High-resolution analysis of replication-associated lesions bypass in time and space

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Maria Beatriz Villagomez Torres

geb. am 29.04.1989, Ecuador

November 2022

Dekan:
1. Berichterstatter:
2. Berichterstatter:

Tag der mündlichen Prüfung:

Declaration

Herewith I, Maria Beatriz Villagomez Torres, declare that the work presented in this thesis is my very own. I confirm that any information or data that have been derived from other sources are referred to and correctly cited in the thesis.

Abstract

To maintain genome stability, DNA must be replicated accurately during every cell cycle. Fork progression and DNA synthesis depend on the coordinated action of the replisome. Nonetheless, replisomes frequently encounter various obstacles and, if not properly resolved, these can slow down or stall the progression of the replication fork. Even though the mechanisms of how the replisome responds to lesions are not completely understood, a few *in vitro* studies have implied mechanistic differences between the synthesis of lagging and leading DNA strands. The current challenge is to understand how the replisome responds to damage *in vivo*; therefore, the aim of this thesis was to comprehend the cellular responses to site-specific lesions found in lagging and leading strands in *Saccharomyces cerevisiae*. Additionally, to gain further insights into damage bypass, post-replication repair was investigated to understand its temporal and spatial organisation across the genome.

To gain insights into how leading and lagging strands are replicated in the presence of polymerase-stalling lesions, the distribution of labour of the main DNA polymerases was investigated and the synthesis of daughter strands analysed under damaging conditions. Utilising a system that inflicts abasic sites at a single locus on chromosome VI in budding yeast while monitoring DNA polymerase usage, first, evidence for the role of Polymerase δ in the synthesis of damaged leading strands was provided. Secondly, by combining Next-Generation Sequencing approaches, I showed that both leading and lagging strands remained mostly un-replicated in the presence of damage. Moreover, it was observed that replication re-started downstream of the lesion, possibly indicating that *de-novo* re-priming events occurred in both strands. Additionally, the presence of abasic sites altered the loading of core replisome proteins at the damaged site. Finally, the experimental approach used during this study generated high genomic instability, which made it difficult to conclude on the dynamics of the replisome under DNA damaging conditions.

To define the distribution of post-replication repair and to understand how cells maintain balance between translesion synthesis and template switching, a doxycycline repression system for depletion of Rad18 was applied. This confirmed that template switching is the preferred pathway over translesion synthesis in bypassing lesions induced by the alkylating agent methyl methanesulfonate. Interestingly, the genome-wide analysis of post-replication tracts indicated biases of template switching and translesion synthesis mostly towards regions within early firing origins as well as a greater contribution on leading strands than on lagging strands.

Zusammenfassung

Zur Aufrechterhaltung der Genomstabilität muss die DNA während jedes Zellzyklus‘ akkurat repliziert werden. Das reibungslose Fortschreiten der Replikationsgabel und die Synthese neuer DNA hängt von der koordinierten Aktion des Replisoms ab. Trotzdem stoßen Replisomen häufig auf Hindernisse, welche das Fortschreiten der Replikationsgabel verlangsamen oder aufhalten, wenn sie nicht überwunden werden können. Die Mechanismen, durch die das Replisom auf Läsionen reagiert, sind bisher nicht vollständig verstanden. Einige *in vitro* Studien deuten auf mechanistische Unterschiede zwischen der Synthese von Leit- und Folgestrang hin. Daher ist es wichtig und notwendig zu verstehen, wie das Replisom *in vivo* auf Schäden im Genom reagiert. Das Ziel der vorliegenden Arbeit war es, die zellulären Reaktionen auf ortsspezifische DNA-Läsionen zu verstehen, die in diskontinuierlichen und kontinuierlichen DNA-Strängen in *Saccharomyces cerevisiae* gefunden wurden. Um weitere Erkenntnisse in die Schadensumgehung zu erhalten, wurde außerdem untersucht, wie die Reparatur nach der Replikation zeitlich und räumlich über das Genom hinweg organisiert ist.

Um zu verstehen, wie Leit- und Folgestrang der DNA in Gegenwart von Läsionen, die Polymerasen behindern, repliziert werden, wurde die Arbeitsverteilung der wichtigsten DNA-Polymerasen untersucht und die Synthese von Tochtersträngen in Gegenwart von DNA-Schäden analysiert. Unter Verwendung eines Systems, das gleichzeitig abasische Stellen an einem einzigen Lokus auf Chromosom VI in Bäckerhefe verursacht als auch die Verwendung von DNA-Polymerasen identifiziert, wurde zunächst ein Nachweis für die Rolle von Polymerase δ bei der Synthese von geschädigten Leitsträngen geliefert. Außerdem wurde durch den Einsatz von Hochdurchsatzsequenzierungs-Methoden gezeigt, dass sowohl kontinuierliche als auch diskontinuierliche DNA-Stränge in Gegenwart von Schäden größtenteils nicht repliziert werden. Darüber hinaus wurde beobachtet, dass die Replikation hinter der Läsion neu gestartet wurde, was auf einen Neustart der Replisomen auf beiden DNA-Strängen hindeutet. Des Weiteren veränderte das Vorhandensein von abasischen Stellen die Assoziation von Replisom-Bestandteilen an der beschädigten Stelle. Schließlich ist zu erwähnen, dass der während dieser Studie verwendete experimentelle Ansatz eine hohe genomische Instabilität erzeugte, die es schwierig machte, finale Rückschlüsse auf die Dynamik des Replisoms unter DNA-schädigenden Bedingungen zu ziehen

Um die Verteilung der Postreplikations-Reparatur zu präzisieren und zu verstehen, wie Zellen das Gleichgewicht zwischen Transläsions-Synthese und Matrizenwechsel erhalten, wurde ein Doxycyclin-abhängiges System zur Depletion von Rad18 angewendet. Dies bestätigte, dass der Matrizenwechsel der bevorzugte Weg gegenüber der Transläsions-Synthese bei der Behebung von Läsionen ist, welche durch das Alkylierungsmittel MMS induziert wurden. Interessanterweise zeigte die genomweite Analyse von Postreplikationsstellen außerdem eine Überrepräsentation des Matrizenwechsels und der

Transläsions-Synthese in Regionen mit frühen Replikationsursprüngen sowie einen größeren Einfluss auf die Leitstrang-Synthese im Vergleich zur Folgestrang-Synthese der DNA.

Acknowledgments

Content

List of figures

List of tables

Chapter 1: Introduction

1.1 Molecular mechanisms of DNA replication and replication stress

DNA replication is a fundamental biological process to maintain genome stability (Mirkin & Mirkin, 2007). This event is highly regulated, and it consists of a coordinated and multistep process that assembles several replication factors to the replication fork in a cell-cycle-dependent manner. Three DNA polymerases are the central enzymes involved in the synthesis of daughter strands. Polymerase α (Pol α), which has a primase activity, and polymerase ε (Pol ε) and polymerase δ (Pol δ) responsible for the bulk synthesis and elongation of the leading and lagging strands, respectively (reviewed in Johansson & Dixon, 2013). A variety of endogenous and exogenous agents capable of inducing DNA lesions constantly challenge genome stability. Thus, replisomes frequently encounter several obstacles to their progression, and if not properly solved, these barriers can slow down or stall fork progression, making them prone to produce DNA breaks (Mazouzi, et al., 2014). In this section, I describe the molecular mechanisms surrounding DNA replication and how this is affected by DNA damage.

1.1.1 DNA replication program

DNA replication is a fundamental biological process that ensures the accurate and complete duplication of the genetic code (Prioleau & MacAlpine, 2016). An essential pre-requisite is that DNA replication must occur with extremely high fidelity and efficiency once per cell cycle in order to prevent the accumulation of genetic alterations that potentially disturb cell survival and viability (Ekundayo & Bleichert, 2019).

In eukaryotes, DNA replication initiates at multiple origins, and then spreads across the genome. Features that define replication origins differ within species. The chromosomal sequences sufficient to act as an origin of replication in *Saccharomyces cerevisiae*, for instance, are three to four DNA sequences of 10-15 base pairs (bp) spread over 100-150 bp that include the highly conserved and essential Autonomously Replicating Sequence (ARS) (reviewed in Barberis, 2010). In the genome of budding yeast, several studies have identified and mapped over 700 potential origin target sites (Xu et al., 2006). Evidence also indicates that yeast chromosomes have more replication origins than needed for timely replicate DNA during S-phase, and only a small number of these seem to be efficient replicators (Barberis, et al., 2010). Moreover, chromosome context and chromatin structure appear to play a key role in controlling the efficient and timely activation of replicating origins, including the

action of different chromatin-binding proteins, chromatin-modifying enzymes, and chromatin remodelling proteins (Knott, et al., 2009).

DNA replication initiation involves two steps, licensing and firing. These steps are closely interconnected but are coupled to separate phases of the cell cycle (Masai, et al., 2010). Licensing occurs during mitotic exit and G1-phase at the origins of replication. It involves the loading of the pre-replicative complex (pre-RCs) composed by the origin recognition scaffold (Orc1-6), the helicase loader (Cdc6 and Cdt1p) and the Mcm2-7 helicase, which binds in the form of a double hexamer encircling double-stranded DNA (dsDNA), but is inactive at this stage (Boos & Ferreira, 2019). Subsequently, firing occurs in early S-phase where Cdc45 and GINS associate with Mcm2-7 thereby activating the CMG helicase complex (Cdc45, Mcm2–7, and GINS) and remodelling the helicase to encircle single-stranded DNA (ssDNA) for DNA unwinding (Reusswig & Pfander, 2019). Consequently, the mature replisome allows each parental strand to synthesise complementary daughter strands. Three main DNA polymerases are in charge of replicating the genome Pol α, Pol ε, and Pol δ. Accompanied by other core proteins such as the clamp loader Replication Factor C (RFC), the ring-shaped DNA sliding clamp Proliferating Cell Nuclear Antigen (PCNA; Pol30 in *S. cerevisiae*) and the single-strand DNA-binding protein Replication Protein A (RPA; RFA in *S. cerevisiae*) (Bai, et al., 2017; Burgers & Kunkel, 2017) (Figure 1.1).

Because DNA polymerases can only extend DNA in direction 5′ to 3′, only the leading strand is synthesised continuously in the direction of the fork movement, while the lagging strand is extended in the opposite direction in a discontinuous manner. The replication of both leading and lagging strands is initiated by Pol α in a process called priming. Pol α will synthesise a primer that contains a short RNA stretch (~10 nucleotides) followed by 10-20 DNA bases (Leman & Noguchi, 2013). When Pol α detaches from the template, RFC very effectively allows the access to the primer 3'-end for PCNA loading, and then Pol δ and Pol ε will continue the synthesis of DNA. Interestingly, the synthesis of the lagging and leading strands are mechanistically different. On the one hand, the recruitment of Pol δ to PCNA accomplishes the synthesis of the lagging strand in the form of Okazaki fragments. As these are discontinuous fragments, their maturation finalises with the action of Fen1 and DNA ligase 1 (in yeast Rad27 and Cdc9, respectively). On the other hand, leading strand is synthesised when RFC associates with Pol ε, and then PCNA loads to dsDNA (Bell & Labib, 2016). Replication is completed when forks from different origins collide with each other or when forks reach both ends of the chromosomes (Wang, et al., 2001).

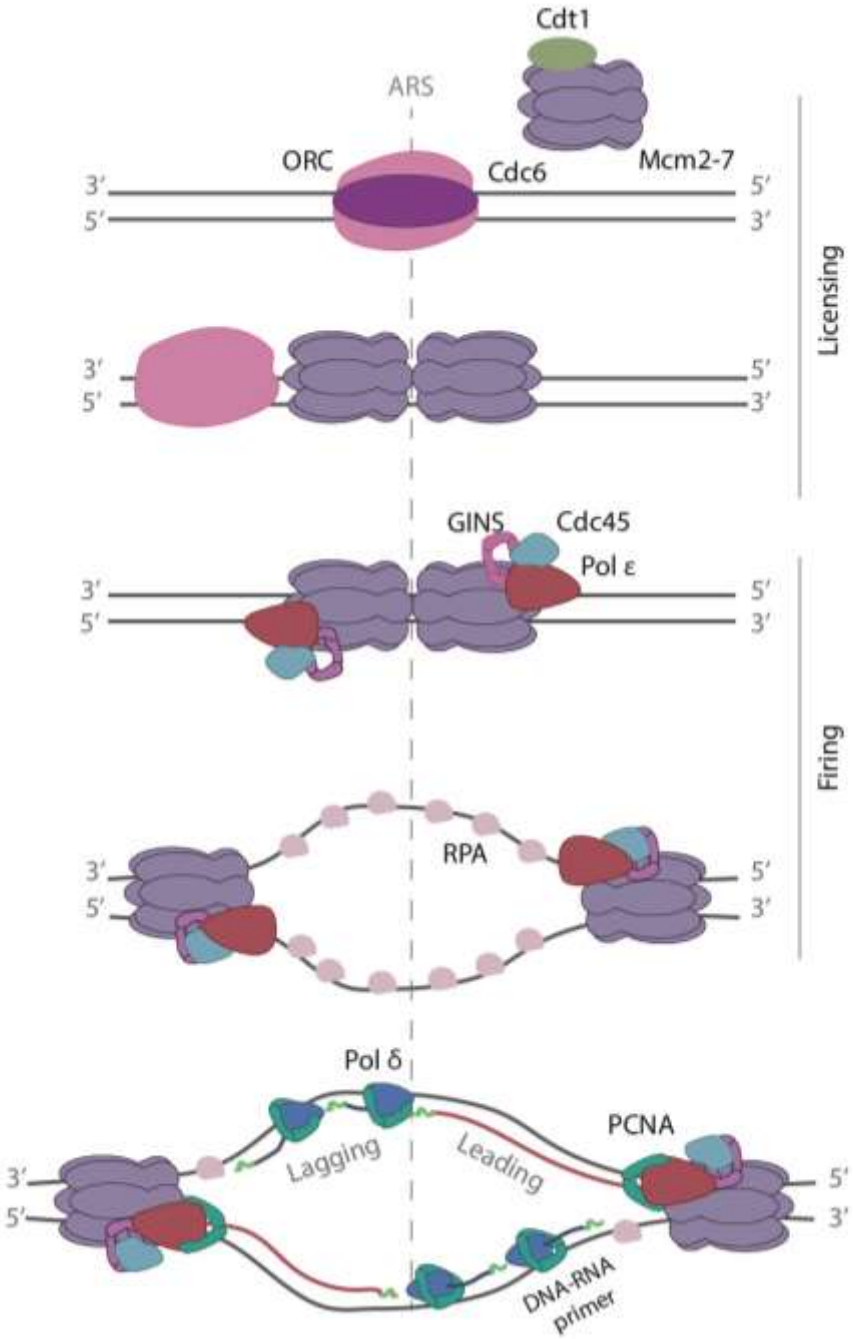

Figure 1.1. Assembly of the eukaryotic replisome.

The licensing of replication initiates with the origin recognition scaffold (Orc1-6) and Cdc6, together they recruit one Cdt1–Mcm2-7 complex, followed by a second one, to form a double Mcm2-7 hexamer at each direction of replication. Other proteins are required for further assembly: Dpb11, Sld2, Sld3, and Sld7 (not shown). The mature replisome formed by Cdc45, Mcm2-7 and GINS (CMG complex) Pol ε as well as DDK and CDK kinase activity prime replication firing. The unwinding of double-stranded DNA by the helicase complex allows the access to parental strands for the synthesis of daughter strands. Pol α starts synthesis by forming a small DNA/RNA primer (green line), which is extended by Pol ε and Pol δ to form the leading (red line) and lagging strands (blue line), respectively (modified from Burgers & Kunkel, 2017).

1.1.2 Problems at the replication fork

Genome replication is an extremely accurate process producing only one error per 10^9–10^{10} nucleotides polymerised (Loeb, 1991). This extraordinary capacity is due to the high fidelity of replicative polymerases to select the correct nucleotide as well as their excellent exonuclease proofreading activity (reviewed in Loeb & Monnat Jr, 2008). However, the intrinsic fidelity of replicative polymerases also means that they are easily stalled by different lesions on the DNA. The DNA replication machinery, in fact, must continuously face a variety of obstacles triggered by intracellular and extracellular sources causing replication stress, which causes slowing-down or the complete stalling of DNA synthesis. A prolonged stalling of the replicative forks and the failure to stabilise them notably can induce their collapse or the dissociation of the replisome from DNA, potentially leading to the formation of DNA double-strand breaks (DSBs) and chromosomal rearrangements (Zeman & Cimprich, 2014).

In humans, there are more than 40.000 lesions per cell at any given time due to metabolism, and naturally occurring depurination and deamination events (Nakamura, et al., 2014), so it is inevitable that the replication fork be constantly challenged by lesions. Most of these lesions can trigger the activation of the damage checkpoint that will promote the activation of the DNA-damage response (DDR) mechanism or damage tolerance events to preserve fork stability and complete DNA replication (Gao, et al., 2017).

Replicative polymerases can stall during replication for several reasons, for instance, the recognition of a damaged template can block their progression. Physical barriers on the template like protein complexes or secondary structures can obstruct the progress of polymerases. Low nucleotide or amino acid pools also cause polymerase stalling. In each case, cells must be able to recognise when a polymerase stalls to resolve the pertinent problem and proceed with DNA replication, however, how polymerases respond to stalling events is not well understood (Edenberg, et al., 2014).

It is known, however, that polymerases-stalling lesions usually result in the continuing unwinding of the parental DNA by the CMG helicase complex, which leads to the formation of stretches of single-stranded DNA (Pacek, 2004). The persistence of ssDNA, coated with RPA, generates the signal for replication stress response. Central to this checkpoint-signalling pathway is the damage sensing protein Ataxia-telangiectasia and Rad3-related, ATR (Mec1 in *S. cerevisiae*). ATR is recruited to stalled replication forks and once activated, phosphorylates several downstream proteins, including Chk1 (Rad53 in *S. cerevisiae)* (Paulsen & Cimprich, 2007). Checkpoint signalling in response to replication stress can be divided in two branches, both initiated from Mec1 and Rad53, but different mediator proteins are responsible for signal transmission: the DNA damage checkpoint (DDC) and the DNA replication checkpoint (DRC). While DDC functions throughout the cell cycle and depends on the checkpoint mediator Rad9 to activate Rad53, the DRC is specific to S-phase and is mediated by Mrc1 and other fork components to signal replication impediments (Pardo, et al., 2017).

The checkpoint activation stabilises forks by preventing the disassembly of the replication machinery, blocking cell cycle progression, downregulating late origin firing, and facilitating the re-start of collapsed forks (Paulsen & Cimprich, 2007) (Figure 1.2). Replication forks that are stabilised by the ATR pathway can re-start after the removal of the source of stress. Nonetheless, other pathways act when the blocking stress cannot be removed. Dormant origins firing, for instance, can rescue replication forks stalled at DNA lesions (McIntosh & Blow, 2012). Second, the replication machinery can re-prime downstream of the lesion leaving behind ssDNA gaps. The gaps then are filled-in by post-replication repair (PRR) pathways using specialised polymerases or the sister chromatid as a template (Mailand, et al., 2013) (detailed in Section 1.2). Overall, these processes allow the completion of replication, preventing prolonged fork stalling.

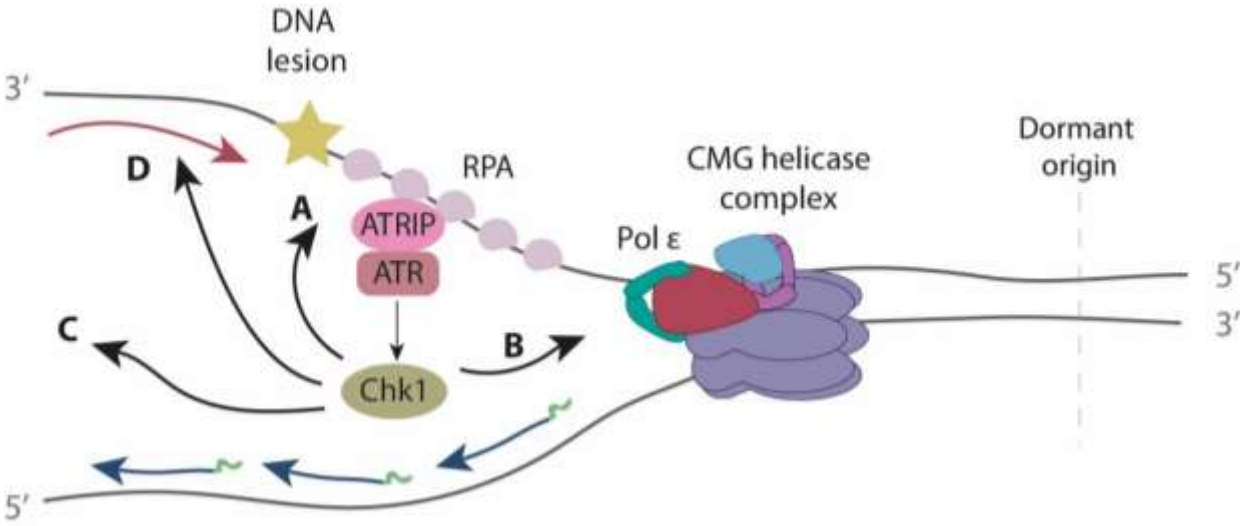

Figure 1.2. The ATR-mediated replication stress response.
ATR (Mec1) and its obligate binding partner ATRIP are activated by the stalled replication fork, where ATR initiates a signalling cascade primarily mediated by the effector kinase Chk1 (Rad53). This response promotes: **A.** fork stabilization and re-start (e.g. firing dormant origins, lesion bypass), **B.** the slowing down S-phase progression, **C.** the prevention of late replication origins firing **D.** the prevention of the disassembly of the replication machinery (modified from Zeman & Cimprich, 2014).

1.1.3 Sources of replication stress

Replication stress arises from various endogenous and exogenous sources that act locally and globally on replication dynamics. Such threats include by-products of cellular metabolism, reactive oxygen species (ROS), nucleotide pool imbalances, shortage of replication factors, ultraviolet (UV) light, and a number of chemical mutagens (Wilhelm, et al., 2020). Other sources that potentially lead to replication stress are linked to the genetic and epigenetic features of specific loci, such as telomeres, centromeres, ribosomal DNA (rDNA) or fragile sites. These genetic regions are intrinsically difficult to replicate due to the presence of repeated sequences that can form secondary structures (Gadaleta & Noguchi, 2017). Next, I described the sources of replication stress relevant to this study.

1.1.3.1 Apurinic/Apyrimidinic (AP) or abasic sites

Thousands of Apurinic/Apyrimidinic (AP) sites are generated in each cell daily, making them the most frequent DNA lesion in cells. Abasic sites are defined by a structure where a base is removed from the DNA while retaining the deoxyribose and the phosphodiester bonds (Kitsera, et al., 2019). These can be formed spontaneously through base hydrolysis and be the result of intermediates of DNA repair pathways as they are formed from the removal of misincorporated or damaged bases (Boiteux & Guillet, 2004). AP sites result from hydrolysis of the N-glycosyl bond enzymatically by the action of glycosylases (Gelot, et al., 2014) (Figure 1.3). In yeast, genetic studies have demonstrated that most spontaneous AP sites originate from the misincorporation of 2´-deoxyuridine, 5´-triphosphate (dUTP) in place of 2′-deoxythymidine-5′-triphosphate (dTTP) during DNA synthesis, subsequent uracil removal by the uracil DNA glycosylase, Ung1, creates a potentially toxic AP site (Kim & Jinks-Robertson, 2010). Finally, abasic sites are able to block DNA polymerases progression, which threatens the completion of DNA replication (Thompson & Cortez, 2020).

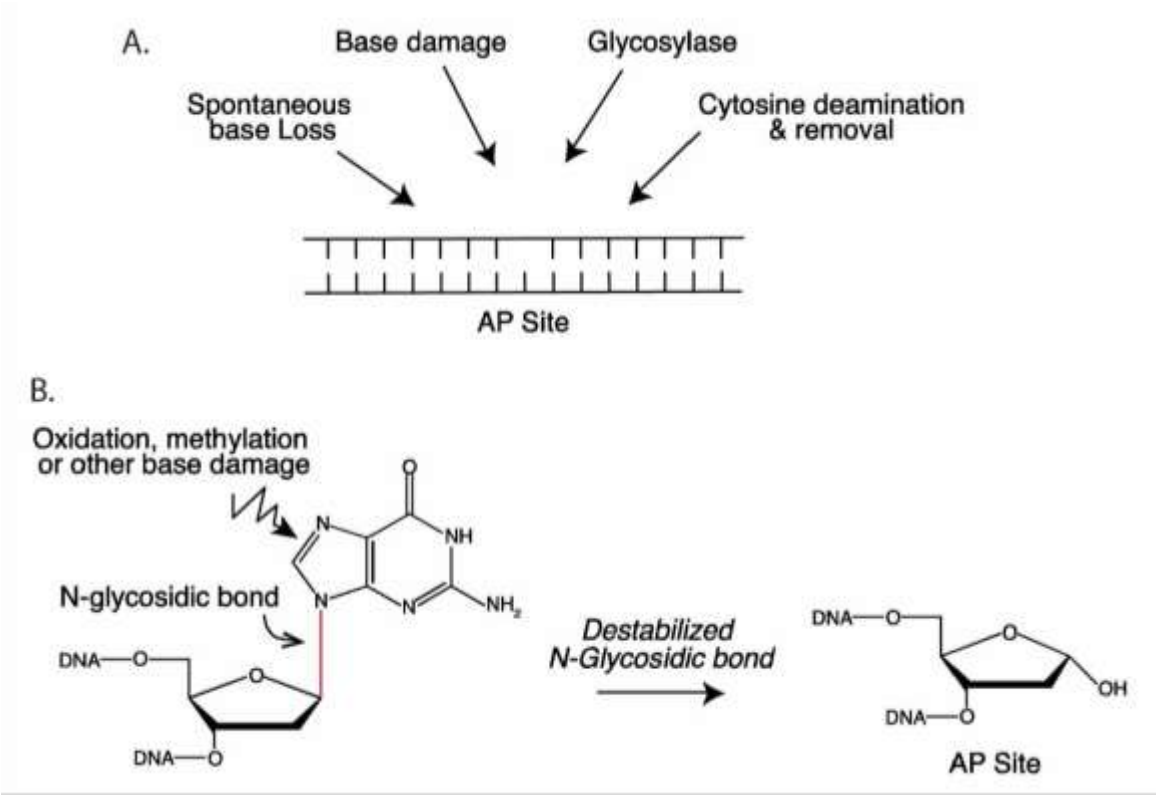

Figure 1.3. Mechanisms of AP site formation.

A. Potential sources of AP site formation. B. Destabilisation of N-glyosidic bond after base damage generates AP site. Abasic sites have a structure where the DNA backbone remains intact but the DNA base is missing (Thompson & Cortez, 2020).

1.1.3.2 Misincorporation of ribonucleotides

While DNA polymerases are highly selective, they misincorporate more than 13,000 ribonucleotides (rNMPs) during each round of DNA replication in budding yeast (Williams, et al., 2016). Ribonucleotide Excision Repair (RER) is the main pathway that recognises and removes these lesions through the action of the specialised enzyme RNase H2 together with the endonucleases Fen1 and Exo1. Importantly, when the replisome encounters a misincorporated ribonucleotide, DNA replication stalls blocking the progression of replicative polymerases (Zeman & Cimprich, 2014). To prevent DNA synthesis from stalling, replicases are replaced by translesion synthesis (TLS) polymerases (described in Section 1.2.2), which have a lower base selectivity (Cerritelli & Crouch, 2016). Alternatively, it has been shown in *S. cerevisiae* that the removal of rNTPs can be initiated by topoisomerase 1 (Top1), in a process known as "salvage pathway", which results in an aberrant and irreversible single-strand DNA breaks with further mutagenic consequences that also can lead to replication stress (Kim, et al., 2011).

1.1.3.3 Nicks, gaps, and ssDNA

Nicks, gaps, and ssDNA represent natural intermediates in several DNA repair pathways, which are tied to replication stress. Indeed, if the replication machinery encounters these lesions, they could be passively converted to DSBs (Zeman & Cimprich, 2014).

1.1.3.4 Bulky DNA lesions

Bulky DNA adducts are indicators of exposure to genotoxic compounds and the result of metabolism and DDR by-products (Ricceri, et al., 2010). The exposure to short-wave UV radiation induces the formation of the bulky adduct, cyclobutane pyrimidine dimer (CPD) on the DNA, which can reduce the speed of the replication fork progression (Elvers*, et al.,* 2011). On the other hand, the exposure of DNA with the alkylating agent methyl methanesulfonate (MMS) modifies both guanine (to 7-methylguanine) and adenine (to 3-methlyladenine) causing base mispairing and replication blocks, respectively (Lundin, et al., 2005).

1.1.4 Repair of abasic sites and bulky adducts

Two main pathways are engaged in repairing abasic sites and bulky adducts, and their activation depends on the type, location, and context of the damage.

1.1.4.1 Base excision repair

The formation of AP sites initiates the base excision repair (BER) pathway. In mammalian cells, BER can be broken in short-patch repair for single-nucleotide lesions and long-patch repair for multiple-nucleotide lesions; however, in yeast, both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe,* long-patch repair has not been yet demonstrated (Memisoglu & Samson, 2000). BER initiates by the removal of the alkylated base by a specific mono-functional DNA N-glycosylase. In *S. cerevisiae*, Mag1 is the only DNA glycosylase known to excise alkylated bases (Admiraal, et al., 2019). The resulting AP site is processed by a AP

endonuclease, Apn1 or Apn2, which excises DNA at the 5′- end of the AP site yielding SSBs and leaving a 3′- hydroxyl and 5′-deoxyribose phosphate (5′-dRP) end (Thompson & Cortez, 2020). The removal of 5′-dRP can be accomplished through the successive action of a Pol (presumably δ or ε) and a 5′-flap endonuclease (Rad27/Fen1) followed by sealing DNA strands with a DNA ligase (Cdc9/Lig1) (Ma, et al., 2008). The repair of abasic sites can also occur in an Apn1-independent manner by the action of a glycosylase-associated AP lyase, which nicks the sugar-phosphate backbone at the 3′- end of the AP site. Ntg1, Ntg2, and Ogg1 have such AP lyase activity (Kim & Jinks-Robertson, 2010) (Figure 1.4).

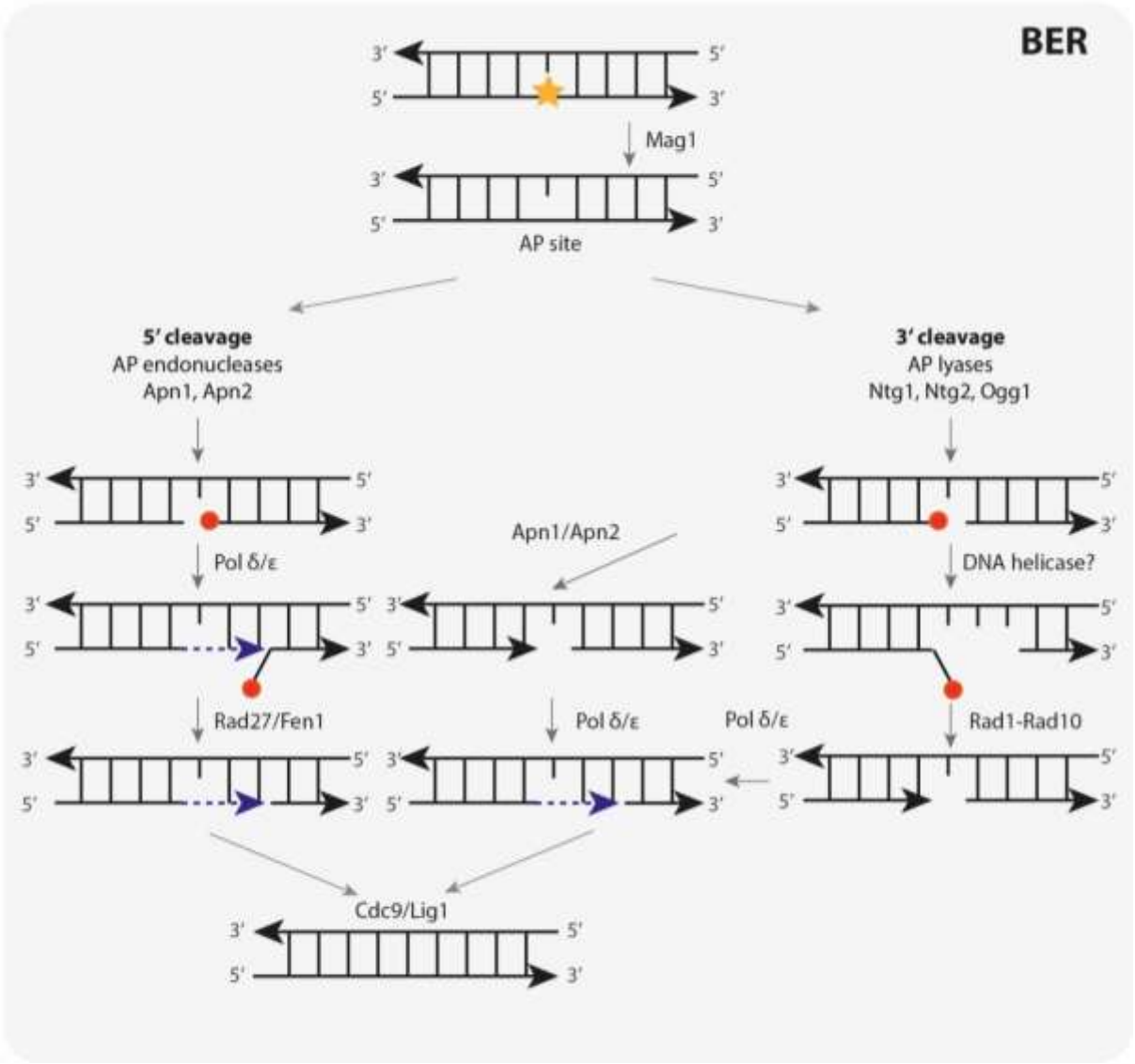

Figure 1.4. Simplified scheme of BER in budding yeast.
BER initiates with the action of a specific DNA glycosylase (Mag1 in budding yeast) that recognises and eliminates the damaged base generating an AP site in DNA. *Left:* The DNA backbone is cleaved at abasic sites by an AP endonuclease, Apn1 or Apn2, resulting in the formation of a 3′-hydroxyl and a 5′- deoxyribose phosphate (5′-dRP). The resulting gap is then filled-in (Pol δ or ε) and ligated (Cdc9). *Right*: In a minor fraction, the AP site can be catalysed by an AP lyase, Ntg1, Ntg2 or Ogg1, forming a 3′-dRP end. The 3′-dRP is excised by Apn1 or Apn2, alternatively, a 3′-flap structure can be formed (unknown DNA helicase) and cleaved (Rad1-Rad10) (modified from Boiteux & Guillet, 2004).

1.1.4.2 Nucleotide excision repair

The nucleotide excision repair (NER) pathway removes bulky and helix-distorting lesions and it is well known for its role in repairing UV-induced pyrimidine dimers. In addition, it has been seen that NER can act as a back-up pathway for the removal of AP sites (Boiteux & Guillet, 2004). In eukaryotes, NER initiates with the incision of a long DNA fragment on both sides of the lesion (around 25-30 nucleotides), followed by repair synthesis and ligation. NER is divided into two sub-pathways that differ in the initial steps of DNA damage recognition: global genomic (GG) and transcription coupled (TC). During the GG-NER, the entire genome is examined for lesions. In contrast, TC-NER is activated when RNA polymerase II (RNAP II) is stalled during transcript elongation, thus dedicated to the rapid removal of lesions on actively transcribed genes. (Marteijn, et al., 2014).

In yeast, GG-NER requires the damage binding factors Rad14, RPA, and the Rad4–Rad23 complex to recognise damage and recruit the transcription factor TFIIH that contains the DNA helicases Rad3 and Rad25. TFIIH unwinds DNA to create a bubble structure that surrounds the DNA lesion, and then two endonucleases, the Rad1–Rad10 complex and Rad2, incise the damaged DNA strand. The excised oligo deoxynucleotide is then released in a complex with TFIIH. Following the release of the damaged DNA, DNA polymerases fill-in the gap. After ligation of the repaired patch, the process of NER is complete (Prakash & Prakash, 2000; Vaughn & Sancar, 2020) (Figure 1.5).

Rad26 plays an essential role in TC-NER; however, this is not solely responsible for TC-NER. Rpb9, a nonessential subunit of RNAP II, is largely responsible for Rad26-independent TC-NER. The Rad26- and Rpb9-dependent TC-NER have different efficiencies in genes with different transcription levels and regions of a gene. For instance, TC-NER appears to be accomplished primarily by the Rad26-dependent mechanism in slowly and moderately transcribed genes. In contrast, for highly transcribed genes, Rpb9 plays a more significant role (Li S. , 2015).

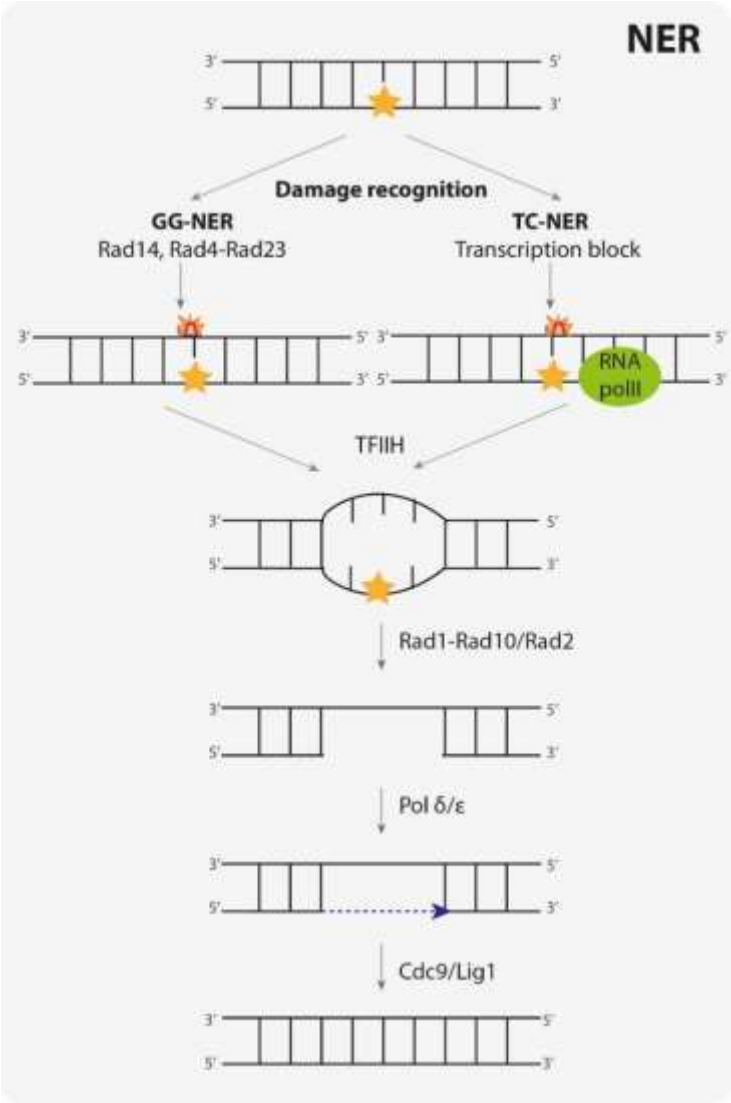

Figure 1.5. Simplified scheme of NER in budding yeast.

Left: In the GG-NER pathway, damage is recognised with the help of accessory proteins (Rad14, Rad4-Rad23, RPA). *Right:* During TC-NER, lesions are indirectly recognised during transcript elongation by RNA Pol II stalling. In both GG- and TC-NER, TFIIH associates and opens the double helix around the lesion. Then around 26 bp of the damaged region is excised (Rad1–Rad10 and Rad2). The resulting gap is then filled-in (Pol δ or ε) and ligated (Cdc9) (modified from Volker, et al., 2001).

1.1.5 Evidence of the re-start of stalled replication forks

The re-start of DNA replication requires a *de-novo* re-priming mechanism downstream of the stalled fork. Remarkably, an increasing number of studies supported that prokaryotes and eukaryotes can re-stablish replication of a stalled or collapsed replication fork. For instance, by combining electron microscopy (EM) and two-dimensional (2D) gel electrophoresis, Lopes, et al., (2006) revealed that UV-irradiated NER-deficient yeast cells accumulated small ssDNA gaps. Possibly resulting from re-priming events downstream of the lesions on both leading and lagging strands, which suggested damage bypass and *de-novo* priming events. Furthermore, by using a reconstituted replication system assembled with purified proteins and a plasmid containing a single bulky adduct, it was discovered

mechanistic distinctions between the lagging- and leading-strand re-priming. Whilst lagging-strand lesions are bypassed immediately, and replication is re-started downstream of the damaged region, the leading-strand polymerase stalling significantly impacts replication fork progression affecting priming events downstream of the lesion. Interestingly, it was also shown that the core replisome alone could re-establish leading-stand synthesis; nonetheless, this is inefficient and highly depends on Pol α and RPA availability (Yeeles, et al., 2015; Taylor & Yeeles, 2018).

Other studies have shown that yeast was viable without Pol ε catalytic domains, which proposed that at least one other polymerase can replicate the leading strand when Pol ε is inactive. Moreover, Pol δ function was found to be indispensable for cell viability indicating a significant role of Pol δ during DNA replication (reviewed in Johnson, et al., 2015). These observations have suggested a critical role for Pol δ in replicating the leading strand in the absence of Pol ε. To examine this hypothesis, Garbacz et al., (2018) determined the contribution of each replicase by mapping the incorporation of ribonucleotides of variants of the three major replicases, Pol α (Pol1 L868M), Pol δ (Pol3 L612M), and Pol ε (Pol2 M644G) in budding yeast. They observed that the mutant Pol3 L612M had a high bias in the bulk leading-strand replication, thereby suggesting that Pol δ replicates both the leading and lagging strands across the genome when Pol ε is absent. Even though Pol ε uncoupling is prevented during unperturbed replication, uncoupling might occur under replication stress conditions, leading to the theory that Pol δ could have a pivotal role in re-establishing coupled leading-strand synthesis (Yeeles, et al., 2017). In fact, some suggested that in the presence of damage, Pol δ briefly takes over leading-strand synthesis until the 3′- end of the leading strand is reconnected with the advancing Pol ε/CMG complex (Garbacz, et al., 2018; Aria & Yeeles, 2019).

1.1.6 Systems to study DNA replication stress

DNA replication stress occurs randomly in the genome, which restricts the precise understanding of the temporal and spatial events of the initial response of the replisome to lesions. A variety of tools have been created to tease apart events resulting from replication through damaged DNA, and the most relevant for this study are described below.

1.1.6.1 Global DNA damage

Most studies aiming to understand how DNA replication deals with perturbations have involved the treatment of cells with damaging agents such as MMS and UV irradiation. These damaging agents can cause a broad spectrum of DNA lesions, which are unknown in some cases. In MMS-treated cells, for instance, DNA replication stress is induced globally (Leman & Noguchi, 2013). The most predominant lesions are base mispairing and replication blocks, which are repaired mainly by BER. Moreover, there is direct evidence that MMS also causes double-strand breaks, which induces both inter- and intra-chromosomal recombination repair responses (reviewed in Lundin, et al., 2005). Even though non-well characterised lesions have constrained the clear understanding of the response of the replisome to

damage, *in vitro*-based cell studies involving genome-wide induction of replication fork (RF) stalling have provided unique hints about the molecular interactions and undergoing mechanisms. In fact, powerful findings have been generated in combination with cutting-edge technologies. A proposed mechanism to explain how damage is handled in cells was obtained by using high-throughput fluorescence microscopy, in which the distribution of RPA was used as a marker for ssDNA on MMS-treated yeast cells. Here, Wong, et al. (2020) suggested that lesions in the replication template generally do not cause fork uncoupling. Still, forks recover efficiently by re-priming without prolonged stalling or accumulation of ssDNA. Furthermore, they stated that damaged daughter-strand gaps, instead of stalled replication forks, are responsible for activating the post-replication repair checkpoint.

Enzymatic approaches to inflict lesions to DNA avoid the problems associated with the uncertainty of the identity of the mutagenic lesion intrinsic to chemical or environmental mutagens. Using a protein with defined activity, such as the altered human uracil DNA glycosylase (hereafter referred as UDG^{Y147A} or UDG*), which removes thymines from DNA, is able to induce the formation of AP sites during normal cell growth (Kavli, et al., 1996). The mutational effects of AP sites have been investigated in *S. cerevisiae* using UDG*, and it was found that the genome-wide formation of abasic sites leads to the saturation of repair pathways, a situation that represents a limitation for this kind of approach (Auerbach, et al., 2005).

1.1.6.2 Site-specific DNA-protein blocks

As mentioned above RF stalling can occur when the DNA replication machinery encounters DNA lesions, secondary structures, or DNA-bound proteins. However, in some cases, RF arrests are programmed by cells in order to orchestrate DNA replication with other cellular processes (Bastia & Zaman, 2014). Many of these well-characterised systems have been exploited to further elucidate how site-specifically stalled RFs are processed *in vivo*.

A tool to induce a DNA replication barrier is the bacterial system, Tus-Ter, which takes advantage of the high-affinity binding of the terminator protein, Tus, to a specific DNA sequence called Ter (Larsen, et al., 2014). The Tus-Ter system has been successfully reconstituted in yeast and mammalian cells (Larsen, et al., 2014; Willis, et al., 2014), and it has proven that when introduced into the genome induces cellular responses similar to those that occur when the replisome encounters certain types of DNA damage (Larsen, et al., 2014). Alternatively, the replication termination sequence 1 (*RTS1*) model has been used as a replication fork barrier tool to probe mechanisms of recombination-dependent re-start (Lambert, et al., 2005). In fission yeast, *RTS1* ensures mating-type switching by regulating the direction of replication of telomeres to centromeres (Eydmann, et al., 2008). Although this protein block system has enabled recapitulation of events occurring in the genome, they differ from events occurring when replication forks encounter damage.

1.1.6.3 Inducing the formation of site-specific abasic sites

By combining the advantages of the methods previously described to study RF stalling events, the Ulrich laboratory has developed a system that allows the induction of site-specific and well-defined lesions (AP sites) in the genome of *S. cerevisiae*. The system referred to as the "AP site system" comprises an inducible construct that regulates the amount of induced damage *in vivo*. Additionally, its experimental layout allows the real-time study of the immediate response of the replisome to damage. As a result, the AP site system can unravel current discrepancies about DNA replication stress. Moreover, the AP site system in combination with molecular biology and deep-sequencing tools such as quantitative-PCR (qPCR) and Chromatin Immunoprecipitation sequencing (ChIP-seq) can identify different factors recruited at the site of damage. Importantly, this tool can extend further studies to the context of chromatin remodelling and post-replication repair pathways.

To target the damage to a specific location in the genome, we took advantage of the bacterial Tet Operator (TetO) sequence and its binding partner the protein Tet Repressor (TetR). Our system integrates an array of 44 TetO sequences (TetO44; ~1.5 kb) in chromosome VI of budding yeast, and by fusing UDG* to TetR, we target lesions specifically to the TetO44 array. The induction of the TetR construct is controlled by the *GAL1/10* promoter. In addition, this construct is fused to an auxin-inducible degron (AID*) to avoid any interference with the replisome passage (Figure 1.6). Finally, to minimize the processing of AP sites into strand breaks, our system was designed in a strain background devoid of the following BER proteins: Apn1, Apn2, Ntg1 and Ntg2 (Elizabeth Colby, unpublished).

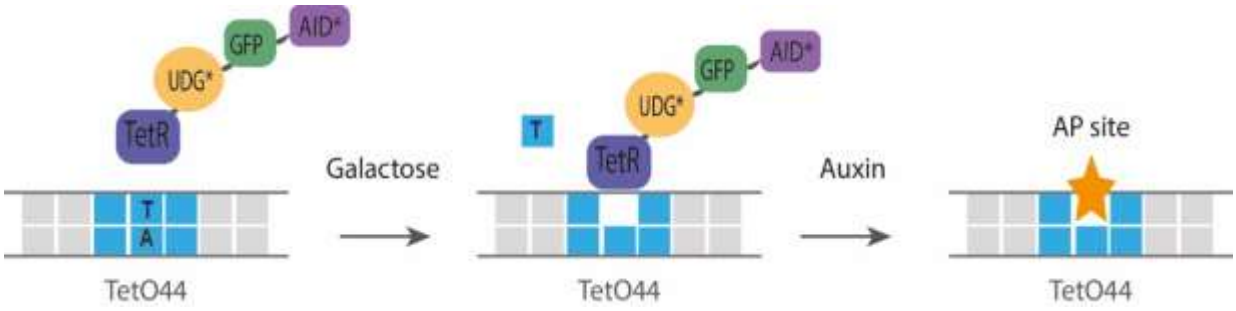

Figure 1.6. Apurinic/apyrimidinic (AP) site system.
The AP site system targets the formation of abasic sites to the Chromosome VI of budding yeast by integrating 44 TetO sequences (TetO44 array). At the same time, the binding partner TetR is fused to a mutated version of the human uracil DNA glycosylase (UDG*) that removes thymines from DNA. To avoid any interference with the replisome passage and control the amount of damage a degron (AID*) is also bonded to TetR. To minimise processing of AP sites into strand breaks, the yeast background lacks the base excision repair (BER) proteins Apn1, Apn2, Ntg1 and Ntg2.

1.1.7 Next-Generation Sequencing tools to study DNA replication

Recently, various high-throughput approaches based on Next-Generation Sequencing (NGS) have been developed to identify missing links between replication, chromatin regulation and genome stability. Below I described tools relevant to this study.

1.1.7.1 Mapping the incorporation of ribonucleotides in genomic DNA

The establishment of methods for mapping the position of ribonucleotides across the genome is based on the mutation on individual DNA polymerases that allow more frequent incorporation of ribonucleotides (Brown & Suo, 2011). The elevated ribonucleotide incorporation into genomic DNA was found to be useful for marking newly synthesised strands *in vivo*. This approach, termed Polymerase usage sequencing (Pu-seq), allows to map the division of labour between Pol α, Pol ε and Pol δ genome-wide (Keszthelyi, et al., 2015; Daigaku, et al., 2015), and there are three other variations of this method (Reijns, et al., 2015; Koh, et al., 2015; Clausen, et al., 2015;).

The mapping of ribonucleotides is achieved by using parallel yeast cells deleted for *RNH201*, the gene encoding the catalytic subunit of ribonuclease H2 (RNase H2). Each strain contains equivalent additional mutation in the gene encoding the catalytic subunit of the main DNA polymerases (*pol1 L868M*, *pol2 M644G*, or *pol3 L612M*), which reduces the efficiency of the steric gate and increases the rate of ribonucleotide incorporation. Once cell cultures are harvested and DNA is extracted, ribonucleotides are nicked at their 3'- or 5'-end by using alkali hydrolysis or the endonuclease RNase HII, respectively. The resulting ends are then captured for NGS library preparation by either GLOE-Seq, emRibo-Seq, Ribose-Seq or HydEnd-Seq. Finally, by comparing the number of ribonucleotide incorporation sites between the Watson and Crick strands and between the parallel strains, a map of Pol α, Pol ε and Pol δ usage across the entire genome can be constructed (Keszthelyi, et al., 2015; Reijns, et al., 2015; Daigaku, et al., 2015; Koh, et al., 2015; Clausen, et al., 2015).

Interestingly, Pu-seq has been applied in combination with the replication barrier *RTS1* in *S. pombe*. Naiman and colleagues (2021) showed that for forks re-started by homologous recombination (HR), both leading and lagging strands are synthesised by Pol δ, even further confirming that Pol δ plays an important role in RF stalling.

1.1.7.2 Genome-wide mapping of single-stranded breaks

The high-resolution identification, mapping and quantification of SSBs can provide a profound understanding of a variety of fundamental biological processes regarding their genesis and as well as their repair (Zilio & Ulrich, 2021). GLOE-Seq (Genome-wide Ligation of 3'-OH Ends), for instance, is a protocol designed for efficiently recovering SSBs genome-wide. GLOE-seq uses an initial step of heat-denaturation on the nicked DNA to make the 3'-OH termini accessible for the ligation of a biotinylated proximal adaptor. The biotinylated ssDNA is fragmented and converted onto dsDNA before the ligation of a distal adaptor. The PCR amplification of these fragments with barcoded primers allows their sequencing on an NGS platform (Sriramachandran, et al., 2020). Notably, GLOE-seq has proven to be superior

in terms of coverage and signal-to-background ratio in comparison to other methods such as emRibo-Seq in which fragmentation and ligation of the distal adaptor precede endonuclease treatment and denaturation (Ding, et al., 2015, Reijns, et al., 2015). As briefly mentioned above, GLOE-Seq has been used to capture the resulting 3'-OH ends of RNase HII treatment for the mapping of embedded ribonucleotides in genomic DNA. In addition, GLOE-Seq can be employed for capturing nicks produced by endonucleases, such as the apurinic endonuclease, Ape1, enabling the genome-wide analysis of abasic sites (Sriramachandran, et al., 2020).

