Transcription as source of genetic heterogeneity in budding yeast

Jonathan Houseley[1] and Baptiste Piguet[1]

[1]Babraham Institute

October 10, 2023

Abstract

Transcription presents challenges to genome stability both directly, by altering genome topology and exposing single-stranded DNA to chemical insults and nucleases, and indirectly by introducing potential obstacles to the DNA replication machinery. Such obstacles include the RNA polymerase holoenzyme itself, DNA bound regulatory factors, G-quadruplexes and RNA::DNA hybrid structures known as R-loops. Here we review the detrimental impacts of transcription on genome stability in budding yeast, as well as the mitigating effects of transcription-coupled DNA repair and of systems that maintain DNA replication fork processivity and integrity. We conclude that the impacts of transcription on DNA replication by the replisome must be very mild with little direct mutagenic outcome under normal conditions. However, transcription can significantly impair the fidelity of replication fork rescue mechanisms, particularly Break Induced Replication (BIR), which is used to restart collapsed replication forks when other means fail. This leads to *de novo* mutations, structural variation and extrachromosomal circular DNA formation that contribute to genetic heterogeneity. By re-analysing published datasets, we show that different classes of genes have different interactions with the replisome, and that highly transcribed environment dependant genes in *S. cerevisiae* tend to be located close to replication origins. We have previously implicated interactions between replication origins and BIR forks in adaptive transcription-induced copy number variation events, which indicates that environment dependent genes are preferentially located in genomic environments more prone to instability, particularly under replication stress.

Transcription as source of genetic heterogeneity in budding yeast

Baptiste Piguet and Jonathan Houseley*

Epigenetics Programme, Babraham Institute, Babraham, CB22 3AT

*Corresponding author: jon.houseley@babraham.ac.uk

ORCIDs: 0000-0001-7077-2266 and 0000-0001-8509-1500

Abstract

Transcription presents challenges to genome stability both directly, by altering genome topology and exposing single-stranded DNA to chemical insults and nucleases, and indirectly by introducing potential obstacles to the DNA replication machinery. Such obstacles include the RNA polymerase holoenzyme itself, DNA bound regulatory factors, G-quadruplexes and RNA::DNA hybrid structures known as R-loops. Here we review the detrimental impacts of transcription on genome stability in budding yeast, as well as the mitigating effects of transcription-coupled DNA repair and of systems that maintain DNA replication fork processivity and integrity. We conclude that the impacts of transcription on DNA replication by the replisome must be very mild with little direct mutagenic outcome under normal conditions. However, transcription can significantly impair the fidelity of replication fork rescue mechanisms, particularly Break Induced Replication (BIR), which is used to restart collapsed replication forks when other means fail. This leads to *de novo* mutations, structural variation and extrachromosomal circular DNA formation that contribute to genetic heterogeneity. By re-analysing published datasets, we show that different classes of genes have different interactions with the replisome, and that highly transcribed environment dependant genes in *S. cerevisiae* tend to be located

close to replication origins. We have previously implicated interactions between replication origins and BIR forks in adaptive transcription-induced copy number variation events, which indicates that environment dependent genes are preferentially located in genomic environments more prone to instability, particularly under replication stress.

Keywords: adaptation; mutagenesis; transcription-replication conflicts; DNA repair; R-loops; yeast; break-induced replication

Main text

At first sight, RNA polymerase II has evolved to function in a rather curious way with respect to genome integrity. Intuitively, we might expect the holoenzyme once loaded to simply follow the helical turns of the template strand, displacing histones and creating a single stranded bubble in a very localised region but causing minimal disruption to surrounding genome topology. This is not how RNA polymerase II operates: it passages DNA through the active site, creating positive supercoils ahead of the holoenzyme and negative supercoils behind that must be either resolved by topoisomerases or allowed to relax when the polymerase dissociates (Tsao et al., 1989). Quite how RNA polymerase achieves this remains unclear – even the bacterial RNA polymerase II working on a naked DNA can induce supercoiling (Janissen et al., 2023) – and although this behaviour facilitates the release of the nascent RNA (which would otherwise become wound around the DNA) it also creates a set of vulnerabilities for mutation:

1. Transcription requires the DNA helix to be transiently melted, and the single-stranded DNA (ssDNA) formed is more sensitive to chemical insults than dsDNA.
2. Supercoiling around the polymerase must be relaxed by topoisomerases, which can be mutagenic when topoisomerase action is not completed.
3. The transcription regulatory machinery and RNA polymerases must be removed to allow passage of replication forks.
4. Negative supercoiling behind the polymerase enhances the formation of non-B DNA structures, such as G-quadruplexes and RNA:DNA hybrids called R-loops, which are considered to be a threat to genome stability.

Many of these vulnerabilities arise from interactions between DNA replication and direct or indirect outcomes of transcription. The replisome moves extremely rapidly and must traverse transcriptional units almost without pause to duplicate chromosomes sufficiently fast not to become the rate limiting step of cell division. Here we review the mechanisms by which transcription may induce DNA damage in the budding yeast *S. cerevisiae* , focusing particularly on interactions with DNA replication, and assess whether genome structure has evolved to enhance genetic heterogeneity at particular times and places by increasing vulnerability to transcription-induced genome instability.

Transcription provides mutational vulnerabilities and opportunities for repair

Although transcription generates torsional stress, yeast RNA polymerases are less sensitive to this than we might expect: even cells lacking both topoisomerase I (Top1) and topoisomerase II (Top2) show defective RNA polymerase I transcription but surprisingly minor effects on RNA polymerase II and RNA polymerase III activity (Brill et al., 1987; Brill and Sternglanz, 1988). However, transcription generates significant supercoiling which is a Top1 substrate, and Top1 errors in the form of short deletions can form up to 50% of mutations in highly transcribed sequences (Lippert et al., 2011; Takahashi et al., 2011). These arise both through mis-processing of trapped Top1 cleavage complexes in tandem repeats and through Top1 cleaving at ribonucleotides (Cho et al., 2013). Ribonucleotides are mistakenly incorporated in the genome during replication but usually removed by RNase H2 (Huang et al., 2015; Williams and Kunkel, 2014; Williams et al., 2013), however processing of ribonucleotides by Top1 results in a non-ligatable end subject to potentially error prone repair (Huang et al., 2015; Kim et al., 2011). Top1-mediated mutations are therefore a significant danger to the integrity of coding sequences, though surprisingly these mutations are biased to the non-template strand (Cho et al., 2015), and Top1 overexpression in yeast has been associated with mutagenesis and increased DNA damage (Sloan et al., 2017).

The passage of RNA polymerase II along the template strand creates a bubble of single stranded DNA (ssDNA) that is also likely to increase mutation on the non-template strand as ssDNA is innately more prone to endogenous chemical damage. For instance, spontaneous depurination and depyrimidination occur four times more frequently in ssDNA than in dsDNA (Billen, 1990), cytosine deamination to uracil forms over 100-fold faster (Lindahl, 1993), and base alkylation occurs more frequently in ssDNA (Fu et al., 2012); indeed recent genome-wide studies have confirmed that the non-template strand is more frequently altered by alkylating agents and through spontaneous cytosine deamination (Mao et al., 2017; Williams et al., 2023).

In contrast, the genetic stability of the template strand is enhanced by transcription due to efficient machinery for repairing DNA lesions that are encountered by RNA polymerase II. Many lesions can block the progression of RNA polymerase including intra and inter-strand crosslinks, DNA-protein crosslinks, cyclopurines and abasic sites (Brooks et al., 2000; Fielden et al., 2018; Jung and Lippard, 2007; Tornaletti et al., 2006). When elongating RNA pol II encounters such impediments, the remodelling factor Rad26 binds to the polymerase and the DNA sequence upstream. The ATPase activity of Rad26 allows a forward translocation of RNA polymerase II, facilitating the bypassing of non-bulky lesions or benign obstacles (Duan et al., 2020; Xu et al., 2017), but when transcription-blocking lesions (TBLs) are encountered the holoenzyme can engage the transcription-coupled nucleotide excision repair (TC-NER) system. In yeast, TC-NER is activated by Rad26 or occasionally the non-essential Rpb9 RNA pol II subunit (reviewed in (Li, 2015)) , depending on expression level (Duan et al., 2020; Li and Smerdon, 2002, 2004). While the RNA polymerase traverses the block or, for severe lesions, is ubiquitinated and degraded, the cell initiates a repair process involving excision of nucleotides surrounding the lesion. Then, DNA polymerases including Polδ and Polε are recruited to fill in the gap and the newly synthesised DNA is ligated before the transcription restarts (reviewed in (Gregersen and Svejstrup, 2018)).

