Published in final edited form as:

Mol Cell. 2020 September 17; 79(6): 1037–1050.e5. doi:10.1016/j.molcel.2020.08.011.

Genetic characterization of three distinct mechanisms supporting RNA-driven DNA repair and modification reveals major role of DNA polymerase ζ

Chance Meers[1,2], **Havva Keskin**[1,3], **Gabor Banyai**[1], **Olga Mazina**[4], **Taehwan Yang**[1], **Alli L. Gombolay**[1], **Kuntal Mukherjee**[1], **Efiyenia I. Kaparos**[1,5], **Gary Newnam**[1], **Alexander Mazin**[4], **Francesca Storici**[1,6,*]

[1]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA.

[2]Currently at Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

[3]Currently at Omega Bio-Tek, Norcross, GA 30071, USA.

[4]Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, Philadelphia, PA 19102, USA.

[5]Currently at School of Medicine, New York University, New York, NY 10016, USA.

[6]Lead Contact

Summary

DNA double-strand breaks (DSBs) are dangerous lesions threatening genomic stability. Fidelity of DSB repair is best achieved by recombination with a homologous template sequence. In yeast, transcript RNA was shown to template DSB repair of DNA. However, molecular pathways of RNA-driven repair processes remain obscure. Utilizing assays of RNA-DNA recombination with and without an induced DSB in yeast DNA, we characterize three forms of RNA-mediated genomic modifications, RNA- and cDNA-templated DSB repair (R-TDR and c-TDR) using an RNA transcript or a DNA copy of the RNA transcript for DSB repair, respectively; and a new mechanism of RNA-templated DNA modification (R-TDM) induced by spontaneous or mutagen-induced breaks. While c-TDR requires reverse transcriptase, translesion DNA polymerase ζ plays a major role in R-TDR and it is essential for R-TDM. This study characterizes mechanisms of

*Correspondence: storici@gatech.edu.

Author Contributions

C.M. conducted most of the *in vivo* experiments and performed most of the statistical analysis of the data. H.K. conducted part of *in vivo* experiments and helped in the data analysis. G.B. performed qRT-PCR and western blot analysis. T.Y., A.L.G. and E.K. worked in the mutant screen. K.M assisted in mutagenic assay experiments. G.N. assisted with lab procedures. A.V.M. together with O.M.M. designed and analyzed *in vitro* experiments. C.M and F.S. designed experiments and wrote the manuscript with input and assistance from all authors.

Declaration of Interests

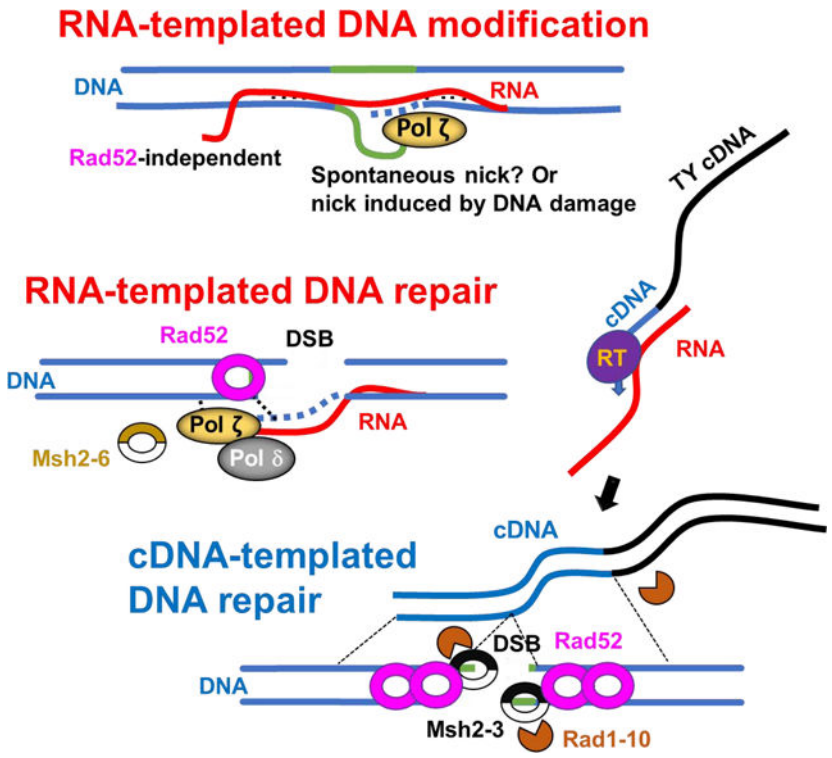
The authors declare no competing financial interest.

RNA-DNA recombination, uncovering a role of Pol ζ in transferring genetic information from transcript RNA to DNA.

eTOC Blurb

Still much is unknown about the roles of RNA in DNA repair and modification. Meers et al. characterize three genetic mechanisms by which RNA can directly or indirectly transfer genetic information to DNA in yeast cells, and uncover that direct RNA-templated DNA recombination is aided by translesion DNA Polymerase ζ.

Graphical Abstract



Keywords

DNA repair; DSB repair; RNA-mediated; RNA-templated; homologous recombination; reverse transcriptase; cDNA-mediated; Ty-less; translesion polymerase; DNA polymerase ζ (Zeta); *REV3*; *REV1*; *RAD5*; *RAD1*; *RAD10*; *MSH2*; *MSH3*; RNA recombination

Introduction

The preservation of genomic integrity is a delicate balancing act between faithful repair of DNA damage guided by endogenous repair systems and erroneous mishaps of these repair pathways. DNA double-stranded breaks (DSBs) are among the most dangerous types of DNA lesions leading to mutations, chromosome rearrangements and/or inhibiting cells ability to divide (Chatterjee and Walker, 2017). DSBs are repaired by two main pathways

including non-homologous end joining (NHEJ) and homologous recombination (HR). NHEJ proceeds by ligating the broken DNA ends together at the expense of frequent addition or deletion of genetic information at the DSB site (Chang et al., 2017). HR involves the exchange of genetic information from a homologous DNA template sequence to the site of DSB (Jasin and Rothstein, 2013; Piazza and Heyer, 2019). However, genetic information is also transiently deposited in the form of RNA. The transfer of genetic information from RNA to DNA is generally considered to be a specialized process exploited by mobile genetic elements and viruses (Goodier, 2016). We have previously shown that RNA can directly template repair of a DSB in an HR-dependent (*RAD52*-dependent) process, and that this process is inhibited by ribonuclease (RNase) H1 and H2, which function to degrade RNA-DNA hybrids. (Keskin et al., 2016; Keskin et al., 2014; Mazina et al., 2017; Meers et al., 2016). However, it was suggested that up to 8 % of the yeast genome is susceptible to RNA-DNA hybrid formation (Wahba et al., 2016). The formation of these hybrid sites is stimulated by various mutants often associated with RNA metabolism (Garcia-Muse and Aguilera, 2019) and/or transcription level, suggesting that many highly transcribed genomic loci are prone to RNA-DNA recombination events. In addition, mobile genetic elements can reverse transcribe transcript RNA to cDNA-intermediates used in DSB repair (Keskin et al., 2014; Meers et al., 2016; Morrish et al., 2002). However, little is known about the molecular mechanisms which support RNA-templated DSB repair and DNA modification, and which polymerase enzyme(s) can use RNA as a template to transfer genetic information to DNA *in vivo*. Here we uncover a major role of translesion DNA polymerase ζ in using RNA as a donor in RNA-DNA recombination. Overall, we characterized three mechanisms by which RNA can indirectly or directly template the repair of a DSB or modify genomic DNA in the absence of an induced DSB. We term these mechanisms cDNA-templated DSB repair (c-TDR), RNA-templated DSB repair (R-TDR), and RNA-templated DNA modification (R-TDM), respectively.