1.1.7.3 Capturing newly synthesised DNA

The incorporation of nucleoside analogues like 5-bromo-20-deoxyuridine (BrdU) into DNA is a powerful tool for *in vivo* studies of DNA replication. The analysis of the immunoprecipitation (IP) of BrdU-labelled DNA by NGS has allowed the genome-wide, sequence-specific tracking of replication and replication fork dynamics under DNA lesions and replication stress (Viggiani, et al., 2010). In yeasts, such studies have been eased by introducing vectors expressing thymidine kinase and a nucleoside transporter to reconstruct the thymidine salvage pathway, which is absent in budding and fission yeast (Haye-Bertolozzi & Aparicio, 2017). Moreover, the use of the thymidine analogue, 5-ethynyl-2-deoxyuridine (EdU), helped to develop a novel NGS method in *S. cerevisiae* (Petryk, et al., 2016). The method named OK-Seq (Okazaki fragments sequencing) permits a high resolution and quantitative analysis of replication forks initiation, progression and termination. OK-Seq relies on the isolation, purification and sequencing of newly synthesised Okazaki fragments marked with EdU, and can reveal the proportions of rightward and leftward moving forks throughout the genome.

1.1.7.4 Detection of DNA-protein interactions

The genome-wide mapping of DNA-protein interactions is essential for almost all aspects of cellular function, including DNA replication. The main tool for studying such mechanisms is ChIP-Seq, which utilises antibodies to detect a protein of interest (PoI) and enriches the DNA fragments bound to these proteins (Park, 2009). Since ChIP-Seq provides the actual DNA sequences of the precipitated fragments, the data obtained is very accurate and of high resolution (Mundade, et al., 2014). An alternative to study DNA-protein interactions is by using the technique CUT&RUN (Cleavage Under Targets and Release Using Nuclease), an enzyme-tethering method in which a protein A-Micrococcal Nuclease (pA-MNase) fusion is targeted to genomic regions via a PoI's antibody. CUT&RUN, in addition to retaining the advantages of ChIP-Seq, overcomes some of its limitations as this greatly reduces background relative to ChIP and requires lower starting cell numbers (Skene & Henikoff, 2017).

Yu and colleagues (2014), by fusing BrdU-Seq and ChIP-Seq, developed a unique method called eSPAN-Seq (enrichment and sequencing of protein-associated nascent strand DNA sequencing). eSPAN-Seq is a potent tool to discern whether a protein is enriched at the

leading or lagging strand. By using this approach in *S. cerevisiae*, the strand specificities of Pol ε and Pol δ were confirmed and it also found that the PCNA clamp is more abundant on the lagging strand than the leading strand.

1.2 DNA-damage tolerance and Post-Replication repair

The genome is particularly vulnerable during S-phase, as many types of lesions block replication fork progression. Most DNA damage removal occurs via conserved excision repair pathways such as NER or BER. Because these pathways need a complementary strand as a template for re-synthesis of the excised region, such repair only operates when the lesion is present in dsDNA. When damage escapes repair and is encountered by the replication fork, it can potentially block the progression of DNA synthesis forming gaps and leaving ssDNA, so at this point, NER or BER are no longer productive. Hence, cells have engineered DNA-damage tolerance (DDT) pathways that become critical during S- and G2-phases and allow the completion of replication and timely cell-cycle progression (Ulrich, 2011; Lehner & Jinks-Robertson, 2014). In this section, I refer to the molecular mechanism of post-replication repair (PRR) and potential factors involved in these pathways.

1.2.1 Activation of Post-Replication Repair

The PCNA ring plays a critical role during DDT; in fact, PRR is directly controlled through PCNA covalent modifications by ubiquitin (Hoege, et al., 2002). As mentioned at the beginning of this chapter, when replication fork stalling uncouples the activity of the replicative helicase and polymerases, the helicase continues to unwind DNA while the polymerases remain stalled, resulting in the accumulation of stretches of ssDNA (Paulsen & Cimprich, 2007). RPA-coated ssDNA activates the Rad6 pathway, which is initiated by the recruitment of Rad18, a PCNA-directed E3 ubiquitin ligase, to the vicinity of the DNA lesion, where it associates with the E2 ubiquitin-conjugating enzyme, Rad6. The Rad6-Rad18 complex catalyses the mono-ubiquitylation of PCNA on the conserved residue, Lys-164, which recruits translesion synthesis polymerases to substitute the activity of the replicative polymerases at the stalled fork (Hoege, et al., 2002; Stelter & Ulrich, 2003).

On the other hand, Rad18 can further associate with another E3 ubiquitin ligase, Rad5, allowing the recruitment of the Mms2-Ubc13 dimer, an ubiquitin-conjugating enzyme, to the vicinity of stalled replication forks. The binding of Mms2-Ubc13 promotes the extension of the K164 mono-ubiquitylated PCNA with K63-linked chains (Ulrich & Jentsch, 2000). The Rad5-mediated poly-ubiquitylation of PCNA directs lesion bypass using the template switching (TS) pathway (Haracska, et al., 2004; Moldovan, et al., 2007) (Figure 1.7).

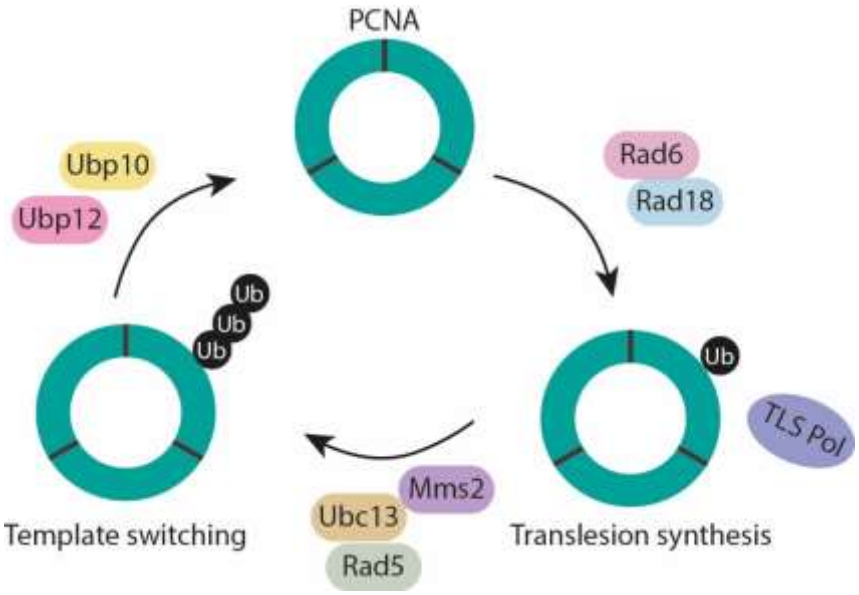

Figure 1.7. PCNA ubiquitylation and activation of DNA-damage bypass.
Post-translational modification of PCNA determines the DNA-damage tolerance pathway that cells use to process a DNA lesion encountered during genome replication. The mono-ubiquitylation triggers TLS whilst poly-ubiquitylation activates TS.

1.2.2 Translesion synthesis and polymerases

Translesion synthesis employs specialised DNA polymerases that transiently replace replicative polymerases at stalled forks to bypass lesions. The best-characterised TLS polymerases in vertebrates are Pol η, Pol ι, Pol κ , Rev1 (Y-family), and Pol ζ (B-family), of which Pol η (Rad30), Rev1 and Pol ζ have been identified in budding yeast. The unique structure of TLS polymerases allows them to synthesise DNA across the lesion. They have a relatively non-restrictive active site that makes them more capable of accommodating damaged or distorted templates than replicative DNA polymerases. Moreover, TLS polymerases lack the proofreading exonuclease domains, which is critical for accurate DNA synthesis. Therefore, TLS can confer damage tolerance but with much higher error rates (Gao, et al., 2017) (Figure 1.8).

In some cases, lesion bypass can be accomplished by one TLS polymerase in a single step. For instance, Pol η acts as an inserter and extender to bypass a CPD. Nevertheless, replication through many lesions requires two steps and two different polymerases, one for nucleotide insertion and the other for extension (reviewed in Prakash, et al., 2005; Sale, 2013). Certain polymerases appear to be predominantly good at the insertion step, such as Pol ι and Rev1, while others are notably more efficient at the extension step, Pol ζ, and to a lesser extent Pol κ (Johnson, et al., 2000). Pol ζ, especially, plays a crucial role even where bypass can be carried out by a single polymerase *in vitro*, as it has been seen to promote an error-free TLS through a number of lesions (Shachar, et al., 2009).

Next, I discuss additional features of the main TLS polymerases.

1.2.2.1 Rev1

The catalytic activity of Rev1 differs from other TLS polymerases, as this is a deoxycytidine monophosphate (dCMP) transferase capable of inserting cytosines opposite an abasic site, and at limited number of adducts, such as N^2-adducts of guanine, and template guanines (reviewed in Lehmann, et al., 2007). Moreover, Rev1 has been involved in replicating sequences prone to forming DNA secondary structures such as G-quadruplexes (Sarkies, et al., 2010). Importantly, Rev1 also has a non-catalytic function responsible for coordinating TLS through the interaction of its well conserved extreme carboxyl terminus with other TLS polymerases, PCNA and ubiquitin (D'Souza, et al., 2009).

1.2.2.2 Pol η

Pol η, also known as Rad30 in yeast, incorporates nucleotides opposite UV-induced CPDs in a relatively error-free manner (McCulloch, et al., 2004). Pol η does not appear to be able to carry out TLS past other major photoproducts such as the pyrimidine-pyrimidone (6-4) (6-4PP). However, it can act on a limited number of different types of damage *in vitro* with less efficiency, like in cisplatin lesions. Moreover, Pol η has not shown to play a role in replicating past 8-oxoguanine or abasic sites (reviewed in Lehmann, et al., 2007). Interestingly, it has been seen that in the absence of Pol η the bypass of CPDs results in the inaccurate action of Pol κ or Pol ι in cooperation with Pol ζ (Ziv, et al., 2009).

1.2.2.3 Pol ζ

It has been proposed that the major role of Pol ζ is to extend rather than insert nucleotides from mismatched primer termini and abnormal structures. Pol ζ lacks the proofreading exonuclease activity that other B-family polymerases, like Pol δ and Pol ε, have. However, Pol ζ has higher fidelity in comparison to Y-family polymerases (Johnson, et al., 2000). The core of Pol ζ is composed of two subunits: Rev3, the catalytic subunit, and Rev7, a subunit that stimulates Rev3 activity (Nelson, et al., 1996). Studies on Rev3-deficient cells have revealed that Pol ζ is involved in the chromosomal maintenance of DNA in addition to the tolerance of various sources of DNA damage, such as cisplatin and UV-light, by TLS and HR (Sonoda, et al., 2003).

1.2.3 Template switching

Despite the efforts to understand the molecular mechanisms of the error-free DDT pathway, template switching, all the details and effector proteins involved in this process have not been completely elucidated. TS presumably utilises the newly synthesised daughter strand as a template to bypass damaged DNA, and two possible mechanisms have been suggested.

1.2.3.1 Replication fork regression model

The replication fork regression model of error-free DDT is thought to be specific when DNA damage is encountered in the leading strand. In this case, the lesion stalls the replication fork, resulting in uncoupling the leading strand synthesis. Meanwhile, the lagging strand

synthesis continues way past the blocked nascent leading strand generating ssDNA. Single-stranded DNA is the substrate for the Rad6-Rad18 complex, which recruits Rad5 to the site of damage. Rad5 unwinds and anneals the parental and nascent strands into a chicken-foot-like structure. The opposite nascent strand then serves as a template to restore the information of the damaged parental strand. The synthesis resumes beyond the point of the lesion, and normal replication is restored (Xu, et al., 2015) (Figure 1.8).

The replication fork regression model implies that this could be activated simply by Rad5 helicase activity and does not necessarily require the poly-ubiquitylation of PCNA for error-free DDT directly. Instead, the poly-ubiquitylated PCNA might function to inhibit the association of unwanted factors to the site of damage that could prevent the regression of the replication fork (Xu, et al., 2015). However, the role of Rad5 in this model has been challenged, Sogo and colleagues (2002) observed by EM that Hydroxyurea-treated (HU) Rad53-deficient cells exhibit extensive single-stranded gaps, hemi-replicated intermediates, and accumulation of Holliday junction alike structures, which suggested the requirement of a Rad53-dependent stabilisation of the replication fork in the reversal mediated by Rad5.

1.2.3.2 Template switch model

The alternative template switch model for error-free DDT relies on recombination events for strand exchange. This model involves factors responsible for homologous recombination, and similarly to the previous model, it results in the formation of specific intermediate structures, such as double Holliday junctions. At first, the template switch model requires strand invasion events, mediated by Rad51, where the switch of the blocked nascent strand template to the newly synthesised sister chromatid promotes the continuation of replication. In this manner, replication can proceed past the damaged DNA region in the parental stand. Then nascent strand can switch back to its original parental strand, and finally normal replication is resumed (Xu, et al., 2015; Branzei & Szakal, 2016) (Figure 1.8).

Recently the Ulrich laboratory proposed a model for the initiation of template switching that depends on the combined action of the nuclease Exo1 and the helicase Pif1. This model states that the accumulation of daughter-strand gaps left behind the replication fork is expanded at their 5′-junction via Exo1's exonucleolytic activity in cooperation with the 9-1-1 checkpoint clamp. Additionally, the recruitment of Pif1, via its interaction with PCNA, contributes to gap expansion at the 3′-junction by generating ssDNA 3' flaps that subsequently undergo nuclease cleavage. The expanded gaps are thought to facilitate the invasion of the damaged strand and thus initiate TS (García-Rodríguez, et al., 2018B).

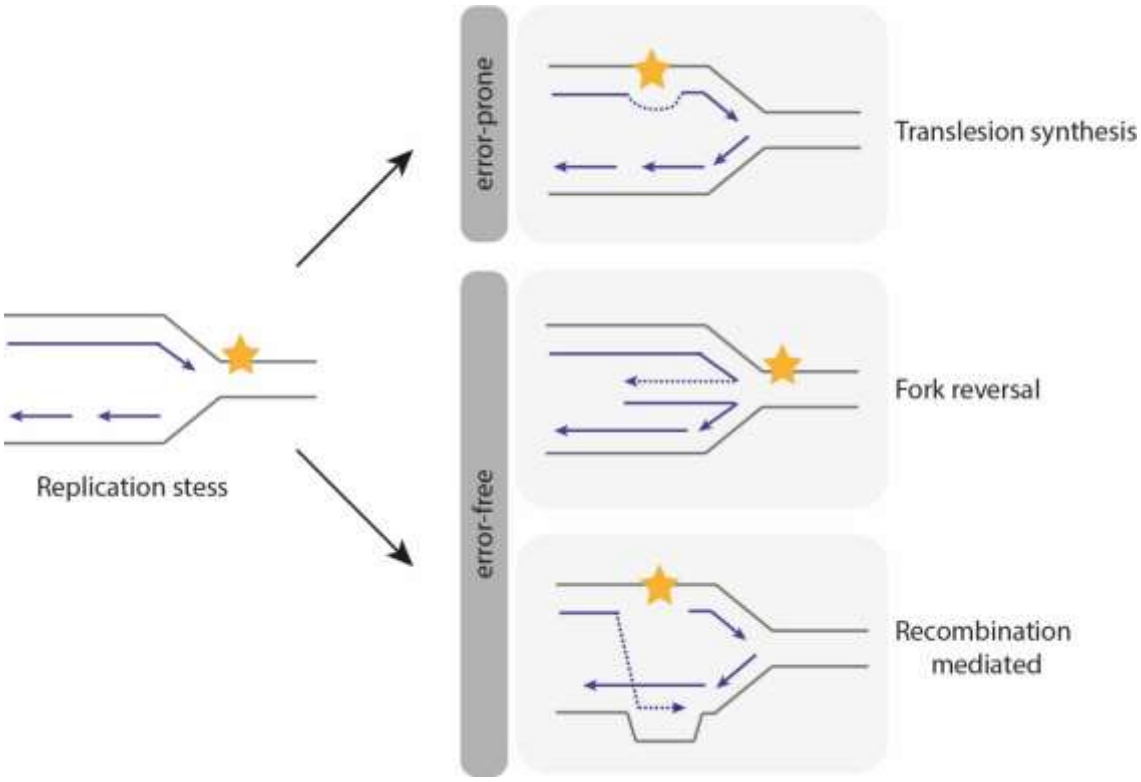

Figure 1.8. DNA-damage tolerance pathways.
The error-prone DDT pathway, translesion synthesis, bypasses lesions with the aid of specialised polymerases. The error-free mechanism is predicted to occur via a replication fork reversion through Rad5 helicase activity, resulting in a chicken-foot structure, or template switching may occur via a homologous recombination-factor mediated strand invasion to switch the template to the newly synthesised sister chromatid.

1.2.4 Factors influencing the choice of a DDT pathway

The Rad6 pathway is essential for cells to tolerate and survive DNA damage and normally operates during S-phase. However, two studies using the temporal expression of Rad18 or Pol η discovered that this pathway could be delayed without significantly affecting cell viability until G2/M-phase of the cell cycle, suggesting that PRR and bulk DNA replication can be separated without compromising their function (Karras & Jentsch, 2010; Daigaku, et al., 2010). Furthermore, electron microscopy studies of UV-irradiated cells suggested that the accumulation of gaps behind the replication forks on both the lagging and leading strands are filled-in a manner that depends on DNA-damage tolerance pathways (Lopes, et al., 2006).

DNA-damage tolerance can proceed through one of the two PRR mechanisms. In terms of maintaining genome stability, translesion synthesis and template switching are pivotal for cells to tolerate and survive DNA damage. TS has greater advantages than TLS as it does not induce further mutations. TLS, on the other hand, has been linked to increasing mutation rates in cells (reviewed in Gao, et al., 2017). However, how cells choose between TLS and TS, and by extension, the switch from mono- to poly-ubiquitylation of PCNA, is not entirely understood. Some studies have implied that poly-ubiquitylation on PCNA is predominant

with increasing amounts of DNA damage and the accumulation of RPA-coated ssDNA (Chang, et al., 2006). Moreover, since Rad18 and Rad5 interact in yeast through the domains required for their homodimerization, one intriguing model suggests that the amount of Rad18 bound to RPA-ssDNA could affect the binding of Rad5. The accumulation of RPA-ssDNA might result in more Rad18-RPA interactions and less Rad18 homodimerization, thereby allowing Rad5 to better compete for binding to Rad18 and promote poly-ubiquitin chains formation (reviewed in Chang & Cimprich, 2009). Notably, the chemical and structural nature of the lesion is frequently a determinant of the repair mechanism employed, which appears to be the case also for DDT pathways. Daigaku, et al., (2010), for instance, observed that UV-irradiation-induced lesions were bypassed preferentially with TLS, whereas TS functions as a back-up.

1.2.5 Tools and methods to study DNA damage bypass

1.2.5.1 The quantification of PRR tracts

Traditional analyses of PRR, such as DNA combing, involve the measurements of the amount and size of DNA synthesised after cells are exposed to genotoxic agents. However, to what extent the newly synthesised DNA is due to normal replication and DDT processes is not clear. To address this issue, the Ulrich Laboratory has established a system in *S. cerevisiae* capable of separating the bulk genome replication from PRR by using a doxycycline-repressible construct for Rad18 (Tet-Rad18 system) (Daigaku, et al., 2010). To detect PRR-associated DNA synthesis, G1-phase synchronised cells are exposed to a damaging agent (MMS, UV-light or HU), and allowed to pass through S-phase. When cells have reached the G2/M arrest, Rad18 is re-expressed in the presence of the thymidine analogue, BrdU. Capturing of the newly synthesised DNA marked by BrdU followed by next-generation sequencing allows the detection and quantification of PRR tracts (Figure 1.9).

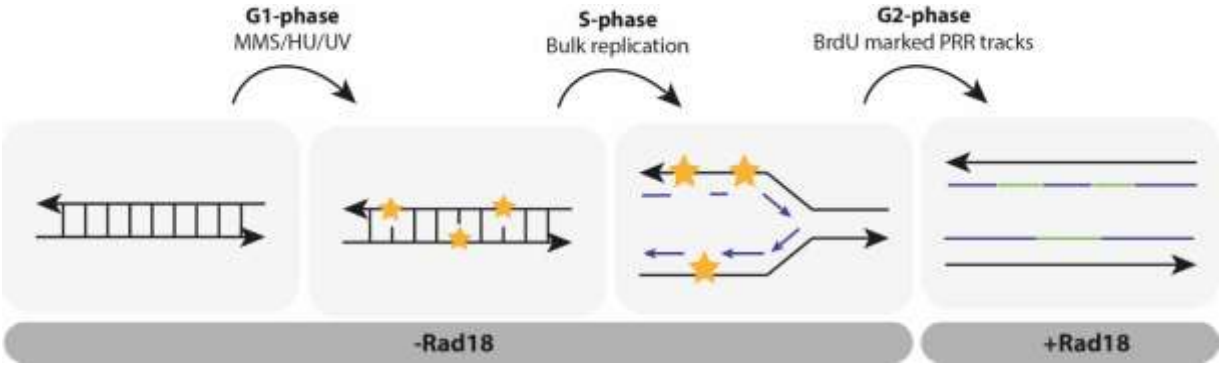

Figure 1.9. Capturing post-replication repair tracts
Tet-Rad18 cells synchronised in G1-phase are exposed to damaging agents before being released into S-phase to permit bulk replication. During G2-phase, Rad18 is re-expressed in the presence of BrdU that mark PRR tracts. The genome-wide distribution of PRR tracts is mapped by BrdU-Seq.

1.2.5.2 The *in vitro* reconstitution of replication forks

Yeeles and colleagues (2015) have successfully engineered a sophisticated *in vitro* tool with purified yeast proteins that resembles an *in vivo* eukaryote replication fork. This tool defines the minimum set of proteins and small-molecule co-factors required to initiate origin-dependent DNA replication and utilises plasmids or linearized DNA as a template to study replication events. This tool has mainly been used to test the role and importance of several proteins of interest in DNA synthesis. For instance, Kurat, et al., (2016), by using a fully chromatinised template, showed that chromatin promotes the regular priming of lagging-strand synthesis by facilitating Pol α function at replication forks. Furthermore, setting the system with several concentrations of the DNA polymerases indicated that Pol δ plays a role in establishing leading-strand synthesis before Pol ε engagement (Yeeles, et al., 2017). Similarly, the integration of a CPD on the template revealed that the replisome alone could re-establish leading-strand synthesis downstream of the lesion by re-priming events. TLS has also been investigated through this system, Guilliam & Yeeles (2020) exploited the efficiency of Pol η in bypassing a CPD and demonstrated that this polymerase facilitates bypass on the leading-strand to re-start uncoupled replication. Additionally, they found that Pol η permits gap filling on the lagging strand.

Even though this potent tool has clarified multiple questions about DNA replication, a disadvantage of using *in vitro* tools is that they may not fully duplicate what occurs *in vivo*. In fact, other factors are not taken into account, such as intermediates of metabolism, which might highly influence the response of the replisome.

1.3 Aims of thesis

Replisomes frequently encounter a variety of obstacles, which can slow them down or stall them altogether. Cells must overcome these obstacles and complete genome replication to maintain a timely cell cycle progression and cell viability. Significant studies have shown evidence of the importance of DNA-damage tolerance pathways during S- and G2-phases, when the replication forks stall. However, the current challenge is to understand the site-specific and quantitative response of the replisome to damage *in vivo*.

During this project I aimed to comprehend how the leading and lagging strands are synthesised in the presence of polymerase-stalling lesions. To assess this, I utilised the AP site system to create site-specific abasic sites in chromosome VI of budding yeast, moreover, the highly controlled experimental layout allowed me to induce DNA damage in a timely manner and closely follow cell cycle progression. Firstly, by additionally integrating the polymerase usage assay, I intended to map the distribution of labour of the three main DNA polymerases in the damaged region. Secondly, in combination with the NGS approaches, ChIP-Seq and BrdU-Seq, I aimed to understand what factors are recruited to the site of damage and how daughter strands are synthesised in the presence of abasic sites.

Furthermore, as part of this project, I focused on DNA-damage tolerance pathways and I aimed to understand how post-replication repair is organised temporally and spatially genome-wide, and how cells maintain the balance between translesion synthesis and template switching. To answer these queries, I used the system previously established in the Ulrich laboratory to quantify post-replication repair tracts jointly with BrdU-Seq on MMS-treated cells.

Chapter 2: Materials and Methods

2.1 Media

2.1.1 Bacterial growth media

The Media lab at the Institute of Molecular Biology prepared the Luria Broth (LB) liquid medium and agar plates.

Luria Broth: mix 0.5% (w/v) bacto-tryptone, 0.25% (w/v) bacto-yeast extract and 170 mM NaCl, adjusted to pH 7.0 with NaOH, sterilised by autoclaving. For the selection of transformed bacterial cells, ampicillin (Amp) or kanamycin (Kan) was added at a final concentration of 100 µg/mL (stock: 100 mg/mL and 30 mg/mL, respectively). LB was stored at 4^{o}C.

2.1.2 Yeast growth media

Yeast Peptone (YP) media, 2% (w/v) YPD agar, 4% (w/v) water-agar, and 20% (w/v) glucose were prepared by the Media lab.

20% glucose (w/v): 20 g of glucose was dissolved in 100 mL of filtered water, sterilised by autoclaving.

20% raffinose (w/v): 20 g of raffinose was dissolved in 100 mL of filtered water, sterilised by autoclaving.

20% galactose (w/v): 20 g of galactose was dissolved in 100 mL of filtered water, sterilised by autoclaving.

Yeast Peptone (YP): 1% (w/v) yeast extract and 2% (w/v) peptone, sterilised by autoclaving.

Yeast Peptone with glucose, raffinose or galactose (YPD, YPRaff or YPGal): 20% glucose, raffinose or galactose was diluted into YP media to a final concentration of 2%.

YPD Agar Plates: 2% YPD agar was melted, poured into petri dishes and allowed to solidify.

2.5x Synthetic complete (SC) media: 12.5 g of ammonium sulfate, 4.25 g of yeast nitrogen base (without amino acids and ammonium sulfate), and 5 g of the desired SC powder were mixed in 1 L of water, stirred for 30 min, and autoclaved.

Liquid SC medium: 200 mL of 2.5x SC medium, 250 mL of distilled water, and 50 mL of 20% (w/v) glucose were mixed.

5-fluoroorotic acid (5-FOA) SC agar: 200 mL 2.5x SC medium, 250 mL of melted 4% water-agar, 50 mL 20% glucose, 8 mL uracil (2 mg/mL), 8 mL adenine (2 mg/mL), and 500 mg of 5-fluoroorotic acid, were mixed, poured into petri dishes and allowed to solidify

Sporulation medium: 1% (w/v) potassium acetate (KOAc) in water, filter-sterilised.

Pre-sporulation medium: 0.8% (w/v) bacto-yeast extract, 0.3% (w/v) bacto-peptone and 10% glucose were mixed and sterilised by autoclaving.

Special sporulation medium: 1% (w/v) potassium acetate, 0.1% (w/v) bacto-yeast extract and 0.05% glucose were mixed and sterilised by autoclaving.

Sorbitol-Tris-EDTA (STE) Buffer: 1.2 M sorbitol, 25 mM Tris-HCl pH 8.0 and 25 mM EDTA (pH 8.0).

STE-Zymolyase solution: 37 µL STE buffer, 2 µL 1M DTT and 1 µL Zymolyase 20T (20 mg/mL) were mixed and prepared fresh.

2.2 Strains

2.2.1 *Escherichia coli* strains

Table 2.1. *Escherichia coli* strains used in this study

ID	Name	Genotype	Purpose	Supplier
14	Top Ten	*F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK λ– rpsL(StrR) endA1 nupG*	Standard Cloning	Invitrogen

2.2.2 *Saccharomyces cerevisiae* strains

Table 2.2. *Saccharomyces cerevisiae* strains used in this study

ID	Name	Genotype	Source
1	DF5	Mat-a/α, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*	Jentsch strain collection
3	DF5	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*	Jentsch strain collection
2222	tetO7-RAD18 pCM244, BrdU-inc	Mat-a, *his3-Δ200,leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(RAD18)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*	Y. Daigaku
2224	tetO7-*RAD18* pCM244, *TLSΔ*, BrdU-inc	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, *rev1*::*URA3*, *rev3*::KanMX, *rad30*::*HIS3*, BrdU-inc::*URA3*	Y. Daigaku
2226	tetO7-*RAD18* pCM244, *ubc13Δ*, BrdU-inc	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, *ubc13*::trp1-1, BrdU-inc::*URA3*	Y. Daigaku

2821	W303 *RAD5+*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	E. Colby
3060	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*	E. Colby
3062	*BERΔ*, Gal-NLS-TetR-GFP-AID*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*	E. Colby
3439	tetO7-*RAD18* pCM244, BrdU-inc, AFB2	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*HIS3*	N. Gralievska
3891	tetO7-*RAD18* pCM244, BrdU-inc, *rpb9Δ, rad26Δ*	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, *rpb9*::HphB, r*ad26*::natNT2	N. Gralievska
3987	tetO7-*RAD18* pCM244, BrdU-inc, *rpb9Δ*	Mat-a, *his3-Δ200,leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, *rpb9*::hphB,	N. Gralievska
4370	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, TetO44-*URA3*	Mat-a*, leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, TetO44-*URA3*::PES4IG	L. Batista
4406	*BERΔ*, Gal-TetR-UDG*-GFP-AID*, AFB2, TetO44-*URA3*, Rfa1-9myc, BrdU-inc	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, TetO44-*URA3*::PES4IG, RFA1-9myc::natNT2, BrdU-inc::*HIS3*	L. Batista
4442	*BERΔ*, G1tag-NLS-TetR-UDG*-GFP-AID*, TetO44-*URA3*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*, TetO44-*URA3*::PES4IG	L. Batista
4468	*BERΔ*, *rnh201Δ*, Pol3 L612M	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*	L. Batista
4469	*BERΔ*, *rnh201Δ*, Pol2 M644G	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *rnh201*::HIS3, *pol2 M644G::POL2*	L. Batista

4689	*BERΔ*, Pol3 L612M, *rnh201Δ*, TetO44-*URA3*	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol3 L612M*::*POL3*, *rnh201*::HIS3, TetO44-*URA3*::PES4IG	This study
4758	*BERΔ*, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2 M644G::POL2*, *rnh201*::HIS3, TetO44-*URA3*::PES4IG	This study
4759	*BERΔ*, Pol3 L612M, *rnh201Δ*, TetO44-*URA3*, G1tag-NLS-TetR-UDG*-GFP-AID*	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4760	*BERΔ*, Pol2 M644G, *rnh201Δ*,TetO44-*URA3* , G1tag-NLS-TetR-UDG*-GFP-AID* (High)	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2 M644G::POL2, rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4761	*BERΔ*, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*, G1tag-NLS-TetR-UDG*-GFP-AID* (Low)	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2 M644G::POL2*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4763	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2	Mat-a*, leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*	This study
4764	*BERΔ*, Pol3 L612M, *rnh201Δ*,TetO44-*URA3*, *ura3Δ*	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, *ura3*::natNT2	This study
4765	*BERΔ*, Pol3 L612M, *rnh201Δ*, TetO44-*URA3*, *ura3Δ*, G1tag-NLS-TetR-UDG*-GFP-AID* (low)	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol3 L612M*::*POL3*, rnh201::*HIS3*, TetO44-*URA3*::PES4IG, *ura3*::natNT2, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4766	*BERΔ*, Pol2 M644G, *rnh201Δ*,	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2*	This study

	TetO44-*URA3*, *ura3Δ*	*M644G::POL2*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, *ura3*::natNT2	
4767	*BERΔ*, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*, *ura3Δ*, G1tag-NLS-TetR-UDG*-GFP-AID* (High)	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2 M644G::POL2*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, *ura3*::natNT2, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4768	*BERΔ*, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*, *ura3Δ*, G1tag-NLS-TetR-UDG*-GFP-AID* (Low)	Mat-α*, leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *pol2 M644G::POL2*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4IG, *ura3*::natNT2, G1tag-NLS-TetR-UDG*-GFP-AID*::*LEU2*	This study
4779	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1, his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*	M. Garbacz
4850	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*	This study
4880	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, TetO44, Rfa1-9myc, BrdU-inc	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, TetO44::PES4IG, Rfa1-9myc::natNT2, BrdU-inc::*HIS3*	This study
4898	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, TetO44-*URA3*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, TetO44-*URA3*::PES4	This study
5081	*BERΔ*, *rnh201Δ*, *ura3Δ*, Pol3 L612M	Mat-α*, leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, apn1::KanMX, apn2::KanMX, ntg1::hphB, ntg2::hphB, *pol3 L612M*::*POL3*, rnh201::*HIS3*, *ura3*::natNT2	This study
5082	*BERΔ*, *rnh201Δ, ura3Δ*, Pol2 M644G	Mat-α, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, *rnh201*::*HIS3*, *pol2 M644G::POL2*, *ura3*::natNT2	This study
5083	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-	This study

	AID*, AFB2, *ura3Δ*	NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2	
5084	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ, rnh201Δ*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *rnh201*::*HIS3*	This study
5103	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol2 M644G	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol2 M644G::POL2*	This study
5104	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol3 L612M	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*	This study
5128	tet-*RAD18*, BrdU-inc, AFB2, Rad26-AID*	Mat-a, *his3-Δ200,leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*HIS3*, RAD26-AID*-6Flag::hphB	This study
5171	tet-*RAD18*, BrdU-inc, AFB2, Rad26-AID*, Rpb9-AID*	Mat-a, *his3-Δ200,leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*HIS3*, RAD26-AID*-6Flag::hphB, RPB9-AID*-GFP::*TRP1*	This study
5172	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2	This study
5193	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol2 M644G, *rnh201Δ*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol2 M644G::POL2*, *rnh201*::*HIS3*	This study
5194	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol3 L612M, *rnh201Δ*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*	This study
5252	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2,	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-	This study

	ura3Δ, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*	NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol2 M644G::POL2*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4	
5253	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol3 L612M, *rnh201Δ*, TetO44-*URA3*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4	This study
5254	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1*	This study
5265	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M, *rnh201Δ*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1*, *rnh201*::*HIS3*	This study
5266	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1*	This study
5267	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol2 M644G	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol2 M644G::POL2*	This study
5268	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2*, ura3Δ*, Pol3 L612M	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*	This study
5271	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M, r*nh201Δ*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1*, *rnh201*::*HIS3*	This study
5272	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*,	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB,, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*,	This study

	Pol2 M644G, *rnh201Δ*	*ura3*::*natNT2*, *pol2 M644G::POL2*, *rnh201*::*HIS3*	
5273	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol3 L612M, *rnh201Δ*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*	This study
5274	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol3 L612M, *rnh201Δ*, TetO44-*URA3*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol3 L612M*::*POL3*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4	This study
5275	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M, *rnh201Δ*, TetO44-*URA3*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1, rnh201*::*HIS3*, TetO44-*URA3*::PES4	This study
5291	tetO7-*RAD18*, pCM244, BrdU-inc, AFB2	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*TRP1*	This study
5292	tetO7-*RAD18*, pCM244, BrdU-inc, AFB2*, rpb9Δ*	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52,* tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*TRP1*, *rpb9*::*HIS3*	This study
5295	tetO7-*RAD18*, pCM244, BrdU-inc, AFB2, AID*-Rpb9	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*TRP1*, AID*-9myc-RPB9::hphB	This study
5297	*BERΔ*, Gal-NLS-TetR-UDG*-GFP-AID*, AFB2, *ura3Δ*, Pol1 L868M, *rnh201Δ*, TetO44-*URA3*	Mat-a, *leu2-3, 112 trp1-1, can1-100 ura3-1 ade2-1 his3-11,15, apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-UDG*-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol1 L868M:: POL1*, *rnh201*::*HIS3*, TetO44-*URA3*::PES4	This study
5307	*BERΔ*, Gal-NLS-TetR-GFP-AID*, AFB2, *ura3Δ*, Pol2 M644G, *rnh201Δ*, TetO44-*URA3*	Mat-a, *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*, *apn1*::KanMX, *apn2*::KanMX, *ntg1*::hphB, *ntg2*::hphB, Gal-NLS-TetR-GFP-AID*::*LEU2*, AFB2::*TRP1*, *ura3*::natNT2, *pol2 M644G::POL2*, rnh201::*HIS3*, TetO44-*URA3*::PES4	This study
5330	tetO7-*RAD18*, pCM244, BrdU-inc, AFB2, AID*-	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*,	This study

	Rpb9, Rad26-AID*	AFB2::*TRP1*, AID*-9myc-RPB9::hphB, *RAD26*-AID*-9myc:: *HisMX*	
5331	tetO7-*RAD18*, pCM244, BrdU-inc, AFB2, *rpb9Δ*, Rad26-AID*	Mat-a, *his3-Δ200, leu2-3,2-112, lys2-801, trp1-1(am), ura3-52*, tetO7(*RAD18*)::KanMX, CMVp(tetR'-SSN6)::*LEU2*, BrdU-inc::*URA3*, AFB2::*TRP1*, *rpb9*::*HIS3*, *RAD26*-AID*-6Flag:: hphB	This study

2.3 Buffers and reagents

2.3.1 General solutions, kits, enzymes and buffers

0.5 M ethylene-diamine-tetraacetic acid (EDTA) pH 8.0, 10x phosphate-buffered saline (PBS), 10x sodium dodecyl sulfate (SDS) running buffer, 20x saline-sodium citrate (SSC), 1 M Tris-HCl pH 7.5, 5 M sodium chloride, and 10x Tris-Borate-EDTA (TBE) were prepared by the Media lab at the Institute of Molecular Biology.

Other commonly used commercial reagents, buffers, and kits are described in the following table:

Table 2.3. Commonly used reagents

Chemicals	Supplier
Methyl methanesulfonate	Sigma-Aldrich - 129925
5-Flouroorotic acid monohydrate	TCR/ Biozol - F595000
Doxycycline hydrochloride	Sigma-Aldrich - D3447
Formaldehyde solution, 36.5-38% in H_2O	Sigma-Aldrich - F8775
5-Bromo-2-deoxyuridine	Sigma-Aldrich - B5002-5G
Propidium Iodide	Sigma-Aldrich - 81845-100MG
Indole-3-acetic acid sodium salt	Biomol - Cay16954-10
Nocodazole	Sigma-Aldrich - M1404-50MG
Sodium azide, NaN3	Sigma-Aldrich - S2002-100G
Ponceau S	Sigma-Aldrich - P3504-50G
Buffers	
4x NuPAGE LDS buffer	Fisher Scientific - 11559166
5x Green GoTaq Reaction Buffer	Promega - M7911
High Sens. RNA ScreenTape Sample Buffer	Agilent Technologies - 5067-5580

High Sens. D1000 ScreenTape Sample Buffer	Agilent Technologies - 5067-5603
Deoxynucleotide (dNTP) Solution Mix	New England Biolabs - N0447L
Kits	
DNA Clean & Concentrator	Zymo Research
MasterPureTM Yeast DNA Purification Kit	Biozym - MPY80200
ChIP DNA Clean & Concentrator™	Zymo Research
ExoSAP-IT™ PCR Product Cleanup Reagent	Thermo Fisher - 78200.200.UL
Agilent Bioanalyzer High-Sensitivity DNA Kit	Agilent Technologies, 5067-4646
Enzymes	
Proteinase K	Sigma-Aldrich - P4850
Q5 DNA polymerase	New England Biolabs - M0491S
Restriction enzymes	New England Biolabs
RNase A	Sigma-Aldrich – 10109169001
RNase HII	New England Biolabs - M0288S
T4 DNA ligase	New England Biolabs - M0202L
T4 DNA Polymerase	New England Biolabs - M0203S
Taq DNA polymerase	Protein Core Facility - IMB
Phusion HF DNA polymerase	New England Biolabs - M0530S
Fast alkaline phosphatase	Thermo Fisher Scientific - EF0651
α-factor Mating Pheromone	Eurogentec Germany - AS602215
a-Factor Mating Pheromone	Zymo Research - Y1004-500
Zymolyase 20T	AMS Biotechnology- 120491-1
Zymolyase 100T	AMS Biotechnology- 120493-1

2.3.2 Reagents for yeast manipulation

100 mM Lithium acetate (LiOAc): 100 mM lithium acetate and 10 mM Tris-HCl pH 7.4 were mixed and autoclaved.

50% (w/v) PEG 3350: 50 g of PEG3350 was dissolved in 100 mL of water, filter-sterilised.

Herring sperm (HT) DNA: 300 mg of DNA sodium salt from herring testes, Type XIV (Sigma-Aldrich) was dissolved in Tris-EDTA (TE) buffer (to 10 mg/mL). This was then sheared by passing through a syringe needle and/or by sonication, until the viscosity decreased considerably. HT DNA was extracted with phenol/chloroform, ethanol precipitated and

resuspended in TE buffer before boiling at 95°C for 5 min and cooled-down on ice for at least 5 min. This procedure was performed by Laura Tomini or Marwah Al-Hushail.

0.1 M tri-sodium citrate: 25.8 g sodium citrate tribasic hydrate in 1 L of ultra-pure water.

50 mM Sodium citrate buffer (pH 7.0): a few drops of 0.1 M citric acid were added to 490 mL of 0.1 M tri-sodium citrate until pH reaches 7.0.

86% glycerol: 86 mL of 100% glycerol and 14 mL of water.

2.3.3 Buffers for protein purification

Trichloroacetic acid (TCA), 100% (w/v): 250 g of TCA powder in 113.5 mL water and stored at 4°C.

TCA ,55% (w/v): 55 mL 100% TCA was filled-up with water to 100 mL. A bottle with special TCA-resistant pouring ring and lid was used. This procedure was performed by Laura Tomini or Marwah Al-Hushail.

NuPAGE 1xLDS-Di-thio-threitol (DTT) buffer: NuPAGE 4x LDS buffer was diluted to 1x in water and supplemented with 100mM DTT.

PBS-Tween (PBS-T): 1x PBS and 0.1% (v/v) Tween 20.

Ponceau S solution: 0.1 % (w/v) Ponceau S in 5 % (v/v) acetic acid.

Blocking solution: 5% (w/v) skim milk powder (Sigma-Aldrich) in PBS-T.

Blotting buffer (Turbo transfer): 3.9 mM glycine, 4.8 mM Tris, 0.13 mM SDS, and 20% (v/v) ethanol were mixed.

Blotting buffer (wet transfer): 192 mM glycine, 25 mM Tris, and 15% (v/v) methanol were mixed.

2.3.4 Buffers and solutions for molecular biology methods

2.3.4.1 DNA analysis

TE buffer: 50 mM Tris-HCl pH 8.0 and 50 mM EDTA.

Y1 buffer (spheroplast method): 1 M sorbitol and 100 mM EDTA was mixed and stored at 4^{o}C until use. 14 mM β-mercaptoethanol was added immediately before use.

Lysis buffer (spheroplast method): 50 mM Tris-HCl pH 8.0 and 50 mM EDTA was mixed and stored at 4^{o}C until use.

6x DNA blue loading buffer: 50% (w/v) sucrose, 0.1% (w/v) bromophenol blue (BPB), and 0.1% (w/v) xylene cyanol F were mixed, made to a total volume of 100 mL with TE, and filter-sterilised.

2.3.4.2 Chromatin immunoprecipitation (ChIP) and Bromodeoxyuridine immunoprecipitation (BrdU-IP)

RIPA buffer: 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0, 140 mM NaCl, 1% (v/v) Triton X-100 (Sigma-Aldrich), and 0.1% (w/v) sodium deoxycholate were dissolved in water. The buffer was filter-sterilised and stored at 4°C.

Washing buffer I: 50 mM HEPES pH 7.5, 1 mM EDTA pH 8.0, 500 mM NaCl, 1% (v/v) Triton X-100, and 0.1% (w/v) sodium deoxycholate were mixed in water. The buffer was filter-sterilised and stored at 4°C.

Washing buffer II: 250 mM LiCl, 0.5% (v/v) IGEPAL-CA630 (Sigma-Aldrich), and 0.5% (w/v) sodium deoxycholate were dissolved in water. The buffer was filter-sterilised and stored at 4°C.

Immunoprecipitation (IP) buffer: 1x PBS and 0.0625% (v/v) Triton-X100. Buffer made fresh each time.

Beads buffer: 1x PBS and 0.02% (v/v) Tween-20. Buffer made fresh each time.

Elution buffer: Tris-EDTA and 1% (v/v) SDS. Buffer made fresh each time.

2.3.4.3 Southern blotting

Alkaline Gel: 1% (w/v) agarose melted in filtered-water. Cooled to 60°C before adding 50 mM NaOH and 1 mM EDTA (pH 8.0).

Alkaline Gel Running Buffer: 50 mM NaOH and 1 mM EDTA (pH 8.0). Buffer made fresh each time.

Neutralization Buffer (Alkaline Gel): 1 M Tris-HCl pH 7.6, and 1.5 M NaCl.

Depurination Buffer (Native Gel): 0.25 M HCl in water.

Denaturation Buffer (Native Gel): 1.5 M NaCl and 0.5 M NaOH in water.

Neutralization Buffer (Native Gel): 1.5 M NaCl and 0.5 M Tris in H_2O, adjusted to pH 7.5 with HCl.

Hybridisation Buffer: 5x Denhardts solution, 0.5% (w/v) SDS, 6x SSC and 50 µg/mL HT DNA (added before use).

Low Stringency Wash Buffer: 2x SSC and 0.1% (w/v) SDS.

High Stringency Wash Buffer: 0.2x SSC and 0.1% (w/v) SDS.

Stripping Solution: 0.1 % SDS in water. Made fresh each time.

2.4 Plasmids

Table 2.4. Plasmids used in this study

ID	Name	Use	Source/Reference
233	pFA-*HIS3*	HIS3MX disruption cassette for gene deletion	H. Ulrich
309	pFA6a	KanMX disruption cassette for gene deletion	M. Knop
452	pFA-KanMX4	KanMX disruption cassette for gene deletion	W. Kramer
841	pFA6a TRP1	Source of TRP1	A. Schürer
1369	pCM244	Tet-regulated yeast expression system: TetR'-SSN6 vector	Euroscarf
1370	pCM245	Tet-regulated yeast expression system: TetR'-SSN6 vector	Euroscarf
1377	pCM225	Tet-regulated yeast expression system: promoter substitution	Euroscarf
1419	pUC-PES4IG	Insertion of constructs into the intergenic region at ARS607	H. Ulrich
1420	pBS-*URA3*	Source of *URA3*	H. Ulrich
1460	p404-BrdU-inc	Integration of BrdU-inc for in vivo BrdU labeling	O. Aparicio
1462	p306-BrdU-inc	Integration of BrdU-inc for in vivo BrdU labeling	O. Aparicio
1633	pFA6a-natNT2	Nourseothricin disruption cassette for gene deletion	J. Diffley
1634	pFA6a-hphNT1	HygromycinB disruption cassette for gene deletion	J. Diffley
2058	PES4IG TetO44-*URA3*	TetO44-*URA3* sequence flanked by PES4IG regions	E. Colby
2190	pHisMx-AID*	HIS3MX: C-terminal AID* tagging	M. Morawska-Onyszczuk
2192	pUC-PES4IG longer	Longer arms of homology for insertion of constructs into the intergenic region at ARS607	E. Colby
2226	YIp128-GAL-NLS-TetR-UDG*-GFP-AID*	TetR-UDG*-AID* fragment with Gal promoter	E. Colby
2227	YIp128-GAL-NLS-TetR-UDG-GFP-AID*	TetR-UDG-AID* fragment with Gal promoter	E. Colby
2228	YIp128-GAL-NLS-TetR-GFP-AID*	TetR-AID* (empty) fragment with Gal promoter	E. Colby
2353	pHyg-AID*-6xFlag	HygromycinB: C-terminal AID*-6xFlag tagging	M. Morawska

2354	pHyg-AID*-GFP	HygromycinB: C-terminal AID*-GFP tagging	M. Morawska
2397	pRS303-ADH-AFB2	HIS3MX: ADH-AFB2 sequence (F-box protein) for AID* construction	R. Wong
2398	pYM 9myc Nat	Nourseothricin: C-terminal 9-myc tagging	R. Wong
2444	YIp204-ADH-AFB2	TRP1: ADH-AFB2 sequence (F-box protein) for AID* construction	R. Wong
2467	pHyg-PCUP1-AID*	HygromycinB: N-terminal AID* tagging	H. Ulrich
2468	pKan-PCUP1-AID*	Kanamycin A: N-terminal AID* tagging	H. Ulrich
3849	YIp128-G1tag-NLS-TetR-TDG-GFP-AID*	TetR-UDG-AID* fragment with G1tag (Sic1 promoter to induce UDG* expression at G1 phase)	L. Batista
3853	p173-*pol2 M644G*	For *pol2 M644G* integration with *URA3* marker	J. Williams
3854	p170-*pol3 L612M*	For *pol3 L612M* integration with *URA3* marker	J. Williams
4319	pDS2-amdSYM	amdSYM disruption cassette for gene deletion	Euroscarf
4433	TetO44-PES4-LongArms	TetO44 sequence (without *URA3*) flanked by long PES4IG arms for popping-out the *URA3* sequence next to TetO array	This study
4434	TetO44-PES4	TetO44 sequence (without *URA3*) flanked by PES4IG arms for popping-out the *URA3* sequence next to TetO array	This study
4955	pTRP-AID*-GFP	TRP1: C-terminal AID*-GFP tagging	This study
5108	pHyg-PCUP1-AID*(N)-9myc repaired	HygromycinB: N-terminal AID*-9xmyc tagging	This study
5495	pYIAL30-pol1L868M	For pol1 L868M integration with *URA3* marker	M. Garbacz

2.5 DNA oligonucleotides

DNA oligonucleotides were purchased from Sigma-Aldrich (Merck) or Integrated DNA Technologies (IDT).

