccgagct-tccggatgtgtcactgt-gtcataaactgagaattg-gtacagacgcgagagtt-ggagtttcgattgctat-tcgcacaaatctga-gagccggaatcactg-catgctttgatcgt-gagtttttgcagagat-ttcttgacaattttcagc-cgacgacatccccgaat-taaaccctatttcattt-gaaagaggatcttcgaat-taggttcctacaaa	AATTTTAGCCA-CAAA	caattaaaa-cacttttttt



4

Translesion synthesis of endogenous lesions protects genome integrity in *C. elegans*

Ivo van Bostelen, Robin van Schendel, Ron Romeijn, Marcel Tijsterman

Abstract

Bases within DNA are frequently damaged, resulting in blocks for transcription and replication. Loss of replication potential and the ensuing replication-associated DNA breaks induce genomic rearrangements that are at the basis of cancer and hereditary disease. The translesion synthesis (TLS) pathway safeguards replication potential by synthesizing DNA opposite damaged bases, however with reduced or no accuracy, thus leading to mutations. TLS has mainly been studied in systems that are exposed to genotoxins because the mutation rate resulting from spontaneous DNA damage is generally too low to be monitored under laboratory conditions. Here, we use *C. elegans* to study the role of TLS in maintaining genetic stability; we whole-genome sequenced mutation accumulation (MA) lines of TLS proficient and deficient animals that were grown under unchallenged conditions. For each genetic background we provide a comprehensive mutation profile that includes base substitutions, microsatellite indels, insertions, deletions, and complex genomic rearrangements. We complement our previous work on *rev-1, polh-1* and *polk-1* mutants with mutations in Polζ and with loss of PCNA ubiquitination at residue K165. In addition, we combined mutations in different TLS proteins to address redundancies in TLS activities: we created animals that are deficient for all Y-family polymerases, and for the first time created an animal that lost all TLS activity. Remarkably, these strains were viable and produced progeny, arguing that TLS is not essential for animal life. Yet, in completely TLS deficient backgrounds we observed increased genomic instability. Furthermore, we found that most endogenous replication blocks form at guanines, which in the absence of TLS lead to DSBs that are repaired via polymerase theta-mediated end-joining.

Introduction

Relentless encounters with endogenous and exogenous DNA damaging agents and the intrinsic instability of DNA call for efficient and faithful mechanisms to maintain the healthy state of the genetic material and thereby safeguard transcriptive and replicative potential [1]. Structural modifications of the bases that are paired between the phosphate-sugar backbones of the DNA molecule form roadblocks for transcription and replication. When replication stalls at the site of these lesions - and this is not resolved - DNA double strand breaks (DSBs) result, which can lead to genomic rearrangements or missegregation and eventually aneuploidy and cell death, which are all linked to cancer and inborn disease [2-4]. The DNA repair pathways Nucleotide Excision Repair (NER) and Base Excision Repair (BER) detect and repair damaged bases to restore DNA to a healthy state [5-7]. Although these pathways provide proficient protection, lesions can remain undetected or inflicted when the DNA is being replicated, leading to stalling of the replicative polymerases δ and ε; the narrow catalytic centers of these very accurate polymerases cannot accommodate structurally altered DNA [8]. The lesions that have the potential to form roadblocks for transcription and replication are diverse in structure and origin. Prominent exogenous sources include solar UV-C and UV-B that induce dimerization of adjacent pyrimidines, forming cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs). These intrastrand crosslinks distort the structure of DNA in such a way that replicative polymerases stall at these sites [9]. UV-A, UV-B and short wavelength visible light can also induce DNA damage indirectly via the formation of reactive oxygen species that cause oxidation of bases [10-13]. In addition, there are many other exogenous sources of DNA damage [14]. However, for a cell within a multicellular organism, endogenously generated base damages, resulting from metabolic processes within the cell, may be a much more prominent problem - because ever-present these strongly contribute to development of age-related disease [15]. The most common endogenous lesions are abasic sites, which form spontaneously or as an intermediate of BER. The second most abundant group of endogenous lesions are those that result from oxygen radicals attacking the DNA, with 7,8-dihydro-8-oxo-2′-deoxyguanosine (8oxoG) as the most often occurring base adduct. The steady-state number of endogenous base damages is estimated to be approximately 50.000 lesions per human cell and each of these lesions has the potential to induce base substitution mutations and stall DNA replication [4,15].

The abundance of endogenous base damages has led to a strong evolutionary selection pressure on mechanisms that sustain replicative potential in the presence of DNA damage. These DNA Damage Tolerance (DDT) pathways do not remove lesions but allow for synthesis of DNA across damaged bases. Hence DDT provides cells with the potential to complete DNA replication in the presence of damage, thereby suppressing the formation of replication-associated DSBs. There are two well-known DDT pathways

described in literature: template switching [16-18] and translesion synthesis [19,20]. Template switching (TS) uses replication of the undamaged strand opposite the lesion as a temporal template for replication and in that way permits bypass of the damage in an error-free way. It is efficiently employed in bacteria (where it is also referred to as damage avoidance) and yeast, yet it is unclear whether this pathway plays a prominent role in higher eukaryotes [2,21,22]. TS relies on complex homologous recombination (HR)-like reactions and is relatively time-consuming. In addition, it is thought that TS may result in genomic rearrangements if HR-like intermediates cannot be resolved properly [18,23]. Translesion Synthesis (TLS) - the other DDT pathway - is mechanistically much simpler and of unambiguous importance in multicellular eukaryotes. TLS employs specialized translesion polymerases that are able to bypass damaged templates. Their wider catalytic center and lack of proofreading capacity allow them to accommodate damaged bases and bulky adducts as templates and synthesize DNA opposite these lesions, enabling the continuation of DNA replication without actual repair [8,24]. In eukaryotes TLS is mediated by polymerases from the Y-family: Polη, Polκ, Polι, and REV1 and by B-family polymerase Polζ that consists of core subunits Rev3 and Rev7 [9,20,25]. Structural differences in their active sites define the functional specificities of TLS polymerases and bypass of some lesions may require the combined efforts of more than one TLS polymerase [10-13,26]. Yeast studies have shown that lesion bypass i) occurs directly at the replication fork by switching the replicative polymerase temporarily for a TLS polymerase, or ii) occurs in a post-replication manner, by filling in the single stranded gaps that remain opposite sites containing a DNA lesion [26-29].

Polη is arguably the best characterized of all TLS polymerases, which is likely because of the increased cancer risk and sensitivity towards sunlight of xeroderma pigmentosum variant (XPV) patients that have a mutation in the Polη encoding gene, highlighting its role in the bypass of sunlight-induced DNA damage [15,30-32]. Polη is also important for the bypass of many endogenous lesions, especially 8oxoG [15,33-35]. Bypass of CPDs and 8oxoG by Polη is remarkably error free, but on undamaged DNA Polη has, as compared to the replicative polymerases, low fidelity.

Polκ is the most conserved TLS polymerase being present in both prokaryotes and eukaryotes. While extensively studied in *E. coli*, the role of Polκ in eukaryotes is less clear as the loss of Polκ does not cause profound changes in spontaneous or damage-induced mutagenesis. As Polκ is able to extend mispaired primer termini and is relatively accurate on undamaged DNA a role of being an TLS extender has been suggested in addition to its specialized role in direct lesion bypass [24]. Similarly to Polη, also Polκ can interact with PCNA and REV1 [35].

A third conserved Y-family TLS polymerase, REV1, was identified by reduced UV-induced mutagenesis (reversionless phenotype) in yeast. Besides, Rev1 mutants have reduced levels of spontaneous mutagenesis, which also suggests a role in bypass of

endogenous lesions [36]. Remarkably, while REV1 shares its basic structure with Polη and Polκ, its catalytic activity is limited to the incorporation of deoxycytidines across undamaged or damaged guanines but also across adenines and abasic sites [37-40]. In addition to direct bypass through catalysis, REV1 also plays a non-catalytic role via interactions with other proteins including unmodified PCNA, K164-ubiquitinated PCNA, Y-family polymerases and the REV7 subunit of B-family TLS polymerase Polζ [35]. Because Rev1 and Rev3 mutant cells have similar phenotypes, the non-catalytic role of Rev1 may especially be important for Polζ activity [35].

Polζ consists of two core subunits: the catalytic subunit Rev3 and the accessory subunit Rev7. The Rev3 gene product contains a catalytic domain homologous to the other members of the B-family Polδ, ε and α [25,41,42]. Accessory subunits Pol31 and Pol32, that are shared with Pol δ, are also part of the Polζ complex [43-46]. As Polζ is able to extend aberrant primer termini it has been proposed to act as an extender polymerase acting after a Y-family polymerase has performed the insertion step during lesion bypass [26,47]. The processes that involve Polζ all carry a high risk of inducing base substitutions.

The high mutation rate and low processivity of TLS polymerases on undamaged DNA calls for a strict regulation of their activity; only allowing access to the DNA when absolutely necessary. Once the replication fork arrests at the site of a blocking lesion the eukaryotic sliding clamp, PCNA, is mono-ubiquitinated at lysine 164 by the E3 ubiquitin ligase RAD18 and the E2 conjugating enzyme RAD6 [48]. This reaction is promoted by the interaction of RAD18 with RPA-coated ssDNA at the site of the lesion. In addition, PCNA-K164 ubiquitination can be performed by CRL4-Cdt2 which may in fact be the more prevalent pathway when endogenous lesions are concerned [49]. Mono-ubiquitination of PCNA increases the interaction of Y-family polymerases with PCNA via the ubiquitin binding motifs at the C-terminus of these polymerases [50]. Yeast and mammalian cells that carry a K164R substitution have defects in TLS, but mammalian cells still retain partial TLS activity [48,51,52]. The remaining TLS activity in mammalian cells may be dependent on a non-catalytic function of Rev1 similar to what has been found in DT40 chicken cells [29]. While different studies show that PCNA ubiquitination is important for somatic hypermutation likely via the recruitment of error-prone TLS polymerases [53-55], it is unknown if PCNA-K164 modification influences spontaneous mutagenesis caused by error-prone TLS of endogenous lesions in multicellular organisms.

While many studies have shown TLS to be mutagenic, it has become clear that TLS actually protects against the formation of highly mutagenic replication-associated chromosome breaks and thereby strongly contributes to safeguarding genomic stability [2]. In previous studies we have described how the Y-family polymerases REV-1, Polη and Polκ contribute to genomic stability under unchallenged conditions [Chapter 3 and 56,57]. Here, we present the contribution of the B-family polymerase Polζ and

the TLS facilitating modification of PCNA in suppressing genome alterations. We also constructed animals that lack either all Y-family polymerases *(rev-1; polh-1; polk-1)* or all known TLS polymerases *(rev-1; polh-1; rev-3; polk-1)*. Our study presents the most comprehensive analysis of how TLS activity affects the stability of a genome under non-challenged circumstances.

Results

Polymerase ζ in *C. elegans* and its contribution to animal fitness

To study loss of Polζ activity we searched for a Rev3 ortholog in the *C. elegans* genome and found open reading frame (ORF) Y37B11A.2. The predicted ORF has two isoforms: Y37B11A.2a and Y37B11A.2b with predicted protein products of 1303 amino acids (aa) and 1019 aa, respectively. Both putative gene products are much smaller than the vertebrate protein REV3L (3130 aa [58,59] or alternatively spliced 3052 aa [60]), but comparable to Rev3 in *S. cerevisiae* (1504 aa) [25]. The CLUSTAL alignment of the REV3 protein sequences from yeast, fruit fly, mouse and human reveal extensive evolutionary conservation especially at the C-terminal polymerase domain (Suppl. fig. 1a).

We next obtained the allele gk919715, generated in the million mutation project [61], which carries a point mutation leading to a stop codon in exon 3 of isoform Y37B11A.2a, and in exon 1 of Y37B11A.2b (Suppl. fig. 1a). We assume this allele to be a null allele, not encoding a viable protein product; without functional catalytic subunit REV-3 there is no functional Polζ. After a number of backcrosses to N2, to remove background mutations, we quantified the brood size and identified a significant ($P<0.0001$, Mann-Whitney test) decrease: WT animals have on average broods of 250-260 progeny [±SEM 7], *rev-3* mutants broods are ~170 [±SEM 15] (Fig. 1a); no significant embryonic lethality is observed (Fig. 1b). In order to test whether TLS is impaired we exposed *rev-3* mutants to UV-C, a bona fide inducer of TLS substrates. We find *rev-3* mutants to have decreased embryonic survival in response to UV-C exposure as compared to WT animals (Fig. 1c).

The contribution of PCNA-K165 ubiquitination to animal fitness

With the isolation of a *rev-3* allele, we complemented a set of mutant strains comprising all TLS polymerases encoded by the *C. elegans* genome. We decided to extend this set with a strain carrying a mutation that in other systems disrupts the efficient recruitment of TLS polymerases to sites of blocked replication, *i.e.* the K164R mutation in PCNA [48,53,54,62]. Using CLUSTAL alignments of PCNA protein sequences from different species we found that K165 in *C. elegans* PCNA corresponds to residue K164 in yeast, fly and mammalian PCNA (Suppl. fig. 1b). We used homology-driven repair of a CRISPR/Cas9-induced DSB to generate K165R in *C. elegans* PCN-1. In all systems tested, this substitution blocks mono- and poly-ubiquitination as well as SUMOylation at this site.

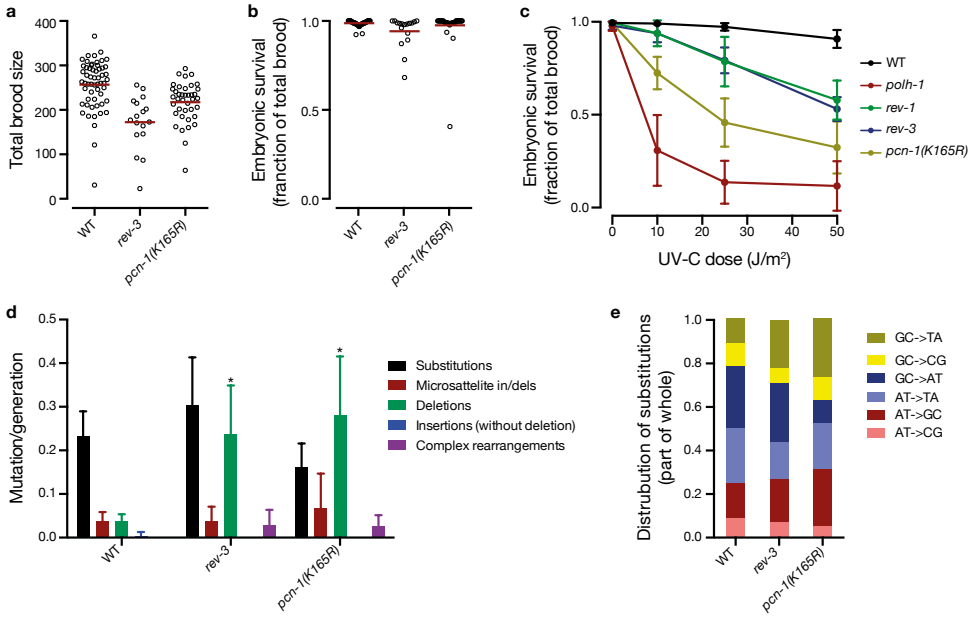


Figure 1: Characterization of rev-1 and pcn-1(K165R) translesion synthesis mutants. *a) Total brood sizes were determined of animals of the indicated genotypes. Each dot represents the brood of a single animal and the red line indicates the mean. b) The embryonic survival of the indicated genotypes was determined. Each dot represents the embryonic survival of the brood of a single animal and the red line indicates the mean. c) Hermaphrodite adult animals of indicated genotypes were exposed to different doses of UV-C and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20 hours time period post irradiation. d) Frequencies of the different mutation categories of the indicated genotypes determined by whole genome sequencing. Significant differences ($P<0.05$) from WT frequencies are indicated with *. e) Distribution of the base substitutions of the indicated genotypes determined by whole genome sequencing.*

We tested animal fitness of homozygous *pcn-1(K165R)* mutant animals by determining their brood size and observed a small but statistically significant decrease as compared to WT animals (217 [±SEM 7] versus 257 [±SEM 7], $P<0.0001$; Fig. 1a). No effect on embryonic viability was observed (Fig.1b). To investigate whether *pcn-1(K165R)* mutant animals display phenotypes associated with a TLS deficiency we exposed animals to UV-C. Here, we observed an increased sensitivity of *pcn-1(K165R)* animals as compared to WT animals. The extent of increased UV-sensitivity is, however, not as severe as for *polh-1* knockout animals, which was assayed as a reference, indicating that without the possibility of K165 ubiquitination UV-C lesions can still be bypassed by POLH-1, although with a significantly lower efficiency. This is consistent with previous studies that show that TLS is not completely abrogated in a PCNA-K164R mutant conditions [48,55,62].

REV-3 and PCNA-K165 modification guard genome stability

Spontaneous base substitutions in yeast are in part dependent on Polζ and PCNA-K164 modification, which suggests that Polζ- and PCNA-K164-dependent TLS on endogenous lesions can be mutagenic [51,63]. Up until now it is unclear whether Polζ-dependent bypass of endogenous lesions in multicellular eukaryotes is similarly mutagenic and even less is known about the influence of PCNA-K164/K165 modification on spontaneous mutagenesis. We found that both mutant backgrounds have reduced broods arguing for an involvement in maintaining genetic stability also under non-challenged growth. To study the influence of these TLS factors on spontaneous mutagenesis we established mutation accumulation lines: starting from a single animal we set up several lines for each mutant background, transferred a few of the progeny every generation to a new culture plate, and after 50 generations we isolated DNA and sequenced the genome of 3 or 4 lines for each genotype. This 'micro-evolution' experiment yields both the spontaneous mutagenic spectrum and rate (mutation/generation) for each genetic background.

Perhaps surprisingly but animals that have grown for many generations without Polζ activity did not display significant differences in the base substitution rate and nature when compared to WT animals (Fig. 1d,e). This observation either argues that REV-3 dependent TLS on spontaneous lesions does not cause substantial numbers of point mutations or that back up mechanisms are equally or more mutagenic. Also, for *pcn-1(K165R)* mutant animals we observed no statistically significant difference for base substitution rate or nature (Fig. 1d,e). From these observations we conclude that bypass of spontaneous lesions dependent on REV-3 and PCNA modification is not a substantial inducer of base substitutions.

We next determined the frequencies of other mutation types, *i.e.* deletions and insertions. While, as expected, we did not observe significant changes in microsatellite instability in either of our mutants, we found a marked increased rate of deletion induction in both the *rev-3* and *pcn-1(K165R)* backgrounds (Fig. 1d). In addition, both mutant backgrounds spawned complex rearrangements, which were not induced in WT animals. We thus conclude that REV-3 and modification of PCNA-K165 are protectors of genomic integrity. Because the formation of deletions and more complex rearrangement do not go together with a drop in rate of base pair mutagenesis, we conclude that TLS on the deletion-causing lesion is without error – apparently the lesions that cause deletions in the mutant background do not contribute to the substitution rate in WT. This notion argues that the type of endogenous damage that require REV3 or PCNA ubiquitination is not without some degree of base pairing capacity.

Loss of all Y-family activity and loss of all TLS activity

Because of the documented redundancy of Y-family TLS polymerases [56,64], and to study the total contribution of Y-family dependent TLS to animal fitness and

spontaneous mutagenesis, we combined knockout mutations in the genes encoding Polη, Polκ and REV1 generating *polh 1;polk 1;rev-1* triple knockout mutant animals (note that the *C. elegans* genome does not encode Polι). To assess the effects of complete loss of bypass potential we also crossed-in the *rev-3* mutation to generate: *rev-1;polh-1;polk-1;rev-3* quadruple knockout animals. These animals allow us to address, for the first time, how complete lack of TLS activity effects animal fitness and spontaneous mutagenesis.

We first examined animal fitness under unchallenged conditions. We found brood size and embryonic survival of the "Y-family polymerase dead" and "TLS-dead" mutants to be lower than WT, yet remarkably mildly affected, and animals also develop normally (Fig. 2a,b). From this we conclude that a multicellular animal (*C. elegans* has ~1000 cells) can proliferate without TLS activity, presumably because the level of spontaneous replication-blocking DNA damage is low and perhaps also because repair pathways exist that can process stalled forks.

Mutation accumulation in double, triple and quadruple TLS mutants

Although the detrimental effects of loss of all bypass potential on animal fitness are apparently limited, the effects on a larger time scale can be more severe, as is evident from the increased levels of spontaneously occurring deletions in the absence of Polη and Polκ [56]. Here, we re-analyze the MA data from *polk-1;polh-1* mutant animals with updated software, complemented with mutation accumulation data for *rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutant animals. Remarkably, we find no significant differences for substitution rates in *polh-1;polk-1, rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutants, arguing that spontaneous mutagenesis in *C. elegans* is not resulting from TLS action. Also, no change in microsatellite instability was observed. However, from the following observation we conclude that TLS is predominantly acting in an error-free manner to bypass spontaneous damage: we observed an approximately 30-fold induction in the rate with which deletions are formed in *polk-1;polh-1, rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutants. In all these backgrounds, also more complex events were found: a collection of inversions, tandem duplications or events that combine different categories of genomic rearrangements. This data further substantiates our earlier conclusion that TLS on spontaneous lesions in not a strong driver of mutagenesis but it is in fact an essential protector of genomic stability. Interestingly, loss of REV-1 and REV-3 does not further increase the genomic instability that is observed in the *polh-1;polk-1:* the Y-family-dead and the TLS-dead mutant animals have a near identical mutational load as *polh-1;polk-1* mutant animals. These observations together with the notion that loss of REV-1 and REV-3 by themselves only marginally increase deletion mutagenesis argue that Polη and Polk act redundantly to bypass spontaneous DNA damage in *C. elegans* in an error-free manner, in which they do not require the action of REV-1, as a scaffold, or REV-3, as an extender.

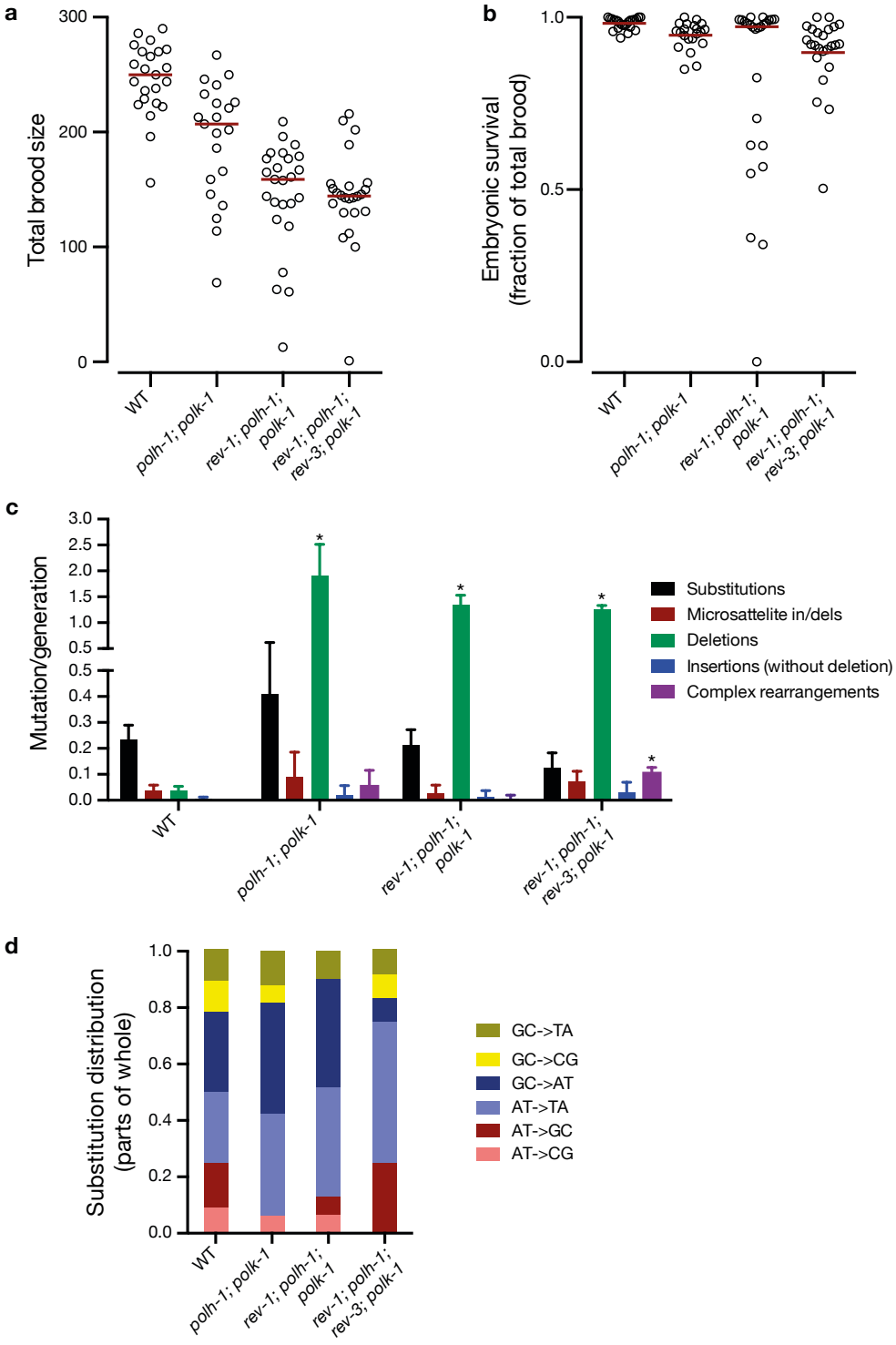

a
Total brood size
300
200
100
0
WT
polh-1; polk-1
rev-1; polh-1; polk-1
rev-1; polh-1; rev-3; polk-1
b
Embryonic survival (fraction of total brood)
1.0
0.5
0.0
WT
polh-1; polk-1
rev-1; polh-1; polk-1
rev-1; polh-1; rev-3; polk-1
c
Mutation/generation
3.0
2.5
2.0
1.5
1.0
0.5
0.5
0.4
0.3
0.2
0.1
0.0
*
Substitutions
Microsattelite in/dels
Deletions
Insertions (without deletion)
Complex rearrangements
WT
polh-1; polk-1
rev-1; polh-1; polk-1
rev-1; polh-1; rev-3; polk-1
d
Substitution distribution (parts of whole)
1.0
0.8
0.6
0.4
0.2
0.0
GC->TA
GC->CG
GC->AT
AT->TA
AT->GC
AT->CG
WT
polh-1; polk-1
rev-1; polh-1; polk-1
rev-1; polh-1; rev-3; polk-1

Figure 2: Characterization of mutants deficient for multiple TLS polymerases.
*a) Total brood sizes were determined of animals of the indicated genotypes. Each dot represents the brood of a single animal and the red line indicates the mean. b) The embryonic survival of the indicated genotypes was determined. Each dot represents the embryonic survival of the brood of a single animal and the red line indicates the mean. c) Frequencies of the different mutation categories of the indicated genotypes determined by whole genome sequencing. Note the scale difference on the y-axis when compared to figure 1d. Significant differences ($P<0.05$) from WT frequencies are indicated with *. d) Distribution of the base substitutions of the indicated genotypes determined by whole genome sequencing.*

The deletion footprints in all TLS mutants are consistent with Polθ dependent repair of replication associated DSBs

Combined with our earlier work we have established increased levels of deletion formation in *polh-1, rev-1, rev-3, pcn-1(K165R)* single mutant animals (not in *polk-1*) and in *polk-1;polh-1* double mutant animals. Further depletion of *rev-1* and/or *rev-3* did not exacerbate mutagenesis. When the size of all deletions is plotted a strikingly specific distribution becomes apparent: the vast majority of deletions are between 50 and 500 base pairs (bp), with a median of 108 bp; a size category we almost never find in WT animals (Fig. 3a & Suppl. table 1). Apart from their size these deletions share other characteristics: i) a subset of deletions in all genetic backgrounds has DNA insertions, which are in sequence often identical to sequence stretches in the immediate flank of the deletion (Fig. 3b & Suppl. table 1), ii) deletions without insertions display at the deletion junctions so-called micro-homology. (Fig. 3c). We determine microhomology by comparing the last nucleotides immediately upstream of the deletion to the most downstream nucleotides of the deletion (further details on the methodology can be found in [65]). Previous studies from our lab have shown that the formation of these deletions is the result of polymerase theta-mediated end joining (TMEJ) of replication-associated double strand breaks (DSBs) [65-67]. The fact that we observe the same deletion signature in the mutants described here, strongly supports that TMEJ is the go to repair pathway in the event of DSBs caused by stalled replication when TLS is not functional.