Although yeast lacking TC-NER components show little phenotype under ideal conditions, the long term importance of this system in maintaining transcriptional homeostasis and suppressing mutations is underlined by the phenotypes of human diseases caused by TC-NER mutations including UV-sensitive syndrome, Xeroderma Pigmentosum and Cockayne syndrome, which is usually caused by mutations in the human orthologue of Rad26 (Lans et al., 2019). Conversely, it has been speculated that one reason why eukaryotes including yeast undergo genome-wide pervasive transcription is to allow RNA polymerase to survey the genome for replisome blocking lesions for repair by TC-NER prior to DNA replication (Ljungman, 2022), and direct evidence exists for this in bacteria as well as for telomeres in budding yeast (Guintini et al., 2022; Martinez et al., 2022). Furthermore, a yeast mutant deficient in global NER but proficient for TC-NER shows increased UV resistance when spurious transcription initiation is licensed, demonstrating the capacity of pervasive transcription to improve genome stability (Selvam et al., 2022).

Transcription therefore has an innately strand biased effect on genome stability, increasing damage on the non-template strand particularly during mutagen exposure, but resolving lesions on the template strand. Transcription does engage repair enzymes that introduce errors and strand breaks which may be mutagenic if encountered by the replisome.

The potential for transcription to impair DNA replication

Transcription does not stop during DNA replication and transcription-replication conflicts must arise since both processes coexist on chromosomes. Furthermore, the transcriptional machinery includes many DNA-bound protein complexes that might impair the movement of the replisomes even if active RNA polymerases are removed.

Whereas bacterial genes tend to be ordered co-directionally with DNA replication (Kunst et al., 1997; McLean et al., 1998; Rocha and Danchin, 2003), in *S. cerevisiae* almost half of the genes are transcribed opposite to the normal direction of DNA replication (referred to as "head-on" genes), so encounters between the replisome and RNA polymerases must be frequent (Garcia-Muse and Aguilera, 2016; Goehring et al., 2023). Also unlike bacteria, there is little evidence that head-on orientated genes in yeast have higher mutation rates so the eukaryotic replisome must be adept at traversing transcriptional units (Sankar et al., 2016). In

theory, DNA damage systems could alleviate transcription-replication collisions in head-on genes, but the DNA damage response is global and delays replication throughout the genome when forks stall, which must represent an emergency response rather than the default behaviour (reviewed in (McClure et al., 2022)). Similarly, while genome-wide profiling of DNA replication has provided evidence that the replisome pauses at highly expressed loci, the delays required for the replisome to wait until transcription completes at each head-on gene would be substantial (often >1 minute given average gene size and RNA pol II holoenzyme speed (Muniz et al., 2021)), whereas minimal pausing was detected (Azvolinsky et al., 2009; Claussin et al., 2022; Kara et al., 2021). Replication profiles show that in S. cerevisiae many replisomes travel 50 kb through gene rich regions in a normal S-phase lasting ˜ 30 minutes (Kara et al., 2021), which is at the limit of travel given the measured yeast fork speed of 1.6-1.9 kb/min (Hodgson et al., 2007; Sekedat et al., 2010), so pauses must be short and infrequent. Therefore, the eukaryotic replisome must efficiently displace RNA polymerases, just as the bacterial replisome has been directly observed to do by electron microscopy (French, 1992).

Although we might imagine that head-on collisions between replication and transcription are inherently recombinogenic, extensive studies of the budding yeast ribosomal DNA show exactly the opposite. The unidirectional replication fork barrier (RFB) formed by Fob1 binds the ribosomal DNA downstream of the massively transcribed 35S region and prevents the replisome meeting oncoming RNA polymerase I head-on (Kobayashi and Horiuchi, 1996). However, replication forks stalled at the RFB are highly recombinogenic (Kobayashi et al., 1998; Stewart and Roeder, 1989), whereas in RFB-defective mutants replication forks proceed seemingly unimpeded head-on through the 35S region while recombination rate is vastly decreased and DNA damage marks are unchanged (Kara et al., 2021; Keszthelyi et al., 2023). This is also true for RNA polymerase II as reversion assays using a head-on or codirectional *lys2* frameshift allele failed to demonstrate any significant difference in the overall rate of reversion between both orientations (Kim et al., 2007). There are exceptions however, since tRNA genes transcribed by RNA polymerase III do impede the replisome in head-on encounters, as detected by both 2D-gels and genome-wide methods and these encounters can be recombinogenic (Azvolinsky et al., 2009; Claussin et al., 2022; Deshpande and Newlon, 1996; Kara et al., 2021; Tran et al., 2017).

Regulatory proteins bound to DNA could also form obstacles for the replisome. The substantial machinery bound at promoters includes transcription factors and transcriptional activators (catalogued in (Rossi et al., 2021)), but as noted above genome-wide studies have revealed only the mildest impacts of these on replisome progression (Azvolinsky et al., 2009; Kara et al., 2021). Helicases including Rrm3 and Pif1 work assist the replisome in removing proteins, easing the path through difficult features including tRNAs (Azvolinsky et al., 2006; Azvolinsky et al., 2009; Claussin et al., 2022; Ivessa et al., 2003; Tran et al., 2017), and it has recently been found that replisome pausing at tRNA is primarily due to bound TFIIIC complex rather than transcription*per se* (Yeung and Smith, 2020). However, the mild replisome pausing at highly expressed RNA polymerase II genes is independent of Rrm3 (Azvolinsky et al., 2009), suggesting that the core replisome is able to remove protein obstacles associated with RNA polymerase II activity.

RNA polymerase II and associated factors therefore have the potential to impede DNA replication, and there is some evidence for this in extremely highly transcribed genes, but during rapid growth under nutrient rich conditions the replisome is very proficient in removing such obstacles, despite transcription being highest under such conditions. RNA polymerase III genes do have an effect on the replisome, but *S. cerevisiae* has dedicated helicases to ensure that replisome pausing at tRNA is minimal.

R-loops as a threat to genome stability

Transcription can result in the formation of R-loops, in which the newly synthesised RNA hybridises with the DNA template strand prior to re-annealing of the transcription bubble to form a stable DNA-RNA hybrid structure (Figure 1A)(reviewed in (Santos-Pereira and Aguilera, 2015)) . R loops have physiological functions in transcription and in accurate chromosome segregation (Boque-Sastre et al., 2015; Kabeche et al., 2018), however, the pathological accumulation of R-loops has been repeatedly shown to increase genetic instability (reviewed in (Brickner et al., 2022; Huertas and Aguilera, 2003; Neil et al., 2018)). Genome-wide profiling

shows that R-loops form more readily following RNA polymerase II pausing near transcription starting sites or transcription termination sites (TSS or TTS, respectively) but also form stochastically, particularly with polymerase stalling or in regions of high GC content, and formation is facilitated by negative supercoiling accumulated behind RNA polymerase (Chan et al., 2014; El Hage et al., 2010; Skourti-Stathaki et al., 2014; Skourti-Stathaki et al., 2011; Wahba et al., 2016).

Budding yeast utilises various mechanisms to ameliorate the impact of R-loops on genome stability. Firstly, coating of nascent mRNA with processing and export factors including THO/TREX reduces R-loop formation and replication fork slowing (Garcia-Rubio et al., 2018; Gomez-Gonzalez et al., 2011; Luna et al., 2019; Wellinger et al., 2006), with THO/TREX mutants having unstable genomes (Prado et al., 1997; San Martin-Alonso et al., 2021; Selvam et al., 2014). Another RNA-binding protein, Npl3, which is involved in mRNA splicing and export in yeast, was shown to decrease R-loop dependent genomic instability and replication stress, and physical proximity of genes to nuclear pores which accelerates mRNA export was shown to reduce R-loop accumulation (Gaillard et al., 2018; Garcia-Benitez et al., 2017; Santos-Pereira et al., 2013). Secondly, helicases can unwind R-loops; the transcription termination factor Sen1 has a particular role in preventing R-loop formation during S phase, with Sen1 depletion causing accumulation of both R-loops and the gH2A histone marker which indicates the concomitant onset of DNA replication stress (Khurana and Oberdoerffer, 2015; Mischo et al., 2011; San Martin-Alonso et al., 2021). Thirdly, both RNase H1 and RNase H2 can degrade the RNA in R-loops. Mutants lacking RNase H2 have a much greater impact on R-loop mediated genome instability (O'Connell et al., 2015; Zimmer and Koshland, 2016), but RNase H1 has an 'on-demand' activity at the critical time in S phase (Lockhart et al., 2019). Fourthly, chromatin modifiers including Rtt109 and FACT counter the formation of R-loops (Canas et al., 2022; Herrera-Moyano et al., 2014).