Results

Constitutive expression of the RNA donor promotes DSB repair by template RNA in *cis*

To better characterize the mechanism regulating RNA-templated DSB repair (R-TDR), we modified the assay that we previously developed to study DSB repair by transcript RNA in *cis* in haploid yeast cells (Keskin et al., 2016). The assay is based on induction of a DSB in an inactive auxotrophic marker gene (*his3*) located on chromosomal DNA, and detection of precise repair of the DSB mediated by a homologous transcript messenger RNA, which is the antisense RNA of the same *his3* gene. To drive transcription of the non-coding antisense *his3* RNA, which serves as homologous template for DSB repair, we replaced the galactose inducible promoter *pGAL1* with the constitutive translation elongation factor EF-1α promoter (*pTEF*) (Figure 1A) to make strains CM-278, 279 (Table S1). In this way, we hypothesized that the template RNA will be actively transcribed at the time of DSB induction, more accurately representing a DSB in an actively transcribed gene. In brief, the genetic assay comprises a *his3* gene interrupted by an artificial intron in the antisense orientation containing a homothallic switching (HO) endonuclease site and is driven by either the inducible *pGAL1* or the constitutive *pTEF* promoter from the antisense orientation (Figure 1A). Galactose induction of the HO endonuclease gene expressed by a *pGAL1*

promoter and located on chromosome III results in a DSB inside the artificial intron within the *his3* gene (Figure 1A). If the spliced antisense transcript RNA driven by either the *pGAL1* or *pTEF* promoter is used to template repair of the DSB by removing the artificial intron from its own *his3* DNA (in *cis*), a functional *HIS3* gene is formed producing histidine prototrophic (His^+) colonies.

Yeast cells can either directly use the antisense RNA as template for repair of the DSB in *his3*, or indirectly by converting the antisense *his3* RNA into a complementary DNA (cDNA) and use this as donor for DSB repair. Suppression of antisense transcription by deletion of its promoter or deletion of the 5'-splice site inside the intronic sequence prevents formation of His^+ colonies generated by either direct or indirect RNA repair of the DSB in *his3* (Keskin et al., 2014; Michelini et al., 2018). Because RNase H1 and H2 quickly degrade RNA/DNA hybrids required for DSB repair by RNA, cells expressing these enzymes repair the *his3* DSB predominately by converting the spliced antisense transcript into a complementary DNA (cDNA) donor for HR (Keskin et al., 2014). The cDNA-templated DSB repair (c-TDR) process was shown to depend on the activity of the yeast retrotransposon Ty because deletion of the *SPT3* gene, a positive regulator of Ty transcription, abolished c-TDR (Keskin et al., 2014) (Figure 1B). When we expressed the *his3* antisense donor from the constitutive promoter *pTEF* in wild-type RNase H cells, we found that the frequency of His^+ colonies dropped over a factor of ten when compared to reproduced results of the inducible *pGAL1* system (Figure 1B). This is likely due to lower expression level of *his3* RNA as illustrated by qRT-PCR results from *pTEF* vs. *pGAL1* (Figure S1A). Markedly, RNase H1 and H2-null cells (*rnh1 rnh201*) in the *pTEF* system show an increased frequency of His^+ colonies that is over a factor of one thousand relative to wild-type RNase H cells. Such high frequency is comparable to that obtained in the *pGAL1* system (Figure 1B). Upon deletion of *SPT3* in RNase H-null cells (*rnh1 rnh201 spt3*), in which R-TDR is the predominant mechanism of DSB repair, the frequency of His^+ colonies remain strikingly high in the *pTEF* system compared to reproduced results from the *pGAL1* system (Figure 1B). Differently from what was observed in wild-type RNase H cells, we found that the frequency of His^+ colonies does not diminish, but even increases (over a factor of three) when the antisense RNA is driven by *pTEF* vs. *pGAL1* in *rnh1 rnh201 spt3* cells (Figure 1B). This opposing trend of His^+ frequencies is not attributed to increased levels of antisense *his3* RNA in *rnh1 rnh201 spt3* cells because transcription levels of *his3* remain higher with *pGAL1* vs. *pTEF* in these mutant cells (Figure S1A). Transformation with DNA oligonucleotides (HIS3.F and HIS3.R, Table S2) designed to repair the DSB in the *his3* gene in a homology-driven manner showed only slight (less than 2-fold) increase of His^+ cells in the *pTEF* vs. *pGAL1* system (*rnh1 rnh201 spt3*) suggesting a possible increase in DNA breakage in the *pTEF* system (Figure S1B). These results suggest that continuous transcription of the antisense donor RNA from the constitutive promoter *pTEF* facilitates recombination of the broken DNA ends with the RNA template in *cis*, increasing the frequency of R-TDR. In contrast, high induction of transcription from *pGAL1* promoter stimulates production of cDNA-intermediates to repair DSBs via c-TDR.

R-TDR requires Rad52 and is independent of NHEJ proteins

Previous results using the inducible antisense *his3* RNA expressed under *pGAL1* showed that R-TDR requires the activity of the recombination enzyme Rad52 to promote inverse strand exchange between DNA and RNA. In contrast, the frequency of R-TDR increased in *rad51*-null or *rad59*-null cells, likely because these mutants suppress competition for DSB repair by the intact sister chromatid (Keskin et al., 2014; Mazina et al., 2017). Here, using the constitutive *his3* RNA expressed under *pTEF*, we found a factor of hundred decrease in the frequency of R-TDR in *spt3*-null cells in the absence of *RAD52* lacking *rnh1 rnh201*, while knock out of *RAD51* or *RAD59* showed no effect. Differently, c-TDR requires all these recombination proteins in the constitutive system (Figure 2A). Recent studies have suggested a possible role of RNA facilitating NHEJ (Chakraborty et al., 2016). In the *cis* assay with *pTEF* (Figure 1A), we found that elimination of NHEJ components (*KU70* or *DNL4*) markedly enhanced the frequency of His^+ colonies by c-TDR and R-TDR, as seen in single *ku70* or *dnl4* mutants and NHEJ mutants generated in a *rnh1 rnh201 spt3* background, but dramatically decreased cell viability following DSB induction (Figure 2B). NHEJ mutants in a *rnh1 rnh201 spt3* background show a factor of 2.2 and 1.8 increase with the *ku70* or *dnl4*-null mutant, respectively. This suggests that loss of NHEJ stimulates R-TDR. In addition, when comparisons are made between NHEJ mutants containing or lacking *SPT3* in a *rnh1 rnh201* background, a factor of 10 decrease is seen in *spt3*-null strains. This would suggest that loss of NHEJ stimulates both c-TDR and R-TDR. We also observed similar effects with the loss of the NHEJ DNA polymerase 4 (*POL4*) (see Figure 4A). *RAD50* variants *R520H T853I* and *R520H T853I D575G* discovered in a random mutagenesis screen to identify mutants with enhanced RNA-templated DSB repair ability showed similar reduced survival and increased frequency of R-TDR (Figure 2C). This is in line with inhibition of the NHEJ pathway. These results suggest that while it is possible that RNA facilitates NHEJ, NHEJ enzymes are not needed for R-TDR. This confirms that transcript RNA can work as a template for DSB repair in *cis* via an HR mechanism.