Table 2.5. G-block used in this study

Name	Sequence (5’-3’)	Application
G-Block	GACGGACGTACTCTAGAGCTTTTCAATTCAATTCATCATTTTTTTTTATATTGGGTAGACAGAGGGACGCCGTCTGCTATATGATATGTATTTAGGTCCGAGTGCTGATAAACTGTAAGGCAAAAGGTA	Pop-out *URA3* sequence next TetO44 array

Table 2.6. DNA oligonucleotides used in this study

ID	Name	Sequence (5'-3')	Use
50	*URA3*-Downstream	CGACGGTACCATGTCGAAAGCTACATATAAG	PCR testing
237	ARS305.3	TTTCAGAGCCTTCTTTGGAG	ChIP-qPCR
665	ARS501-Forward (I)	AAGCAAATTGCAGAAGGTTATGAA	ChIP-qPCR
666	ARS501-Reverse (I)	TTCAAGGCTCTAGCATATGAAACG	ChIP-qPCR
784	*APN2* KO	GCTGGCTCAACGATCTGTC	PCR testing
899	*URA3* up III	AATAGCTTGGCAGCAACAGG	PCR testing
1004	PES4IG probe-Reverse	AAAGGACCCGGGTAGACAGAGGGACGCCGT	Southern blot
1006	*APN1* KO Test	AGGGTAGGTTAGATGGGTTC	PCR testing
1545	hphNT1	AAGGAGGGTATTCTGGGCC	PCR testing
1636	pLAU44 TetO array (2)	GTCATTCTGCTAGAGCTCCCTG	PCR testing
1681	pLAU44 TetO (19)	GTCCCGACTCTAGCATCGTG	PCR testing
1781	PES4IG-Forward	GCTTTGCAGGAGAAGAGTCCGGAG	PCR testing/ Southern blot
1782	PES4IG Test-Downstream	GTGCCACTAGAGACGAAAATGTGC	PCR testing
2228	ChrI probe-Reverse	TGGCAGTTCCGTGAGGCTATTGC	Southern blot
2229	ChrI probe-Forward	TTTGCAGCCCACCGATTTCCAGC	Southern blot
2386	PES4IG Southern-Reverse	GCTGCCCGAACTACTTTATGC	Southern blot
2387	PES4IG Southern-Forward	GGCAGCTGTTACATAGAATCTC	Southern blot
3143	ARS607	CGGCTCGTGCATTAAGCTTG	ChIP-qPCR
3144	ARS607 set2	TGCCGCACGCCAAACATTGC	ChIP-qPCR
3147	ARS607-2kb-Downstream	GGGCACAGGATAAAGAAGAAAGAGC	ChIP-qPCR

3148	ARS607-2kb-Downstream	CCATCTATCTGAATCTTGAACGGTC	ChIP-qPCR
3322	ARS305 a	CACCGGACAGTACATGAAAC	ChIP-qPCR
3790	Distal adaptor	CGAGATCTACACTCTTTCCCTACACGACGC TCTTCCGATCT	GLOE-seq
3791	Distal adaptor	GACTGGAGTTCAGACGTGTGCTCTTCCGAT CT	GLOE-seq
3792	Complementary strand synthesis	GATCGGAAGAGCACACGTCTGAACTCCAG TC	GLOE-seq
3898	Proximal adaptor	CTACACGACGCTCTTCCGATCTNNNNN*N-NH_2 (*: phosphorothioate bond, IDT code *; NH_2: 3'-amino modification, IDT code /3AmMO/)	GLOE-seq
3899	Proximal adaptor	PO_4-AGATCGGAAGAGCGTCGTGTAGGGAAAGA GTGTAGATCTCGTTTT-Bio (PO_4: 5'-phosphorylation, IDT code /5Phos/; T-Bio: 3'-biotin-dT, IDT code /3BiodT/)	GLOE-seq
4029	TetO44array-Forward- set2	AGGGAGCTCTAGCAGAATGA	ChIP-qPCR
4030	TetO44array-Upstream	GAGCTTCTTGCGGTCTCTATC	ChIP-qPCR
4032	PES4IG-Downstream	GCTGCAAGCTCGTTAGGATTTA	ChIP-qPCR
4033	PES4IG-Upstream	CCGGGTCCTTTTAAGAATTTTTTC	ChIP-qPCR
4407	*RNH201*-KO-Reverse	TGTCGAAAAACCTTGAAAACAACTACTGCA CACCAAATTGATACGATTAACGTACGCTGC AGGTCGAC	KO cassette
4408	*RNH201*-KO-Forward	ACATATGTAGTATTACATGAAGATATATAG TATGTGCAAACTGGAGGTGAATCGATGAA TTCGAGCTCG	KO cassette
4409	*RNH201*-KO-Forward	TGGATGATGTAAACAGGCAG	PCR testing
4640	*POL2*-seq-Reverse	ACATTGAGAAAGGGTACCGG	PCR testing/ Sequencing
4641	*POL2*-seq-Forward	CCTTTGAATTCGTAACGCCT	PCR testing/ Sequencing
4642	*POL3*-seq-Reverse	ATCTACAAAATGGCGATAGT	PCR testing/ Sequencing
4643	*POL3*-seq-Forward	GGTACCAAGATCCATAGCTT	PCR testing/ Sequencing
4900	*NTG1* Upstream	GCAAACGCTGTGTCAAGACG	PCR testing
4901	HygromycinB-Downstream	TTCGAGCGGAGGCATCCG	PCR testing
4902	*NTG2* Downstream	TTCTGGCTGGTTCATTTTGGGC	PCR testing

4903	*RNH201* Upstream	ATATCCAACGGAGTGATGCACG	PCR testing
4988	*URA3*-KO-Downstream	GCACAGAACAAAAACCTGCAGGAAACGAAGATAAATCCGTACGCTGCAGGTCGAC	PCR Knockout
4989	*URA3*-KO-Upstream	GTGAGTTTAGTATACATGCATTTACTTATAATACAGTTTTATCGATGAATTCGAGCTCG	PCR Knockout
4990	*URA3*-KO	GATTCGGTAATCTCCGAGCAGAAG	PCR testing
5008	*URA3*-Flanking-regions-KO-Downstream	GCATATTTATGGTGAAGGATAAGTTTTGACCATCAAAGAACGTACGCTGCAGGTCGAC	PCR Knockout
5009	*URA3*-Flanking-regions-KO-Upstream	GAAGCTTTTTCTTTCCAATTTTTTTTTTTTCGTCATTATAATCGATGAATTCGAGCTCG	PCR Knockout
5010	*URA3*-KO-flanking_Test-Upstream	AACGGTTCATCATCTCATGGATCTGCAC	PCR testing
5030	*URA3*-KO with amdSYM	ATTGTTGAGCGATAGCAGCTCT	PCR testing
5033	*AFB2*-Reverse	ATGAGGAACCAAATTGAAATCCGCAAA	PCR testing
5034	*TRP1*-Forward	GGAGGGCATTGGTGACTATTGAG	PCR testing
5035	*AFB2*-Forward	TCAACTCCGGATCCCCCATCAC	PCR testing
5125	TetO44-PES4 gBlock-Reverse	ACCTTTTGCCTTACAGTTTATC	Amplification of G-Block for *URA3* next TetO deletion
5126	TetO44-PES4 gBlock-Forward	ACGGACGTACTCTAGAGCT	Amplification of G-Block for *URA3* next TetO deletion
5135	TetO44-PES4 gBlock	CGTCCCTCTGTCTACCCAATATAAA	PCR testing of *URA3* pop-out with gBlock
5136	PES4IG Test II - Reverse	TCACTGCTCCAACATCCCCAT	PCR testing
5154	PES4IG Test III - Reverse	CTCCGTTTCTGCTGGGCCT	PCR testing
5155	TetO44-Forward	AGAGACGGAGCCAAACTCTAGC	PCR testing of *URA3* pop-out when using pHU#4434
5156	PES4IG Test - Forward	CAGTGACAAAAAAATCACTGGGCC	PCR testing

5558	*RAD26* C-terminal AID*-6Flag-Forward	TGGGTGCTTGATGAAGAATTTAGGAATAACAATGCTTCAGGTACCCCTAAAGATCCAGCCAAACC	PCR tagging
5559	*RAD26* C-terminal AID*-6Flag-Reverse	CGTCAAGGGTACTCTGTCCATTCTAAGTTTGATGAATTCGAGCTCGTTAA	PCR tagging
5560	*RAD26* Upstream	GAGGCATCCAAGGAAGCTCGCCAAG	PCR testing
5616	*RPB9* C-terminal AID*-9myc-Forward	TTACTTCAGATCAAAAAAACAAAAGGACGCAGTTTTCACAGGTACCCCTAAAGATCCAGCCAAACC	PCR tagging
5617	*RPB9* C-terminal AID*-9myc-Reverse	TTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGAGCACGCAGGCATTTGCTCGGCATG	PCR tagging
5618	*RPB9*-Upstream	AATGTCACTCTCGGGAAAATGT	PCR testing
5619	*RPB9* C-terminal AID*-GFP-Reverse	TTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGAGCACGCCGCATAGGCCACTAGTGGAT	PCR tagging
5626	*RPB9* C-terminal AID*-GFP-Reverse_2	TTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGAGCACCGCATCTGTGCGGTATTTCACACC	PCR tagging
5655	*POL1*-Forward	GCCTGAAAAATTCGGTAGAGGAAATTCG	PCR testing/ Sequencing
5656	*POL1*-Reverse	CTGCAGGCATGTTTTTACCACCT	PCR testing/ Sequencing
5696	*RPB9*-KO-Forward	TCTAGCCAAAAGAGCAAGTTAAACTCCCCTTAAAACTGCCGGATCCCCGGGTTAATTAA	KO cassette
5697	*RPB9*-KO-Reverse	CCATTTTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGGAATTCGAGCTCGTTTAAAC	KO cassette
5698	*RPB9*-2-Upstream	CCGAGTTCTTGTTTGCCATTAG	PCR testing
5711	*RPB9*-KO-Forward-2	CTAGCCAAAAGAGCAAGTTAAACTCCCCTTAAAACTGCTCGTATGTGTAATTGAAGAAAGATACG	KO cassette
5712	*RPB9*-KO-Reverse-2	CCCATTTTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGAGCACCTTGGGTACTCTGAGCGA	KO cassette
5713	*RPB9*-Downstream	TAGGCTGGTGTAAATAGTAGTGC	PCR testing
5714	*RPB9*-KO-Forward-3	GAACGAAAGTTACAAAATCTAGCCAAAAGAGCAAGTTAAACTCCCCTCGGATCCCCGGGTTAATTAA	KO cassette
5715	*RPB9*-KO-Reverse-3	CCCATTTTCTCTCCCTCTGTCATTAATTTTGAAAGTTCGTTGAGCACGAATTCGAGCTCGTTTAAAC	KO cassette
5722	*POL1*-Forward-4	TTAGGTTTAGCTGCACTGGCAAAACAG	PCR testing
5723	*POL1*-Reverse-4	TCAGTTTTCATCACCTTCTTAACTTCACGT	PCR testing

5733	*RPB9*-KO-Forward-4	AAGTTACAAAATCTAGCCAAAAGAGCAAGTTAAACTCCCCCGTACGCTGCAGGTCGAC	KO cassette
5734	*RPB9*-KO-Reverse-4	TCCTTTTGCCGGTAGATTAGGCTGGTGTAAATAGTAGTGCATCGATGAATTCGAGCTCG	KO cassette
5735	*RPB9*-3-Upstream	GTCGTAGAATTGAAACGGGTTGTGATAG	PCR testing
5751	*RPB9*-KO-Flanking_Forward	TACATCATCCTTATGTGATGGTGCTGTTGTAGTCGTAGAACGTACGCTGCAGGTCGAC	KO cassette
5752	*RPB9*-KO-Flanking_Reverse	AGACGTATCTCTCGTCCATTTTCATTACATTTTCATCTTTCTTCATCGATGAATTCGAGCTCG	KO cassette
5753	*RPB9*-4-Upstream	GGTGCAGCTTCAACGAAT	PCR testing
5786	*RPB9* N-terminal AID*-9myc Forward	AGCCAAAAGAGCAAGTTAAACTCCCCTTAAAACTGCTATGTTAGCTTGCCTCGTCCCCGC	PCR tagging
5787	*RPB9* N-terminal AID*-9myc-Reverse	GGGTACAACATATTGTTGCAGTCACGACAAAATCTAAACGTAGTGAGCTCTGGCACCCGCTCC	PCR tagging
5788	N-tagging AID* Test-Forward	AACGGATCCACTAGCCTCGAG	PCR testing
5789	*RPB9*-2-Downstream	CCCGAGAGTGACATTTGGGAC	PCR testing
5790	*RPB9*-KO-Forward-5	AAGTTACAAAATCTAGCCAAAAGAGCAAGTTAAACTCCCCTCGTACGCTGCAGGTCGACGG	KO cassette
5791	*RPB9*-5-Upstream	GATGGTGCTGTTGTAGTCGTAG	PCR testing
5812	*RAD26* C-terminal AID*-9myc-Reverse	GGTAAAGTTACGTCAAGGGTACTCTGTCCATTCTAAGTTTGAATTCGAGCTCGTTTAAAC	PCR tagging
5813	*RAD26* Upstream-2	AAGAGGCATCCAAGGAAGCT	PCR testing
5862	*URA3*-Forward	TTTTGCGAGGCATATTTATGGTGAAGGA	PCR testing
5863	*URA3*-Reverse	TAGTATATTCTCCAGTAGCTAGGGAGCC	PCR testing
5864	BrdU-inc-Forward	TGACCCACGCATGTATCTATCTCATTT	PCR testing
5865	BrdU-inc-Reverse	CAGAGTGCCAGCCCTGG	PCR testing
5866	*POL3*-Forward	CGCCTACCTGCCTTTAAGGC	PCR testing
5867	*POL3*-Reverse	TCTTCAAAGCCAATTGTCTACCATTTAAAACATCT	PCR testing
5868	*POL2*-Forward	TTAGGAGTGATTTGAAGAATGAATTCAAGATAGATCCT	PCR testing
5869	*POL2*-Reverse	TGCGTAACTTAGTTCATCAAATGTCAAAACTTTCT	PCR testing
5994	ChrI probe-Forward	CTCTCGGACCTCTCAGATACCACG	Southern blot

5995	ChrI probe-Reverse	CTAGTGGCGTCTGCGCC	Southern blot

2.6 Antibodies

Table 2.7. Primary antibodies used in this study

ID	Name	Host	Clonal type	Dilution	Supplier
11	BrdU	Rat	Monoclonal	1% Milk-PBS-T, 1:500	Abcam
49	GFP	Mouse mAb/IgG1 kappa	Monoclonal	5% Milk-PBS-T, 1:1000	Roche
99	c-Myc	Mouse mAb/IgG1	Monoclonal	5% Milk-PBS-T, 1:1000	IMB - PPCF
100	c-Myc	Mouse mAb/IgG1	Monoclonal	5% Milk-PBS-T, 1:1000	Invitrogen
137	Phosphoglycerate kinase 1 (PGK1)	Mouse mAb/IgG1 kappa	Monoclonal	5% Milk-PBS-T, 1:1000	Invitrogen
150	Pol30 R37/38 (serum) (PCNA)	Rabbit	Polyclonal	5% Milk-PBS-T, 1:1000	H. Ulrich
166	Rad18 (yeast)	Rabbit	Polyclonal	5% Milk-PBS-T, 1:500	H. Ulrich
253	Tubulin YL1/2	Rat	Monoclonal	5% Milk-PBS-T, 1:1000	-
297	Rad53	Mouse	Monoclonal	5% Milk-PBS-T, 1:500	Susan Gasser

Table 2.8. Secondary antibodies used in this study

Name	Host	Dilution	Supplier
Anti-Mouse immunoglobulins/HRP	Goat	5% Milk-PBS-T, 1:5000	DAKO
Anti-Rabbit immunoglobulins/HR	Goat	5% Milk-PBS-T, 1:5000	DAKO
Anti-Rat immunoglobulins/HRP	Goat	5% Milk-PBS-T, 1:5000	DAKO

2.7 Determination of cell density

Cell number and density of yeast and bacterial cultures were determined by measuring OD600 of 1:10 dilutions of cells on Eppendorf Biophotometer, assuming that 1 OD600 of exponential YPD culture equals to ~10 x 10^6 cells/mL.

2.8 *Escherichia coli* manipulation

Bacterial cultures were grown at 37°C in LB media complemented with appropriate antibiotic.

2.8.1 Preparation of chemically competent cells

A 2 mL overnight culture of Top Ten *E. coli* was diluted 100-fold in 200 mL of fresh LB medium and grown at 25°C until 0.4-0.6 OD600. Cells were put on ice for 10 min and harvested at 2500 g for 10 min at 4°C. The cell pellet was washed according to Zymo Research's Mix & Go! *E. coli* Transformation Kit's instructions. Aliquots of cells were snap-frozen in liquid nitrogen and stored at -80°C for one time use. This procedure was performed by Laura Tomini or Marwah Al-Hushail.

2.8.2 Transformation using competent cells

A 200 µL aliquot of competent Top Ten *E. coli* was thawed and mixed with 50-100 ng of plasmid DNA. Cells were kept on ice between 15-30 min before heat-shock at 42°C for 1 min. The mix was further incubated on ice for 2 min. If ampicillin was used as a selection marker, cells were directly plated on Amp LB agar plates after centrifugation at 3000 g. If another antibiotic was used as selection marker, cells were further incubated for 1 h at 37°C with shaking prior to plating.

2.8.3 Isolation of plasmid DNA

Single bacterial clones were inoculated overnight in 2 mL LB liquid medium supplemented with corresponding antibiotic. Plasmid DNA was extracted using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) following the manufacturer's manual. The DNA was quantified using a Nanodrop spectrophotometer.

2.8.4 Cloning

DNA cloning was carried out using standard techniques using restriction enzymes (NEB), FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific), Quick Blunting Kit (NEB) and Quick Ligation Kit (NEB). Phusion DNA polymerase (NEB) was used for PCR cloning reactions.

2.9 Yeast manipulation

2.9.1 Cultivation and storage

All strains were grown at 30°C in YP or SC medium supplemented with appropriate amino acids. Unless otherwise stated, glucose was used as the carbon source. Liquid cultures were incubated with constant shaking at 180-220 rpm. For harvesting small volumes of culture, tabletop centrifuges (Thermo Scientific) were used, meanwhile large volumes of cells were collected in a SORVALL RC 6+ (Thermo Scientific) at 3000-5000 rpm.

To induce expression from the *GAL1/10* promoter, cells in an YPD or SC starter culture were grown in YPRaff overnight. The following morning cells were pelleted and then shifted to YPGal. For the induction of the Tet-off, cells inoculated in YPD were supplemented with doxycycline at a final concentration of 2 µg/mL.

For the long-term yeast storage, all yeast strains were saved at -80°C by using 1 part of 86% glycerol (v/v) and 4 parts of saturated overnight yeast culture. The cryogenic vials containing this mixture were placed at -80°C without flash freezing.

2.9.2 Mating and tetrad dissection

Overnight grown haploid strain cultures with opposite mating types were mixed on fresh YPD agar plates and incubated for 3-5 hours at 30°C. Zygotes were picked up on YPD plates using the MSM micromanipulator (Singer) before further incubation for 2-3 days at 30°C. When possible, the resulting a/α cells were streaked on selective agar plates to ensure proper isolation of diploids. Diploid cells were grown overnight in a liquid YPD medium. The next day, 500 µL of the culture was transferred to a 1.5 mL Eppendorf, cells were pelleted and washed five times with sterile water and once with sporulation medium. These were resuspended in 2 mL of sporulation medium and incubated for 3 days at 25°C. Alternatively, for difficult mating, cells were initially resuspended in pre-sporulation medium and incubated overnight at 25°C, the next day cells were pelleted, washed five times with sterile water, resuspended in special sporulation medium and incubated for 3-5 days at 25°C.

To digest the ascus wall and separate the tetrads, 20 µL of the sporulated cultures were mixed with 20 µL of STE-Zymolyase solution and incubated for 15 min at room temperature. At least 10 µL of the mixture was very gently streaked on a side of an YPD agar plate. Tetrads were dissected using the MSM micromanipulator (Singer). Finally, tetrads containing the desired mutations were identified by replica plating on different selection plates and additional genomic confirmation by colony PCR was performed. The mating type of the haploids was determined by complementation of a rare auxotrophic marker with tester strains.

2.9.3 PCR-based gene tagging or deletion

The tagging or deletion of specific genes on yeast was generated by creating PCR cassettes. For gene deletions, the PCR cassettes were amplified using oligonucleotides carrying around 40 bp homology at the 5'- and 3'- ends of the desired gene and around 20 bp homology at the 5'- and 3'- ends of the selection marker sequence. For tagging a gene on the N- or C-terminus, an oligonucleotide pair was used that shared homology with the 5'- or 3'- end of the desired gene. The PCRs were performed using a high-fidelity polymerase (Phusion HF DNA polymerase).

2.9.4 PCR cassette transformation

Exponentially growing yeast cells (2-3 OD600) were pelleted at 1500 g for 1 min. Washed once with sterile water and then washed with 1 mL of 100 mM LiOAc. After centrifugation, 240 µL of 50% PEG 3350, 36 µL 1 M LiOAc, 5 µL of single-stranded HT DNA and 1-2 µg of linearized plasmid (mixed with water to a volume of 76 µL) were added in that order to the pelleted cells. The mixture was vortexed and incubated on a thermo-block at 30°C for 30 min. Next, the cells were heat-shocked for 20 min at 42°C. Cells were centrifuged at 1500 g for 1 min and resuspended in 100 µL of sterile water and then plated on the appropriate selective medium. In the case of using selection markers with antibiotics, cells were recovered in 2 mL of YPD for 3-5 hours at 30^{o}C, before plating. Colonies were visible after 3-4 days and these were checked for correct genomic integration by colony PCR.

2.9.5 Cell synchronisation

Exponentially growing MATa cells (OD600 0.8-1) were synchronised at G1-phase by incubation at 30^{o}C for 1.5 to 2 hours with 10 mg/mL of α-factor with constant shaking. Analogously, MATα cells were synchronised with a-factor at a final concentration of 0.75 µg/mL, according to the manufacturer's recommendations (Zymo Research). The a-factor stock was 100 µg/mL diluted in 0.5 mg/mL Bovine Serum Albumin (BSA). To released cells into the S-phase, these were washed 2-3 times with room temperature medium. After cells were collected, inactivation was carried out immediately in 1% (w/v) sodium azide.

2.9.6 Growth curve assays

Overnight saturated cells were diluted to a density of 0.1 OD600 and allowed to grow exponentially for 4 to 6 hours at 30^{o}C with constant shaking. Thereafter, 100 µL of each strain culture was taken to a 96-well plate, plain medium was included as a blank, to measure cell densities using a Tecan SPARK 20M. The samples were transferred to a new 96-well plate at a dilution of 0.02 OD, and then the plate was left in Tecan SPARK 20M for 16 hours at 30^{o}C in constant shaking, the OD was measured every hour. A growth curve of the OD600 readings against time was plotted to compare growth rates between strains.

2.9.7 Analysis of cell cycle progression by flow cytometry

1 mL of yeast cultures inactivated in 1% (w/v) sodium azide were fixed in 70% ethanol overnight at 4°C or for at least 10 min at room temperature. Fixed cells were pelleted and washed three times in 50 mM sodium citrate buffer pH 7.0. These were then resuspended in 1 mL of the same buffer and 10 µL of RNase A (10 mg/mL) was added prior incubation at 55°C for 1 h. Next, Proteinase K was added to a final concentration of 80 mg/mL and incubated at 55°C for 1 h. Cells were then stained with 32 mg/mL propidium iodide and sonicated on a Branson sonifier for 10 s. Finally, the DNA content was analysed on a FACSVerse or BD LSRFORTESSA SORP cytometer (BD Biosciences), and data was visualised using the FlowJo v10 software.

2.9.8 Spot assay

The equivalent of 0.1 OD600 exponentially growing cells were serially diluted (10-fold) and spotted on freshly prepared YPD plates supplemented with or without MMS in different concentrations. After spotting, plates were incubated for two days at 30°C.

2.9.9 Survival assay

Survival assays were carried out by plating a known number of exponentially growing cells on YPD agar. Culture dilutions were determined depending on the strain, usually 100 µL of 1:1x10^3 dilution cultures were used. Colonies were counted three days after incubation at 30°C. For each strain survival was normalised to the unchallenged condition.

2.10 Protein analysis

2.10.1 Whole cell protein extraction

1-2 OD600 of cells were harvested and resuspended in 500 µL of ice-cold water alongside 74 µL 2 M NaOH and 6 µL β-mercaptoethanol. The mixture was incubated on ice for 20 min before 80 µL of 55% (w/v) TCA was added, followed by further incubation on ice for 10 min. Samples were centrifuged at 21000 g for 10 min at 4°C. The supernatant was discarded, and the pellet was resuspended in 40 µL of NuPAGE 1xLDS-DTT buffer. Finally, samples were heated for 10 min at 65°C and loaded onto poly-acrylamide gels.

2.10.2 Poly-acrylamide gel electrophoresis (SDS-PAGE)

To separate proteins, 10-15 µL of each sample resuspended in NuPAGE 1XLDS-DTT buffer was run in pre-casted gels using 1x SDS running buffer in an electrophoresis chamber (Mini-PROTEAN™ system or CriterionTM vertical electrophoresis cell, Bio-RAD). SDS-PAGE was performed using 4-15% Mini-PROTEANTM TGX Stain-freeTM (10-or 15-wells, Bio-Rad) or 4-15% CriterionTM TGX Stain-FreeTM (26-wells, Bio-Rad) pre-casted gels. The pre-stained PageRuler (Thermo Fisher Scientific) was used as a size ladder.

2.10.3 Western blotting with semi-dry Trans-Blot Turbo system

Gels were blotted into nitrocellulose membranes (Bio-Rad) using the Trans-Blot Turbo system according to the manufacturer's instructions. The mini or mid Nitrocellulose Transfer Kit (Bio-Rad) pre-soaked on transfer buffer Turbo Blot (Bio-Rad) was utilised for electro-blotting. The transfer stack was built from anode to cathode in the following order: six layers of Wypall X60 cloths (Kimberly-Clark), a nitrocellulose membrane (BioRad, pore size 0.2 mm), the electrophoresed gel and another six layers of Wypall X60 cloths. The electro-blotting of protein gels was performed using Bio-Rad's mid molecular weight blotting programme (Mini gels: 7 min at 1.3 A and up to 25 V; Midi gels: 7 min at 2.5 A and up to 25 V). After transfer, membranes were briefly rinsed with water, stained with Ponceau, and blocked in 5% milk in PBS-T for at least 30 min at room temperature. Membranes were then washed twice with PBS-T for 5 min, and then incubated in constant rotation with the relevant primary antibody overnight at 4°C. The next day, membranes were washed three times with PBS-T for 5 min to remove the unbound antibody and then incubated with the secondary antibody for 1 hour with constant rotation at room temperature. Three more washes with PBS-T were performed to remove the unspecific antibody signal and finally blots were scanned with the Fusion FX (Vilber) using Amersham ECL Select Western Blotting Detection Reagents (VWR International).

2.10.4 Western blotting by wet transfer

Alternatively, electro-blotting was done by wet transfer using an electrophoresis chamber and stacking the transfer sandwich in a Mini-Trans-Blot cell (Bio-Rad). The stack was assembled from bottom to top as follow: three layers of Whatman filter paper (Thermo Fisher Scientific), a nitrocellulose membrane (Bio-Rad), the gel, and another three layers of Whatman filter paper. The sandwich was placed vertically in the chamber with the membrane side facing the anode. Transfer onto the membrane was done by applying a constant voltage of 100 V for 1 hour. After the transfer, the western blotting was carried out as described in Section 2.10.3

2.10.5 Stripping membrane

In order to incubate the membrane with other primary antibodies, the membrane was washed with water for 5 min, incubated with 0.2 M NaOH for 5 min, and washed with water for another 5 min. Finally, the membrane was blocked with 5% milk in PBS-T for at least 30 min at room temperature before incubating with primary antibody overnight at 4^{o}C.

2.10.6 Dot-blotting

To check the incorporation of BrdU in yeast, DNA was extracted using either the spheroplast method or the MasterPure Yeast DNA Purification kit (described below). DNA was quantified using the Multiskan GO Microplate spectrophotometer (Thermo Scientific). Around 500 ηg/µL-1 µg/µL of DNA were denature at 95^{o}C for 10 min and then incubated on ice for 5

min. 2 µL of single-stranded DNA was loaded onto a Nitrocellulose Hybond-N+ membrane (Thermo Fisher Scientific) and allowed to dry at room temperature for at least 20 min. The DNA was cross-linked to the membrane with UV at 1200 J/m^2 using a Stratagene-UV Stratalinker. The membrane was quickly washed with PBS-T and blocked with 5% Milk-PBS-T for 10 min before being incubated with anti-BrdU antibody (Abcam) at 4°C overnight. The next day, the membrane was washed and treated with a secondary antibody as described in Section 2.10.3

2.11 Molecular biology methods

2.11.1 DNA extraction

During all DNA extraction protocols, concentration of DNA was measured with a Qubit fluorimeter 2.0 using a Qubit dsDNA HS (High-sensitivity) kit following the manufacturer's protocol. Additionally, the 230, 260 and 280 nm absorbance was measured on a Multiskan GO Microplate spectrophotometer (Thermo Scientific).

2.11.1.1 Spheroplast method

This protocol was used for big volumes of yeast culture and the resulting DNA was mainly used for GLOE-Seq library preparation, BrdU-IP, or Southern blotting purposes.

300 OD600 of cells were harvested in a SORVALL RC 6+ (Thermo Scientific) at 5000 rpm for 5 min and washed once in 25 mL 10 mM Tris-HCl pH 8.5. Pelleted cells were resuspended in 4 mL of Y1 buffer supplemented with 400 µL Zymolyase 100T, and incubated in a water-bath at 30°C for 1 hour. Spheroplast formation was checked under the microscope by mixing 5 µL cells with 5 µL 20% SDS. When spheroplasting was close to 100% lysis, samples were pelleted at 2000 g for 5 min. After washing once with Y1 buffer without β-mercaptoethanol, the cells were resuspended in 4.75 mL lysis buffer supplemented with 250 µL 20% SDS and incubated in water-bath at 37°C for 1 hour. Next, the debris were removed by centrifugation at 5000 g for 10 min and the supernatant was transferred to a fresh 50 mL tube. To precipitate proteins, 1.6 mL of 5 M KOAc was added to the supernatant, mixed by inversion and incubated on ice for 45-60 min. Samples were then centrifuged at 17000 g for 15 min and the supernatant was transferred to another 50 mL tube, this step was repeated two more times until most debris were removed.

DNA was precipitated by adding 19.5 mL 100% ethanol, followed by centrifugation at 8000 g for 15 min. DNA was washed with 25 mL 70% ethanol and centrifuged at 8000 g for 10 min. The cleaning step was repeated with 2 mL 70% ethanol and the entire volume was transferred to a 2 mL Eppendorf tube. After DNA centrifugation at 21000 g for 10 min at 4°C, pellets were air-dried and resuspended in 245 µL of TE overnight at 4°C. Next day, DNA was incubated with 5 µL RNase A (10 mg/mL) and 5 µL 5 M NaCl at 55°C for 1 hour. DNA was then precipitated by using 1.5 mL 100% ethanol, spun-down at 21000 g for 10 min, cleaned

once with 1 mL of 70% ethanol and recovered by centrifugation. Finally, DNA was air-dried before resuspension with 250 µL of 10 mM Tris-HCl pH 8.5 at 4°C overnight.

2.11.1.2 Genomic-tip columns

This protocol was used for big volumes of yeast culture and the resulting DNA was mainly used for GLOE-Seq and Southern blotting purposes.

The starting material was around 400 OD600 cells harvested in a SORVALL RC 6+ (Thermo Scientific) at 5000 rpm for 5 min. Cells were washed once in 25 mL 10 mM Tris-HCl pH 8.5 before following the QIAGEN genomic DNA purification procedure according to the manufacturer's handbook. The buffers were made as described by manufacturer and the QIAGEN Genomic-tip 100/G columns were used to collect DNA. Before starting with DNA extraction, all buffers were brought to room temperature and β-mercaptoethanol was added fresh to Y1 buffer.

2.11.1.3 Fast method

This protocol was used for small volumes of saturated yeast cultures and the resulting DNA was used mostly for PCR purposes.

1-2 mL of saturated yeast culture or single colonies were used for DNA extraction using the MasterPure Yeast DNA Purification kit (Epicentre Biotechnologies) as per manufacturer's instructions.

2.11.2 Polymerase Chain Reaction (PCR)

2.11.2.1 Colony PCR

Single yeast colonies were picked with 10 µL tips and transferred into PCR tubes containing 20 µL 0.25% SDS. The PCR tubes were vortexed and boiled at 95°C for 10 min in a Professional TRIO 48 cycler (Biometra) before spinning them down for 10 s.

The PCR mix was set as follow: 10 µL 5x Green GoTaq Reaction Buffer (Promega), 1 µL 10 mM dNTPs mix, 1 µL of each oligonucleotide (10 mM), 0.25 µL Taq DNA polymerase, 2 µL 20% Triton-100, 1 µL of above described boiled supernatant, and water to final volume of 50 µL. The PCR reactions were run on a Professional TRIO 48 cycler (Biometra) with the following thermal program:

Step 1: 95°C for 2 min

Step 2: 95°C for 30 s

Step 3: 50-65°C for 30 s (annealing temperature depending on oligonucleotides)

Step 4: 72°C for 0.5-2 min (extension time depending on fragment size)

Step 5: Repeat step 2-4 for 34 more times

Step 6: 72°C for 5-10 min (final extension)

Step 7: 20°C forever

2.11.2.2 High-fidelity PCR

PCR cassettes were produced with Phusion HF DNA polymerase (NEB). The following volumes were added to each PCR reaction: 10 µL Phusion GC Buffer Pack, 1 µL 10 mM dNTPs mix, 2.5 µL of each oligonucleotide (10 mM), 0.5 µL Phusion DNA polymerase, < 250 ŋg DNA and water to final volume of 50 µL. The PCR reactions were run on a Professional TRIO 48 cycler (Biometra) with the following thermal cycle:

Step 1: 98°C for 2 min

Step 2: 98°C for 10 s

Step 3: 45-72°C for 30 s (annealing temperature depending on oligonucleotides)

Step 4: 72°C for 0.5-2 min (15-30 s per kb)

Step 5: Repeat step 2-4 for 34 more times

Step 6: 72°C for 5-10 min (final extension)

Step 7: 20°C forever

2.11.2.3 Quantitative PCR

Enrichment of the genomic locus of interest in ChIP samples was analysed by qPCR. For one reaction the qPCR was set up as follow: 5 µL 2x SYBR Green PCR Master Mix (Applied Biosystems), 1 µL each oligonucleotides (10 mM), and 3 µL genomic material. Quadruplicates of each sample was loaded into a 384-well MicroAmp optical reaction plates (Applied Biosystems). The plates were closed with transparent adhesive PCR seals (VWR International), vortexed, and briefly centrifuged at 1000 g. The qPCR reactions were run on a Viia7 system cycler with a QuantStudioTM Real-Time PCR software (Life Technologies) using the following program:

Step 1: 95°C for 15 min

Step 2: 95°C for 15 s

Step 3: 55°C for 30 s

Step 4: 72°C for 30 s

Step 5: Repeat step 2-4 for 40 times

Step 6: 95°C for 15 s

Table 2.9. Oligonucleotide pair for qPCR

Genomic locus	Oligonucleotide pair
Origin ARS607	ID 3143: CGGCTCGTGCATTAAGCTTG ID 3144: TGCCGCACGCCAAACATTGC
Origin ARS305	ID 237: TTTCAGAGCCTTCTTTGGAG ID 238: CAAACTCCGTTTTTAGCCCC
Origin ARS501	ID 665: AAGCAAATTGCAGAAGGTTATGAA ID 666: TTCAAGGCTCTAGCATATGAAACG

Fold enrichment of each genomic locus was analysed using the percent input (% Input) method, which normalized the input chromatin for each immunoprecipitation. The following equations were used to get the % Input for each IP.

$$\textbf{Adjusted input} = \text{Ct (Input)} - \log2 \text{ (percentage of used input)}$$

$$\textbf{\% Input} = 100 \times 2^{(\text{Ct (ChIP)} - \text{Adjusted Input})}$$

Finally, the data were plotted as heatmaps using a script created by Dr. Nicola Zilio available at https://github.com/helle-ulrich-lab/plot-heatmap-chip.

2.11.3 Sanger sequencing

For Sanger DNA sequencing, 50-700 ng of PCR product or plasmid DNA was mixed with 1 µL of corresponding oligonucleotide (10 mM) and water to a final volume of 7 µL. Samples were sent to StarSEQ GmbH (Mainz) for sequencing. The sequencing results were checked with SnapGene.

2.11.4 Agarose gel electrophoresis

DNA or PCR products were mixed with 6x Blue DNA loading dye in a 5:1 ratio, and then loaded on a solidified 1% agarose gel containing 1x SYBR Safe DNA Stain (Life Technologies). The agarose gel was run in TBE buffer at 111 V for 30 min in a horizontal gel electrophoresis apparatus (Bio-Rad). New England Biolab's 100 bp or 1 kb DNA ladder was used as a size standard. DNA was visualised with Fusion FX (Vilber).

2.11.5 Chromatin immunoprecipitation sequencing

Around 50 OD600 of yeast cells were collected per each ChIP in 50 mL Falcon tubes. Cells were immediately cross-linked with 1% formaldehyde (Sigma-Aldrich) for 10 min at room temperature, and then quenched with 125 mM glycine for 5 min at room temperature. Next, cells were harvested at 2000 g for 2 min at 4°C and resuspended in 50 mL ice-cold PBS. After cells were washed twice with ice-cold PBS, they were resuspended in 800 µL RIPA buffer (without Triton-100) supplemented with an EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich). This resuspension was transferred to 1.5 mL screw-cap tubes (VWR International)

containing 350 µL 0.5 mm zirconia beads (Carl Roth). These samples were either used fresh or snapped-frozen on dry-ice for storage at –80°C.

The cryo-lysis was performed in a Precellys 24/ Evolution tissue homogenizer (Bertin Instruments) using dry-ice and a built-in lysis settings (Hard: 12 cycles of lysis for 20 s at 6800 rpm and 30 s pause). After lysis, samples were transferred into 15 mL bioruptor tubes (Diagenode) containing 200 µL sonication beads for DNA shearing using a Bioruptor Pico (20 cycles: 30 s on and 30 s off. After 10 cycles a 20 min pause). The cell debris were pelleted by centrifugation at 21000 g for 20 min at 4°C and the supernatant was transferred into 1.5 mL DNA LoBind tubes. At this step, samples could be snapped-frozen on dry-ice and stored at –80°C.

Next, 600 µL of the sonicated sample was incubated with 2-5 µg of antibody for 2 h at 4°C with constant shaking (60 µL of the resuspension was left aside as input sample). Afterwards, 60 µL of ProteinG MagBead MX beads (GenScript), previously washed and blocked (washed in RIPA buffer and blocked in 5 mg/mL BSA, 200 µg/mL glycogen and 250 µg/mL yeast tRNA overnight at 4°C), were added to the samples and incubated for another 2 h at 4°C with constant rotation. Beads were pelleted and the supernatant was discarded. The beads were then washed with ice-cold RIPA buffer, Washing buffer I, Washing buffer II and TE buffer. All the washes were performed for 10 min each at 4°C with constant rotation. After the fourth wash, the beads were pelleted and the remaining TE buffer was removed. Next, 100 µL Elution buffer and 1 µL of RNase A (10 mg/ml) were added to the beads and incubated at 50°C for 2 h. Then, 1 µL Proteinase K was added and samples were again incubated at 50°C for 2 h. Finally, the cross-link was reversed by incubating the samples at 65°C overnight.

The next day, samples were cooled-down at room temperature, the beads were spun-down (1000 g for 1 min), and the supernatant was collected by using a magnetic rack. To purify the DNA from the IPs and inputs, the ChIP-DNA Clean & Concentrator Kit was used according to manufacturer's instructions and 20 µL of nucleases-free water was used for samples' elution. DNA concentration was measured using the Qubit dsDNA HS Kit (Invitrogen), and lastly samples were stored at -20°C until submitting for library preparation to the Genomics Facility at the Institute of Molecular Biology.

2.11.6 Bromodeoxyuridine Immunoprecipitation

100 OD600 of cells grown in the presence BrdU were collected in 50 mL Falcon tubes. DNA was extracted using the spheroplast method (Section 2.11.1.1) and this was quantified using the Qubit dsDNA HS Kit. Before proceeding with immunoprecipitation, the incorporation of BrdU was confirmed by dot-blotting (Section 2.10.6). 6-8 µg of DNA was sheared using a Bioruptor Pico (8 cycles of 30 s on and 90 s off, with a 10 min pause after 4 cycles). The desired DNA size was between 250-300 bp, which was confirmed using a High-Sensitivity D1000 ScreenTape or a Bioanalyzer High- Sensitivity DNA chip on a Tape Station 2200 or 2100 Bioanlyzer (Agilent Technologies), respectively. Sonicated DNA was transferred to 1.5 mL DNA LoBind tubes, recovered with a 1:1 v/v ratio of AMPure XP beads (Beckman Coulter)

and resuspended in 27 µL nuclease-free water. For input samples, 60 ηg of the sheared DNA was transferred to 1.5 mL LoBind tubes and kept aside until DNA purification.

For BrdU-IP, 6 µg sonicated DNA was combined with 5 µL 10x PBS for genomic DNA (Invitrogen) and water to a volume of 50 µL. Sonicated DNA was heat-denatured for 10 min at 95°C and immediately incubated on ice for 5 min. After a brief centrifugation, 4 µg of anti-BrdU antibody (Abcam) was added to each tube and these were incubated on a rotating wheel for 1 h at 4°C in a dark room. Next, samples were briefly spun-down and 50 µL Protein G-Dynabeads (Life Technologies), previously washed and resuspended with Beads buffer, was added and incubated on a rotating wheel for 45 min at 4°C in a dark room. After this time, beads were pelleted and the supernatant was discarded before proceeding with three washes in 1 mL ice-cold IP buffer and one wash in 1 mL ice-cold TE. Beads were resuspended in 100 µL Elution buffer and incubated at 65°C for 15 min. Tubes were briefly centrifuged and supernatant was immediately collected using a magnetic rack. Purification of ssDNA was carried out using the DNA Clean & Concentrator kit (Zymo Research) according to manufacturer's instructions. 17 µL nucleases-free water was used for samples' elution. DNA concentration was measured using the Qubit ssDNA Kit (Invitrogen), and lastly samples were stored at -20^{o}C until submitting for library preparation to the Genomics Facility at the Institute of Molecular Biology.

2.11.7 Southern blotting

2.11.7.1 Denaturing gel

After extraction (Section 2.11.1), DNA was digested overnight with the relevant restriction enzyme at the manufacturer's indicated temperature using 1 unit of each enzyme (NEB) for every 1 µg DNA. AMPureXP beads (Beckman Coulter) at 1:1 v/v ratio was used to recover the digested DNA and 15 µL nuclease-free water was used for DNA elution. 1.5 µg DNA per sample was used for loading on an alkaline gel. For abasic sites protection, 0.1 M fresh $NaBH_4$ (1 M stock made fresh) was added to DNA and incubated at 37^{o}C for 5 min. For alkali treatment, 0.2 M NaOH (5 M stock made fresh) was added and incubated at 65^{o}C for 30 min. All samples were treated with EDTA pH 8.0 to a final concentration of 10 mM followed by the addition 6x Gel Loading Dye, no SDS (NEB). Samples were loaded onto an alkaline gel, run at 120 V for 10 min and then run at 60 V for 18 h using alkaline running buffer in a cold-room. Next, the gel was rinsed once in water, twice in neutralising buffer (alkaline gel) for 25 min and once more with 10x SSC for 10 min before assembling the transfer sandwich (Section 2.11.7.3).

2.11.7.2 Native gel

After DNA was extracted (as described in 2.11.1), 2 µg DNA was digested overnight with the relevant restriction enzyme at the manufacturer's indicated temperature using 1 unit of each enzyme (NEB) for every 1 µg DNA in a final volume of 25 µL. Next, 5 µL of 6x DNA blue loading dye was added to the digested DNA before loading onto a 1% agarose gel. The gel

was initially run at 120 V for 10 min and then run at 60 V for 18 h using 1x TBE buffer in a cold-room. The next day, the gel was briefly rinsed in water followed by one wash in 1 L depurination buffer with gentle shaking for 10 min (or until the bromophenol blue turned yellow). Next, the gel was rinsed twice in water for 5 min, washed in 1 L denaturing solution on shaker for 30 min (or until the bromophenol blue turned purple). The gel was then rinsed twice in water for 5 min, and washed in 1 L neutralising solution with gentle shaking for 30 min. Finally, the gel was quickly rinsed first in water and then once in 10x SSC for 10 min before assembling the transfer sandwich (Section2.11.7.3).

2.11.7.3 Assembling transfer sandwich

DNA was transferred overnight from the gel to a nitrocellulose Hybond-N+ membrane (Thermo Fisher Scientific) by setting up a Southern blot as follows. A 3MM blotting paper (Whatmann) was pre-soaked in 10x SSC and placed on top of a platform positioned on the middle of a tray filled with 10x SSC. The gel was placed upside down, followed by the membrane, which was previously soaked in 10x SSC. The following layers consisted of 25x 3MM blotting paper, followed by 12x thicker papers (550 g/m^2), and a stack of paper towels compressed by a weight of 800 g. The following day the membrane was rinsed (DNA on top) with water for 5 min. After letting the membrane dry at room temperature for at least an hour. The DNA was cross-linked to the membrane using a UV dose of 1200 J/m^2 using a Stratagene-UV Stratalinker.

2.11.7.4 Probes labelling

The Klenow Based: Random Primed DNA Labelling kit (Sigma-Aldrich) was used as per the manufacturer's instructions using α-^{32}P dCTP that was less than 2 weeks old (Perkin Elmer). Unincorporated nucleotides were removed using Probe-Quant G-50 columns (Fisher Scientific). All probes were denatured prior to use by incubating at 95^{o}C for 10 min followed by 5 min on ice.

2.11.7.5 Hybridisation

The membrane was placed into a hybridisation tube containing 20 mL pre-warmed hybridisation buffer and left rotated at 68^{o}C for at least 2 h. The buffer was replaced with 10 mL fresh hybridisation buffer or PerfectHyb™ Plus Hybridization Buffer (Sigma-Aldrich), 100 µg/mL HT-DNA and the probe. Hybridisation was performed overnight at 68^{o}C.

2.11.7.6 Washing blots

The hybridisation buffer was removed and the blot was briefly rinsed with 20 mL low-stringency wash buffer. After two washes in 20 mL low-stringency washing buffer for 5 min, these were repeated twice more in 20 mL high-stringency wash buffer for 10 min. All wash steps occurred at 68^{o}C with rotation.

2.11.7.7 Stripping blots

Southern blots were stripped by adding 50 mL boiling 0.1% SDS to the membrane and incubated at 68°C for 15 min with rotation. This step was repeated at least twice before the blots were hybridised again as described in Section 2.11.7.5.

2.11.7.8 Radioactivity detection

The membrane was dried at room temperature before wrapping it with plastic film and exposing it to the white side of a phosphoimager screen for 2 - 5 nights. The membrane was developed using a Typhoon FLA on the "Phosphoimaging" setup with a 100 µm resolution.

2.11.7.9 Southern blot analysis

Images were analysed and quantified using ImageJ. Bands of equal area were defined over the full-length TetO and chromosome 1 (Chr. 1) control bands. To standardise DNA loading, intensity volume of each TetO band was divided by intensity volume of the corresponding Chr. 1 control band. The ratios expressed as a percentage gave an indication of the intactness of the full-length TetO band.

2.12 Deep-sequencing methods

2.12.1 Library preparation for Genome-wide Ligation of 3′-OH Ends sequencing (GLOE-Seq)

DNA (5 µg) extracted using the spheroplast method described in Section 2.11.1.1 was treated with RNase HII (NEB) for 2 h at 37°C and recovered with AMPureXP beads (Beckman Coulter) at a 1:1 v/v ratio. 2.5 µg of nicked DNA was used for library preparation and the GLOE-Seq protocol was followed according to Sriramachandran *et al.* (2020).

2.12.2 Polymerase usage sequencing

A modified version of already published methods was used (Keszthelyi , *et al.*, 2015; Reijns*, et al.,* 2015; Daigaku*, et al.,* 2015; Koh*, et al.,* 2015; Clausen*, et al.,* 2015). The background of the parallel strains used during this assay included the deletion of the gene RNH201, in addition to a mutation on the gene encoding the catalytic subunit of DNA Pol α (Pol1 L8686M), Pol ε (Pol2 M644G), or Pol δ (Pol3 L612M). After the DNA was extracted (as described in 2.11.1.1), 5 µg DNA was treated with RNase HII (NEB) followed by NGS library preparation using GLOE-seq (Section 2.12.1).

2.12.3 ChIP-Seq and BrdU- Seq library preparation

The Genomics Core Facility at the Institute of Molecular Biology prepared the ChIP-Seq and BrdU-Seq libraries and sequenced them. Strand-specific DNA libraries were performed using the Accel-NGS® 1S Plus DNA Library Kit according to the manufacturer's indications. For BrdU-IP libraries, the final library PCR extension was done with the uracil-tolerant DNA polymerase KAPA HiFi (Roche). The library pools were sequenced on a NextSeq500

sequencer using a High Output 75 Cycle kit (paired-end). Around 5 million reads for each sample were used for the analysis.

2.12.4 Deep-sequencing data analysis

The Bioinformatics Core Facility at the Institute of Molecular Biology analysed the ChIP-seq, BrdU-seq, and GLOE-seq data. The analysis pipeline included: trimming (Trimmomatic), mapping to the yeast genome sacCer3 (bowtie2), BAM indexing (samtools) and normalising BIGWIG tracks (bedtools).

Chapter 3: Results I - Distribution of labour between replicative DNA polymerases in damage conditions

3.1 Background

The process of duplicating the genome is highly regulated and involves a multi-protein replication machinery. Obstacles frequently challenge the replisome progression, and these potentially can stall the DNA replicases, Pol ε and Pol δ, at the template. Thus, in order to preserve genome integrity, a decision needs to be made to resolve the lesion and re-start the stalled fork (Edenberg, et al., 2014). The study of daughter strand synthesis with the emergence of DNA combing assays made it clear that replication fork re-start pathways operate in eukaryotic cells (Tourrière, et al., 2005). However, the understanding behind the mechanisms of replication re-start and the response of the DNA polymerases to damage is still unclear. Multiples *in vitro* studies sustain that replication forks can be rescued by a *de-novo* synthesis of a primer downstream of the damaged region, although through different mechanisms on both the leading and lagging strands (Heller & Marians, 2006; Van, et al., 2010; Taylor & Yeeles, 2018). The canonical model for discontinuous lagging-strand synthesis has explained how stalled forks are re-started by constant cycles of priming events (McInerney & O'Donnell, 2004; Lopes, et al., 2006). In contrast, a pathway to re-prime a stalled-leading-strand synthesis has been the subject of much debate. Studies using a reconstituted yeast replisome indicated that leading strand re-priming by Pol α is inefficient, although it can occur at a subset of stalled forks (Taylor & Yeeles, 2019). Moreover, it was proposed that Pol δ has a temporal and critical role in re-establishing coupled leading strand synthesis (Johnson, et al., 2015; Yeeles, et al., 2017). Altogether, these observations suggested different roles of the main DNA replicases during damage. Additionally, as most of these finding relied on biochemical assays, it is possible that other cellular factors promote the re-start of DNA replication and re-priming *in vivo*.

In this chapter, I aimed to establish and validate a system to study the labour of the main DNA polymerases in replicating a damaged template *in vivo*. To assess this, I constructed strains that endogenously and chronically targeted damage in *S. cerevisiae*. First, I provided initial evidence of the contribution of Pol ε and Pol δ on the synthesis of a template with polymerase-stalling lesions. Yet, I found that this system generated elevated levels of genomic instability. Consequently, in the second part of this chapter, I intended to re-establish and validate the experimental approach to better control the amount of damage induced into cells.

3.2 Experimental approach

I combined two tactics to analyse the usage of the replicative polymerases in the synthesis of damaged DNA. First, I utilised the Apurinic/apyrimidinic (AP) site system (Section 1.1.6.3) to form abasic sites in a TetO array introduced in *S. cerevisiae*. In addition, I integrated the genomic background needed for the mapping of ribonucleotides incorporated on genomic DNA (Section 1.1.7.1).

3.2.1 System to create site-specific abasic sites in budding yeast

Elizabeth Colby, a former member of the Ulrich laboratory, constructed the AP site system. This system induces the controlled formation of abasic sites by taking advantage of the bacterial Tet Operator (TetO) sequence and its binding partner the protein Tet Repressor (TetR).

Site-specific location of damage: The targeting of abasic sites to chromosome VI of budding yeast was achieved by the integration of an array of 44 TetO sequences (TetO44 array). This array was designed to have 10 bp of sequence heterology between each 19 bp operator sequence, therefore reducing the chance recombination events (1542 bp). Originally, an early- (ARS607) and a late- (ARS501) firing origin were chosen as sites for TetO44 array incorporation, however, during this study, I only introduced it near the early firing origin ARS607. Moreover, the TetO44 array was inserted alongside a functional *URA3* gene for the selection of positive clones (Figure 3.1A).

Controlling damage targeting: The formation of abasic sites was achieved by fusing the TetR to a mutant of the human uracil DNA glycosylase enzyme, referred as UDG* (Kavli, et al., 1996); on this manner abasic sites were formed within the integrated TetO array. Along with UDG*, the construct comprises an N-terminal nuclear localisation signal (NLS), a C-terminal Super-glow GFP-tag, the ADH1 terminator sequence, and a functional *LEU2* gene that served for the selection of positive clones. To avoid any interference with the replisome passage and inhibit further damage formation, the construct was fused to the auxin-inducible degron system (AID*) for controlled proteasomal degradation (Nishimura, el al, 2009). Originally the system was created using the galactose promoter, *GAL1/10*, to modulate the expression of the TetR-UDG* construct in cells. The *GAL1/10* promoter is induced by a medium shift to galactose and can be subsequently repressed by shifting the cells into medium containing glucose (Figure 3.1B-C). During this chapter, I explored using two promoters to express UDG*, the previously mentioned galactose promoter and a G1-tag (*SIC1*) promoter to chronically induce the formation of abasic sites every G1-phase (Figure 3.1B-C).