Break point analysis

Next, we examined whether the base composition at deletion junctions could provide hints towards the underlying spontaneous lesion. Previous work, analyzing deletions formed at replication-blocking G-quadruplexes, revealed that at least one of the breakpoints maps very close, if not immediately adjacent, to the replication impediment [68]. We rationalized that this may also occur at spontaneous lesions in a TLS-compromised animal: one could argue that the nascent strand, which is extended right up to the blocking lesion is a substrate for TMEJ and thus defines one end of the deletion. According to this logic, the first deleted base at the junction will frequently

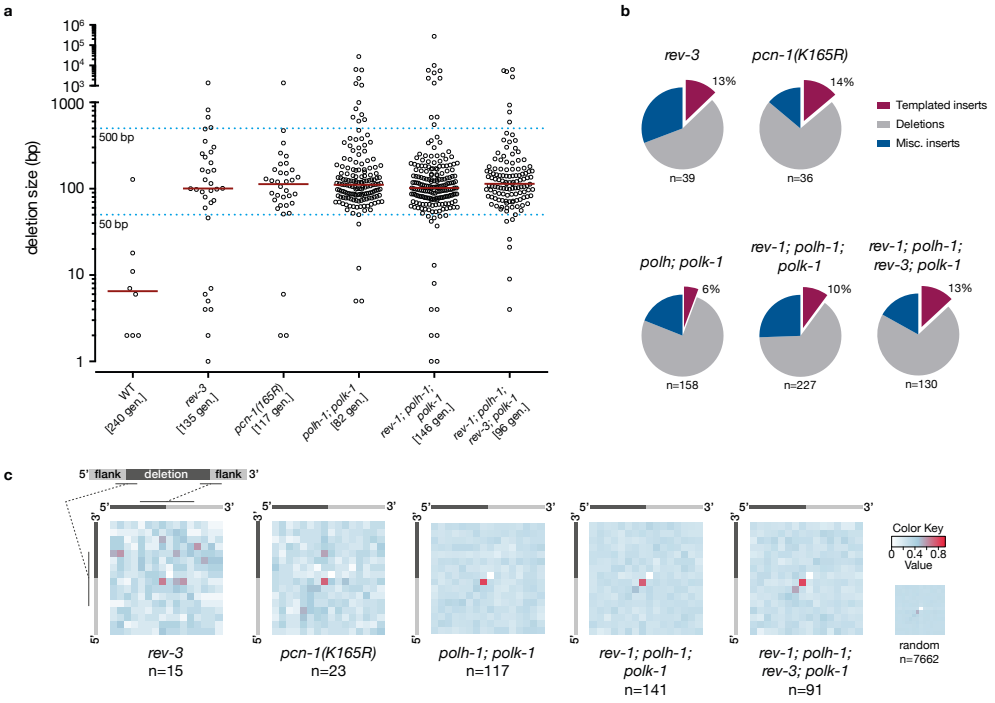


Figure 3: Footprint of deletions that occur spontaneously in TLS deficient mutants. *a) Size representation of all deletions found in the mutation accumulation data of the indicated genotypes. Red bars indicate means. b) Distribution of deletion alleles: the majority are simple deletions (in grey), without containing insertions in between both breakpoints. Some deletions have small miscellaneous inserts of 1 to 6 bp (in blue), whereas others (in magenta) have an insert that is templated from the flank of the break. c) Heat map representations of micro-homology at deletion breakpoints of the indicated genotypes. The bases that flank the right and left deletion breakpoints and are either retained or lost in the deletion alleles are plotted against each other. The heat map, representing the indicated number of deletions, reveals overrepresentation of 1 nucleotide of micro-homology in all mutants. A heat map for a simulated set of random deletions (n=7662) with random distribution is displayed on the right.*

represent the nucleotide that is complementary to the damaged base. To assess this idea, we have determined the normalized base distribution at the deletion junctions.

The total numbers of deletions in the *rev-3* and *pcn-1(K165R)* mutants were too low (n=34 in both cases) for a statistically founded analysis of base distribution effects at deletion junctions. However, the number of deletions derived from of *polk-1;polh-1, rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutant animals are sufficiently substantial to address questions about the most prevalent endogenous damage that requires TLS action. In all three mutants we detect a non-random distribution, *i.e.* a significant enrichment for cytosines at the -1 position (Fig. 4a-d). From this observation we infer that the first nucleotide that cannot be incorporated at a replication impediment in

polk-1;polh-1, *rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutant animals is prevalently a cytosine, arguing that TLS at guanines is required to suppress replication-associated DSBs. This may point to N2-dG and/or 8oxoG lesions as the most important source of spontaneous mutagenesis. Damaged guanines as the primary source for the deletion junctions also explains the apparent enrichment of cytosines at the more downstream positions -2 until -6 because TMEJ itself frequently results in loss of a few nucleotides at one or either side of a DSB [69]. We also analyzed a subset of deletions in which the influence of TMEJ is expected to be reduced: because simple deletions mostly originate from TMEJ using microhomology this feature perturbs the correct annotation of the deletion junction (e.g. three different annotations are possible for a deletion with 2 bases of microhomology). To overcome this issue, we quantified the base distributions using only those deletions that contain small inserts as for these cases the use of homology is not a disturbing factor. Using this subset, the enrichment for cytosines at position -1 indeed becomes more pronounced.

Discussion

For this study we set out to complement our previous work on the contribution of TLS on spontaneous mutagenesis. We find that Polζ and PCNA ubiquitination at K165 both contribute to TLS of spontaneous DNA damage and thereby protect genome stability. We also assayed animals that lacked all Y-family polymerases animals that in addition lacked Polζ and thereby are considered to be devoid of all TLS activity. Surprisingly, we found that TLS is not essential for animal life as TLS-dead animals proliferate and produce morphologically normal progeny. We found that Y-family-dead and TLS-dead animals over generations accumulate mutations at the same rate and with the same characteristics as *polh-1;polk-1* mutant animals, while the increase in mutation accumulation in *rev-1* or *rev-3* animals is only minor, which argues that Polη and Polk i) act redundantly to bypass the vast majority of spontaneous DNA damage in *C. elegans*, ii) act in an error-free manner, and iii) for most lesions do not require the action of REV-1 or REV-3. The comprehensive mutation profiles we present for different TLS compromised *C. elegans* strains consistently shows that TLS is anti-mutagenic and we find that the most common endogenous replication blocks reside at guanines, which in the absence of TLS generate replication-associated DSBs that are repaired by polymerase theta-mediated end-joining.

When the first eukaryotic factors of translesion synthesis (Rev1, Rev3 and Rev7) were characterized in yeast they were found to promote both induced mutagenesis and spontaneous mutagenesis [Reviewed in 36]. The major causes for why TLS is mutagenic lies in the "hard-to-read" character of bulky lesions, the wider catalytic centers of TLS polymerases (allowing for less stringent base pairing), and the lack of proofreading. In bacteria and yeast, TS provides an error free alternative to TLS, which may explain

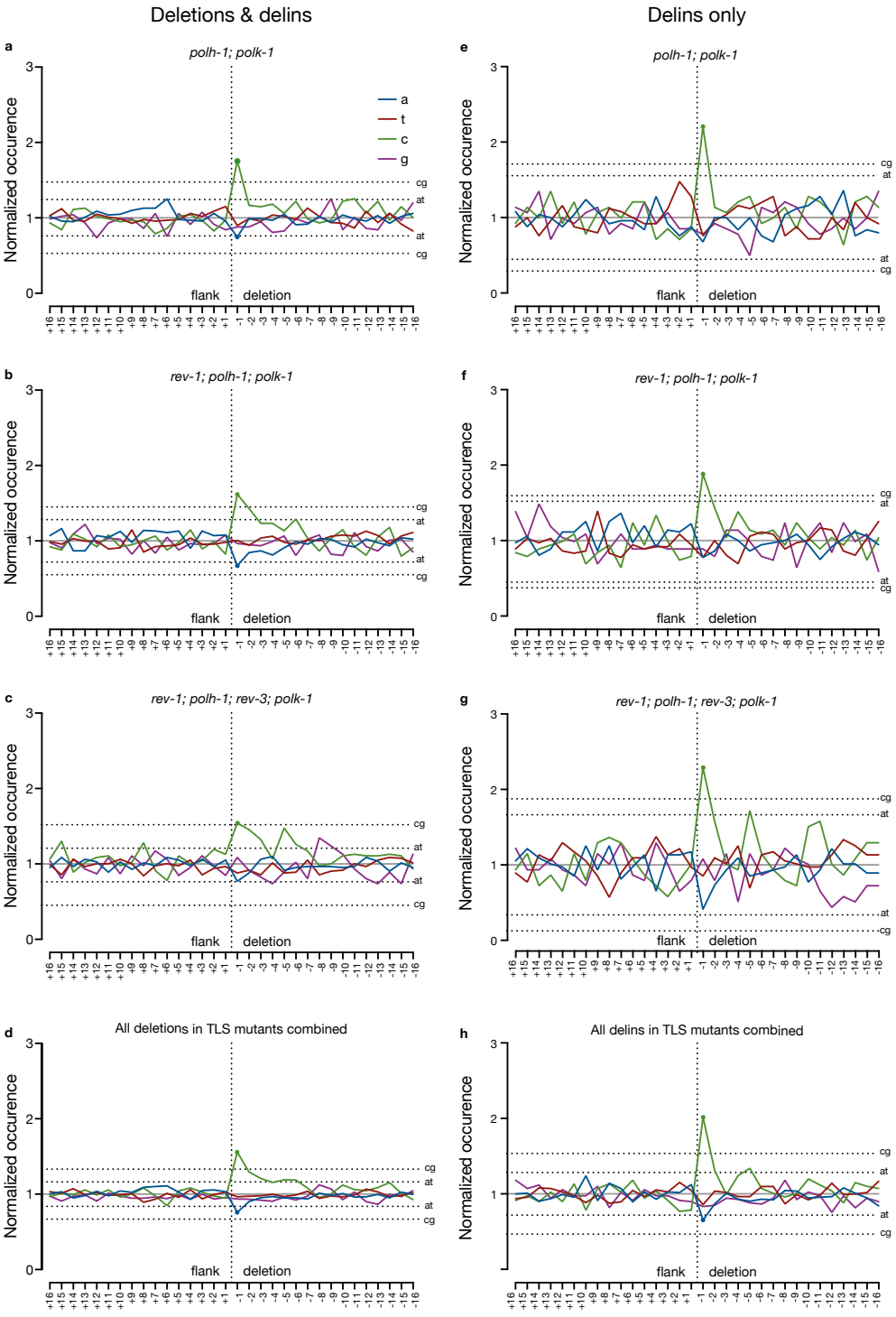

Deletions & delins
Delins only
a
polh-1; polk-1
b
rev-1; polh-1; polk-1
c
rev-1; polh-1; rev-3; polk-1
d
All deletions in TLS mutants combined
e
polh-1; polk-1
f
rev-1; polh-1; polk-1
g
rev-1; polh-1; rev-3; polk-1
h
All delins in TLS mutants combined
Normalized occurence
a
t
c
g
cg
at
flank
deletion

Figure 4: Break point analysis. *a-d) Base composition at the 5' and 3' junctions of deletions and delins of the indicated genotypes. The flanking sequences have positive numbers; the deleted sequences have negative, -1 being the first nucleotide within the deletion. The dotted lines are drawn at +/-3.5x SD of the indicated bases and represent the variation of these bases along the complete sequence (+16 until -16). Significant deviations (>3.5xSD or <-3.5xSD) are indicated with a solid circle. e-h) Base composition at the 5' and 3' junctions of only delins of the indicated genotypes. The flanking sequences have positive numbers; the deleted sequences have negative, -1 being the first nucleotide within the deletion. The dotted lines are drawn at 3.5x SD of the indicated bases and represent the variation of these bases along the complete sequence (+16 until -16). Significant deviations (>3.5xSD or <-3.5xSD) are indicated with a solid circle.*

the lower mutation rate in TLS deficient mutants [21]. While such mechanism may be present in higher eukaryotes also, the importance of it remains unclear [22]. In yeast TS is mediated by Rad5. We have, however, not observed any effect on UV-C sensitivity of the knockout of the homologous gene in *C. elegans*, either by itself or in a Polη deficient background (Suppl. fig. 2). In the multicellular *C. elegans* model we find that loss of TLS functionality does not lead to a strong reduction in spontaneous base substitutions levels. Instead, we find that in absence of TLS the overall mutagenic consequence is strongly increased as small deletions (50 to 500 bp) manifest, which by virtue of being ORF disrupting have a more detrimental outcome than TLS–induced SNVs.

Loss of Polζ in *C. elegans* only mildly affects animal fitness under unchallenged conditions, while Polζ is essential for mammalian development: missense mutations in the human REV3L gene can cause Mobius syndrome which is associated with developmental abnormalities [70], and Rev3 knockout mice are embryonic lethal [71]. Because Polζ is not essential for survival in *C. elegans*, we were able to study this TLS polymerase in the context of a multicellular eukaryote, finding conserved roles for Polζ in the bypass of UV-C induced and spontaneous replication blocking lesions and an overall anti-mutagenic effect of REV-3 dependent TLS. An explanation for the essential function of mammalian Polζ could be that it has additional features for which the origin lies in the N-terminal part of the protein that is not present in yeast and *C. elegans*. The mammalian REV3L gene produces a protein roughly twice the size of the *C. elegans* protein, which may indicate more complex or diverse functions in higher eukaryotes. Alternatively, the genome of mammals is 30 times as large and perhaps also the growth conditions of cultured cells (high oxygen) may provide further demand on TLS activity.

As expected, we show that modification of *C. elegans* PCNA at K165 (equivalent to K164 in other organisms) is an important but not essential promotor of TLS. In the Y-family polymerase deficient mutants we observe a much higher frequency of deletion formation than in the *pcn-1(K165R)* mutant and also the UV-C sensitivity of the latter is not as severe as observed for a polh 1 mutant, which is consistent with studies in yeast and MEFs that show TLS is still partially functional in PCNA-K164R mutants [48,62]. These observations may reflect that TLS is generally less efficient in the *pcn-1(K165R)* mutant or that a specific subset of a yet unknown type of lesions require PCNA-K165

modification. Two distinct sub-pathways of TLS have been suggested, one requiring PCNA-K164 ubiquitination and the other requires regulation by Rev1 [29], explaining why TLS is still functional in the *pcn-1(K165R)* mutant. It would thus be interesting to assay *rev-1; pcn-1(K165R)* double mutant animals. Other modifications of PCNA-K164 are linked to different damage tolerance or repair pathways: RAD5-dependent poly-ubiquitination is thought to induce TS and SUMOylation is linked to suppression of HR [48,72,73]. We anticipated a potential changed mutation profile because of possible deregulation of these two other DNA damage response pathways, but the mutation profile of *pcn-1(K165R)* mutant animals is conform that of a TLS mutant, hence we propose that PCNA-K165 modification in response to spontaneous DNA damage specifically leads to activation of TLS.

Previous work has shown that the specific footprint of small deletions resulting from TLS impairment is the result of TMEJ of replication-associated DSBs [65-68]. Although the rate of deletion induction is different for different TLS mutant backgrounds [57,66], the characteristics are remarkably similar. Deletion junction analysis in *polk-1;polh-1*, *rev-1;polh-1;polk-1* and *rev-1;polh-1;rev-3;polk-1* mutant animals all point to damaged guanines as primary substrates for TLS activity. Interestingly, OGG1, the primary glycosylase that in many other biological systems removes 8oxoG via BER [Reviewed in 74] is not encoded by the *C. elegans* genome, which may provide an explanation for this strong base effect.

TLS may be especially important during *C. elegans* embryogenesis, because embryos are extremely sensitive to cell cycle delays and thus need a quick fix for stalled replication [56,75,76]. The frequency of deletion formation in the TLS-dead mutants may also provide an indication for how often a replication-fork runs into an insurmountable block for the replicative polymerases in the context of a multicellular organism. In fact, if one assumes that in a TLS-dead mutant context every lesion that cannot be bypassed forms a deletion of >50bp, it becomes clear how rare spontaneous replication blocks are. In an estimated 10-15 rounds of replication per generation we find 1 deletion. For a genome of 100.000.000 bases, this means that only 1 in ~10^9 bases requires TLS for its replication. Because replication-associated breaks are such rare events, they may not induce DNA damage signaling, causing them to be highly mutagenic and easily transmitted to next generations [67].

From our study we conclude that TLS not only protects replication potential, it also protects genomic stability and is *grosso modo* anti-mutagenic. Spontaneous damages that require TLS are not strongly mutagenic, in fact, in the absence of error-free TLS, damaged guanines cause genomic changes that are more deleterious than point mutations. This predominantly beneficial outcome of TLS activity provides evolutionary conservation of these specialized (so-called mutagenic) polymerases in all living organisms [20,77].

Materials & Methods

General culturing and strains used

All strains were cultured according to standard methods [78]. The N2 Bristol strain was used as WT control. The *C. elegans* REV-3 ortholog was found after using the S. cerevisiae nucleotide sequence as an input to perform a BLAST-search on the *C. elegans* genome. The ORF is currently known as Y37B11A.2a and both the ORF and protein sequence show significant homology to REV3 protein sequences from other organisms (Suppl. fig. 1). The allele gk919715, generated in the million mutation project [61], introduces an early stop in the ORF. This *rev-3(gk919715)* allele and *rev-1(gk455794)*, were obtained from the Caenorhabditis Genetics Center, Minnesota, USA. The *polh-1(lf0031)* and *polk-1(0029)* alleles were isolated in our own laboratory [56] and for this study CRISPR/cas9 mediated genome editing was employed to induce a the point mutation in the *pcn-1* gene that induced the K165>R substitution (see below) yielding the *pcn-1(lf0222)* allele, also referred to as *pcn-1(K165R)*. Combinations of the different alleles were made by crossing single and double mutants.

CRISPR/Cas9 mediated targeted genome editing of pcn-1

First, the genomic target in pcn-1 was determined. Protein sequences were aligned and it was established that amino acid residue K165 aligned to the residue K164 in human, mice and yeast. A CRISPR target at the genomic location corresponding to this residue was designed using online software (crispr.mit.edu): GAACACGATGCCAGCCTTGG. Cloning of the vectors was done as described in [79-84]. For the induction of the necessary K165R aa substitution a repair substrate was made by PCR-amplifying a 3 kb fragment around the genomic location using primers: ttataaaattcgtcaaaataagttgc (forward) and ttgagcaaaaagcgtaaaacc (reverse) on genomic DNA. After cloning the amplicon into pGEMT, site directed mutagenesis was performed with primers: cgatgccagccctggtggcggtgatg (forward) and catcaccgccaccagggctggcatcg (reverse) and the Quick-change II kit of Agilent Technologies.

Plasmids were injected using standard *C. elegans* microinjection procedures. Briefly, 1 day before injection, L4 animals were transferred to new plates and cultured at 15 degrees. Gonads of young adults were injected with a solution containing 30 ng/µl, pDD162 (Peft-3::Cas9 (Addgene 47549; ref. [85]), 100 ng/µl pMB70 (u6::sgRNA with pcn-1 target, 30 ng/µl pGEMT-pcn1K165R-HRtemplate, 10 ng/µl pGH8, 2.5 ng/µl pCFJ90, 5 ng/µl pCFJ104. Progeny (F1) animals that express mCherry were picked to new plates 3–4 days post injection and allowed to produce offspring. Of each F1 plate 10 F2 animals were pooled, lysed and genotyped. Genotyping was done by PCR amplification a 448 bp product around the CRISPR/Cas9 target site with primers: GCGACTCGATCATCTTCACA (forward) and CCATTCTCCTCGATCGGATA (reverse). The induced mutation does not only

induce a K165R amino acid substitution, it also creates a recognition site for the BstNI restriction enzyme. Digestion with BstNI enzyme of the WT genomic sequence results in an uncut fragment of 448 bp while the mutant sequence results in two fragments of 206 bp and 242 bp.

Brood size & embryonic survival

To determine brood sizes, L4 animals were singled on OP50 plates. Every day the mother was transferred to a new plate and one day later embryos and larvae on the plate were quantified. In each experiment broods of at least 20 mothers were determined.

UV-C survival assays

To measure germline sensitivity to UV-C, staged young adults (one day post L4) were transferred to empty NGM plates and exposed to different doses of UV light. Per dose and genotype 3 plates with 3 adults were transferred to fresh NGM plates with OP50 and allowed to lay eggs for 20 hours. Subsequently, adults were discarded and the brood on the plate was allowed to hatch. 24 hours later the numbers of non-hatched eggs and surviving progeny were determined. The UV-C source was predominantly 254 nm (Philips). Before every UV-C exposure the irradiance of our light source was determined using an International Light photometer (model: IL1400BL, ser. nr.: 7819). This varied slightly between experiments with an average of 26,0 (±1,0) μW*cm-2 (equals 0,260 J*m-2*s-1).

Mutation accumulation & bio-informatic analysis

Mutation accumulation (MA) lines were established by transferring single F1 animals that originated from a homogenous parent, starting 6 clonal MA lines. Of each line three worms of the next generations were transferred to a new plate, marking every generations. MA lines were maintained for approximately 50 generations. Not all lines grew at the same speed because some of them developed growth retardation. At the end of MA, of each line three animals were singled to new plates and propagated to obtain full clonal plates for DNA isolation. Worms were washed off in M9 and to remove bacteria from the intestines the worms were incubated on a shaker for 2h at RT. Subsequently, genomic DNA was isolated using a Blood and Tissue Culture Kit (Qiagen). Whole genome sequencing and bio-informatics were performed as described in [69]. The methods and bio-informatics that were used to create the heat maps and the break point analysis were performed as described in [65].

References

[1] Lindahl T. Instability and decay of the primary structure of DNA. Nature 1993;362:709–15. doi:10.1038/362709a0.

[2] Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome. Transl Cancer Res 2013;2:107–29. doi:10.3978/j.issn.2218-676X.2013.04.01.

[3] Barnes DE, Lindahl T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu Rev Genet 2004;38:445–76. doi:10.1146/annurev.genet.38.072902.092448.

[4] Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. Toxicol Sci 2011;120 Suppl 1:S130–45. doi:10.1093/toxsci/kfq371.

[5] Baute J, Depicker A. Base excision repair and its role in maintaining genome stability. Crit Rev Biochem Mol Biol 2008;43:239–76. doi:10.1080/10409230802309905.

[6] Krokan HE, Bjoras M. Base Excision Repair. Cold Spring Harb Perspect Biol 2013;5:a012583–3. doi:10.1101/cshperspect.a012583.

[7] Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JHJ. Understanding nucleotide excision repair and its roles in cancer and ageing. Nature Reviews Molecular Cell Biology 2014;15:465–81. doi:10.1038/nrm3822.

[8] Branzei D, Foiani M. Maintaining genome stability at the replication fork. Nature Reviews Molecular Cell Biology 2010;11:208–19. doi:10.1038/nrm2852.

[9] Beukers R, Eker APM, Lohman PHM. 50 years thymine dimer. DNA Repair 2008;7:530–43. doi:10.1016/j.dnarep.2007.11.010.

[10] Kielbassa C, Roza L, Epe B. Wavelength dependence of oxidative DNA damage induced by UV and visible light. Carcinogenesis 1997;18:811–6.

[11] Kuluncsics Z, Perdiz D, Brulay E, Muel B, Sage E. Wavelength dependence of ultraviolet-induced DNA damage distribution: involvement of direct or indirect mechanisms and possible artefacts. J Photochem Photobiol B, Biol 1999;49:71–80. doi:10.1016/S1011-1344(99)00034-2.

[12] Zhang X, Rosenstein BS, Wang Y, Lebwohl M, Mitchell DM, Wei H. Induction of 8-oxo-7,8-dihydro-2'-deoxyguanosine by ultraviolet radiation in calf thymus DNA and HeLa cells. Photochem Photobiol 1997;65:119–24.

[13] Godley BF, Shamsi FA, Liang F-Q, Jarrett SG, Davies S, Boulton M. Blue light induces mitochondrial DNA damage and free radical production in epithelial cells. J Biol Chem 2005;280:21061–6. doi:10.1074/jbc.M502194200.

[14] Friedberg EC. A brief history of the DNA repair field. Cell Res 2008;18:3–7. doi:10.1073/pnas.80.15.4639.

[15] Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell 2017;168:644–56. doi:10.1016/j.cell.2017.01.002.

[16] Blastyák A, Pintér L, Unk I, Prakash L, Prakash S, Haracska L. Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell 2007;28:167–75. doi:10.1016/j.molcel.2007.07.030.

[17] Atkinson J, McGlynn P. Replication fork reversal and the maintenance of genome stability. Nucleic Acids Res 2009;37:3475–92. doi:10.1093/nar/gkp244.

[18] Branzei D, Foiani M. Template switching: from replication fork repair to genome rearrangements. Cell 2007;131:1228–30. doi:10.1016/j.cell.2007.12.007.

[19] Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harb Perspect Biol 2013;5:a012708. doi:10.1101/cshperspect.a012708.

[20] Sale JE, Lehmann AR, Woodgate R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Reviews Molecular Cell Biology 2012;13:141–52. doi:10.1038/nrm3289.

[21] Fuchs RP. Tolerance of lesions in E. coli: Chronological competition between Translesion Synthesis and Damage Avoidance. DNA Repair 2016;44:51–8. doi:10.1016/j.dnarep.2016.05.006.

[22] Unk I, Hajdu I, Blastyák A, Haracska L. Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair 2010;9:257–67. doi:10.1016/j.dnarep.2009.12.013.

[23] Lee JA, Carvalho C, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell 2007.

[24] Prakash S, Johnson RE, Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem 2005;74:317–53. doi:10.1146/annurev.biochem.74.082803.133250.

[25] Gan GN, Wittschieben JP, Wittschieben BØ, Wood RD. DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res 2008;18:174–83. doi:10.1038/cr.2007.117.

[26] Prakash S, Prakash L. Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. Genes & Development 2002;16:1872–83. doi:10.1101/gad.1009802.

[27] Jansen JG, Fousteri MI, de Wind N. Send in the clamps: control of DNA translesion synthesis in eukaryotes. Mol Cell 2007;28:522–9. doi:10.1016/j.molcel.2007.11.005.

[28] Temviriyanukul P, van Hees-Stuivenberg S, Delbos F, Jacobs H, de Wind N, Jansen JG. Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair 2012;11:550–8. doi:10.1016/j.dnarep.2012.03.007.