3’ non-homologous tail removal by Rad1–10 and Msh2–3 is dispensable for R-TDR

To further characterize molecular components of R-TDR, we investigated the role of DNA clippases, which are important players in DSB repair by HR (Ivanov and Haber, 1995; Lyndaker and Alani, 2009). We discovered that loss of *RAD1–10* strongly decreases the frequency of c-TDR but not R-TDR (Figure 2D). In wild-type RNase H cells, the frequency of His^+ colonies dramatically decreased upon knockout of *RAD1* or *RAD10*. While in *rnh1 rnh201 spt3* cells, in which repair is directly templated by transcript-RNA (R-TDR), knock out of *RAD1–10* shows only minor reduction in His^+ frequency. This suggests that the *RAD1–10* clippase is important for c-TDR but only has a minor impact on R-TDR (Figure 2D), possibly to remove the intronic 3′-DNA tail. Therefore, loss of either *RAD1–10* or *SPT3* (Keskin et al., 2014) allows separation of R-TDR from c-TDR, as both inhibit c-TDR. Similar to DSB repair by *his3* cDNA, DSB repair by a *his3* linear DNA molecule generated by PCR was strongly reduced in *rad1*-null cells (Figure S2). These results show that *RAD1–10* is important for DSB repair by a homologous DNA template provided in *trans* likely to remove 3′ tails of cDNA or PCR product that are not used in the HR repair process (Lyndaker and Alani, 2009). In line with the above results, we found that knock out of *MSH2* or *MSH3* mismatch repair (MMR) genes, which are reported to act in the recognition

of 3′ tails and recruitment of *RAD1–10* (Lyndaker and Alani, 2009; Sugawara et al., 1997), decreased the frequency of His^+ colonies in wild-type RNase H cells by a factor of ten. This reduction in the frequency of His^+ colonies was not seen for knock out of *MSH6* (Figure 2E), which functions in MMR but is not involved in 3′ tail removal (Paques and Haber, 1997). Markedly different results were obtained when these MMR genes were deleted in the *rnh1 rnh201* and *rnh1 rnh201 spt3* background. Knock out of *MSH3* or *MSH6* in a *rnh1 rnh201* background reduced the His^+ frequency by a factor of two, while deletion of *MSH2* reduced the frequency over a factor of nine (Figure 2E). However, in *rnh1 rnh201 spt3* background, only *MSH2* and *MSH6* deletion reduced the frequency of His^+ colonies, while *MSH3* deletion had no effect on R-DTR. Interestingly, the *msh6*-null mutation specifically reduced the frequency of R-TDR, possibly suggesting a mutagenic role of R-TDR. We conclude that *MSH3* is important solely in c-TDR, likely with support for *MSH2* and their role in removal of 3′ tails. However, *MSH2* and *MSH6* may play a role in the removal of mismatch bases possibly following reverse transcription.

R-TDR is independent of the Ty retrotransposon

To determine whether R-TDR requires the presence of the Ty retrotransposon for DNA repair synthesis by Ty reverse transcriptase (RT) on the RNA template, we cloned the *cis* system described in (Figure 1A), containing an I-SceI endonuclease cut site in place of the HO cut site, onto a yeast centromeric plasmid (HKb-67). We then engineered a *Saccharomyces paradoxus* strain lacking endogenous Ty activity (Garfinkel et al., 2005) (Ty-less, DG-2204, Table S1) with an integrated copy of the I-SceI endonuclease gene under the *pGAL1* promoter. In this strain (HK-692, 696), we constructed the *rnh1*, *rnh201*, or *rnh1 rnh201* null mutation(s). As a control, the same genetic engineering was done in a *S. cerevisiae* strain containing active Ty (FRO-767), in which we also constructed the knockout mutants *rnh1*, *rnh201*, or *rnh1 rnh201*, as well as *spt3* and *rnh1 rnh201 spt3* (Table S1). We then introduced the HKb-67 plasmid into these *S. paradoxus* and *S. cerevisiae* strains. The presence or absence of the Ty RT protein in the wild-type *S. cerevisiae* and *S. paradoxus* strains (HK-809, 812 and HK-815, 817) (Table S1) was confirmed by Western blot analysis (Figure S3A). *S. cerevisiae* and *S. paradoxus* cells of all the above genotypes containing the *cis* system on plasmid HKb-67 were plated on galactose medium to induce the DSB in *his3*. We then examined the frequency of His^+ colonies formed for these strains. Results obtained for all *S. cerevisiae* strains containing the plasmid-*cis* system with I-SceI endonuclease were in line with published data (Keskin et al., 2014) and data shown in Figure 1B. *S. paradoxus* Ty-less wild-type, *rnh1* or *rnh201* strains behaved similarly to the corresponding *S. cerevisiae* strains containing the *spt3*-null allele (Figure 3A). We were unable to detect any His^+ colony in these strains, demonstrating that c-TDR does not occur in Ty-less strains. On the contrary, His^+ colonies were detected in the *S. paradoxus* strain containing *rnh1 rnh201* mutations. This is similar to the *S. cerevisiae rnh1 rnh201 spt3* strain (Figure 3A). Transformation with DNA oligonucleotides (HIS3.F and HIS3.R, Table S2) designed to repair the DSB in the *his3* gene showed no significant difference for *rnh1 rnh201* cells compared to all other genotypes of *S. paradoxus* cells (Figure S3B and Table S10), demonstrating that the I-SceI DSB stimulates HR to similar extent in these cells. These results prove that while c-TDR requires Ty, the Ty retrotransposon is not required for R-TDR, and R-TDR must proceed with DNA polymerase.