Yeast background: The W303 strain corrected for the *rad5-G535R* mutation background was used in this study. To minimise processing of AP sites into strand breaks, the system was designed in a strain background that lacked the base excision repair (*BERΔ*) proteins, Apn1, Apn2, Ntg1 and Ntg2. Moreover, unless stated I also deleted the endogenous *URA3* gene located in Chromosome V, reason behind this step is detailed in Section 4.2.2.1.

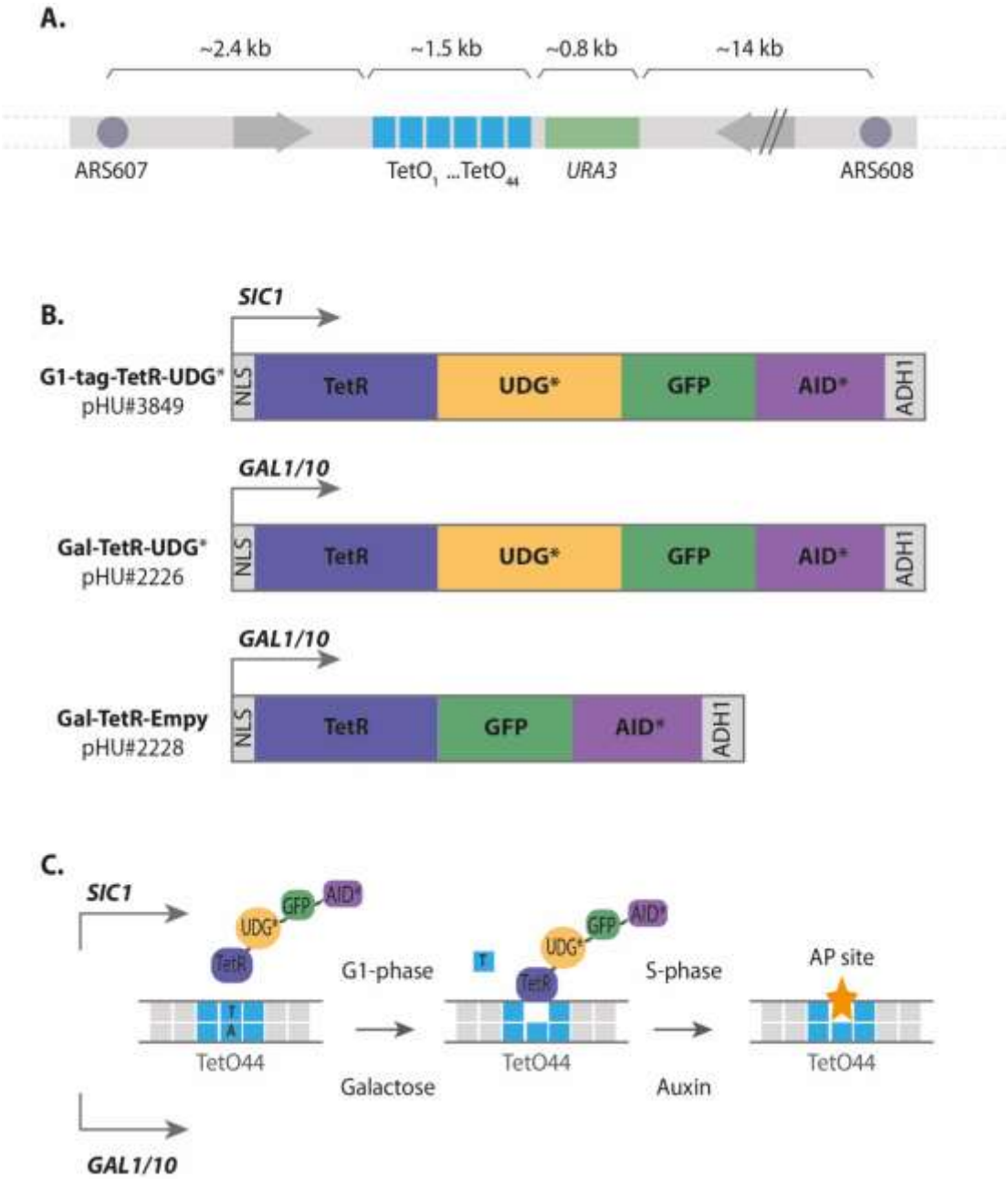

Figure 3.1. Features of the AP site system.
A. TetO44 array (~1.5 kb) was integrated downstream of the early origin of replication ARS607. The late origin of replication ARS608 is located ~14 kb downstream of the TetO44 array. B. TetR constructs used in this chapter C. Schematic induction of abasic sites formation in the TetO44 array by using the G1-tag-TetR-UDG* (*SIC1*) and the Gal-TetR-UDG* (*GAL1/10*) construct.

3.2.2 Mapping ribonucleotides incorporated into DNA

To study the distribution of labour of Pol α, Pol ε and Pol δ in the synthesis of DNA, I used already published principles to map the position of ribonucleotides incorporated into DNA by NGS (Keszthelyi, et al., 2015; Daigaku, et al., 2015; Reijns, et al., 2015; Koh, et al., 2015; Clausen, et al., 2015) (described in detail in Section 1.1.7.1). The strains needed for this assay had a deletion of the gene *RNH201* and each strain contained equivalent mutation in the gene encoding the catalytic subunit of Pol α, Pol ε and Pol δ, which reduced the efficiency

of their steric gate and increased the frequency of incorporation of ribonucleotides into DNA. This method exploits a comparison of ribonucleotide incorporation between mutant strains pairs, thus mapping the relative contribution of these polymerases to DNA replication (Table 3.1, Figure 3.2A).

Table 3.1. Rate of ribonucleotide incorporation in wild-type and polymerases-mutated cells

Wild-type	Polymerase-mutant
Pol α: 1:625	Pol1 L868M: 1:40
Pol ε: 1:1250	Pol2 M644G: 1:100
Pol δ: 1:5000	Pol3 L612M: 1:500

To track directly the genome-wide contribution of polymerases, I used the NGS tool, GLOE-Seq (Sriramachandran, et al., 2020). This was accomplished by treating genomic DNA with RNase HII to generate nicks at the 5′- end of ribonucleotides, followed by ligation of a biotinylated adaptor to the 3′-hydroxyl end of the deoxynucleotide immediately upstream of the ribonucleotide (Reijns, et al., 2015) (Figure 3.2B). This approach allowed me to map the site of ribonucleotide incorporation in the Watson and the Crick strand for each strain. By comparing the number of ribonucleotide incorporation sites between the Watson and Crick strands and between the strain pairs, a map DNA polymerase usage was constructed.

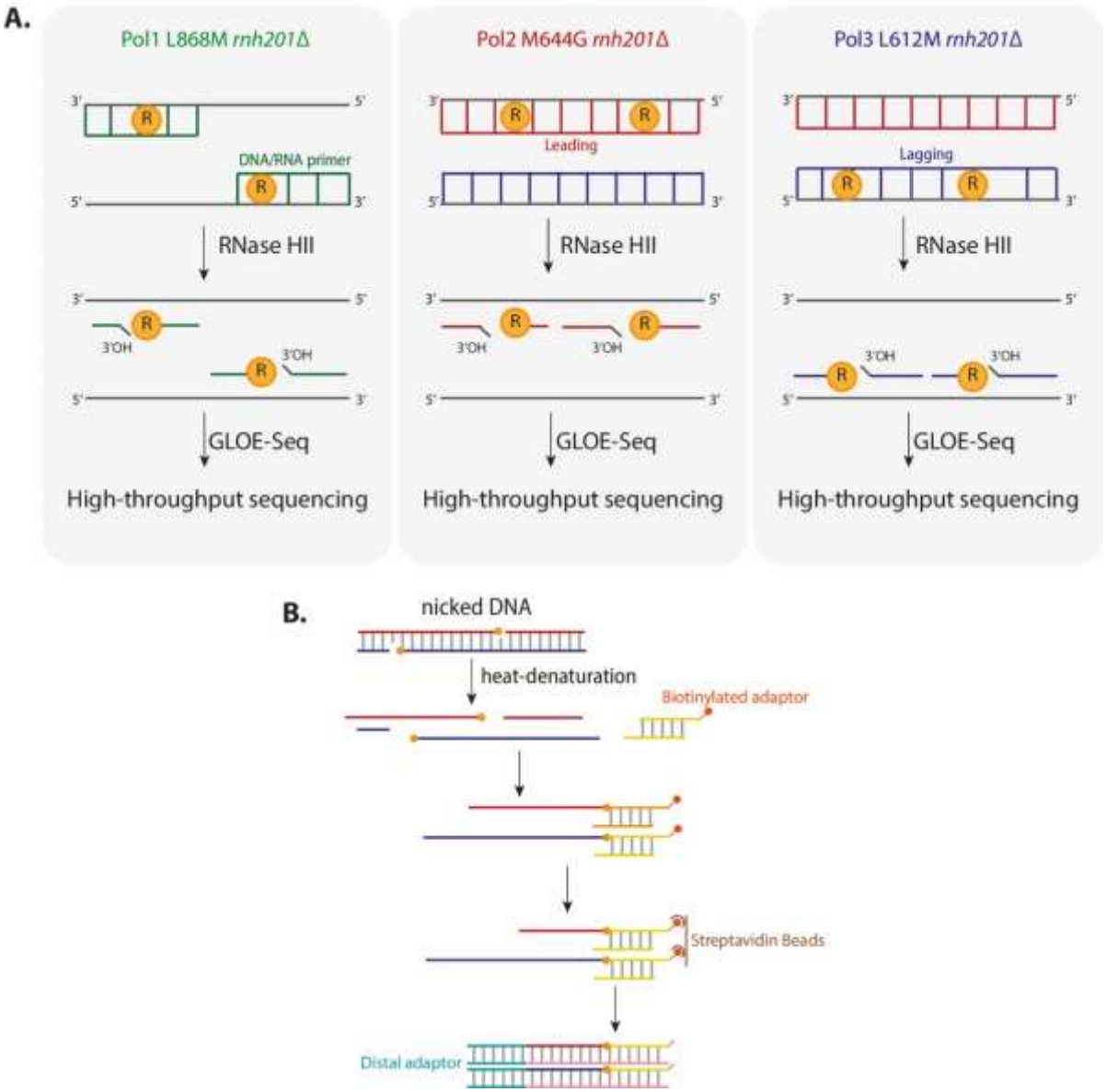

Figure 3.2. Next-generation sequencing tools for mapping ribonucleotides embedded in DNA
A. The division of labour of the DNA polymerases, Pol α, Pol ε and Pol δ was accomplished by identifying and mapping ribonucleotides across the genome by NGS. First, three parallel strains harbouring genetic properties able to introduce ribonucleotides into DNA more frequently were constructed. These ribonucleotides were converted into SSBs by RNase HII treatment, which were then captured by GLOE-Seq for NGS analysis B. For GLOE-Seq, DNA was initially heat-denatured to capture 3'-OH termini with a biotinylated proximal adaptor. The biotinylated ssDNA was recovered with streptavidin beads prior synthesis of its complementary strand. dsDNA was ligated to a distal adaptor, which served for PCR with barcoded oligos and sequencing on a NGS platform (modified from Sriramachandran, et al., 2020).

3.3 A system to induce chronic damage and measure labour of DNA replicases

Strains with deletion of the *RNH201* gene and mutations in the gene encoding the catalytic subunit of Pol ε (Pol2 M644G) and Pol δ (Pol3 L612M) have been characterised in the past (Nick McElhinny, et al., 2010A). In order to study the usage of DNA polymerase in a damaged template, I constructed similar strains and included the background necessary for the AP site system. In addition, to analyse and compare replication events occurring in the absence of abasic sites in the TetO44 array, I constructed Non-damage control strains that lacked the TetR-UDG*construct.

For this study, I needed the following strains:

Table 3.2. Strains needed to assess the labour of DNA polymerases on a damaged template

			Yeast background
1	Non-damage	Pol ε	*BERΔ*, TetO44-*URA3*, Pol2 M644G*, rnh201Δ, ura3Δ*
2	Damaged	Pol ε	*BERΔ*, TetO44-*URA3*, TetR-UDG*-GFP-AID*, Pol2 M644G, *rnh201Δ, ura3Δ*
3	Non-damage	Pol δ	*BERΔ*, TetO44-*URA3*, Pol3 L612M , *rnh201Δ, ura3Δ*
4	Damaged	Pol δ	*BERΔ,* TetO44-*URA3*, TetR- UDG*-GFP- AID*, Pol3 L612M, *rnh201Δ, ura3Δ*

In agreement with Koh, et al., (2015), Keszthelyi, et al., (2015) and Orebaugh, et al., (2018), I planned to collect cells grown to stationary phase in order to accumulate ribonucleotides in genomic DNA. Using the *GAL1/10* promoter to control UDG* expression was not an option, because AP sites would have formed only once after induction. Furthermore, there was the risk that the damage be repaired in early stages of cell growth (e.g. by TLS or TS), which would have masked the real labour of the DNA replicases in a damaged template. Instead, abasic sites needed to be accumulated chronically during cell growth. Thus, I decided to restrict UDG* expression to G1-phase by using a G1-tag promoter, which consisted of a fragment of the cyclin-dependent kinase inhibitor, Sic1 (G1-tag-TetR-UDG*) (Figure 3.1B).

In *S. cerevisiae,* Sic1 is known to control G1- to S-phase transition by blocking the activity of Cdk1, a kinase required for DNA replication. In addition, multisite phosphorylation of Sic1 is thought to time its ubiquitylation and destruction, and by extension, the timing of S-phase entry (Verma, et al., 1997; Tripodi, et al., 2007). It follows that the transition to S-phase will degrade our G1-tag-TetR-UDG* construct, circumventing the use of auxin degron system.

3.3.1 Formation of abasic sites using the construct G1-tag-TetR-UDG*

The AP site system was originally established to induce the expression of UDG* using the *GAL1/10* promoter, as it has a faster induction time and reduced leakiness in comparison to other promoters such as the *CUP1* promoter (Peng, et al., 2015). In her doctoral thesis, Elizabeth Colby confirmed that the Gal-TetR-UDG* construct was able to induce abasic sites within the TetO44 array. She found that abasic sites were formed up to 10 kb away from the array, and prominently she observed that different levels of UDG* were a significant factor for how far they extended. Low UDG* levels, for instance, formed abasic sites at distances up to 5 kb away from the TetO44 array. Moreover, it was seen that AP sites were not spread evenly within the TetO44 array.

In this project, abasic sites needed to be kept intact to be able to accumulate ribonucleotides within the damaged region. Thus, I exchanged the *GAL1/10* promoter to induce UDG* to the G1-tag promoter and tested if abasic sites are formed comparably to what was previously observed by Colby. By following the GFP tag on the TetR-UDG* constructs, I checked if the *SIC1* promoter expressed UDG* chronically.

Illustrated in Figure 3.3, I confirmed that UDG* was persistently expressed in asynchronous cells (Line 3). In addition, in order to compare the acute induction of UDG*using the Gal-promoter and the G1-tag-promoter, I incubated the cells for 2 h with galactose or α-factor (Line 2 & 4), respectively, and I observed that UDG* was induced at similar levels with both promoters.

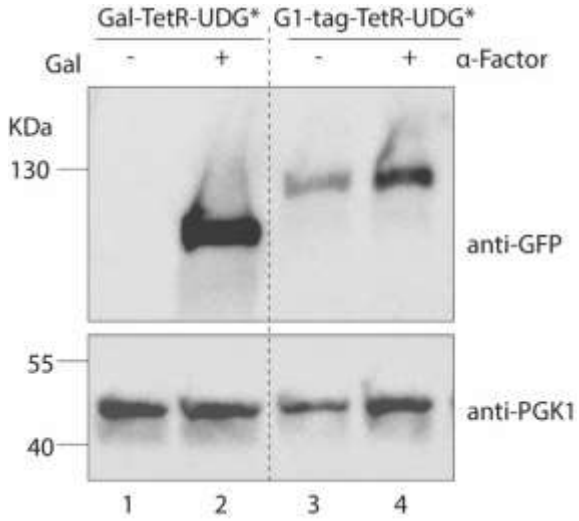

Figure 3.3 UDG* levels using different TetR promoters
Western blotting indicating the expression of the TetR-UDG* construct using the *GAL1/10* and G1-tag promoter. Cells were treated with and without galactose or α-factor (2 hours), respectively. *Strain ID:4442.*

Base on the premise that abasic sites are converted into strand breaks upon strong alkaline conditions (Strauss, et al., 1997), next, I performed an alkaline Southern blot to examine whether the G1-tag-TetR-UDG* construct efficiently formed abasic sites in the TetO44 array. The Southern blot started with the extraction of DNA from cells with chronic and acute UDG* expression. The DNA was then digested with enzymes that cleaved either side of the TetO44 array and posteriorly DNA was incubated with NaOH to cleave abasic sites. To discriminate the presence of SSBs due to the formation of *in vivo* by-products of repair pathways, I treated my samples prior NaOH with sodium borohydride ($NaBH_4$). Sodium borohydride provided protection to abasic sites alkaline cleavage by reducing their aldehyde group to an alcohol (Mazumder, et al., 1996) (Figure 3.4A). Furthermore, I used a sample where UDG* was not expressed as a control devoid of abasic sites. This was the control also used to quantify the intactness of the TetO44 array upon induction of UDG*.

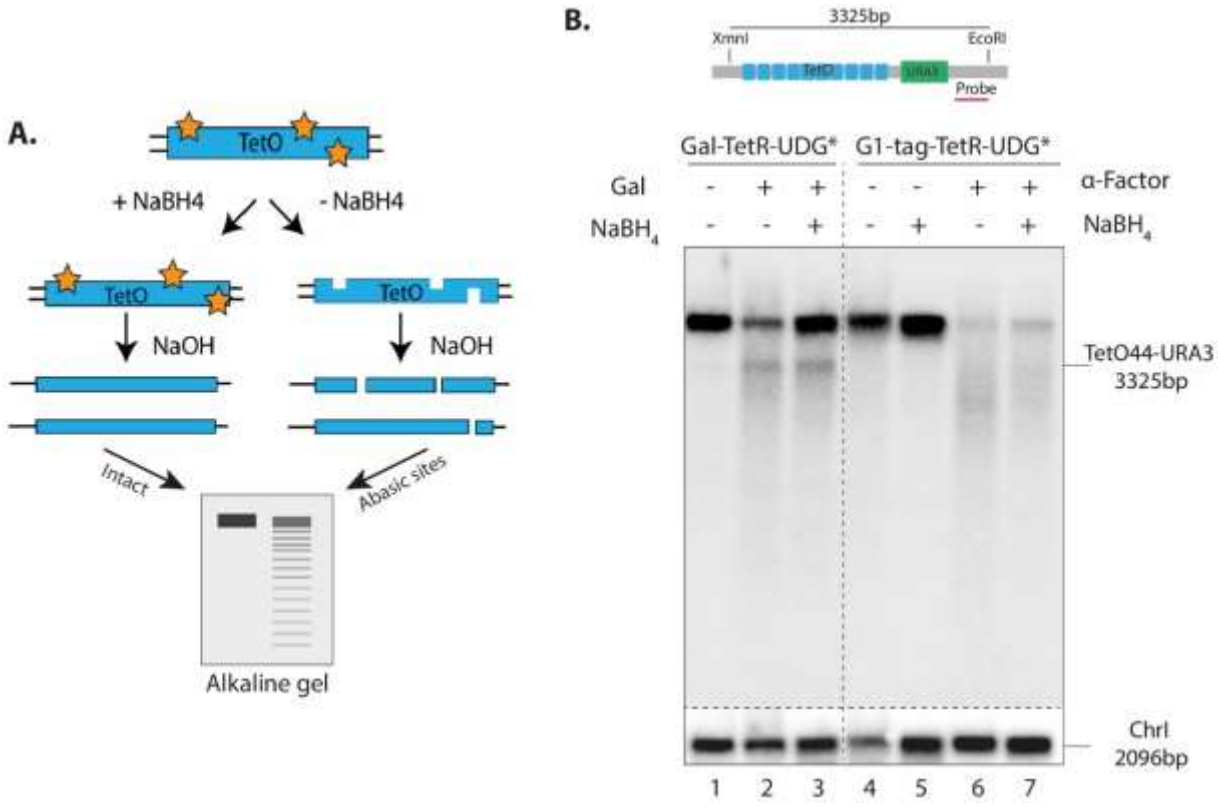

Figure 3.4 Inducing abasic sites at the TetO44 array using the G1-tag-TetR-UDG*
A. Scheme of alkaline Southern blot. **B.** Alkaline Southern blot of DNA digested with XmnI/EcoRI and treated with NaBH4 (0.1 M, 37°C, 5 min) where indicated. Then samples were treated with NaOH (0.2 M, 65°C, 30 min). 2 µg DNA was loaded per lane on a 1% alkaline agarose gel. The blot was first probed for the TetO44 array, and then probed for the control region on Chr1. Representative figure of two repetitions. *Strain ID:4442.*

Table 3.3. Percentage of intactness of the TetO44 array after expression of different TetR constructs

	TetR Construct	% TetO44 intactness
1	Control	96
2	Gal-TetR-UDG*	47
3	Gal-TetR-UDG*	83
4	G1-tag-TetR-UDG*	72
5	G1-tag-TetR-UDG*	86
6	G1-tag-TetR-UDG*	11
7	G1-tag-TetR-UDG*	13

The alkaline gel in Figure 3.4B showed the formation of a smear in both chronic and acute UDG* expressing samples without $NaBH_4$ treatment (Line 4 and 6), confirming that the G1-tag-TetR-UDG* construct efficiently induced the formation of AP sites at the TetO44 array. The quantification of the TetO44 array band upon chronic and acute UDG* expression was 72 and 11% (Table 3.3), respectively, clearly confirming greater damage of the TetO44 array using the acute induction. Remarkably, the acute G1-tag-TetR-UDG* expression formed a smear even in the presence of $NaBH_4$ (Line 7), possibly indicating the presence of single-stranded breaks *in vivo*, or the presence of abasic sites too close to each other that provoked the formation of strand breaks before the alkaline treatment. Thus, during this study the acute induction of the G1-tag-TetR-UDG* construct was not used.

3.3.2 Strains construction

As abasic sites were efficiently formed using the G1-tag-TetR-UDG* construct, I proceeded with the construction of the strains mentioned in Table 3.2. I tried different cell mating and genetic transformation approaches to obtain the desired background. First, I crossed different a- and α-MAT haploid cells with different combinations of the required genetic features; however, none of these strategies produced zygotes. I also attempted to do consecutive alkali transformations in combination with PEG to stimulate plasmid DNA uptake through heat-shock. Only one sequence of transformations allowed me to obtain the genetic background needed (order of transformation as presented in Table 3.4).

All the clones selected for this study had a full-length TetO44 array, confirmed by Sanger sequencing (not shown). In addition, different protein levels of UDG* were seen to play a significant factor in how abasic sites spread within the TetO44 array (Elizabeth Colby, unpublished), so I chose clones with different level of UDG*, which I defined as low and high UDG* clones (Figure 3.5).

Table 3.4. Strains to assess the labour of DNA polymerases - G1-tag-TetR-UDG*

ID	Level of UDG*	Yeast background
4766	Non-damage	*BERΔ*, Pol2 M644G, *rnh201Δ,* TetO44-*URA3, ura3Δ*
4768	Low-UDG*	*BERΔ*, Pol2 M644G, *rnh201Δ,* TetO44-*URA3*, G1-tag-TetR- UDG*-GFP-AID*, *ura3Δ*
4767	High-UDG*	*BERΔ*, Pol2 M644G, *rnh201Δ,* TetO44-*URA3*, G1-tag-TetR- UDG*-GFP-AID*, *ura3Δ*
4764	Non-damage	*BERΔ*, Pol3 L612M , *rnh201Δ,* TetO44-*URA3*, *ura3Δ*
4765	Low-UDG*	*BERΔ,* Pol3 L612M, *rnh201Δ,* TetO44-*URA3*, G1-tag-TetR- UDG*-GFP-AID*, *ura3Δ*
4765	High-UDG*	*BERΔ,* Pol3 L612M, *rnh201Δ,* TetO44-*URA3*, G1-tag-TetR- UDG*-GFP-AID*, *ura3Δ*

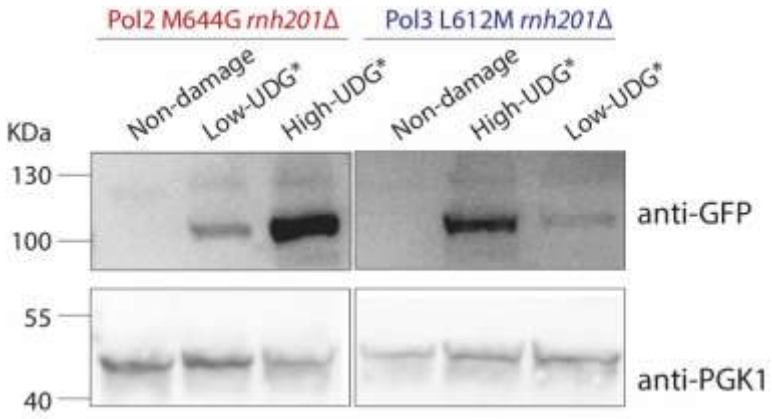

Figure 3.5 UDG* levels of strains used in this study
Western blotting indicating the level of the G1-tag-TetR-UDG* construct in unsynchronised exponentially growing cells. *Strain ID: 4766, 4768, 4767,4764, 4765.*

3.3.2.1 Characterisation of strains

A few studies have documented that cells with mutations in the catalytic subunit of the DNA polymerases, Pol α, Pol ε and Pol δ, exhibit slow growth rates and sensitivity to damage (Pavlov, et al., 2001; Li, et al., 2005; Venkatesan, et al., 2006; Nick McElhinny, et al., 2010A). Thus, to examine possible consequences of unrepaired ribonucleotides incorporated into DNA *in vivo*, I compared several properties on the six yeast strains (Table 3.4) including cell doubling time, colony formation, and cell cycle progression.

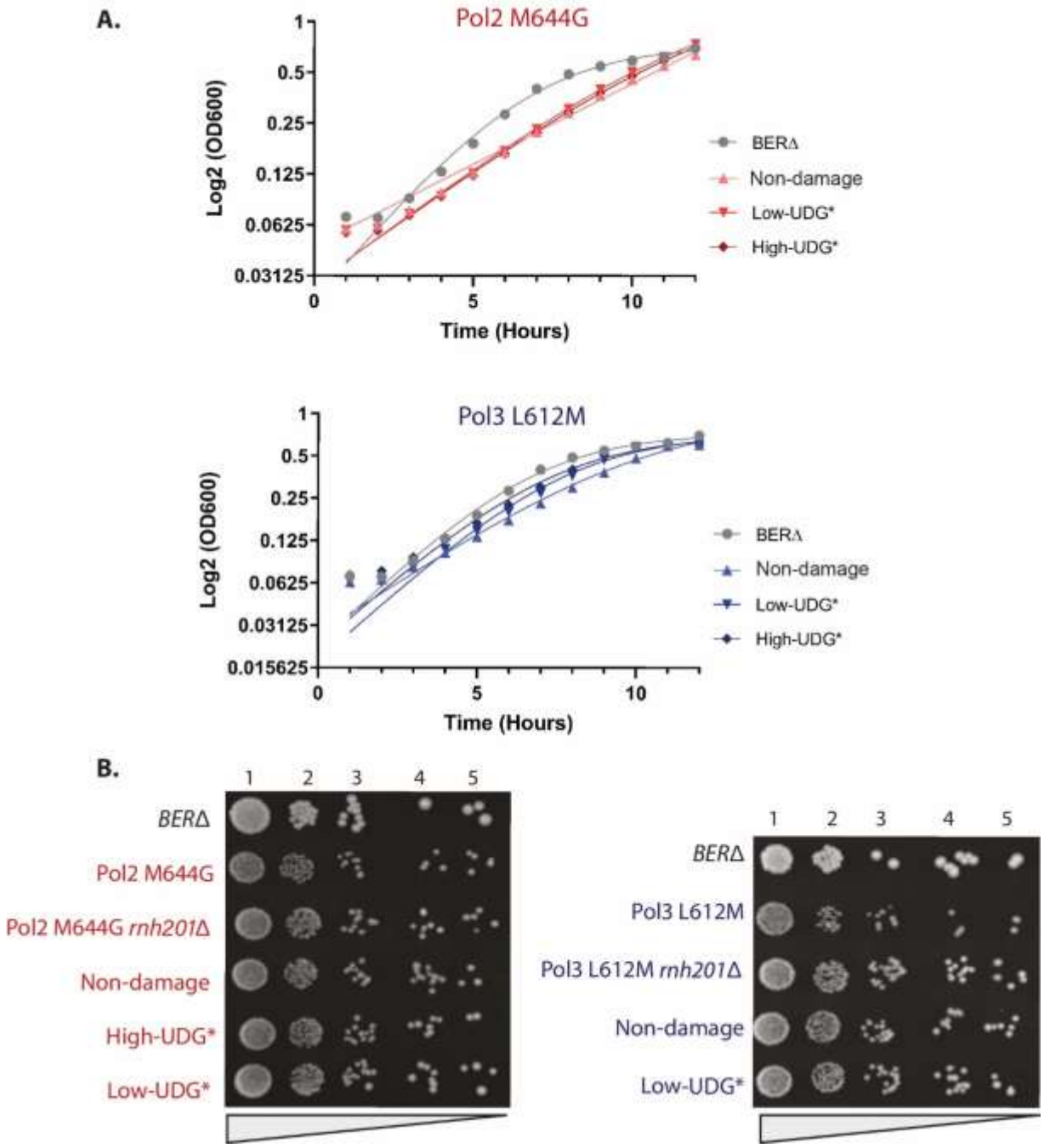

Figure 3.6. Characterisation of cell growth and colony formation of polymerase-mutant strains.
A. To compare cell growth of the different polymerase-mutant strains, OD600 readings were taken every one hour overnight in a Tecan SPARK 20M. Growth curves were plotted using the average of three replicates per run. B. Spot assays were performed using 10-fold serial dilutions of exponential growing cultures (1: 1:12, 2: 1:125, 3: 1:250, 4: 1:2500, 5:1:5000). *Strain ID: 4766, 4768, 4767,4764, 4765.*

Yeast strains enclosing the genetic background of AP site system (*BERΔ*) showed to have a longer doubling time (~2 hours) than wild-type W303 cells (~1.5 hours) (not shown). In addition, I observed that while Pol δ-mutant strains grew at similar rates than the *BERΔ* strain, Pol ε-mutant strains grew slightly slower. Indeed, the calculation of their doubling time was ~2.3 and ~2.7 hours, respectively, which indicated longer growth effects when Pol ε is mutated (Figure 3.6A). Moreover, I performed a spot assay to compare colony formation in these strains and it seemed that both Pol ε (Pol2 M644G) and Pol δ (Pol3 L612M) cells, either with or without an expressing UDG*, colonies size was heterogeneous and mostly smaller compared to the *BERΔ* strain (Figure 3.6B).

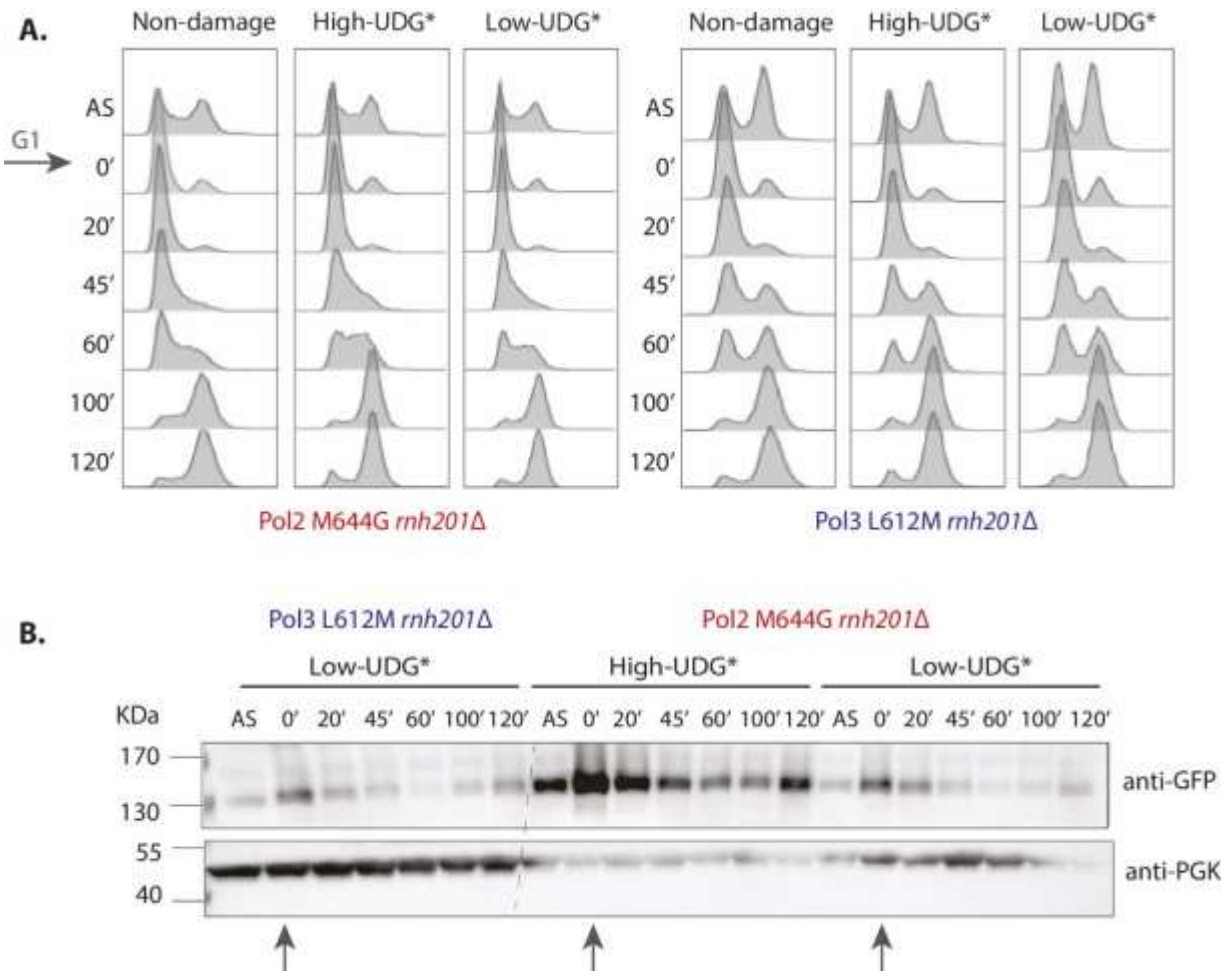

Figure 3.7. Cell cycle progression and chronic expression of UDG*

A. Analysis of cell cycle progression of G1-phase synchronised cells using a-factor (0.75 µg/ml) for 2 hours before release into cell cycle. **B.** Western blotting indicating the levels of UDG* throughout cell cycle. Grey arrows indicate cells in G1-phase and peaks of UDG* expression. During this setup an acute expression of UDG* was used. *Strain ID: 4766, 4768, 4767,4764, 4765.*

It was noted differences between Pol ε and Pol δ cell cycle profiles. In accordance with previous observations by Nick McElhinny, et al., (2010A), Pol2 M644G cells stalled longer in S-phase compared to Pol3 L612M cells (Figure 3.7A). This accumulation in S-phase might be

related to the slightly lower catalytic activity of Pol2 M644G, which could have altered its partitioning between polymerase and exonuclease activity. Another explanation could be related to observations of elevated dNTP pools on Pol2 M644G strains (Nick McElhinny, et al., 2010B), which together with S-phase delay can be signs of DNA replication stress. Furthermore, shown in Figure 3.7B, the strain with low expression of UDG*, its levels peaked at G1-phase and faded away in S- and G2-phases. On the contrary in high expressing UDG* stains, although the signal was higher in G1-phase, this did not decline completely over time.

3.3.3 The usage of DNA polymerases in the TetO44 array

The characterisation of the polymerase-mutant strains did not show a cell growth, colony morphology, or a cell cycle progression profile that indicated a highly compromised cell viability. Thus, I continued with the analysis of the usage of DNA replicases in a damaged template. For this, cell cultures were grown to stationary phase before collection and DNA extraction. I used RNase HII to create nicks at the 5'- end of the ribonucleotide incorporated into the genome of my strains. Then, I captured the upstream 3'-OH end for library preparation by GLOE-Seq. Library pools were sequenced on a NextSeq500 sequencer using a High-Output 75 Cycle kit. Around 5 million reads per sample were used to calculate the ratio of the contribution of Pol δ and Pol ε from each strain pair, in this manner determining the distribution of labour of each DNA polymerase in the leading and lagging strands of the region of interests, the TetO44 array. All sequencing results were analysed by the Bioinformatics Facility at IMB.

3.3.3.1 The ratio of contribution Pol δ and Pol ε to determine their distribution of labour

To determine the distribution of labour of Pol δ and Pol ε, GLOE-Seq data was processed as described in Section 2.12.4 and the ratio of the contribution of Pol δ and Pol ε was calculated similarly to Pu-Seq (Keszthelyi, et al., 2015).

The sequencing results for each strain pair produced four data sets containing the bin reads count on both the Watson and Crick strand from each Pol ε- and Pol δ-mutant strain. To enable a direct comparison between the strain pairs, reads count within each position (x) for each data set were normalised (N) to the sum of all the counts in the given data set. Then under the assumption that each strand was synthesised by either Pol δ or Pol ε, the relative usage of Pol δ and Pol ε on the Watson and Crick strand was calculated.

The ratio of contribution (R) of Pol δ synthesis was calculated as follows:

$$R\text{Watson}(x)\delta = N\text{Watson}(x)\delta \,/\, N\text{Watson}(x)\delta + N\text{Watson}(x)\varepsilon$$

$$R\text{Crick}(x)\delta = N\text{Crick}\,(x)\delta \,/\, N\text{Crick}\,(x)\delta + N\text{Crick}(x)\varepsilon$$

As this result is a relative number, the ratio of Pol ε was calculated as follows:

$$R\text{Watson}(x)\varepsilon = 1 - R\text{Watson}(x)\delta$$

$$R\text{Crick}(x)\varepsilon = 1 - R\text{Crick}(x)\delta$$

Therefore if the ratio of contribution (R) of Pol δ and Pol ε was,

- Equal to 0.5: There was equal usage of both Pol δ and Pol ε
- More than 0.5: Pol δ was preferentially used
- Less than 0.5: Pol ε was preferentially used

3.3.3.2 Genome-wide usage of DNA polymerases

Previous studies have supported the model wherein Pol α and Pol δ primarily replicate lagging strands while Pol ε is mainly responsible for the replication of leading strands (Larrea, et al., 2010; Lujan, et al., 2014). Thus, to confirm this statement further and validate my system, I determined the division of labour between Pol ε and Pol δ across the entire genome by analysing the sequencing results from the strain pair "Non-damage Pol2 M644G" and "Non-damage Pol3 L612M". The ratio of contribution of Pol δ and Pol ε was calculated on a position 5 kb region from ~400 consensus ORIs distributed across the genome of budding yeast.

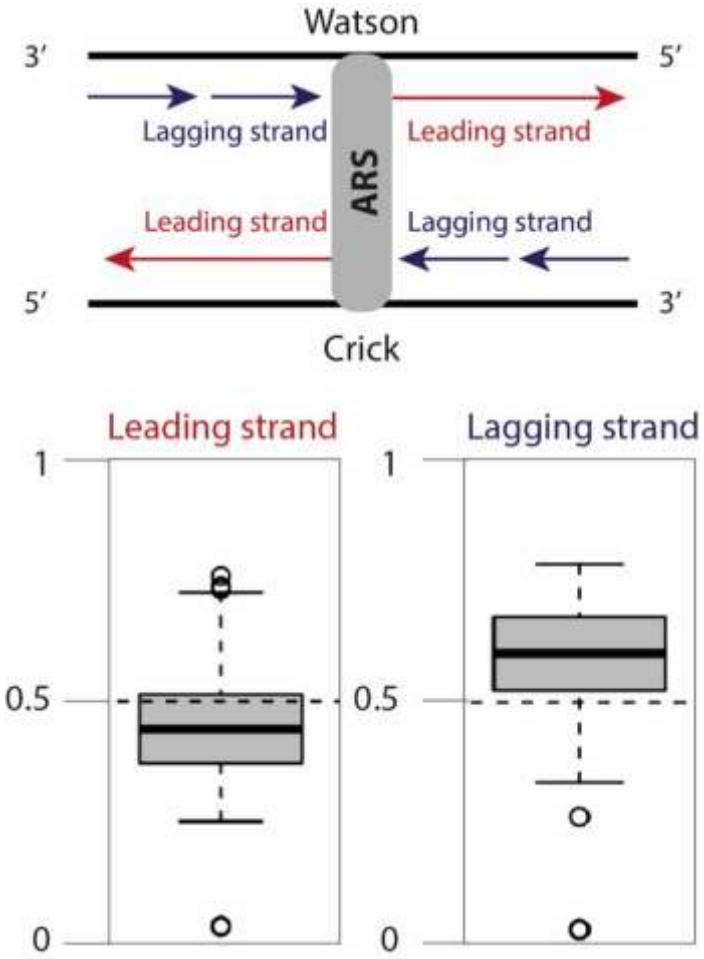

Figure 3.8. The genome-wide usage of Pol ε and Pol δ
Meta-analysis of the DNA polymerases usage of consensus origins of replication (ORIs) (400). The ratio of contribution (R) of Pol δ and Pol ε was calculated from both the Watson and Crick strands at a distance of 5 kb from the ORIs. *R=0.5: equal usage for both Pol δ and Pol ε. R>0.5: preferential usage of Pol δ. R<0.5: preferential usage of Pol ε.* Representative figure of four repetitions. *Plot provided by Giuseppe Petrosino (Bioinformatics CF). Strains ID: 4766 and 4764. Project ID in Annexes.*

I confirmed that the division of labour is broadly maintained across an entire genome, by Pol ε and Pol δ synthesising the leading and lagging strands, respectively (Figure 3.8).

3.3.3.3 Evidence of Pol δ role in leading strand synthesis under damaged conditions

Then, I checked how the TetO44 array was synthesised in the presence of abasic sites, so the ratio of contribution of Pol δ and Pol ε usage was calculated on three different regions within the damaged locus: upstream of the array (-2 kb), the TetO44 array (~1.5 kb) and downstream of the array (+2 kb) (Figure 3.9A).

The following parallel strain pairs were compared for the analysis:

Non-damage Pol3 L612M *rnh201Δ*	VS.	Non-damage Pol2 M644G *rnh201Δ*
Low-UDG* Pol3 L612M rnh201Δ	VS.	Low-UDG* Pol2 M644G *rnh201Δ*
High-UDG* Pol3 L612M *rnh201Δ*	VS.	High-UDG* Pol2 M644G *rnh201Δ*

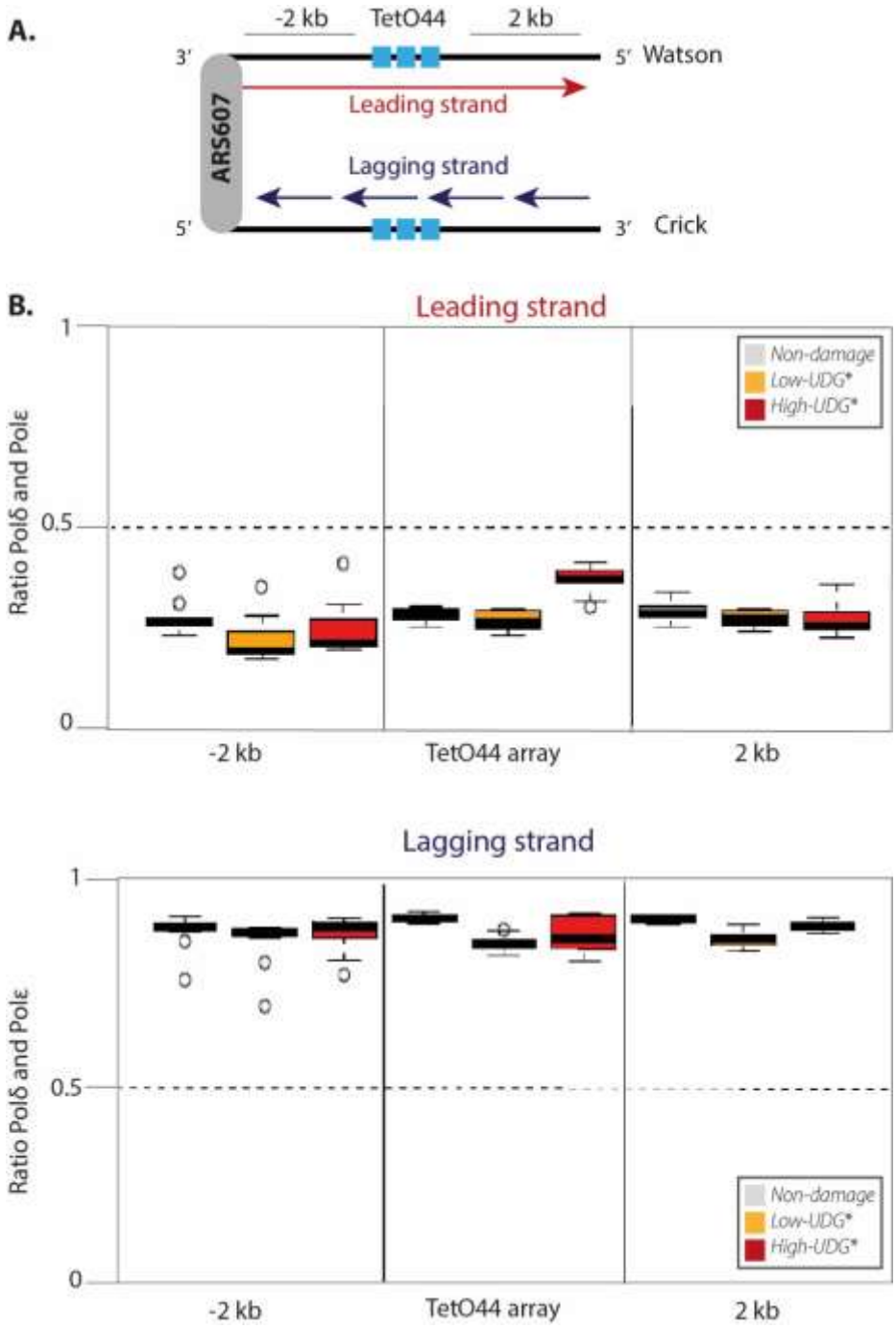

Figure 3.9. Analysis of polymerase usage in the TetO44 array and adjacent regions (-/+ 2 kb).
A. Scheme of the expected replication directionality of Watson and Crick strands at the TetO44 array. **B**. Bar plots indicating the ratio of contribution (R) of Pol δ and Pol ε usage in the strain pairs Non-damage (grey), Low-UDG*(yellow), and High-UDG*(red). *R=0.5: equal usage for both Pol δ and Pol ε. R>0.5: preferential usage of Pol δ. R<0.5: preferential usage of Pol* ε. Representative figure of four repetitions. *Plots provided by Giuseppe Petrosino (Bioinformatics CF). Strain ID: 4766, 4768, 4767,4764, 4765. Project ID in Annexes.*

Pol δ usage in strains with or without damage showed that this polymerase was indeed mainly in charge of synthesising the lagging strand. In addition, the contribution of Pol δ was similar through all the strain pairs, this suggested that the usage of Pol δ is the same in an undamaged and a damaged template (Figure 3.9B). Congruently with Taylor & Yeeles (2018), these findings could imply that the stalling of Pol δ on the lagging strand did not affect ongoing replisome progression because of the occurrence of repeated priming events downstream of the abasic sites. Importantly, this trend was observed in four different biological replicates of the experiment.

On the other hand, the contribution of Pol ε was not only biased towards the leading strand either when UDG* was expressed or not, but it appeared to be rather variable. In the TetO44 array, specially, the contribution of Pol ε in the High-UDG* strains demonstrated a lower contribution of Pol ε usage than in the unchallenged conditions. This indicated preferential role of Pol δ in synthesising the leading strand under high amounts of damage (Figure 3.9B). Even though this trend was observed only in two of four biological replicates, these results could suggest that Pol δ facilitated uncoupled tolerance of polymerase-stalling lesions on the leading strand.

3.3.4 Chronic induction of abasic sites caused the shortening of the TetO44 array

Out of four independent analysis of polymerase usage that were performed, I observed discrepancies amongst the replicates. To identify the possible reasons, I went a step back and observed if the count of sequenced reads distribution within the TetO44 array was similar between strain pairs and replicates. Here, I noticed that, in all the replicates, fewer reads mapped to the TetO44 array than adjacent regions (+/- 2 kb) in both Pol δ- and Pol ε-mutant strains with and without UDG* induction (Figure 3.10). This could be due a low efficiency of *in vivo* replication of the TetO44 array, or due the repetitive nature of the TetO44 array represented a technical challenge for deep-sequencing (Treangen & Salzberg, 2011).

In the regions of interest, I also observed that some areas completely lacked reads (Figure 3.10B). Lack usage of one of the DNA polymerase could explain this observation. That said, previous findings in *S. pombe* showed that stalled replication forks are recombinogenic, inducing translocations and deletions (Lambert, et al., 2005). Thus, the lack of read counts could also mean that the chronic induction of abasic sites triggered the deletion of regions of the TetO44 array. The reasoning behind this idea was that a persistently stalled fork may provoke strand breaks, which, when repaired by a mutagenic recombination repair pathway (e.g. salvage HR), would result in the shortening of the TetO44 array.

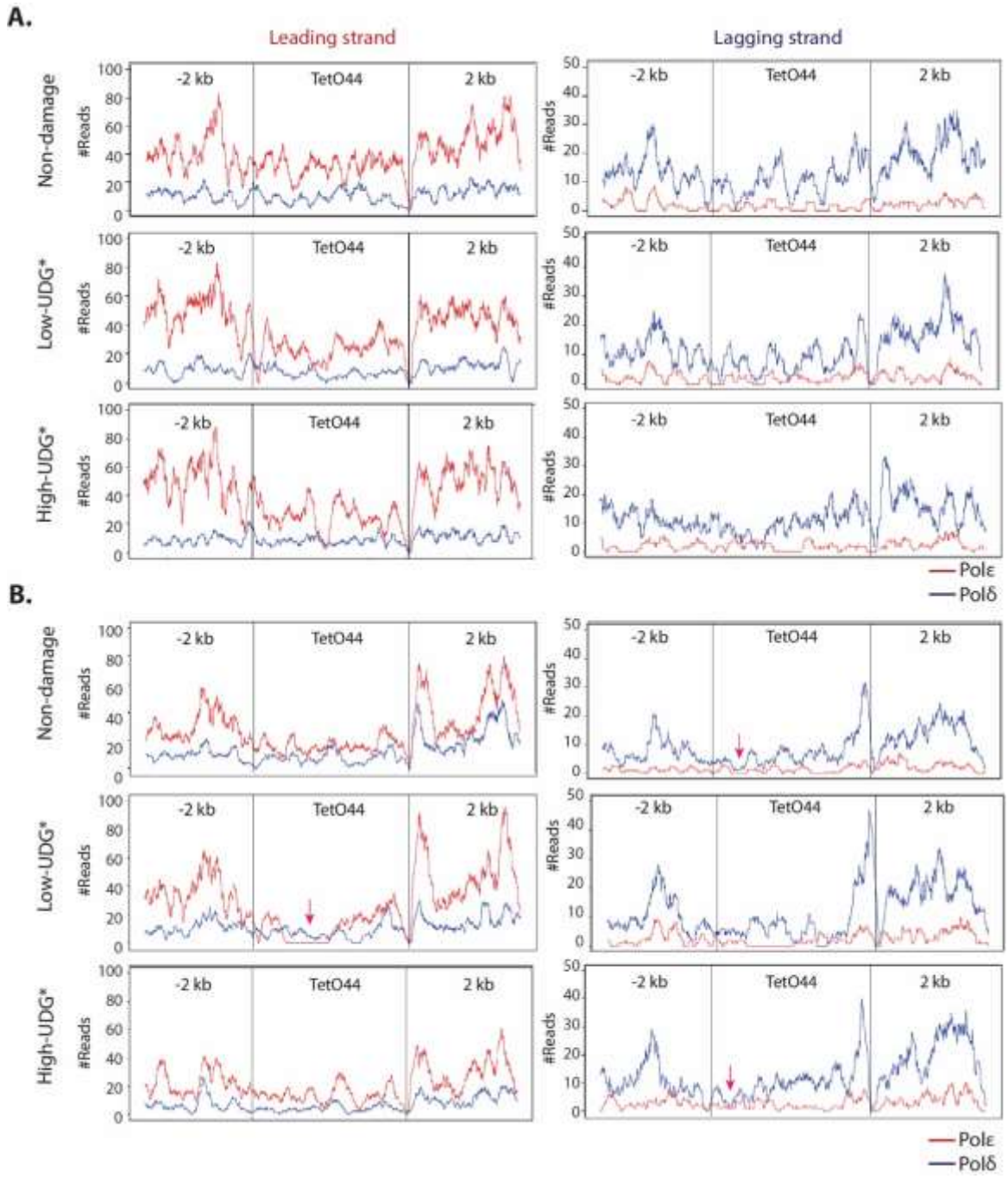

Figure 3.10 Distribution of sequencing read counts for the polymerase usage analysis

Normalised number of reads for each Pol δ- and Pol ε-mutant strain sets in Non-damage, Low-UDG* and High-UDG* conditions. **A**. Replicate 1. **B**. Replicate 3. Pink arrows indicate regions without reads. *Plots provided by Giuseppe Petrosino (Bioinformatics CF). Strain ID: 4766, 4768, 4767,4764, 4765. Project ID in Annexes.*

In order to investigate if the TetO44 array shortened during the chronic induction of UDG*, I performed PCR for three regions within the array. As illustrated in Figure 3.11A, the products showed that the size of the TetO44 array varied from one strain to the other. In fact, in some of the replicates, fragments of the TetO44 were not amplified at all (not shown), possibly indicating that the oligo binding site at the array was absent, even further suggesting deletions of the TetO44 sequence occurred.