This multi-layered defence against R-loop formation and persistence is not surprising given the genome instability that results when cells with increased R-loops undergo DNA replication. However, the mechanism by which R-loops cause genome instability remains unclear as replicative helicases can unwind RNA-DNA helices *in vitro* and replication forks can overcome DNA–RNA hybrids formed in co-directional transcription units (Garcia-Rubio et al., 2018; Hamperl et al., 2017; Shin and Kelman, 2006). Moreover, hyperrecombination phenotypes caused by R-loops in yeast depend on Histone H3S10 phosphorylation suggesting an indirect mechanism involving chromatin (Garcia-Pichardo et al., 2017). One possible explanation is that in head-on encounters the RNA polymerase becomes trapped between the R-loop and the replisome, particularly in the absence of Sen1; this could impair normal mechanisms for removing the RNA polymerase holoenzyme from DNA in front of the replisome and form a potent roadblock (Felipe-Abrio et al., 2015; Zardoni et al., 2021). Alternatively, as noted above ssDNA is more prone to base damage, which would therefore accumulate on the non-template strand in transcriptional R-loops. However, experiments to address this mechanism discovered lesions arising from the replication bubble but there was no dependence on R-loops (Williams et al., 2023). Finally, R-loops can be targeted by the TC-NER or global genome NER pathway, resulting in cleavage and removal of the RNA:DNA duplex followed by gap repair (Crossley et al., 2023; Sollier et al., 2014), which forms a short-lived ssDNA gap that could in theory be processed into a DSB by DNA replication.

Negative supercoiling behind RNA polymerase can also promote the formation of G-quadruplexes - non-B form four-stranded nucleic acid structures made between several guanosines, often stabilised by monovalent cations (Selvam et al., 2014). These structures can form throughout the *S. cerevisiae* genome and particularly at telomeres though they are not required for telomere function (Esnault et al., 2023; Skourti-Stathaki et al., 2014). Like R-loops, these structures have been implicated in replication fork stalling and leading to genome instability, and are removed by the helicase Pif1 (Lopes et al., 2011; Paeschke et al., 2011; Piazza et al., 2010).

Interruption of replisome progression by R-loops and G-quadruplexes has the potential to be recombinogenic, and may contribute to genetic heterogeneity even in unperturbed cells in a transcription-dependent manner. However, R-loops and G-quadruplexes are common – the fact that these species can be detected in genome wide assays even in wild-type yeast, indicates that at any given site they are present in a significant fraction

of cells. If each R-loop even delayed the replisome, let alone caused a time-consuming and mutagenic repair process, this would dramatically slow DNA replication irrespective of mutational burden. In reality, even RNase H1/H2 double mutants in which R loops are highly stabilised are viable and do not show a substantial growth defect (Zhao et al., 2018), suggesting that at most only a tiny fraction of R-loops actually perturb replisome progression.

Mechanisms for ensuring replisome processivity

The replisome includes separate polymerase machineries for the leading and lagging strands, and the replicative helicase complex (Cdc45-MCM-GINS or CMG) which removes obstacles and unwinds DNA in front of the polymerases, as well as numerous auxiliary and regulatory factors (reviewed in (Pellegrini, 2023)). In unperturbed replication, the polymerase machineries are connected to the CMG to ensure that the replisome travels as a unit and to allow regulation of fork progression rate (Gambus et al., 2006; Jones et al., 2021; Simon et al., 2014).

When the replisome encounters an obstacle, either the CMG complex is uncoupled and continues along the template while one or both of the polymerases pause, or the whole replisome pauses. The former outcome results in the characteristic signature of replication stress (Figure 1B) - an accumulation of single stranded DNA ahead of the replication fork - observed with treatments that cause base lesions (eg: MMS) or impair DNA polymerase activity (eg: HU). Persistent accumulation of ssDNA around the replication fork is recognised by the kinase Mec1 (the yeast homolog of ATR) via Ddc2 and the 9-1-1 complex, which then phosphorylates histone H2A at S129 in neighbouring chromatin, forming a focus of gH2A (Feng et al., 2006; Lopes et al., 2006; Namiki and Zou, 2006; Pardo et al., 2017; Puddu et al., 2011).

For base lesions such as 8-oxo-2'-deoxyguanosine (8-oxo-dG), the DNA polymerases in the replisome are replaced with trans-lesion synthesis enzymes Rad30 (Polη), Rev3/7 (Polζ) and/or Rev1, allowing DNA replication to continue past the lesion while (hopefully) inserting the correct base, with the lesion itself being left in place for NER or BER to cure post-replication (reviewed in (Powers and Washington, 2018; Zhao and Washington, 2017)) (Figure 1B). Alternatively, template switching can be used to bypass lesions. In this process, the stalled polymerase temporarily switches to using the other newly synthesised DNA strand as a template to bypass obstacles (reviewed in (Ripley et al., 2020)) (Figure 1B). Both strategies aim to restart efficient DNA synthesis on both strands at the replication fork, resulting in a functional replication fork that is no longer associated with the replicative helicase, an equivalent situation to a fork paused for example by a short pulse of HU (Petermann et al., 2010). In both cases, replication appears to resume as normal suggesting that these replication forks re-couple to the CMG helicase, probably facilitated by the slower progression of CMG that is uncoupled from the replicative polymerases (Graham et al., 2017; Petermann et al., 2010).

The CMG helicase can traverse almost all replisome obstacles, and can switch from single and double stranded binding modes when uncoupled and back to single stranded binding once re-coupled(Wasserman et al., 2019). However, programmed replication fork barriers can block CMG progression resulting in a stable fork arrest, personified by the RFB in *S. cerevisiae* ribosomal DNA and *RTS1* at the *S. pombe*mating-type locus (reviewed in (Labib and Hodgson, 2007)). This arrest is mediated not by a physical block to the replisome but by the Fork Protection Complex (FPC), which travels with the CMG complex (reviewed in (Shyian and Shore, 2021)). Similarly, the DNA replication checkpoint, which responds to global replication stress, can pause and stabilise replication forks without uncoupling also through the FPC (Noguchi et al., 2004; Pardo et al., 2017), though at least in human cells and likely in yeast the forks become progressively less stable with time (Petermann et al., 2010).

Only if bypass and restart mechanisms fail, and additionally the fork is not resolved by an oncoming replication fork from another origin, must further more potentially mutagenic options be explored. Once disconnected from the CMG helicase, a stalled fork gains the capability to reverse by re-annealing of the nascent leading and lagging strands to yield a 4-pronged structure akin to a Holliday Junction and capable of migration over significant distances, a remodelling process assisted by Rad5 in budding yeast (Blastyak et al.,

2007; Toth et al., 2022; Unk et al., 2010) (Figure 1C). Reversed replication forks gain a free double stranded end, which can be resected and initiate homologous recombination (Cotta-Ramusino et al., 2005; Lemacon et al., 2017). Homologous recombination can restart the replication process, albeit in a different form known as Break-Induced Replication (BIR) in which the leading strand is copied as a migrating D-loop, then the lagging strand copied from the newly synthesised leading strand (Kramara et al., 2018; Liu and Malkova, 2022; Malkova and Ira, 2013) (Figure 1C). Therefore, whereas in canonical DNA replication both newly formed DNA helices contain one strand from the original DNA template (known as 'semi conservative replication'), in BIR both strands of the daughter DNA helix are newly synthesised ('conservative replication'). BIR uses a different DNA polymerase complement and relies on other helicases for processivity (notably Pif1) (Kramara et al., 2018; Liu and Malkova, 2022; Lydeard et al., 2010); movement of this complex is slow and error prone when forced to proceed over long distances, but in almost all instances this problem is avoided as the BIR fork is resolved by a replication fork coming from the other direction (Liu et al., 2021; Mayle et al., 2015). Because BIR is innately slow and error prone, replication forks are actively protected from resection so that transient fork pausing can be overcome with a non-recombinational system that likely does not involve a loss of the replisome structure or function (Brambati et al., 2018). Single strand gaps are also converted to double strand breaks when encountered by the replisome and could enter a similar recombinational restart pathway (Vrtis et al., 2021). However, this must again be a last resort, and mammalian cells use PARP to guard against this outcome and it is speculated that yeast use the FPC, at least for Top1-mediated damage (Ray Chaudhuri et al., 2012; Westhorpe et al., 2020).

BIR forks have much more potential for introducing mutations than normal replication forks (Deem et al., 2011; Pardo and Aguilera, 2012; Sakofsky et al., 2014); the lack of the CMG complex makes these prone to further stalling and recombinational repair if unresolved (Liu et al., 2021; Smith et al., 2007). Furthermore, proper sister chromatid cohesion requires replisome factors including the FPC, meaning that recombinational repair during BIR is much more likely to result in non-allelic homologous recombination (van Schie and de Lange, 2021). Template switching based on microhomology is also more frequent, resulting in complex structural variations showing multiple switches with more or less homology at break points (Pardo and Aguilera, 2012; Sakofsky et al., 2015). These properties make the slow progression of BIR forks actively advantageous to the cell by providing the highest chance of resolution by a normal replication fork coming from the other direction and limiting the chances of mutation.

Therefore, stalled forks are protected from introducing further mutations at multiple levels. Firstly, these forks are protected from processing by the recombination machinery for as long as possible, secondly the recombination events are tightly restrained and thirdly the low processivity of BIR forks favours resolution by oncoming forks.

Known examples of transcription-induced genetic change

Effective DNA repair of spontaneous damage and the excellent processivity of the replisome form a multi-layered defence against potential mutations despite the considerable strain imposed by transcription. In keeping with this, mutation accumulation studies find only minimal evidence of a link between mutation and transcription (Chen and Zhang, 2014; Zhu et al., 2014), but clear examples of transcription-induced mutation have been reported and mechanistic drivers have been elucidated.