[29] Edmunds CE, Simpson LJ, Sale JE. PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell 2008;30:519–29. doi:10.1016/j.molcel.2008.03.024.

[30] Johnson RE, Prakash S, Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 1999;283:1001–4.

[31] Johnson RE, Kondratick CM, Prakash S, Prakash L. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 1999;285:263–5. doi:10.1126/science.285.5425.263.

[32] McDonald JP, Levine AS, Woodgate R. The Saccharomyces cerevisiae RAD30 Gene, a Homologue of Escherichia coli dinB and umuC, Is DNA Damage Inducible and Functions in a Novel Error-Free Postreplication Repair Mechanism. Genetics 1997;147:1557–68.

[33] Haracska L, Yu SL, Johnson RE, Prakash L, Prakash S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat Genet 2000;25:458–61. doi:10.1038/78169.

[34] Lee D-H, Pfeifer GP. Translesion synthesis of 7,8-dihydro-8-oxo-2'-deoxyguanosine by DNA polymerase eta in vivo. Mutat Res 2008;641:19–26. doi:10.1016/j.mrfmmm.2008.02.006.

[35] Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff RV, Walker GC. Eukaryotic translesion polymerases and their roles and regulation in

DNA damage tolerance. Microbiol Mol Biol Rev 2009;73:134–54. doi:10.1128/MMBR.00034-08.
[36] Lawrence CW. Cellular functions of DNA polymerase zeta and Rev1 protein. Adv Protein Chem 2004;69:167–203. doi:10.1016/S0065-3233(04)69006-1.
[37] Nelson JR, Lawrence CW, Hinkle DC. Deoxycytidyl transferase activity of yeast REV1 protein. Nature 1996;382:729–31. doi:10.1038/382729a0.
[38] Haracska L, Unk I, Johnson RE, Johansson E, Burgers PM, Prakash S, et al. Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. Genes & Development 2001;15:945–54. doi:10.1101/gad.882301.
[39] Zhang Y, Wu X, Rechkoblit O, Geacintov NE, Taylor J-S, Wang Z. Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Res 2002;30:1630–8. doi:10.1093/nar/30.7.1630.
[40] Washington MT, Minko IG, Johnson RE, Haracska L, Harris TM, Lloyd RS, et al. Efficient and error-free replication past a minor-groove N2-guanine adduct by the sequential action of yeast Rev1 and DNA polymerase zeta. Mol Cell Biol 2004;24:6900–6. doi:10.1128/MCB.24.16.6900-6906.2004.
[41] Morrison A, Christensen RB, Alley J, Beck AK, Bernstine EG, Lemontt JF, et al. REV3, a Saccharomyces cerevisiae gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. J Bacteriol 1989;171:5659–67.
[42] Nelson JR, Lawrence CW, Hinkle DC. Thymine-thymine dimer bypass by yeast DNA polymerase zeta. Science 1996;272:1646–9. doi:10.1126/science.272.5268.1646.
[43] Baranovskiy AG, Lada AG, Siebler HM, Zhang Y, Pavlov YI, Tahirov TH. DNA Polymerase and Switch by Sharing Accessory Subunits of DNA Polymerase. Journal of Biological Chemistry 2012;287:17281–7. doi:10.1074/jbc.M112.351122.
[44] Johnson RE, Prakash L, Prakash S. Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. Proc Natl Acad Sci U S A 2012;109:12455–60. doi:10.1073/pnas.1206052109.
[45] Makarova AV, Stodola JL, Burgers PM. A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. Nucleic Acids Res 2012;40:11618–26. doi:10.1093/nar/gks948.
[46] Lee Y-S, Gregory MT, Yang W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc Natl Acad Sci U S A 2014;111:2954–9. doi:10.1073/pnas.1324001111.
[47] Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L. Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. Nature 2000;406:1015–9. doi:10.1038/35023030.
[48] Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 2002;419:135–41. doi:10.1038/nature00991.
[49] Terai K, Abbas T, Jazaeri AA, Dutta A. CRL4Cdt2 E3 Ubiquitin Ligase Monoubiquitinates PCNA to Promote Translesion DNA Synthesis. Mol Cell 2010;37:143–9. doi:10.1016/j.molcel.2009.12.018.
[50] Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, et al. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science 2005;310:1821–4. doi:10.1126/science.1120615.
[51] Stelter P, Ulrich HD. Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature 2003;425:188–91. doi:10.1038/nature01965.
[52] Haracska L, Torres-Ramos CA, Johnson RE, Prakash S, Prakash L. Opposing Effects of Ubiquitin Conjugation and SUMO Modification of PCNA on Replicational Bypass of DNA Lesions in Saccharomyces cerevisiae. Mol Cell Biol 2004;24:4267–74. doi:10.1128/MCB.24.10.4267-4274.2004.
[53] Arakawa H, Moldovan G-L, Saribasak H, Saribasak NN, Jentsch S, Buerstedde J-M. A role for PCNA ubiquitination in immunoglobulin hypermutation. PLoS Biol 2006;4:e366. doi:10.1371/journal.pbio.0040366.
[54] Langerak P, Nygren AOH, Krijger PHL, van den Berk PCM, Jacobs H. A/T mutagenesis in hypermutated immunoglobulin genes strongly depends on PCNAK164 modification. J Exp Med 2007;204:1989–98. doi:10.1084/jem.20070902.
[55] Krijger PHL, van den Berk PCM, Wit N, Langerak P, Jansen JG, Reynaud C-A, et al. PCNA ubiquitination-independent activation of polymerase η during somatic hypermutation and DNA damage tolerance. DNA Repair 2011;10:1051–9. doi:10.1016/j.dnarep.2011.08.005.
[56] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.
[57] van Bostelen I, Tijsterman M. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light. DNA Repair 2017. doi:10.1371/journal.pbio.1001378.
[58] Gibbs PE, McGregor WG, Maher VM, Nisson P, Lawrence CW. A human homolog of the Saccharomyces cerevisiae REV3 gene, which encodes the catalytic subunit of DNA polymerase zeta. Proc Natl Acad Sci U S A 1998;95:6876–80.
[59] Lin W, Wu X, Wang Z. A full-length cDNA of hREV3 is predicted to encode DNA polymerase ζ for damage-induced mutagenesis in humans. Mutation Research/DNA Repair 1999;433:89–98. doi:10.1016/S0921-8777(98)00065-2.
[60] Morelli C, Mungall AJ, Negrini M, Barbanti-Brodano G, Croce CM. Alternative splicing, genomic structure, and fine chromosome localization of REV3L. Cytogenet Cell Genet 1998;83:18–20. doi:10.1159/000015157.
[61] Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, et al. The million mutation project: a new approach to genetics in Caenorhabditis elegans. Genome Research 2013;23:1749–62. doi:10.1101/gr.157651.113.
[62] Hendel A, Krijger PHL, Diamant N, Goren Z, Langerak P, Kim J, et al. PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLoS Genet 2011;7:e1002262. doi:10.1371/journal.pgen.1002262.
[63] Roche H, Gietz RD, Kunz BA. Specificity of the yeast rev3 delta antimutator and REV3 dependency of the mutator resulting from a defect (rad1 delta) in nucleotide excision repair. Genetics 1994;137:637–46.
[64] Jansen JG, Temviriyanukul P, Wit N, Delbos F, Reynaud C-A, Jacobs H, et al. Redundancy of mammalian Y family DNA polymerases in cellular responses to genomic DNA lesions induced by ultraviolet light. Nucleic Acids Res 2014;42:11071–82. doi:10.1093/nar/gku779.
[65] van Schendel R, van Heteren J, Welten R, Tijsterman

M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.
[66] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.
[67] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.
[68] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.
[69] van Schendel R, Roerink SF, Portegijs V, van den Heuvel S, Tijsterman M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat Comms 2015;6:7394. doi:10.1038/ncomms8394.
[70] Tomas-Roca L, Tsaalbi-Shtylik A, Jansen JG, Singh MK, Epstein JA, Altunoglu U, et al. De novo mutations in PLXND1 and REV3L cause Möbius syndrome. Nat Comms 2015;6:7199. doi:10.1038/ncomms8199.
[71] Bemark M, Khamlichi AA, Davies SL, Neuberger MS. Disruption of mouse polymerase zeta (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. Curr Biol 2000;10:1213–6. doi:10.1016/S0960-9822(00)00724-7.
[72] Gali H, Juhasz S, Morocz M, Hajdu I, Fatyol K, Szukacsov V, et al. Role of SUMO modification of human PCNA at stalled replication fork. Nucleic Acids Res 2012;40:6049–59. doi:10.1093/nar/gks256.
[73] Moldovan G-L, Dejsuphong D, Petalcorin MIR, Hofmann K, Takeda S, Boulton SJ, et al. Inhibition of homologous recombination by the PCNA-interacting protein PARI. Mol Cell 2012;45:75–86. doi:10.1016/j.molcel.2011.11.010.
[74] Boiteux S, Coste F, Castaing B. Repair of 8-oxo-7,8-dihydroguanine in prokaryotic and eukaryotic cells: Properties and biological roles of the Fpg and OGG1 DNA N-glycosylases. Free Radic Biol Med 2016. doi:10.1016/j.freeradbiomed.2016.11.042.
[75] Brauchle M, Baumer K, Gönczy P. Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in *C. elegans* embryos. Curr Biol 2003;13:819–27. doi:10.1016/S0960-9822(03)00295-1.
[76] Holway AH, Kim S-H, La Volpe A, Michael WM. Checkpoint silencing during the DNA damage response in Caenorhabditis elegansembryos. J Cell Biol 2006;172:999–1008. doi:10.1093/nar/gkh710.
[77] Jarosz DF, Beuning PJ, Cohen SE, Walker GC. Y-family DNA polymerases in Escherichia coli. Trends Microbiol 2007;15:70–7. doi:10.1016/j.tim.2006.12.004.
[78] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.
[79] Lo T-W, Pickle CS, Lin S, Ralston EJ, Gurling M, Schartner CM, et al. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics 2013;195:331–48. doi:10.1534/genetics.113.155382.
[80] Waaijers S, Portegijs V, Kerver J, Lemmens BBLG, Tijsterman M, van den Heuvel S, et al. CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. Genetics 2013;195:1187–91. doi:10.1534/genetics.113.156299.
[81] Chiu H, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. Genetics 2013;195:1167–71. doi:10.1534/genetics.113.155879.
[82] Cho SW, Lee J, Carroll D, Kim J-S, Lee J. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics 2013;195:1177–80. doi:10.1534/genetics.113.155853.
[83] Katic I, Großhans H. Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. Genetics 2013;195:1173–6. doi:10.1534/genetics.113.155754.
[84] Tzur YB, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiácovo MP. Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. Genetics 2013;195:1181–5. doi:10.1534/genetics.113.156075.
[85] Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nature Methods 2013;10:1028–34. doi:10.1038/nmeth.2641.

Supplemental information

Supplemental table 1 is only available in an electronic format and can be requested by contacting the author: ivo.vanbostelen@gmail.com

Supplemental figure 1: Protein alignments of REV3 and PCNA

CLUSTAL O(1.2.2) multiple sequence alignment

Sc = Saccharomyces cerevisiae
Ce = Caenorhabditis elegans
Dm = Drosophila melanogaster
Hs = Homo sapiens
Mm = Mus musculus