The analysis of these PCR fragments by Sanger sequencing allowed me to analyse and compare the sequence of TetO44 array in the presence and absence of damage. As seen in Table 3.5, I confirmed that the length of the TetO44 array was reduced in the strains that chronically accumulated abasic sites, which implied the occurrence of DNA strand breaks and likely posterior activation of a mutagenic repair pathway (Figure 3.11B; Table 3.3). Moreover, the TetO44 array length in the Non-damage control strains was not altered, hence I concluded that loss of sequence from the TetO44 array was due the constant induction of UDG* rather than the polymerase-mutant background.

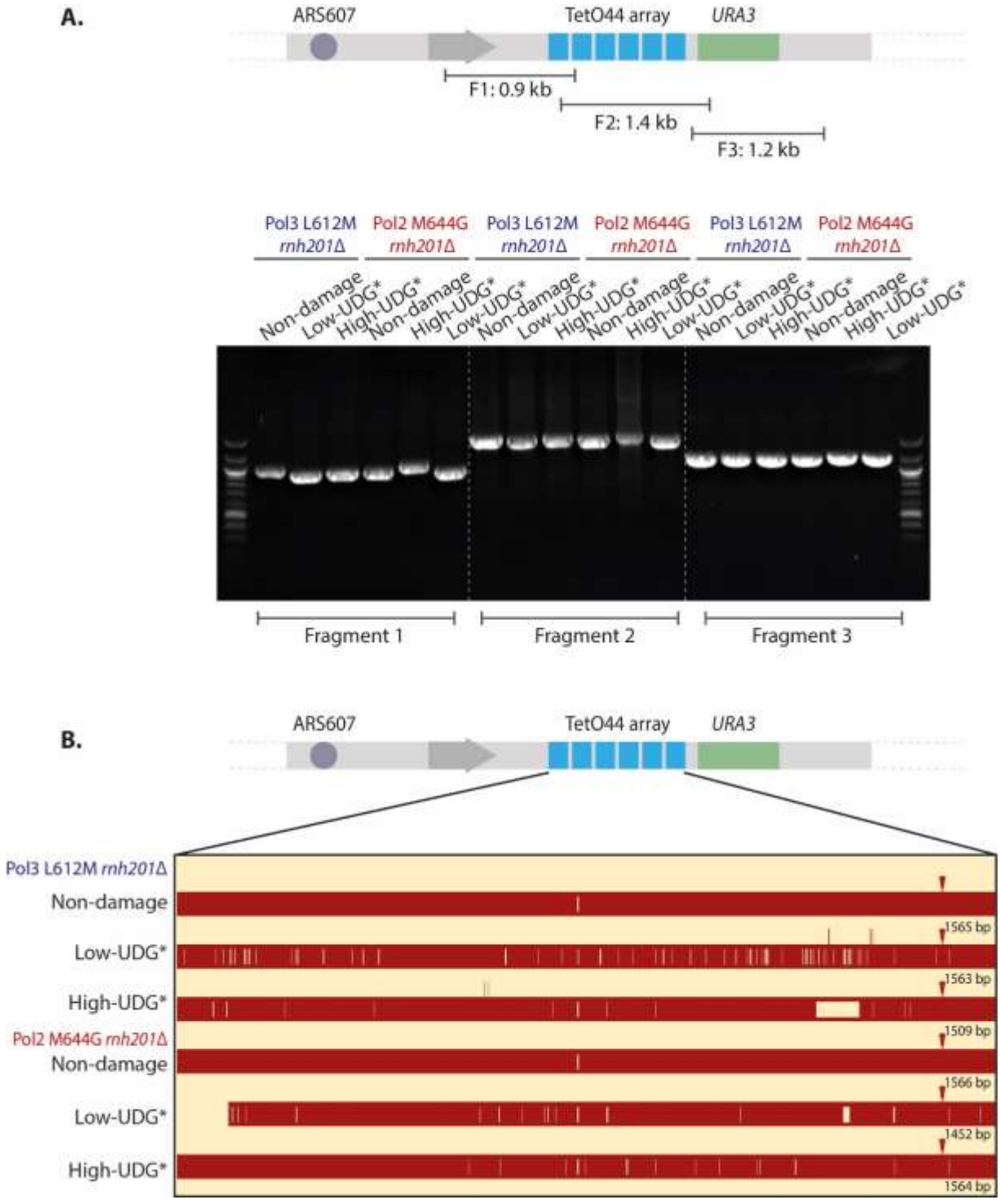

Figure 3.11. PCR and Sanger sequencing of the TetO44 array

A. Scheme of the amplified regions within the TetO44 array and visualisation of PCR amplicons on an agarose gel generated from the input DNA used for the polymerase usage analysis. **B.** Homology maps showing the sequences of the TetO44 arrays of the polymerase-mutant strains with and without damage. Maps were created using Sanger sequencing results of amplified fragments and illustrated in SnapGene V5.0. Representative figure of four repetitions. *Strain ID: 4766, 4768, 4767,4764, 4765*

Table 3.5. Features of the TetO44 array sequence

	ID	Level of UDG*	Size (bp)	Gaps	Mismatches
Pol3 L612M *rnh201Δ*	4764	Non-damage	1565	1	2
	4765	Low-UDG*	1563	5	67
	4765	High-UDG*	1509	5	12
Pol2 M644G *rnh201Δ*	4766	Non-damage	1566	1	2
	4768	Low-UDG*	1452	3	19
	4767	High-UDG*	1564	2	14

These results indicated that the genomic sequence around, and including the TetO44 array for the strains pairs used in the polymerase usage analysis was not the same, which meant that the comparison of TetO44 arrays was not accurate between sets or replicates.

3.4 System to timely induce damage and measure labour of DNA replicases

3.4.1 Formation of abasic sites using the construct Gal-TetR-UDG*

In Section 3.3 I showed that the chronic induction of abasic sites caused high levels of genomic instability in my cells, thus, I decided to try a different experimental approach to investigate the distribution of labour of the DNA polymerases in a damaged template. I used a method that allowed me to control the amount of damage induced in the region of interest, TetO44 array, by using the *GAL1/10* promoter to express the UDG* construct (Gal-TetR-UDG*). This construct permited the study of replication events occurring in one cell cycle as well as it allowed the timely induction of AP sites formation in G1-phase synchronised cells (Figure 3.12). However, for the purpose of this project, the risk of studying replication events of a single cell cycle was that the polymerase-mutant strains might not accumulate ribonucleotides at the site of damage sufficiently to be detected by NGS.

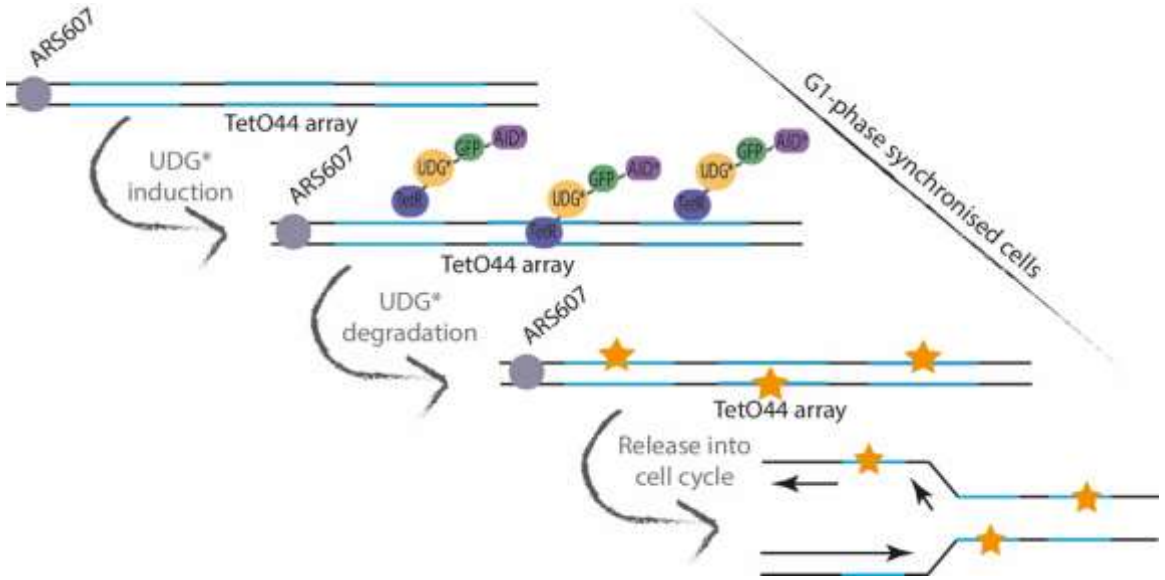

Figure 3.12. Scheme of the induction of abasic sites using the Gal-TetR-UDG* construct
The induction of UDG* in exponentially growing cells synchronised in G1-phase was achieved by changing the carbon source in the medium to galactose for 2 hours. The Gal-TetR-UDG* construct was subsequently targeted for degradation by shifting the cells into medium containing glucose supplemented with auxin for 1 hour. Then α-factor was washed away and cells were release into the cell cycle.

3.4.2 Strains construction and characterisation

I created yeast strains with the AP site system and the polymerase-mutant background, but instead of the G1-tag-TetR-UDG*, I used the Gal-TetR-UDG* construct to form abasic sites in the TetO44 array. Moreover, since I wanted to study how re-priming events occur downstream of the damaged region, I included strain pair with equivalent polymerase mutant for Pol α (Pol1 L868M). In addition, I constructed Non-damage control strains carrying a TetR construct without the UDG* enzyme, referred as "Gal-TetR-Empty" construct (Figure 3.1B).

To combine the background of the AP site system and the polymerase usage systems in yeast, I tried four different strategies. Two of them involved different mating steps and the other two included specific order of genetic transformations. One of the crossing strategies formed zygotes, which subsequently were subjected to sporulation conditions for tetrads formation. Even though different approaches to sporulate zygotes were used, none produced tetrads. Favourably, one of the strategies that involved consecutive transformation produced the require background for this study (the sequence of genetic transformations follows the same order shown in Table 3.6).

The strains used in this study were:

Table 3.6. Strains to assess the labour of DNA polymerases - Gal-TetR-UDG*

ID	Condition	Construct	Yeast background
5275	Non-damage	Gal-TetR-Empty	*BER*Δ, Gal-TetR-GFP-AID**, AFB2, *ura3*Δ, Pol1 L868M, *rnh201*Δ, TetO44-*URA3*
5297	Damage	Gal-TetR-UDG*	*BER*Δ, Gal-TetR-UDG*-GFP-AID**, AFB2, *ura3*Δ, Pol1 L868M, *rnh201*Δ, TetO44-*URA3*
5307	Non-damage	Gal-TetR-Empty	*BER*Δ, Gal-TetR-GFP-AID**, AFB2, *ura3*Δ, Pol2 M644G, *rnh201*Δ, TetO44-*URA3*
5252	Damage	Gal-TetR-UDG*	*BER*Δ, Gal-TetR- UDG*-GFP-AID**, AFB2, *ura3*Δ, Pol2 M644G, *rnh201*Δ, TetO44-*URA3*
5274	Non-damage	Gal-TetR-Empty	*BER*Δ, Gal-TetR-GFP-AID**, AFB2, *ura3*Δ, Pol3 L612M, *rnh201*Δ, TetO44-*URA3*
5253	Damage	Gal-TetR-UDG*	*BER*Δ, Gal-TetR-UDG*-GFP-AID**, AFB2, *ura3*Δ, Pol3 L612M, *rnh201*Δ, TetO44-*URA3*

To examine possible consequences on cell viability due to the genetic background, I compared several properties on the six yeast strains. First, the analysis for cell growth under non-TetR-induction conditions indicated that both damaged and non-damaged strains carrying the Pol ε mutant grew slightly slower in comparison to strains with mutated Pol α and Pol δ (Figure 3.13A-B). Additionally, I analysed viability by survival and a spot assay under two hours TetR-induction in G1-phase synchronised cells. As illustrated in Figure 3.13C and Figure 3.13D, no major differences among the TetR-UDG* stains survival were detected compared to the Non-damage strains.

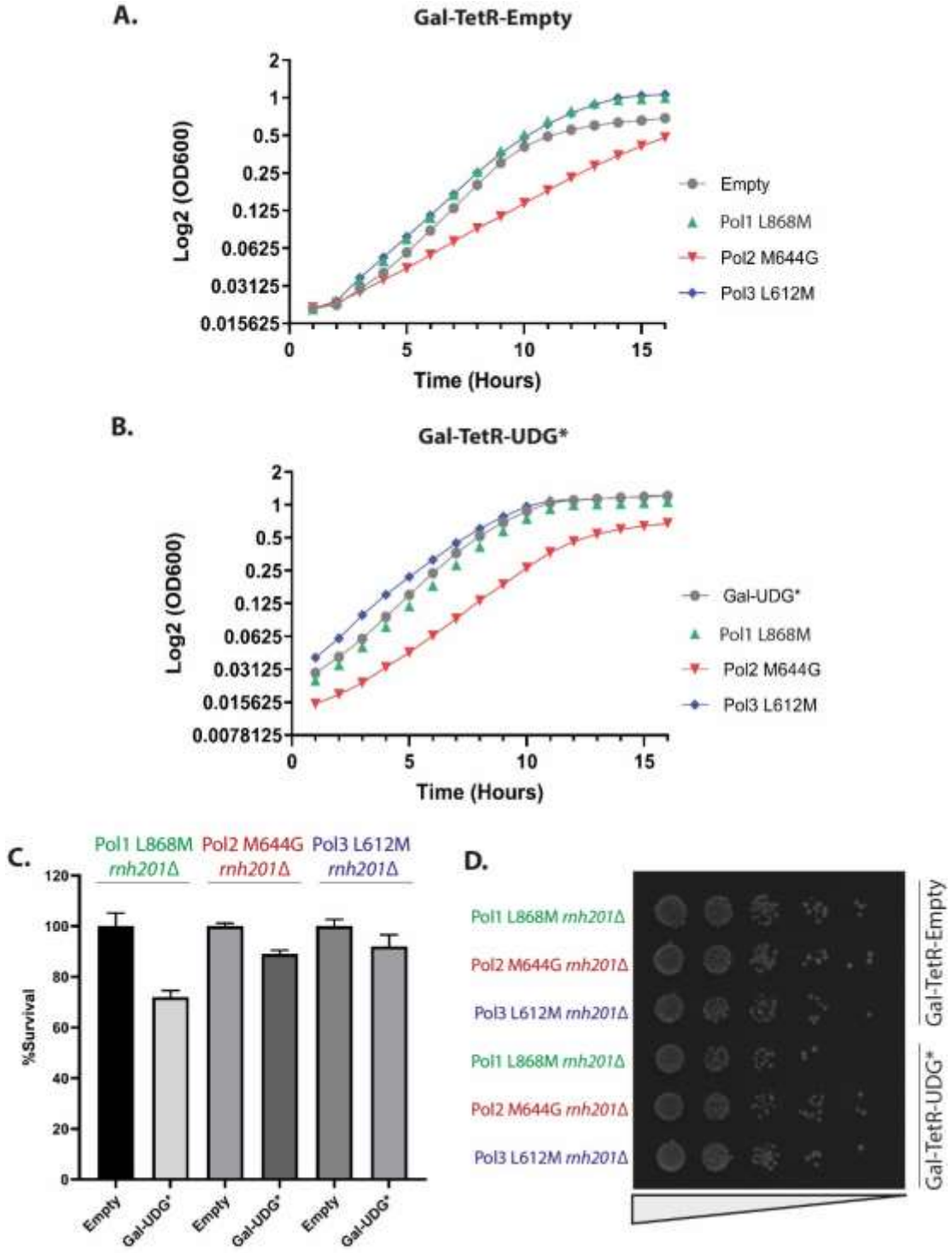

Figure 3.13 Characterisation of cell growth and colony formation of polymerase-mutant strains.
To compare cell growth of polymerase-mutant strains upon non-TetR-induction, OD600 readings were taken every one hour overnight in a Tecan SPARK 20M. Growth curves were plotted using the average of three replicates per run. **A.** Gal-TetR-Empty, and **B.** Gal-TetR-UDG*. **C.** Cell viability upon 2-hour TetR-induction. To calculate the viability percentage, strains with Gal-TetR-Empty construct were used as control (100%). **D.** Spot assays upon 2-hour TetR-induction were performed using 10-fold serial dilutions of 0.5 OD G1-phase synchronised cells (1: 1:12, 2: 1:125, 3: 1:250, 4: 1:2500, 5:1:5000). Representative figure of three independent repetitions. *Strain ID: 5275, 5307, 5274, 5297. 5252, 5253.*

Next, I analysed how cells progressed into the cell cycle following the experimental layout described in Figure 3.14A.

I observed a longer stalling in G1- and S-phase for cells with damage in comparison to the non-damaged cells, which could reflect the activation of DNA damage checkpoints (Branzei & Foiani, 2009; Kidane, et al., 2014). Furthermore, with both construct, it was evident that cells carrying a mutated Pol α progressed faster into the cell cycle compared to Pol2 M644G, and Pol3 L612M strains. In fact, strains carrying a mutated Pol ε spent the longest time in S-phase (Figure 3.14B). As mentioned in section 3.3.2.1, this could be a sign of DNA replication stress.

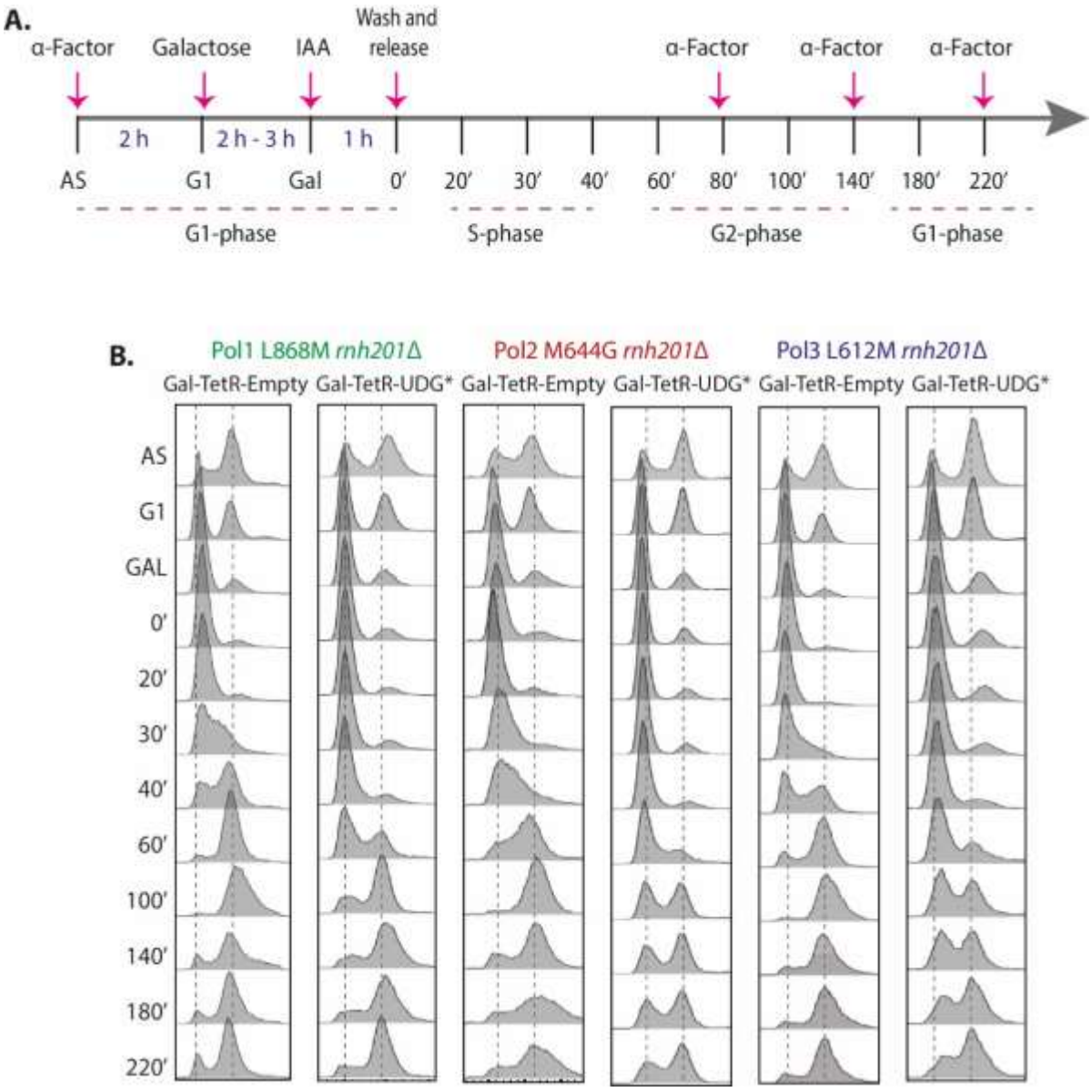

Figure 3.14. Cell cycle progression of polymerase-mutant strains expressing the Gal-TetR-Empty and Gal-TetR-UDG* constructs

A. Experimental layout. Cells growing in YPD media were first synchronised in G1-phase with α-factor (2 hours), then cells were shifted to YP-Gal to express the TetR construct (2 hours). The construct was degraded by shifting the cells back into YPD media supplemented with auxin for 1 hour. α-factor was washed away and the cells were released into the cell cycle. To avoid cells passing to the next S-phase, pulses of α-factor were added in late G2-phase. Different time-points were collected for further analysis. B. Analysis of cell cycle progression. Representative figure of three independent repetitions. *Strain ID: 5275, 5307, 5274, 5297. 5252, 5253.*

Overall, the characterisation of the strains using the galactose promoter to induce the formation of abasic sites in the TetO44 array showed similarities with the strains created in Section 3.3. Thus, before continuing with the analysis of the DNA polymerase usage, I first standardised the experimental conditions by testing different times of UDG* expression, to try to avoid creating high genomic instability that could lead to the shortening of the TetO44 array.

3.4.3 Standardisation of experimental conditions

By using alkaline Southern blots, Elizabeth Colby previous found that expressing UDG* between 1.5-2 hours was sufficient to detect abasic sites. An induction time of 4 hours, however, was too long for *BER*Δ mutant strains, as it produced high amounts of SSBs, probably *in vivo*, as well as a slow progression through S-phase and high Rad53 phosphorylation (pRad53). An induction time of 1 hour, on the contrary, showed fewer abasic sites and mostly an intact TetO44 array. Based on these results, to determine the experimental conditions in my background I tested 2 and 3 hours of induction of the Gal-TetR-UDG* construct in the three polymerase-mutant strains, using the experimental layout described in Figure 3.14A.

I, first, observed how cells progressed in the cell cycle by taking different time-points after release from G1-phase. As expected, cells where UDG* was expressed for 3 hours accumulated longer in G1-, S-, and G2-phase than the ones where induction was for 2 hours. These observations suggested that there were major difficulties in processing the damage due to higher amounts of abasic sites. Moreover, cells with both 2 and 3 hours of UDG* expression passed to the next G1-phase (time-point 220'), possibly reflecting that cells resolved damage (Figure 3.15A).

In parallel, I looked at the activation of the DDR by checking the level of phosphorylation of the essential intermediate kinase in the checkpoint pathway, Rad53. In addition to the small subunit of the ribonucleotide reductase complex, Rnr4, which catalyses the rate limiting step in the biosynthesis of all four deoxyribonucleoside triphosphates and is highly induced in response to DNA damage (reviewed in Tsaponina & Chabes, 2013). The induction of UDG* for 3 hours activated the damage checkpoint just after abasic sites were formed (time-point Gal) throughout G2-phase (time-point 100') (Figure 3.15C). When UDG* was induced for 2 hours, the damage checkpoint activation mostly was seen in G2-phase (time-point 100') (Figure 3.15B).

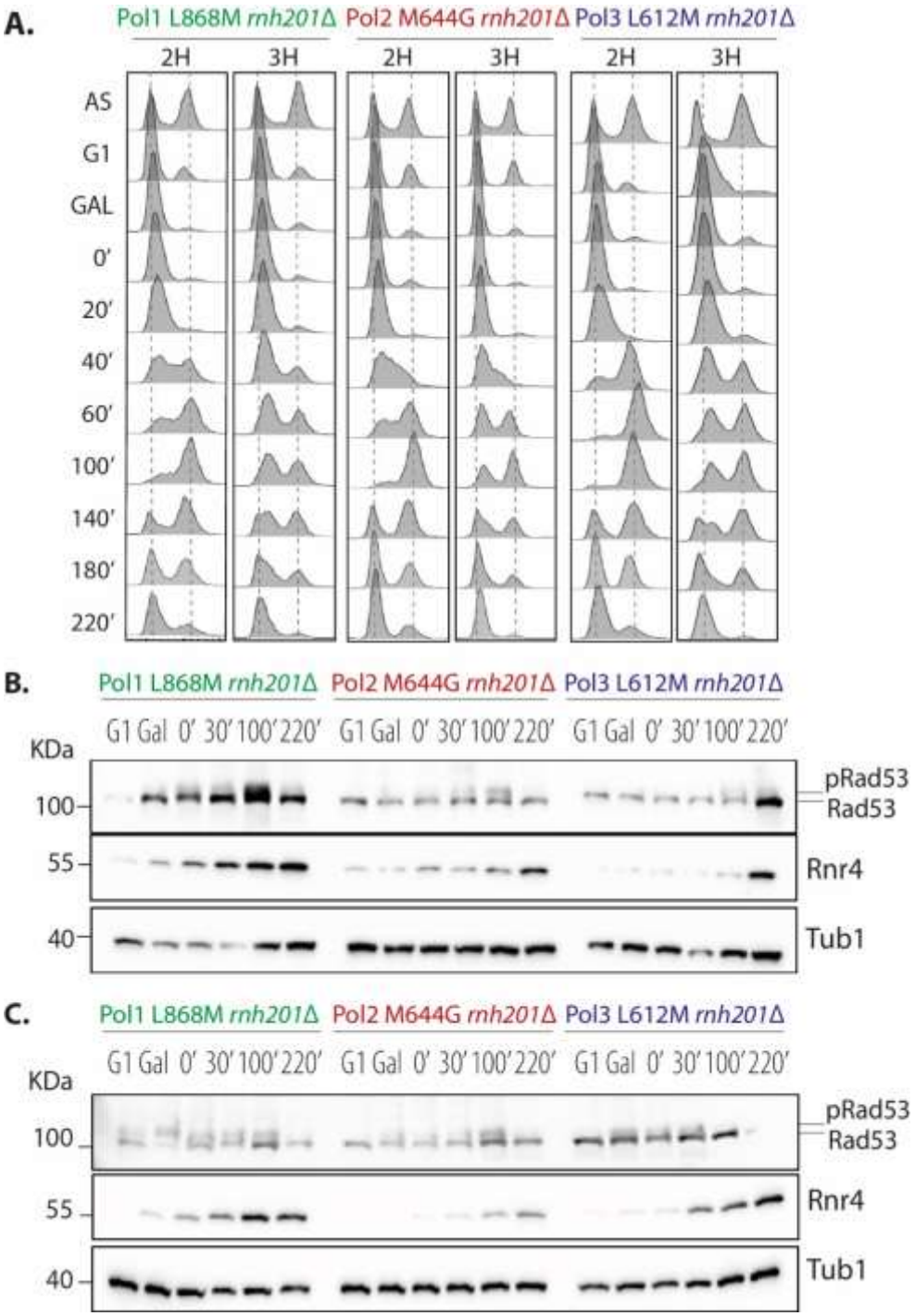

Figure 3.15. Cell cycle progression and checkpoint activation of 2 and 3 hours UDG* induced cells.
Cells treated as describe in Figure 3.14A. UDG* was induced for 2 and 3 hours. Different time-points were collected for further analysis. **A.** Analysis of cell cycle progression. **B-C.** Western blotting indicating DNA damage checkpoint activation. *Strain ID: 5297, 5252, 5253.*

As the induction of the Gal-TetR-UDG* construct for 2 hours showed fewer effects on cell cycle progression, along with a reduced activation of the checkpoint response compared to a 3 hour induction, I decided to further validate this condition and check the intactness of the TetO44 array by Southern blotting.

Importantly, a conventional PCR (Figure 3.16) and subsequent homology analysis of Sanger sequenced fragments (not shown) indicated a stable length of the TetO44 array throughout the time-course in all the polymerase-mutant strains with both induction times.

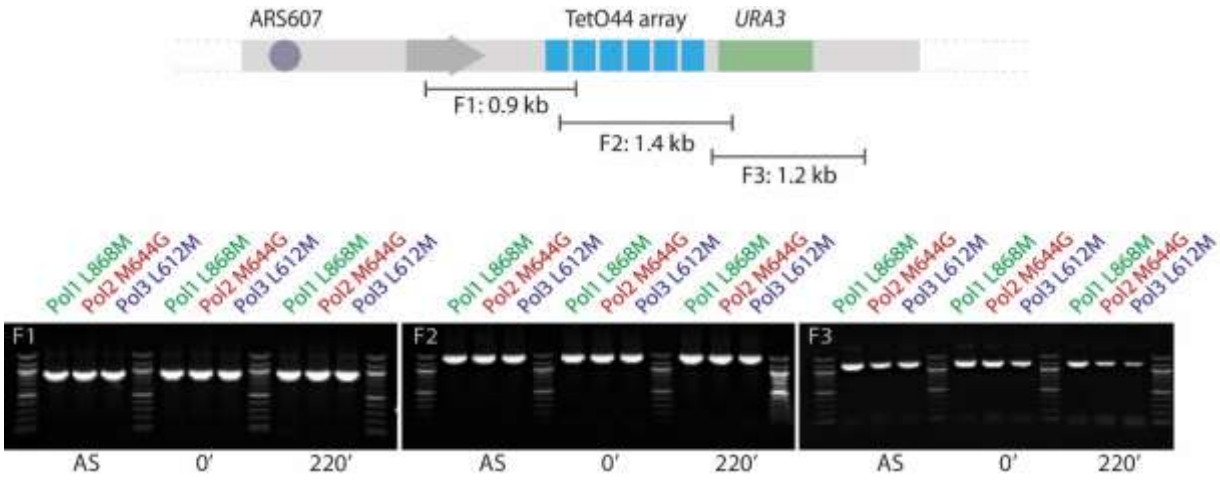

Figure 3.16. PCR of TetO44 array
Different time-points were taken before and after UDG* induction. These samples were the input for the PCR amplification of three fragments within the TetO44 array. These were visualised on a 1% agarose gel. Figure showing amplification results of cells with UDG* induction for 2 hours. *Strain ID: 5297, 5252, 5253.*

3.4.4 Abasic sites on TetO44 array are processed into double-stand breaks

To evaluate the intactness of the TetO44 array after UDG* induction for 2 hours using the *GAL1/10* promoter, I performed a native Southern blot to examine the existence of DSBs. The premise was that an abasic site on the template could be processed into a gap or SSB by an excision-repair based step and the resulting nick converted into a DSB.

As previously shown by Elizabeth Colby, abasic sites were found up to 10 kb away from the TetO44 array, thus, in order to have a complete perspective of the effect of AP sites on the stability of the array, I analysed a fragment of ~11 kb (Figure 3.17). I collected samples as described on Figure 3.14A before DNA was extracted and restricted digested with BssSI-v2. The resulting fragments were separated by a native gel electrophoresis followed by Southern blot analysis. I included an undamaged TetO44 array as a non-breakage control (-) as well as a TetO44 array with an artificial double-stranded break control (+) created by additionally digesting with PvuI.

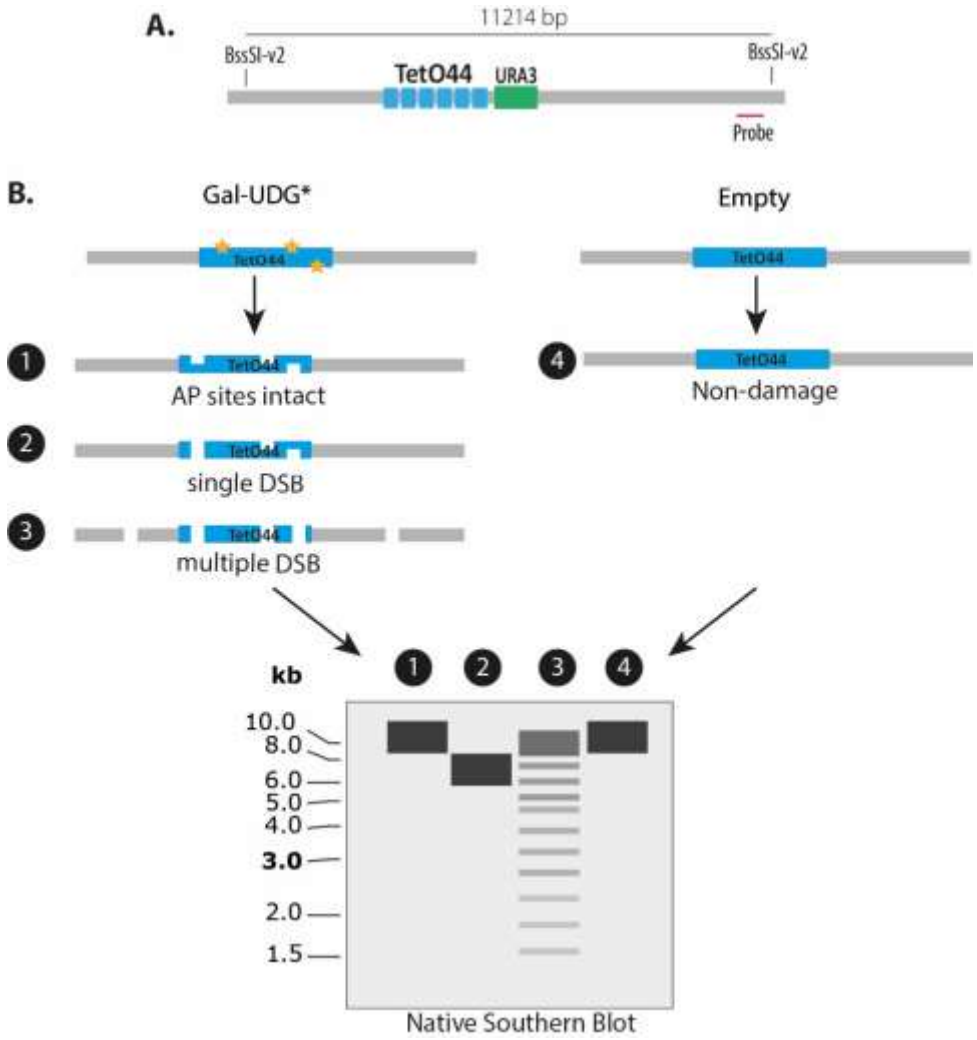

Figure 3.17. Scheme of visualisation of DSBs in native Southern blot

A. Input DNA was digested with BssSI-v2, which allowed the study of a region of 11.2 kb within the TetO44 array, the radioactive-labelled ssDNA probe used for the native Southern blot hybridised downstream of the array. B. Scheme of the visualisation of dsDNA breaks in a native gel upon Gal-TetR-UDG* induction. Non-damaged strains (Gal-TetR-Empty) were expected to keep an intact band.

Although the induction of the Gal-TetR-Empty construct showed different protein levels amongst polymerase-mutant strains, these differences did not affect cells progressing into the cell cycle (Figure 3.18A-B). Moreover, as expected, none of the Non-damage strains presented formation of DSBs, as the TetO44 array band in the native Southern blot remained intact. I observed, however, a second band that ran at a lower molecular size in the Polε-mutant strain. As this band was seen throughout the time-course, I discarded that this was an effect of the induction of the TetR construct. Instead, one possible reason could be that the Polε-mutant strain had a mix cell population containing TetO44 arrays with different lengths, or this could also be a direct effect of the genomic instability generated by the misincorporated ribonucleotides in the genome (Figure 3.18C, Table 3.7).

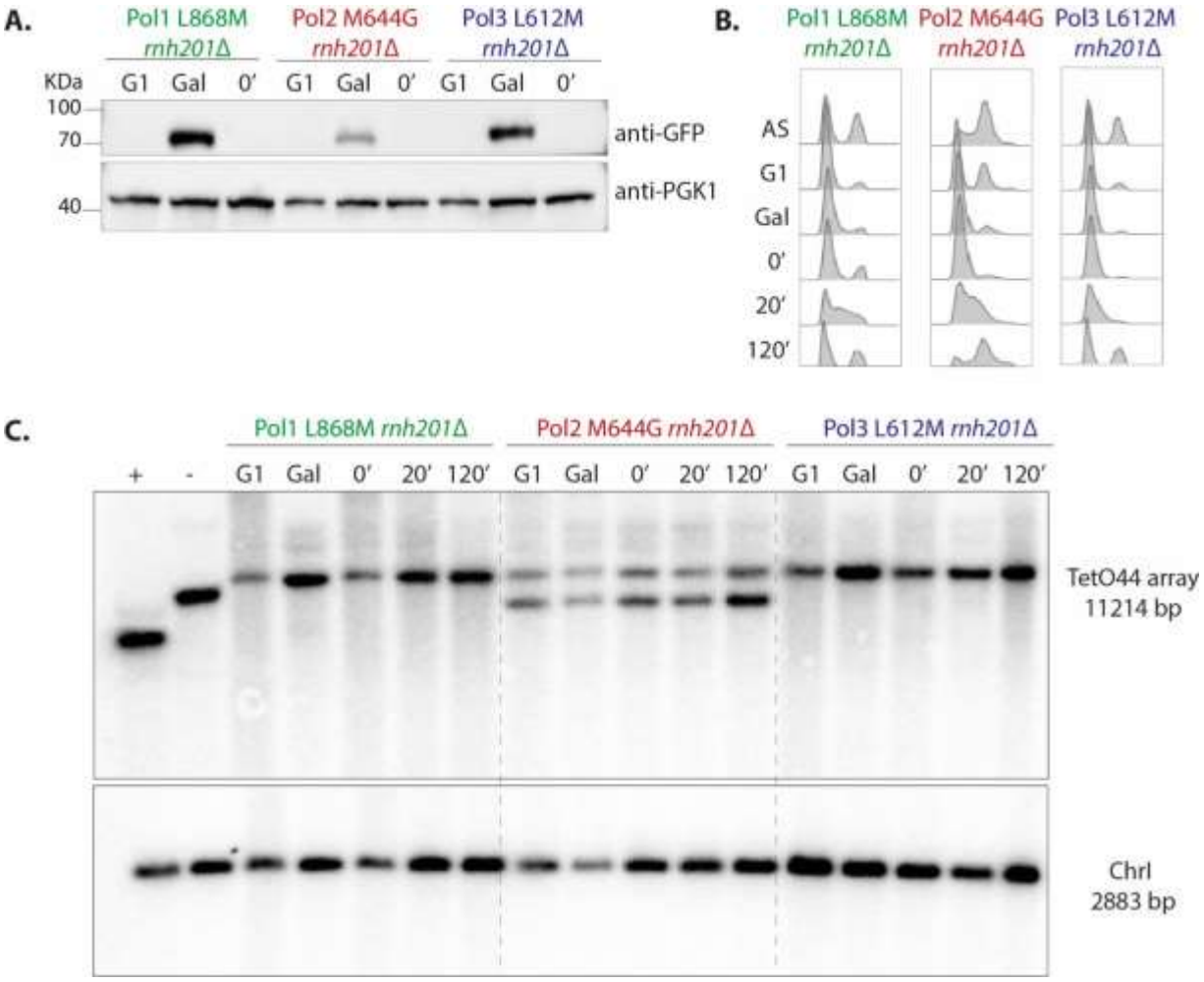

Figure 3.18. Native Southern blot of polymerase-mutant strains under non-damaged conditions
Samples treated as in Figure 3.14A were taken for **A.** Western blotting indicating Gal-TetR-Empty construct levels. **B.** Analysis of cell cycle progression. **C.** Native Southern blot. Membrane was first probed for the TetO44 array, and then probed for the control region on Chr1. *Strains ID: 5275, 5307, 5274.*

Similar to Section 3.4.3, when the UDG* construct was induced, I observed that the Pol ε-mutant strain accumulated longer in S-phase compared to mutated Pol α and Pol δ strains (Figure 3.19A-B). Remarkably, the Southern blot showed that the TetO44 array band started being degraded just after the Gal-TetR-UDG* construct was expressed (time-point Gal) in the three polymerase-mutant strains (Figure 3.19C). The quantification of the band intactness indicated that this decrease happened over the time-course (Table 3.7). Almost all the full-length TetO44 band appeared to be degraded by the end of the time-course, this possibly indicated the existence of multiple double-strand breaks within the array. Alternatively, taking into consideration that this assay was based on the average of all cells in the population, it could be speculated that there was at least one DSB per cell occurring randomly in the region within the TetO44 array.

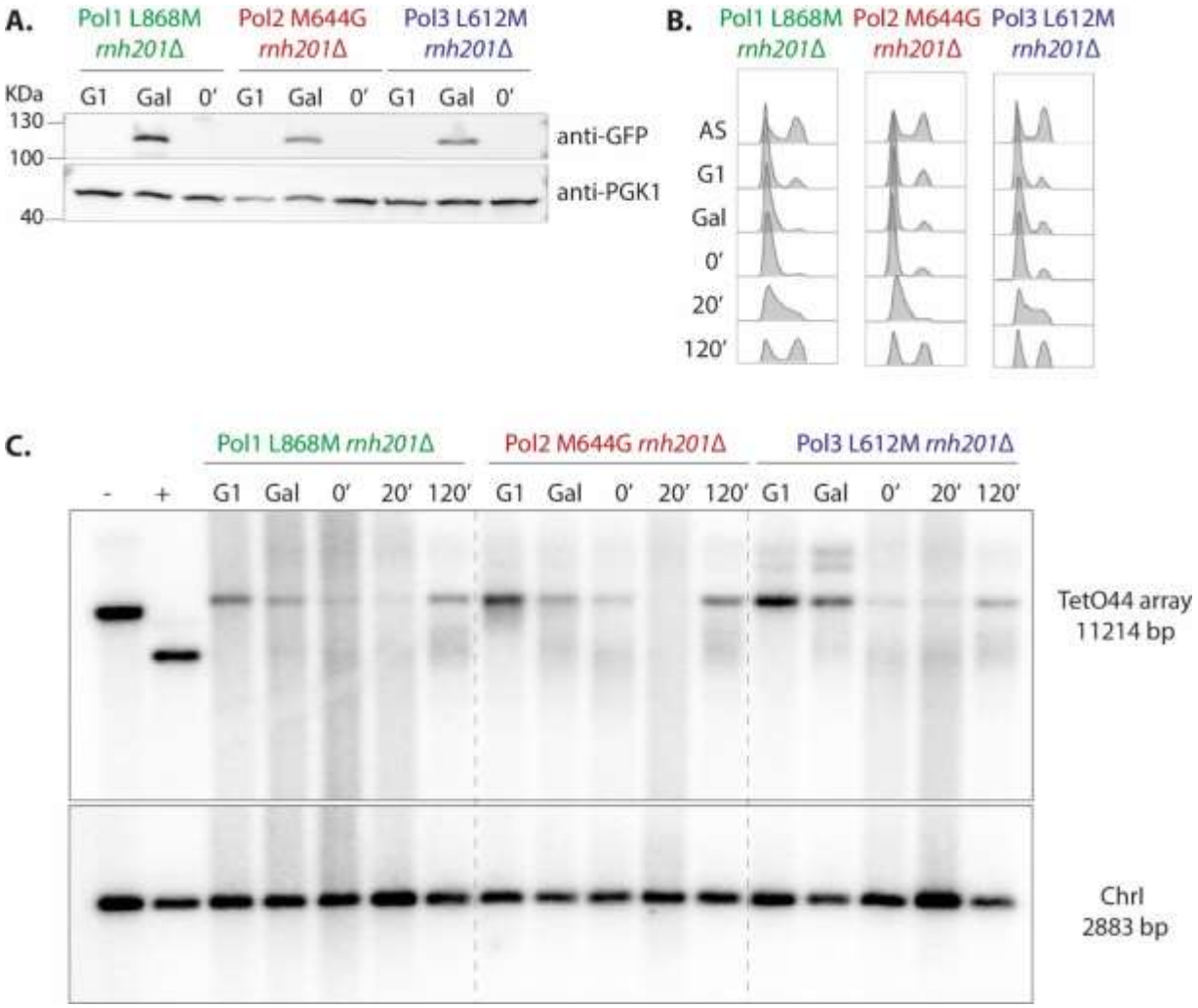

Figure 3.19. Native Southern blot of polymerase-mutant strains under damaged conditions
Samples treated as in Figure 3.14A were taken for **A.** Western blotting indicating Gal-TetR-UDG* construct levels. **B.** Analysis of cell cycle progression. **C.** Native Southern blot. Membrane was first probed for the TetO44 array, and then probed for the control region on Chr1. *Strain ID: 5297, 5252, 5253.*

Altogether, these results indicated that abasic sites were processed into DSBs. In fact, it can be suggested that the DSBs could have formed independently of the response of the replication machinery to abasic sites in S-phase, as the presence of DSBs were seen from early G1-phase upon UDG* induction (time-point Gal).

Table 3.7. Percentage of intactness of the TetO44 array after expression of different TetR constructs

		% Intactness TetO44	
		Gal-TetR-UDG*	Gal-TetR-Empty
	-	100	100
Pol1 L868M *rnh201Δ*	G1	100	100
	Gal	86	99
	0	74	112
	20	29	100
	120	36	106
Pol2 M644G *rnh201Δ*	G1	100	100
	Gal	65	97
	0	33	69
	20	20	65
	120	32	72
Pol3 L612M *rnh201Δ*	G1	100	100
	Gal	95	118
	0	21	119
	20	15	115
	120	28	111

Because the induction of UDG* in the polymerase-mutant strains led to the formation of strand breaks within the TetO44 array, I decided not to proceed with the analysis of the distribution of labour of the replicative DNA polymerases.

3.5 Discussion

During this chapter, I intended to create a system capable of studying the distribution of labour of the main DNA polymerases on a damaged template in the model organism *S. cerevisiae*. I used the AP site system, which regulates the formation of abasic sites in a site-specific manner within a cell by taking advantage of the TetO/TetR system. In addition, I exploited ribonucleotides-permissive forms of Polα, Polε or Polδ in RNase H2-defective strains and used deep sequencing to map ribonucleotides with single-nucleotide resolution. By using this complex system, polymerases usage sequencing screen showed evidence of the role of DNA polymerase δ in the synthesis of damaged leading strands. However, I also found that the AP site system generated high genomic instability leading to the formation of double-strand breaks within the TetO44 array. This made it difficult to conclude on the actual distribution of labour of the DNA polymerases on damaged DNA, and raised the need for a system able to induce abasic sites in a less acute manner to avoid further processing into DSBs.

3.5.1 Role of DNA polymerases in a damage template

Damaged nucleobases, abasic sites, and bulky adducts such as pyrimidine dimers do not pose a block for the CMG unwinding action. However, the replicative polymerases are typically intolerant to these DNA lesions (Guilliam, 2021). While the stalling of Pol δ, allows for the continued synthesis of a lagging strand by downstream reinitiation events, the stalling of Pol ε causes uncoupling of leading strand synthesis (Taylor & Yeeles, 2019). In order to complete replication, TLS polymerases can be recruited to the stalled nascent strands and synthesise DNA across the lesion. However, these polymerases are widely inaccurate and poorly adapted to bulk synthesis. Another option for the completion of DNA replication is the action of a free replicative polymerase able to re-engage with the nascent leading strands and re-couple synthesis with the CMG (reviewed in Guilliam & Yeeles, 2020). By using different versions of the NGS tool to map incorporated ribonucleotides into genomic DNA in undamaged conditions, it was shown that Pol δ can synthesise up to 180 nt on the leading strand (Garbacz, et al., 2018 & Zhou, et al., 2019), making it the best candidate to re-establish replication for a damaged leading strand. Taylor & Yeeles (2018) were the first to indicate that Pol δ indeed is critical in re-coupling the synthesis of leading strand following a CPD lesion. Congruently to these findings, I also observed evidence in favour of the usage of Pol *δ* in leading strand synthesis upon DNA damage (Section 3.3.3.3).

In section 3.3.4, I reported that the chronic induction of abasic sites shortened the TetO44 array, possibly due to a mutagenic-recombination repair pathway. A possible explanation could be that the accumulation of abasic sites provoked the collapsed of the replication fork, forming DSBs. Interestingly, previous studies revealed that the repair of these breaks and the re-start of replication can be achieved by the error-prone break-induced replication (BIR). During BIR, Pol δ was also found to synthesise both the leading and lagging strand, with no requirement for Pol ε (Donnianni, et al., 2019). This data would be consistent with

my own observations, in which BIR aided the re-initiation of the stalled fork. However, with the limited data in this study, it is challenging to draw a definitive conclusion about the role of Pol δ and BIR. Thus, it would be worthwhile to investigate these assumptions by additional approaches.

3.5.2 Top1 processing ribonucleotides in DNA

Yeasts devoid of RNase H2 are viable and contain abundant rNMPs incorporated in their DNA (Nick McElhinny, et al., 2010A). In the absence of RER, alternative pathways that might induce mild replication stress and checkpoint activation usually remove misincorporated ribonucleotides. However, these are also associated with higher mutation rates due to the formation of short deletions and introduction of repetitive sequences (Kim, et al., 2011). These phenotypes are linked to the cleavage of ribonucleotides by Top1, which generates a blocked 3'- end that could be lethal or mutagenic, if not processed (Cerritelli & Crouch, 2016). Recently, Li and colleagues (2019) found that Apn2 is specialised in removing multiple types of adducts from 3'- ends, principally those linked to Top1 cleavage at rNMPs. As the AP site system was established in cells that lack the major AP endonucleases and lyases (*apn1Δ apn2Δ ntg1Δ ntg2Δ)* it could be suggested that the study of the distribution of labour of the DNA polymerases caused unexpected additional genomic instability, if Top 1 participated in the processing of incorporated ribonucleotides.

3.5.3 Abasic sites processed into double-strand breaks

Abasic sites are cytotoxic in many ways; first, they are a source of replication stress, as the progression of DNA polymerases is blocked in their presence. Second, abasic sites can yield the formation of single-base-pair substitution in yeast cells. Thrid, due to their reactive chemistry, they are intermediates in the formation of lesions that are more challenging to repair, such as DSBs (Boiteux & Guillet, 2004).

In this chapter, I showed that the formation of abasic sites either with a chronic or acute induction of the enzyme UDG* in the genome of budding yeast led to the formation of DSBs within the TetO44 array (Section 3.3.4 & 3.4.4). Likewise, the high accumulation of AP sites resulting in the formation of DSBs has been seen in *Helicobacter pylori in vivo*. Kidane and colleagues (2014) proposed that these DSBs arise from the exhaustion of BER components and the removal of clustered damaged bases by Ogg1 in G1-phase, in which the error-prone pathway of non-homologous end-joining (NHEJ) most likely repaired the resulting DSBs. During Ogg1-BER, Ogg1 excises damaged bases followed by β-elimination on the resulting abasic site yielding a 3′-blocked SSBs (Ba & Boldogh, 2018). These lesions are considered more toxic than AP sites themselves. However, such blocked ends are usually processed by the activity of Apn1 and Apn2 in yeast (Boiteux & Guillet, 2004). As the AP site system was built in the *BERΔ* quadruple mutant background, it is possible that Ogg1 was in charge of processing abasic sites and, in the absence of Apn1 and Apn2, an unprocessed blocked end could have led to DSBs formation.