We suggest that stalling of DNA replication forks either spontaneously or due to a replication-transcription conflict will occasionally require fork restart by BIR, but *de novo* mutation only arises when poorly processive BIR forks encounter obstacles resulting (for example) from transcription (illustrated in Figure 2A,B). In other words, two separable events are required for a mutagenic outcome. Direct measurements of BIR fork progression reveal them to be profoundly impaired and prone to template switching by head-on encounters with transcription, as well as epigenetic marks to which the replisome is insensitive (Che et al., 2015; Liu et al., 2021).

The best characterised example of transcription-induced recombination in yeast occurs at the ribosomal DNA. Under the standard model, replication forks stalled at the ribosomal DNA RFB are cleaved, resected

then undergo non-allelic homologous recombination to cause CNV if cohesin is removed by the transcription of non-coding RNAs to facilitate recombination between repeats (Kobayashi and Ganley, 2005; Kobayashi et al., 2004). However, replication forks are stably paused at the RFB and it is unclear why these would be cleaved prior to resolution by oncoming replication forks (Carr and Lambert, 2013). Nonetheless, rDNA hyper-recombination occurs in mutants lacking polymerase component Pol32 or deacetylase Hst3/Hst4, which should only impact the processivity of BIR forks, indicating that BIR is involved (Che et al., 2015; Houseley and Tollervey, 2011; Ide et al., 2013; Jack et al., 2015). Replication forks paused at the RFB are stable (Calzada et al., 2005), but the prominent double strand break signal from the RFB indicates that replication fork reversal occurs, so resection and reinvasion of this end is likely to occur at some frequency (Burkhalter and Sogo, 2004; Kara et al., 2021; Zhu et al., 2019). The BIR events initiated are likely short-lived due to resolution by oncoming forks, but could easily encounter oncoming RNA polymerase I and be forced into template switching. This would require the BIR fork to traverse the RFB, but given that the RFB is mediated by FPC which is bound to CMG and that fork reversal dissociates the polymerases from CMG, this is plausible (Figure 2C).

Recently, we demonstrated that copper resistance acquisition in yeast arises through transcriptionally induced copy number amplification of the *CUP1* gene (Hull et al., 2017; Whale et al., 2022). The mechanism proved extremely complex, requiring a BIR event that becomes error-prone on encountering an epigenetic scar at a low-efficiency replication origin upstream of the gene itself. Transcriptional dependence requires the activator complex Mediator and the mRNA export complex TREX-2 which are recruited during activation of inducible genes such as *CUP1* . The only detectable TREX-2-dependent replication fork stalling at the locus occurred at the inefficient replication origin, suggesting that the primary impact of transcription is to prevent the successful firing of this origin. The unsuccessful origin firing leaves a scar which causes BIR forks repairing local replisome stalling to undergo template switching, resulting in copy number amplification (Figure 2D). Importantly, this behaviour was not specific to the *CUP1* gene as the *SFA1* gene, which bestows formaldehyde resistance and like *CUP1* has an inefficient replication origin just upstream, can undergo an equivalent transcriptionally-induced gene amplification through BIR (Hull et al., 2017).

It is worth noting that genetic changes induced by both these systems are not restricted to chromosomes. Both ribosomal DNA and *CUP1*loci are susceptible to transcription-induced extrachromosomal circular DNA (eccDNA) formation. These species, particularly extrachromosomal ribosomal DNA circles (ERCs), accumulate to massive levels in aged cells, adding 30-40% to genome size (Cruz et al., 2018; Hull et al., 2019; Sinclair and Guarente, 1997), and ability of eccDNA to reintegrate in chromosomal DNA provides an additional pathway by which major transcription-induced genetic changes can arise (Beverley et al., 1984; Brewer et al., 2015; Demeke et al., 2015; Galeote et al., 2011)

Interpretation of many studies in this area is complex due to the use of the extraordinarily highly expressed *GAL1* promoter; which is one of the few for which we can detect transcription-induced pausing of the replisome by TrAEL-seq (Kara et al., 2021). This may be sufficient to directly induce local BIR events at some frequency, and the pattern of transcription-induced mutations induced in a *lys2* reporter are consistent with those reported to occur during BIR (Datta and Jinks-Robertson, 1995). However, whether those mutations arise through conflicts between the BIR fork and transcription is hard to resolve if the rate of BIR induction is not necessarily constant, though follow-on studies showed a direct correlation between transcriptional strength and reversion of *lys2* mutations, as well as induction of recombination events (Kim et al., 2007; Saxe et al., 2000), all consistent with BIR.

Overall, where it is possible to separate the individual contributions of direct impacts of transcription on the replisome and on BIR, the evidence is consistent with mechanisms in which transcription-replisome conflicts are at most the initiator, whereas the actual mutations are caused by further conflicts impairing the processivity of BIR. We suggest that the danger posed by increased sensitivity of BIR forks to transcription-related obstacles is the reason why transcription is globally down-regulated in response to DNA replication stress (reviewed in (Silva and Ideker, 2019)). Notably however, for the ribosomal DNA,*CUP1* and *SFA1* transcription-induced genomic instability relies on specific properties of the locus, raising the question of

whether some genes are configured in a manner that might be more prone to mutation or genome rearrangement.

Replication patterns vary with gene function and expression

To address this point, we re-analysed our published TrAEL-seq datasets for increased replication fork density indicative of fork stalling (Kara et al., 2021; Whale et al., 2022), across gene sets categorised as 'housekeeping' or 'environment dependent' as we hypothesised that environment-dependent genes might be configured to evolve more readily than housekeeping genes. Although mutants of the transcriptional activators SAGA and TFIID affect steady state mRNA levels of all genes, Huisinga and Pugh noted that RNA pol II genes were more responsive to one or the other (Huisinga and Pugh, 2004), then Donczew *et al*refined this categorisation into TFIID dependent ('housekeeping') and Coactivator Redundant (CR, 'environment dependent') sets (Donczew et al., 2020). We stratified genes in quartiles for transcription based on NET-seq data (Churchman and Weissman, 2011), then subdivided each quartile into TFIID or CR genes (Figure 3A). Few genes in the lowest quartile are reliably designated as SAGA or CR dependent so this quartile was not subdivided.

For TFIID genes, replication forks moving head-on to the direction of transcription accumulate slightly across the entire transcribed region indicating that replisome progression is retarded (Figure 3B, top left), whereas signal from replisomes moving co-directionally with RNA polymerase II is reduced, indicating that replisome movement is accelerated (Figure 3B, top right). Curiously, these effects are equivalent across expression quartiles and therefore likely reflect a sensitivity of the replisome to transcription units rather than transcription-replisome conflicts.

In contrast, head-on CR genes show a transcription-dependent increase in signal from the TES to the TSS that would be consistent with the replisome being increasingly retarded by either direct encounters with RNA polymerase II or indirect features associated with transcription such as R-loops, while co-directional replisome progression is largely unaffected (Figure 3B, middle). The TrAEL-seq signal also increases dramatically in the 10kb upstream of the TSS (Figure 3C), which would be consistent with replication origins being more frequently located upstream of highly expressed CR genes. We therefore measured distances from each TSS to the nearest replication origin: this does not differ from random for TFIID genes but is significantly closer for highly expressed CR genes, with the majority of these genes having a replication origin within 10kb (Figure 3D). This is very interesting given our recent observation that CNV events triggered by expression of the *CUP1* gene depend on a closely adjacent replication origin (Whale et al., 2022); it should be noted that the *CUP1* locus was excluded from the analysis presented here because of high copy number.

The TFIID and SAGA gene sets were originally classified as 'housekeeping' or 'environment dependent', and GO analysis of the TFIID and CR gene sets remains in accord with this, the former being dominated by translation and the latter by metabolic genes (glycolysis and metabolite biosynthesis), which are used in a more environmentally dependent manner. This analysis indicates that the replisome tends to interact with transcription in head-on CR genes but not TFIID genes, skewing the potential for transcription induced mutation and CNV towards environmentally responsive genes. Furthermore, highly expressed CR genes have evolved to lie close to replication origins, which have the capacity to induce copy number variation mechanisms of the type we observed at *CUP1* .

Outlook

Yeasts are colonising organisms: a few cells arriving in a compatible environment must rapidly reproduce, likely in competition with other yeasts and microorganisms, then cells or spores from this population colonise new environments to complete the life cycle. The capacity for adaptive evolution is required to survive environmental differences and emerging threats from competing organisms as well as, where the colony forms in a living organism, host defences. However, populations arise from a few cells and have inherently low genetic heterogeneity, which limits evolutionary potential unless genetic heterogeneity is acquired during colony growth. Yet wild-type yeast cells in the lab have low enough rates of all mutation classes that populations grown from single cells are essentially clonal (Serero et al., 2014).