To further validate the role of the Ty retrotransposon in c-TDR and not in R-TDR, we overexpressed a Ty1 element from the *pGAL1* promoter and introduced it in our strains either on a plasmid (pGTyClaI) or by integrating it into the yeast genome. We hypothesized that the overexpression of Ty would increase repair by cDNA but not RNA. We confirmed Ty overexpression by Western blot (Figure S3C). Indeed, overexpression of Ty from pGTyClaI in the *pGAL1* (Figure S3D) or *pTEF* (Figure S3E) system, and from the integrated Ty in the *pTEF* system (Figure 3B) strongly increased levels of c-TDR but had no impact on DSB repair by DNA oligos (Figure S3F). This suggests specificity of Ty RT to convert RNA into cDNA in our DSB repair assay but not to amplify the sequence of DNA oligos. Moreover, we observed that the overexpression of Ty did not promote R-TDR because it did not result in increased His^+ frequency in a *rnh1 rnh201 rad1* background that is deficient in repair by cDNA (Figure 3B). While c-TDR was inhibited over a factor of twenty-five by loss of *rad1* in wild-type cells when Ty was overexpressed, only less than a factor of two decrease of the His^+ frequency was observed in *rnh1 rnh201 rad1* cells overexpressing Ty (Figure 3B). This shows that even overexpression of Ty RT in a c-TDR inhibited background like *rad1*-null cannot stimulate R-TDR, and suggests that R-TDR proceeds without the aid of a bona fide RT. Overall, these results demonstrate that c-TDR is Ty RT-driven, while R-TDR does not proceed using a bona fide RT like Ty RT, and likely proceeds with a DNA polymerase.

DNA polymerase ζ promotes RNA-templated DSB repair

We next sought to determine which DNA polymerase(s) is responsible for R-TDR. DNA polymerase δ plays a major role in DSB repair by HR (McVey et al., 2016). Primer extension experiments showed that DNA Pol δ contains some RT activity but has low processivity on RNA templates (Storici et al., 2007). In yeast, there are four specialized polymerases that are associated with replication of damaged DNA: Pol4, η, ζ and deoxycytidyl transferase encoded by the *REV1* gene, which forms a complex with Pol ζ subunits (Kawasaki and Sugino, 2001) (Nelson et al., 1996). Pol4 works in NHEJ (McVey et al., 2016), while the translesion polymerases η, ζ and Rev1 can bypass a variety of unnatural or modified nucleotides, including ribonucleotide tracts in DNA (Lazzaro et al., 2012; Makarova and Burgers, 2015; Su et al., 2019). *RAD5* has recently been shown to recruit DNA Pol ζ to repair ssDNA gaps at stressed DNA replication forks (Gallo et al., 2019). Here we investigated the role of these specialized polymerases in R-TDR. Genetic disruptions showed that elimination of the translesion DNA polymerase ζ pathway (*rev3*, *rev1* or *rad5*-null mutation) drops the frequency of R-TDR by a factor of 2.5 in a *rnh1 rnh201 rad1* background, in which c-TDR is strongly impaired (Figure 4A). Loss of Pol η (*rad30-null*) shows no impact on R-TDR in the *rnh1 rnh201 rad1* background (Figure 4A). Surprisingly, elimination of *pol32*, an accessory subunit of Pol δ and ζ showed only minor decrease in His^+ frequency in the same genetic background. In line with results shown in (Figure 2B), knock out of the NHEJ-associated Pol4 (*pol4*-null) elevated the frequency of R-TDR by a factor of five (Figure 4A). Complementary results are seen in NHEJ mutants lacking *rev3* with a factor of 2.8 to 5.8 decrease in *dnl4* and *ku70* mutants, respectively (Figure 4B). In addition to results impairing c-TDR in a *rnh1 rnh201 rad1* background, we tested the impact of Pol ζ mutants in a *rnh1 rnh201 spt3* background. Similarly, we find that loss of Pol ζ results in a significant decrease in the frequency of His^+ colonies (Figure 4C).

These findings support a predominant role of Pol ζ in R-TDR. To examine whether Pol ζ catalytic activity was responsible for R-TDR, we constructed a low fidelity mutant of the catalytic subunit *rev3* L979F (Stone et al., 2009). The low fidelity mutant of Pol ζ showed a minor decrease in the frequency of repair as compared to wild-type Pol ζ (Figure 4D). To examine whether overexpression of the wild-type catalytic subunit of Pol ζ, *REV3*, would increase the frequency of R-TDR, we integrated a copy of the *REV3* gene under the galactose inducible promoter *pGAL1* at the *CAN1* locus in *rnh1 rnh201* cells (Table S1). Indeed, overexpression of *REV3* gene resulted in a modest but significant increase in the frequency of His^+ colonies (Figure 4E). These data show that overexpression of the catalytic subunit of Pol ζ promotes R-TDR. Overall, our results support a dominant function of Pol ζ in R-TDR.

Rad52-independent RNA-DNA recombination in the absence of an induced DSB

Upon replacement of the inducible *pGAL1* with the constitutive *pTEF* promoter, we unexpectedly detected abundant formation of His^+ colonies in the *rnh1 rnh201*, *rnh1 rnh201 spt3* and *rnh1 rnh201 rad1* backgrounds without induction of the DSB (Figure 5A).His^+ colonies were detected by growing cells in glucose containing medium, suppressing activation of the HO endonuclease driven by the *pGAL1* promoter. This observation was not the result of leaky HO endonuclease expression in glucose medium as knock out of the *HO* endonuclease gene did not reduce the frequency of His^+ colonies in similar conditions (Figure 5B). We termed this mechanism RNA-templated DNA modification (R-TDM). In addition, this repair is dependent on the RNA template as loss of either the antisense promoter or removal of the intronic branch site strongly reduces the frequency of His^+ colonies (Figures 5B and S4). We hypothesized that formation of an RNA-DNA hybrid between the antisense *his3* transcript and the *his3* DNA gene may recruit MMR, nucleotide excision repair (NER) or base excision repair (BER) nucleases resulting in cleavage within the *his3* locus, and recombination by RNA. We investigated the role of previously reported nucleases (*MLH1* and *RAD1*) involved in processing of R-loops (Freudenreich, 2018; Sollier et al., 2014; Su and Freudenreich, 2017), along with other nucleases involved in DNA repair and found no difference in the frequency of His^+ colonies (Figure 5C).

We then examined the impact of two DNA damaging agents in their ability to stimulate R-TDM at the *his3* locus, bleomycin, which generates DSB (Moseley, 1989), and methyl-methanesulfonate (MMS), which results in base alkylation damage that can be processed to single or double-stranded breaks (Ma et al., 2008). While bleomycin was not found to stimulate R-TDM, possibly due to rare DSB formation mainly outside of the *his3* locus, treatment with MMS showed significant increase in the frequency of His^+ colonies (Figure 5D). Importantly, these His^+ colonies were not due to mutations of *his3*, e.g. generating new splice sites in the sense orientation, allowing splicing from the sense transcript to produce the His^+ phenotype while the DNA retained the intron. In fact, all His^+ clones examined had precisely lost the intron (20/20) in the DNA and 19/20 had perfect His^+ sequence. One sample (1/20) had a G to C transversion (typical of MMS mutagenesis) resulting in V79L mutation but also had precisely removed the intron. These results support a role of either spontaneous breaks or breaks that are occurring during the repair of DNA damage as triggers for R-TDM.