Furthermore, as NER proteins are present in our background, it can be implied that this repair pathway could also have participated in the processing of abasic sites by the action of Rad1-Rad10 (Boiteux & Guillet, 2004). The premise behind this possibility is that NER, in the process of removing damage, nicked the DNA forming a gap, which was then converted into a DSB. Thus, it would be interesting to elucidate the participation of NER in the formation of DSBs.

Kidane and colleagues (2014) also suggested that AP sites lead to the accumulation of DSBs in S- and G2-phases by "colliding" with replication forks. Equally, a more direct connection between fork arrest and error-prone HR has been established in fission yeast, where a single collapsed fork caused mutations and large-scale genomic changes, including deletions and translocations (Iraqui, et al., 2012). When UDG* was induced, I observed the presence of DSBs in early G1-phase and a pronounced effect during S-phase (Section 3.4.4), thus I cannot rule out that some of the DSBs originated when a replication fork encountered damage.

3.5.4 Abasic site formation is biased towards the lagging strand

Interestingly, Suzuki and colleagues (2019), by studying the effects of site-specifically modified oligodeoxynucleotides containing a single natural abasic site or a chemically synthesized tetrahydrofuran (THF) in human cells, found an unequal mutagenicity when the lesion was located in leading or lagging strands. They observed that not only base substitution but also deletions were more predominantly located on lagging strands. Therefore, it would be interesting to explore whether UDG*-induced damage favours mutagenesis mostly on leading or lagging strand.

3.5.5 Replication fork breakage and re-start

Even though I found that abasic sites were processed into DSBs upon the induction of UDG*, most cells progressed through the cell cycle and reached the next G1-phase (Figure 3.14). In fact, survival assays did not show pronounced effect on cells viability (Figure 3.13). Altogether, these observations indicated that damage must have been mostly resolved. I suggested that DSBs were processed with mutagenic-HR mechanisms as the integrity of the TetO44 array was altered. Fork breakage is not uncommon; any repair process that involves DNA strand cutting can generate a DSB, if the incision happens in ssDNA (Cortez, 2019). HR has been suggested to be vital for the repair and re-start of replication forks following replisome disassembly (McGlynn & Lloyd, 2002). Moreover, despite the fact that AP sites are primarily repaired by the BER pathway, with NER acting as a back-up (Torres-Ramos, et. al., 2000), several studies have shown that in the absence of excision pathways, the repair of abasic sites is accomplished by a recombination repair mechanism. For instance, in budding yeast, the *apn1Δ ntg1Δ ntg2Δ* triple mutant (*BERΔ*) and *ntg1Δ ntg2Δ apn1Δ rad1Δ* quadruple mutant (*NERΔ*) exhibited an 18- and a 170-fold increase in genomic recombination rate, respectively, indicating that recombination is indeed involved in the processing of lesions normally repaired by the BER (Swanson, et. al., 1999).

Little is known about the identity of the recombination repair factors involved in the re-start of replication forks. Here I described possible pathways that could have resolve DSBs in my genomic background:

Cohesin-dependent sister-chromatid exchange (SCE): Spontaneous DSBs arising from replication failures were thought to be repaired by cohesion-dependent sister-chromatid exchange (SCE). Cohesin consists of two SMC (structural maintenance of chromosomes, Smc1 and Smc3) and two non-SMC components (Scc1 and Scc3). These are deposited around chromatin in G1-phase to hold sister chromatid together through replication until anaphase and were found to additionally be loaded in the proximity of the DSB and favour HR post-replication (Cortés-Ledesma & Aguilera, 2006). This mechanism, however, raises the question of whether there is a specific machinery that would favour SCE over other forms of post-replicative repair.

Break-induced replication: BIR can repair broken forks to ensure completion of DNA replication, and it can efficiently repair DSBs at which the homology with the donor molecule is limited. BIR has been seen to act during S- and G2-phases to repair forks encountering a single strand nick. Nevertheless, BIR-induced DNA synthesis is inaccurate with frameshift mutation rate of 2,800-fold compared to spontaneous events (Deem, et al., 2011; Costes & Lambert, 2013). According to my observations, it is possible that BIR had a role in resolving DSBs created by abasic sites but further confirmation is needed.

Previously suggested by Kidane, *et. al.*, (2014), the error-prone pathway of NHEJ may be involved in the repair of DBSs in G1-phase.

"Toxic" NHEJ: There is evidence that the end-joining of DNA ends at collapsed replication forks is mediated by NHEJ, which results in the induction of copious numbers of toxic genomic alterations (Balmus, et al., 2019; Neal, et. al., 2020).

3.5.6 Unrepaired DSBs and cells adaptation to damage

Induction of Gal-TetR-UDG* either for 2 or 3 hours triggered the activation of the checkpoint DNA-damage response as indicated by high levels of pRad53 in late G2-phase. Remarkably, I also observed that Rad53 was de-phosphorylated in the next G1-phase (Figure 3.15B-C). If it is truth that the formation of abasic sites led to the random processing of a single DSB within the TetO44 array (Figure 3.19), it would not be far-fetched to speculate that cells adapted to damage. Previous studies in budding yeast have shown that a single unrepaired DSB does not permanently arrest cells in G2/M phase, instead cells can escape from G2/M arrest despite the continued presence of the broken chromosome (Pellicioli, et al., 2001). Adaptation and recovery are brought about by inactivating upstream elements of the DNA checkpoint rather than downstream effectors that block cell cycle progression. Thus, adaptation is always accompanied by the disappearance of pRad53, which is consistent with my findings. Interestingly, RPA and pRad52 have been seen to serve as monitors of the presence of unrepaired ssDNA and thus control a cell's decision to adapt (Eun Lee, et al., 2003). Therefore, it would be interesting to investigate this assumption further.

3.5.7 Cytotoxic effects of UDG* induction

The capacity of UDG* to form abasic sites *in vivo* was studied by Otterlei, et al., (2000). They found that after a 2.5 hours induction of the enzyme, around ~6 AP sites were formed in 100 kb. Taking this data, it can be assumed that in the AP site system, upon UDG* expression at least 1 abasic site was formed in a 10 kb ratio within the TetO44. However, the high affinity of TetR for the TetO sequence (Das, et al., 2016) could have resulted in highly specific targeting of UDG*, possibly meaning that a greater number abasic sites clustered only in the TetO44 array. Previous studies in bacterial cells have shown that clustered DNA damage formed by ionising radiation were converted to lethal DSBs during DNA repair (Blaisdell, Harrison, & Wallace, 2001). This finding could confirm that the accumulation of AP sites within the TetO44 array indeed led to genomic instability.

In other study, Auerbach, et al. (2005) examined the mutational effects of UDG* in *S. cerevisiae*. They found that UDG* expression enhanced the mutation frequencies to 150-fold, most of which were single-base-pair substitutions, in addition to a specific 290-bp deletion, and a minority of small deletions and insertions. Correspondingly, Sanger sequencing of the TetO44 array sequences upon chronic UDG* expression (Table 3.5) revealed the presence of mostly single-base-pair substitutions, followed by small deletions and insertions. All these results indicated UDG* high levels of genome instability. Therefore, if the AP system is to be ultimately established, great caution should be taken with respect to the cytotoxic effect of UDG*.

3.5.8 Abasic sites tolerance by translesion synthesis

AP sites are unambiguously mutagenic in yeast; they are capable of yielding a variety of base-pair substitutions at the site of the lesion, most likely reflecting the action of TLS polymerases. Investigations behind the mechanisms of TLS have indicated that AP sites can efficiently be bypassed when Pol δ, Pol η or Rev1 act as inserters with Pol ζ as extender (reviewed in Boiteux & Guillet, 2004). Guilliam & Yeeles (2020) extended these findings by showing that despite the great ability to tolerate certain lesions, Pol δ also stalls upon encountering bulky or distorting lesions. In these instances, the specialised TLS polymerases are required for direct synthesis across the damage prior to recoupling, in a process termed “on the fly” TLS. They also found that leading and lagging-strand TLS are mechanistically similar, where Pol η plays a pivotal role. Even though these findings provide great insight into the role of TLS in bypassing polymerase-stalling lesions, it would be exciting to confirm if this also occurs *in vivo.* In fact, as the AP site system can provide information about replication in both time and space, it would be able to discriminate between “on the fly” from gap-filling TLS. However, it is pivotal first to find the experimental conditions under which AP sites are not processed into DSBs.

Chapter 4: Results II - Synthesis of DNA daughter strands in the presence of AP sites

4.1 Background

The response of the replisome to nucleotide-base lesions depends on whether damage is located on the lagging or leading strand. On the one hand, the synthesis of the lagging strand has been seen to resume downstream of the lesion. Conversely, damage in the leading strand represents a major block for replication re-start (reviewed in Yeeles, et al., 2013). The study of UV-irradiated *S. cerevisiae* cells showed that ssDNA gaps accumulated along replicative duplexes, which indicated that replication is indeed re-started even on leading strands (Lopes, et al., 2006). Nevertheless, how the leading and lagging strands achieve this is not completely clear. The tool eSPAN was created to study DNA replication-linked processes and has helped to elucidate the strand-specific association of replication proteins on stalled forks. For instance, Yu and colleagues (2014) found that at HU-stalled forks, the PCNA clamp partly disassociated from the lagging strand. Furthermore, other single-molecule analyses of DNA replication have suggested that several proteins that are not part of the core replication machinery promote efficient damage bypass and replication re-start (Petermann & Helleday, 2010). Even though these findings have provided significant insights about how the replisome responds to damage and how DNA replication is resumed after lesions, further work is needed to clarify the mechanistic distinctions between leading and lagging strand synthesis in the presence of damage.

During this chapter, I evaluated how leading and lagging strands are synthesised in the presence of polymerase-stalling lesions. To achieve this, I took advantage of the AP site system to create damage site-specifically, and I combined it with the NGS method, BrdU-ssSeq, to study the synthesis of daughter strands. Moreover, I investigated the association of core replisome proteins, PCNA and RPA, at abasic-site-stalled replication forks by ChIP-ssSeq. Here, I showed that both the leading and lagging strand remained mostly un-replicated in the presence of abasic sites, and similarly to previous findings, I observed that replication re-started downstream of the lesion. However, I also showed that abasic sites formation led to greater genomic instability linked to DSBs formation and DNA resection.

4.2 Experimental approach

4.2.1 System to create site-specific abasic sites in budding yeast

To create abasic sites in the chromosome VI of budding yeast, I used the AP site system detailed in Section 3.2.1. The specific experimental conditions that I used during this chapter are described hereafter.

Controlling damage targeting: The *GAL1/10* promoter was used to induce expression of the TetR-UDG* construct (Figure 3.1B) and to form AP sites in the TetO44 array.

Yeast background: The W303 strain corrected for the *rad5-G535R* mutation background was used in this study. I introduced the features needed for the AP site system: the quadruple *apn1Δ apn2Δ ntg1Δ ntg2Δ* mutant, the Gal-TetR-UDG*-GFP-AID* construct, the auxin receptor F-box protein, AFB2, and the TetO44 array. Additionally, as *S. cerevisiae* lacks the thymidine salvage pathway for efficient cellular uptake and incorporation of thymidine analogues into DNA, I integrated the Herpes simplex virus thymidine kinase (*HSV–TK*) and human equilibrative nucleoside transporter (*hENT1*) sequences to enable the incorporation of BrdU (Viggiani & Aparicio, 2006). For immunoprecipitation studies, I tagged the protein of interest at its C-terminus using the 9-myc tag. Finally, unlike Chapter 3, the endogenous *URA3* gene located on Chromosome V was not deleted because a free auxotrophic or antibiotic selection marker was not available to accomplish this.

4.2.2 Standardising conditions for the study of daughter strands synthesis

4.2.2.1 Strategy to examine replication events downstream of the TetO44 array

The genetic background of the AP site system has two *URA3* gene sequences, one endogenous non-functional located in chromosome V and the other functional one, usually used as a selection marker for introducing the TetO44 array, in chromosome VI. If NGS analysis is performed in this background, the discrimination of reads coming from one chromosome or the other is impossible. In fact, it is pivotal to unequivocally identify the origin of the sequencing data for the functional *URA3* because it is located downstream of the TetO44 array, and therefore it represents valuable information on replication re-priming and re-start events after damage bypass. This problem can be solved by deleting the endogenous *URA3* sequence similarly to what was done in the polymerase-mutant strains in Chapter 3. However, in the current background (Section 4.2.1), a free auxotrophic or antibiotic selection marker was not available to perform this task.

To delete the endogenous *URA3* gene, I tried an alternative strategy based on the dominant marker *amdSYM*. *amdSYM* should be a recyclable marker that confers *S. cerevisiae* the ability to use acetamide as sole nitrogen source, and it does not require a specific genetic background for the strain to be selected (Solis-Escalante, et al., 2013). Nonetheless, the strain background that I employed was not capable of surviving solely on acetamide even when *amdSYM* was introduced (not shown).

Therefore, I opted to remove the functional *URA3* sequence located downstream of the TetO44 array. I considered using a strategy that allowed me to modify already constructed strains with no selection markers available and that did not require an auxotrophic marker. To achieve this, I tried using short synthetic sequences (G-block), and different cloning strategies to remove the *URA3* sequence.

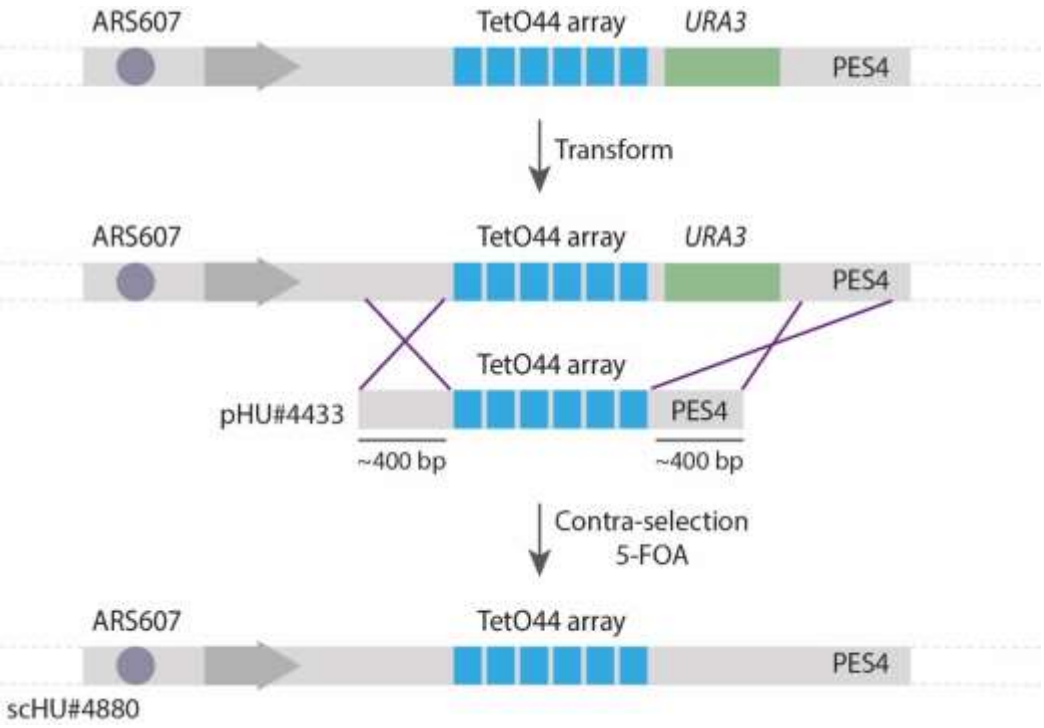

Figure 4.1. Scheme to remove the *URA3* sequence downstream of TetO44 array
Yeast strains carrying the TetO44 array on ChrVI were transformed with plasmid encompassing the exact TetO44 sequence and ~400 bp homology at each adjacent region on the PES4 locus (pHU#4433). Positive clones were selected by growth on 5-FOA and confirmation of correct TetO44 array length was done by conventional PCR.

I succeeded to remove the *URA3* sequence by cloning a plasmid with homology to the TetO44 array and PES4 locus sequences, but without the *URA3* sequence. By transforming this plasmid into a strain that carried already an integrated TetO44-*URA3* sequence, the *URA3* sequence could be efficiently removed while maintaining an intact length of TetO44 array by selection with medium containing 5-FOA (Figure 4.1).

4.2.2.2 Standardising the induction time of the Gal-TetR-UDG* construct

Elizabeth Colby found that the expression of Gal-TetR-UDG* for 2 hours was sufficient to see formation of abasic sites mainly in the TetO44 array. Moreover, an induction time of 4 hours showed to be too damaging for cells. Here, to determine the experimental conditions to study the synthesis of nascent leading and lagging strands, I assessed the amount of damage I should use. Hence, I induced the Gal-TetR-UDG* construct for 2, 2.5 and 3 h using the

experimental layout described in Figure 4.2A, and checked how the induction time influenced checkpoint activation and cell cycle progression.

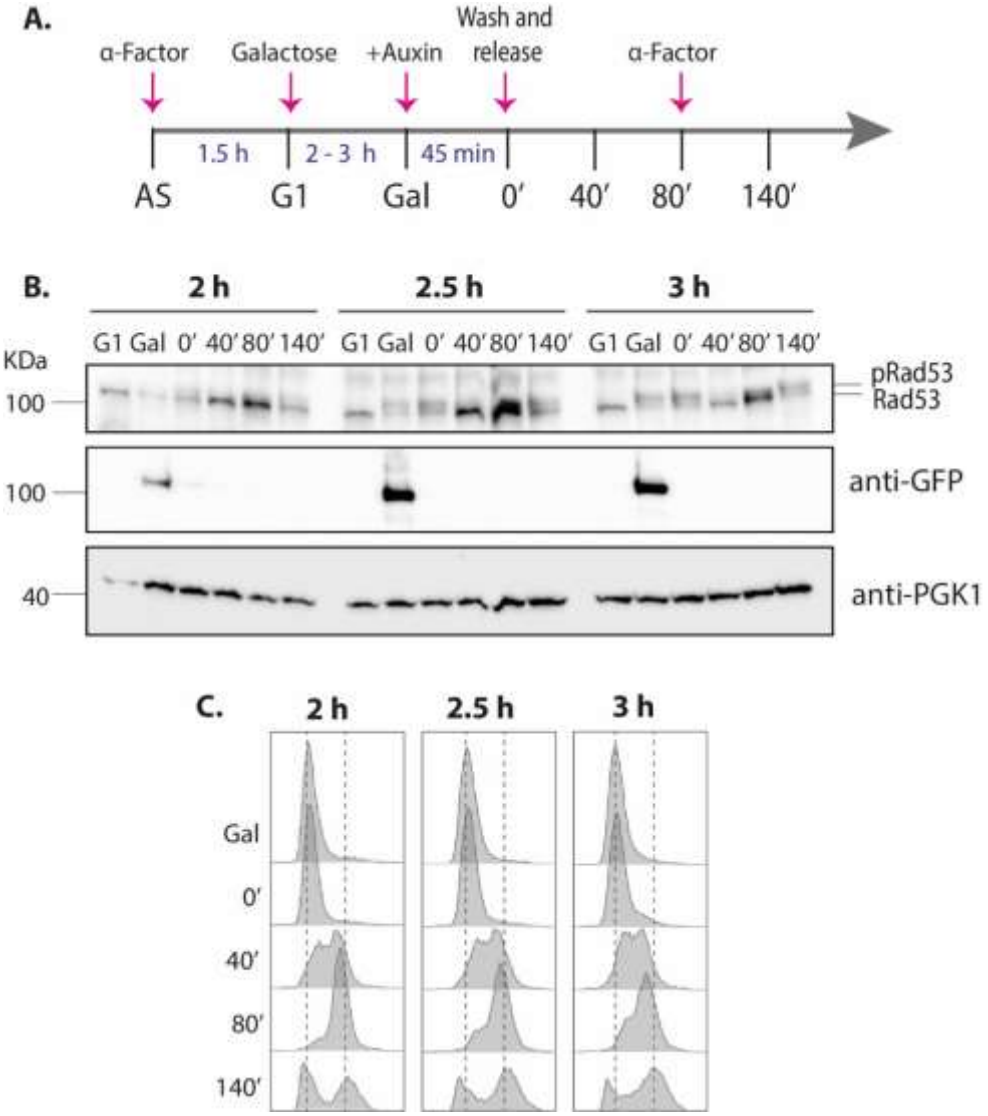

Figure 4.2. Testing different induction times for Gal-TetR-UDG*

A. Exponentially growing cells in YPD media were synchronised in G1-phase with α-factor for 1.5 hours. While in the G1-phase, UDG* was induced for 2, 2.5 and 3 hours by growing cells in YP-Gal. The construct was degraded by shifting cells back to YPD supplemented with auxin for 45 min. α-factor was washed out and cells were released into the cell cycle. To avoid cells progressing through the next S-phase, a pulse of α-factor was added in late G2-phase. Different time-points were collected for further analysis. **B.** Western blotting indicating the induction of UDG* **C.** Analysis of cell cycle progression. *Strain ID: 4880.*

As shown in Figure 4.2B, I observed that the levels of Gal-TetR-UDG* increased proportionately with the time of induction. Similarly, the levels of pRad53 increased with longer times of UDG* induction. Cell cycle profiles showed that cells where UDG* was induced for 2 h progressed slightly faster through the cell cycle than with longer UDG* expression times. Notably, expression of the TetR construct for 2.5 and 3 h did not show striking differences either for checkpoint activation or cell cycle progression (Figure 4.2C).

In light of these results, for subsequent experiments, I analysed replication events following induction of UDG* for 2 and 2.5 h.

4.2.2.3 Long setup to detect replication events within the TetO44 array

I reasoned that if TetO44 replication was re-established, this most likely occurred post-replication in late G2-phase, which meant that I needed to capture replication events at this stage. As the doubling time of cells with similar genetic background (Sections 3.3.2.1 & 0) varied from 1.5 to 2.7 h, I took time-points up to 4.5 h after release from G1-phase, by which time I expected genome replication to be completed. Additionally, to avoid capturing replication events from subsequent cell cycles, pulses of α-factor were added in early G2-phase to block cells progressing into the next S-phase (Figure 4.3).

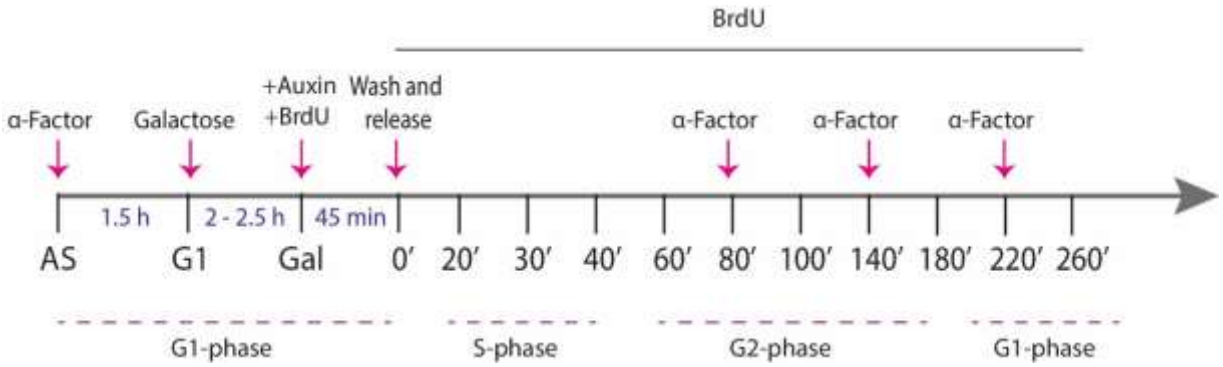

Figure 4.3. Experimental setup to study DNA synthesis in a damaged template
Cells exponentially growing in YPD media were synchronised in G1-phase with α-factor for 1.5 hours. While in G1-phase, UDG* was induced for 2 and 2.5 hours by shifting cells into YP-Gal. The construct was degraded by shifting cells back into YPD supplemented with auxin for 45 min. α-factor was washed out and cells were released into the cell cycle in YPD supplemented with 0.4 mg/mL BrdU. To avoid cells progressing into the next S-phase, pulses of α-factor were added in G2-phase. Different time-points were collected up to 4.5 hours for further analysis.

4.3 Synthesis of leading and lagging strands on damaged conditions

In this chapter, I aimed to elucidate to the following questions:

1. How are leading and lagging strands synthesised in the presence of abasic sites?
2. Is the replication of the TetO44 array re-established?
3. Is replication re-started downstream of a damaged template (i.e. TetO44 array)?
4. How PCNA and RPA associate to a damaged template?

In order to answer these questions, I used the AP site system to create abasic sites in the TetO44 array. I also utilised an experimental layout that allowed for collection of samples at different phases of the cell cycle, in this manner studying replication in time and space. Samples were processed for further deep-sequencing analysis with strand-specific BrdU-Seq, which permitted the tracking the synthesis of daughter strands specifically at the TetO44 array. In parallel, to investigate the dynamics of the core replisome proteins, PCNA and RPA within the array, I performed chromatin immunoprecipitation followed by quantitative PCR and deep-sequencing.

4.3.1 Chronic activation of the checkpoint DNA-damage response

First, I examined cells progression into the cell cycle and checkpoint activation upon damage.

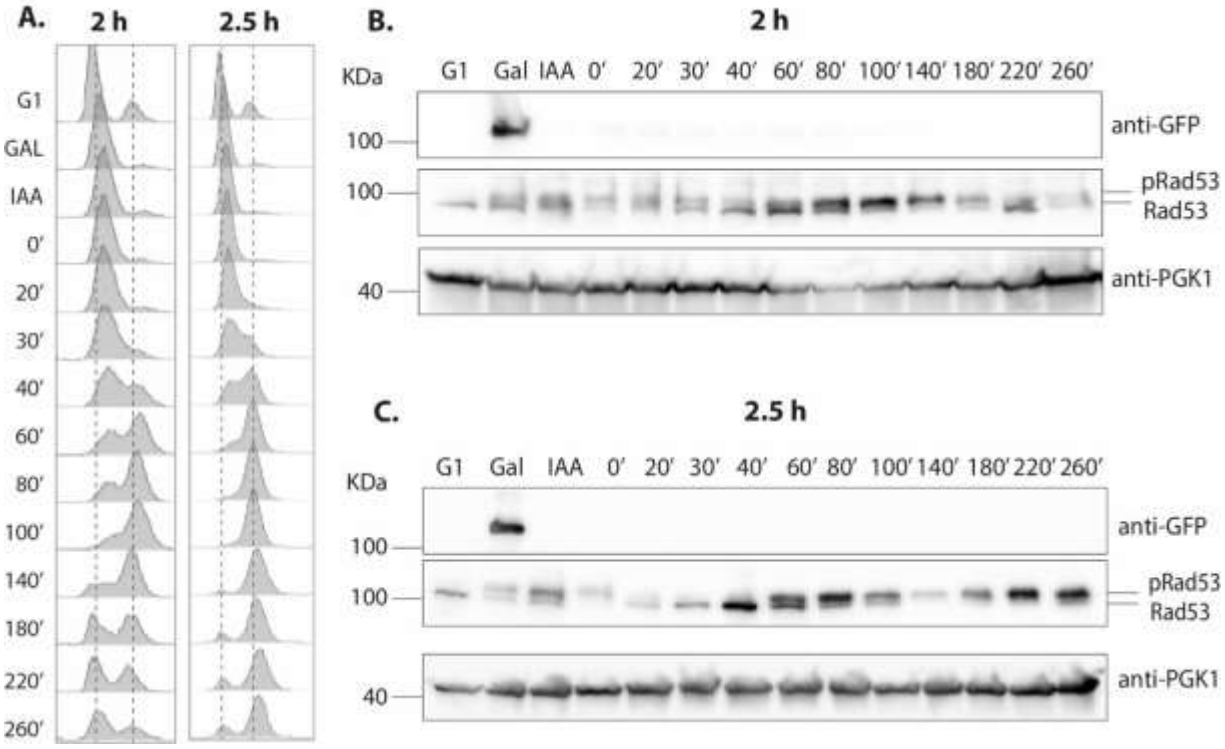

Figure 4.4. Cell cycle progression and checkpoint activation under 2 and 2.5 hours UDG* expression
Time-points were collected as indicated in Figure 4.3 for further analysis **A.** Analysis of cell cycle progression. **B-C.** Western blotting indicating checkpoint activation and induction of UDG* for 2 **(B.)** and 2.5 **(C.)** hours. *Strain ID: 4880.*

I detected very distinct cell cycle progression profiles for cells where UDG* was induced for 2- and 2.5 h.

On the one hand, cells with a 2 h UDG* induction accumulated appreciably in S-phase, there was not obvious accumulation in G2-phase and most cells passed to the next G1-phase (Figure 4.4A). Moreover, I observed chronical checkpoint activation from G1- to G2-phase, as exemplified by pRad53, which decreased considerably when cells entered into a next cell cycle (Figure 4.4B). Therefore, it appeared that abasic sites were repaired and DNA replication was resumed.

On the other hand, cells with a 2.5 h UDG* induction did not accumulated in S-phase, rather they stalled extensively in G2-phase. In fact, only a small fraction of cells passed to the next cell cycle (Figure 4.4A). The damage repair checkpoint was activated soon after UDG* induction, Rad53 was then de-phosphorylated in S-phase and re-phosphorylated extensively in G2-phase (Figure 4.4C). According to Uzunova and colleagues (2014), the de-phosphorylation of Rad53 in S-phase might be a sign of cells adaptation to damage, which generally occurs when cells meet impediments for coping with persistent damage during DNA replication, and it provides opportunities to repair damage in subsequent cell cycle phases.

4.3.2 Leading and lagging strand synthesis in the presence of abasic sites

I focused on studying how daughter strands were synthesised at the origin ARS607 and regions surrounding the TetO44 array following induction of UDG* for 2 and 2.5 h. For this, I immunoprecipitated the BrdU incorporated into genomic DNA to generate libraries for high-throughput sequencing by the IMB's Genomics Core Facility. Sequencing reads were trimmed and mapped to the yeast genome (sacCer3) by the IMB's Bioinformatics Core Facility as indicated in Section 2.12.4.

From the assay performed on cells with a 2 h UDG* expression, I observed three interesting behaviours. First, BrdU was not incorporated in the TetO44 array at any time-point, other than some minor incorporation in the lagging strand for the time-points corresponding to mid- and late-S-phase (40'-80'). Based on these results, I concluded that, on a damaged template, while the leading strand remained un-replicated, replication must be re-started for the lagging strand, though not successfully. Second, BrdU was incorporated downstream of the damaged region, possibly indicating re-priming and replication re-started after abasic sites were bypassed. However, I could not discard the possibility that DNA synthesis was due to the firing of the late origin ARS608 located ~15 kb downstream of the array. Third, I noticed that the replication of the TetO44 array was not re-established, as BrdU was not incorporated even 4.5 h after cells were released into the cell cycle (Figure 4.5A). Overall, the above-described analysis showed that the TetO44 array remained un-replicated in the presence of damage, but interestingly cells progressed through to the next G1-phase (Figure 4.4A). This behaviour could hint adaptation to damage, that is, if the damage was not resolved in G1-phase it would be in the following cell cycle (Pellicioli, et al., 2001). Therefore, it could be interesting to analyse additional time-points, past the 4.5 h used in these experiments, to confirm this hypothesis.

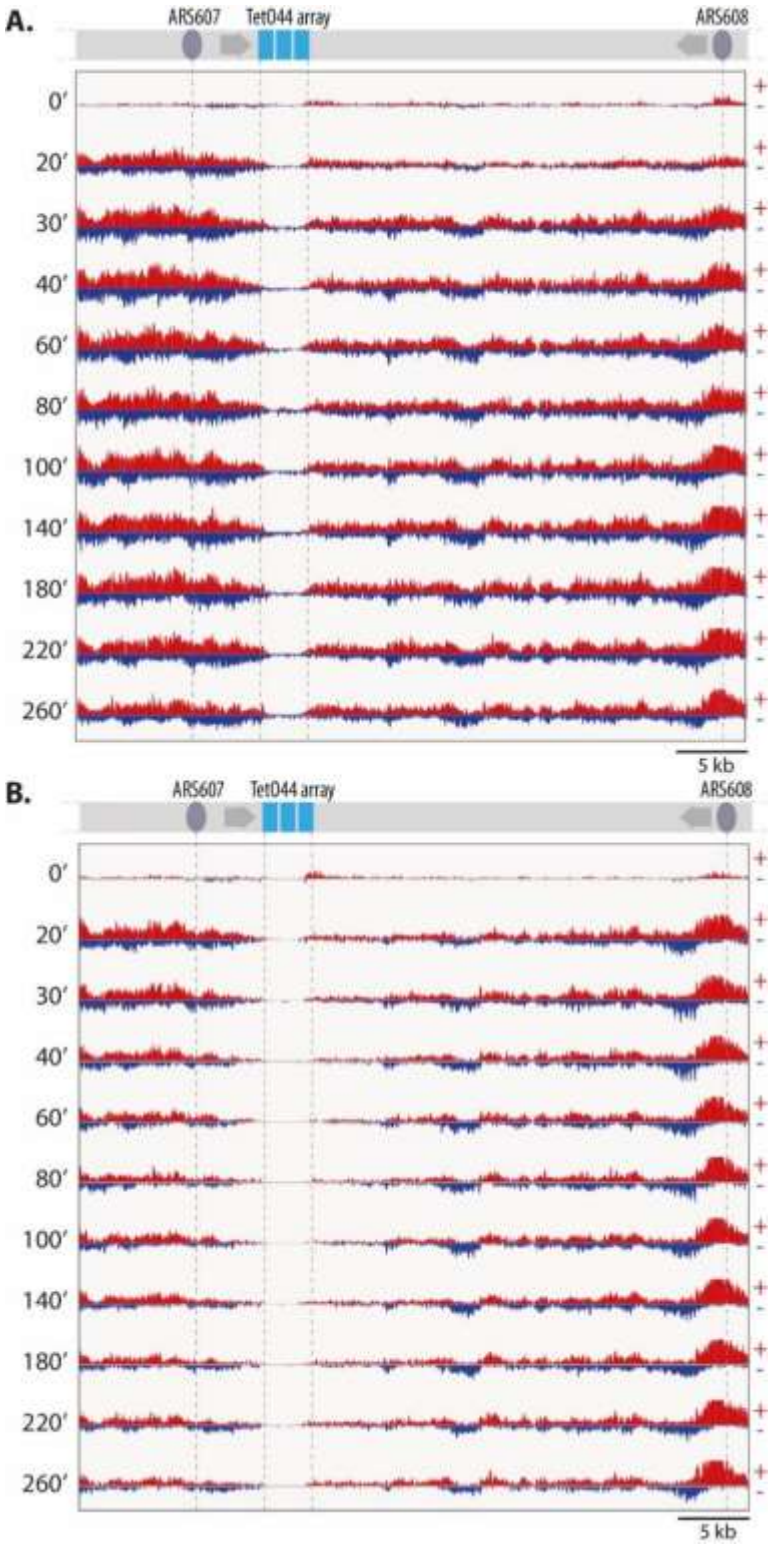

Figure 4.5. Synthesis of daughter strands within the TetO44 array.

BrdU incorporation profiles at the early firing origin ARS607 (ChrVI) obtained by BrdU-ssSeq **A.** 2-hour and **B.** 2.5-hour UDG* induction. Red (+) indicates leading strand. Blue (-) indicates lagging strand. DNA libraries were generated using the Accel-NGS® 1S Plus DNA Library Kit. Library pools were sequenced on a NextSeq500 sequencer using a High-Output 75 Cycle kit. 5 million reads of each sample were used. The sequencing reads were mapped to the yeast genome (sacCer3). IGV v2.13.2 was used to illustrate the sequencing reads (Data range:-12/12. Peak height: 65). *Strain ID: 4880. Project ID in Annexes.*

When UDG* was induced for 2.5 h, as more abasic sites were most likely formed within the TetO44 array, synthesis of daughter stands was very much disturbed. I observed three interesting behaviours. First, BrdU was not incorporated in neither the leading nor lagging strands, denoting that DNA remained un-replicated. Second, even though BrdU was incorporated downstream of the damaged region, this occurred very weakly in both strands. In fact, the incorporated BrdU seemed to be degraded in time-points corresponding to early G2-phase (40'-100'), possibly indicating that newly synthesised DNA was being resected. Similarly, it cannot be ruled out that DNA synthesis was due to the firing of the late origin ARS608. Moreover, BrdU was not incorporated in the array after 4.5 h, thus, replication of the TetO44 array was not re-established (Figure 4.5B). Interestingly, after 2.5 h of UDG* induction, the TetO44 array remained un-replicated and cells stalled in G2-phase (Figure 4.4A). This result perhaps suggested that high amounts of abasic sites represented too much damage for cells to adapt, meaning that they would stall in G2-phase until damage is resolved or until the replisome collapses.

4.3.3 Synthesis of DNA from active forks in non-damaged locus

Next, I investigated how daughter strands are synthesised at firing origins not related to ARS607 as a non-damage scenario. I focused in the early firing origin ARS305 in chromosome III and the late firing origin ARS501 in chromosome V.

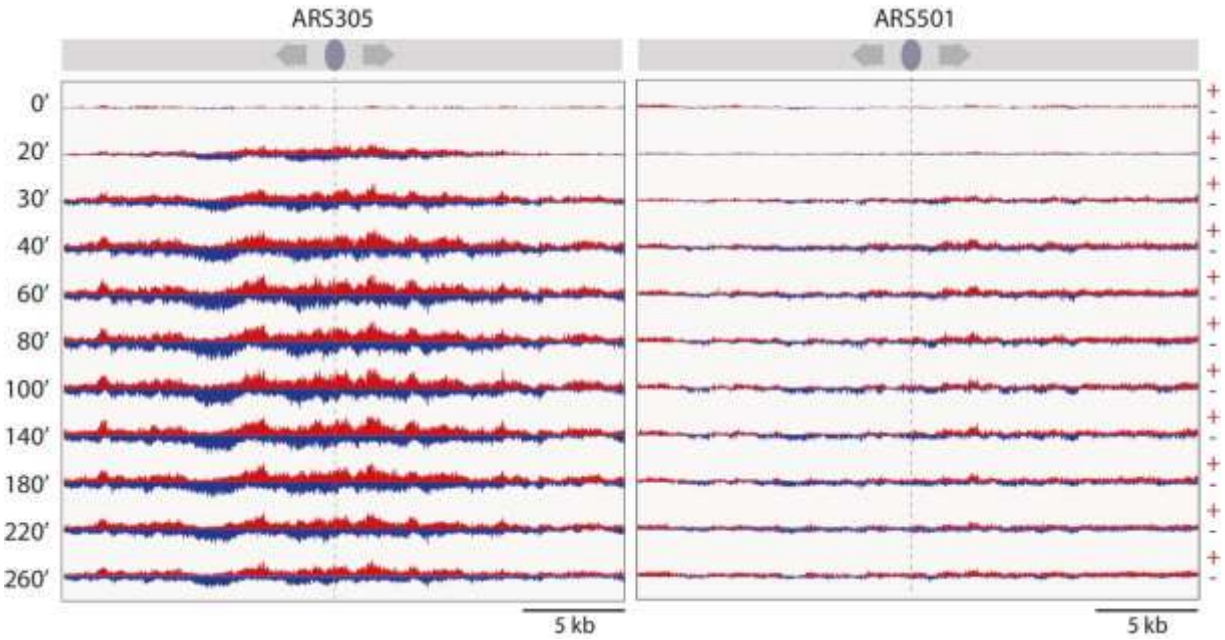

Figure 4.6. Synthesis of daughter strands in an early and a late firing origin.
BrdU incorporation profiles of the early firing origin ARS305 (ChrIII) and the late firing origin ARS501 (ChrV) obtained by BrdU-ssSeq. UDG* was expressed for 2 hours. Red (+) indicates leading strands. Blue (-) indicates lagging strands. DNA libraries performed using the Accel-NGS® 1S Plus DNA Library Kit. Library pools were sequenced on a NextSeq500 sequencer using a High-Output 75 Cycle kit. 5 million reads of each sample were used. The sequencing reads were mapped to the yeast genome (sacCer3). IGV v2.13.2 was used to illustrate the sequencing reads (Data range:-20/20. Peak height: 65). *Strain ID: 4880. Project ID in Annexes.*

As illustrated in Figure 4.6, I observed an obvious difference between early and late firing origins. The incorporation of BrdU in S-phase suggested that origin firing occurred as a continuum throughout S-phase, with the early origin ARS305 firing at time-point 20'. Moreover, as expected, at the late origin ARS501 was seen a delayed synthesis of DNA. This results confirmed that daughter strand synthesis was disturbed only upon the presence of abasic sites.

4.4 Dynamics of core replisome proteins during DNA damage encountering

The analysis of daughter strand synthesis showed that BrdU incorporated in the TetO44 array was being degraded overtime, possibly suggesting that newly synthesised DNA was being resected (Figure 4.5). To confirm this, I checked the enrichment and distribution of RPA (RFA in budding yeast) and PCNA by ChIP. In order to do this, I tagged the largest subunit of the RFA complex, RFA1, in its C-termini with the 9-myc tag and followed the protocol described in Section 2.11.5. The ChIPs were used either for qPCR or for library preparation and subsequently strand-specific high-throughput sequencing. The sequencing results were trimmed and mapped to the yeast genome (sacCer3) by the IMB's Bioinformatics Core Facility as indicated in Section 2.12.4.

4.4.1 Enrichment of PCNA and RFA1 in early and late firing origins

To compare the enrichment of PCNA and RFA1 at the origins of replication ARS501, ARS305 and ARS607, I performed a ChIP-qPCR and calculated the amount of DNA recovered in the IP as a percentage of input, calculated as indicated in Section 2.11.2.3.

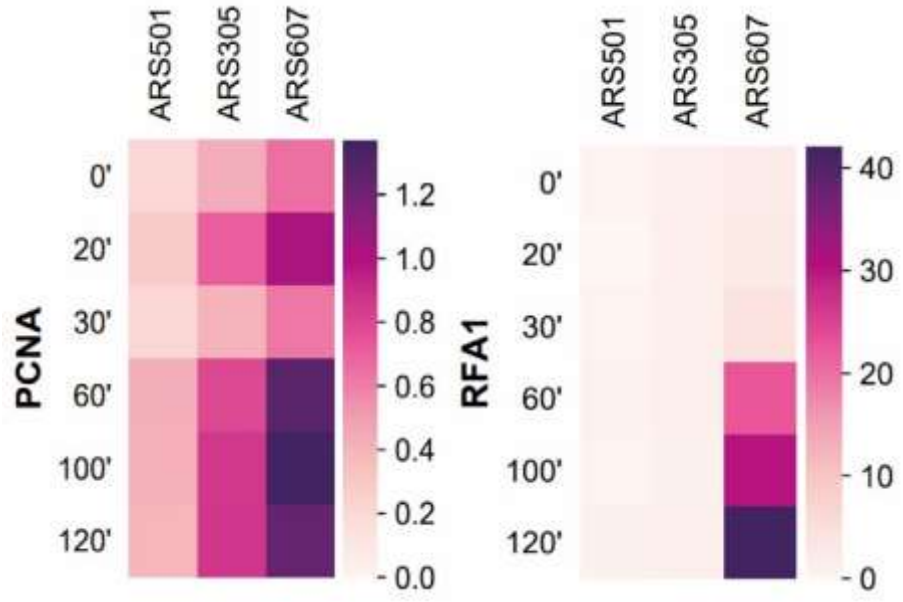

Figure 4.7. Enrichment heatmaps of PCNA and RFA1 on different origins of replication.
Time-points were collected as indicated in Figure 4.3 and then samples were processed for ChIP-qPCR. Heatmaps illustrate the amount of DNA recovered in the IP calculated as % input for PCNA and RFA1 on the early firing origins ARS305 and ARS607 and the late firing origin ARS501. *Strain ID: 4880.*

Illustrated in Figure 4.7, PCNA was much more enriched at a damaged locus (ARS607) than in a non-damaged locus (ARS305), probably reflecting the loading of PCNA not only for DNA replication but also for DDR. Moreover, the high enrichment of RFA1 at the later time-point at the ARS607 origin also likely indicated the response of the replisome to damage, which could ultimately led to resection events.

4.4.2 PCNA loaded at high levels at the site of damage

The ChIP-Seq results for PCNA detailed its distribution within the TetO44 array in time and space.

Shown in Figure 4.8, PCNA was strongly bound to the TetO44 array, especially in early time-points (G1- to S-phase) following induction of UDG* either for 2 or 2.5 h. Peculiarly, PCNA peaked just after abasic sites were formed. PCNA is pivotal for damage repair, thus a possible explanation for this observation is the activation of NER for AP site repair. The main role of PCNA in NER is during the re-synthesis of the excised DNA fragments, when it interacts with DNA Pol δ/ε (Strzalka & Ziemienowicz, 2011). Interestingly, PCNA also accumulated on the TetO44 array during S- and G2-phases, perhaps in preparation for DNA synthesis or for bypassing damage and binding of specialised translesion synthesis polymerases (Hedglin, et al., 2016).

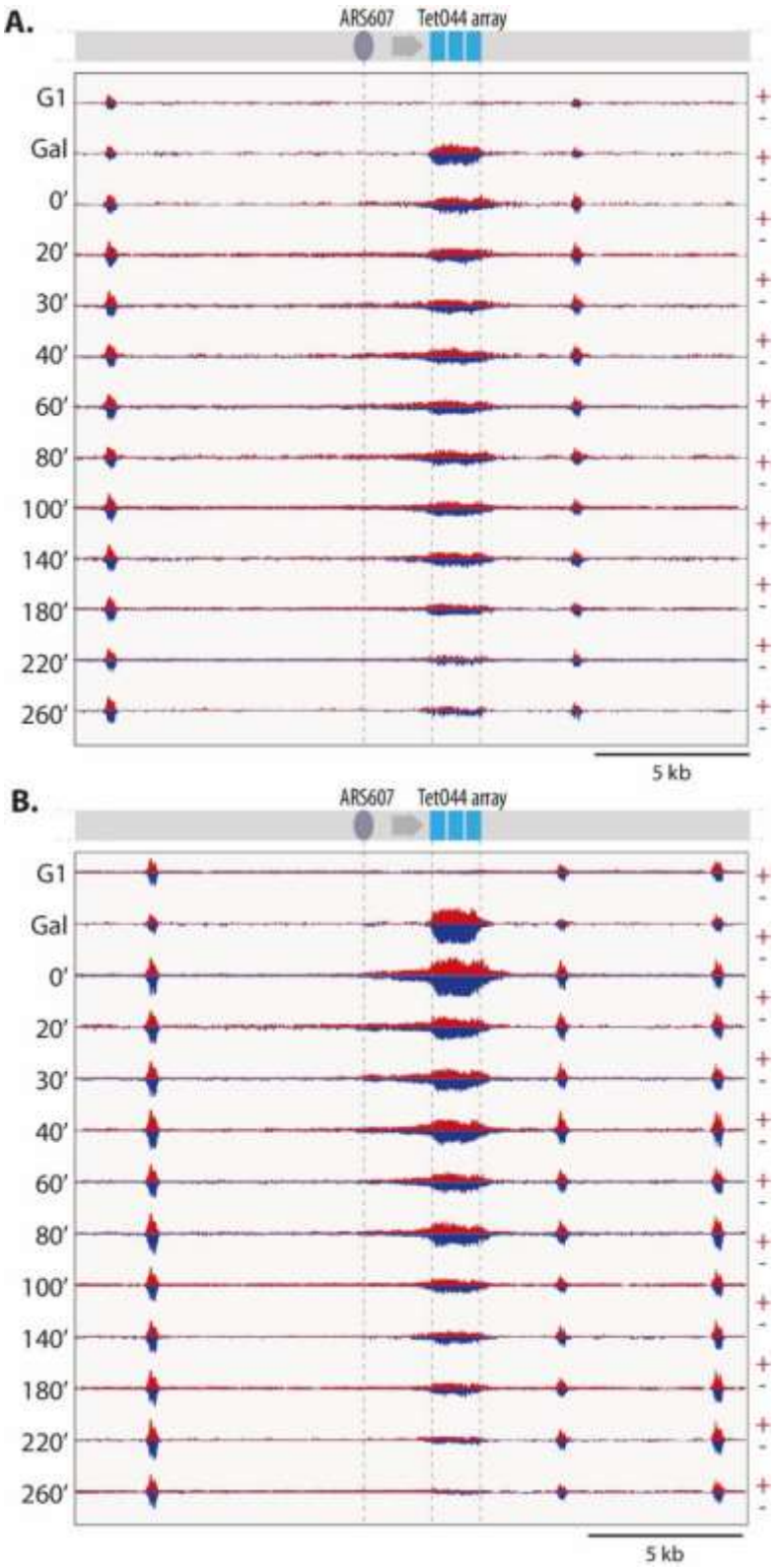

Figure 4.8. Time and spatial distribution of PCNA within the TetO44 array.
PCNA distribution profiles at the early firing origin ARS607 (ChrVI) obtained by ChIP-ssSeq **A.** 2-hour and **B.** 2.5-hour UDG* induction. Red (+) indicates leading strands. Blue (-) indicates lagging strands. DNA libraries were generated using the Accel-NGS® 1S Plus DNA Library Kit. Library pools were sequenced on a NextSeq500 sequencer using a High-Output 75 Cycle kit. 5 million reads of each sample were used. The sequencing reads were mapped to the yeast genome (sacCer3). IGV v2.13.2 was used to illustrate the sequencing reads (Data range:-50/50. Peak height: 50). *Strain ID: 4880. Project ID in Annexes.*

4.4.3 RPA coats the gap formed in the TetO44 array

To detail the distribution of RPA within the TetO44 array in time and space, I observed the binding of RFA1 by ChIP-ssSeq.

The sequencing results showed that RFA1 coated the TetO44 array as well as its surrounding regions on both the leading and lagging strands throughout the time-course. Interestingly, RFA1 spread in the direction of replication. These observation suggested that the DNA within the array was being resected, thereby provoking a gradual expansion of ssDNA gap. I found that RFA1 coated DNA up to a range of 15 kb from the TetO44 array. Moreover, the induction time of UDG* indicated that the longer the construct was expressed the larger was the amount of RFA1 coating the damaged region (Figure 4.9).

There is good evidence in literature for specific enzymatic activities contributing to an expansion of daughter-strand gaps overtime. In budding yeast, this is mediated by the exonuclease Exo1, which is responsible for long-range resection of DSB ends. The helicase Pif1 also has an auxiliary role in gap expansion, known to interact with PCNA, could further explain the high enrichment of PCNA on the TetO44 array (Figure 4.7 & Figure 4.8) (García-Rodríguez, et al., 2018A, García-Rodríguez, et al., 2018B).