We suggest that the yeast genome has evolved to undergo mutation linked to transcription of environmentally responsive genes, since focusing mutations on such genes has a higher chance of yielding an adaptive genetic change. However, under ideal conditions this mutation rate is very low so transcriptionally-induced mutations are hard to detect, and the replisome is regulated to minimise replication fork stalling even under adverse conditions (Duch et al., 2018; Duch et al., 2013). However, long term exposure to environmental toxins such as ethanol does increase replication stress (Voordeckers et al., 2020), and we predict that genetic heterogeneity will then emerge in the population, particularly at environmentally responsive genes through increasing use and decreasing processivity of BIR forks.

Acknowledgements

We thank Alex Whale for critical reading of this manuscript. BP is funded by a Cambridge Commonwealth, European & International Trust Scholarship, JH is funded by the BBSRC [BI Epigenetics ISP; BBS/E/B/000C0523].

References

Azvolinsky, A., Dunaway, S., Torres, J.Z., Bessler, J.B., and Zakian, V.A. (2006). The S. cerevisiae Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. Genes Dev *20* , 3104-3116.

Azvolinsky, A., Giresi, P.G., Lieb, J.D., and Zakian, V.A. (2009). Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae. Mol Cell*34* , 722-734.

Beverley, S.M., Coderre, J.A., Santi, D.V., and Schimke, R.T. (1984). Unstable DNA amplifications in methotrexate-resistant Leishmania consist of extrachromosomal circles which relocalize during stabilization. Cell*38* , 431-439.

Billen, D. (1990). Spontaneous DNA damage and its significance for the "negligible dose" controversy in radiation protection. Radiat Res*124* , 242-245.

Blastyak, A., Pinter, L., Unk, I., Prakash, L., Prakash, S., and Haracska, L. (2007). Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell *28* , 167-175.

Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S., Villanueva, A., Esteller, M., and Guil, S. (2015). Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. PNAS *112* , 5785-5790.

Brambati, A., Zardoni, L., Achar, Y.J., Piccini, D., Galanti, L., Colosio, A., Foiani, M., and Liberi, G. (2018). Dormant origins and fork protection mechanisms rescue sister forks arrested by transcription. Nucleic Acids Res *46* , 1227-1239.

Brewer, B.J., Payen, C., Di Rienzi, S.C., Higgins, M.M., Ong, G., Dunham, M.J., and Raghuraman, M.K. (2015). Origin-Dependent Inverted-Repeat Amplification: Tests of a Model for Inverted DNA Amplification. PLoS Genet *11* , e1005699.

Brickner, J.R., Garzon, J.L., and Cimprich, K.A. (2022). Walking a tightrope: The complex balancing act of R-loops in genome stability. Mol Cell *82* , 2267-2297.

Brill, S.J., DiNardo, S., Voelkel-Meiman, K., and Sternglanz, R. (1987). Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. Nature *326* , 414-416.

Brill, S.J., and Sternglanz, R. (1988). Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. Cell *54* , 403-411.

Brooks, P.J., Wise, D.S., Berry, D.A., Kosmoski, J.V., Smerdon, M.J., Somers, R.L., Mackie, H., Spoonde, A.Y., Ackerman, E.J., Coleman, K.*, et al.* (2000). The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine

is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. J Biol Chem *275* , 22355-22362.

Burkhalter, M.D., and Sogo, J.M. (2004). rDNA enhancer affects replication initiation and mitotic recombination: Fob1 mediates nucleolytic processing independently of replication. Mol Cell *15* , 409-421.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. (2005). Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. Genes Dev *19* , 1905-1919.

Canas, J.C., Garcia-Rubio, M.L., Garcia, A., Antequera, F., Gomez-Gonzalez, B., and Aguilera, A. (2022). A role for the Saccharomyces cerevisiae Rtt109 histone acetyltransferase in R-loop homeostasis and associated genome instability. Genetics *222* .

Carr, A.M., and Lambert, S. (2013). Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. J Mol Biol *425* , 4733-4744.

Chan, Y.A., Aristizabal, M.J., Lu, P.Y., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., and Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. PLoS Genet *10* , e1004288.

Che, J., Smith, S., Kim, Y.J., Shim, E.Y., Myung, K., and Lee, S.E. (2015). Hyper-Acetylation of Histone H3K56 Limits Break-Induced Replication by Inhibiting Extensive Repair Synthesis. PLoS Genet*11* , e1004990.

Chen, X., and Zhang, J. (2014). Yeast mutation accumulation experiment supports elevated mutation rates at highly transcribed sites. PNAS*111* , E4062.

Cho, J.E., Kim, N., and Jinks-Robertson, S. (2015). Topoisomerase 1-dependent deletions initiated by incision at ribonucleotides are biased to the non-transcribed strand of a highly activated reporter. Nucleic Acids Res *43* , 9306-9313.

Cho, J.E., Kim, N., Li, Y.C., and Jinks-Robertson, S. (2013). Two distinct mechanisms of Topoisomerase 1-dependent mutagenesis in yeast. DNA Repair (Amst) *12* , 205-211.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature*469* , 368-373.

Claussin, C., Vazquez, J., and Whitehouse, I. (2022). Single-molecule mapping of replisome progression. Mol Cell *82* , 1372-1382 e1374.

Cotta-Ramusino, C., Fachinetti, D., Lucca, C., Doksani, Y., Lopes, M., Sogo, J., and Foiani, M. (2005). Exo1 processes stalled replication forks and counteracts fork reversal in checkpoint-defective cells. Mol Cell *17* , 153-159.

Crossley, M.P., Song, C., Bocek, M.J., Choi, J.H., Kousorous, J., Sathirachinda, A., Lin, C., Brickner, J.R., Bai, G., Lans, H.*, et al.* (2023). R-loop-derived cytoplasmic RNA-DNA hybrids activate an immune response. Nature *613* , 187-194.

Cruz, C., Della Rosa, M., Krueger, C., Gao, Q., Horkai, D., King, M., Field, L., and Houseley, J. (2018). Tri-methylation of histone H3 lysine 4 facilitates gene expression in ageing cells. eLife *7* , e34081.

Datta, A., and Jinks-Robertson, S. (1995). Association of increased spontaneous mutation rates with high levels of transcription in yeast. Science *268* , 1616-1619.

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., Chabes, A., and Malkova, A. (2011). Break-induced replication is highly inaccurate. PLoS Biol *9* , e1000594.

Demeke, M.M., Foulquie-Moreno, M.R., Dumortier, F., and Thevelein, J.M. (2015). Rapid evolution of recombinant Saccharomyces cerevisiae for Xylose fermentation through formation of extra-chromosomal circular DNA. PLoS Genet *11* , e1005010.

Deshpande, A.M., and Newlon, C.S. (1996). DNA replication fork pause sites dependent on transcription. Science *272* , 1030-1033.

Donczew, R., Warfield, L., Pacheco, D., Erijman, A., and Hahn, S. (2020). Two roles for the yeast transcription coactivator SAGA and a set of genes redundantly regulated by TFIID and SAGA. eLife *9* .

Duan, M., Selvam, K., Wyrick, J.J., and Mao, P. (2020). Genome-wide role of Rad26 in promoting transcription-coupled nucleotide excision repair in yeast chromatin. PNAS *117* , 18608-18616.

Duch, A., Canal, B., Barroso, S.I., Garcia-Rubio, M., Seisenbacher, G., Aguilera, A., de Nadal, E., and Posas, F. (2018). Multiple signaling kinases target Mrc1 to prevent genomic instability triggered by transcription-replication conflicts. Nature communications *9* , 379.

Duch, A., Felipe-Abrio, I., Barroso, S., Yaakov, G., Garcia-Rubio, M., Aguilera, A., de Nadal, E., and Posas, F. (2013). Coordinated control of replication and transcription by a SAPK protects genomic integrity. Nature *493* , 116-119.

El Hage, A., French, S.L., Beyer, A.L., and Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev *24* , 1546-1558.

Esnault, C., Magat, T., Zine El Aabidine, A., Garcia-Oliver, E., Cucchiarini, A., Bouchouika, S., Lleres, D., Goerke, L., Luo, Y., Verga, D., *et al.* (2023). G4access identifies G-quadruplexes and their associations with open chromatin and imprinting control regions. Nat Genet *55* , 1359-1369.

Felipe-Abrio, I., Lafuente-Barquero, J., Garcia-Rubio, M.L., and Aguilera, A. (2015). RNA polymerase II contributes to preventing transcription-mediated replication fork stalls. EMBO J *34* , 236-250.

Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P., and Kohtz, J.D. (2006). The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev *20* , 1470-1484.

Fielden, J., Ruggiano, A., Popovic, M., and Ramadan, K. (2018). DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. DNA Repair (Amst) *71* , 198-204.

French, S. (1992). Consequences of replication fork movement through transcription units in vivo. Science *258* , 1362-1365.

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

Gaillard, H., Garcia-Benitez, F., and Aguilera, A. (2018). Gene gating at nuclear pores prevents the formation of R loops. Mol Cell Oncol*5* , e1405140.