Like R-TDR, R-TDM is independent of cDNA-mediated repair because loss of *SPT3* or *RAD1* does not decrease the His^+ frequency in the absence of the induced DSB in *his3* (Figures 5B and 5C). Additionally, overexpression of Ty in strains lacking the *HO* gene (*ho*-null) did not result in increased frequency of His^+ colonies both in the wild-type and *rnh1 rnh201* backgrounds (Figure 5E). We hypothesized that the R-TDM mechanism, which is independent of a DSB, requires formation of an RNA-DNA hybrid (R-loops) during transcription (Rondon and Aguilera, 2019). Rad52 is strongly required for R-TDR by promoting formation of RNA-DNA heteroduplex in an inverse RNA-strand exchange reaction by forming an active complex with dsDNA in the proximity of the DNA DSB (Mazina et al., 2017). In contrast, to promote R-loop formation, Rad52 needs to form an active complex with RNA and to promote formation of RNA-DNA heteroduplex in intact dsDNA without breaks. We tested the ability of Rad52 to stimulate R-loop formation *in vitro* and found that while Rad52 can promote D-loop formation with DNA, it cannot promote R-loop formation with RNA (Figure 6). In line with the biochemical data, loss of *rad52* was inconsequential to the frequency of R-TDM in the yeast cells (Figure 5B). Surprisingly, we found that elimination of *rad59* but not *rad51* resulted in decrease but not elimination of R-TDM (Figure 5B). These results demonstrate that in the absence of an induced DSB, transcript RNA has the capacity to recombine with homologous DNA sequences and mediate DNA modifications even without the catalytic support of a recombination protein. The data also uncover a novel mechanism of Rad52-independent recombination in yeast.

DNA polymerase ζ is essential for RNA-DNA recombination in the absence of an induced DSB

While Ty RT is not required for R-TDM, as shown by knockout of *SPT3* or overexpression of Ty (Figure 5B and 5E), Pol ζ is essential. Knockout of the *REV3* gene strongly reduces the frequency of His^+ colonies in *rnh1 rnh201 spt3* cells grown in galactose (Figures 4A). In line with R-TDR results, loss of the DNA Pol ζ translesion synthesis pathways (*rev1*, *rev3* or *rad5*-null) strongly reduces the frequency of R-TDM (Figure 5F). Furthermore, the low fidelity DNA synthesis derivative *rev3*-L979F strongly diminishes the frequency of His^+ colonies in *rnh1 rnh201* cells grown in glucose with no DSB induction (Figure 5F). Sequencing of the *HIS3* locus from several His^+ colonies isolated from wild-type *REV3* found 53/54 with correct *HIS3* sequence. In contrast, only 40/54 in the *rev3*-L979F mutant had correct *HIS3* sequence (Table S8) showing a significant increase in the frequency of mutation (*P*-value 0.004 obtained using Mann-Whitney *U*-test). This result supports a synthesis role by Pol z in R-TDM. Lastly, overexpression of the *REV3* gene coding for the catalytic subunit of Pol ζ results in an increase (a factor of two) in the frequency of His^+ colonies in *rnh1 rnh201* cells lacking the *HO* gene (*ho*-null) (Figure 5G). Overall, these findings demonstrate an essential function of Pol z in R-TDM, and point to Pol ζ being the RT in both R-TDM and R-TDR.

Discussion

Mechanism of RNA-templated DSB repair (R-TDR) driven by Pol ζ

Understanding the potential of RNA to recombine with DNA has been difficult given RNA sequence resemblance to the DNA sequence from which it is generated. However, large

scale RNA-mediated genome rearrangements have been observed in ciliates (Burns et al., 2016; Nowacki et al., 2008). In addition, it has been proposed that RNA-templated mechanisms may drive somatic hypermutation (Steele, 2017). We previously demonstrated that RNA is recombinogenic in the absence of RNase H1 and H2 and is aided by Rad52 through RNA-DNA annealing or an inverse RNA strand exchange with homologous dsDNA ends (Keskin et al., 2014; Mazina et al., 2017). In addition, this process does not require extensive end resection (Mazina et al., 2017). Here we expanded on this model to find limited dependence on Rad1–10, Msh2–3 clippases suggesting R-TDR can proceed with limited 3′-nonhomologous tail removal. We hypothesize that extensive resection is favored in c-TDR facilitating recombination between the broken DNA ends and the cDNA molecule. In contrast, during R-TDR, extensive resection is not needed as the donor RNA is already localized to the DSB site and can efficiently form RNA:DNA heteroduplex even at blunt ended DNA structures by Rad52 in an inverse RNA strand exchange (Mazina et al., 2017). Removal of the 3′ flap generated in the artificial intron could be attributed to limited tail excision by the 3'-to-5' proofreading activity of DNA polymerase δ (Paques and Haber, 1997). We exploit this difference in the preference for more extensive end clipping for c-TDR vs. limited end clipping in R-TDR to distinguish between these repair pathways. This is in conjunction with *spt3*-null mutants, which inhibits transcription of endogenous Ty elements. Recently, Rad52 was shown to also limit resection in budding yeast (Yan et al., 2019), which may aid in R-TDR. Surprisingly, we found little effect of *rad59*-null or *rad51*-null mutants in R-TDR when the antisense RNA is driven by the *pTEF* promoter. This is contrary to our previous work exploiting the galactose inducible donor system showing a stimulation in the frequency of R-TDR in the absence of either *rad51* or *rad59* (Keskin et al., 2014; Mazina et al., 2017). We propose that loss of either *rad51* or *rad59* inhibits competition with sister chromatid recombination. However, this competition may be less favorable for the sister chromatid in the *pTEF* constitutive system, as the donor RNA is already actively transcribed at the time the DSB is induced. In contrast to effects seen in HR mutants, NHEJ mutants show a dramatically increased frequency of R-TDR events but a marked loss of cell viability following DSB induction.

Remarkably, we discovered a unique role of translesion DNA polymerase ζ in mediating R-TDR. The current understanding of DNA translesion synthesis pathways in yeast suggests that replication stalling results in ubiquitination of the PCNA clamp loader, switching replicative polymerases to translesion polymerases (Pol η, ζ) and bypass of the damaged DNA (Martin and Wood, 2019; Plosky and Woodgate, 2004). We find reductions in the frequency of R-TDR in the absence of translesion polymerases, predominantly DNA Pol ζ in *rnh1 rnh201 rad1* and *rnh1 rnh201 spt3* backgrounds, in which c-TDR is inactive. Recombination-associated DNA synthesis is largely considered to proceed with the high fidelity replicative polymerase δ (Holmes and Haber, 1999; Li et al., 2009; Maloisel et al., 2008) but lagging strand DNA polymerase α and translesion DNA polymerases are also associated with DSB repair (Hirano and Sugimoto, 2006; Holmes and Haber, 1999; Rattray et al., 2002; Sneeden et al., 2013). Earlier work has shown that yeast replicative polymerases (δ and α) contain minimal RT activity (Storici et al., 2007), in addition to documented RT activity of human Pol η (Su et al., 2019) and yeast Pol ζ in the bypass of multiple embedded ribonucleotides in DNA (Lazzaro et al., 2012). Beyond the role of yeast Pol ζ in the bypass

of multiple embedded ribonucleotides, we provide *in vivo* data suggesting Pol ζ can reverse transcribe transcript RNA at sites of DNA damage. However, the fidelity of DNA polymerase ζ is known to be significantly lower than that of replicative polymerases with a preference for base substitutions (Zhong et al., 2006). Interestingly, this supports our results showing that *MSH6* is important for R-TDR, possibly in the repair of mismatches generated by DNA Pol ζ, but not c-TDR, which is Pol ζ-independent. This may suggest a high mutagenic nature of R-TDR, possibly in part explaining the higher mutation rate seen in transcribed regions (Jinks-Robertson and Bhagwat, 2014). We propose that during R-TDR, DNA polymerase δ encounters donor RNA annealed to the 3′ end of the DSB and this results in polymerase switching to DNA Pol ζ providing increased RT capabilities driving repair of the DSB aided directly by an RNA template (Figure 7A).