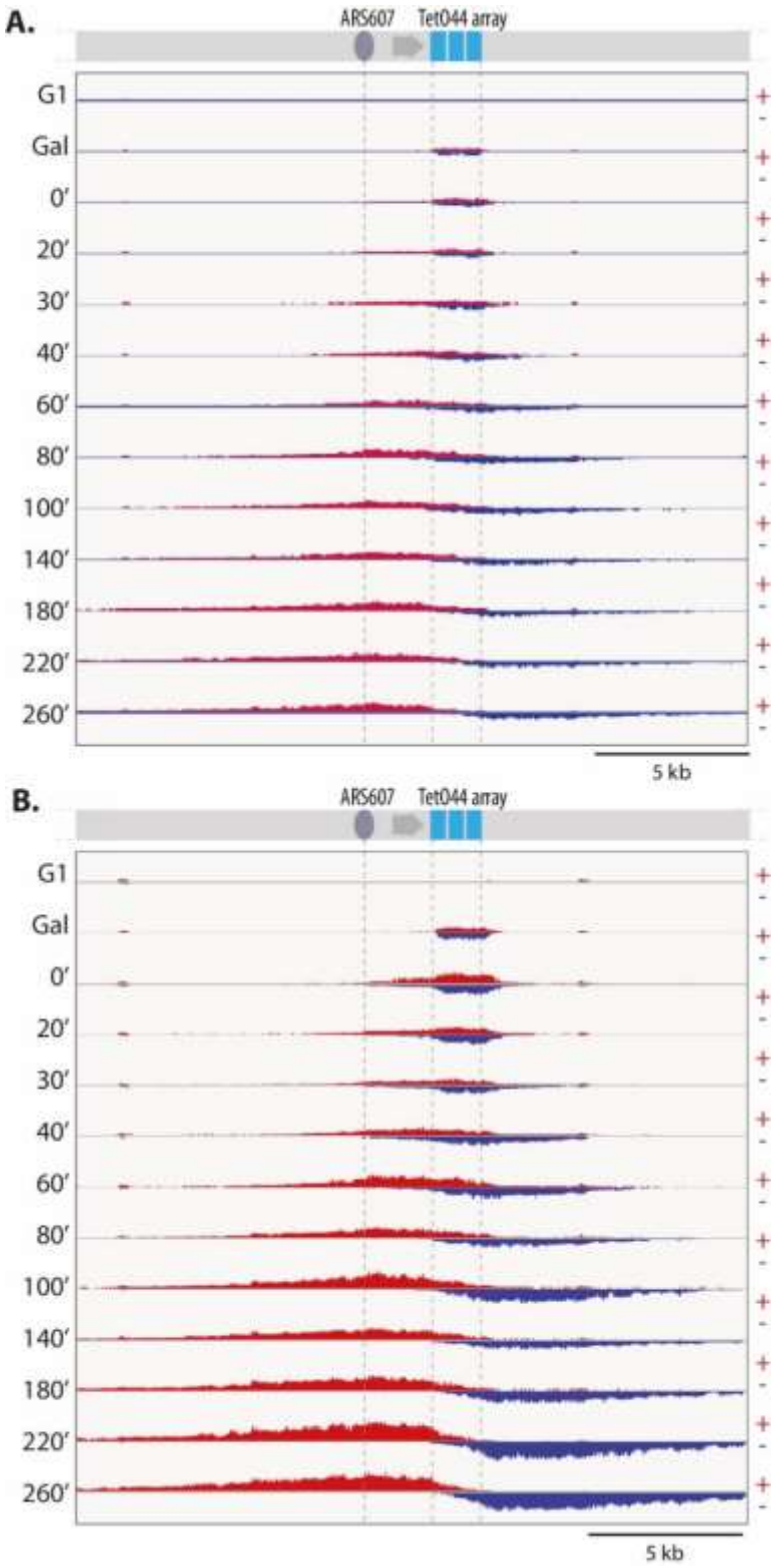

Figure 4.9. Time and spatial distribution of RFA1 within the TetO44 array.
RFA1 distribution profiles at the early firing origin ARS607 (ChrVI) obtained by ChIP-ssSeq **A.** 2-hour and **B.** 2.5-hour UDG* induction. Red (+) indicates leading strands. Blue (-) indicates lagging strands. IGV v2.13.2 was used to illustrate the sequencing reads (Data range:-150/150. Peak height: 40. DNA libraries were generated using the Accel-NGS® 1S Plus DNA Library Kit. Library pools were sequenced on a NextSeq500 sequencer using a High-Output 75 Cycle kit. 5 million reads of each sample were used. The sequencing reads were mapped to the yeast genome (sacCer3). *Strain ID: 4880. Project ID in Annexes.*

4.5 AP sites generate instability of TetO44 array

Replication stress is a highly dynamic chain of events that starts with the acute arrest of forks with fully assembled replisomes. When replication stress persists, stalled forks collapse, which causes the dissociation and/or impairs the modification of replisome components (Lambert & Carr, 2005). This latter has been documented in mammalian cells, where, upon damage, a high amount of RPA was loaded on chromatin to levels that were several fold higher than undamaged cells. In fact, it was found that this accumulation of RPA on chromatin was linked to the conversion of stalled forks to DSBs (Toledo, et al., 2013). Similarly, my results also potentially reflected the impair loading of the core replisome components PCNA and RPA, in the presence of abasic sites (Figure 4.8 & Figure 4.9). Therefore, in order to examine if the stalled fork led to DSBs formation, I checked the intactness of the TetO44 array.

4.5.1 Abasic sites are processed into double-stand breaks

To examine the presence of DSBs at the TetO44 array after AP sites formation, I performed a native Southern blot. I achieved this by digesting the input DNA with BssSI-v2. The resulting fragments were separated on a gel electrophoresis run followed by Southern blot analysis. I included a TetO44 array with an artificial DSB control (+) by additionally digesting it with XmaI (Figure 4.10A).

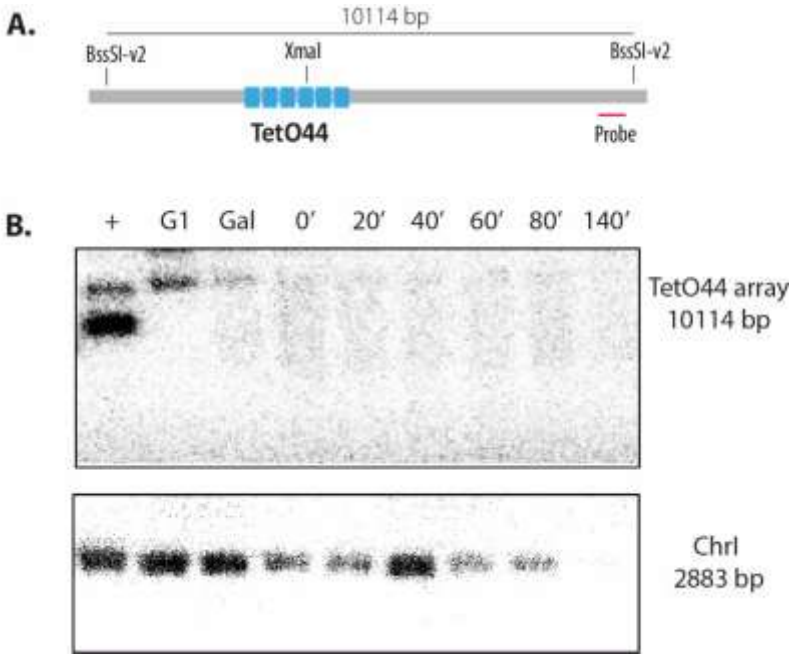

Figure 4.10. Native Southern blot to check DSBs in the TetO44 array under damage conditions
The samples used for the BrdU-ssSeq assay were used for examining DSBs by native Southern blotting. **A.** Input DNA was digested with BssSI-v2, which allowed to study a region of 10.1 kb within the TetO44 array. The radioactively labelled ssDNA probe hybridised downstream of the array. **B.** Native Southern blot. The blot was first probed for the TetO44 array, and then probed for the control region on Chr1. *Strain ID: 4880.*

The Southern blot revealed the degradation of the TetO44 array band just after abasic sites were formed, which indicated the presence of DSBs at this locus in early G1-phase (Figure 4.10B). However, I was not able to discriminate whether multiple breaks were formed or at least one DSB per cell randomly formed in the array.

This result confirmed that the formation of abasic sites immediately led to their processing into DSBs, and possibly justifying that the DDR directed to the high enrichment of RPA and PCNA at the site of damage.

4.5.2 Quantitative comparison of genomic DNA over the time-course

In a BrdU-ssSeq experiment, the libraries generated from the input DNA act as reference for how well a locus can be potentially represented in the IP. Consequently, I used the sequencing information from the input DNA from the BrdU-ssSeq experiment described in Section 4.3 to examine the integrity of the template during the time-course after UDG* was induced for 2 h.

By comparing the sequencing data for the Watson and Crick strands in the AP site system setup, I observed that the sequenced reads count within the TetO44 array decreased over the time-course, with the Watson strand being the most affected. In fact, I noticed that the sequenced reads count at the time-point 260' dropped to 20% compared to time-point G1 (Figure 4.11A). Interestingly, this trend was only seen for the fork stalled at the early origin ARS607 and not for the active fork at early (ARS305) and late firing origin (ARS501) (Figure 4.11B-C).

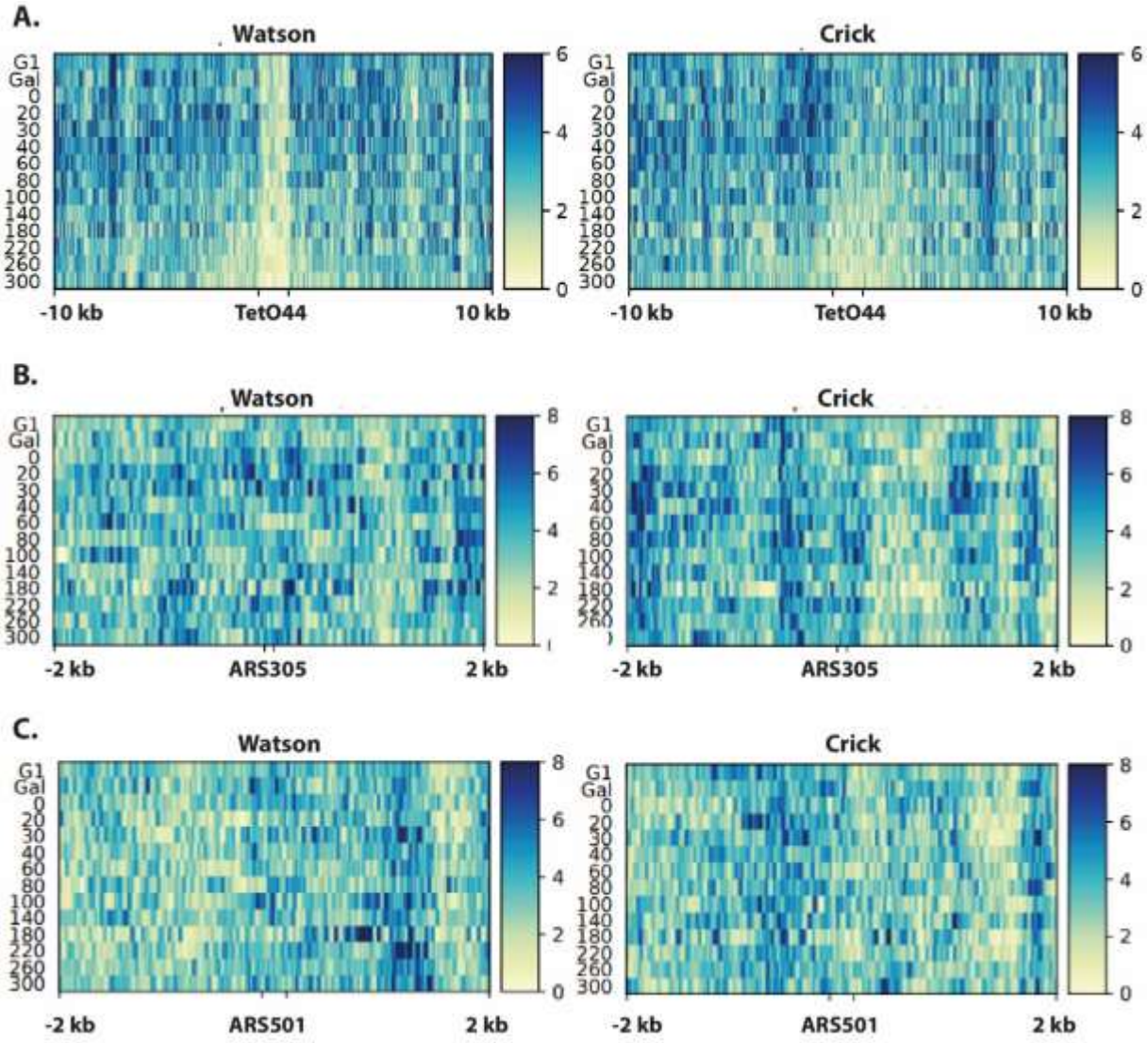

Figure 4.11. Heatmaps representing sequencing depth of input samples.
The sequencing depth of the input samples used in Section 4.3 is shown as heatmaps for template integrity analysis. **A.** TetO44 array and regions -/+10kb. **B.** Early firing origin ARS305 and regions -/+2kb. **C.** Late firing origin ARS501 and regions -/+2kb. *Plot provided by Giuseppe Petrosino (Bioinformatics CF). Strain ID: 4880. Project ID in Annexes.*

Altogether these results further suggested that daughter strands were not replicated and/or strand resection occurred.

4.6 Discussion

In this chapter, I found that in the presence of AP sites at the TetO44 array, the leading and lagging strands remained un-replicated. Even though some attempts of lagging strand synthesis were detected, these were inefficient, and its occurrence seemed to be dependent on the amount of damage induced. Similarly to previous studies, I found that replication re-started downstream of the lesions in both strands, most likely due to re-priming events. Moreover, I observed that the lack of synthesis in the TetO44 array led to the formation of an ssDNA gap, which gradually expanded in the direction of the movement of the replisome. In parallel, I found that the presence of AP sites altered the loading levels of PCNA and RPA at the site of damage. Finally, I showed that the induction of abasic sites provoked the formation of DSBs at the TetO44 array.

4.6.1 Daughter-strand gap expansion

In Section 4.3.2, I showed that upon replication fork stalling, the TetO44 array remained un-replicated, forming stretches of single-stranded DNA. The increased RPA binding within the damaged region up to a range of 15 kb suggested that the ssDNA gap expanded gradually over time. Additionally, I found that abasic sites were processed into DSBs (Section 4.5.1), indicating that ssDNA could be a manifestation that emerged from the resection of DSB ends (Wong, Petriukov, & Ulrich, 2021). A possible explanation for these observations is that DSBs were formed from nicks on the parental DNA strands as intermediates of NER in G1-phase. Even though damage bypass and checkpoint signalling were expected to counteract such events, extended regions of ssDNA might simply be more vulnerable and therefore led to further instability.

Limoli and colleagues (2002) described the detrimental degradation of nascent DNA in the context of stalled forks. By studying UV-irradiated XPV cells (human *Xeroderma pigmentosum* variant) deficient in the TLS Pol η, they found that Mre11 accumulated and co-localised with the damage marker γH2AX in a replication-dependent manner. Analogously to what I showed in this chapter, they observed that DNA damage created daughter-strand gaps that later led to the formation into DSBs. Interestingly, it was too found that the exonucleolytic activity of Mre11 and Exo1 participated in ssDNA gap extension in this situation (Federico, et al., 2020). In the context of the AP site system, perhaps Mre11 and Exo1 also had a role in resecting DNA, thus it would be interesting to confirm this hypothesis further.

In budding yeast, an excess of ssDNA at arrested forks activates Rad53, which prevents replication fork breakdown and inhibits Exo1 activity. However, when the activity of Exo1 is deregulated, it can lead to excessive resection and the formation of long ssDNA gaps (García-Rodríguez, et al., 2018A). I observed a high degree of DNA resection at the TetO44 array after UDG* was expressed, situation that occurred possibly due to Exo1 deregulation. The exact molecular steps involved in this process remain unclear. In fact, previous reports in fission yeast indicated that uncontrolled resection of nascent DNA by Exo1 at a paused

replication fork generates a structure that cannot be rescued by a converging fork (Saada, et al., 2017). Provided that, when UDG* was expressed for 2.5 h, cells strongly stalled in G2-phase (Figure 4.4A), it could be an indication of the presence of an irreparable damage.

The helicase Pif1 has also been found to be a promoting factor for long-range resection in the Exo1 branch (Jimeno, et al., 2018). In fact, PCNA helps recruiting Pif1 to the sites of damage and could also serve to enhance its processivity during gap extension (Moldovan, et al., 2007). Moreover, Pif1 has been revealed to play an important role in several aspects of genome maintenance, due to its link to template switching and break-induce repair (García-Rodríguez, et al., 2018B; Li, et al., 2021). It is therefore tempting to speculate that Pif1 could also be involved in the long-range resection of daughter strands for the repair of abasic sites probably by TS or BIR.

4.6.2 Different functions of PCNA at the TetO44 array

I indicated that upon abasic sites formation, PCNA was loaded to the site of damage, that is, the TetO44 array and neighbouring regions, from early G1-phase until G2-phase (Figure 4.8). These findings were not surprising because PCNA tethering is a common theme in damage repair. This is well exemplified by the large of number of repair proteins in the PCNA interactome (reviewed in Cortez, 2019). For instance, the loading of PCNA in early G1-phase was most likely due to the recognition of the lesions and repair pathway e.g. NER. During replication, cells could have circumvented the damage by using a group of specialised TLS polymerases, which also require PCNA (Essens, et al., 2005). On the contrary, as mentioned above, PCNA loading onto the TetO44 array could have aided the recruitment of Pif1 to resect daughter strands.

Even though the loading of PCNA to the TetO44 array showed no clear bias between leading and lagging strands (Figure 4.8), previous findings revealed that upon DNA damage PCNA was mainly unloaded from lagging strands (Yu, et al., 2014). Yu and colleagues (2014) suggested that PCNA unloading could be to facilitate the loading of the 9-1-1 complex and thus promote checkpoint DDR. Another possible explanation is that PCNA unloading facilitated the recruitment of TLS polymerases to the leading strands of the stalled forks. On the contrary PCNA unloading could be a mechanism whereby checkpoint kinases regulate fork stability under replication stress. More studies are needed to check which, if any, of these possibilities applies, and ultimately gain insight into the strand-specific regulation of PCNA under DNA damage.

4.6.3 Checkpoint response and amount of DNA damage

Elizabeth Colby confirmed that the time of UDG* induction is correlated with the amount of abasic sites formed within the TetO44 array. Thus, it was expected that expressing UDG* for 2 h should have generated attenuated cell effects compared to longer expression times. An example of these effects was seen through RPA loading, where a 2.5 h UDG* induction showed a stronger RPA signal than a 2 h UDG*induction (Section 4.4.3). Moreover, in Section

4.3.1, I described that a 2 h UDG* induction provoked the stalling of cells in S-phase however most passed to the next cell cycle. On the contrary, a 2.5 h UDG* induction caused the chronic stalling of cells in G2-phase. Altogether these results suggested that different damage responses could be activated depending on the amount of AP sites formed within the TetO44 array.

Interestingly, Bantele and colleagues (2019) already suggested that the DNA damage checkpoint is not a single pathway, but rather at least two distinct ones that can be discriminated by their ability to quantify DNA damage. For instance, a high load of DNA damage triggers an immediate, Mrc1-dependent checkpoint response, whilst low-levels of DNA lesions are presumably skipped by the replisome and channelled into the Rad9 branch. The amount of ssDNA gaps has been proposed to modulate the checkpoint response, thus controlling cell-wide processes, such as cell cycle progression or DNA replication. (Wong, Petriukov, & Ulrich, 2021; García-Rodríguez, et al., 2018A). According to these reports, it is possible to assume that the amount of damage formed at the TetO44 array using different induction times could have determined the mechanism to resolve damage in my strains.

4.6.4 The high enrichment of RPA inhibits re-priming

In Section 4.3.2 I showed that inducing expression of UDG* for 2 h, BrdU was transiently incorporated into the TetO44 array only on the lagging strand, possibly reflecting attempts to re-start synthesis. Even though synthesis was inefficient, the canonical model for discontinuous lagging-strand synthesis might be the mechanism behind these observations.

On the contrary, in addition to remaining un-replicated, RPA looked to be coating the leading strand more than the lagging strand at the TetO44 array (Figure 4.11). Interestingly, Taylor & Yeeles (2018) indicated that the re-start of a leading strand downstream of a single CPD is highly influenced by the availability of RPA, since re-priming by Pol α is only possible at reduced RPA levels. This discovery suggested that the lack of leading-strand-synthesis re-start at the TetO44 array in my strains might have been due to the high level of RPA loading inhibiting the action of Pol α.

Furthermore, the incorporation of BrdU at regions downstream the TetO44 array showed that replication was indeed re-started, this observation was prominent especially when UDG* was induced for 2 h. For longer expression times, the weak incorporation of BrdU indicated that re-start events occurred, but in an inefficient manner, suggesting that re-priming may eventually have become ineffective due to an extremely high density of lesions.

Chapter 5: Results III - Genome-wide analysis of post-replication repair

5.1 Background

Lesions encountered during DNA replication are often bypassed by DNA-damage tolerance mechanisms. DDT broadly includes the translesion synthesis and template switching pathways, and it is well known that the mono-ubiquitylation and poly-ubiquitylation of PCNA commands the activation of TLS or TS, respectively (Gao et al., 2017). The nature of the replication challenge as well as the extent of fork stalling caused by different concentrations of a genotoxic agent have seen to be key determinants for the pathway choice (reviewed in Quinet, et al., 2021). However, how cells precisely choose between these two pathways is still unclear. During the last years, the Ulrich laboratory has dedicated extensive efforts to elucidate the mechanisms behind post-replication repair. Davies, et al., (2008), for instance, indicated that damage-induced PCNA ubiquitylation requires the ssDNA-binding protein, RPA, for the recruitment of Rad18 to ssDNA. Moreover, Daigaku, et al., (2010) demonstrated that both Rad6 pathway branches are separable in time and space from genome replication. Furthermore, García-Rodríguez, et al., (2018A) showed that exonuclease activity of Exo1 is needed in preparation for strand invasion during TS.

As a continuation of the work initiated by Nataliia Gralievska, during this chapter, I intended to further characterise how PRR is organised temporally and spatially in the genome of budding yeast. In addition, I aimed to understand how cells maintain the balance between TLS and TS upon treatment with the alkylating agent MMS. In order to do this, I quantified PRR tracts distributed across the genome by using the doxycycline-repressible Rad18 system and the NGS tool BrdU-ssSeq. Here, I confirmed that TS is the preferred pathway over TLS for bypassing lesions induced by MMS. Interestingly, I found enrichment biases of PRR tracts mostly towards regions around early firing origins. Finally, I observed a greater contribution of TLS and TS on leading strands than on lagging strands.

5.2 Experimental approach

The quantification of PRR tracts was performed with a previously established system in our laboratory in which a doxycycline-repressible construct for Rad18 is used to separate the bulk genome replication from post-replication repair (Daigaku, *et al*, 2010). To be able to study the genome-wide distribution of PRR tracts, I collected samples upon Rad18 re-expression and continued with the protocol described in Section 2.11.6 for the immunoprecipitation of BrdU. Then these samples were input for library preparation and subsequently NextSeq500 sequencing by the Genomics facility. Sequencing reads were trimmed, mapped to the yeast genome (sacCer3), and analysed by the Bioinformatics facility.

5.2.1 The doxycycline-repressible Rad18 system

Yasukazu Daigaku, a former postdoc in the Ulrich laboratory, created the doxycycline-repressible Rad18 system. This system takes advantage of the bacterial Tet Operator (TetO) sequence and its binding partner the protein Tet Repressor (TetR). The experimental conditions used during this chapter are described as follows.

Yeast background: The DF5 yeast background was used during this study. For the purpose of BrdU incorporation, genes encoding exogenous thymidine kinase (*HSV–TK*) and a nucleoside transporter *(hENT1*) were introduced on an integrative plasmid (Viggiani & Aparicio, 2006). Further features were combined by subsequent transformations or by mating and tetrad dissection.

Tetracycline-regulated activator/repressor Rad18 system: The dual Tet-Rad18 system was constructed by replacement of the *RAD18* promoter with a TetO7 array. Moreover, the system introduces a construct with a reverse TetR (TetR') fused to a VP16 activator and a general transcription repressor, Ssn6, to render Rad18 repressible by doxycycline (Bellí, et al., 1998).

Experimental layout: Saturated yeast cultures were taken into liquid YPD supplemented with doxycycline and left overnight on rotating incubator to repress Rad18. Next day, cells were arrested in G1-phase and treated with MMS before released into S-phase. Then, upon the re-expression of Rad18 and in the presence of a thymidine analogue, BrdU, PRR tracts were marked and captured for BrdU-ssSeq (Figure 5.1).

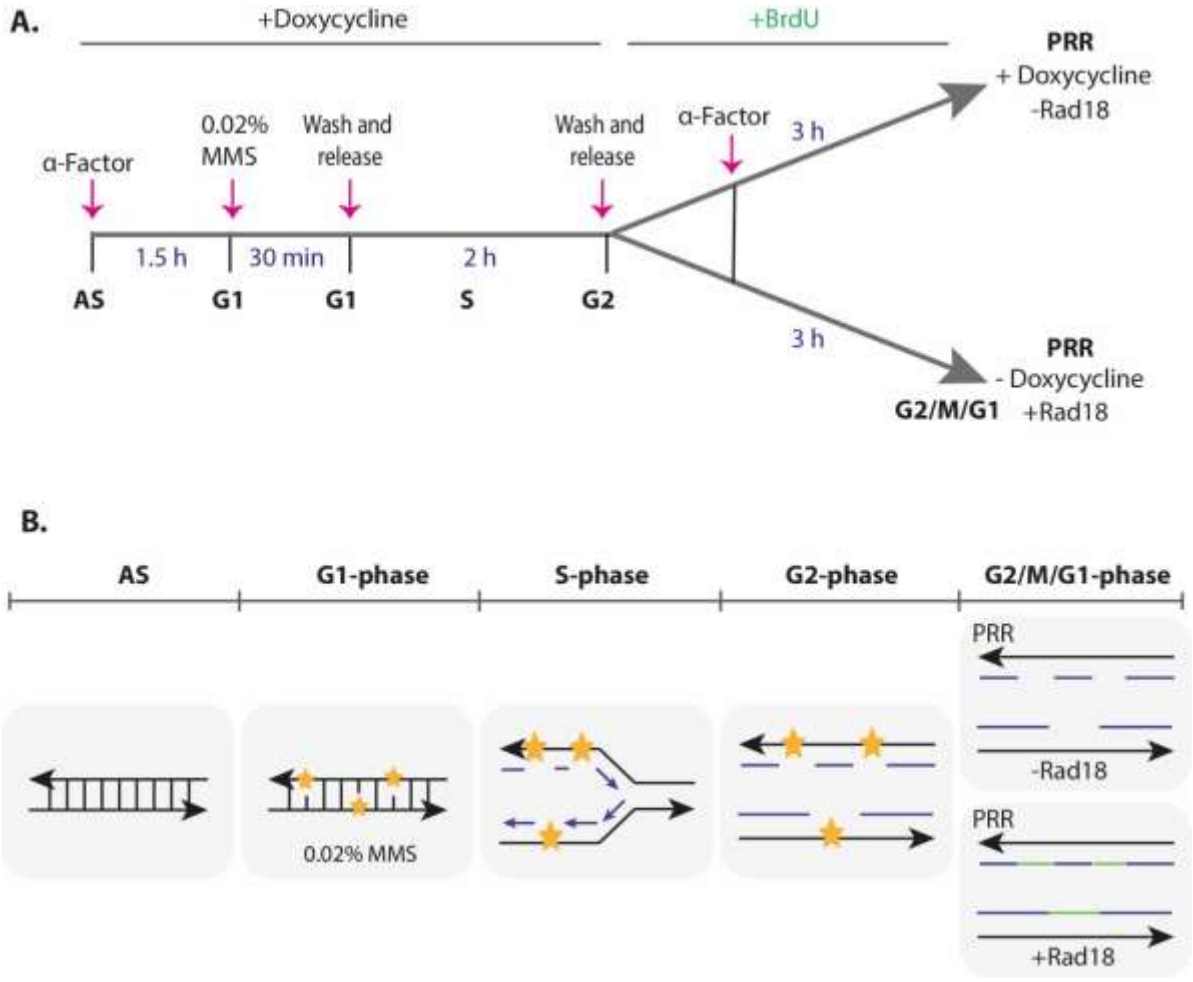

Figure 5.1. Quantification and visualisation of PRR tracts

A. 0.001-0.01 OD saturated yeast cultures were first taken into clean YPD media supplemented with 2 µg/mL doxycycline an incubated at 30°C overnight. Next day, cells grown to 0.5 OD were synchronised in G1-phase by adding 5 µg/mL α-factor (1.5 hours), then cells were treated with 0.02% MMS (30 min). Cells were released into S-phase (2 hours) by washing out α-factor twice with warm media supplemented with doxycycline. To take samples with and without Rad18 re-expression, equal volumes of cell culture were divided in two flasks. For re-expression of Rad18, doxycycline was washed out twice with warm media and cells were resuspended in warm YPD containing 0.4 mg/mL BrdU. For repression of Rad18 the washing step was skipped and only 0.4 mg/mL BrdU was added to the culture. α-factor was added 1 hour after release to stop cells passing to next S-phase After 3 hours, approximately 200 OD cells were collected for further analysis with BrdU-ssSeq. **B.** Scheme of events occurring during the experimental layout. *AS:* Undamaged dsDNA. *G1-phase:* MMS treatment induced lesions in DNA. *S-phase:* Bulk replication and replication stress. *G2-phase:* Lesions formed ssDNA gaps. *+Rad18:* PRR filled-in the ssDNA gaps. *–Rad18:* PRR is not active thus, gaps are not filled-in. Blue lines: bulk replication tracts. Green lines: post-replication repair tracts.

5.3 Previous findings

Nataliia Gralievska initiated the genome-wide analysis of post-replication repair project. First, she intended to determine if PRR tracts were specifically enriched in a certain region of the genome. In order to accomplish this, she followed the experimental layout described in Figure 5.1A, and the distribution of PRR tracts was examined in: regions with different replication timings (early to late firing origins), coding and non-coding regions (ORFs and IRs), and specific DNA structures regions (R-loops and G-quadruplexes).

Next, I describe some of her findings:

Early and late firing origins: Nataliia found enrichment biases of PRR tracts towards early replication origins, probably reflecting that PRR preferentially acts on regions near early and not late origis of replication. On the contrary, this finding also suggested that late origins of replication were blocked upon damage, or perhaps these regions were repaired by other mechanism even before genome bulk replication initiated (Figure 5.2).

- Role of BER: By observing the PRR-tract distribution obtained from Mag1-deficient cells, Nataliia aimed to determine if base excision repair had a role in repairing the damage induced by MMS. From this experiment two interesting observations were found. First, it was observed that the enrichment bias of PRR tracts towards early origins was lost. Second, it was seen that MMS-treated *mag1Δ* cells stalled in S-phase and only upon Rad18 re-expression cells progressed to G2-phase. Altogether, these findings suggested that both BER and PRR were essential for damage repair

Open reading frames (ORFs): It was observed enrichment of PRR tracts on the template strands where RNA Pol traverses, which meant that lesions were located on the opposite strand (non-template) where PRR acted after replication. This finding led to the assumption that PRR did not have a role at ORFs possibly because another mechanism had removed lesions maybe even before genome bulk replication. As TC-NER plays a role in DNA damage repair at moderate and highly transcribed genes, it was thought that lesions in ORFs were resolved preferentially by TC-NER. However, the role of TC-NER was not confirmed.

Intergenic regions (IRs): It was observed the enrichment of PRR tracts at IRs, which corresponded to template strands neighbouring ORFs. As IRs are regions in the genome that locate between two mRNA coding genes, it was suggested that PRR was preferentially active when damage was located in non-coding templates.

Centromeres: It was found a Rad18-dependent enrichment of PRR tracts at regions near centromeres.

R-loops and G-quadruplexes: There was no PRR-tract enrichment in neither R-loops nor G-quadruplexes structures.

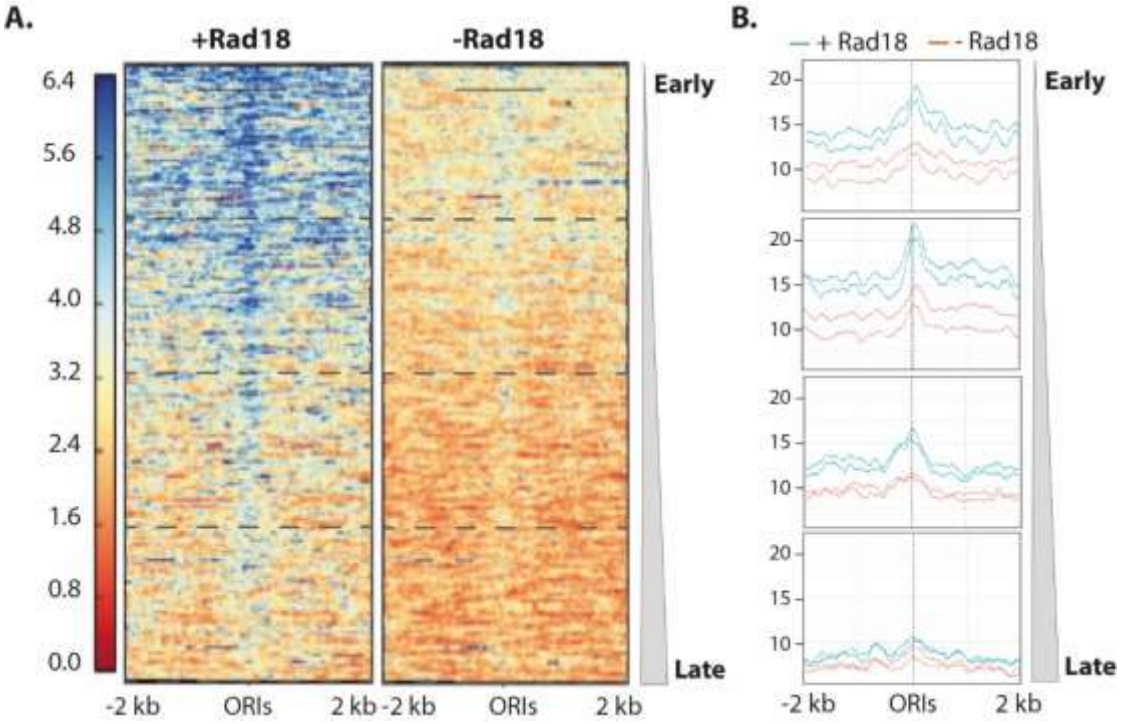

Figure 5.2. Quantification and visualisation of PRR tracts biases towards early origins of replication
Genome-wide coverage of PRR tracts of samples with and without Rad18 re-expression obtained by BrdU-Seq. The experimental layout used during this assay is described in Figure 5.1A. **A.** Heatmaps and **B.** line plots representing the enrichment of PRR tracts around ORIs sorted according to replication timing in four quartiles (from early to late firing ORIs). The signal extended for 2 kb to the 3'- and 5'-sides from the ORIs. Representative figure of two repetitions. *Plots provided by Nastasja Kreim (Bioinformatics CF). Strain ID: 2222 (taken from Nataliia Gralievska, unpublished).*

5.4 Contribution of TLS and TS in time and space

Previous studies have showed that the contribution of translesion synthesis and template switching differs according to the nature of the replication challenge. For instance, TS-defective cells are sensitive to MMS treatment, whilst TLS-deficient cells are not, suggesting that TS is the preferable pathway to tolerate MMS-induced lesions (Xiao, et al., 2000).

To get further insights into the contribution of TLS and TS in MMS-treated cells, I investigated if TLS and TS are contributing to lesions bypass in a time-dependent manner. In order to accomplish this, I analysed the genome-wide distribution of PRR tracts of TLS-deficient (*rev1Δ, rev3Δ, rad30Δ*) and TS-deficient (*ubc13Δ*) strains. In this manner, the *TLSΔ* strain should be devoid of TLS activity, and the *TSΔ* strain should prevent PCNA poly-ubiquitylation and thereby error-free damage bypass (Daigaku, et al., 2010). I followed the experimental layout described in Figure 5.3A and to capture the timely contribution of TLS and TS, I collected samples 1, 2 and 3 hours after Rad18 was re-expressed (Figure 5.3D).

Before proceeding with the BrdU-ssSeq assay, I examined cell cycle progression on both strains with and without Rad18 re-expression. Illustrated in Figure 5.3B, I observed that *TLSΔ* and *TSΔ* strains had not a measurable effect during S-phase. In addition, upon the re-expression of Rad18, it was seen that after 3 h both TLS- and TS-deficient cells progressed to the next G1-phase, which implied that damage was being resolved. Moreover, it was seen that the repression of Rad18 highly reduced the incorporation of BrdU into genomic DNA (Figure 5.3C).

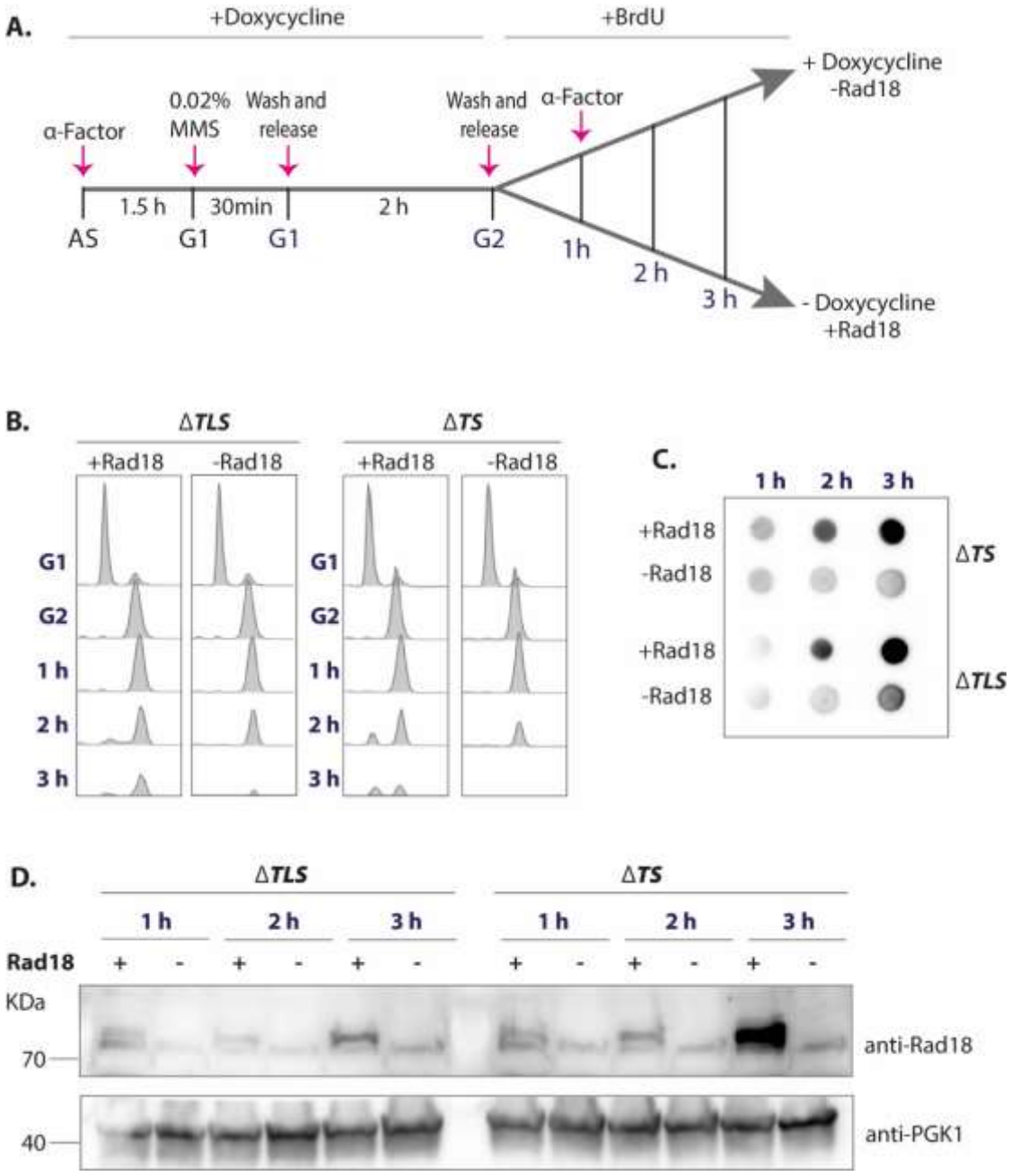

Figure 5.3. Tet-Rad18 system in *TLSΔ* and *TSΔ* strains

A. Cells grown in YPD supplemented with 2 µg/mL doxycycline were synchronised in G1-phase by adding 5 µg/mL α-factor (1.5 hours), then cells were treated with 0.02% MMS (30 min). Cells were wash out and released into S-phase (2 hours). For re-expression of Rad18, doxycycline was washed out twice and cells were resuspended in YPD containing 0.4 mg/mL BrdU. α-factor was added 1 hour after release to stop cells going to next S-phase. Samples after 1, 2 and 3 hours with and without Rad18 re-expression were collected for further analysis with BrdU-ssSeq. B. Analysis of cell cycle progression. C. Dot blot for detection of BrdU incorporation. D. Western blotting indicating Rad18 re-expression. Representative figure of three repetitions. *Strains ID: 2224 and 2226.*

I decided to examine if *TLSΔ* and *TSΔ* strains showed PRR-tract enrichment biases towards early origins of replication, thus I displayed PRR tracts according to replication timing in four quartiles. In order to do this, the Bioinformatics Core Facility at IMB first created sequencing depth normalised coverage tracts for all data as well as for individual strands by using Deeptools (Ramirez, et al., 2016). Then the enrichment around ORIs was calculated using quantification by Deeptools and further processed using custom R scripts. The ORIs were separated according to their timing and divided into quartiles with the earliest replication ORIs represented in the first quartile and the latest replication ORIs in the fourth quartile.

Similar to Figure 5.2, I observed a high enrichment of PRR tracts at early replication origins. In addition, I detected PRR tracts at 2 h after Rad18 re-expression in TLS-deficient cells. In contrast PRR tracts were observed only after 3 h Rad18 re-expression in TS-deficient strains. These results not only confirmed that TS was the preferred pathway to bypass MMS-induced lesions, but also it suggested that there was a timely-contribution of TS and TLS (Figure 5.4).

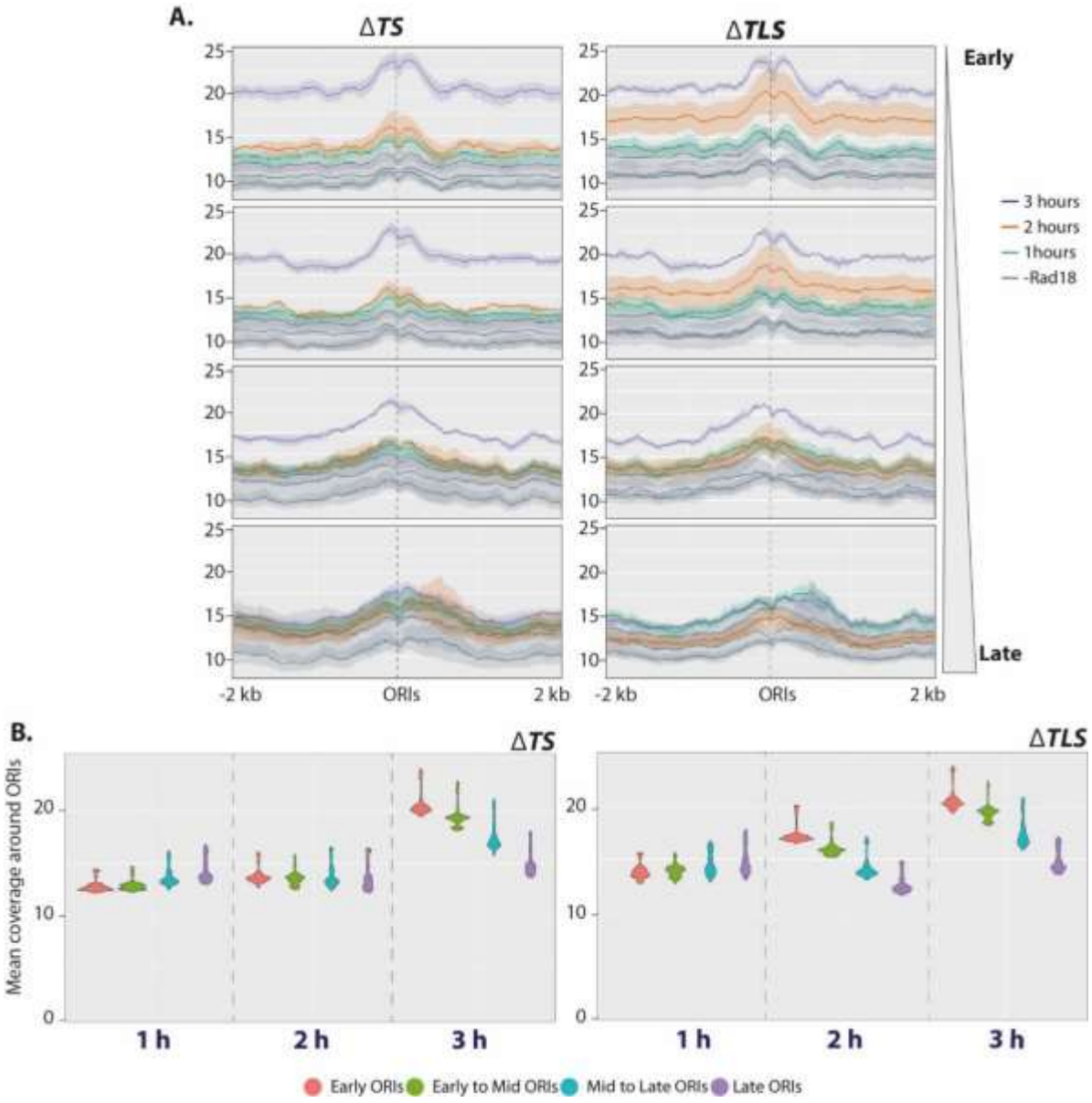

Figure 5.4. Quantification and visualisation of PRR tracts on *TLSΔ* and *TSΔ* strains

Genome-wide coverage of PRR tracts of samples with and without Rad18 re-expression obtained by BrdU-ssSeq. The experimental layout used during this assay is described in Figure 5.3A. **A.** Line plots representing the enrichment of PRR tracts around ORIs sorted according to replication timing in four quartiles (from early to late firing ORIs). **B.** Violin plots of the mean enrichment per timing quartile smoothed using a 5 bp window. The signal extended for 2 kb to the 3'- and 5'-sides from the ORIs. Coloured background indicates standard error calculated from three independent experiments. *Plots provided by Nastasja Kreim (Bioinformatics CF). Strains ID: 2224 and 2226. Project ID in Annexes.*

Then PRR tracts were then sorted strand-specifically to examine if there were biases of TLS and TS towards leading and lagging strands (Figure 5.5A). Remarkably, I showed that both TLS and TS had a greater contribution on leading strands than on lagging strands. This trend was seen mainly in regions within the earliest firing origins and prominently at 2 h after Rad18 re-expression on both *TLSΔ* and *TSΔ* strains (Figure 5.5B-C).

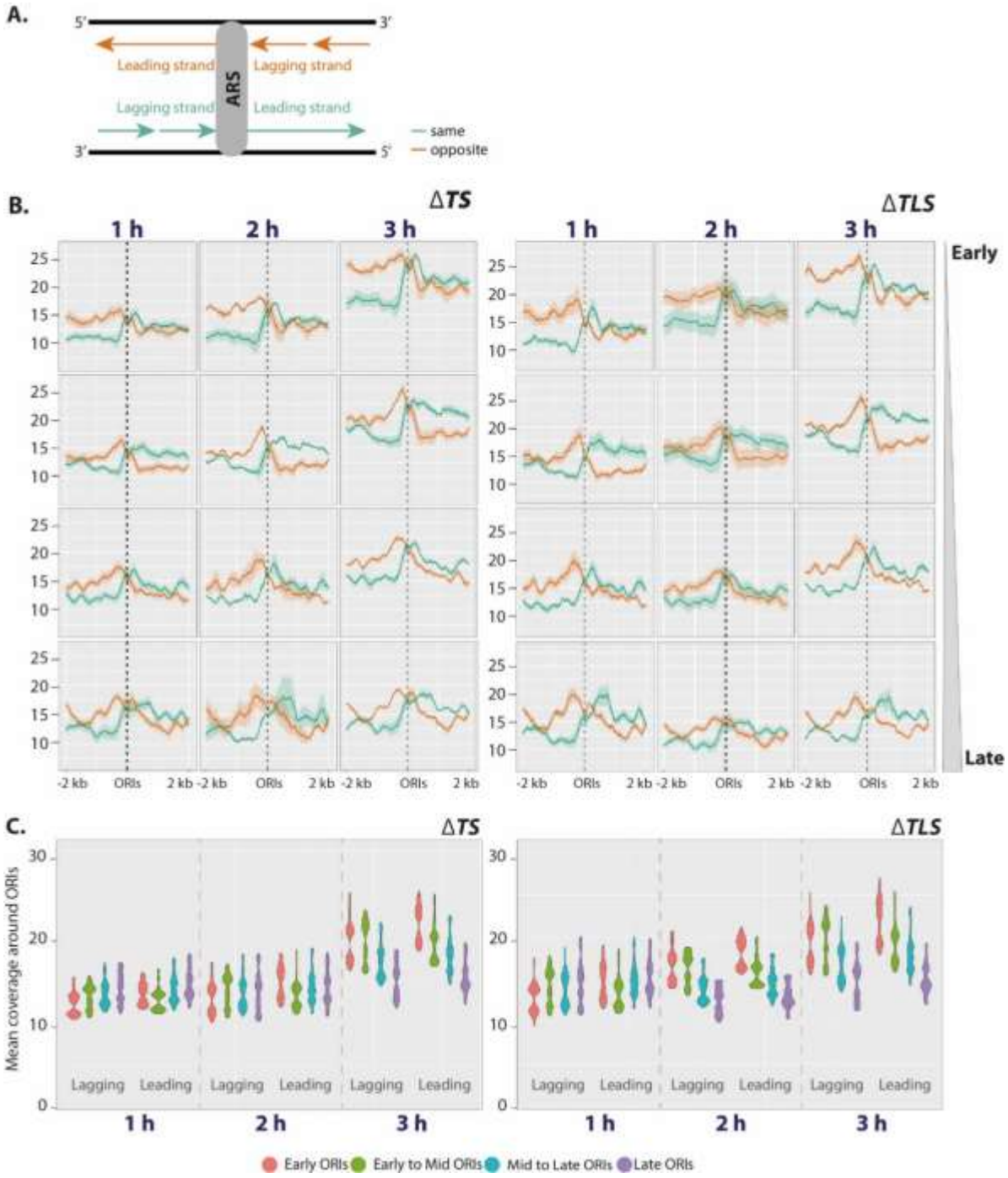

Figure 5.5. Strand-specific quantification and visualisation of PRR tracts on *TLSΔ* and *TSΔ* strains
Genome-wide coverage of strand-specific PRR tracts of samples with and without RAD18 re-expression obtained by BrdU-ssSeq on *TLSΔ* and *TSΔ* strains. The experimental layout used during this assay is described in Figure 5.3A. **A.** Schematic representation of DNA replication in the vicinity of replication origin, depicting the leading and lagging strands orientations. **A.** Line plots representing the enrichment of PRR tracts around ORIs sorted according to replication timing in four quartiles (from early to late firing ORIs). **B.** Violin plots of the mean enrichment per timing quartile smoothed using a 5 bp window. The signal extended for 2 kb to the 3'- and 5'-sides from the ORIs. Coloured background indicates standard error calculated from three independent experiments. *Plots provided by Nastasja Kreim (Bioinformatics CF). Strains ID: 2224 and 2226. Project ID in Annexes.*

5.5 Discussion

In this chapter, I showed the time-dependent contribution of template switching and translesion synthesis in the bypass of MMS lesions. Similar to previous studies, I confirmed that TS is the preferred pathway over TLS when using this alkylating agent. Interestingly, the genome-wide analysis of PRR tracts indicated the role of TS and TLS mostly in regions within early firing origins. Furthermore, I found that TS and TLS had a greater contribution on leading strands than on lagging strands.

5.5.1 The type of lesion determines the choice between TLS and TS

Even though it has been seen that lesions provoked by a certain genotoxic agent, predominantly used TLS or TS to bypass damage, it does not mean that a certain agent defines the choice of a PRR pathway. In fact, the choice between TLS and TS is highly influenced by the nature of the DNA lesion. For instance, UV-induced lesions are bypassed predominantly via TLS (Daigaku, et al., 2010). However, it was also determined that the UV-irradiation photoproduct, thymine-thymine pyrimidine (6-4) pyrimidone is largely bypassed (~95%) by TS (Masłowska, et al., 2019). Similarly, in this chapter, I found that MMS lesions were more efficiently bypassed by TS (Figure 5.4), however this does not exclude the contribution of TLS. As briefly mentioned in Section 1.1.6.1, MMS can cause a broad spectrum of DNA lesions, and some could be more prone to be repaired by TLS.

5.5.2 Evidence of cross-talk between factors involved in TLS and TS

Interestingly, it has been proposed the link between TLS polymerases and TS/HR repair. McIlwraith and colleagues (2005), by using a combination of fractionated human cell extracts and purified polymerases, uncovered an unexpected property for Pol η in promoting primer extension of an invading strand during TS/HR. They suggested that after replication-fork collapse and Rad51-mediated strand invasion, human Pol η might extend DNA synthesis at the sites of stalled or broken replication forks and like TLS, interactions with PCNA were likely involved. It would be interest to investigate the cross-talk between TLS and TS *in vivo*.