a) REV3 protein sequence

Features:
M start site of C. elegans REV-3 isoform a
M start site of C. elegans REV-3 isoform b
Q Q > Ochre stop in allele rev-3(gk919715). Q298 in isoform a, Q14 in isoform b.

```
REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      MAAAGEAIDGVYSVRLVIADFYMEKPQFGMDPCYSELRGKEIKRVPVIRVFGGNSRGQKT    60
REV3_Hs      ----------MFSVRIVTADYYMASPLQGLDTCQSPLTQAPVKKVPVVRVFGATPAGQKT    50
REV3_Mm      ----------MFSVRIVTADYYMASPLPGLDTCQSPLTQLPVKKVPVVRVFGATPAGQKT    50

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      CMHVHGVFPYLYIPYDKKDFESLERGILQMAMHLDKAINISLGQGSSNAQHVFKIQLVKG    120
REV3_Hs      CLHLHGIFPYLYVPYDGYGQQ-PESYLSQMAFSIDRALNVALGNPSSTAQHVFKVSLVSG    109
REV3_Mm      CLHLHGIFPYLYVPYDGYGQQ-PESYLSQMAFSIDRALNVALGNPSSTAQHVFKVSLVSG    109

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      IPFYGYHRVEHQFLKIYMFNPRFVRRAANLLQSGAILSKNFSPHESHVPYILQFMIDYNL    180
REV3_Hs      MPFYGYHEKERHFMKIYLYNPYTMVKRICELLQSGAIMNKFYQPHEAHIPYLLQLFIDYNL    169
REV3_Mm      MPFYGYHEKERHFMKIYLYNPAMVKRICELLQSGAIMNKCYQPHEAHIPYLLQLFIDYNL    169

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      YGMSYVHVPLEVLKFRRNHDDD--VI--------------------PYANVKQAQLLDITT    219
REV3_Hs      YGMNLINLAA--VKFRKARRKSNTLHATGSCKNHLSGNSLADTLFRWEQDEIPSSLILEG    227
REV3_Mm      YGMNLINLAA--VKFRKARRKGNASHATGLFKHQLSGNSPAGTLFRWEEDEIPSSLLLEG    227

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      AKKVACSALEVDVSSNFILNRFQLVAKSKSNHTNPGIEAIWNDEKLRRQKLVEKHTDAGD    279
REV3_Hs      VEPQSTCELEVDAVAADILNRLDIEAQ---IGGNPGLQAIWEDEKQRRRNRNETSQMSQP    284
REV3_Mm      VEPLSTCELEVDAVAADILNRLDIEAQ---IGGNPGLQAIWEDEKQRRRNRNESSQISQP    284
```

```
REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      EE-KAEAVPVLELPPTQERHQIEIAESDIF---------YRT---------ALESKLMTL    320
REV3_Hs      ESQDHRFVPATESEKKFQKRLQEILKQNDFSVTLSGSVDYSDGSQEFSAELTLHSEVLSP    344
REV3_Mm      ESQDCRFVPATESEKQFQKRLQEVLKQNDFSVTLSGSVDYSNGSQEFSAELTLHSEILSP    344

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ----------------------------EQSTLSE-----------------QTLSDQTI    335
REV3_Hs      EMLQCTPANMVEVHKDKESSKGHTRHKVEEALINEEAILNLMENSQTFQPLTQRLSESPV    404
REV3_Mm      EMLPCSPANMIEVHKDTDLSKGNTKHKVEEALINEEAILNLIENSQTFQPLTQRLSETPV    404

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      L----PQATMQTTMPG-----TKAQKRRFNL-QKLLANAVYPEECSQDQQQL--------    377
REV3_Hs      FMDSSPDEALVHLLAGLESDGYRGERNRMPSPCRSFGNNKYPQNSDDEENEPQIEKEEME    464
REV3_Mm      FMGSSPDESLVHLLAGLESDGYQGEKNRMPLPCHSFGESQNPQNSDDEENEPQIEKEEME    464

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ---LVNASFIQ-NHVTCGYSSSVSLST-----SKDESDDLDE-TVVDEELIL-SL--TQP    424
REV3_Hs      LSLVMSQRWDSNIEEHCAKKRSLCRNTHRSSTEDDDSSSGEEMEWSDNSLLLASLSIPQL    524
REV3_Mm      LSVVMSQRWDSDIEEHCAKKRSLCRNAHRSSTEEDDSSSEEEMEWTDNSLLFANLSIPQL    524

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      HGAIPH--DATLREE------------DL---------------------------------    439
REV3_Hs      DGTADENSDNPLNNENSRTHSSVIATSKLSVKPSIFHKDAATLEPSSSAKITFQCKHTSA    584
REV3_Mm      DGTADENSDNPLNNENSRAHSSVIATSKLSVRPSIFHKDAATLEPPSSAKITFQCKHTSA    584

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ---ELLDALQLLEEQN--ESESHVDLDSSLAPLSQHKKFEL--------------------    475
REV3_Hs      LSSHVLNKEDLIEDLSQTNKNTEKGLDNSVTSFTNESTYSMKYPGSLSSTVHSENSHKEN    644
REV3_Mm      LSSHVLNKDGLTEDLSQPNS-TEKGRDNS-VTFTKESTYSMKYSGSLSSTVHSDNSHKEI    642

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      --TPELLDKETAATAALFDEDVDSDEDAD---------------------------------    502
REV3_Hs      SKK-EILPV-SSCESSIFDYEEDIPSVTRQVPSRKYTNIRKIEKDSPFIHMHRHPNENTL    702
REV3_Mm      CKKDKSLPV-SSCESSVFDYEEDIPSVTRQVPSRKYSNMRKIEKDASCIHVNRHISETIL    701

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ---------------------------------------------------QETRHDFST    511
REV3_Hs      GKNSFNFSDLNHSKNKVSSEGNEKGNSTALSSLFPSSFTENCELLSCSGENRTMVHSLNS    762
REV3_Mm      GKNSFNFADLNHSKRKLSSEGNEKGNSTSLSGVFPSSLTENCDLLPSSGENRSMAHSLES    761

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      VLDDV--DELLLKLTQS------QPAESKELKASSKLPQIDGADDRLQRTPIKSISSKSK    563
REV3_Hs      TADESGLNKLKIRYEEFQEHKTEKPSLSQQAAHYMFFPSVVLS-NCLTR-PQKLSPVTYK    820
REV3_Mm      ITDESGLNKLKIRYEEFQEHKMEKPSLSQQAAHYMFFPSVVLS-NCLTR-PQKLSPVTYK    819

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      SSPSKTPT-TPIGQKSLPK-------SPR-----------TPKTSAAKKYAPLALTIGSS    604
REV3_Hs      LQPGNKPSRLKLNKRKLAGHQETSTKSSETGSTKDNFIQNNPCNSNPEKDNA----LASD    876
REV3_Mm      LQSGNKPSRLKLNKKKLIGLQETSTKSTETGATKDSCTHNDLYTGASEKENG----LSSD    875
```

```
REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      SSKKSNDEFAGRPSNPRLSLQLDQGTGTGTLRPE--------------------------    638
REV3_Hs      LTKTTRGAFENKTPT---DGFIDCHFGDGTLETEQSFGLYGNKYTLRAKRKVNYETEDSE    933
REV3_Mm      SAKATHGTFENKPPT---EHFIDCHFGDGSLEAEQSFGLYGNKYTLRAKRKVNYETEDSE    932


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ------------------------------------------------------------    638
REV3_Hs      SSFVTHNSKISLPHPMEIGESLDGTLKSRKRRKMSKKLPPVIIKYIIINRFRGRKNMLVK    993
REV3_Mm      SSFVTQNSKISLPHPMEIGENLDGTLKSRKRRKMSKKLPPVIIKYIIINRFRGRKNMLVK    992


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ------------------------------------------------------------    638
REV3_Hs      LGKIDSKEKQVILTEEKMELYKKLAPLKDFWPKVPDSPATKYPIYPLTPKKSHRRKSKHK    1053
REV3_Mm      LGKIDSKEKQVILTEEKMELYKKLAPLKDFWPKVPDSPATKYPIYPLTPKKSHRRKSKHK    1052


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      --------------------ISLRKKLAMSEMRRKSFEDSFVLLKNDCTPVRST------    672
REV3_Hs      SAKKKTGKQQRTNNENIKRTLSFRKKRSHAILS--PPSPSYNAETEDCDLNYSDVMSKLG    1111
REV3_Mm      SAKKKPGKQHRTNSENIKRTLSFRKKRTHAVLS--PPSPSYIAETEDCDLSYSDVMSKLG    1110

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ---RRSTSNLDKTHIICSLT--P---------------------------------------    690
REV3_Hs      FLSERSTSPINSSPPRCWSPTDPRAEEIMAAAEKEAMLFKGPNVYK-KTVNSRIGKTSRA    1170
REV3_Mm      FLSERSTSPINSSPPRCWSPTDPRAEEIMAAAEKESMLFKGPNVYNTKTVSPRVGKASRA    1170


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ------------------RDRNPGLSDMFETEDGKQLPPKKVVRKT--------------    718
REV3_Hs      RAQIKKSKAKLANPSIVTKKRNKRNQTNKLVDDGKKKPRAKQK--TNEKGTSRKHTTLKD    1228
REV3_Mm      RAQVKKSKARLANSSVVTNKRNKRNQTTKLVDDGKKKPRAKQKQRANEKSLSRKHAIPAD    1230


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      -----------RWSTRNQDIE----------------SLPKAGCEIERPHRS---------    743
REV3_Hs      EKIKSQSGAEVKFVLKHQNVSEFASSSGGSQLLFKQKDMPLMGSAVDHPLSASLPTGINA    1288
REV3_Mm      EKMKPHSEAE--LTPNHQSVSELTSSS-GAQALSKQKEMSQTGPAVDHPLPPAQPTGISA    1287


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ----EGSALDELKPRRSARHKVN--SANPDECSSEIQTTGPRVTTTSLDRPQKKARLS--    795
REV3_Hs      QQKLSGCFSSFLESKKSVDLQTFPSSRDDLHPSVVCNSIGPGVSKINVQRPHNQSAMFTL    1348
REV3_Mm      QQRLSNCFSSFLESKKSVDLRTFPSSRDDSHSSVVYSSIGPGISKINIQRSHNQSAMFTR    1347


REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ------------------------------------------------------------    795
REV3_Hs      KESTLIQKNIFDLSNHLSQVAQNTQISSGMSSKIEDNANNIQRNYLSSIGKLSEYRNSLE    1408
REV3_Mm      KETTLIQKSIFDLSNHLSQVAQSTQVCSGIISPKTEESSSTQKNCGSSMGKLNEYRSSLE    1407

REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ---------------------------------------------QSPKENTKT-------    804
REV3_Hs      SKLDQAYTPNFLHCKDSQQQIVCIAEQSKHSETCSPGNTASEESQMPNNCFVTSLRSPIK    1468
REV3_Mm      SKPEQVCAPNFLHCKDSQQQTVSVSEQSKTSETCSPGNAASEESQTP-NCFVTSLKSPIK    1466
```

```
REV3_Sc      ------------------------------------------------------------    0
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      --------------------------SM----------------------------NGTV    810
REV3_Hs      QIAWEQKQRGFILDMSNFKPERVKPRSLSEAISQTKALSQCKNRNVSTPSAFGEGQSGLA    1528
REV3_Mm      QIAWEQKQRGFILDMSNFKPEKVKQRSLSEAISQTKALSQCKNQNVSTPSVFGEGQSGLA    1526


REV3_Sc      ----------------------------------------------------------MSRE    4
REV3_Ce      ------------------------------------------------------------    0
REV3_Dm      ALEK-------------ATKDSSSNSESPHQQENSVSEQIEYLES----KPKKSDETARS    853
REV3_Hs      VLKELLQKRQQKAQNANTTQDPLS---NKHQPNKNISGSLEHNKANKRTRSVTSPRKPRT    1585
REV3_Mm      VLKELLQKRQQKAQSTNVVQDST----STHQPDKNISVSNEHKKANKRTRPVTSPRKPRT    1582


REV3_Sc      SNDTIQSDTVRSSSKSDYFRIQL---------NNQDYYMSKPTFL---------DPSHGE    46
REV3_Ce      ------------------------------------MNFCIRNFICEHSQEIPGPMDTVY    24
REV3_Dm      CDEKLQRELIP----QEPAGISP-G--------DSANSTEEITFSPCHDEAIESDTESDY    900
REV3_Hs      PRSTKQKEKIPKLLKVDSLNLQNSSQLDNSVSDDSPIFFSDPGFESCYSLEDSLSPEHNY    1645
REV3_Mm      PRRTKPKEQTPRRLKVDPLNLQTSGHLDNSLSDDSPILFSDPGFESCYSLEDSLSPEHNY    1642
                                                        *           .

REV3_Sc      SLPLNQFSQ-------------------------------------------------      55
REV3_Ce      NFRRNKIPVFHIY----------------------GVTDTY-------------------    43
REV3_Dm      --IVTKLRK---------------------------------------------------    907
REV3_Hs      NFDINTIGQTGFCSFYSGSQFVPADQNLPQKFLSDAVQDLFPGQAIEKNEFLSHDNQKCD    1705
REV3_Mm      NFDINTIGQTGFCSFYSGSQFVPADQNLPQKFLSDAVQDLFPGQAIDKSELLSHDRQSCS    1702
                 . :

REV3_Sc      ------------------------------------------------------------    55
REV3_Ce      ------------------------------------------------------------    43
REV3_Dm      ---TPNLKRLRWSIRSELLNKQFTPSSGIRPPETETTP----QLSPKSNESNTPELMRSF    960
REV3_Hs      EDKHHTTDSASWIRSGTLSPEIFE-KSTIDSNENRRHNQWKNSFHPLT--TRSNSIMDSF    1762
REV3_Mm      EEKHHVSDSSPWIRASTLNPELFE-KVAMDNNENHRHSQWKNSFHPLT--SHSNSIMESF    1759


REV3_Sc      ----------------------------------------------------VPNIRVFG    63
REV3_Ce      ----QKACLHIHGVL-----------------------------------------------    54
REV3_Dm      YEH-------------------SLI-------------------------VNSPSVFSDFL    977
REV3_Hs      CVQQAEDCLSEKSRLNRSSVSKEVFLSLPQPNNSDWIQGHTRKEMGQSLDSANTSFTAIL    1822
REV3_Mm      CVQQAENCLTEKSRLNRSSVSKEVFLSLPQANSSDWIQGHNRKEADQSLHSANTSFTTIL    1819


REV3_Sc      ALPTGHQVLCHVHGILPYMFIKYDGQITDTSTLRHQRCAQVHKTLEVKIRASFKRKKDDK    123
REV3_Ce      ----------------PYLVLRVGNKVTPSVLAA-M-RAKVNKGIEKEIETSTGK--T--    92
REV3_Dm      DSPEIH----------------------MDSPRSA-P-PS----------------------    993
REV3_Hs      SSPDGELVDVA-CEDLELYVSRNNDMLTPTPDSS-P-RSTS-----SP---SQSK--N--    1867
REV3_Mm      SSPDGELVDAA-SEDLELYVSRNNDVLTPTPDSS-P-RSTS-----SP---LQSK--N--    1864
                                       :       :

REV3_Sc      HDLAGDK---LGNLN-FVADVSVVKGIPFYGYHVGWNLFYKISLLNPSCLSRISELIRDG    179
REV3_Ce      NKFSVDYVY---KMESFSSRS-------LYGYQDEEEDFVRVYFSSPWYLKKATHSLGKE    142
REV3_Dm      ---PDSNSFVIAPLELPPSYDEV--VSGSRKMDIPEYEFQKPYYSNPSDVSKVT-EVGFL    1047
REV3_Hs      GSFTPRTANILKPLMSPPSREEI--MATLLDHDLSETIYQEPFCSNPSDVPEKPREIGGR    1925
REV3_Mm      GSFTPRTAHILKPLMSPPSREEI--VATLLDHDLSEAIYQEPFCSNPSDVPEKPREIGGR    1922
                      :    :             .     : .    .*  : .    :

REV3_Sc      KIFGKKFEIYESHIPYLLQWTADFNLFGC-----SWINVDRCYFRSPVLNSIL-DIDKLT    233
REV3_Ce      VIDKPLFQPYEAHLPFHLQFFIDNSIFG-MDNIYLKKVKFRMDSRDEDATLIYKDLSVGD    201
REV3_Dm      VLHIPG------NKLNDCDPFQ--SILGNDRGLASWRRRQLIAI----------------    1083
REV3_Hs      LLMVET------RLANDLAEFE--GDFS-LEGLRLWKTAFSAMTQNPR--------PGSP    1968
REV3_Mm      LLMVET------RLPNDLIEFE--GDFS-LEGLRLWKTAFSAMTQNPR--------PGSP    1965
              :          .             :.

REV3_Sc      INDDLQLLLDRF-----CDFKCNVLSRRDFPRVGNGLIEIDILPQ---FIKNREKLQHRD    285
REV3_Ce      IRN-NELLLSPYEKKTTCHVECDALVTDI---------LNMEMQADNVHSSNPGLEYLWR    251
REV3_Dm      --G-GLAMLQRHRG---EQKVREYFSTQQ---------RIAIEPAQ--LAPTWQEAKIWL    1126
REV3_Hs      LRS-GQGVVNK------GSSNSPKMVEDK---------KIVIMPCK--CAPSRQLVQVWL    2010
REV3_Mm      LRN-GQAVVNK------ESSNSHKMVEDK---------KIVIMPCK--YAPSRQLVQAWL    2007
                    ::.             :                       .
```

```
REV3_Sc      IHHDFLEKLGDISDIPVKPYVSSARDMINELTMQREELSLKEYKEPP-------ETKRHV    338
REV3_Ce      EEKERCDALG----IELKDVFNSYEPRKSTICPQERDMLRTA------------------    289
REV3_Dm      KAKELLRQREE-PKKSS-------DDIDSPI----------KIKRQ--------------    1154
REV3_Hs      QAKEEYERSKKLPKTKPTGVVKSAENFSSSVNPDDKPVVPPKMDVSPCILPTTAHTKEDV    2070
REV3_Mm      QAKEEYERSKKLPKTELTPVTKSAENVSPSLNPGDTCAVSPQVDKCPHTLSSSAHTKEEV    2067
               ::                          :

REV3_Sc      SGHQ----WQSSGEFEAFYKKAQHKTSTFDGQIPNFENFIDKNQKFSAINTPYEALPQLW    394
REV3_Ce      --RKMAKKFRQERCMSEQL-DDLMVTRIAESQQTT--------------SSVGASDTTIW    332
REV3_Dm      -KITMMLQAEE-----------------GDGGSGDEDAG-EELDCSLSLT--------PL    1187
REV3_Hs      DNSQIALQAPTTGCSQTAS-ESQMLPPVASASDPEKDED-DDDNYYISYSSPDSPVIPPW    2128
REV3_Mm      SKSQIALQTSTTGCSQTLL-------AAASAAVPEEDED-DNDNCYVSYSSPDSPGIPPW    2119
                                        ..                .

REV3_Sc      PRLPQIEINNNSMQDKKNDDQVNA-----------------SFTEYEICGVDNE--NEGV    435
REV3_Ce      EQPGEDPMEEDRIGREKTPEELDAERQKEEDREEAAEDDDNDPKNQEAEMTMFADSQKPV    392
REV3_Dm      SQAKDK-------------------------------CKATPTS----SKARETGKSRL    1211
REV3_Hs      QQPISPDSKA-----------LN------------GDDRPSSPVE-ELPSLAFENFLKPI    2164
REV3_Mm      QQAASPDFRS-----------LN------------GDDRHSSPGK-ELCSLAVENFLKPI    2155
              :  .                                        .           . :

REV3_Sc      -KGSNIKSRSYSWLPESIASPKDSTILLDHQTKYHNTINFSMDCAMTQNMASKRKLRSSV    494
REV3_Ce      ERDEAMDVD-------------------------------------------------    401
REV3_Dm      KRGTR-----------------------------------------------------    1216
REV3_Hs      KDGIQKSPCSEPQEPLVISPI-------------------------------------    2185
REV3_Mm      KDGIQKSSCSESWEPQVISPI-------------------------------------    2176


REV3_Sc      SANKTSLLSRKRKKVMAAGLRYGKR-AFVYGEPPFGYQDILNKLEDEGFPKIDYKDPFFS    553
REV3_Ce      ----------------------------------------EEVEDD-------------    407
REV3_Dm      --------------------------LS----------FIGSQDEE-------------    1226
REV3_Hs      ----------------NTRARTGKCESLCFHSTPIIQRKLLERLPEA-------------    2216
REV3_Mm      ----------------HARARTGKWDPLCLHSTPVMQRKFLEKLPEA-------------    2207
                                                     :

REV3_Sc      NPVDLENKPYAYAGKRFEISSTHVSTRIPV-------QFGGETVSVYNKPTFDMFSSWKY    606
REV3_Ce      ----EELNDKGT-GEVIEISSDGDDDQDIVQWITTKSAYDQNFISM-N----------EF    451
REV3_Dm      --------PPSSQSSEQSVSSS--AAQAELDRSSFLRQLEGS--SQ-DRQHD-----LSF    1268
REV3_Hs      ----PGLSPLSTEPKTQKLSNKKGSNTDTLRRVLLT-QAKNQFAAV-NTPQK-----ETS    2265
REV3_Mm      ----TGLSPLSVEPKTQKLYNKKGSDADGLRRVLLTTQVENQFAAV-NTPKK-----ETS    2257
                       .   .  .: .        :           .  :  :

REV3_Sc      ALKPPTYDAVQKWYNKVP---SMGNKKTESQ--ISMHTPHSKFLYKFASDVSGKQKRKKS    661
REV3_Ce      LFDAKKIENQRQFSREMTACSVMTDETVVAAAEHSQPVEQTRLLSQTSIGEPSTSEKSVE    511
REV3_Dm      GLSHATLDNTFGFKVN--------------------------------L--ENLQQAKAD    1294
REV3_Hs      QIDGPSLNNTYGFKVS--------------------------------I--QNLQEAKAL    2291
REV3_Mm      QIDGPSLNNTYGFKVS--------------------------------I--QNLQEAKAL    2283
              :.  . :    :  .                                      .: .

REV3_Sc      SVHDSLTHLTLEIHANTRSDKIPDPAIDEVSMIIWCLEEETFPLDLDI-AYEGIMIVHKA    720
REV3_Ce      SDIVGLCVASLELL----------------------VDTKMPMPDAASSEIVSVSLAIYN    549
REV3_Dm      IDCNHLTIITLEVFVSTRGDLQPDPMHDEIRCLFYAIEHSLPDEKLPS-KACGYIMVNTV    1353
REV3_Hs      HEIQNLTLISVELHARTRRDLEPDPEFDPICALFYCISSDTPLPDTEKTELTGVIVIDKD    2351
REV3_Mm      HEIQNLTLISVELHARTRRDLQPDPEFDPICALFYCISSDTPLPDTEKTELTGVIVIDKD    2343
                  *   ::*:                       :. .    .       .  :

REV3_Sc      SE---------DSTFPTKIQHCINEIPVMFYESEFEMFEALTDLVLLLDPDILSGFEIHN    771
REV3_Ce      DVCRCSSPKLHVLLTTLPVGPAYSDRRVVHMETELEMLEEVAKWIVQYDVDVMIGYETIR    609
REV3_Dm      Q-------DL-LSEGPFH--GIDRDIEVQVVTSEAEAFEALLALCERWDADIYAGYEIEM    1403
REV3_Hs      KT--VFSQDI-RYQTPLLIRSGITGLEVTYAADEKALFHEIANIIKRYDPDILLGYEIQM    2408
REV3_Mm      KT--VTHQDI-RSQTPLLIRSGITGLEVTYAADEKALFQEITNIIKRYDPDILLGYEIQM    2400
              .                          *     *   :. :       * *:  *:*

REV3_Sc      FSWGYIIERCQKIHQFDIVRELARVKCQIK----TKLSDTWGYAHSSGIMITGRHMINIW    827
REV3_Ce      LSWGFFFRRIKLL-GSRISM--DRALIDAYEDHIE-----VDDQEITVAPPKGRLLVSVW    661
REV3_Dm      SSWGYVIDRAKHL-CFNIAPLLSRVPTQKVRDFVDEDREQFTD-LDVEMKLCGRILLDVW    1461
REV3_Hs      HSWGYLLQRAAAL-SIDLCRMISRVPDDKIENRFAAERDEYGSYTMSEINIVGRITLNLW    2467
REV3_Mm      HSWGYLLQRAAAL-SVDLCQMISRVPDDKIENRFAAERDDYGSDTMSEINIVGRITLNLW    2459
              ***:.: *   :    :     *.  :                          **  :.:*
```

```
REV3_Sc      RALRSDVNLTQYTIESAAFNILHKRLPHFSFESLTNMWNAKKSTTELKTVLNYWLSRAQI    887
REV3_Ce      KVVRSDLALRNYDLGSAVANVLRKKIPMLDNAALMRRIKGERSAIRN-DVHLHLLKLSSL    720
REV3_Dm      RLMRSEIALTSYTFENVMYHILHKRCPWHTAKSLTEWFGSP--CTRW-IVMEYYLERVRG    1518
REV3_Hs      RIMRNEVALTNYTFENVSFHVLHQRFPLFTFRVLSDWFDNKTDLYRW-KMVDHYVSRVRG    2526
REV3_Mm      RIMRNEVALTNYTFENVSFHVLHQRFPLFTFRVLSDWFDNKTDLYRW-KMVDHYVSRVRG    2518
             : :*.:: * .* : ..  .:*::: *      *           .   :  : :.

REV3_Sc      NIQLLRKQDYIARNIEQARLIGIDFHSVYYRGSQFKVESFLIRICKSESFILLSPGKKDV    947
REV3_Ce      NISLLTEMNWFLKNAEMARVYGIQFHEVWTRGSQLRVESMLLRLAHRMNFVAPSITHLQR    780
REV3_Dm      TLTLLDQLDLLGRTSEMAKLIGIQFYEVLSRGSQFRVESMMLRIAKPKNLVPLSPSVQAR    1578
REV3_Hs      NLQMLEQLDLIGKTSEMARLFGIQFLHVLTRGSQYRVESMMLRIAKPMNYIPVTPSVQQR    2586
REV3_Mm      NLQMLEQLDLIGKTSEMARLFGIQFLHVLTRGSQYRVESMMLRIAKPMNYIPVTPSIQQR    2578
             .: :* : : : :. * *:: **:*  *  **** :***:::*:.:  . :  :

REV3_Sc      RKQKALECVPLVMEPESAFYKSPLIVLDFQSLYPSIMIGYNYCYSTMIGRVREINLT---    1004
REV3_Ce      NMMGSPEQLQLILEPQSKVYFDPVIVLDFQSLYPSMVIAYNYCYSTILGKIGNLVQMNDE    840
REV3_Dm      AHMRAPEYLALIMEPQSRFYADPLIVLDFQSLYPSMIIAYNYCFSTCLGRVEHLGGSS--    1636
REV3_Hs      SQMRAPQCVPLIMEPESRFYSNSVLVLDFQSLYPSIVIAYNYCFSTCLGHVENLGKYD--    2644
REV3_Mm      SQMRAPQCVPLIMEPESRFYSNSVLVLDFQSLYPSIVIAYNYCFSTCLGHVENLGKYD--    2636
                 : : : *::**:* .* . ::**********::*.****:** :*:: .:

REV3_Sc      -----ENNLGVSKFSLPRNILALLK--NDVTIAPNGVVYAKTSVRKSTLSKMLTDILDVR    1057
REV3_Ce      SRNREEIVLGAIKYHPSKDDIVKLVAYKEVCASPLASMFVKKSKREGVMPLLLREILAAR    900
REV3_Dm      -----PFEFGASQLRVSRQMLQKLLEHDLVTVSPCGVVFVKREVREGILPRMLTEILDTR    1691
REV3_Hs      -----EFKFGCTSLRVPPDLLYQVR--HDITVSPNGVAFVKPSVRKGVLPRMLEEILKTR    2697
REV3_Mm      -----EFKFGCTSLRVPPDLLYQIR--HDVTVSPNGVAFVKPSVRKGVLPRMLEEILKTR    2689
                   :*  .     : :  :   . :  :* .  :.* . *:. :  :* :** .*

REV3_Sc      VMIKKTMNEIGDDNTTLKRLLNNKQLALKLLANVTYGYTSASFSGRMPCSDLADSIVQTG    1117
REV3_Ce      IMVKSAMKRTK--NKKLKRILDARQLALKLVANVSYGYTAANWSGRMPCAELADAILGKG    958
REV3_Dm      QMVKQSMKLHK-DSSALQRILHSRQLGLKLMANVTYGYTAANFSGRMPSVEVGDSVVSKG    1750
REV3_Hs      FMVKQSMKAYK-QDRALSRMLDARQLGLKLIANVTFGYTSANFSGRMPCIEVGDSIVHKA    2756
REV3_Mm      LMVKQSMKSYK-QDRALSRMLNARQLGLKLIANVTFGYTAANFSGRMPCIEVGDSIVHKA    2748
              *:*.:*:     .  *.*:*. :**.***:***::***:*.:*****. ::.*::: ..

REV3_Sc      RETLEKAIDIIEKDETWNAKVVYGDTDSLFVYLPGKTAIEAFSIGHAMAERVTQNNPKPI    1177
REV3_Ce      RETLERSIEMVQRGDYGGAEVIYGDTDSMFVLVRGASVEEAFEIGRRIVDDVTNSNPDPV    1018
REV3_Dm      RETLERAIKLVENNEEWKVRVVYGDTDSMFVLVPGRNRAEAFRIGEEIAKAVTEMNPQPV    1810
REV3_Hs      RETLERAIKLVNDTKKWGARVVYGDTDSMFVLLKGATKEQSFKIGQEIAEAVTATNPKPV    2816
REV3_Mm      RETLERAIKLVNDTKKWGARVVYGDTDSMFVLLKGATKEQSFKIGQEIAEAVTATNPRPV    2808
             *****::*.:::  .   ..*:******:** : * .  ::* **. :.. **  ** *:

REV3_Sc      FLKFEKVYHPSILISKKRYVGFSYESPSQTLPIFDAKGIETVRRDGIPAQQKIIEKCIRL    1237
REV3_Ce      VLKLEKVYKGCVLETKKRYAGWMYEHEND-EGSLDAKGIETVRRDTCPIVAEVLEKSLGL    1077
REV3_Dm      KLKLEKVYQPCMLQTKKRYVGYMYETADQEQPVYEAKGIETVRRDGCPAVAKMLEKVLRI    1870
REV3_Hs      KLKFEKVYLPCVLQTKKRYVGYMYETLDQKDPVFDAKGIETVRRDSCPAVSKILERSLKL    2876
REV3_Mm      KLKFEKVYLPCVLQTKKRYVGYMYETLDQKEPVFDAKGIETVRRDSCPAVSKILERSLKL    2868
              **:****  .:* :****.*: **  .:     :********** *   :::*: : :

REV3_Sc      LFQTKDLSKIKKYLQNEFFKIQIGKVSAQDFCFAKEVKLGA-YKSEKTAPAGAVVVKRRI    1296
REV3_Ce      IFSQNWKTFI-TYLNTVV--LNLPQENFSKFVFCKEYRGDYSA--RAMVP-QKKIAEARI    1131
REV3_Dm      LFETQDVSKIKAYVCRQFTKLLSGRANLQDLIFAKEFRGLNGYKPTACVP-ALELTRKWM    1929
REV3_Hs      LFETRDISLIKQYVQRQCMKLLEGKASIQDFIFAKEYRGSFSYKPGACVP-ALELTRKML    2935
REV3_Mm      LFETRDISLIKQYVQRQCMKLVEGKASIQDFIFAKEYRGSFSYRPGACVP-ALELTRKML    2927
             :*. .  : *  *:      :   : . ..: *.** :          .*    :..  :

REV3_Sc      NEDHRAEPQYKERIPYLVVKGKQGQLLRERCVSPEEFLEGENLELDSEYYINKILIPPLD    1356
REV3_Ce      RKCPSHITLRGERVPYVIIDGVTGSTVYSCVRSIDQFARNPEYKINTFYYLNAHILAALR    1191
REV3_Dm      QKDPRHVPRRGERVPFIIVNGPPGMQLIRLVRSPHDILANEGHKINAIYYITKAIIPPLN    1989
REV3_Hs      TYDRRSEPQVGERVPYVIIYGTPGVPLIQLVRRPVEVLQDPTLRLNATYYITKQILPPLA    2995
REV3_Mm      AYDRRSEPRVGERVPYVIIYGTPGLPLIQLIRRPAEVLQDPTLRLNATYYITKQILPPLA    2987
                        **:*:::: *  *  :        :.      .::: **:.  ::  *

REV3_Sc      RLFNLIGINVGNWAQEIVKSKRASTTTTKVEN----ITRVGT---SATCCNCGEELTKIC    1409
REV3_Ce      RVTDLIPMKIDFL--------------------------------PFAAEQCFVSDCSRIG    1220
REV3_Dm      RCLLLIGANVHDWFASLPRKLLMTPAVGTANELAGARGAKSTISQYFSTTSCVIDCGRQT    2049
REV3_Hs      RIFSLIGIDVFSWYHELPRIHKATSS-----SRSEPEGRKGTISQYFTTLHCPV-CDDLT    3049
REV3_Mm      RIFSLIGIDVFSWYQELPRIQKATSS-----SRSELEGRKGTISQYFTTLHCPV-CDDLT    3041
             *   **  .:                                    :
```

```
REV3_Sc      SLQLCDDCLEKRSTTTLSFLIKKLKRQKEYQTLKTVCRTCSYRYTSDAGIENDHIASKCN    1469
REV3_Ce      KTPWCIECEQNPEELALALVQLGRE-SRARSQIIQYCTSCQSSSQR----LGEDQVVDCA    1275
REV3_Dm      KAGICPDCLKNATTCVVVLSDKTARLERGYQLTRQICQACCGRLG----------SLQCD    2099
REV3_Hs      QHGICSKCRSQPQHVAVILNQEIRELERQQEQLVKICKNCTGCFDR---------HIPCV    3100
REV3_Mm      QHGICSKCRSQPQHVAIILNQEIRELERKQEQLIKICRNCTGSFDR---------HIPCV    3092
             .   * .* .:    .: :     . .:  .     *  *                 *

REV3_Sc      SYDCPVFYSRVKAERYLRDNQSVQRE-EALISLNDW   1504
REV3_Ce      NFTCLLRQTMSIMDRSRTASVLT---AHKMF*----   1303
REV3_Dm      SLDCPVLYVLEGKRRELQQIEHWNKLLEHHF-----   2130
REV3_Hs      SLNCPVLFKLSRVNRELSKAPYLRQLLDQF------   3130
REV3_Mm      SLNCPVLFKLSRVNRELSKAPYLRQLLDQF------   3122
             .  * :        *            .
```

b) PCNA protein sequence

Feature:
K = K164/165>R mutation

```
POL30/PCNA_Sc      MLEAKFEEASLFKRIIDGFKDCVQLVNFQCKEDGIIAQAVDDSRVLLVSLEIGVEAFQEY 60
pcn-1_Ce           MFEAKLANAGLLKKIVESIKDLVTDAPFDCSETAMSLQAMDSSHVALVSLKLEVGLFDTY 60
PCNA_Dm            MFEARLGQATILKKILDAIKDLLNEATFDCSDSGIQLQAMDNSHVSLVSLTLRSDGFDKF 60
PCNA_Hs            MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLVQLTLRSEGFDTY 60
PCNA_Mm            MFEARLIQGSILKKVLEALKDLINEACWDVSSGGVNLQSMDSSHVSLVQLTLRSEGFDTY 60
                   *:**:: :. ::*::::.:** :  . :: .. .:  *::*.*:* **.* :    *: :

POL30/PCNA_Sc      RCDHPVTLGMDLTSLSKILRCGNNTDTLTLIADN-TPDSIILLFEDTKKDRIAEYSLKLM 119
pcn-1_Ce           RCDRTINLGLSLANMSKALKCANNDDTCMLKYEENEGDSIIFTFADPKRDKTQDVTVKMM 120
PCNA_Dm            RCDRNLSMGMNLGSMAKILKCANNEDNVTMKAQDN-ADTVTIMFESANQEKVSDYEMKLM 119
PCNA_Hs            RCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDN-ADTLALVFEAPNQEKVSDYEMKLM 119
PCNA_Mm            RCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDN-ADTLALVFEAPNQEKVSDYEMKLM 119
                   ***: : :*:.* .::* *:*. * *   :  ::   *:: : *   ::::  :  :*:*

POL30/PCNA_Sc      DIDADFLKIEELQYDSTLSLPSSEFSKIVRDLSQLSDSINIMITKETIKFVADGDIGSGS 179
pcn-1_Ce           DIDSEHLGIPDQDYAVVCEMPAGEFQKTCKDLSTFSDSLNITATKAGIVFTGKGDIGSSV 180
PCNA_Dm            NLDQEHLGIPETDFSCVVRMPAMEFARICRDLAQFSESVVICCTKEGVKFSASGDVGTAN 179
PCNA_Hs            DLDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNGN 179
PCNA_Mm            DLDVEQLGIPEQEYSCVIKMPSGEFARICRDLSHIGDAVVISCAKNGVKFSASGELGNGN 179
                   ::* : * * : ::  .  :*: ** :  :**: :.::: *  :*  : * ..*::*..

POL30/PCNA_Sc      VIIKPFVDMEHPETSIKLEMDQPVDLTFGAKYLLDIIKGSSLSDRVGIRLSSEAPALFQF 239
pcn-1_Ce           VTYSPSSNTDDETEAVTLEVKDPVNVNFSIKYMNQFTKATALSDRVRLSLCNDVPVVVEY 240
PCNA_Dm            IKLAQTGSVDKEEEAVIIEMQEPVTLTFACRYLNAFTKATPLSTQVQLSMCADVPLVVEY 239
PCNA_Hs            IKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKATPLSSTVTLSMSADVPLVVEY 239
PCNA_Mm            IKLSQTSNVDKEEEAVTIEMNEPVHLTFALRYLNFFTKATPLSPTVTLSMSADVPLVVEY 239
                   :      . :.   :: :*:.:** :.*. :*:  : *.: **  * : :. :.* :.::

POL30/PCNA_Sc      DLK-SGFLQFFLAPKFNDEE---  258
pcn-1_Ce           PIEENGYLRFYLAPKIDDDENMD  263
PCNA_Dm            AIKDLGHIRYYLAPKIEDNET--  260
PCNA_Hs            KIADMGHLKYYLAPKIEDEEGS-  261
PCNA_Mm            KIADMGHLKYYLAPKIEDEEAS-  261
                    :   *.::::****::*:*
```

Supplemental figure S2

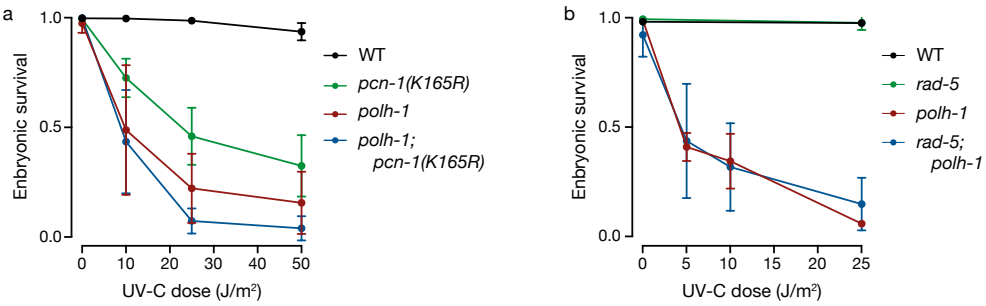


Figure S2. *a+b) Hermaphrodite adult animals of indicated genotypes were exposed to different doses of UV-C and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20 hours time period post irradiation.*

5

A novel assay to study mutation induction of interstrand crosslink repair in *C. elegans*

Ivo van Bostelen, Robin van Schendel, Dennis Brugman & Marcel Tijsterman

Abstract

Interstrand crosslinks (ICLs) are covalent bonds between bases of complementary DNA strands. ICLs are absolute blocks for DNA transcription and replication as they prevent strand separation. These lesions are not only proven cytotoxic, they also have the potential to induce mutations, which range from simple base substitutions to complex genomic rearrangements. One of the most successful systems to unravel a potentially universal mechanism of crosslink repair is the Xenopus egg extract, in which replication and repair of plasmids that contain a single DNA crosslink can be monitored at the molecular level. Ideally, this *in vitro* system is complemented by analogous *in vivo* experiments. We set out to develop such an assay, which is here presented. We demonstrate that the *C. elegans* germline can be used to monitor the repair of the ICL-containing plasmids (pICL) that were previously used in Xenopus egg extracts. We use this novel method to assay different DNA repair deficient backgrounds. We find error free repair and bypass to be affected by defects in nucleotide excision repair (XPA) and translesion synthesis (Polζ, Polη, and PCNA), while mutagenic outcomes are, in part, dependent on polymerase theta-mediated end joining. The established role of these factors in the DNA damage response argues for the validity of this new assay that provides new opportunities to study ICL repair mechanisms at the nucleotide level *in vivo*.

Introduction

Interstrand crosslinks (ICLs) are one of the most toxic types of DNA damage and are potentially very mutagenic. The crosslinking of paired DNA strands through a covalent bond between bases in opposing strands forms an absolute block for transcription and replication because it prevents strand separation. The strong toxicity of ICLs is illustrated by the effective treatment of cancer with agents that induce ICLs - especially replicating cells have great difficulty with ICLs - but also explains why nitrogen mustard has been a very powerful chemical weapon [1]. ICLs are induced by a variety of man-made chemicals such as cisplatin and nitrogen mustards. Although naturally occurring sources of ICLs are very rare they may arise from i) by-products of lipid peroxidation within the cell, specifically aldehydes, ii) the presence of abasic sites, or iii) exposure to natural psoralens [2]. When nitrogen mustard was used as the first chemotherapeutic in 1946 to treat cancer its mode of action was still completely unknown and it took decades before scientists started to understand how cells respond to ICLs and how repair of these lesions occurs [3]. Studies of the genetic disorder Fanconi anemia (FA) have provided great insight in the molecular factors that are involved in crosslink repair. Patients with this syndrome have - among other symptoms - progressive bone marrow failure and greatly increased cancer risk [4]. Moreover, FA cells are hypersensitive to ICL-inducing drugs [5]. Today, nineteen FANC genes are identified in patients. All gene products act in the same ICL repair pathway. Seven of them, FANCA, B, C, E, F, G, and L, constitute the FA core complex that has ubiquitin ligase activity by means of the FANCL subunit [6]. The core complex interacts with FANCM, a structure specific DNA binding protein, and when ICLs are recognized during S-phase the core complex ubiquitinates a complex of two other FANC proteins: FANCD2-FANCI. Ubiquitylated FANCD2/I is stabilized on DNA and promotes both checkpoint activation and the recruitment of repair factors that mediate ICL repair [7,8].

Detailed knowledge about the molecular steps that resolve crosslinks during DNA replication has come from studying the replication of plasmids that contain a single cisplatin crosslink in *Xenopus laevis* egg extracts. Studies from the Walter lab have revealed a replication-dependent mechanisms of ICL repair where two replication forks converge on a single cisplatin crosslink [9]. For this pathway it was shown that FANCD2-FANCI ubiquitination controls the incision steps necessary for ICL unhooking [10]. This generates two substrates for different repair reactions: on the one hand a gapped structure that contains a replication blocking lesion, on the other hand a DNA double strand break. While the gapped substrate requires TLS (to bypass the lesion) and NER (to remove the lesion), the DSB is repaired later via Rad51-dependent recombination [11]. Although it is not completely understood which polymerases play a role in the insertion of nucleotides directly opposite a lesion, it has become clear that in Xenopus egg extracts the TLS polymerases REV1 and Polζ collaborate to extend the nascent DNA strand after

this insertion step [9]. In addition, sequencing of repair products of replication-coupled ICL repair has shown a REV1 dependency for mutagenic products [12].