Galeote, V., Bigey, F., Beyne, E., Novo, M., Legras, J.L., Casaregola, S., and Dequin, S. (2011). Amplification of a Zygosaccharomyces bailii DNA segment in wine yeast genomes by extrachromosomal circular DNA formation. PLoS One *6* , e17872.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol *8* , 358-366.

Garcia-Benitez, F., Gaillard, H., and Aguilera, A. (2017). Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. PNAS *114* , 10942-10947.

Garcia-Muse, T., and Aguilera, A. (2016). Transcription-replication conflicts: how they occur and how they are resolved. Nat Rev Mol Cell Biol *17* , 553-563.

Garcia-Pichardo, D., Canas, J.C., Garcia-Rubio, M.L., Gomez-Gonzalez, B., Rondon, A.G., and Aguilera, A. (2017). Histone Mutants Separate R Loop Formation from Genome Instability Induction. Mol Cell *66* , 597-609 e595.

Garcia-Rubio, M., Aguilera, P., Lafuente-Barquero, J., Ruiz, J.F., Simon, M.N., Geli, V., Rondon, A.G., and Aguilera, A. (2018). Yra1-bound RNA-DNA hybrids cause orientation-independent transcription-replication collisions and telomere instability. Genes Dev *32* , 965-977.

Goehring, L., Huang, T.T., and Smith, D.J. (2023). Transcription-Replication Conflicts as a Source of Genome Instability. Annu Rev Genet.

Gomez-Gonzalez, B., Garcia-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marin, A., Foiani, M., and Aguilera, A. (2011). Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. EMBO J *30* , 3106-3119.

Graham, J.E., Marians, K.J., and Kowalczykowski, S.C. (2017). Independent and Stochastic Action of DNA Polymerases in the Replisome. Cell *169* , 1201-1213 e1217.

Gregersen, L.H., and Svejstrup, J.Q. (2018). The Cellular Response to Transcription-Blocking DNA Damage. Trends in biochemical sciences*43* , 327-341.

Guintini, L., Paille, A., Graf, M., Luke, B., Wellinger, R.J., and Conconi, A. (2022). Transcription of ncRNAs promotes repair of UV induced DNA lesions in Saccharomyces cerevisiae subtelomeres. PLoS Genet*18* , e1010167.

Hamperl, S., Bocek, M.J., Saldivar, J.C., Swigut, T., and Cimprich, K.A. (2017). Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. Cell *170* , 774-786 e719.

Herrera-Moyano, E., Mergui, X., Garcia-Rubio, M.L., Barroso, S., and Aguilera, A. (2014). The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. Genes Dev *28* , 735-748.

Hodgson, B., Calzada, A., and Labib, K. (2007). Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase. Molecular biology of the cell *18* , 3894-3902.

Houseley, J., and Tollervey, D. (2011). Repeat expansion in the budding yeast ribosomal DNA can occur independently of the canonical homologous recombination machinery. Nucleic Acids Res *39* , 8778-8791.

Huang, S.Y., Ghosh, S., and Pommier, Y. (2015). Topoisomerase I alone is sufficient to produce short DNA deletions and can also reverse nicks at ribonucleotide sites. J Biol Chem *290* , 14068-14076.

Huertas, P., and Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell *12* , 711-721.

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell *13* , 573-585.

Hull, R.M., Cruz, C., Jack, C.V., and Houseley, J. (2017). Environmental change drives accelerated adaptation through stimulated copy number variation. PLoS Biol *15* , e2001333.

Hull, R.M., King, M., Pizza, G., Krueger, F., Vergara, X., and Houseley, J. (2019). Transcription-induced formation of extrachromosomal DNA during yeast ageing. PLoS Biol *17* , e3000471.

Ide, S., Saka, K., and Kobayashi, T. (2013). Rtt109 prevents hyper-amplification of ribosomal RNA genes through histone modification in budding yeast. PLoS Genet *9* , e1003410.

Ivessa, A.S., Lenzmeier, B.A., Bessler, J.B., Goudsouzian, L.K., Schnakenberg, S.L., and Zakian, V.A. (2003). The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. Mol Cell *12* , 1525-1536.

Jack, C.V., Cruz, C., Hull, R.M., Keller, M.A., Ralser, M., and Houseley, J. (2015). Regulation of ribosomal DNA amplification by the TOR pathway. PNAS *112* , 9674-9679.

Janissen, R., Barth, R., Polinder, M., Torre, J.v.d., and Dekker, C. (2023). Single-molecule visualization of twin-supercoiled domains generated during transcription. bioRxiv, 2023.2008.2025.554779.

Jones, M.L., Baris, Y., Taylor, M.R.G., and Yeeles, J.T.P. (2021). Structure of a human replisome shows the organisation and interactions of a DNA replication machine. EMBO J *40* , e108819.

Jung, Y., and Lippard, S.J. (2007). Direct cellular responses to platinum-induced DNA damage. Chemical reviews *107* , 1387-1407.

Kabeche, L., Nguyen, H.D., Buisson, R., and Zou, L. (2018). A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science *359* , 108-114.

Kara, N., Krueger, F., Rugg-Gunn, P., and Houseley, J. (2021). Genome-wide analysis of DNA replication and DNA double-strand breaks using TrAEL-seq. PLoS Biol *19* , e3000886.

Keszthelyi, A., Whale, A., Houseley, J., and Baxter, J. (2023). The Fork Protection Complex protects long replicons from DNA damage at the cost of genome instability induced by DNA topological stress. bioRxiv, 2023.2008.2004.551986.

Khurana, S., and Oberdoerffer, P. (2015). Replication Stress: A Lifetime of Epigenetic Change. Genes (Basel) *6* , 858-877.

Kim, N., Abdulovic, A.L., Gealy, R., Lippert, M.J., and Jinks-Robertson, S. (2007). Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. DNA Repair (Amst) *6* , 1285-1296.

Kim, N., Huang, S.N., Williams, J.S., Li, Y.C., Clark, A.B., Cho, J.E., Kunkel, T.A., Pommier, Y., and Jinks-Robertson, S. (2011). Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. Science*332* , 1561-1564.

Kobayashi, T., and Ganley, A.R. (2005). Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. Science*309* , 1581-1584.

Kobayashi, T., Heck, D.J., Nomura, M., and Horiuchi, T. (1998). Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes Dev *12* , 3821-3830.

Kobayashi, T., and Horiuchi, T. (1996). A yeast gene product, Fob1 protein, required for both replication fork blocking and recombinational hotspot activities. Genes Cells *1* , 465-474.

Kobayashi, T., Horiuchi, T., Tongaonkar, P., Vu, L., and Nomura, M. (2004). SIR2 regulates recombination between different rDNA repeats, but not recombination within individual rRNA genes in yeast. Cell*117* , 441-453.

Kramara, J., Osia, B., and Malkova, A. (2018). Break-Induced Replication: The Where, The Why, and The How. Trends Genet *34* , 518-531.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S.*, et al.* (1997). The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature *390* , 249-256.

Labib, K., and Hodgson, B. (2007). Replication fork barriers: pausing for a break or stalling for time? EMBO Rep *8* , 346-353.

Lans, H., Hoeijmakers, J.H.J., Vermeulen, W., and Marteijn, J.A. (2019). The DNA damage response to transcription stress. Nat Rev Mol Cell Biol*20* , 766-784.

Lemacon, D., Jackson, J., Quinet, A., Brickner, J.R., Li, S., Yazinski, S., You, Z., Ira, G., Zou, L., Mosammaparast, N.*, et al.* (2017). MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. Nature communications *8* , 860.

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

Li, S., and Smerdon, M.J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in Saccharomyces cerevisiae. EMBO J*21* , 5921-5929.

Li, S., and Smerdon, M.J. (2004). Dissecting transcription-coupled and global genomic repair in the chromatin of yeast GAL1-10 genes. J Biol Chem *279* , 14418-14426.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362* , 709-715.

Lippert, M.J., Kim, N., Cho, J.E., Larson, R.P., Schoenly, N.E., O'Shea, S.H., and Jinks-Robertson, S. (2011). Role for topoisomerase 1 in transcription-associated mutagenesis in yeast. PNAS *108* , 698-703.

Liu, L., and Malkova, A. (2022). Break-induced replication: unraveling each step. Trends Genet *38* , 752-765.

Liu, L., Yan, Z., Osia, B.A., Twarowski, J., Sun, L., Kramara, J., Lee, R.S., Kumar, S., Elango, R., Li, H., *et al.* (2021). Tracking break-induced replication shows that it stalls at roadblocks. Nature*590* , 655-659.

Ljungman, M. (2022). Transcription and genome integrity. DNA Repair (Amst) *118* , 103373.

Lockhart, A., Pires, V.B., Bento, F., Kellner, V., Luke-Glaser, S., Yakoub, G., Ulrich, H.D., and Luke, B. (2019). RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. Cell reports*29* , 2890-2900 e2895.

Lopes, J., Piazza, A., Bermejo, R., Kriegsman, B., Colosio, A., Teulade-Fichou, M.P., Foiani, M., and Nicolas, A. (2011). G-quadruplex-induced instability during leading-strand replication. EMBO J *30* , 4033-4046.