5.5.3 Kinetics of TLS and TS

The stalling of DNA polymerases and RPA binding to the ssDNA triggers PCNA ubiquitylation upon DNA damage (Davies, et al., 2008). While different types of lesions are able to prompt these effects, it is unclear whether lesions induce different ratios of ubiquitylation that determine the choice between TLS and TS, or that PCNA ubiquitylation is the permissive but not the deciding factor for TLS or TS.

Che and colleagues (2021) speculated that the availability and kinetics of TLS polymerases are key determinants for the choice of a tolerance pathway. It was seen that TLS polymerases are capable of inserting bases opposite different DNA lesions with different

efficiencies and accuracies (Prakash, et al., 2005). Moreover, it is notable that different TLS polymerases are not temporally equally available *in vivo*. For instance, the abundance of Rev1 and Pol η changes throughout the cell cycle, peaking in G2/M phase (Waters & Walker, 2006; Plachta, et al., 2015). In fact, it was reported that different levels of regulation including cellular concentrations of TLS polymerases or certain post-translational modifications may also promote TLS (Sabbioneda, et al., 2007). Interestingly, Rev1-Pol ζ are able to slowly bypass TT(6-4) lesions, however these lesions are mostly tolerated by TS (Masłowska, et al., 2019). This observation led to hypothesise that the speed of TLS polymerases mediating damage could be a determinant for cells choosing the error-prone pathway. Overall, these findings indicate that TLS polymerase kinetics might influence the decision for damage bypass.

On the other hand, the kinetics of TS is likely determined by the expansion of daughter strand gaps by Exo1 and Pif1 (García-Rodríguez, et al., 2018A; García-Rodríguez, et al., 2018B) in addition to the binding of RPA and Rad51 to ssDNA, which should not be drastically altered with different types of DNA lesions.

During this project, we additionally suggested that other factors, such as chromatin state (heterochromatin and euchromatin) and chromosomic regions (centromeres) might determine the spatial distribution of PRR tracts (Section 5.3).

5.5.4 Bypassing damage in leading strands

By using a reconstitute replisome, Guilliam & Yeeles (2021) showed that yeast replisome is inherently tolerant to leading-strand thymine glycol and 8-oxoG lesions, with damage bypass happening through a replicase switch mode. They suggested that upon damage, Pol ε disengages from leading strands promoting a switch to Pol δ at the stalled strand. Pol δ endorses efficient lesion bypass independently of TLS polymerases (Guilliam & Yeeles, 2021). However, when studying the bypass of a single CPD located in the leading-strand, Guilliam & Yeeles (2020A) observed that Rad6-Rad18 complex bound to exposed RPA-coated ssDNA facilitated Pol η recruitment and TLS. Even though in the absence of Pol δ, Pol η promoted CPD bypass slowly and inefficiently.

Interestingly, I showed biases of PRR tracts towards leading strands (Figure 5.5), yet the explanation behind this in not clear. Similar findings were observed in *E. coli*, where sub-lethal doses of the damaging agent, 5-azacytidine, induced TS especially on leading strands (Laranjo, et al., 2019A). In fact, Laranjo, et al. (2017B) indicated that the Exo I alone is responsible for the leading strand bias in wild-type cells and that its interaction with single-stranded breaks is necessary for this effect. These findings support even further the importance of Exo1 in the expansion of ssDNA gaps in preparation for strand invasion (García-Rodríguez, et al., 2018A). In fact, it would be great to explore further if Exo1 promoted gap expansion in leadings strands in preparation for TS by using the Tet-Rad18 system.

Chapter 6: General conclusions and future perspectives

Throughout this project, I investigated different aspects of DNA replication stress and DNA-damage tolerance. Following, I summarised the findings of this study, I discussed possible experimental improvements to the AP site system. Finally, in line with replication stress and bypass tolerance, I outlined potential topics and strategies that would be interesting to explore further.

6.1 Synthesis of leading-and lagging strands in the presence of abasic sites

In Chapter 3 and 4, by taking advantage of the AP site system in combination with NGS tools, I suggested that the replicative polymerases, Pol ε and Pol δ, contribute to replication and respond to damage differently. Moreover, I implied that the response of the replisome to nucleotide-base lesions depends on whether damage is located on the lagging or leading strands.

My observations included:

- Pol δ contributed to the synthesis of damaged leading strands
- Leading and lagging strands remained un-replicated at the site of damage
- When attempts of lagging strand synthesis were detected, these occurred inefficiently and dependently on the amount of abasic sites
- Replication re-started downstream of the lesions, possibly indicating that *de-novo* re-priming events occurred in both leading and lagging strands

Interestingly, by studying the dynamics of replisome components at the site of damage, I found impair loading of the core replisome proteins, PCNA and RPA.

- The PCNA clamp was highly enriched at the site of damage, possibly as part of the damage repair response
- RPA highly coated the ssDNA gap left behind the replication fork, which gradually expanded overtime, possibly indicating that newly synthesised DNA was being resected

I also found that the AP site system generated high genomic instability leading to the formation of DSBs within the region of interest. This made it difficult to conclude on the actual response of the replisome to damage, and raised the need for a system able to induce abasic sites in a less acute manner to avoid further processing into DSBs.

6.2 Template switching is the preferred pathway to bypass MMS lesions

In Chapter 5, by using the Tet-Rad18 system, I showed that TS was the preferred pathway over TLS when using MMS as damaging agent. However, this observation did not discard the importance of TLS in certain MMS lesions. Interestingly, the genome-wide analysis of PRR tracts exhibited TS and TLS contribution biases towards regions within early firing origins. Furthermore, strand-specific analysis indicated that TS and TLS had greater contribution in leading strands than in lagging strands.

6.3 Controlling the amount of abasic sites formation

In the course of this project, I observed that the AP site system, besides inducing the formation of abasic sites, led to the formation of DSBs. Next, I suggest possible improvements to the AP site system.

Inducing damage with enzyme CDG: Previous studies have confirmed the cytotoxic effect of the enzyme UDG*, in which its expression enhanced the mutation frequencies to 150-fold in a wild-type strain (Auerbach, et al., 2005). Perhaps exchanging the UDG* for a less mutagenic enzyme could reduce the genomic instability in the AP site system background. For instance, the expression of CDG (cytosine DNA glycosylase) only enhances the mutation frequencies to 7-fold in a wild-type strain. However, by using CDG instead of thymines, cytosines will be removed from DNA to form AP sites. Moreover, CDG is not as efficient as UDG* in targeting dsDNA (Otterlei, et al., 2000).

dCas9 for specific targeting of damage: Because the expression of Gal-UDG* construct was found to form abasic sites up to a 10 kb distance from the TetO44 array (Elizabeth Colby, unpublished), replacing the TetR protein for a more site-specific targeting approach could limit the amount of damage. Similarly to House, et al. (2020) the catalytically inactive Cas9 (dCas9) with an appropriate guide RNA could be used to specifically target UDG* to chromosome VI for the formation of abasic sites.

Using doxycycline to stop TetR and TetO binding: The AP site system comprises the fusion of the AID*-degron system to the TetR construct. After 45 minutes auxin is added to the media, it is seen that the TetR construct is degraded by the proteasome. However, it is probable that the binding of TetR to the TetO array is not completely abolished. In fact, it has been seen that TetO/TetR interaction provokes a replication block, which potentially induces a local DNA damage repair response, with the recruitment of proteins involved in double-strand break repair (Beuzer, et al., 2014). Therefore, doxycycline could be used to stop the TetR to bind the operator sequence. In addition, treating cells with doxycycline for only 5 minutes is enough to achieve this.

Shorter lengths of TetO array: During the establishment of the AP site system 12, 44 and 109 TetO sequences were tested to optimise the proper length of TetO array. Elizabeth Colby found that the length of the array was a compromise between high specificity of targeting (longer array better) and focusing abasic sites to a precise location (shorter array better).

Therefore, if a TetO12 array was used, maybe less abasic sites could be formed and probably their spread to regions within the TetO array could also be reduced.

6.4 Cutting-edge tools to study replication stress and lesion bypass

Here, I described possible strategies and tools that could be used in combination with the AP site (improved version) and Tet-Rad18 system to further continue the study of lesion bypass.

Strand-specific damage targeting: Similarly to the *in vitro* system created by Yeeles, et al., (2015) and Taylor & Yeeles (2018) to study replication events in a damage scenario. It would be interesting to investigate the strand-specific response of the replisome to damage by targeting the formation of abasic sites to either leading or lagging strands. This is possible by using the genome-editing tool CRISPR/Cas9. Briefly mentioned above, to target UDG* it would be necessary to use a catalytically inactivated Cas9 (dCas9) construct. Moreover, to guide AP site formation to either leading or lagging, it would be necessary to use a complementary-guide RNA. Whinn, et al. (2019) described an approach that targets stand-specific damage already.

Single-molecule mapping: Population-based sequencing methods to study DNA replication are broadly insensitive to coincident events affecting the replisome. Single-molecule based methods overcome many of the limitations of genome-wide approaches. A great example of a single-molecule analysis is Replicon-Seq. This method combines the excision of full-length replicons by controlled nuclease cleavage at replication forks, which are then sequenced by long-read Nanopore. This approach maps the ends of nascent DNA molecules with nucleotide resolution and allows the precise localisation of replisomes along the genome (Claussin, et al., 2022).

Single-cell assays: Similar to Wong, et al. (2020) powerful single-cell analysis using fluorescence microscopy can allow the study of the involvement of candidate factors during lesion bypass.

Strand-specific protein association: Yu and colleagues (2014), by fusing BrdU-Seq and ChIP-Seq developed a unique method called eSPAN-Seq. This potent tool can discern whether a protein is enriched at the leading or lagging strand. By using this method, important protein interaction and modifications can be measured during damage bypass.

Using the approaches outlined above a more complete picture of how lesion bypass operates across the genome could be gathered including:

- It would gain important insights about the strand-specific spatiotemporal kinetics of proteins recruited to the site of damage during and post-replication
- Results would reflect important improvements in sensitivity and background discrimination

- Investigations could be extended to understand the mutual influences between chromatin and lesion bypass
- It could confirm the involvement of candidate factors in damage bypass

Appendix

7.1 Annexes

Table 7.1 Relevant data sets for each of the figures where genomic data is shown

ID	Project name	Biological replicates	Referred in	Data set
-	sfb_ulrich_2019_05_sriramachandran_GLOEseq_Pol	1	Not shown	-
1351	sfb_ulrich_2020_02_villagomz_Gloeseq_Puseq	3	Figure 3.8 Figure 3.9 Figure 3.10	Set 1 Set 1 Set 1 & 3
1663	sfb_ulrich_2021_03_villagomez_BrdU_timecourse	1	Figure 4.5 Figure 4.6 Figure 4.11	Set 1 Set 1 Set 1
1716	sfb_ulrich_2021_05_villagomez_BrdU_timecourse	1		
1665	sfb_ulrich_2021_02_villagomez_ChIP_timecourse	1	Figure 4.8 Figure 4.9	Set 1 Set 1
1717	sfb_ulrich_2021_04_villagomez_ChIP_timecourse	1		
1747	imb_ulrich_2021_10_villagomez_BrdU_PRR	3	Figure 5.4 Figure 5.5	All replicates were ploted together

7.2 Abbreviations

2D	Two-dimensional
3’-OH	3’- hydroxyl end
5'-dRP	5′-deoxyribose phosphate
5-FOA	5-fluoroorotic acid
6-4PP	Pyrimidine-pyrimidone (6-4)
9-1-1	Rad9, Hus1, Rad1
AID*	Auxin-inducible degron
Amp	Ampicillin
ARS	Autonomously replicating sequence
BER	Base excision repair
BIR	Break-induce repair
BPB	Bromophenol blue
BrdU	5-bromo-2-deoxyuridine
BSA	Bovine serum albumin
CDG	Cytosine DNA glycosylase
Cdk	Cyclin-dependent kinase
ChIP	Chromatin immunoprecipitation
CMG	Cdc45, Mcm2–7, and GINS
CPD	Cyclobutane-pyrimidine dimer
CUT&RUN	Cleavage under targets and release using nuclease
DDC	DNA damage checkpoint
DDR	DNA damage response
DDT	DNA-damage tolerance
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphates
Dox	Doxycycline
DRC	DNA replication checkpoint
DSB	Double-stranded break
dsDNA	Double-stranded DNA
DTT	Di-thio-threitol
dTTP	2′-deoxythymidine-5′-triphosphate

dUTP	2´-deoxyuridine, 5´-triphosphate
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-ligating enzyme
EDTA	Ethylene-diamine-tetraacetic acid
EdU	5-ethynyl-2-deoxyuridine
EM	Electron microscopy
emRibo-Seq	Embedded ribonucleotide sequencing
GFP	Green fluorescent protein
GG-NER	Global genomic nucleotide excision repair
GLOE-Seq	Genome-wide Ligation of 3′-OH Ends
h	Hours
HR	Homologous recombination
HT-DNA	Herring sperm DNA
HU	Hydroxyurea
HydEnd-Seq	Hydrolytic end sequencing
IP	Immunoprecipitation
Kan	Kanamycin
KOAc	Potassium acetate
LB	Luria broth
LiOAc	Lithium acetate
MCM	Minichromosome maintenance
Min	Minutes
MMS	Methyl methanesulfonate
Mnase	Micrococcal nuclease
$NaBH_4$	Sodium borohydride
NER	Nucleotide excision repair
NGS	Next-Generation Sequencing
NHEJ	Non-homologous end-joining
NLS	Nuclear localisation signal
OK-Seq	Okazaki fragments sequencing
PBS	Phosphate-buffered saline

PBS-Tween	Phosphate-buffered saline-Tween
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PoI	Protein of interest
Pol	Polymerase
Pol α	Polymerase α
Pol δ	Polymerase δ
Pol ε	Polymerase ε
pRad53	Rad53 phosphorylation
pre-RC	Pre-replicative complex
PRR	Post-replication repair
Pu-seq	Polymerase usage sequencing
qPCR	Quantitative PCR
RF	Replication fork
RFA	Replication factor A
RFA1	Replication factor A protein 1
RFC	Replication factor c
Ribose-Seq	Embedded ribonucleotides sequencing
RNAPII	RNA polymerase II
RNase H2	Ribonuclease H2
rNMP	Ribonucleoside monophosphate
Rnr4	Ribonucleotide reductase complex 4
RPA	Replication protein A
RPA1	Replication protein A protein 1
RTS1	Replication termination sequence 1
SC	Synthetic media
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate–poly-acrylamide gel electrophoresis
S	Seconds
Seq	Sequencing
SMC	Structural maintenance of chromosomes
SSB	Single-stranded break

SSC	Saline-sodium citrate
ssDNA	Single-stranded DNA
ssSeq	Strand-specific sequencing
STE	Sorbitol-tris-edta
TBE	Tris-borate-EDTA
TCA	Trichloroacetic acid
TC-NER	Transcription coupled nucleotide excision repair
TE	Tris-EDTA
TetO	Tetracycline-inducible promoter/operator
TetR	Tetracycline-inducible repressor
TLS	Translesion synthesis
Top1	Topoisomerase 1
TS	Template switching
Tus	Terminus utilization substance
UDG	Uracil DNA glycosylase UDG
UDG*/ UDG^{Y147A}	Mutated human uracil DNA glycosylase
XPV	Human *Xeroderma pigmentosum* variant
YP	Yeast peptone
YPD	Yeast peptone dextrose
YP-Gal	Yeast peptone galactose
YP-Raff	Yeast peptone raffinose

7.3 Curriculum vitae

References

Admiraal, Eyler, Baldwin, Brines, Lohans, Schofield, & O'Brien. (2019). Expansion of base excision repair compensates for a lack of DNA repair by oxidative dealkylation in budding yeast. *J. Biol. Chem, 294*(37), 13629 –13637.

Aria, V., & Yeeles, J. (2019). Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. *Mol Cell., 73*(2), 199–211.

Auerbach, P., Bennett, R., Bailey, E., Krokan, H., & Demple, B. (2005). Mutagenic specificity of endogenously generated abasic sites in Saccharomyces cerevisiae chromosomal DNA. *PNAS, 102*(49), 17711–17716.

Ba, X., & Boldogh, I. (2018). 8-Oxoguanine DNA glycosylase 1: Beyond repair of the oxidatively modified base lesions. *Redox Biology, 14*, 669-678.

Bai, Yuan, Sun, Georgescu, O'Donnell, & Li. (2017). Architecture of the Saccharomyces cerevisiae Replisome. *Adv Exp Med Biol., 1042*, 207–228.

Balmus, G., Pilger, D., Coates, J., Demir, M., Sczaniecka-Clift, M., Barros, A., . . . Jackson, S. (2019). ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. *Nature Communications, 10*.

Bantele, S., Lisby, M., & Pfander, B. (2019). Quantitative sensing and signalling of single-stranded DNA during the DNA damage response. *Nature Communications, 10*(944).

Barberis, Spiesser, & Klipp. (2010). Replication Origins and Timing of Temporal Replication in Budding Yeast: How to Solve the Conundrum? *Curr Genomics, 11*(3), 199–211.

Bastia, D., & Zaman, S. (2014). Mechanism and physiological significance of programmed replication termination. *Semin Cell Dev Biol, 30*, 165-73.

Bell, & Labib. (2016). Chromosome Duplication in Saccharomyces cerevisiae. *Genetics, 203*(3), 1027–1067.

Bellí, G., Garí, E., Piedrafita, L., Aldea, M., & Herrero, E. (1998). An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Research, 26*(4), 942–947.

Beuzer, P., Quivy, J., & Almouzni, G. (2014). Establishment of a replication fork barrier following induction of DNA binding in mammalian cells. *Cell Cycle, 13*(10), 1607–1616.

Blaisdell, J., Harrison, L., & Wallace, S. (2001). Base Excision Repair Processing of Radiation-induced Clustered DNA Lesions. *Radiation Protection Dosimetry, 97*(1).

Blastyak, A., Pintér, L., Unk, I., Prakash, L., Prakash, S., & Haracska, L. (2007). Yeast Rad5 Protein Required for Postreplication Repair Has a DNA Helicase Activity Specific for Replication Fork Regression. *Molecular Cell, 28*, 167–175.

Boiteux, & Guillet. (2004). Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae. *DNA Repair, 3*(1), 1-12.

Boos, & Ferreira. (2019). Origin Firing Regulations to Control Genome Replication Timing. *Genes (Basel), 10*(3), 199.

Branzei, D., & Foiani, M. (2009). The checkpoint response to replication stress. *DNA repair, 8*(9), 1038-1046.

Branzei, D., & Szakal, B. (2016). DNA damage tolerance by recombination: Molecular pathways and DNA structures. *DNA Repair (Amst), 44*, 68–75.

Brown, J., & Suo, Z. (2011). Unlocking the Sugar "Steric Gate" of DNA Polymerases. *Biochemistry, 50*(7), 1135–1142.

Burgers, P., & Kunkel, T. (2017). Eukaryotic DNA Replication Fork. *Annu Rev Biochem, 86*, 417-438.

Cerritelli, & Crouch. (2016). The Balancing Act of Ribonucleotides in DNA. *Cell, 41*(5), 434-445.

Chang, D., & Cimprich, K. (2009). DNA damage tolerance: when it's OK to make mistakes. *Nature Chem. Bio., 5*(2), 82-90.

Chang, D., Lupardus, P., & Cimprich, K. (2006). Monoubiquitination of proliferating cell nuclear antigen induced by stalled replication requires uncoupling of DNA polymerase and mini-chromosome maintenance helicase activities. *J. Biol. Chem., 281*, 32081–32088.

Che, J., Hong, X., & Rao, H. (2021). PCNA Ubiquitylation: Instructive or Permissive to DNA Damage Tolerance Pathways? *Biomolecules, 11*(10), 1543.

Clausen, A., Lujan, S., Burkholder, A., Orebaugh, C., Williams,, J., Clausen, M., . . . Kunkel, T. (2015). Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nature Structural & Molecular Biology volume, 22*, 185–191.

Claussin, C., Vazquez, J., & Whitehouse, I. (2022). Single-molecule mapping of replisome progression. *Molecular Cell, 82*, 1372-1382.

Cortés-Ledesma, F., & Aguilera, A. (2006). Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. *EMBO Rep., 7*(9), 919–926.

Cortez, D. (2019). Replication-Coupled DNA Repair. *Molecular Cell, 74*(5), 866-876.

Costes, A., & Lambert, S. (2013). Homologous Recombination as a Replication Fork Escort: Fork-Protection and Recovery. *Biomolecules, 3*(1), 39-71.

Daigaku, Y., Davies, A., & Ulrich, H. (2010). Ubiquitin-dependent DNA damage bypass is separable from genome replication. *Nature, 465*(7300), 951–955.

Daigaku, Y., Keszthelyi, A., Müller, C., Miyabe, I., Brooks, T., Retkute, R., . . . Nieduszynsk, C. (2015). A global profile of replicative polymerase usage. *Nature Structural & Molecular Biology, 22*, 192–198.

Das, A., Tenenbaum, L., & Berkhout, B. (2016). Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther., 16*(3), 156–167.

Davies, A., Huttner, D., Daigaku, Y., Chen, S., & Ulrich, H. (2008). Activation of Ubiquitin-Dependent DNA Damage Bypass Is Mediated by Replication Protein A. *Mol Cell., 29*(5-2), 625–636.

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., . . . Malkova, A. (2011). Break-Induced Replication Is Highly Inaccurate. *Plos Biology, 9*(2), 1-14.

Ding, J., Taylor, M., Jackson, A., & Reijins, M. (2015). Genome-wide mapping of embedded ribonucleotides and other non-canonical nucleotides using emRiboSeq and EndoSeq. *Nat Protoc., 10*(9), 1433–1444.

Donnianni, R., Zhou, Z., Lujan, S., Al-Zain, A., Garcia, V., Glancy, E., . . . Symington, L. (2019). DNA Polymerase Delta Synthesizes Both Strands during Break-Induced Replication. *Molecular Cell, 76*, 371–381.

D'Souza, S., Waters, L., & Walker, G. (2009). Novel Conserved Motifs in Rev1 C-terminus are Required for Mutagenic DNA Damage Tolerance. *DNA Repair (Amst), 7*(9), 1455–1470.

Edenberg, E., Downey, M., & Tocyski, D. (2014). Polymerase Stalling during Replication, Transcription and Translation. *Current Biology, 24*.

Ekundayo, B., & Bleichert, F. (2019). Origins of DNA replication. *PLOS Genetics*, 12.

Elvers, Johansson, Groth, Erixon, & Helleday. (2011). UV stalled replication forks restart by re-priming in human fibroblasts. *Nucleic Acids Research, 39*(16), 7049–7057.

Essens, J., Theil, A., Baldeyon, C., van Cappellen, W., Houtsmuller, A., Kanaar, R., & Vermeulen, W. (2005). Nuclear Dynamics of PCNA in DNA Replication and Repair. *Mol Cell Biol., 25*(21), 9350–9359.

Eun Lee, S., Pellicioli, A., Vaze, M., Sugawara, N., Malkova, A., Foiani, M., & Haber, J. (2003). Yeast Rad52 and Rad51 Recombination Proteins Define a Second Pathway of DNA Damage Assessment in Response to a Single Double-Strand Break. *Mol Cell Biol., 23*(23), 8913–8923.

Eydmann, T., Sommariva, E., Inagawa, T., Mian, S., Klar, A., & Dalgaard, J. (2008). Rtf1-Mediated Eukaryotic Site-Specific Replication Termination. *Genetics, 180*, 27-39.

Federico, M., Siri, O., Calzetta, N., Paviolo, N., de la Vega, B., Martino, J., . . . Gottifredi, V. (2020). Unscheduled MRE11 activity triggers cell death but not chromosome instability in polymerase eta-depleted cells subjected to UV irradiation. *Oncogene, 39*, 3952–3964.

Gadaleta, & Noguchi. (2017). Regulation of DNA replication through natural impediments in the eukaryotic. *Genes, 8*, 98.

Gao, Y., Mutter-Rottmayer, E., Zlatanou, A., Vaziri, C., & Yang, Y. (2017). Mechanisms of Post-Replication DNA Repair. *Genes (Basel), 8*(2), 64.

Garbacz, Lujan, Burkholder, Cox, Wu, Haber, & Kunkel. (2018). Evidence that DNA polymerase δ contributes to initiating leading strand DNA replication in Saccharomyces cerevisiae. *Nat Commun., 9*, 1-11.

García-Rodríguez, N., Morawska, M., Wong, R., Daigaku, Y., & Ulrich, H. (2018A). Spatial separation between replisome- and template-induced replication stress signaling. *EMBO Journal, 37*.

García-Rodríguez, N., Wong, R., & Ulrich, H. (2018B). The helicase Pif1 functions in the template switching pathway of DNA damage bypass. *Nucleic Acids Res., 46*(16), 8347–8356.

Gelot, Magdalou, & Lopez. (2014). Replication Stress in Mammalian Cells and Its Consequences for Mitosis. *Genes, 6*(2), 267-298.

Guilliam. (2021). Mechanisms for Maintaining Eukaryotic Replisome Progression in the Presence of DNA Damage. *Front. Mol. Biosci., 8*.

Guilliam, T., & Yeeles, J. (2020). An updated perspective on the polymerase division of labor during eukaryotic DNA replication. *Critical Reviews in Biochemistry and Molecular Biology, 55*(5), 469-481.

Guilliam, T., & Yeeles, J. (2020). Reconstitution of Translesion Synthesis Reveals a Mechanism of Eukaryotic DNA Replication Restart. *Nat Struct Mol Biol., 27*(5), 450-460.

Guilliam, T., & Yeeles, J. (2021). The eukaryotic replisome tolerates leading-strand base damage by replicase switching. *EMBO J., 40*(5).

Haracska, L., Torres-Ramos, C., Johnson, R., Prakash , S., & Prakash , L. (2004). Opposing effects of ubiquitin conjugation and SUMO modification of PCNA on replicational bypass of DNA lesions in Saccharomyces cerevisiae. *Mol. Cell Biol., 24*, 4267–4274.

Haye-Bertolozzi, J., & Aparicio, O. (2017). Quantitative Bromodeoxyuridine Immunoprecipitation Analyzed by High-Throughput Sequencing (qBrdU-Seq or QBrdU). In M. Muzi-Falconi, & G. Brown, *Genome Instability* (pp. 209-225). New York: Humana Press.

Hedglin, M., Pandeya, B., & Benkovic, S. (2016). Stability of the human polymerase δ holoenzyme and its implications in lagging strand DNA synthesis. *Proc. Natl Acad. Sci., 113*, 1777–1786.

Heller, R., & Marians, K. (2006). Replication fork reactivation downstream of a blocked nascent leading strand. *Nature, 439*, 557–562.

Hoege, C., Pfander, B., Moldovan, G., Pyrowolakis, G., & Jentsch, S. (2002). RAD6 -Dependent DNA Repair Is Linked to Modification of PCNA by Ubiquitin and SUMO. *Nature, 419*, 135-141.

House, N., Parasuram, R., Layer, J., & Price, B. (2020). Site-specific targeting of a light activated dCas9-KillerRed fusion protein generates transient, localized regions of oxidative DNA damage. *PLoS One, 15*(12).

Iraqui, I., Chekkal, Y., Jmari, N., Pietrobon, V., Fréon, K., Costes, A., & Lambert, S. (2012). Recovery of Arrested Replication Forks by Homologous Recombination Is Error-Prone. *Plos Genetics, 8*(10), 1-15.

Jimeno, S., Camarillo, R., Mejias-Navarro, F., Fernandes-Avila, M., Soria-Bretones, I., Prados-Carvajal, R., & Huertas, P. (2018). The Helicase PIF1 Facilitates Resection over Sequences Prone to Forming G4 Structures. *Cell Reports, 24*, 3262–3273.

Johansson, E., & Dixon, N. (2013). Replicative DNA Polymerases. *Cold Spring Harb Perspect., 329*(1), 85-93.

Johnson, Klassen, Prakash, & Prakash. (2015). A major role of DNA polymerase δ in replication of both the leading and lagging DNA strands. *Mol Cell., 59*(2), 163–175.

Johnson, R., Washington, M., Haracska, L., Prakash, S., & Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature, 406*.

Karras, G., & Jentsch, S. (2010). Pathway Operates Uncoupled from the Replication Fork and Is Functional Beyond S Phase. *Cell, 141*(2), 255-267.

Kavli, B., Slupphaug, G., Mol, C., Arvai, A., Petersen, S., Tainer, J., & Krokan, H. (1996). Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO, 15*(13), 3442-3447.

Keszthelyi, A., Daigaku, Y., Ptasińska, K., Miyabe , I., & Carr, A. (2015). Mapping ribonucleotides in genomic DNA and exploring replication dynamics by polymerase usage sequencing (Pu-seq). *Nature Protocols, 10*, 1786–1801.

Kidane, D., Murphy, D., & Sweasy, J. (2014). Accumulation of abasic sites induces genomic instability in normal human gastric epithelial cells during Helicobacter pylori infection. *Oncogenesis, 3*(128).

Kim, & Jinks-Robertson. (2010). Abasic Sites in the Transcribed Strand of Yeast DNA Are Removed by Transcription-Coupled Nucleotide Excision Repair. *Mol Cell Biol., 30*(13), 3206–3215.

Kim, Huang, Williams, Li, Clark, Cho, . . . Jinks-Robertson. (2011). Mutagenic Processing of Ribonucleotides in DNA by Yeast Topoisomerase I. *Science, 332*(6037), 1561-1564.

Kitsera, Rodriguez-Alvarez, Emmert, Carell, & Khobta. (2019). Nucleotide excision repair of abasic DNA lesions. *Nucleic Acids Research, 47*(16), 8537–8547.

Knott, Viggiani, Tavaré, & Aparicio. (2009). Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae. *Genes Dev., 23*, 1077–1090.

Koh, K., Balachander, S., Hesselber, J., & Storici, F. (2015). Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nature Methods, 12*, 251–257.

Kurat, C., Yeeles, J., Patel, H., Early, A., & Diffley, J. (2016). Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Molecular Cell, 65*, 117-130.

Lambert, S., & Carr, A. (2005). Checkpoint responses to replication fork barriers. *Biochimie, 87*, 591–602.

Lambert, S., Watson, A., Sheedy, D., Martin, B., & Carr, A. (2005). Gross Chromosomal Rearrangements and Elevated Recombination at an Inducible Site-Specific Replication Fork Barrier. *Cell, 121*(5), 689-702.

Laranjo, L., Gross, S., Zeiger, D., & Lovett, S. (2017). SSB recruitment of Exonuclease I aborts template-switching in Escherichia coli. *DNA Repair (Amst), 57*, 12-16.

Laranjo, L., Klaric, J., Pearlman, L., & Lovett, S. (2019). Stimulation of Replication Template-Switching by DNA-Protein Crosslinks. *Genes (Basel), 10*(1), 14.

Larrea, A., Lujan, S., Nick McElhinny, S., & Kunkel, T. (2010). Genome-wide model for the normal eukaryotic DNA replication fork. *PNAS, 107*(41), 17674-17679.

Larsen, N., Hickson, I., & Mankouri, H. (2014). Tus-Ter as a tool to study site-specific DNA replication perturbation in eukaryotes. *Cell Cycle, 13*(19), 2994–2998.

Larsen, N., Sass, E., Suski, C., Mankouri , H., & Hickson, I. (2014). The Escherichia coli Tus–Ter replication fork barrier causes site-specific DNA replication perturbation in yeast. *Nature Communications*.

Lehmann, A., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J., . . . Green, C. (2007). Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA repair, 6*, 891–899.

Lehner, K., & Jinks-Robertson, S. (2014). Shared Genetic Pathways Contribute to the Tolerance of Endogenous and Low-Dose Exogenous DNA Damage in Yeast. *Genetics, 198*(2), 519–530.

Leman, & Noguchi. (2013). The Replication Fork: Understanding the Eukaryotic Replication Machinery and the Challenges to Genome Duplication. *Genes (Basel), 4*(1), 1-32.

Leman, A., & Noguchi, E. (2013). The Replication Fork: Understanding the Eukaryotic Replication Machinery and the Challenges to Genome Duplication. *Genes, 4*, 1-32.

Li, F., Want, Q., Seol, J., Che, J., Lu, X., Shim, E., . . . Niu, H. (2019). Apn2 resolves blocked 3′ ends and suppresses Top1-induced mutagenesis at genomic rNMP sites. *Nature Structural & Molecular Biology, 26*, 155–163.

Li, L., Murphy, K., Kanevets, U., & Reha-Krantz, L. (2005). Sensitivity to Phosphonoacetic Acid: A New Phenotype to Probe DNA Polymerase δ in Saccharomyces cerevisiae. *Genetics, 170*(2), 569–580.

Li, S. (2015). Transcription coupled nucleotide excision repair in the yeast Saccharomyces cerevisiae: The ambiguous role of Rad26. *DNA Repair, 36*, 43-48.

Li, S., Wang, H., Jehi, S., Li, J., Liu, S., Truong, L., . . . Wu, X. (2021). PIF1 helicase promotes break-induced replication in mammalian cells. *EMBO J., 40*(8).

Limoli, C., Giedzinski, E., Bonner, W., & Cleaver, J. (2002). UV-induced replication arrest in the xeroderma pigmentosum variant leads to DNA double-strand breaks, γ-H2AX formation, and Mre11 relocalization. *PNAS, 99*(1), 233-238.

Loeb, L. (1991). Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res., 51*, 3075–3079.

Loeb, L., & Monnat Jr , R. (2008). DNA polymerases and human disease. *Nature Reviews Genetics, 9*, 594–604.

Lopes, M., Foiani , M., & Sogo, J. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell, 21*, 15-27.

Lujan, S., Clausen, A., Clarck, A., MacAlpine, H., MacAlpine, D., Malc, E., . . . Kunkel, T. (2014). Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Research, 24*, 1751–1764.

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A., & Helleday, T. (2005). Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Research, 33*(12), 3799–3811.

Ma, Resnick, & Gordenin. (2008). Apn1 and Apn2 endonucleases prevent accumulation of repair-associated DNA breaks in budding yeast as revealed by direct chromosomal analysis. *Nucleic Acids Res., 36*(6), 1836–1846.

Mailand, Gibbs-Seymour, & Bekker-Jensen. (2013). Regulation of PCNA-protein interactions for genome stability. *Nat. Rev. Mol. Cell. Biol., 14*, 269–282.

Marteijn, L. V. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature Reviews Molecular Cell Biology, 15*, 465–481.

Masai, H., Matsumoto, S., You, Z., & Yoshizawa-Suga, N. (2010). Eukaryotic Chromosome DNA Replication: Where, When, and How? *Annual Review of Biochemistry, 79*, 89-130.

Masłowska, K., Laureti, L., & Pagès, V. (2019). iDamage: a method to integrate modified DNA into the yeast genome. *Nucleic Acids Res., 47*(20), 124.

Mazouzi , A., Velimezi, G., & Loizou, J. (2014). DNA replication stress: Causes, resolution and disease. *Experimental Cell Research*, 329 (1), 85-93.

Mazumder, A., Neamati, N., Pilon, A., Sunder, S., & Pommier, Y. (1996). Chemical trapping of ternary complexes of human immunodeficiency virus type 1 integrase, divalent metal, and DNA substrates containing an abasic site. Implications for the role of lysine 136 in DNA binding. *Journal of Biological Chemistry, 271*(44), 27330-27338.

McCulloch, S., Kokoska, R., Masutani, C., Iwai, S., Hanaoka, F., & Kunkel, T. (2004). Preferential cis–syn thymine dimer bypass by DNA polymerase η occurs with biased fidelity. *Nature, 428*, 97-100.

McGlynn, P., & Lloyd, R. (2002). Recombinational repair and restart of damaged replication forks. *Nature Reviews Molecular Cell Biology, 3*, 859–870.

McIlwraith, M., Vaisman, A., Liu, Y., Fanning, E., Woodgate, R., & West, S. (2005). Human DNA Polymerase η Promotes DNA Synthesis from Strand Invasion Intermediates of Homologous Recombination. *Molecular Cell, 20*(5), 783-792.

McInerney, P., & O'Donnell, M. (2004). Functional uncoupling of twin polymerases: mechanism of polymerase dissociation from a lagging-strand block. *J Biol Chem, 279*(20), 21543-51.

McIntosh, & Blow. (2012). Dormant origins, the licensing checkpoint, and the response to replicative stresses. *Cold Spring Harb. Perspect. Biol, 4*, 012955.

Memisoglu, & Samson. (2000). Base excision repair in yeast and mammals. *Elsevier, 451*(1-2), 39-51.

Mirkin, E., & Mirkin, S. (2007). Replication Fork Stalling at Natural Impediments. *Microbiol Mol Biol Rev, 71*(1).

Moldovan, G., Pfander, B., & Jentsch, S. (2007). PCNA, the Maestro of the Replication Fork. *Cell, 129*(4), 665-679.

Mundade, R., Ozer, H., Wei, H., Prabhu, L., & Lu, T. (2014). Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. *Cell Cycle, 13*(18), 2847-2852.

Naiman, K., Campillo-Funollet, E., Watson, A., Budden, A., Miyabe, I., & Carr, A. (2021). Replication dynamics of recombination-dependent replication forks. *Nature Communications, 12*.

Nakamura, J., Mutlu, E., Sharma, V., Collins, L., Bodnar, W., Yu, R., . . . Swenberg, J. (2014). The Endogenous Exposome. *DNA Repair (Amst), 19*, 3-13.

Neal, J., Dunger, K., Geith, K., & Meek, K. (2020). Deciphering the role of distinct DNA-PK phosphorylations at collapsed replication forks. *DNA Repair, 94*.

Nelson, J., Lawrence, C., & Hinkle, D. (1996). Thymine-thymine dimer bypass by yeast DNA polymerase ζ. *Science, 272*, 1646–1649.

Newlon, T. (1993). The structure and function of yeast ARS elements. *Curr. Opin. Genet. Dev., 3*, 752–758.

Nick McElhinny, S., Kumar, D., Clark, A., Watt, D., Watts, B., Lundström, E., . . . Kunkel, T. (2010A). Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol*(10), 774-81.

Nick McElhinny, S., Watts, B., Kumar, D., Watt, D., Lundströmb, E., Burgers, P., . . . Kunkel, T. (2010B). Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *PNAS, 107*(11), 4949–4954.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., & Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nature Methods, 6*, 917–922.

Noguchi. (2010). The DNA Replication Checkpoint and Preserving Genomic Integrity During DNA Synthesis. *Nature Education, 3*(9), 46.

Orebaugh, C., Lujan, S., Burkholder, A., Clausen, A., & Kunkel, T. (2018). Mapping Ribonucleotides Incorporated into DNA by Hydrolytic End-Sequencing. In M. B. Muzi-Falconi, *Genome Instability. Methods in Molecular Biology* (Vol. 1672, pp. 329–345). New York, NY.: Humana Press.

Otterlei, M., Kavli, B., Standal, R., Skjelbred, C., Bharati, S., & Krokan, H. (2000). Repair of chromosomal abasic sites in vivo involves at least three different repair pathways. *EMBO, 19*(20), 5542–5551.

Pacek, W. (2004). A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *EMBO J, 23*.

Pardo, B., Crabbé, L., & Pasero, P. (2017). Signaling pathways of replication stress in yeast. *FEMS Yeast Research, 17*(2).

Park, P. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews, 10*, 669-680.

Paulsen, & Cimprich. (2007). The ATR pathway: Fine-tuning the fork. *DNA repair, 6*(7), 953-966.

Pavlov, Y., Shcherbakova, P., & Kunkel, T. (2001). In Vivo Consequences of Putative Active Site Mutations in Yeast DNA Polymerases α, ε, δ and ζ. *Genetics, 159*, 47-64.

Pellicioli, A., Eun Lee, S., Lucca, C., Foiani, M., & Haber, J. (2001). Regulation of Saccharomyces Rad53 Checkpoint Kinase during Adaptation from DNA Damage–Induced G2/M Arrest. *Molecular Cell, 7*(2), 293-300.

Peng, B., Williams, T., Henry, M., Nielsen, L., & Vickers, C. (2015). Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microb Cell Fact., 14*.

Petermann, E., & Helleday, T. (2010). Pathways of mammalian replication fork restart. *Nature Reviews Molecular Cell Biology, 11*, 683–687.

Petryk, N., Kahli, M., d'Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., . . . Hyrien, O. (2016). Replication landscape of the human genome. *Nature Communications, 7*.

Plachta, M., Halas, A., McIntyre, J., & Sledziewska-Gojska, E. (2015). The steady-state level and stability of TLS polymerase eta are cell cycle dependent in the yeast S. cerevisiae. *DNA Repair (Amst), 29*, 147-153.

Prakash, S., & Prakash, L. (2000). Nucleotide excision repair in yeast. *Mutation Research, 451*, 13-24.

Prakash, S., Johnson, R., & Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annual Review of Biochemistry, 74*, 317-353.

Prioleau, M., & MacAlpine, D. (2016). DNA replication origins—where do we begin? *Genes & Dev.*, 1683-1697.

Quinet, A., Tirman, S., Cybulla, E., Meroni, A., & Vindigni, A. (2021). To skip or not to skip: choosing repriming to tolerate DNA damage. *Mol Cell., 81*(4), 649–658.

Ramirez, F., Devon, R., Björn , G., Vivek, B., Richter, A., Heyne, S., . . . Manke, T. (2016). deepTools2: A Next Generation Web Server for Deep-Sequencing Data Analysis. *Nucleic Acids Research, 44*(1), 160-165.

Reijns, M., Kemp, H., Ding, J., Marion de Procé, S., Jackson, A., & Taylor, M. (2015). Lagging-strand replication shapes the mutational landscape of the genome. *Nature, 518*, 502–506.

Reusswig, & Pfander. (2019). Control of Eukaryotic DNA Replication Initiation—Mechanisms to Ensure Smooth Transitions. *Genes (Basel), 10*(2), 99.

Ricceri, Godschalk, Peluso, Phillips, Agudo, Georgiadis, . . . Vineis. (2010). Bulky DNA adducts in white blood cells: a pooled analysis of 3600 subjects. *Cancer Epidemiol Biomarkers Prev., 19*(12), 3174–3181.

S. Prakash, L. P. (2000). Nucleotide excision repair in yeast. *Mutat Res, 30*(45), 13-24.

Saada, A., Teixeira-Silva, A., Iraqui, I., Costes, A., Hardy, J., Paoletti, G., . . . Lambert, S. (2017). Unprotected Replication Forks Are Converted into Mitotic Sister Chromatid Bridges. *Mol Cell., 66*(3), 398-410.

Sabbioneda, S., Bortolomai, I., Giannattasio, M., Plevani, P., & Muzi-Falconi, M. (2007). Yeast Rev1 is cell cycle regulated, phosphorylated in response to DNA damage and its binding to chromosomes is dependent upon MEC1. *DNA Repair, 6*, 121–127.

Sale, J. (2013). Translesion DNA Synthesis and Mutagenesis in Eukaryotes. *Cold Spring Harb Perspect Biol., 5*(3).

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

Shachar, S., Ziv , O., Avkin , S., Adar , S., Wittschieben , J., Reissner , T., . . . Livneh, Z. (2009). Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J., 28*(4), 383–393.

Skene, P., & Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, 1-35.

Sogo, J., Lopes, M., & Foiani, M. (2002). Fork Reversal and ssDNA Accumulation at Stalled Replication Forks Owing to Checkpoint Defects. *Science, 297*(5581), 599-602.

Solis-Escalante, D., Kuijpers, N., Bongaerts, N., Bolat, I., Bosman, L., Pronk, J., . . . Daran-Lapujade, P. (2013). amdSYM, a new dominant recyclable marker cassette for Saccharomyces cerevisiae. *FEMS Yeast Res, 13*(1), 126-139.

Sonoda, E., Okada, T., Zhao, G., Tateishi, S., Araki, K., Yamaizumi, M., . . . Takeda, S. (2003). Multiple roles of Rev3, the catalytic subunit of polζ in maintaining genome stability in vertebrates. *EMBO J., 22*(12), 3188–3197.

Sriramachandran, A., Petrosino, G., Méndez-Lago, M., Schäfer, A., Batista-Nascimento, L., Zilio, N., & Ulrich, H. (2020). Genome-wide Nucleotide-Resolution Mapping of DNA Replication Patterns, Single-Strand Breaks, and Lesions by GLOE-Seq. *Mol Cell., 78*(5), 975–985.

Stelter, P., & Ulrich, H. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature, 425*, 188–191.

Strauss, P., Beard, W., Patterson, T., & Wilson, S. (1997). Substrate binding by human apurinic/apyrimidinic endonuclease indicates a Briggs-Haldane mechanism. *J. Biol. Chem., 272*, 1302-7.

Strzalka, W., & Ziemienowicz, A. (2011). Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation. *Ann Bot., 107*(7), 1127–1140.

Suzuki, T., Katayama, Y., Komatsu, Y., & Kamiya, H. (2019). Large deletions and untargeted substitutions induced by abasic site analog on leading versus lagging strand templates in human cells. *Mutagenesis, 34*(5-6), 421–429.

Swanson, R., Morey, N., Doetsch, P., & Jinks-Robertson, S. (1999). Overlapping Specificities of Base Excision Repair, Nucleotide Excision Repair, Recombination, and Translesion Synthesis Pathways for DNA Base Damage in Saccharomyces cerevisiae. *Molecular and Cellular biology, 19*(4), 2929-2935.

Taylor, M., & Yeeles, J. (2019). Dynamics of Replication Fork Progression Following Helicase–Polymerase Uncoupling in Eukaryotes. *J Mol Biol., 431*(10), 2040–2049.

Taylor, M., & Yeeles, J. (2018). The Initial Response of a Eukaryotic Replisome to DNA Damage. *Molecular Cell, 70*(6), 1067-1080.

Thompson, & Cortez. (2020). New Insights into Abasic Site Repair and Tolerance. *DNA Repair (Amst)*.

Toledo, L., Altmeyer, M., Rask, M., Lukas, C., Larsen, D., Povlsen, L., . . . Lukas, J. (2013). ATR Prohibits Replication Catastrophe by Preventing Global Exhaustion of RPA. *Cell, 155*, 1088–1103.

Torres-Ramos, C., Johnson, R., Prakash, L., & Prakash, S. (2000). Evidence for the Involvement of Nucleotide Excision Repair in the Removal of Abasic Sites in Yeast. *Mol Cell Biol., 20*(10), 3522–3528.

Tourrière, H., Versini, G., Cordón-Preciado, V., Alabert, C., & Pasero, P. (2005). Mrc1 and Tof1 Promote Replication Fork Progression and Recovery Independently of Rad53. *Molecular Cell, 19*(5), 699-706.

Treangen, T., & Salzberg, S. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet., 13*(1), 36-46.

Tripodi, F., Zinzalla, V., Vanoni, M., Alberghina, L., & Coccetti, P. (2007). In CK2 inactivated cells the cyclin dependent kinase inhibitor Sic1 is involved in cell-cycle arrest before the onset of S phase. *Biochemical and Biophysical Research Communications, 359*(4), 921–7.

Tsaponina, O., & Chabes, A. (2013). Pre-activation of the genome integrity checkpoint increases DNA damage tolerance. *Nucleic Acids Research, 41*(22), 10371–10378.

Ulrich, H. (2011). Timing and spacing of ubiquitin-dependent DNA damage bypass. *FEBS Letters, 585*(18), 2861-2867.

Ulrich, H., & Jentsch, S. (2000). Two RING finger proteins mediate cooperation between ubiquitin-conjugating enzymes in DNA repair. *Embo. J., 19*, 3388–3397.

Uzunova, S., Zarkov, A., Ivanova, A., Stoynov, S., & Nedelcheva-Veleva, M. (2014). The subunits of the S-phase checkpoint complex Mrc1/Tof1/Csm3: dynamics and interdependence. *Cell Div., 9*(4).

Van, C., Yan, S., Michael, M., Waga, S., & Cimprich, K. (2010). Continued primer synthesis at stalled replication forks contributes to checkpoint activation. *J Cell Biol., 189*(2), 233–246.

Vaughn, & Sancar. (2020). Mechanisms and Maps of Nucleotide Excision Repair. *The Royal Society of Chemistry, 2*(15), 1-23.

Venkatesan, R., Hsu, J., Lawrence, N., Preston, B., & Loeb, L. (2006). Mutator Phenotypes Caused by Substitution at a Conserved Motif A Residue in Eukaryotic DNA Polymerase δ. *Journal of Biological Chemistry, 281*(7), 4486-4494.

Verma, R., Annan, R., Huddleston, M., Carr, S., Reynard, G., & Deshaies, R. (1997). Phosphorylation of Sic1p by G1 Cdk Required for Its Degradation and Entry into S Phase. *Science, 278*(5337), 455-460.

Viggiani, C., & Aparicio, O. (2006). New vectors for simplified construction of BrdU-Incorporating strains of Saccharomyces cerevisiae. *Yeast, 23*, 1045–1051.

Viggiani, C., Knott, S., & Aparicio, O. (2010). Genome-Wide Analysis of DNA Synthesis by BrdU Immunoprecipitation on Tiling Microarrays (BrdU-IP-chip) in Saccharomyces cerevisiae. *Cold Spring Harb Protoc., 2010*(2), 1-10.

Volker, M., Mone, M., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., . . . Mullenders, L. (2001). Sequential Assembly of the Nucleotide Excision Repair Factors In Vivo. *Molecular Cell, 8*, 213-224.

Wang, Vujcic, & Kowalski. (2001). DNA Replication Forks Pause at Silent Origins near the HML Locus in Budding Yeast. *Mol Cell Biol., 21*(15), 4938–4948.

Waters, L., & Walker, G. (2006). The critical mutagenic translesion DNA polymerase Rev1 is highly expressed during G2/M phase rather than S phase. *Proc Natl Acad Sci U S A., 103*(24), 8971–8976.

Whinn, K., Kaur, G., Lewis, J., Schauer, G., Mueller, S., Jergic, S., . . . Ghodke, H. (2019). Nuclease dead Cas9 is a programmable roadblock for DNA replication. *Scientific Reports volume, 9*.

Wilhelm, Said, & Naim. (2020). DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. *Genes, 11*, 642.

Williams, Lujan, & Kunkel. (2016). Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat. Rev. Mol. Cell Biol, 17*, 350–363.

Willis, N., Chandramouly, G., Huang , B., Kwok, A., Follonier, C., & Deng, C. (2014). BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature, 510*, 556–559.

Wong, R., García-Rodríguez, N., Zilio, N., Hanulová, M., & Ulrich, H. (2020). Processing of DNA Polymerase-Blocking Lesions during Genome Replication Is Spatially and Temporally Segregated from Replication Forks. *Molecular Cell, 77*(1), 3-16.

Wong, R., Petriukov, K., & Ulrich, H. (2021). Daughter-strand gaps in DNA replication – substrates of lesion processing and initiators of distress signalling. *DNA Repair, 105*.

Xiao, W., Broomfield, S., & Hanna, M. (2000). The Saccharomyces cerevisiae RAD6 group is composed of an error-prone and two error-free postreplication repair pathways. *Genetics, 155*(4), 1633–1641.

Xu, Aparicio, Aparicio, & Tavare. (2006). Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in S. cerevisiae. *BMC Genomics*, 276.

Xu, X., Blackwell, S., Lin, A., Li, F., Qin, Z., & Xiao, W. (2015). Error-free DNA-damage tolerance in Saccharomyces cerevisiae. *Mutation Research, 764*, 43-50.

Yeeles, Deegan, Janska, Early, & Diffley. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature, 519*, 431–435.

Yeeles, J., Poli, J., Marians, K., & Pasero, P. (2013). Rescuing Stalled or Damaged Replication Forks. *Cold Spring Harb Perspect Biol., 5*(5).

Yeeles, Janska, Early, & Diffley. (2017). How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol Cell., 65*(1), 105-116.

Yu, C., Gan, H., Han, J., Zhou, Z., Jia, S., Chabes, A., . . . Zhang, Z. (2014). Strand-Specific Analysis Shows Protein Binding at Replication Forks and PCNA Unloading from Lagging Strands when Forks Stall. *Mol Cell., 56*(4), 551–563.

Zeman, & Cimprich. (2014). Causes and Consequences of Replication Stress. *Nat Cell Biol., 16*(1), 2-9.

Zhou, Z., Lujan, S., Burkholder, A., Garbacz, M., & Kunkel, T. (2019). Roles for DNA polymerase δ in initiating and terminating leading strand DNA replication. *Nature Communications, 10*.

Zilio, N., & Ulrich, H. (2021). Exploring the SSBreakome: genome-wide mapping of DNAsingle-strand breaks by next-generation sequencing. *The FEBS Journal, 288*, 3948–3961.

Ziv, O., Geacintov, N., Nakajima, S., Yasui, A., & Livneh, Z. (2009). DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients. *Proc Natl Acad Sci, 106*, 11552–11557.