While it was first thought that this mechanism is the predominant pathway for any ICL, a recent study showed that for psoralen crosslinks and crosslinks originating from abasic sites another mechanism is more relevant, one that avoids the generation of a DSB. Here, the DNA backbone is not incised but one of the two N-glycosyl bonds that form the crosslink is cleaved by the NEIL3 DNA glycosylase. The two resulting daughter molecules have single stranded DNA gaps that are subsequently filled by REV1 mediated TLS. This pathway, which is independent of FAND2-FANCI, highlights that ICL repair is flexible and dependent on the structure of the crosslink [13]. FA factors are believed to be specific for ICL repair that occurs when DNA is being replicated and will not act outside this context. The recognition and repair of ICLs independent of replication is not fully understood but studies have pointed towards nucleotide excision repair (NER) factors and DNA mismatch repair machinery in the recognition of ICLs and in the initiation of repair. The efficiency with which ICLs are repaired outside S-phase is dependent on the level in which they distort the DNA helix structure; lesions that strongly disrupt the normal structure of the DNA helix, like cisplatin crosslinks, are more readily recognized and therefore also removed more efficiently [14]. A specific role might be reserved for XPA in the recognition or stabilization of repair intermediates, especially in cisplatin crosslinks [15-17]. TLS is important in both replication dependent and independent ICL repair but the mechanisms may differ [18,19]. Roles of REV1, Polζ and Polη are not entirely clear, but Polκ and PCNA-K164 ubiquitination functions in replication-independent repair in Xenopus egg extracts and a similar role for Polκ was confirmed in mammalian cells [20].

The cell-free model system of Xenopus has provided many valuable insights in the mechanism of ICL repair, but the generality of it is yet unclear, so are many questions yet unanswered and potentially difficult to answer *in vitro*. Here, we employ the model organism *C. elegans* and a newly designed assay to study how ICLs are repaired and mutations are induced during cell divisions and embryogenesis. We are especially interested in the dual roles of TLS in both the induction of and protection against mutation induction. The *C. elegans* model has previously allowed us to study the role of TLS in response to exogenous and endogenous DNA damage in the form of mono-adducts and UV-induced damage [21-23]. The strong conservation of DNA repair activities also make *C. elegans* well suited to study ICL repair in the context of the DNA damage response [24,25]. The mutagenicity of ICL-inducing agents has been studied in different genetic backgrounds in *C. elegans*. Exposure to mechlorethamine, which induces mostly non-bulky guanine mono-adducts but also 5-7% ICLs [26], induces a modest increase in base substitutions but a marked increase in the frequency of insertions, deletions and other structural rearrangements [27]. Exposure to cisplatin,

which mainly induces intrastrand crosslinks and a small percentage of ICLs (<5%) [28], causes a substantial increase in base substitutions, which is more pronounced in *xpf-1* mutants, as expected from a role for the encoded protein in both NER and ICL repair. Cisplatin also induce deletions (especially in the size range of 3-20 bp), insertions and other rearrangements. The deletion junction characteristics argue for the involvement of an alternative end-joining mechanism that uses microhomology to promote repair. In addition, complex rearrangements were found after exposure to cisplatin and mechlorethamine consistent with persistent replication fork stalling and subsequent DSB formation [27]. Photo-activated psoralens (UV/TMP) also generate ICLs and this agent has been widely used to generate knockouts in *C. elegans* [29,30]. In depth analysis of large numbers of UV/TMP-generated deletion alleles in *C. elegans* has shown that the DSB generated after exposure are repaired via Polθ dependent alternative end-joining [31]. Finally, exposure to the ICL-inducing agent mitomycin C (MMC) did not affect base substitutions rates but instead caused deletions. Approximately 50% MMC deletions are small (<20 bp), but the size range extends to >300Mb [32]. Most, if not all of these chemicals, that produce categorically different types of mutations do not only induce ICLs but also mono-adducts or intra-strand crosslinks that are also substrates for TLS. In order to discriminate the type of lesions that is responsible for the different mutational outcomes we developed an *in vivo* assay that monitors repair of one well-defined ICL in *C. elegans* germ cells.

In this assay, we made use of the fact that transgenes can be efficiently introduced in *C. elegans* by microinjection into the gonadal syncytium of a young adult hermaphrodite (P0 generation). Then, during early embryogenesis, multiple copies of the exogenous DNA concatenate to form a high molecular weight extrachromosomal array. A subset of the next generation (F1) will carry this array and a subset of these transgenic F1 animals will produce transgenic offspring, thus providing a stable transgenic line (TGL) [33]. We used this biology to monitor the fate of injected plasmids that contain one ICL. Transgenic animals that contain the DNA surrounding the ICL were analyzed to determine the repair footprints at the nucleotide level. Here we present the methodology and provide preliminary results obtained in different DNA damage response deficient strains.

Materials & Methods

General culturing and strains used

All strains were cultured according to standard methods as described in [34]. The N2 Bristol strain was used as WT control. The strains with alleles *rev-1(gk455794)*, *rev-3(gk919715)*, *polh-1(ok3317)*, *xpa-1(ok0698)*, *fcd-2(tm1298)* were obtained from the Caenorhabditis Genetics Center, Minnesota, USA. The *polh-1(lf0031)*, *polk-1(lf0029)* and *pcn-1(K165R)* alleles were engineered in our laboratory [21 & chapter 4 of this thesis].

Cisplatin and UV/TMP survival assays

To measure germline sensitivity to photo-activated psoralen, staged L4 animals were first treated with TMP (Sigma-Aldrich, T6137, stock: 2,5 mg/ml in acetone) at a concentration of 10 µg/ml in M9 buffer for 1 hour at RT on a rotor. Animals were then transferred to unseeded 6 cm NGM plates; ~200 animals per dose of UV-A (source: predominantly 366 nm, GE lighting F8T5 BLB U.S.A.). The irradiance of our source was determined using a Blak-ray® long wave ultraviolet meter (model: J221, ser. #12994). Measurements varied slightly between experiments with an average of 140 (±10) µW*cm-2 (equals 1,40 J*m-2*s-1). The exact dose was determined by varying the exposure time. Per dose and genotype 12 exposed animals were transferred and equally divided over four fresh OP50 seeded 6 cm NGM plates and allowed to produce offspring for 48 hours. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later the numbers of non-hatched eggs and surviving progeny were determined.

Cisplatin germline sensitivity was performed as follows. Staged L4 populations were soaked for 3 hours in M9 containing cisplatin (Accord Healthcare BV, 1 mg/ml) at indicated doses. After treatment for each dose 4 plates (6 cm, NGM seeded with OP50) with 3 L4 stage animals were allowed to produce offspring for 48 hours at 20°C. Subsequently adults were discarded and the brood on the plate was allowed to hatch. 24 hours later non-hatched eggs and surviving progeny were quantified.

Micro-injections of control, pICL and phenotypic markers

Plasmid injections were performed according to standard *C. elegans* microinjection procedures. In short: L4 animals were picked to fresh OP50 seeded NGM plates and incubated at 15°C, 20-24h pre-injection. Plasmid solutions were injected into the gonadal syncytium of the young adult animals (generation P0). pICL and the control plasmid (pCON, same sequence but without crosslink) are the plasmids described in [9] and were supplied to us by Dr. Puck Knipscheer. Two injection mixes were made. Control injection mix: 10 ng/µl pCON with phenotypic markers 100 ng/µl pRF4 (rol-6(su1006)), 10 ng/µl pGH8 (Prab-3::mCherry::unc-54utr), 2.5 ng/µl 1 pCFJ90 (Pmyo-2::mCherry::unc-54utr) and 5 ng/µl pCFJ104 (Pmyo-3::mCherry::unc-54) [mCherry plasmids are described in 35] in sterile ultrapure ddH2O, and ICL injection mix: 10 ng/µl pICL with phenotypic markers 100 ng/µl pRF4 (rol-6(su1006)), 10 ng/µl pGH8 (Prab-3::mCherry::unc-54utr), 2.5 ng/µl 1 pCFJ90 (Pmyo-2::mCherry::unc-54utr) and 5 ng/µl pCFJ104 (Pmyo-3::mCherry::unc-54) in sterile ultrapure ddH2O.

Selection of transgenic F1 and transgenic lines & making worm lysates

Progeny animals (F1 generation) that expressed mCherry were singled to new plates 3–4 days post injection and allowed reproduce for 48 hours. After generating progeny, single F1's were lysed in 15 µl SWLB (50mM KCL, 10mM Tris-HCL pH 8.3, 2.5 mM MgCl2.6H2O,

0,45% NP40, 0,45% Tween20) and heated for 60°C for 1 hour and 90°C for 15 minutes. The progeny of the F1's, the F2 generation, was screened for mCherry expression too. When F2's expressed mCherry, this was identified a transgenic line (TGL) and 5 mCherry positive F2's are lysed together in one reaction in the same way as the F1 animals.

PCR reaction on pICL, SapI digest & gel electrophoresis

In order to amplify the sequences surrounding the site of the crosslink we performed a nested PCR to achieve high specificity and yield using GoTaq© G2 DNA Polymerase (Promega). For the first (external) PCR we used forward primer ATGCCCTGGCTCACAAATAC, and reverse primer AACCGTATTACCGCCTTTGA, which will produce a product of 1048 bp in the case of repair of the ICL without insertions or deletions. For this external PCR 1.0 µl lysis (from F1 or TGL) was used in a reaction mix of 13.8 µl sqH2O, 0.8 µl Forward primer, 0.8 µl Rev primer, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega) and PCR program 3 min at 95°C; 35 x (20 sec at 95°C, 30 sec at 54°C, 1 min at 72°C); 3 min at 72°C. For internal PCR 1.0 µl PCR product from external PCR was used as substrate, with the forward primer GACATATGGGAGGGCAAATC and reverse primer AATTTGTGATGCTATTGCTTTATTTG to generate a product of 889bp. Internal PCR reaction: 13.8 µl sqH2O, 0.8 µl forward primer, 0.8 µl reverse primer, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega). The following PCR program was used for the internal PCR: 3 min. 95°C; 35 x (20 sec at 95°C, 30 sec at 54°C, 1 min at 72°C); 3 min at 72°C. Final PCR products were then analyzed by gel electrophoresis to assess presence of PCR products and the presence of larger deletions/insertions. When the crosslink was repaired error free a restriction site will be present for SapI. To analyze the level of mutation induction internal PCR products were digested by SapI (#R0569L, NE Biolabs) and analyzed by gel electrophoresis to estimate error-free repaired and resistant (mutated) fractions.

Pooling and cloning pICL PCR products

Since each F1 or TGL likely has more than one copy of pICL (or pCON) in their extrachromosomal array the PCR products needed to be cloned to be able to analyze single repair products on nucleotide level with Sanger sequencing. PCR products of approximately 20 F1's were pooled and ligated into pGEM-T following manufacturer's instructions. For the PCR products of the TGLs, 4 or 5 reactions were pooled and ligated into pGEM-T according to manufacturer's protocol. Ligation mix of each pool was used in one transformation reaction in competent *E. coli* DH5α and subsequently 90% and 10% of the reaction was cultured on two selective AXI plates (Ampicillin 100 µg/ml, X-gal 5.0g/ml and IPTG 0,5 mM) for blue/white screening.

Colony PCR

White colonies were picked into sterile ddH2O and incubated at RT for 1h on a shaker. The PCR reaction was performed using 1.0 µl from the colony-ddH2O mix. Forward primer: GTAAAACGACGGCCAG and reverse primer: CAGGAAACAGCTATGAC were used in a mix of 13.8 µl sqH2O, 2.0 µl GoTaq White 5x PCR Buffer (Promega), 2.0 µl GoTaq Green 5x PCR buffer (Promega), 0.4 µl dNTPs 10 mM and 0.2 µl GoTaq-Polymerase 5u/µl (Promega). The following PCR program was used: 3 min at 95°C; 35 x (20 sec at 95°C, 30 sec at 48°C, 1 min at 72°C); 3 min at 72°C. This generates a product of 772bp in the case of repair of the ICL without insertions or deletions. PCR products were analyzed by gel electrophoreses, to confirm proper product yield before Sanger sequencing.

Sanger sequencing and analysis of pICL repair products

For Sanger sequencing of the ICL repair products the sequence primer GTAAAACGACGGCCAG was used. Sequences were mapped to the original sequence of the pICL plasmid with the Sequence Analyzer 8 program developed in our laboratory. For detailed analysis and mapping of complex deletions with insertions to the original pICL sequence the online BLAST tool was used [36].

Results & Discussion

TLS and NER deficiency cause sensitivity towards crosslinking-inducing agents

Although potentially mutagenic, TLS of mono-adducts protects the genome from more severe genomic insults such as deletions that arise when DNA replication is permanently blocked and DNA double strand breaks arise as a consequence. Previously, we have reported on the requirement for functional TLS to maintain a stable genome in *C. elegans* and on the requirement for Polθ-dependent end-joining in case TLS is dysfunctional [21-23]. Data obtained in other systems demonstrated a role for the TLS polymerases REV1 and Polζ in replication-associated ICL repair [13,18]. Polη has been implicated because Polη deficient animals are hypersensitive to cisplatin, an ICL-inducing agent [21], and Polκ, as well as the NER factor XPA, has been suggested to act in a replication-independent mode of repair. [15-17,20].

To provide a context for studying ICL repair in *C. elegans* in molecular detail, we first tested the involvement of TLS factors by exposing young adults of different TLS deficient backgrounds to ICL-inducing agents and quantified the survival of their offspring. Following cisplatin exposure, the most pronounced sensitivity is observed for TLS mutants *polh-1* and *rev-1*: at a dose of 200 µM, which is not toxic for WT animals, we observe almost complete embryonic lethality (Fig.1). At this dose we do not observe any significant sensitivity for the other two TLS polymerase mutants *polk-1* and *rev-3* (REV-3 is the catalytic subunit of Polζ). Especially, the absence of sensitivity of *rev-3*

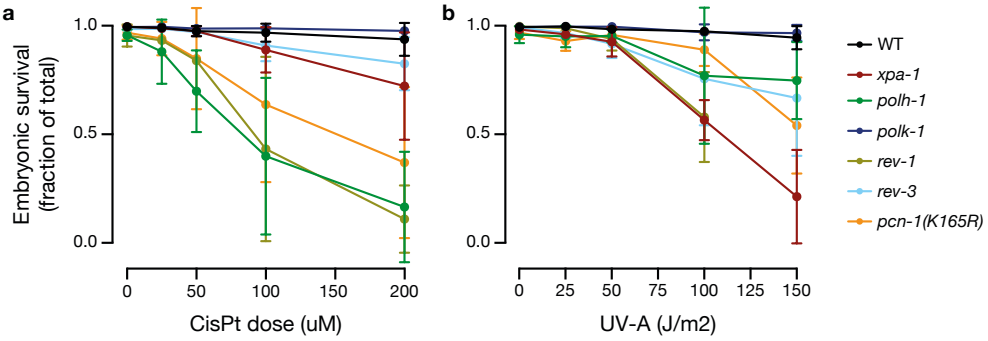


Figure 1. TLS mutants are hypersensitive to crosslinking agents. *Hermaphrodite adult animals of indicated genotypes were exposed to different doses of Cisplatin (a) or UV-TMP (b) and the embryonic survival of the progeny, as a fraction of the total brood, was determined for a 20h time period post exposure.*

mutant animals was surprising: Budzowska *et al.* have described that REV1 performs extension across cisplatin lesions yet in close collaboration with Polζ (thus REV3) [12]. While species-specific mechanistic changes can be argued, an alternative interpretation is that *C. elegans* sensitivity to crosslink-inducing agents do not necessarily read out ICL repair, but instead, translesion synthesis activity across mono-adducts or intrastrand crosslinks - cisplatin induces many more mono-adducts than ICLs: <5% of lesions are ICLs [37]. The sensitivity of *polh-1* mutants may also be seen in that context: while Polη is the most versatile of TLS polymerases and could be a good candidate for the insertion step directly opposite the unhooked crosslink in ICL repair, it likely also is the TLS polymerase to bypass mono-adducts. In keeping with this notion, we find that modification of PCNA at K165 makes worms more sensitive to cisplatin, which augments elaborate data obtained in other species that mono-ubiquitination of PCNA at K165 is a central step in activation of TLS. Yet, *pcn-1(K164R)* animals are not as sensitive to cisplatin as *polh-1* or *rev-1* mutant animals arguing that mono-ubiquination is not essential, or at least not for all lesions. We found no immediate genetic indications for profound replication-independent repair in *C. elegans* as loss of Polκ and XPA does not (or very moderately) sensitize animals to cisplatin exposure in this assay.

We next tested sensitivity of animals exposed to UV-A/TMP. UV-A/TMP induces relatively more crosslinks (*i.e.* psoralens) than cisplatin: up to 40% of all lesions [38]. Here, we did not observe any sensitivity in *polk-1* mutants, intermediate sensitivity in the *polh-1, rev-3* and *pcn-1(K165R)* single mutants and the strongest sensitivity in *xpa-1* and *rev-1* mutants. These results, that are very different from the cisplatin data make clear that animal sensitivity assays without further context is very limited with respect to providing insight into the mechanism of ICL repair. All known ICL-inducing agents also induce mono-adducts and many proteins involved in ICL repair are often

also involved in the response to those. Animal or cell sensitivity can be very useful to establish complementation groups, to categorize pathways, but to address the question on to which lesions these pathways act more specific assays are required.

A novel assay to establish mutagenicity of a single cisplatin crosslink in *C. elegans*

One of the best options to study replication and repair of ICLs is to assay synthetic plasmids that carry a single cisplatin crosslink. This has been done in cell free Xenopus egg extracts, a widely-used model system to study the dynamics of DNA replication and replication-coupled DNA repair. Here we describe how plasmids containing a single ICL can be introduced in the model organism *C. elegans* and their *in vivo* fate be monitored. Transgenes can efficiently be introduced in *C. elegans* through microinjection of DNA into the gonadal syncytium of a young adult hermaphrodite. During oogenesis the exogenously provided DNA fragments will recombine and form arrays, which, at low frequency, can start to behave as a chromosome, *i.e.* be replicated and passed on to progeny cells each cell cycle [33,39]. As a result, a subset of the F1 generation carries such an extrachromosomal array. We mixed the cisplatin crosslink-containing plasmids used in [9] with non-damaged plasmids that encode the mCherry marker and select mCherry positive transgenic F1 animals. In order for the transgenic array to be expressed in multiple cells of the animal it needs to be replicated during embryonic development. The ICLs thus need to be repaired for the host plasmid to become part of a stably transmitted extrachromosomal array. We rationalized that the extrachromosomal arrays in transgenic F1 animals would contain specific footprints of ICL repair, and because dozens of plasmid copies can be enclosed in one array, multiple ICL repair footprints can be obtained in a single transgenic F1. We injected WT animals either with a mix of pICL and mCherry marker plasmids, or with a control mix of pCON (the control plasmid that has the exact same sequence as pICL yet is without the crosslink) and the marker plasmids. All mCherry expressing F1 animals were isolated and allowed to produce offspring before their DNA was extracted. To allow for the analysis of the sequence of the ICL repair products we amplified a locus of approximately 700 nucleotides encompassing the original location of the ICL (Fig. 2b). The sequence at the site of the ICL is constructed in such a way that error-free repair results in the formation of a recognition site for the SapI restriction enzyme (Fig. 2a). This characteristic can be used to estimate the relative level of mutagenic repair at the crosslink versus the level of error-free repair. After PCR amplification the PCR products were incubated with SapI enzyme. When this protocol is performed on transgenic F1's that were injected with pCON control plasmids the PCR product of 727 bp is digested into two fragments of 383 and 344 bp (Fig. 2c, top right). Importantly, this control demonstrates that the SapI digest is almost complete: little or no uncut PCR product is visible on gel. Similar results were obtained for non-transgenic F1's from pCON-injected animals (data not shown);

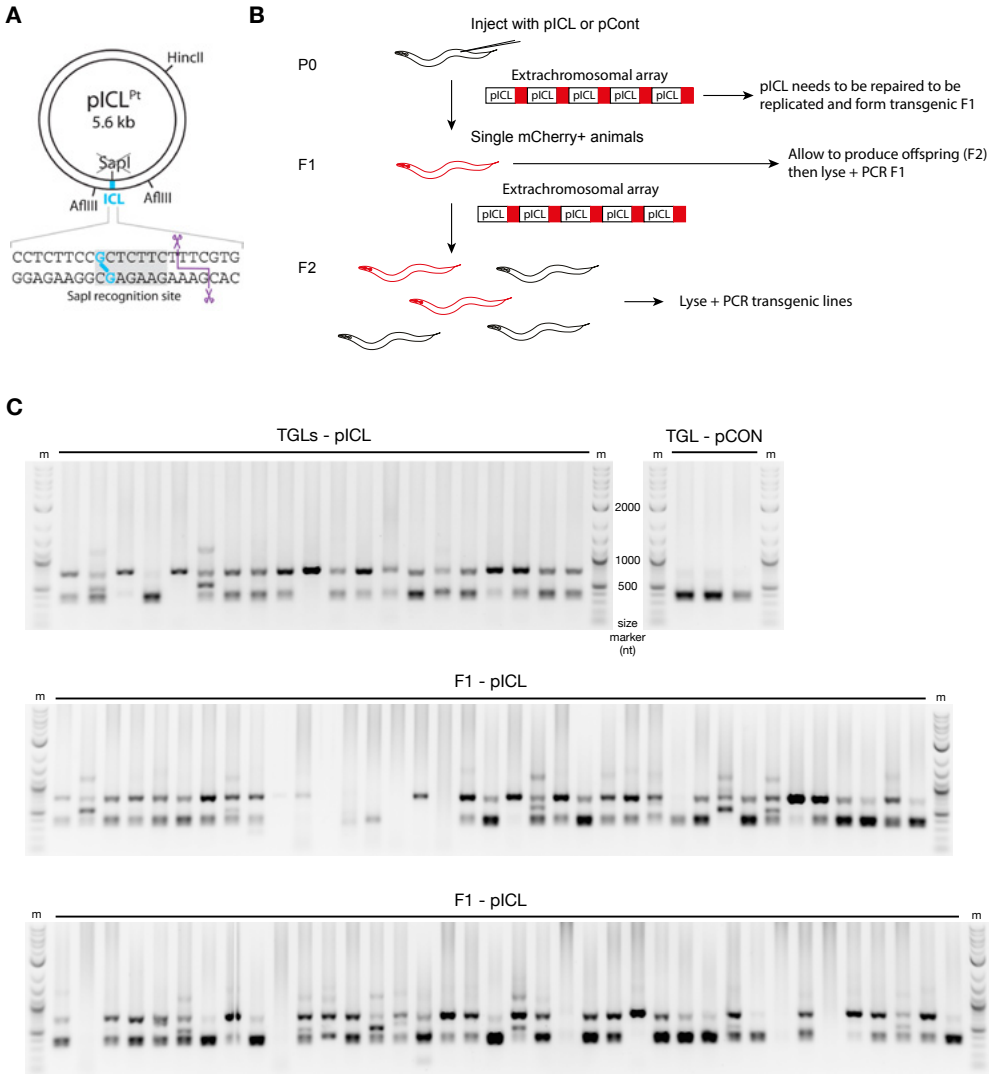


Figure 2. Experimental set-up and proof of principle. *a) Graphic representation of the pICL plasmid used in this study (taken from [10] with permission from AAAS/Science). b) Schematic of the pICL injection protocol in C. elegans. c) Gel electrophoresis images of pICL PCR products treated with SapI restriction enzyme. After injection of WT P0 animals with pICL-mCherry mix or pCON-mCherry mix F1 animals were screened. DNA was extracted from mCherry positive F1's and transgenic lines (TGLs) and used as template. pCON and error-free repaired pICL will result in a PCR product of 727 bp long that is cut into two fragments of 383 and 344 bp by SapI digestion, while mutation at the site of the crosslink result in a SapI resistant fraction of the PCR product. Larger insertions and deletions induce clearly visible changes in the size of PCR products.*

these mCherry expressing animals did not inherit the extrachromosomal array in their germline progenitor cells, and thus didn't pass the array onto subsequent generations.

Error-free repair of the pICL plasmid will also produce a PCR product (727 bp) that is susceptible to SapI digestion, however, a base pair mutation at the site of the crosslink will disrupt the SapI recognition site and prevents SapI cutting, thus resulting a digestion resistant (727 bp) fragment. The same holds for small deletions and insertions that do not induce a recognizable size change of the PCR products. Larger deletions or insertions will result in a PCR fragment of a different size yet also disrupt the SapI recognition sequence. We found that injection of the pICL mix into WT animals produced transgenic progeny indicative of both error-free and error-prone ICL repair. As expected, single transgenic lines (as well as F1's) often show multiple different repair products (Fig. 2c). Our results show that cisplatin ICLs are highly mutagenic, even in a genetic background that is fully proficient in DNA repair.

Characterization of cisplatin ICL repair products in WT animals

Having multiple ICL repair product in a single F1 precludes an immediate quantitative analysis. To solve this caveat and also to study the ICL repair products at nucleotide level we cloned the PCR products and sequenced the clones (see Table 1 and Materials & Methods for experimental details). Taking notice of potential (limited) data skewing because of preferential amplification of smaller than wildtype products, we constructed libraries of ICL repair products from transgenic F1 and transgenic lines. As expected from the data presented in figure 2, we found that in WT animals approximately half of all identified sequences were without a mutation, thus resulting from error-free repair or bypass (Fig. 3a).