Mechanism of RNA-mediated DNA modification (R-TDM)

In addition to the ability of RNA to transfer information back to DNA following a DSB, we found that RNA can transfer genetic information back to DNA in the absence of RNase H1 and H2 without the induction of a DSB. This process of R-TDM is independent of the major HR protein in yeast, Rad52, highlighting a unique form of recombination in yeast. Rad52-independent mitotic recombination events have previously been observed in budding yeast with gene conversion events partially dependent on Rad59 (Coic et al., 2008; Haber and Hearn, 1985). R-TDM may explain these Rad52-independent mitotic gene conversion events. We hypothesize that in the absence of RNase H1 and H2, R-loops form with the transcript RNA and can recombine with DNA in a Rad52-independent manner (Figure 7B). This RNA-DNA recombination in the absence of an induced DSB is strongly dependent on DNA Pol ζ translesion synthesis pathway (*REV1*, *REV3* and *RAD5*). We found no role of nucleases involved in cleavage of R-loops structures (*RAD1, MSH2, RAD2* or *MLH1*) (Su and Freudenreich, 2017). Results following MMS treatment suggest that either spontaneous nicks/DSBs or breaks occurring during the repair of DNA damage may trigger initiation of R-TDM. In addition, reports have suggested a possible role of R-loops in the initiation of origin-independent replication events in *E. coli* (Wimberly et al., 2013), yeast (Stuckey et al., 2015) and human mitochondria (Posse et al., 2019). It is possible that the donor RNA or fragments of the donor RNA may be incorporated during DNA synthesis leading to RNA-mediated DNA modification events similar to oligonucleotide incorporation during replication (Rodriguez et al., 2012), but different from the inclusion of ribonucleotides by DNA polymerases during DNA synthesis (Nava et al., 2020). Altogether, our findings that active transcription from the *pTEF* constitutive system in a RNase H-defective background promotes R-TDR and R-TDM highlight the intriguing possibility that loci susceptible to RNA/DNA hybrid or R-loop formation, because of local elevated transcription levels (El Hage et al., 2014; Wahba et al., 2016) or other genetic or environmental factors (Figure 5D) (Chan et al., 2014; Garcia-Muse and Aguilera, 2019), may be genomic areas prone to both R-TDR or R-TDM.

Mechanism of cDNA-templated DSB repair (c-TDR)

Mobile genetic elements and their impact on genome diversification are areas of intense investigation (Feschotte and Pritham, 2007). However, information on how these mobile elements affect genome stability is scarce. Reports have demonstrated cDNA-mediated DSB

repair events in yeast, mice and human cells (Derr et al., 1991; Keskin et al., 2014; Melamed et al., 1992; Ono et al., 2015; Onozawa et al., 2014). Others have shown a role of cDNA recombination events as drivers of copy number variations in human neuronal cells, leading to the hypothesis of a "recording" and "playback" of preferred gene variants (Lee et al., 2018). Previous reports have demonstrated strong requirements for *RAD52* in cDNA-mediated recombination between Ty elements in yeast but found that *RAD1* was not important for this process (Nevo-Caspi and Kupiec, 1996). This is contrary to our results showing a strong dependence on *RAD1–10* for cDNA-templated HR between genomic *his3* and *HIS3* cDNA. We also find a strong dependence on *MSH2–3* for cDNA-templated DSB repair but not *MSH6*. This likely highlights a requirement for removal of nonhomologous ends during HR. Furthermore, we found that increased expression of the template RNA from the *pGAL1* system stimulated the frequency of c-TDR, uncovering a role of elevated transcriptional level coupled with cDNA synthesis by Ty retrotransposon in c-TDR. In addition, we show also overexpression of Ty1 stimulates DSB repair at the *his3* locus via c-TDR. However, it is unknown how this *his3* RNA is captured and reverse transcribed by Ty. This may resemble events during early embryogenesis when genome-wide demethylation and derepression of retroelements occurs (Reik et al., 2001; Surani, 2001), during which retroelement and cDNA insertions at CRISPR/Cas9 editing site in mice embryos are observed (Jeon et al., 2019). Studies of Ty RT have indicated Ty1 RT transferring from normal Ty1 template ends to various tRNA templates (Mules et al., 1998). This has also been reported for non-LTR retroelements like those of LINE-1 elements in humans to poly-A tracts (Dombroski et al., 1994; Esnault et al., 2000). Therefore, c-TDR may require removal of much longer non-homologous tails than R-TDR, possibly explaining its stronger dependence on the activity of *RAD1–10* and *MSH2–3* (Figure 7C). It would be interesting to understand the mechanism and rules which govern cDNA amplification of genomic transcripts, and their role on genome in/stability.

Limitations

While the findings of this study rely on the removal of RNase H1 and H2 to detect both R-TDM and R-TDR, these processes may also occur in genetic loci prone to RNA/DNA hybrid formation, as described above, and are likely underrepresented in this study because the genetic assay employed requires a recombination event between a spliced antisense transcript RNA and DNA. Intronic sequences have been shown to prevent R-loops formation in highly expressed genes in both humans and yeast (Bonnet et al., 2017). Moreover, recombination between the *his3* sense transcript RNA containing the inverted intron, or the unspliced *his3* antisense transcript and the *his3* DNA can also occur but cannot be detected in our genetic assay. While transcript RNA from the antisense transcript RNA is initiating DSB repair in wild-type cells of our *cis* system, the repair process is likely not fast enough to be completed before RNase H1 and H2 start cleaving the RNA annealed to the broken DNA ends. Our inability to detect DSB repair by transcript RNA in wild-type cells of our *cis* system is not due to lack of such repair mechanism when RNase H1 and H2 are functional, but possibly to some limitations of the experimental system we adopted. Thus, we believe that our results underestimate the ability of RNA to recombine with DNA.

Continuous advancement in technology for genome engineering may help unravel the relationship between RNA and DNA in genome stability. By exploiting the efficient and modular clustered regularly interspaced short palindromic repeats (CRISPR) and the CRISPR-associated protein 9 (Cas9) system (Makarova et al., 2020), it will be possible to develop novel genetic designs that can help to better understand the properties of R-TDR, R-TDM and c-TDR. Our work demonstrates that constitutive transcription before DSB induction significantly enhances R-TDR in *rnh1 rnh201 spt3* cells. In addition, we find that R-TDR and R-TDM may be a mutagenic process driven by Pol ζ, and hypothesize that small bits of information like single nucleotide polymorphisms may transfer from RNA to DNA. New site-specific DNA/RNA editing, and mutagenesis techniques will be of great interest for exploring this possibility. Overall, we show that RNA molecules can modify a genomic DNA sequence at a homologous genomic locus in yeast DNA through a variety of different mechanisms interacting with multiple DNA repair pathways. Our findings illustrate a powerful role of RNA in directly (R-TDR and R-TDM) and indirectly (c-TDR) templating genomic modifications as a driver of genome in/stability.