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

Luna, R., Rondon, A.G., Perez-Calero, C., Salas-Armenteros, I., and Aguilera, A. (2019). The THO Complex as a Paradigm for the Prevention of Cotranscriptional R-Loops. Cold Spring Harb Symp Quant Biol *84* , 105-114.

Lydeard, J.R., Lipkin-Moore, Z., Sheu, Y.J., Stillman, B., Burgers, P.M., and Haber, J.E. (2010). Break-induced replication requires all essential DNA replication factors except those specific for pre-RC assembly. Genes Dev *24* , 1133-1144.

Malkova, A., and Ira, G. (2013). Break-induced replication: functions and molecular mechanism. Current opinion in genetics & development*23* , 271-279.

Mao, P., Brown, A.J., Malc, E.P., Mieczkowski, P.A., Smerdon, M.J., Roberts, S.A., and Wyrick, J.J. (2017). Genome-wide maps of alkylation damage, repair, and mutagenesis in yeast reveal mechanisms of mutational heterogeneity. Genome research *27* , 1674-1684.

Martinez, B., Bharati, B.K., Epshtein, V., and Nudler, E. (2022). Pervasive Transcription-coupled DNA repair in E. coli. Nature communications *13* , 1702.

Mayle, R., Campbell, I.M., Beck, C.R., Yu, Y., Wilson, M., Shaw, C.A., Bjergbaek, L., Lupski, J.R., and Ira, G. (2015). DNA REPAIR. Mus81 and converging forks limit the mutagenicity of replication fork breakage. Science *349* , 742-747.

McClure, A.W., Canal, B., and Diffley, J.F.X. (2022). A DNA replication fork-centric view of the budding yeast DNA damage response. DNA Repair (Amst) *119* , 103393.

McLean, M.J., Wolfe, K.H., and Devine, K.M. (1998). Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. J Mol Evol *47* , 691-696.

Mischo, H.E., Gomez-Gonzalez, B., Grzechnik, P., Rondon, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol Cell *41* , 21-32.

Muniz, L., Nicolas, E., and Trouche, D. (2021). RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. EMBO J *40* , e105740.

Namiki, Y., and Zou, L. (2006). ATRIP associates with replication protein A-coated ssDNA through multiple interactions. PNAS *103* , 580-585.

Neil, A.J., Liang, M.U., Khristich, A.N., Shah, K.A., and Mirkin, S.M. (2018). RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)n repeats via break-induced replication. Nucleic Acids Res*46* , 3487-3497.

Noguchi, E., Noguchi, C., McDonald, W.H., Yates, J.R., 3rd, and Russell, P. (2004). Swi1 and Swi3 are components of a replication fork protection complex in fission yeast. Mol Cell Biol *24* , 8342-8355.

O'Connell, K., Jinks-Robertson, S., and Petes, T.D. (2015). Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity. Genetics *201* , 963-975.

Paeschke, K., Capra, J.A., and Zakian, V.A. (2011). DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. Cell *145* , 678-691.

Pardo, B., and Aguilera, A. (2012). Complex chromosomal rearrangements mediated by break-induced replication involve structure-selective endonucleases. PLoS Genet *8* , e1002979.

Pardo, B., Crabbe, L., and Pasero, P. (2017). Signaling pathways of replication stress in yeast. FEMS yeast research *17* .

Pellegrini, L. (2023). The CMG DNA helicase and the core replisome. Curr Opin Struct Biol *81* , 102612.

Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. Mol Cell *37* , 492-502.

Piazza, A., Boule, J.B., Lopes, J., Mingo, K., Largy, E., Teulade-Fichou, M.P., and Nicolas, A. (2010). Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in Saccharomyces cerevisiae. Nucleic Acids Res *38* , 4337-4348.

Powers, K.T., and Washington, M.T. (2018). Eukaryotic translesion synthesis: Choosing the right tool for the job. DNA Repair (Amst)*71* , 127-134.

Prado, F., Piruat, J.I., and Aguilera, A. (1997). Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. EMBO J *16* , 2826-2835.

Puddu, F., Piergiovanni, G., Plevani, P., and Muzi-Falconi, M. (2011). Sensing of replication stress and Mec1 activation act through two independent pathways involving the 9-1-1 complex and DNA polymerase epsilon. PLoS Genet *7* , e1002022.

Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K.J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., and Lopes, M. (2012). Topoisomerase I poisoning results in PARP-mediated replication fork reversal. Nat Struct Mol Biol *19* , 417-423.

Ripley, B.M., Gildenberg, M.S., and Washington, M.T. (2020). Control of DNA Damage Bypass by Ubiquitylation of PCNA. Genes (Basel) *11* .

Rocha, E.P., and Danchin, A. (2003). Gene essentiality determines chromosome organisation in bacteria. Nucleic Acids Res *31* , 6570-6577.

Rossi, M.J., Kuntala, P.K., Lai, W.K.M., Yamada, N., Badjatia, N., Mittal, C., Kuzu, G., Bocklund, K., Farrell, N.P., Blanda, T.R., *et al.* (2021). A high-resolution protein architecture of the budding yeast genome. Nature *592* , 309-314.

Sakofsky, C.J., Ayyar, S., Deem, A.K., Chung, W.H., Ira, G., and Malkova, A. (2015). Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. Mol Cell *60* , 860-872.

Sakofsky, C.J., Roberts, S.A., Malc, E., Mieczkowski, P.A., Resnick, M.A., Gordenin, D.A., and Malkova, A. (2014). Break-induced replication is a source of mutation clusters underlying kataegis. Cell reports *7* , 1640-1648.

San Martin-Alonso, M., Soler-Oliva, M.E., Garcia-Rubio, M., Garcia-Muse, T., and Aguilera, A. (2021). Harmful R-loops are prevented via different cell cycle-specific mechanisms. Nature communications *12* , 4451.

Sankar, T.S., Wastuwidyaningtyas, B.D., Dong, Y., Lewis, S.A., and Wang, J.D. (2016). The nature of mutations induced by replication-transcription collisions. Nature *535* , 178-181.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat Rev Genet *16* , 583-597.

Santos-Pereira, J.M., Herrero, A.B., Garcia-Rubio, M.L., Marin, A., Moreno, S., and Aguilera, A. (2013). The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. Genes Dev *27* , 2445-2458.

Saxe, D., Datta, A., and Jinks-Robertson, S. (2000). Stimulation of mitotic recombination events by high levels of RNA polymerase II transcription in yeast. Mol Cell Biol *20* , 5404-5414.

Sekedat, M.D., Fenyo, D., Rogers, R.S., Tackett, A.J., Aitchison, J.D., and Chait, B.T. (2010). GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. Mol Syst Biol*6* , 353.

Selvam, K., Plummer, D.A., Mao, P., and Wyrick, J.J. (2022). Set2 histone methyltransferase regulates transcription coupled-nucleotide excision repair in yeast. PLoS Genet *18* , e1010085.

Selvam, S., Koirala, D., Yu, Z., and Mao, H. (2014). Quantification of topological coupling between DNA superhelicity and G-quadruplex formation. Journal of the American Chemical Society *136* , 13967-13970.

Serero, A., Jubin, C., Loeillet, S., Legoix-Ne, P., and Nicolas, A.G. (2014). Mutational landscape of yeast mutator strains. PNAS *111* , 1897-1902.

Shin, J.H., and Kelman, Z. (2006). The replicative helicases of bacteria, archaea, and eukarya can unwind RNA-DNA hybrid substrates. J Biol Chem *281* , 26914-26921.

Shyian, M., and Shore, D. (2021). Approaching Protein Barriers: Emerging Mechanisms of Replication Pausing in Eukaryotes. Front Cell Dev Biol*9* , 672510.

Silva, E., and Ideker, T. (2019). Transcriptional responses to DNA damage. DNA Repair (Amst) *79* , 40-49.

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinkovic, D., *et al.* (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature *510* , 293-297.

Sinclair, D.A., and Guarente, L. (1997). Extrachromosomal rDNA circles–a cause of aging in yeast. Cell *91* , 1033-1042.

Skourti-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N.J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature *516* , 436-439.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell *42* , 794-805.

Sloan, R., Huang, S.N., Pommier, Y., and Jinks-Robertson, S. (2017). Effects of camptothecin or TOP1 overexpression on genetic stability in Saccharomyces cerevisiae. DNA Repair (Amst) *59* , 69-75.

Smith, C.E., Llorente, B., and Symington, L.S. (2007). Template switching during break-induced replication. Nature *447* , 102-105.

Sollier, J., Stork, C.T., Garcia-Rubio, M.L., Paulsen, R.D., Aguilera, A., and Cimprich, K.A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. Mol Cell*56* , 777-785.

Stewart, S.E., and Roeder, G.S. (1989). Transcription by RNA polymerase I stimulates mitotic recombination in Saccharomyces cerevisiae. Mol Cell Biol *9* , 3464-3472.