		Transgenic F1 animals				Transgenic lines					
Genotype	Injected P0 animals	Transgenic animals singled	Pools prepared for cloning	Clones sequenced[1]	Repair products identified	Transgenic animals singled	Pools prepared for cloning	Clones sequenced[1]	Repair products identified	Total pools	Repair products total[2]
WT	17	107	5	120	106	21	5	120	86	10	192
fcd-2	2	43	2	48	42	4	1	24	10	3	52
xpa-1	5	109	7	168	114	19	5	120	42	12	156
polh-1	8	55	3	72	55	8	2	48	33	5	88
rev-3	8	88	4	96	71	2	TLGs did not produce offspring			4	71
rev-1	30	29	3	72	69	7	No data available			3	69
polk-1	11	25	2	48	42	3	TLGs did not produce offspring			2	42
pcn-1(K165R)	10	88	4	96	76	1	TLGs did not produce offspring			4	76

Table 1. Detailed data on ICL injections. *Detailed data on ICL injections. The table lists the number of injected animals for the indicated genotypes and the resulting F1's, transgenic lines, pools of PCR products, and properly sequenced pICL repair products. [1]Each pool was used for a pGEM-T ligation reaction and E. coli transformation. Of each plate 24 white colonies were picked into ddH2O and this was used as substrate for a colony PCR; then amplification products were sequenced. [2]These numbers are lower than the [total pools] x 24, because not each PCR and/or sequence reaction was successful.*

The mutagenic repair events were given different classifications: single nucleotide variations (SNVs) at or close to the ICL site, inserts without deletions of the original sequence, deletions smaller than 50 bp (with or without insertions), and deletions bigger than 50bp (with or without insertions). SNVs make up ~7% of all repair/bypass products. Insertions without loss of the original sequence are very rare; we only observe a single case in all 192 sequenced repair products. Deletions are much more abundant and are categorized in two distinct size ranges: ~25% of the repair products are small deletions

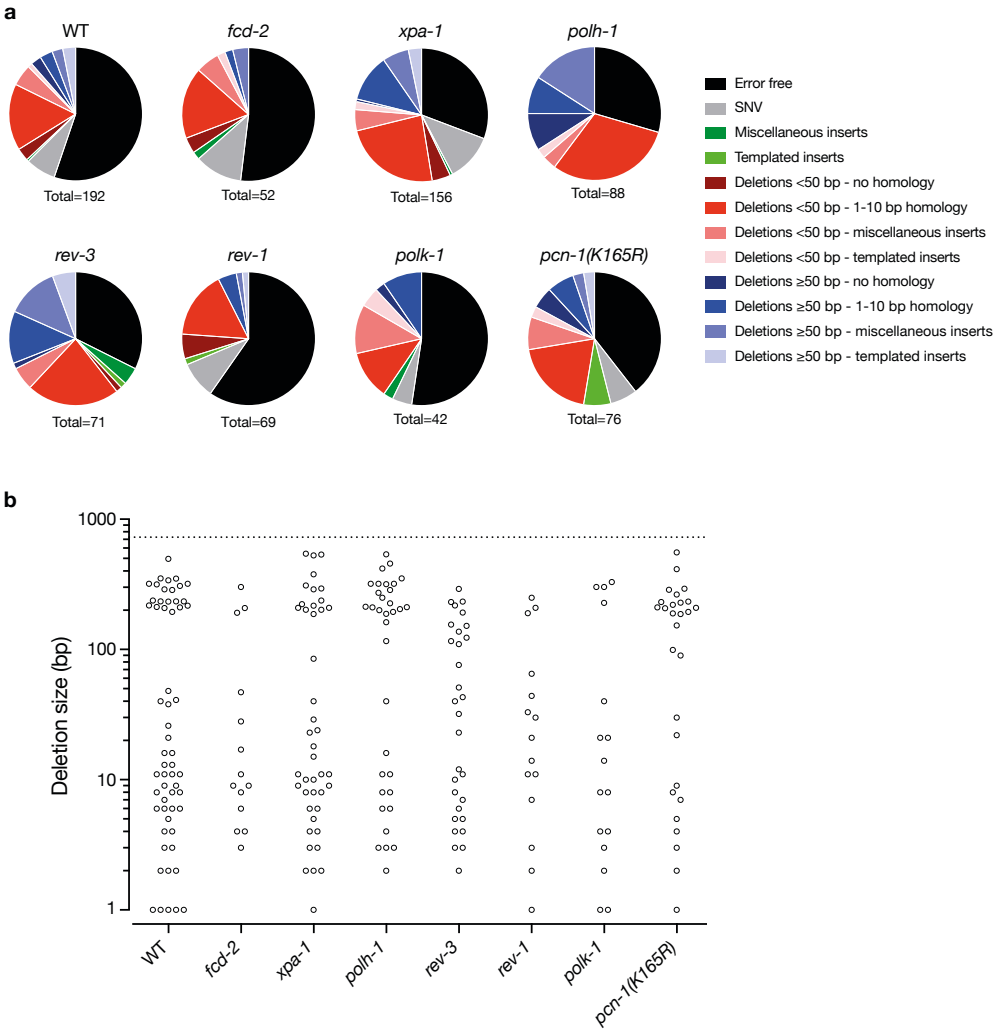


Figure 3. Mutational spectra observed in different DNA response deficient backgrounds. *a) All pICL repair products obtained after sequencing are presented for the indicated genotypes. Categories of repair products were based on observations in WT. b) Sizes of unique deletions observed in the indicated genotypes.*

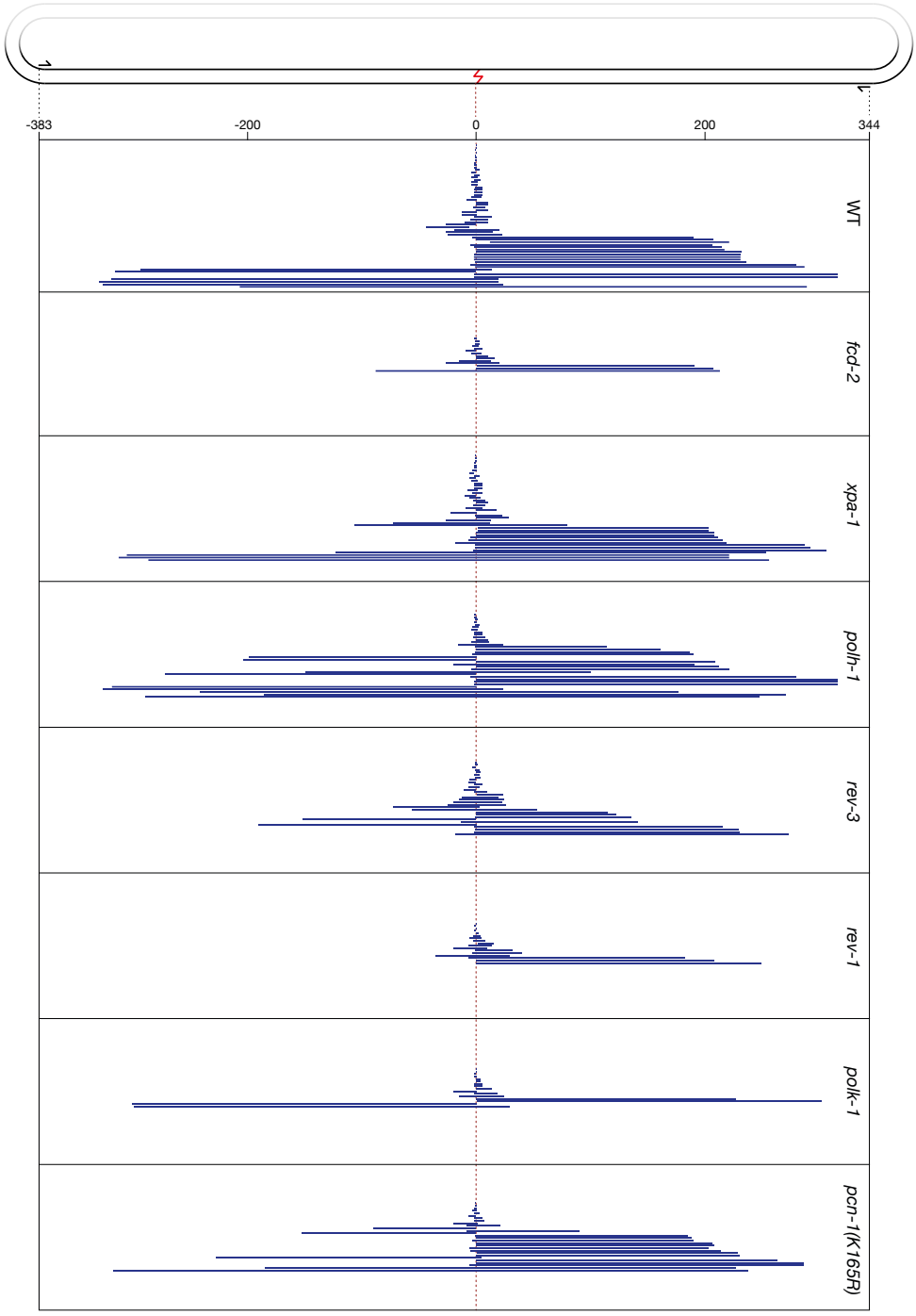
of 1-50 nt and ~11% are deletions >50 nt (Fig. 3 & 4). When studying the 1-50 deletions in further detail we observe characteristic subcategories: deletions where no homology is found at the break ends (~3%), deletions with 1-10 nt homology at the break ends (~16%), deletions with miscellaneous inserts (~3%), and deletions with inserts templated from sequences flanking the deletion (~3%). Within the collection of larger deletions (>50 nt) these same subcategories are found: deletions without homology (~3%), deletions with 1-10 nt homology (~3%), deletions with miscellaneous inserts (~3%), and deletions with inserts templated from sequences flanking the deletion (~3%).

The single nucleotide substitutions are best explained by mutagenic TLS across the unhooked crosslink. A mechanisms that has also been described in Xenopus experiments where in approximately 3% of ICL repair events SNVs are induced by mutagenic TLS [9]. We previously found that persistent replication blocks result in 50-200 bp deletions with significant micro-homology at the deletion junctions together with the occasional presence of templated inserts; a result of polymerase theta mediated end-joining (TMEJ) of replication associated DNA breaks [22,31,40,41]. Also here, we observe deletions of similar size and with other characteristics of TMEJ. This outcome could result from the inability to TLS past the unhooked crosslink. Interestingly, we also observed a high frequency of deletions that are smaller than 50 bp in size - a size category that never substantially contributes to the deletions resulting from persistent replication blocks present in one DNA strand (G-quadruplex structures, mono-adducts and UV-lesions), suggesting that these small deletions are specific for ICLs, a notion that is supported by the observation of similar deletion sizes in *C. elegans* exposed to the ICL-inducing agents mechlorethamine, cisplatin and MMC [27,32]. Interestingly, these smaller deletions also show characteristics of TMEJ. We recently found that one of the two junctions of deletions resulting from persistent replication blocks is exactly at the position of the replication block, best illustrated by deletions at G-quadruplex sites [40,41]. We here find that the vast majority of deletions originating from ICLs also have one junction immediately flanking the lesion (Figure 4), which may point towards a repair intermediate where one of the crosslinked bases block polymerase action.

Multiple DNA damage response pathways are involved in cisplatin ICL repair

The wide spectrum of repair outcomes found in WT animals suggests the involvement of multiple DNA repair pathways. We thus tested a set of genetic backgrounds deficient in different aspects of the DNA damage response: TLS *(polh-1, polk-1, rev-1, rev-3* and *pcn-1(K165R))*; NER *(xpa-1)*; and ICL repair (*fcd-2*). Experimental details for the different

Figure 4. Unique deletions in relation to the location of the crosslink. *On the top a schematic representation of the pICL plasmid with primer sites indicated by small arrows. Unique deletions isolated from the indicated genotypes are mapped to the pICL sequence. Each bar represents a deletion and the red dotted line indicates the location of the crosslink.*


1
2
-383
-200
0
200
344
WT
fcd-2
xpa-1
polh-1
rev-3
rev-1
polk-1
pcn-1(K165R)

strains can be found in Table 1 and in the section Material & Methods. All the factors analyzed did not affect transgenic array formation in the absence of the ICL lesion (Robin van Schendel, M.T. unpublished observation).

We observed a significant shift towards mutagenic repair in strains that lack the TLS polymerases Polζ *(rev-3*) and Polη (*polh-1*) and in animals deficient for the NER protein XPA-1 (Table 2), arguing that both NER and TLS acts to repair ICL in this context. Together with the broad spectrum of outcomes in wild type, this result validates the assay we here present as a means to study ICL repair *in vivo* at the nucleotide level. Below, we will briefly discuss the preliminary data derived in the different genetic backgrounds. While these results are preliminary, they can serve as a platform to guide future research.

FANCD2

Although *C. elegans fcd-2* mutants are sensitive to ICL-inducing agents [25,31], the loss of this ICL repair factor does not have a significant effect on the mutagenicity and the kind of mutations that occur at a well-defined ICL in exogenously provided plasmids (Table 2, Fig. 3 & 4). At present we do not have an explanation for this outcome apart from suggesting that the type of ICL-repair we assay is independent of the Fanconi anemia pathway.

XPA

We find that deficiency for the NER factor XPA (*xpa-1*) results in a statistically significant shift towards more mutagenic repair in comparison with error-free repair (Table 2). Especially deletion frequencies are increased (Fig. 3 & 4). These findings support a role of XPA in error-free repair of ICL repair in *C. elegans*. One explanation for our observation is that loss of or incomplete replication-independent ICL repair leads to more replication blocks that subsequently cause the formation of DNA breaks and deletions. Although there are many indications that multiple NER proteins are involved in ICL repair, their role in this pathway is not fully understood [17]. XPA probably plays a role in damage recognition or unwinding of the DNA around the ICL because it binds junctions between single- and double stranded DNA [15], and it seems to do this especially in replication-independent ICL repair [16]. In addition, NER may play a role in the removal of the unhooked crosslink that remains after TLS has progressed.

Translesion synthesis

We did not observe any effect when *polk-1* was mutated (Table 2, Fig, 3 & 4), as could be expected from the results of the cisplatin and UV-TMP sensitivity experiments presented in Figure 1. Interestingly, the ICL repair spectrum derived from *rev-1* mutant animals is also indistinguishable from that derived from wild type animals (Table 2, Fig, 3 & 4), while these mutant animals are highly sensitive to ICL-inducing

	Repair products		Chi-square test	
Genotype	Error-free (n)	Mutagenic (n)	p-values	Significantly diffent from WT (p<0.0014)
WT	106	86	-	-
fcd-2	27	25	0,633790794	FALSE
xpa-1	48	108	8,34203E-10	TRUE
polh-1	26	62	1,2911E-06	TRUE
rev-3	23	48	0,000110767	TRUE
rev-1	40	27	0,459551655	FALSE
polk-1	22	20	0,712519073	FALSE
pcn-1(K165R)	30	46	0,005807866	FALSE

Table 2. Statistical analyses of error-free and mutagenic repair in different genotypes as compared to WT. *The numbers of error-free repair vs. mutagenic repair products in all genetic backgrounds were used to determine statistic differences of the indicated genetic backgrounds. P-values were calculated with Chi-square tests. In order to control for repeated testing Bonferroni correction was performed on the original α=0.01: seven different genotypes were each compared to WT leading to a new α=0,01/7=0.0014. A genotype is significantly different from WT when p<α.*

agents (Fig. 1). It may be that REV-1 is especially important for TLS of mono-adducts that are induced upon exposure to these agents. It is also possible that one of the other TLS polymerases can take over, resulting in less efficient TLS and high cytotoxicity, but no substantial changes in mutation induction. Our *in vivo* findings are different from observations in Xenopus egg extracts where it was shown that the induction of SNVs close to the ICL is dependent on REV1 [12]. Animals that contain the *pcn-1(K165R)* allele had a similar mutation profile as WT animals (Table 2, Fig. 3 & 4), while being more sensitive to ICL-inducing agents. Mono-ubiquitination of PCNA is a general activation signal for TLS and therefore the effect may be similar as loss of REV-1, and specifically acting on mono-adducts or reduced TLS (causing cytotoxicity), but no change in mutagenicity.

The most striking changes are observed in *polh-1* and *rev-3* mutant animals (Table 2, Fig. 3 & 4). A complete lack of SNVs and a substantial increased large deletion frequency is observed in both these mutants. Thus, contrary to REV1, a role of Polζ appears to be conserved between *C. elegans* and Xenopus [12]. Our findings suggest a role for Polζ and Polη TLS polymerases in the mutagenic bypass of unhooked crosslinks. Our previous research demonstrated that loss of TLS causes persistent replications blocks that in turn lead to DSBs. Repair of these breaks produces deletions with a TMEJ footprint and this is also what we observe here [22,23]. Specifically, the increased frequency of 50-200 deletions is in line with this notion. In such a scenario, replication at the unhooked crosslink is completely blocked ultimately leading to deletion formation at such sites.

Conclusions & Perspectives

ICL-inducing agents are the most widely-used class of chemotherapeutic. They are so effective because DNA crosslinks are specifically hard to tolerate in rapidly cycling cells. However, despite intense research efforts over the last two decades, how cells act to repair these harmful lesions is still incompletely understood. Studying the cellular responses to ICLs may help to develop novel and improved chemotherapeutics. In addition, cancer cells can adopt to the presence of ICLs, for example by upregulating DNA damage response pathways leading to chemoresistance, a major hurdle in the treatment of cancer. Previous work has firmly established a role for TLS, FA and HR pathways working together in ICL repair, thereby contributing to DNA damage tolerance [Reviewed in 42]. With the *in vivo* assay we present here, we wished to expand the toolbox for research into the molecular repair processes of DNA interstrand crosslinks, and by comparing the outcomes to NGS data of animals exposed to ICL-inducing agents (R. van Schendel and M.T. unpublished results) provide more clarity into substrate specificity of different repair mechanisms.

The here-described assay still requires further development. For instance, to monitor reduced array formation we can co-inject an undamaged plasmid that is identical to the ICL-containing plasmid yet carrying a limited number of SNVs, as well as titrating the number of repair events to less than one per animal - at this moment we fail to monitor drop-out effects where mutant animals may produce less transgenic F1's, or F1's carry relatively low number of copies of the pICL plasmid in the transgenic array. Another valuable development is NGS of PCR products of large pools of animals to provide stronger statistical power.

Future and ongoing work is directed to expand on the analysis here presented, also studying plasmids that carry a psoralen crosslink [13]. Taking the latest advances in genome editing in consideration it may become possible to insert small pieces of DNA that carry a crosslink or other lesion of choice directly into the genome as has recently been done in *E. coli* [43].

In conclusion, we here present a novel assay to study the repair and mutagenesis of a single crosslinks at a known site in the nematode *C. elegans* in which many ICL repair factors are conserved [24,25]. With this assay the *in vitro* studies performed in Xenopus egg extract which have provided a detailed model for ICL repair can be paralleled, and other questions specific to an *in vivo* context can be addressed.

Acknowledgements

Several strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Programs . This work is supported by the European Research Council (ERC) and the Netherlands Organization for Scientific Research (NWO). We are grateful to Dr. Puck Knipscheer for providing us with the pICL and pCON plasmids.

References

[1] Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer 2011;11:467–80. doi:10.1038/nrc3088.

[2] Scharer OD. DNA Interstrand Crosslinks: Natural and Drug-Induced DNA Adducts that Induce Unique Cellular Responses. ChemBioChem 2005.

[3] Goodman LS, Wintrobe MM, Dameshek W, Goodman MJ, GILMAN A, McLennan MT. Nitrogen mustard therapy: Use of Methyl-Bis(Beta-Chloroethyl) amine Hydrochloride and Tris(Beta-Chloroethyl) amine Hydrochloride for Hodgkin's Disease, Lymphosarcoma, Leukemia and Certain Allied and Miscellaneous Disorders. Journal of the American Medical Association 1946;132:126–32. doi:10.1001/jama.1946.02870380008004.

[4] Nalepa G, Clapp DW. Fanconi anaemia and cancer: an intricate relationship. Nat Rev Cancer 2018;18:168–85. doi:10.1038/nrc.2017.116.

[5] Niedernhofer LJ, Lalai AS, Hoeijmakers JHJ. Fanconi anemia (cross)linked to DNA repair. Cell 2005;123:1191–8. doi:10.1016/j.cell.2005.12.009.

[6] Ceccaldi R, Sarangi P, D'Andrea AD. The Fanconi anaemia pathway: new players and new functions. Nature Reviews Molecular Cell Biology 2016;17:337–49. doi:10.1038/nrm.2016.48.

[7] Walden H, Deans AJ. The Fanconi anemia DNA repair pathway: structural and functional insights into a complex disorder. Annu Rev Biophys 2014;43:257–78. doi:10.1146/annurev-biophys-051013-022737.

[8] Muniandy PA, Liu J, Majumdar A, Liu S-T, Seidman MM. DNA interstrand crosslink repair in mammalian cells: step by step. Crit Rev Biochem Mol Biol 2010;45:23–49. doi:10.3109/10409230903501819.

[9] Räschle M, Knipscheer P, Knipsheer P, Enoiu M, Angelov T, Sun J, et al. Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 2008;134:969–80. doi:10.1016/j.cell.2008.08.030.

[10] Knipscheer P, Räschle M, Smogorzewska A, Enoiu M, Ho TV, Schärer OD, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science 2009;326:1698–701. doi:10.1126/science.1182372.

[11] Long DT, Räschle M, Joukov V, Walter JC. Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science 2011;333:84–7. doi:10.1126/science.1204258.

[12] Budzowska M, Graham TGW, Sobeck A, Waga S, Walter JC. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. Embo J 2015;34:1971–85. doi:10.15252/embj.201490878.

[13] Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. Cell 2016;167:498–511.e14. doi:10.1016/j.cell.2016.09.008.

[14] Williams HL, Gottesman ME, Gautier J. The differences between ICL repair during and outside of S phase. Trends in Biochemical Sciences 2013;38:386–93. doi:10.1016/j.tibs.2013.05.004.

[15] Camenisch U, Dip R, Schumacher SB, Schuler B, Naegeli H. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. Nat Struct Mol Biol 2006;13:278–84. doi:10.1038/nsmb1061.

[16] Shen X, Do H, Li Y, Chung W-H, Tomasz M, de Winter JP, et al. Recruitment of fanconi anemia and breast cancer proteins to DNA damage sites is differentially governed by replication. Mol Cell 2009;35:716–23. doi:10.1016/j.molcel.2009.06.034.

[17] Wood RD. Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ Mol Mutagen 2010;51:520–6. doi:10.1002/em.20569.

[18] Sharma S, Canman CE. REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. Environ Mol Mutagen 2012;53:725–40. doi:10.1002/em.21736.

[19] Roy U, Schärer OD. Involvement of translesion synthesis DNA polymerases in DNA interstrand crosslink repair. DNA Repair 2016;44:33–41. doi:10.1016/j.dnarep.2016.05.004.

[20] Williams HL, Gottesman ME, Gautier J. Replication-independent repair of DNA interstrand crosslinks. Mol Cell 2012;47:140–7. doi:10.1016/j.molcel.2012.05.001.

[21] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.

[22] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.

[23] van Bostelen I, Tijsterman M. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light. DNA Repair 2017. doi:10.1371/journal.pbio.1001378.

[24] Youds JL, Barber LJ, Boulton SJ. *C. elegans*: a model of Fanconi anemia and ICL repair. Mutat Res 2009;668:103–16. doi:10.1016/j.mrfmmm.2008.11.007.

[25] Collis SJ, Barber LJ, Ward JD, Martin JS, Boulton SJ. *C. elegans* FANCD2 responds to replication stress and functions in interstrand cross-link repair. DNA Repair 2006;5:1398–406. doi:10.1016/j.dnarep.2006.06.010.

[26] Povirk LF, Shuker DE. DNA damage and mutagenesis induced by nitrogen mustards. Mutat Res 1994;318:205–26.

[27] Meier B, Cooke SL, Weiss J, Bailly AP, Alexandrov LB, Marshall J, et al. *C. elegans*whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. Genome Research 2014;24:1624–36. doi:10.1101/gr.175547.114.

[28] Lemaire MA, Schwartz A, Rahmouni AR, Leng M. Interstrand cross-links are preferentially formed at the d(GC) sites in the reaction between cis-diamminedichloroplatinum (II) and DNA. Proc Natl Acad Sci U S A 1991;88:1982–5. doi:10.1073/pnas.1718185115.

[29] Yandell MD, Edgar LG, Wood WB. Trimethylpsoralen induces small deletion mutations in Caenorhabditis elegans. Proc Natl Acad Sci U S A 1994;91:1381–5. doi:10.1073/pnas.91.4.1381.

[30] Barstead RJ, Moerman DG. *C. elegans* deletion mutant screening. Methods Mol Biol 2006;351:51–8. doi:10.1385/1-59745-151-7:51.

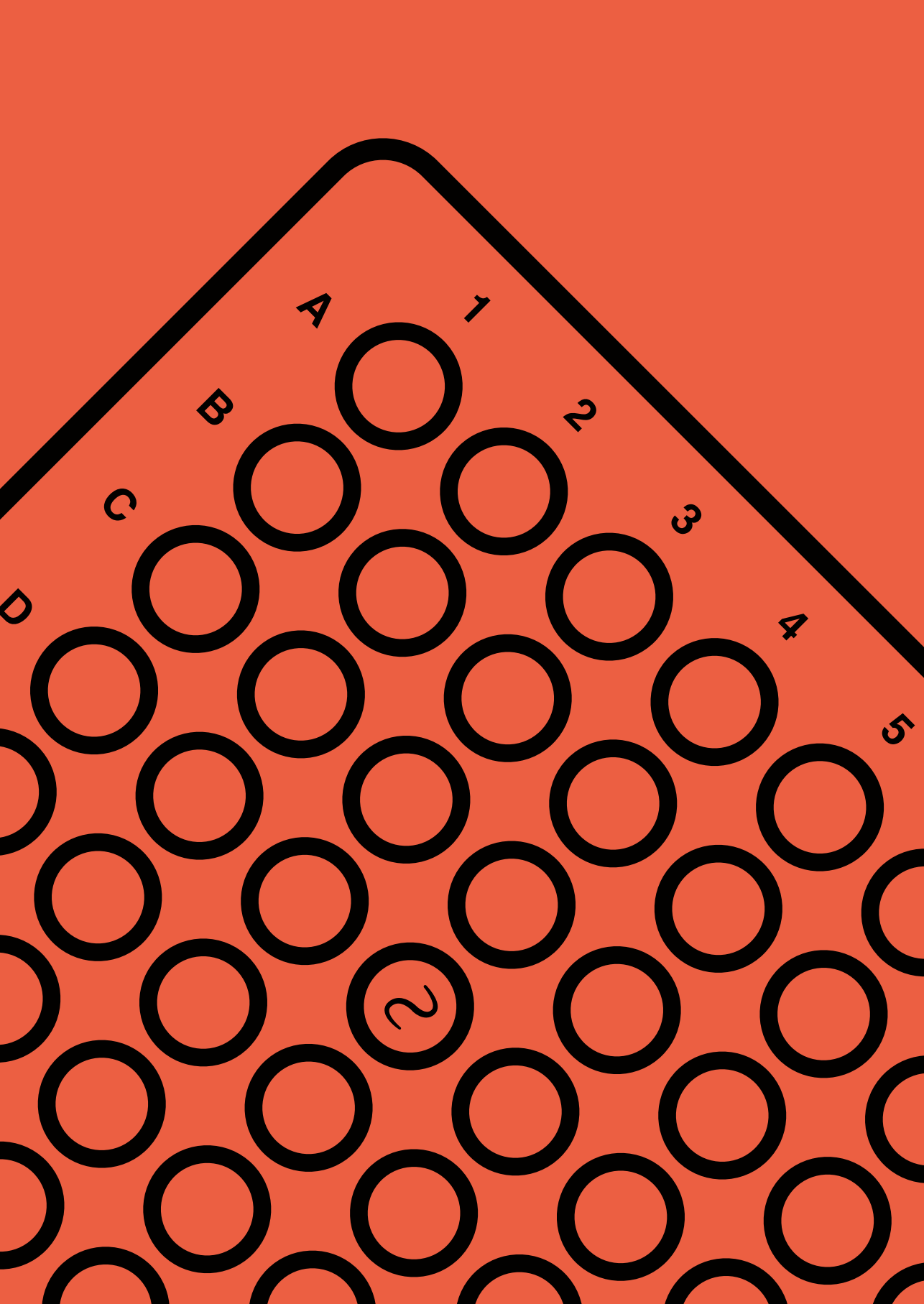
[31] van Schendel R, van Heteren J, Welten R, Tijsterman M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.

[32] Tam AS, Chu JSC, Rose AM. Genome-Wide Mutational Signature of the Chemotherapeutic Agent Mitomycin C in Caenorhabditis elegans. G3 (Bethesda) 2015;6:133–40. doi:10.1534/g3.115.021915.

[33] Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. Extrachromosomal DNA transformation of Caenorhabditis elegans. Mol Cell Biol 1985;5:3484–96.