STAR★Methods

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Francesca Storici (storici@gatech.edu).

Materials Availability—All unique/stable reagents generated in this study are available from the Lead Contact.

Data and Code Availability—Customized shell scripts for variant calling are available on GitHub at https://github.com/agombolay/Variant-Calling. The DNA-seq dataset generated during the current study is available in NCBI's Sequence Read Archive via BioProject "PRJNA656525". All data generated in this study are available from the Lead Contact.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Strain Construction—The yeast strains used in this work are listed in Table S1 and derive from FRO-767 (Storici et al., 2007). This strain contains the site-specific homothallic switching endonuclease in the middle of the *LEU2* gene on chromosome III under the galactose inducible promoter (*pGAL1*). We developed an experiment yeast system consisting of a *his3* gene located on chromosome III containing in artificial intron in the antisense orientation with an homothallic switching endonuclease driven by expression of either *pGAL1* (Keskin et al., 2014) or the constitutive translational elongation factor EF-1α *pTEF* promoter. The native *HIS3* promoter drives transcription from the sense orientation but does not result in functional His3 protein as the artificial intron inserted in the antisense orientation cannot be splicing in the sense orientation. The yeast cells are auxotrophic for histidine (His^-) and do not grow on media without histidine. Following galactose induction of the homothallic switching endonuclease and subsequent double-stranded break (DSB) inside of the artificial intron, if the full length antisense *his3* transcript is used for repair of the DSB (in *cis*), this will result in functional *HIS3* gene and protein and growth on media

lacking histidine. Accurate repair of functional *HIS3* by ligation of the broken ends at the exon-exon junction via non-homologous end joining (NHEJ) is inefficient in this system (<0.1 out of 10^7 viable cells) (data not shown). Strains CM-278, CM-279, CM-280, CM-281, CM-282, CM-283, CM-284, CM-286 were derived from YS-526, YS-527, YS-528, YS-529, YS-530, YS-531 as detailed in Table S1, by single-step replacement of the *Kluyveromyces lactis URA3* (*KlURA3*) marker gene with the promoter of *TEF1* EF-1α amplified with short flanking homologies from plasmid p414-TEF-Cas9 (DiCarlo et al., 2013). Deletion mutants derived from CM-278, CM-279, CM-280, CM-281, CM-282, CM-283, CM-284, CM-286 were performed by single-step replacement of the opening reading frame of the gene of choice with either the *kanMX4*, *hygMX4*, *natMX4* and/or *KlURA3* unless otherwise indicated and confirmed by PCR. All site-specific modifications or insertions were confirmed by PCR and sequenced. *Saccharomyces paradoxus* strains derive from DG-2204, kindly provided by Dr. David Garfinkel (Garfinkel et al., 2005). HK-692, HK-696 were constructed by the *delitto perfetto* method (Storici et al., 2001) by inserting a GSKU (GAL1-*I-SceI KanMx4 KlURA3*) cassette with long homology arms to the *LEU2* locus, disrupting *leu2*. The *KanMx4* and *klURA3* markers of the GSKU cassette were then removed leaving Gal-I-SceI inserted in the *leu2* locus. Deletion mutants derived from HK-692 and HK696 were made as described previously by single-step replacement of the open reading frame of *RNH1* and/or *RNH201*. YCp50pK-Gal-*his3*-AI-*ISce-I* was constructed by PCR amplification of Gal-*his3*-AI-*ISce-I* from genomic DNA (HK-654) with primers adding EcoRI and MluI restriction sites at the ends and subsequent cloning of the EcoRI and MluI-digested PCR product into YCp50pK. The construct was verified by sequencing. YCp50pK-Gal-his3-AI-ISce-I plasmid was used to detect RNA-templated DSB repair events in *S. paradoxus* and *S. cerevisiae* strains in experiments shown in Figure 4B. *S. cerevisiae* FRO-767 strain was used to insert Gal-I-SceI after pop-out of a GSKU cassette at the *leu2* locus, as described above for *S. paradoxus* strains. Strains HK-687 and HK-688 were generated. Successive replacement of *HIS3* ORF with *TRP1* generated strains HK-699 and HK-701, in which all deletion mutants were constructed by replacement of chosen opening reading frames with either the *kanMX4*, *hygMX4* and/or *natMX4* marker gene. Integrated Ty overexpression strains were constructed via the *delitto perfetto* approach. The CORE cassette was inserted into the *CAN1* locus (CM-1099 and CM-1100) and removed by PCR amplified fragments containing either *pGAL1* from BDG102 plasmid or *pGAL1-Ty1* from pGTyClaI plasmid, which were kindly provided by Dr. David Garfinkel (Garfinkel et al., 1988), to construct CM-1093, CM-1095, CM-1099 and CM1100, respectively. Deletion mutants derived from these strains were constructed by replacement of the open reading frame of the gene of choice with either the *kanMX4*, *hygMX4*, *natMX4* and/or *KlURA3*. The *rad50* mutants were created by the *delitto perfetto* method by insertion of a CORE cassette into *rad50* to generate CM-1352, and successive CORE replacement with a PCR product containing the R520H and T853I mutations or the R520H, T853I and D575G mutations to generate CM-1370 and CM-1372 respectively. The constructs were confirmed by sequence analysis. The *rev3* L979F mutant was also created by the *delitto perfetto* method by insertion of a CORE cassette into the *REV3* gene to generate CM-1165 and CM-1166. The CORE cassette was replaced with the sequence of a dsDNA oligonucleotide containing the L979F mutation. The constructs were verified by sequencing. The *pGAL1-REV3* integration strains were constructed by replacement of opening reading frame of *rnh1*

with *natMX4* and *rnh201* with *hygMX4* generating CM-1173 and CM-1175. The CORE cassette was then replaced with *pGAL1-REV3* and confirmed by PCR and sequencing.