Takahashi, D.T., Burguiere-Slezak, G., Van der Kemp, P.A., and Boiteux, S. (2011). Topoisomerase 1 provokes the formation of short deletions in repeated sequences upon high transcription in Saccharomyces cerevisiae. PNAS *108* , 692-697.

Tornaletti, S., Maeda, L.S., and Hanawalt, P.C. (2006). Transcription arrest at an abasic site in the transcribed strand of template DNA. Chem Res Toxicol *19* , 1215-1220.

Toth, R., Balogh, D., Pinter, L., Jaksa, G., Szeplaki, B., Graf, A., Gyorfy, Z., Enyedi, M.Z., Kiss, E., Haracska, L.*, et al.* (2022). The Rad5 Helicase and RING Domains Contribute to Genome Stability through their Independent Catalytic Activities. J Mol Biol *434* , 167437.

Tran, P.L.T., Pohl, T.J., Chen, C.F., Chan, A., Pott, S., and Zakian, V.A. (2017). PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. Nature communications *8* , 15025.

Tsao, Y.P., Wu, H.Y., and Liu, L.F. (1989). Transcription-driven supercoiling of DNA: direct biochemical evidence from in vitro studies. Cell *56* , 111-118.

Unk, I., Hajdu, I., Blastyak, A., and Haracska, L. (2010). Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair (Amst) *9* , 257-267.

van Schie, J.J.M., and de Lange, J. (2021). The Interplay of Cohesin and the Replisome at Processive and Stressed DNA Replication Forks. Cells*10* .

Voordeckers, K., Colding, C., Grasso, L., Pardo, B., Hoes, L., Kominek, J., Gielens, K., Dekoster, K., Gordon, J., Van der Zande, E.*, et al.* (2020). Ethanol exposure increases mutation rate through error-prone polymerases. Nature communications *11* , 3664.

Vrtis, K.B., Dewar, J.M., Chistol, G., Wu, R.A., Graham, T.G.W., and Walter, J.C. (2021). Single-strand DNA breaks cause replisome disassembly. Mol Cell *81* , 1309-1318 e1306.

Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., and Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes Dev *30* , 1327-1338.

Wasserman, M.R., Schauer, G.D., O'Donnell, M.E., and Liu, S. (2019). Replication Fork Activation Is Enabled by a Single-Stranded DNA Gate in CMG Helicase. Cell *178* , 600-611 e616.

Wellinger, R.E., Prado, F., and Aguilera, A. (2006). Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. Mol Cell Biol *26* , 3327-3334.

Westhorpe, R., Keszthelyi, A., Minchell, N.E., Jones, D., and Baxter, J. (2020). Separable functions of Tof1/Timeless in intra-S-checkpoint signalling, replisome stability and DNA topological stress. Nucleic Acids Res *48* , 12169-12187.

Whale, A.J., King, M., Hull, R.M., Krueger, F., and Houseley, J. (2022). Stimulation of adaptive gene amplification by origin firing under replication fork constraint. Nucleic Acids Res *50* , 915-936.

Williams, J.D., Zhu, D., Garcia-Rubio, M., Shaltz, S., Aguilera, A., and Jinks-Robertson, S. (2023). Spontaneous deamination of cytosine to uracil is biased to the non-transcribed DNA strand in yeast. DNA Repair (Amst) *126* , 103489.

Williams, J.S., and Kunkel, T.A. (2014). Ribonucleotides in DNA: origins, repair and consequences. DNA Repair (Amst) *19* , 27-37.

Williams, J.S., Smith, D.J., Marjavaara, L., Lujan, S.A., Chabes, A., and Kunkel, T.A. (2013). Topoisomerase 1-mediated removal of ribonucleotides from nascent leading-strand DNA. Mol Cell *49* , 1010-1015.

Xu, J., Lahiri, I., Wang, W., Wier, A., Cianfrocco, M.A., Chong, J., Hare, A.A., Dervan, P.B., DiMaio, F., Leschziner, A.E.*, et al.*(2017). Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. Nature *551* , 653-657.

Yeung, R., and Smith, D.J. (2020). Determinants of Replication-Fork Pausing at tRNA Genes in Saccharomyces cerevisiae. Genetics *214* , 825-838.

Zardoni, L., Nardini, E., Brambati, A., Lucca, C., Choudhary, R., Loperfido, F., Sabbioneda, S., and Liberi, G. (2021). Elongating RNA polymerase II and RNA:DNA hybrids hinder fork progression and gene expression at sites of head-on replication-transcription collisions. Nucleic Acids Res *49* , 12769-12784.

Zhao, H., Zhu, M., Limbo, O., and Russell, P. (2018). RNase H eliminates R-loops that disrupt DNA replication but is nonessential for efficient DSB repair. EMBO Rep *19* .

Zhao, L., and Washington, M.T. (2017). Translesion Synthesis: Insights into the Selection and Switching of DNA Polymerases. Genes (Basel)*8* .

Zhu, Y., Biernacka, A., Pardo, B., Dojer, N., Forey, R., Skrzypczak, M., Fongang, B., Nde, J., Yousefi, R., Pasero, P.*, et al.* (2019). qDSB-Seq is a general method for genome-wide quantification of DNA double-strand breaks using sequencing. Nature communications *10* , 2313.

Zhu, Y.O., Siegal, M.L., Hall, D.W., and Petrov, D.A. (2014). Precise estimates of mutation rate and spectrum in yeast. PNAS *111* , E2310-2318.

Zimmer, A.D., and Koshland, D. (2016). Differential roles of the RNases H in preventing chromosome instability. PNAS *113* , 12220-12225.

Figure Legends

Figure 1: Replication fork impediments and processing pathways

A: Graphical representation of a co-transcriptional R-loop, formed by nascent DNA becomin bound to the template strand behind the polymerase.

B: Replication pass a polymerase blocking lesion: first, the leading strand polymerase stalls but the replisome continues forming a region of ssDNA ahead of the fork. The polymerase can be restarted either by trans-lesion synthesis or by template switching to use the newly synthesised lagging strand as a template.

C: Initiation of BIR: a stalled replication fork reverses, is resected then invades the template DNA aided by the homologous recombination machinery. Replication can then restart in a migrating D-loop with uncoupled synthesis of the lagging strand, using Pif1 as a helicase to improve processivity.

Figure 2: Mechanisms of genome instability arising through BIR

A: In normal replication forks, CMG is the helicase that unwinds dsDNA. It has the ability to remove many types of impediments, including RNA-DNA hybrids, RNA polymerases and other DNA-bound proteins with minimal pausing, thereby allowing a fast progression of the replication with short pausing times.

B: Break-induced replication (BIR) forks are far more prone to template-switching and to instigate non-allelic recombination, likely because the helicase activity of Pif1 alone is far less than that of the CMG complex with assistance from Pif1 and Rrm3, making BIR more prone to template-switch when encountering impediments.

C: Proposed model for transcription-induced recombination at the ribosomal DNA locus in budding yeast. Stabilised arrested forks can sometimes erroneously be resumed by BIR; this could cause the encounter of switching-prone BIR forks and elongating RNA pol I, leading to template-switching, the formation of extrachromosomal ribosomal DNA circles and copy-number variation of ribosomal DNA repeats.

D: Proposed model for *CUP1* copy-number variation. When encountering the scar left by a collapsed low-efficiency replication origin inside the *CUP1* repeats, a BIR fork could template-switch to another *CUP1* repeat and drive the formation of extrachromosomal CUP1 circles and copy-number variation of *CUP1*repeats. Note that the replisome stalling event that precipitates BIR does not need to be caused by transcription in this model, it is the collapse of the low-efficiency origin that is caused by transcription.

Figure 3: Replisome interactions differ between core and variable genes

A: Gene expression in quartiles separated by gene type. NET-seq reads in the first 500bp of each ORF for cells in YPD were used to rank all genes by expression (Churchman and Weissman, 2011); genes were then separated into 4 expression quartiles of 1,663 genes. Within the 3 higher quartiles, genes were separated into TFIID and CR regulated sets based on (Donczew et al., 2020), unassigned genes were discarded. n= (left to right) 1189, 366, 1414, 131, 1364, 96, 1664. B: TrAEL-seq metaplots of cells growing on YPD for TFIID and CR genes in quartiles described in A, as well as the unstratified lowest quartile. In each case, TrAEL-seq reads representing forks moving head-on or co-directional to the direction of transcription are considered separately. Metaplots are calculated in 20 bp bins over the transcriptional start site (TSS) to transcriptional end site (TES), and for 10kb regions either side. Bins from multi-copy regions (mitochondrial DNA, ribosomal DNA, *CUP1* ,*ENA* , telomeres and sub-telomeres, Ty elements and LTRs) were excluded. Data is an average of 4 wild-type datasets, shaded region shows standard deviation across datasets. C: Metaplot as in B but for 30kb either side of TSS/TES. D: Plot of the distance from TSS to nearest ARS for each gene separated by categories in B. Random distribution represents distance from 2000 random sites to nearest ARS. Whiskers indicate 10-90%. Analysis by one-way ANOVA, *** = $p<0.0001$, ns = not significant $p>0.05$.