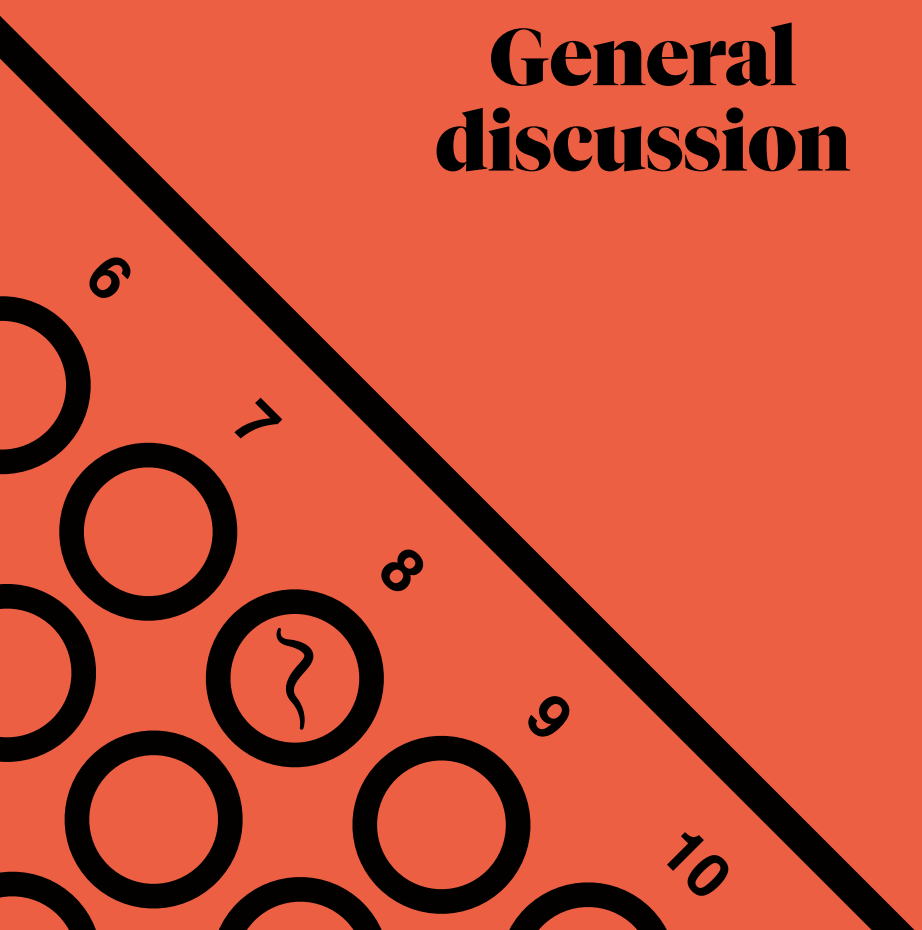
[34] Brenner S. The genetics of Caenorhabditis elegans. Genetics 1974;77:71–94.

[35] Frøkjær-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, et al. Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet 2008;40:1375–83. doi:10.1038/ng.248.
[36] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10. doi:10.1016/S0022-2836(05)80360-2.
[37] Eastman A. Reevaluation of interaction of cis-dichloro (ethylenediamine) platinum (II) with DNA. Biochemistry 1986. doi:10.1021/bi00361a026.
[38] Dronkert ML, Kanaar R. Repair of DNA interstrand cross-links. Mutat Res 2001;486:217–47. doi:10.1016/S0921-8777(01)00092-1.
[39] Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. Embo J 1991;10:3959–70.
[40] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.
[41] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.
[42] Haynes B, Saadat N, Myung B, Shekhar MPV. Crosstalk between translesion synthesis, Fanconi anemia network, and homologous recombination repair pathways in interstrand DNA crosslink repair and development of chemoresistance. Mutat Res Rev Mutat Res 2015;763:258–66. doi:10.1016/j.mrrev.2014.11.005.
[43] Pagès V, Fuchs RP. Inserting Site-Specific DNA Lesions into Whole Genomes. Methods Mol Biol 2018;1672:107–18. doi:10.1007/978-1-4939-7306-4_9.


A
1
B
2
C
3
D
4
5
2

6

General discussion



The work presented in this thesis revolves around the question: *What is the contribution of translesion DNA synthesis to genome stability?* I have approached this question from different angles. Chapter 2 shows how translesion DNA synthesis (TLS) protects genome stability, together with nucleotide excision repair (NER) and polymerase theta-mediated end joining (TMEJ), in the face of the continuous genotoxic insults of natural light. Chapter 3 and 4 show to which extent the different TLS polymerases counteract spontaneous mutagenesis and that TMEJ acts to repair DSBs that arise from persistent replication blocks in TLS compromised animals. Finally, Chapter 5 presents a novel assay to study ICL repair *in vivo*, which was applied to answer the question: *To what extend does TLS contribute to error-free or mutagenic ICL repair?* As with all research, answers give rise to more questions. Next, I will discuss the open ends of my research and some options for future endeavors.

The model organism *C. elegans*

All research presented in this thesis employs the model organism *C. elegans*; a fairly novel tool to study genome stability and DNA damage response. Over recent years, I helped to develop and have used a micro-evolution assay, which is utilizing the availability of genetic knockouts, rapid embryonic development and generational turnover, and relatively compact genome of *C. elegans*, in combination with whole genome sequencing (WGS) and bioinformatics, to study the genome wide accumulation of spontaneous mutations over many generations in an unbiased way. A similar approach was first performed in wild type animals [1,2], and over the last years others and I have applied it to a variety of mutants with defects in different DNA damage response pathways [3-5]. Due to the high resolution of *C. elegans* WGS data, a whole range of genomic alterations can be obtained, ranging from single base substitutions to complex genomic rearrangements, as well as the frequencies with which these mutations arise. As the approach does not rely on exogenous sources of DNA damage, it assays spontaneous mutagenesis, *e.g.* the mutagenic effect of endogenous DNA lesions, arguably the most relevant type of DNA damage organisms need to deal with. The vast majority of genomic rearrangements that were observed in TLS compromised animals are 50–500 deletions with occasional templated inserts and frequent use of microhomology, the signature genomic scar that is left after DSB repair by TMEJ. Although rare, also more complex rearrangements were found in TLS compromised mutants, which form by (a) yet unknown mechanism(s). Up until now the total number of complex rearrangements that were captured is small, but the increasing amount of mutation accumulation data in TLS deficient strains may provide clues toward identifying this mechanism in the future.

A recent technical advance that has been of great benefit to genetic research is the development of CRISPR/Cas9 that allows for targeted and efficient genome editing, which I used in chapter 4, were I generated a specific missense mutation in the gene

encoding the *C. elegans* PCNA ortholog. Employing CRIPSR/Cas9 to engineer any mutant of choice and assaying these mutants with the established mutation-accumulation pipeline is a powerful method to study any DNA damage response factors' effect on mutagenesis.

For the development of the crosslink assay (Ch. 5) I exploited the possibility to introduce foreign DNA to *C. elegans* via micro-injection, and with this assay a new technique to the *C. elegans* toolbox was added. The assay, however, still needs further development and optimization. On the short term, also for validating purposes, other known ICL repair mutants will need to be tested, PolQ mutants being specifically interesting as it was shown that PolQ was required to repair UV/TMP-induced DNA damage [6]; these experiments are currently ongoing. Theoretically, any DNA lesion of interest can be integrated into a similar plasmid and assayed using this pipeline, for example G-quadruplex structures (recently tested in Xenopus egg extracts [7]), or psoralen crosslinks.

Delineating DNA repair and damage tolerance in a model organism such as *C. elegans* is not a principal goal. I employed the many advantages of this model organism in order to, via extrapolation, shed light on the molecular mechanism acting in human cells. Nevertheless, *C. elegans* has some limitations that leave several questions unanswered. Firstly, not all TLS factors are conserved - the *C. elegans* genome does not encode Polι and also the *C. elegans* REV-1 and REV-3 proteins differ significantly when compared to their human orthologs. In higher eukaryotes additional functions are described for REV1 (in homologous recombination and replication through G-quadruplex sequences), which could be reserved for vertebrate REV1 and dependent on C-terminal parts of the protein that are lacking in yeast and *C. elegans* [8-11]. Furthermore, the data I present in chapter 2 and 4 suggest that 8oxoG is a frequent DNA lesion that requires TLS for bypass to protect replication potential. In higher eukaryotes this lesion is repaired by OGG1 dependent base excision repair (BER) [12], but the *C. elegans* genome does not encode OGG1. This raises the question how the tradeoff between OGG1-dependent repair and bypass is in other organisms, and may suggest that *C. elegans* have a greater dependence on TLS than organisms that have OGG1 to repair 8oxoG lesions.

Indeed, several studies have shown that TLS is especially important during *C. elegans* embryogenesis, perhaps because embryos are extremely sensitive to cell cycle delays and thus need a quick fix for stalled replication [13-15]. During this developmental stage HR intermediates are toxic, and also TMEJ may act to suppress them. A substantial part of the mutations that accumulate over generations possibly forms during embryonic development. As an alternative DNA damage tolerance pathway unicellular bacteria and yeast efficiently employ template switching (TS), but so far there is no evidence that such a pathway exists in *C. elegans* (the presence of this pathway in other multicellular eukaryotes is also under debate). This all adds up to the notion that *C. elegans* relies

on TLS as the primary *modus operandi* to deal with replication blocking DNA damage, and this model system thus provides an excellent opportunity to study the effect of these molecular mechanisms that are also employed by human cells on genome stability. For example, future genetic studies using *C. elegans* as a model could focus on the contribution of TLS on evolution, or it may aid in the search for factors that have synthetic lethal interactions with TLS.

Regulation of TLS & polymerase selection

Because TLS polymerases have very low fidelity compared to the replicative DNA polymerases their activity needs tight control. On the other hand, TLS needs to happen quickly, which requires rapid and efficient recruitment to sites of base damage of the specific TLS polymerase that can bypass the replication impediment. It is a generally held conception that different TLS polymerase are proficient in bypassing specific lesions: Polη is especially proficient at bypassing UV-induced CPDs, for example. However, it is not yet clear how the appropriate TLS polymerase is selected. Bypass may be a simple trial and error mechanism, or, there may be yet-unknown sensing mechanisms in play. Clearly, there are still many unknowns and the regulation of TLS remains an interesting research topic for future study.

A well-studied and established TLS regulation pathway is the one acting through ubiquitination of PCNA at residue K164. Using CRISPR/Cas9, I engineered a mutation in the equivalent residue of the *C. elegans* PCNA encoding gene resulting a PCNA-K165R allele, and subsequently corroborated what has been described in other organisms: PCNA ubiquitination is important for TLS activation but not essential. The remaining TLS activity in mammalian cells may be dependent on a non-catalytic function of REV-1, similar to what has been found in DT40 chicken cells [16,17]. The data presented in chapter 3 also argues for a regulatory role for REV-1 in *C. elegans*, and it would thus be interesting to study TLS activity in a *rev-1; pcn1(K165R)* double mutant. Likely, there are several more regulatory mechanisms that contribute to TLS. It is now possible to induce specific amino acid changes in factors that are predicted to function in these mechanisms using CRISPR/Cas9, and subsequently use these mutants in mutation accumulation analysis. In order to match specific lesions to specific TLS polymerases, the assay we describe in chapter 5 may be of use in future research. Our preliminary data already suggests that TLS factors Polζ, Polη, and PCNA are important during the bypass step in cisplatin crosslink repair. Similar plasmids can be generated with other lesions to investigate lesion specific requirements. A reason why this is of special interest is that many chemotherapies induce replication-blocking DNA lesions, and therefore TLS may play an important role in acquiring chemoresistance. Another promising direction for future research will be to identify small molecule inhibitors of TLS. It is suggested that cancer cells have upregulated TLS to deal with increased replication

stress. Certain cancers may also rely more on TLS because of deficiencies in DNA repair pathways. A deeper understanding of the characteristics of TLS factors, the regulation of TLS, and lesion-specific requirements may support identifying efficient new cancer therapies. Several studies have already shown promising results in this regard [18-21].

TLS is anti-mutagenic

The first studies of TLS in bacteria and yeast identified TLS as a mutagenic process: in TLS deficient cells background levels of mutagenesis and UV-induced mutagenesis are reduced. Initially, the chief benefit of TLS was thought to be guarding replication potential, and the tradeoff was increased mutagenesis. Indeed, TLS polymerases are much less accurate then replicative DNA polymerases and the bypassed base lesions can be hard to read due to the alteration of the chemical structure of the base. Finally, TLS polymerases lack proofreading. Thus, three aspects contribute to nucleotide substitutions at damaged bases. Yeast and bacteria employ an alternative error-free DNA damage tolerance pathway that uses the newly synthesized undamaged sister chromatid to replicate past a lesion, generally referred to as template switching (TS). However, in multicellular eukaryotes this mechanism does not seem to be widely used and TLS appears to be the principal damage tolerance pathway. By now there are many studies that challenge the dogma that TLS is *grosso modo* mutagenic and actually show that - in addition to guarding replication potential - TLS benefits genome stability. This is exemplified by the increased cancer risk and sensitivity to sunlight of *Xeroderma pigmentosum* variant (XP-V) patients that have a mutation in the Polη encoding gene [22-24]. Even when mutations are induced by TLS - as we show for REV-1 dependent TLS in chapter 3 - the net result is anti-mutagenic because the formation of much more detrimental 50-500 bp deletions and complex genomic rearrangements is prevented. Also, a lot of TLS activity on specific lesions is error-free, as is exemplified by the accurate bypass of thymine-thymine dimers and cisplatin lesions [25,26].

One important unanswered question is: *What determines the size of the deletions that result from persistent replication blocks?* The work presented in this thesis and in other recent studies shows that the breaks that form in TLS compromised *C. elegans* are repaired through an end-joining mechanism mediated by polymerase theta [4]. DNA breaks at persistent replication blocking G-quadruplex structures are also repaired by TMEJ and result in deletions of the same size (50-500 bp) [3]. However, TMEJ also acts in repair at sites of transposon-induced DNA breaks, CRIPSR/Cas9-induced DSBs, and during ICL repair, and in all these cases the resulting deletions are significantly smaller [5,6]. This suggests that the steps that lead to the formation of the DSB determine the size of the deletion, and not the break repair mechanism TMEJ. In fact, another recently published study demonstrated that single unresolved G-quadruplexed DNA structures can persist through multiple mitoses. Unsuccessful replication across a G-quadruplex

causes single-strand DNA gaps that result in DSBs in subsequent cell divisions [27]. Because the mutagenic outcome in animals that cannot resolve G-quadruplexes are very similar to that observed in TLS defective mutants, it is highly likely that the same mechanism acts there. If so, the size of the deletion is determined by the size of the single-strand gap at the site of the damaged base that remains after the first round of replication. In lagging strand replication, the size and deposition of Okazaki fragments may be determining, or conversely, during leading strand replication repriming 50-500 bp downstream of the lesion may occur. It could also be that the replication fork permanently stalls and another fork can approach from the opposite site, arresting 50-500 bp from the lesion and the initially blocked fork. Ongoing work is looking into which of these (or possibly other) molecular mechanisms determine the size of deletions at persistent replication blocks.

Concluding, the dual functions of TLS in guarding both replication potential and genome stability explain the strong conservation of this pathway throughout all known life forms. This work supports the notion that TLS strongly suppresses genomic instability, as it prevents replication blocks conversion into DSBs that are especially mutagenic. At high levels DSBs resulting from blocked replication or other sources will result in cell death or senescence because essential cell cycle checkpoints are activated. However, under unperturbed conditions these rare replication-associated DNA breaks are unlikely to activate cell cycle checkpoints or induce senescence but will be repaired by mutagenic DSB repair pathways, resulting in mutations that are then transmitted to subsequent generations. Translesion DNA synthesis is beneficial to the health of individual organisms and to species and life in general because it aids ongoing proliferation and growth eventhough DNA damage is ever-present and unavoidable, and in the face of these continuous threats to genomic integrity, TLS guards against the formation of the mutations that otherwise result in cancer, ageing and inborn disease.

References

[1] Denver DR, Dolan PC, Wilhelm LJ, Sung W, Lucas-Lledó JI, Howe DK, et al. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci USA 2009;106:16310–4. doi:10.1073/pnas.0904895106.

[2] Weber KP, De S, Kozarewa I, Turner DJ, Babu MM, de Bono M. Whole genome sequencing highlights genetic changes associated with laboratory domestication of *C. elegans*. PLoS ONE 2010;5:e13922. doi:10.1371/journal.pone.0013922.

[3] Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Comms 2014;5:3216. doi:10.1038/ncomms4216.

[4] Roerink SF, van Schendel R, Tijsterman M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. Genome Research 2014;24:954–62. doi:10.1101/gr.170431.113.

[5] van Schendel R, Roerink SF, Portegijs V, van den Heuvel S, Tijsterman M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat Comms 2015;6:7394. doi:10.1038/ncomms8394.

[6] van Schendel R, van Heteren J, Welten R, Tijsterman M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining. PLoS Genet 2016;12:e1006368. doi:10.1371/journal.pgen.1006368.

[7] Castillo Bosch P, Segura-Bayona S, Koole W, van Heteren JT, Dewar JM, Tijsterman M, et al. FANCJ promotes DNA synthesis through G-quadruplex structures. Embo J 2014;33:2521–33. doi:10.15252/embj.201488663.

[8] Sarkies P, Reams C, Simpson LJ, Sale JE. Epigenetic instability due to defective replication of structured DNA. Mol Cell 2010;40:703–13. doi:10.1016/j.molcel.2010.11.009.

[9] Schiavone D, Guilbaud G, Murat P, Papadopoulou C, Sarkies P, Prioleau M-N, et al. Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. Embo J 2014;33:2507–20. doi:10.15252/embj.201488398.

[10] Okada T, Sonoda E, Yoshimura M, Kawano Y, Saya H, Kohzaki M, et al. Multiple Roles of Vertebrate REV Genes in DNA Repair and Recombination. Mol Cell Biol 2005;25:6103–11. doi:10.1128/MCB.25.14.6103-6111.2005.

[11] Sharma S, Hicks JK, Chute CL, Brennan JR, Ahn J-Y, Glover TW, et al. REV1 and polymerase ζ facilitate homologous recombination repair. Nucleic Acids Res 2012;40:682–91. doi:10.1093/nar/gkr769.

[12] Boiteux S, Coste F, Castaing B. Repair of 8-oxo-7,8-dihydroguanine in prokaryotic and eukaryotic cells: Properties and biological roles of the Fpg and OGG1 DNA N-glycosylases. Free Radic Biol Med 2016. doi:10.1016/j.freeradbiomed.2016.11.042.

[13] Roerink SF, Koole W, Stapel LC, Romeijn RJ, Tijsterman M. A Broad Requirement for TLS Polymerases η and κ, and Interacting Sumoylation and Nuclear Pore Proteins, in Lesion Bypass during *C. elegans* Embryogenesis. PLoS Genet 2012;8:e1002800. doi:10.1371/journal.pgen.1002800.

[14] Brauchle M, Baumer K, Gönczy P. Differential activation of the DNA replication checkpoint contributes to asynchrony of cell division in *C. elegans* embryos. Curr Biol 2003;13:819–27. doi:10.1016/S0960-9822(03)00295-1.

[15] Holway AH, Kim S-H, La Volpe A, Michael WM. Checkpoint silencing during the DNA damage response in Caenorhabditis elegansembryos. J Cell Biol 2006;172:999–1008. doi:10.1093/nar/gkh710.

[16] Edmunds CE, Simpson LJ, Sale JE. PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Mol Cell 2008;30:519–29. doi:10.1016/j.molcel.2008.03.024.

[17] Arakawa H, Moldovan G-L, Saribasak H, Saribasak NN, Jentsch S, Buerstedde J-M. A role for PCNA ubiquitination in immunoglobulin hypermutation. PLoS Biol 2006;4:e366. doi:10.1371/journal.pbio.0040366.

[18] Choi J-S, Kim S, Motea E, Berdis A. Inhibiting translesion DNA synthesis as an approach to combat drug resistance to DNA damaging agents. Oncotarget 2017;8:40804–16. doi:10.18632/oncotarget.17254.

[19] Choi J-S, Kim CS, Berdis A. Inhibition of Translesion DNA Synthesis as a Novel Therapeutic Strategy to Treat Brain Cancer. Cancer Res 2018;78:1083–96. doi:10.1158/0008-5472.CAN-17-2464.

[20] Yamanaka K, Chatterjee N, Hemann MT, Walker GC. Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? PLoS Genet 2017;13:e1006842. doi:10.1371/journal.pgen.1006842.

[21] Zafar MK, Maddukuri L, Ketkar A, Penthala NR, Reed MR, Eddy S, et al. A Small-Molecule Inhibitor of Human DNA Polymerase η Potentiates the Effects of Cisplatin in Tumor Cells. Biochemistry 2018;57:1262–73. doi:10.1021/acs.biochem.7b01176.

[22] Johnson RE, Kondratick CM, Prakash S, Prakash L. hRAD30 mutations in the variant form of xeroderma pigmentosum. Science 1999;285:263–5. doi:10.1126/science.285.5425.263.

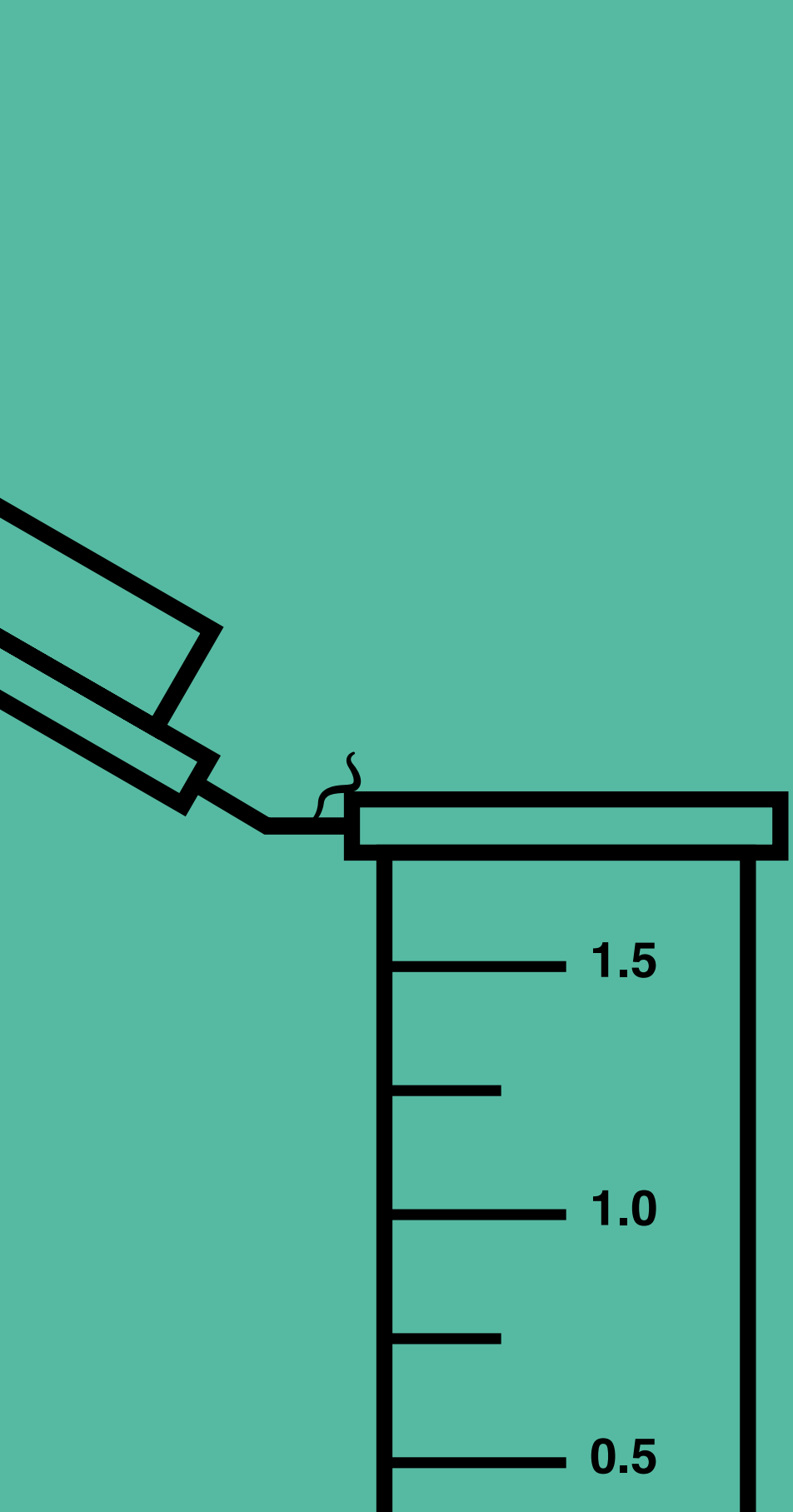
[23] McDonald JP, Levine AS, Woodgate R. The Saccharomyces cerevisiae RAD30 Gene, a Homologue of Escherichia coli dinB and umuC, Is DNA Damage Inducible and Functions in a Novel Error-Free Postreplication Repair Mechanism. Genetics 1997;147:1557–68.

[24] De Palma A, Morren M-A, Ged C, Pouvelle C, Taïeb A, Aoufouchi S, et al. Diagnosis of Xeroderma pigmentosum variant in a young patient with two novel mutations in the POLH gene. Am J Med Genet A 2017. doi:10.1002/ajmg.a.38340.

[25] Johnson RE, Prakash S, Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 1999;283:1001–4.

[26] Budzowska M, Graham TGW, Sobeck A, Waga S, Walter JC. Regulation of the Rev1-pol ζ complex during bypass of a DNA interstrand cross-link. Embo J 2015;34:1971–85. doi:10.15252/embj.201490878.

[27] Lemmens B, van Schendel R, Tijsterman M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. Nat Comms 2015;6:8909. doi:10.1038/ncomms9909.


1.5
1.0
0.5

Addendum

Summary

DNA damage is unavoidable due to the intrinsic chemical reactivity of DNA and environmental factors. These insults to the structure of DNA drive the accumulation of mutations, and are thereby causative to pathologies like cancer, ageing and inborn genetic disease. It is therefore of no surprise that organisms evolved a broad range of molecular pathways with great mechanistic variety to counteract the deleterious effects of a multitude of DNA lesions, that range from chromosomal breaks to small chemical alterations to the structure of the bases of DNA. One subset of damage has had our keenest interest in this thesis: lesions that have the potential to form roadblocks for one of life's essential molecular mechanisms: DNA replication. Loss of efficient and faithful DNA replication leads to genomic instability, reduced proliferative ability, and even cell death. One key molecular pathway that guards replication potential is translesion synthesis (TLS). TLS employs specialized TLS polymerases that are able to synthesize DNA opposite damaged DNA bases that cannot be traversed by the replicative DNA polymerases. By doing so, TLS allows for ongoing DNA replication in the face of DNA damage. Considering that TLS polymerases have a much lower fidelity than the replicative polymerases, and the fact that damaged template bases can be hard to read (because normal base paring is affected) it seems likely that this is a mutagenic process, which is still a widely held notion. However, in this thesis we demonstrate that the net effect of TLS is anti-mutagenic, because TLS prevents the formation of replication-associated DNA breaks and ensuing genomic rearrangements.

In the general introduction in **chapter 1** I discuss the relevant DNA damage response pathways, provide a brief history of genetic research techniques, and introduce the model organism *C. elegans* that is used in all experiments described in this thesis. The nematode *C. elegans* is still a rather novel tool to study genome stability and DNA damage response. I have helped to develop and used a micro-evolution assay, which utilizes the wide accessibility of genetic knockouts, the fast growth and compact genome of *C. elegans*, and whole genome sequencing (WGS) technologies supported by bioinformatics. In chapter 3 and 4, we used this assay to study the genome-wide accumulation of spontaneous mutations over many generations in an unbiased way.

In the first experimental chapter of this thesis, **chapter 2**, I study TLS within the context of the DNA damage response (DDR). The DDR is a complex cellular reaction to genomic insult, composed of signaling pathways and DNA repair networks. I describe a surprisingly simple linear order of events for how *C. elegans* deals with DNA lesions that are induced by light. Separately inactivating nucleotide excision repair (NER) or TLS induces sensitivity to UV-light, but loss of repair of replication associated DNA breaks (alt-EJ) does not. However, an exorbitant phenomenon emerges when these three pathways are inactivated in combination in an animal: *C. elegans* mutants that lack NER, TLS, and alt-EJ grow at normal rates in the dark, but cannot proliferate

when grown in plain light and become sterile. Even very low levels of normal daylight already suppress growth in these triple mutant animals. I conclude that NER and TLS operate to suppress the formation of lethal DNA breaks that require polymerase theta-mediated end joining for their repair. My observations in this chapter also demonstrate the enormous genotoxicity of light, and why multiple layers of protection are needed against such a ubiquitous environmental threat to the integrity of the DNA.

In **chapter 3** I describe the contribution of the Y-family polymerase REV-1 on long-term stability of an animal genome. I have found that REV1 both stimulates and suppresses spontaneous mutagenesis during unperturbed propagation. Because REV-1 stimulates translesion synthesis, it prevents the persistence of single strand DNA gaps that are converted to small deletions by alternative end joining of ensuing double-strand breaks. My findings oppose the current dogma, as the action of REV-1 during unperturbed growth is predominantly anti-mutagenic: REV1 prevents the accumulation of detrimental deletions at the cost of single nucleotide substitutions. I also show that only 1 lesion per ~10^{10} bases, requires REV1 action. I conclude that ensuring replication progression outweighs near-perfect conservation of genetic information in animal cells.

The notion that TLS is not accurate has mainly been established in systems that are exposed to genotoxins because the mutation rate resulting from spontaneous DNA damage is generally too low to be monitored under laboratory conditions. **Chapter 4** builds on chapter 3 and describes the combined roles of all TLS polymerases in *C. elegans* in maintaining genetic stability. I discuss genome wide mutation accumulation data from lines of TLS proficient and deficient animals that were grown under unchallenged conditions, and for each genetic background I provide comprehensive mutation profiles that range from simple base substitutions to complex genomic rearrangements. Previous work is complemented with mutants without functional Polζ and with loss of PCNA ubiquitination at residue K165, an important TLS-activating modification of PCNA. Also, mutations were combined to disrupt redundancy in TLS activities: I crossed animals to combine deficiency for all Y-family polymerases, and for the first time created an animal without any TLS activity. Extraordinarily, these strains could still proliferate and TLS is apparently not essential for animal life. Nonetheless, TLS deficient mutants have increased genomic instability. In addition, the large dataset allowed me to determine that most endogenous replication blocks form at guanines, which in the absence of TLS lead to DSBs that are repaired via TMEJ.

Chapter 5 is the last experimental chapter of this thesis. There, I present a novel *in vivo* assay to study repair of a specific type of DNA lesions: interstrand crosslinks (ICLs). ICLs are covalent bonds between bases of complementary DNA strands and present absolute blocks for DNA transcription and replication. They are highly cytotoxic and have the potential to induce varying types of mutations, ranging from simple base

substitutions to large and complex genomic rearrangements. *Xenopus* egg extract have been, and are being used extensively and successfully to unravel the mechanisms of crosslink repair. We set out to develop an *in vivo* assay in *C. elegans* to complement such *in vitro* assays. In this chapter, I demonstrate that the *C. elegans* germline can be used to monitor the repair of the ICL-containing plasmids that were previously used in Xenopus egg extracts. I then use this novel method to assay different DNA repair deficient backgrounds and find error free repair and bypass to be affected by defects in nucleotide excision repair (XPA) and translesion synthesis (Polζ, Polη, and PCNA), while mutagenic outcomes are partially dependent on TMEJ. These factors have known roles in the DNA damage response, which argues for the validity of this new methodology. With this assay I thus contributed a new opportunity to study ICL repair *in vivo*.

In the general discussion in **chapter 6** I have placed my findings in a broader perspective and I describe which important questions are still unanswered and may be of interest for future research. I finally conclude that the strong evolutionary conservation of TLS is explained by the dual functions of TLS: sustaining replication potential and guarding genome stability. Translesion DNA synthesis is beneficial to the health of individual organisms, species, and life in general, since it supports continuous reproduction and growth. Although DNA damage is always present and unavoidable, TLS guards against the formation of mutations that would otherwise lead to cancer, aging and congenital disease.

Nederlandse samenvatting

De studies die ik in dit proefschrift beschrijf hebben een centraal thema: *Translaesie DNA-synthese* (TLS). TLS is een verdedigingsmechanisme dat in alle bekende levende organismen actief is en het beschermt tegen diverse negatieve gevolgen van beschadigingen aan het genetisch materiaal, het DNA. In al mijn experimenten heb ik gebruik gemaakt van het modelorganisme *C. elegans*. Omdat niet iedereen ingewijd is in moleculair-biologische onderwerpen als DNA & cellen, DNA-schade & herstel, en modelorganismen zoals de nematode *C. elegans*, zal ik hier eerst een beknopte introductie[1] geven, gevolgd door een samenvatting van de door mij verkregen inzichten.

Cellen & DNA

Een volwassen mens is opgebouwd uit ongeveer $37x10^{12}$ cellen die zijn ontstaan zijn uit een enkele bevruchte eicel. 37.000 miljard cellen dus, met een grote verscheidenheid: hersencellen, spiercellen, stamcellen, etc. Elk weefsel is opgebouwd uit specifieke cellen en alle weefsels samen vormen een mens, en uiteraard geldt praktisch hetzelfde voor goudvissen, brandnetels, wolven, champignons, wormen en elke andere meercellige op aarde. Zoals ons lichaam organen bevat, zo bevat een cel organellen met verschillende functies. Er zijn bijvoorbeeld organellen die zorgen voor de energievoorziening van de cel en anderen nemen de eiwitproductie voor hun rekening. Het organel wat voor mij het meest interessant is, is de celkern: een compartiment dat het DNA bevat, het materiaal waarop de instructies gecodeerd staan zijn om een organisme te bouwen en te onderhouden.

Een DNA-molecuul bestaat uit twee complementaire strengen die de bekende "dubbele helix" vormen. De bouwstenen van het DNA zijn vier nucleotiden: Adenine (A), Thymine (T), Guanine (G) en Cytosine (C) - waarbij A altijd met T paart, en G altijd met C. Celkernen van levende organismen bevatten enorm lange DNA-moleculen in de vorm van chromosomen. De wonderlijke organisatie hiervan wordt duidelijk als je beseft dat al het DNA van één cel achter elkaar wel 1 á 2 meter lang is, terwijl een celkern een gemiddelde doorsnede heeft van slechts 0,005 millimeter. Elke menselijke cel heeft 23 paar chromosomen en elk paar bestaat uit één chromosoom van de vader en één van de moeder. Alle 46 chromosomen samen vormen het menselijk genoom. Sinds ongeveer 40 jaar is de wetenschap in staat om de volgorde van nucleotiden in DNA te bepalen. Dit noemen we DNA-sequencen en de reeks van A's, T's, G's en C's noemen we een sequentie. In 2003 is er een imposante en belangrijke mijlpaal bereikt toen de sequentie van het volledige menselijke genoom werd gepubliceerd. Mede hierdoor weten we ondertussen dat de sequentie van het totale menselijke genoom ongeveer 20.000 genen bevat, en elk gen bevat de instructie voor de productie van een eiwit. Eiwitten, ook wel proteïnes genoemd, zijn grote moleculen die uiteenlopende essentiële functies vervullen in ons lichaam. Er zijn bijvoorbeeld enzymen die zorgen voor de vertering van ons voedsel;

[1] Voor een gedetailleerde wetenschappelijke Engelstalige introductie verwijs ik graag naar Hoofdstuk 1: *General Introduction.*

er zijn proteïnes die structuur aan onze cellen en ons lichaam geven; en er zijn weer anderen die een belangrijke rol spelen als hormoon, in de communicatie tussen cellen op afstand.

DNA-schade, herstel & mutaties

Om van een enkele bevruchte eicel uit te kunnen groeien tot een volwassen mens moeten cellen delen en voordat een cel kan delen dient het volledige genetische materiaal gedupliceerd te worden. Dit proces noemen we DNA-replicatie en het wordt uitgevoerd door specifieke eiwitten die we polymerases noemen. Bij DNA-replicatie wordt de dubbele helix van het DNA opengeritst zodat er twee enkele strengen ontstaan. Vervolgens synthetiseert een polymerase een nieuwe streng complementair aan de bestaande streng. Tegenover een A wordt altijd een T ingebouwd en omgekeerd, tegenover een G wordt altijd een C ingebouwd en *vice versa*. De replicatieve polymerases in onze cellen zijn erg efficiënt en ontzettend accuraat: de meeste schattingen zeggen dat slechts 1 per 1.000.000 nucleotiden verkeerd wordt ingebouwd. Echter, bij een genoom van 6,4 miljard nucleotiden resulteert dit alsnog in een grote hoeveelheid fouten, zeker als je beseft dat dat genoom vele malen gerepliceerd moet worden om van 1 naar 37.000 miljard cellen te gaan.

Daarnaast is schade aan het DNA onvermijdelijk omdat DNA van zichzelf een nogal instabiel molecuul is en doordat omgevingsfactoren schade kunnen toebrengen aan het DNA. Bekende voorbeelden van externe bronnen van DNA-schade zijn: UV-straling in zonlicht, chemicaliën in sigarettenrook en röntgenstraling. Verder zorgen ook metabole processen in een cel voor schade aan het DNA. De diverse oorzaken resulteren in verschillende types DNA-laesies: van chromosomale breuken waarbij het DNA-molecuul volledig gebroken is, ook wel dubbelstrengs breuken (DSBs) genoemd, tot kleine chemische veranderingen aan de moleculaire structuur van de nucleotiden. Organismen hebben een uitgebreide reeks aan 'verdedigingslinies' ontwikkeld om de nadelige effecten van een DNA-schade tegen te gaan en in de meeste gevallen zijn deze herstelmechanismes in staat om de schade foutloos te repareren. Desondanks veroorzaakt DNA-schade (of het foutieve herstel ervan) in sommige gevallen veranderingen van het genetisch materiaal, of het kan essentiële processen zoals DNA-replicatie blokkeren. Genetische veranderingen die ontstaan door DNA-schade, of foutjes tijdens DNA-replicatie, of andere processen, noemen we mutaties.

In dit proefschrift heeft een specifiek type DNA-schade mijn grootste interesse: DNA-laesies die wegversperringen vormen voor DNA-replicatie. De polymerases die normaal het DNA dupliceren zijn ontzettend efficiënt en vrijwel foutloos als het DNA onbeschadigd is, maar ze komen in de problemen als de structuur van het DNA afwijkt: het polymerase loopt dan vast en uiteindelijk blijven er stukken DNA over die niet gedupliceerd zijn of er kunnen breuken in het DNA ontstaan, welke

vervolgens leiden tot mutaties, verlies van grotere stukken DNA, verminderde groei & vruchtbaarheid, en uiteindelijk zelfs celdood. In meercellige organismen is het belangrijkste verdedigingsmechanisme dat hiertegen beschermt Translaesie DNA-synthese (TLS). TLS maakt gebruik van gespecialiseerde TLS-polymerases die, anders dan replicatieve polymerases, wél beschadigd DNA kunnen repliceren. Door dit te doen, zorgt TLS voor continuïteit van DNA-replicatie. Hierbij dient wel gezegd te worden dat TLS-polymerases iets vaker fouten maken dan de replicatieve polymerases en beschadigd DNA kan moeilijk te 'lezen' zijn, bijvoorbeeld doordat normale paring tussen A en T, of G en C, wordt beïnvloed. Hierdoor lijkt het logisch dat TLS resulteert in de vorming van nieuwe mutaties en dit is inderdaad nog steeds een wijdverbreid begrip. In dit proefschrift laat ik echter zien dat het netto-effect van TLS anti-mutageen is, omdat het de vorming van DNA-breuken en de resulterende genomische instabiliteit voorkomt. Het belang van TLS voor de menselijke gezondheid wordt goed geïllustreerd door patiënten met de genetische aandoening *Xeroderma pigmentosum* variant. Bij deze mensen functioneert TLS niet goed en daardoor zijn ze extreem gevoelig voor zonlicht en hebben ze zo'n 1000-maal meer kans op het ontwikkelen van huidkanker dan de gemiddelde bevolking. Kortom, TLS heeft twee grote voordelen: 1) het zorgt ervoor dat DNA-replicatie kan doorgaan in de aanwezigheid van DNA-schade; 2) het onderdrukt de vorming van nieuwe mutaties en beschermt daardoor de integriteit van het genoom.

Het modelorganisme *C. elegans*

In alle experimenten die ik beschrijf in dit proefschrift gebruik ik de nematode *C. elegans*. Dit kleine wormpje, ongeveer 1 mm lang en zo dun als een menselijke haar, komt voor in de bodem, verspreid over de wereld. Meer dan 50 jaar geleden is het door onderzoeker Sydney Brenner voor het eerst gebruikt als modelorganisme in de ontwikkelingsbiologie. Vooralsnog is *C. elegans* nog steeds een vrij nieuw model om de stabiliteit van het genoom en DNA-schadeherstel te bestuderen, maar wel een modelorganisme met grote voordelen. Ten eerste lijken wij mensen meer op deze wormen dan je verwacht. Friedrich Nietzsche schreef in 1883 al (vrij vertaald): "Je hebt je ontwikkeld van worm naar mens, maar van binnen ben je nog steeds grotendeels worm." Hoewel hij het niet had over genetica of moleculaire biologie, weten we nu dat het ook op deze velden van toepassing is. Vrijwel alle herstelmechanismen voor DNA-schade zijn geconserveerd gebleven tijdens de evolutie en dus functioneren ze zowel in *C. elegans* als in de mens, op nagenoeg dezelfde manier.

Verder is *C. elegans* zo geschikt als modelorganisme vanwege zijn eenvoud. Een volwassen dier bestaat uit slechts ongeveer 1.000 cellen, en het genoom is ~30 maal compacter dan dat van de mens. Hiernaast is *C. elegans* erg gemakkelijk te kweken en te onderhouden in het laboratorium. De worm is hermafrodiet, kan zichzelf bevruchten, en zeer snel voortplanten: binnen enkele dagen produceert een volwassen dier ~300

kinderen. Heel soms worden ook mannetjes geboren. Dit kunnen we in ons voordeel gebruiken, doordat we deze mannetjes kunnen kruisen met hermafrodieten. Als we dit doen met een mannetje met mutatie X en een hermafrodiet met mutatie Y, kunnen we nageslacht genereren dat beiden mutaties X+Y heeft. Dit is een groot voordeel bij genetisch onderzoek zoals het mijne.

Tijdens mijn promotieonderzoek is de CRISPR/Cas9 techniek beschikbaar gekomen en toegepast in *C. elegans*. Met deze techniek heb ik specifieke genetische aanpassingen kunnen maken, door genen uit te schakelen die een rol spelen bij het herstel van DNA-schade.

Anderzijds heb ik meegewerkt aan de ontwikkeling van een micro-evolutie-assay, dat ook gebruik maakt van de diverse voordelen van *C. elegans* als model: brede toegankelijkheid van genetische knock-outs, snelle groei en het compacte genoom, in combinatie met het sequencen van het gehele genoom en bio-informatica. Dit assay hebben we gebruikt om de accumulatie van spontane mutaties gedurende vele generaties op een onbevooroordeelde manier te bestuderen.

Dit proefschrift

Zoals ik eerder beschreef zijn er vele verschillende beschermingsmechanismes aanwezig in cellen die bescherming bieden bij DNA-schade. Deze mechanismes vormen samen een netwerk dat als geheel wordt aangeduid als *DNA damage response* (DDR). De DDR is een complexe cellulaire reactie bestaande uit signaleringsroutes en DNA-herstelnetwerken. In **hoofdstuk 2** bestudeer ik *Translaesie DNA-synthese* als onderdeel van de DDR. Ik beschrijf een verrassend eenvoudige lineaire volgorde van gebeurtenissen voor hoe *C. elegans* omgaat met DNA-schade veroorzaakt door licht. Als ik TLS en twee andere DNA-schadeherstelroutes tegelijkertijd inactiveer, kunnen de dieren in het donker nog op normale snelheid groeien en zich voorplanten, maar in de aanwezigheid van normaal daglicht kunnen ze eenvoudigweg niet groeien en produceren ze geen nageslacht meer. Zelfs zeer lage niveaus van normaal daglicht onderdrukken de groei al. Ik concludeer in deze studie dat TLS en de andere twee mechanismes dodelijke DNA-breuken, als gevolg van blootstelling aan licht, tegengaan. Mijn observaties in dit hoofdstuk laten ook zien hoe sterk licht het DNA kan beschadigen, en waarom er dus meerdere beschermingslagen nodig zijn tegen een dergelijke, veel aanwezige, bron van DNA-schade.

In **hoofdstuk 3** beschrijf ik de bijdrage van het TLS-polymerase genaamd 'REV-1' aan genoom stabiliteit, over een periode van ongeveer een half jaar, wat gelijk staat aan zo'n 50 generaties. Ik beschrijf hier dat REV-1 spontane mutagenese zowel bevordert als onderdrukt. Omdat REV1 TLS stimuleert, voorkomt het verlies van kleine stukjes sequenties uit het genoom, ook wel deleties genoemd. Het huidige dogma is dat TLS de vorming van mutaties stimuleert. Mijn bevindingen gaan in tegen dit dogma, omdat

de werking van REV-1 onder normale omstandigheden overwegend anti-mutageen is. Hiermee voorkomt REV-1 de accumulatie van deleties, ten koste van puntmutaties[2]. Ik laat ook zien dat slechts één beschadigde nucleotide per ~10^{10} nucleotiden, REV1-activiteit vereist. De conclusie is dat voortgang van DNA-replicatie belangrijker is dan het voorkomen van enkele nucleotidesubstituties.

Het idee dat TLS niet accuraat is, is voornamelijk vastgesteld na blootstelling aan genotoxides, zoals bijvoorbeeld UV-licht, omdat het aantal mutaties zonder deze blootstelling meestal te laag is om te detecteren. In **hoofdstuk 4** beschrijf ik de rol van alle vier TLS-polymerases tezamen bij het beschermen van genetische stabiliteit. We hebben de accumulatie van mutaties in TLS-deficiënte dieren die werden gekweekt onder normale omstandigheden gevolgd, wederom over ongeveer 50 generaties. Voor elke genetische achtergrond hebben we uitgebreide mutatieprofielen kunnen maken, met daarin mutaties die uiteenlopen van eenvoudige puntmutaties tot complexe genomische veranderingen. Mutanten waarbij alle bekende TLS-polymerases uitgeschakeld zijn, kunnen nog steeds groeien en planten zich voort, en blijkbaar is TLS dus niet essentieel. Niettemin hebben TLS-deficiënte mutanten verhoogde genomische instabiliteit. Verder stelde de grote dataset mij in staat te bepalen dat de meest voorkomende replicatieblokkades zich vormen op het nucleotide G, die bij afwezigheid van TLS leiden tot DSBs.

In **hoofdstuk 5** presenteer ik een nieuwe test om reparatie van een specifieke DNA-beschadiging te bestuderen, genaamd *interstrand crosslinks* (ICLs). ICLs zijn sterke verbindingen tussen nucleotiden van complementaire DNA-strengen en vormen absolute blokkades voor DNA-replicatie. Ze vormen een groot probleem voor levende organismes en het herstelproces is complex en nog niet in detail beschreven. Het doel hier was om een assay te ontwikkelen dat herstel van ICLs zichtbaar maakt in een meercellig organisme (*in vivo*) om bestaand biochemische (*in vitro*) assay aan te vullen. In dit hoofdstuk laat ik zien dat *C. elegans* hiervoor inderdaad kan worden gebruikt. We zijn erin geslaagd om stukjes DNA met een ICL te introduceren in levende dieren om vervolgens de sequentie te bepalen van het geïntroduceerde stukje DNA nadat DNA-herstel-processen in de cellen van *C. elegans* de ICL gerepareerd hebben. Door dit te doen kunnen we zien of er wel of geen mutatie is ontstaan op de locatie van ICL. We kunnen ook direct zien wat voor mutatie er is ontstaan. Vervolgens gebruik ik deze nieuwe methode om verschillende DNA-reparatie-deficiënte achtergronden te testen. Onze eerste voorlopige bevindingen laten zien dat deze nieuwe methodologie werkt en met deze test komt er een nieuwe tool beschikbaar om ICL-reparatie *in vivo* te bestuderen.

In de algemene discussie in **hoofdstuk 6** heb ik mijn bevindingen in een breder perspectief geplaatst en beschrijf ik welke belangrijke vragen nog onbeantwoord zijn en interessant zijn voor toekomstig onderzoek. Gebaseerd op mijn eigen onderzoek concludeer ik uiteindelijk dat de sterke evolutionaire conservatie van TLS verklaard

[2] Puntmutaties zijn een specifiek soort mutaties waarbij het ene nucleotide veranderd is in een ander nucleotide. Bijvoorbeeld een verandering van een G naar een T. In de meeste gevallen zijn puntmutaties veel minder schadelijk dan deleties.

wordt door de dubbele functies van TLS: het bewaken van zowel replicatiepotentieel als genoomstabiliteit. In alle studies zie ik dat TLS genomische instabiliteit onderdrukt, omdat het de conversie van replicatieblokkades naar DSBs voorkomt. Als TLS niet goed functioneert ontstaan er DSBs die vervolgens resulteren in grotere en meer schadelijke mutaties die worden doorgegeven aan volgende generaties. Translaesie-DNA-synthese is gunstig voor de gezondheid van individuele organismen, soorten, en het leven in het algemeen, omdat het voortdurende voorplanting en groei ondersteunt. Hoewel DNA-schade altijd aanwezig en onvermijdelijk is, bewaakt TLS tegen de vorming van de mutaties die anders leiden tot kanker, veroudering en aangeboren ziekte.

Acknowledgements | Dankwoord

As I write this, the final hurdle of this voyage called a PhD is in sight: my defense on December 3rd. On the twisting path there have been many roadblocks: disappointments, negative results, or other struggles. It can be hard to put those aside, forget frustrating projects and celebrate the victories, but now I am delighted about the end product. And although the journey has been challenging, it was never lonely. I feel fortunate that so many have helped me bypass the barriers I encountered.

Marcel, your enthusiasm for science is inspiring. Your drive has motivated me to persevere even when projects were seemingly not going anywhere. I owe many thanks to all other members of the Tijsterman group: Nick, Bennie, Wouter, Sophie, Robin, Ron, Jane, Jordi, Evelina, Maartje, Juliette, Joost & Hanneke. I miss your wise lessons, silly jokes and creative ideas. Juliette, you followed me as junior PhD candidate, but you are much wiser than I will ever be. Thank you for standing beside me as paranymph. I am also grateful to all other people from the department of Human Genetics, who all contributed to scientific excellence and a warm working environment.

Ik kan nogal snel afgeleid zijn, of beter: ik laat me graag afleiden. Mijn vele lieve vrienden (die ik hier niet allemaal kan noemen) zijn daar ontzettend goed in. Tegen allen zeg ik: Proost! Beste maat, Eric, bedankt voor al die jaren vriendschap en voor het strakke design van dit boek. Dank ook aan mijn cadeaufamilie: Mieke, Cor, Bas, Parmila, Marieke, Peter en kids. Dan mijn wijze broertje, Toine, ik voel me gesterkt met jou naast mij als paranymph! Lieve papa en mama, Boeddha had gelijk toen hij zei: Ook al draag je je ouders voor 100 jaar op je schouders dan nog heb je ze niet genoeg bedankt voor alles dat ze voor je gedaan hebben. Ik ben jullie voor zoveel meer dankbaar dan een set goede genen. Bedankt voor jullie steun, zelfs in lastige tijden. Ik beloof het, nu ben ik echt klaar met studeren. Liefste Jantine, wat ben ik gelukkig dat we zo'n goede symbiose zijn. Bedankt dat je mij achter mijn broek aangezeten hebt om dit boek af te maken. Ik ben blij met jou aan mijn zijde tijdens het laatste hoofdstuk van dit verhaal en de eerste hoofdstukken van nieuwe avonturen.

Salut,
Ivo

Curriculum vitae

Ivo van Bostelen was born on October 1st 1983 in Leiderdorp and grew up in Roelofarendsveen. In 2002, he successfully completed pre-university education (VWO) with a Nature and Health profile at the Bonaventura college in Leiden. With a keen interest in biology and chemistry Ivo then decided to study biotechnology at Wageningen University & Research center. When he completed his bachelor, he decided to specialize in cellular and molecular biology for his master of science. In 2008, Ivo got his first experience with scientific research by doing a short internship at the molecular biology department in the group of Prof. dr. Ton Bisseling in Wageningen, where he studied the molecular factors involved in plant-bacteria symbiosis. Shortly after, early in 2009, Ivo moved to the lab of Prof. dr. Monica Colaiácovo at the genetics department of Harvard Medical School in Boston USA, for a six months internship. At this prestigious environment he studied chromosomal paring and genome instability during meiosis. The work in Boston did not only get hem hooked on science, he also there got infected with the *C. elegans* bug. Upon return in the Netherlands Ivo decided to move somewhat more towards a medical profile; he found a great place to perform his final Master project at the department of experimental oncology in the group of Prof. dr. Susanne Lens, in the Utrecht Medical Center. Here, he got the chance to study the cellular processes and molecular factors that guide cell division in human cell cultures. Ivo successfully completed this project, and thereby his Master in Cellular and Molecular biotechnology in the summer of 2010. In March 2011, Ivo started his PhD work under the supervision of Prof. dr. Marcel Tijsterman at Leiden University Medical Center, at the former department of Toxicogenetics. After several years, Toxicogenetics became part of the department of Human Genetics where Ivo worked until 2017. As of August 2017, Ivo has made the transition from academia to industry and now works as a Project Leader at MRC Holland in Amsterdam, where he develops diagnostic tests for a wide range of human genetic diseases.

Publications

Y.B. Tzur, C.E. de Carvalho, S. Nadarajan, **I. van Bostelen**, Y. Gu, D.S. Chu, I.M. Cheeseman & M.P. Colaiácovo. LAB-1 targets PP1 and restricts Aurora B kinase upon entrance into meiosis to promote sister chromatid cohesion.
Published in PLoS Biology August 2012.

I. van Bostelen & M. Tijsterman. Combined loss of three DNA damage response pathways renders *C. elegans* intolerant to light.
Published in DNA Repair April 2017.

I. van Bostelen, R. van Schendel, S. Roerink & M. Tijsterman
Suppression of genome instability by Y-family TLS polymerase REV-1 in *C. elegans*.
Manuscript in preparation.

I. van Bostelen, R. van Schendel, R. Romeijn, M. Tijsterman. Translesion DNA synthesis of endogenous lesions protects genome integrity in *C. elegans*.
Manuscript in preparation.

R. van Schendel, **I. van Bostelen**, & M. Tijsterman. Mutagenic consequences of repair of cisplatin and psoralen interstrand crosslinks in *C. elegans*.
Manuscript in preparation.