Media preparation—Synthetic dropout, rich YPD (1 % yeast extract, 2 % peptone, 2 % dextrose) and YPGal (1 % yeast extract, 2 % peptone, 2 % galactose) solid and liquid media have been prepared according to standard protocols. Liquid YPLac (1 % yeast extract, 2 % peptone, 2.7 % (v/v) lactic acid)

METHOD DETAILS

Fluctuation Assay—Quantitative fluctuation assay in liquid culture was performed to determine DNA repair frequencies as indicated by histidine prototrophic growth. Selected strains were grown in flasks of 50 mL YPLac liquid medium for 24h at 30 °C. The density of the cultures was determined by counting cells using a hemocytometer and various concentrations of cells were plated depending on the genotype. In general, per each fluctuation assay, 10^3 cells per sample were plated to YPD solid medium to determine cell survival before DSB induction, grown for 2 days at 30 °C and counted. $10^{5–6}$ cells per sample were plated to YPGal medium to determine survival after DSB induction, grown for 2–3 days at 30 °C and counted. In strains lacking *rnh1 rnh201*, $10^{7–8}$ cells were plated to His^- medium to determine the number of initial His^+ cells before DSB induction, grown for 2 days at 30 °C and counted. $10^{7–8}$ cells were plated to YPGal, grown for 2 days at 30 °C and subsequently replica plated to medium lacking histidine, grown for 2–4 days at 30°C and counted to determine the number of His^+ colonies. The frequency of RNA-mediated repair was calculated by dividing the number of His^+ colonies on His^- medium by the number of colonies on YPGal medium and normalized to 10^6 or 10^7 viable cells. The survival was calculated by dividing the number of colonies grown on YPGal medium by the number of cells plated on the same medium.

For experiments using either empty vector (BDG102) or Ty overexpression vector (pGTyClaI), fluctuation assays were performed as described above but cells were grown in 50 mL Ura^-Lac medium instead of YPLac medium to maintain plasmid selection. Cells were then plated to Ura^- medium instead of YPD to determine survival, and on Ura^- Gal medium to induce the DSB. 10^7 or 10^8 cells plates on Ura^-Gal were replica plated to His^- medium to determine the repair frequency. The frequency of His^+ colonies was calculated by dividing the number of His^+ colonies on His^- medium by the number of colonies on Ura^- Gal medium and normalizing to 10^7 viable cells. The survival was calculated by dividing the number of colonies on SC-Ura^- Gal medium by the number of cells plated on the same medium.

Experiments using the Ty-less strains were conducted with a plasmid carrying *his3* cassette for the *cis* assay (YCp50pK-Gal-his3-AI-I-SceI). The strains were transformed with YCp50pK-Gal-his3-AI-I-SceI and transformant cells were selected on Ura^- medium. Transformant strains were grown in flasks of 50 mL YPLac liquid medium for 24h at 30 °C. The density of the cultures was determined by counting cells using a hemocytometer and various concentrations of cells were plated depending on genotype. 10^3 cells were plated to YPD and Ura^- media and grown for 2 days at 30 °C to determine survival and plasmid

stability. 10^4 cells were plated to YPGal medium and grown for 2 days at 30 °C to determine survival frequency following DSB. 10^7 cells were plated to YPGal and grown for 2 days at 30 °C and replica plated to His^- medium to determine the frequency of RNA-mediated DSB repair events. The frequency of RNA-mediated repair was calculated by dividing the number of His^+ colonies grown on His^- medium by the number of colonies on YPGal medium and normalizing to 10^7 viable cells. The survival was calculated by dividing the number of colonies on YPGal medium by the number of cells plated on the same medium.

To determine the frequency of RNA-mediated DNA modification events without a DSB, experiments were conducted by either deleting the homothallic switching (*HO*) endonuclease gene or were grown in glucose to repress *HO* transcription. Results from deletion or repression of *HO* were found to be similar. Cells were grown in flasks of 50 mL YPLac liquid medium for 24h at 30 °C. The density of the cultures was determined by counting cells using a hemocytometer and 10^7 cells were plated to His^- medium. The frequency of RNA-mediated repair was calculated by dividing the number of His^+ colonies grown on His^- medium by the number of colonies on YPD medium and normalizing to 10^7 viable cells.

Mutagenesis Screen—An ethyl methanesulfonate (EMS) random mutagenesis screen was carried out in *rnh1 rnh201 spt3* cells of YS-486 and YS-487 containing the galactose inducible RNA-templated DSB repair genetic assay. Cells were treated with 1 % EMS, counted and plated to YPD. Individual colonies were transferred to 96-well microtiter plate with 200μL of water and then transferred to YPD, YPGal and His^- medium with non-mutagenized *rnh1 rnh201 spt3* cells as a control. Cells were grown for 2 days at 30 °C. The YPGal plate was then replica plated to His^- medium. Isolates showing a higher frequency of His^+ papillae were isolated and used in a fluctuation assay to determine the frequency of DSB repair by RNA. Genomic DNA was extracted from mutants showing a higher frequency of His^+ colonies relative to the non-mutagenized strain using the QIAGEN Genomic DNA Buffer Set. Next generation sequencing libraries were prepared using the Nextera XT DNA Library Preparation Kit. The library in which we found the *RAD50* mutations was termed CM10. HiSeq 2500 with Rapid Run mode supporting paired-end (2 × 100 cycle) sequencing was performed and sequenced reads were trimmed based on quality with Trim Galore (Martin, 2011). The reads were then aligned to the reference yeast genome (sacCer3) using Bowtie2 (Langmead and Salzberg, 2012). Based on the aligned reads, variants were called using GATK (Van der Auwera et al., 2013). Variants were filtered based on quality (>= 30) and depth of coverage (>= 25). Variants present in both the control (sample NOT treated with EMS) and all cases (samples treated with EMS) were excluded from the analysis. Variants with mutations in genes associated with DNA repair (e.g. *rad50 R520H T853I*) were independently constructed by *delitto perfetto* method in strain CM-280 and CM-282. Frequencies of variants were calculated by fluctuation assay described above. *rad50 D575G* is the result of error during PCR amplification during construction of *rnh1 rnh201 rad50 R520H T853I* mutants.

Mutagen Assay—To determine if chemical mutagens (Bleomycin or Methyl-methanesulfonate) could stimulate R-TDM, cells were inoculated into 200 mL YPD liquid

medium for 4h at 30 °C. 50 mL of cells were transferred to individual tubes and chemical mutagens were added at respective concentrations and grown for 18h at 30 °C. Cells were plated to His^- and YPD medium to determine survival and the frequency of His^+ colonies formed.

Oligonucleotide and PCR product transformations—Transformation by oligonucleotides (1 nmol) was performed as described (Storici et al., 2007). For experiments presented in Figures S1B and S3D oligonucleotides HIS3.F and HIS3.R (Table S2) were used; only oligo HIS3.F for used for experiments presented in Figure S3A. Induction of the homothallic switching endonuclease DSB was done by incubating cells in 2 % galactose medium for 3 h and plating cells to His^- medium to determine repair frequencies. 6 μg of PCR product with homology to *HIS3*, generated using primers HIS3.205F and HIS3.205R, were used in transformation experiments presented in Figure S2.

Western Blot Analysis—Whole cell protein extracts were isolated by collecting $5 \times 10^8 – 1 \times 10^9$ cells per sample. The *S. paradoxus* and *S. cerevisiae* cultures were grown in YPD media to OD600 0.5, whereas for overexpression experiments containing *S. cerevisiae* cells expressing Ty under the *pGAL1* promoter cells were grown in YPLac medium overnight to OD600 0.3. Then, galactose at 2 % (v/w) final concentration was added to the medium and cells were shacked in the incubator at 30 °C for 6h (Keskin et al., 2014). After harvesting by centrifugation, cells were washed with PBS (Corning, 21–040-CV) and kept at −80 °C until processing. Cells were resuspended in lysis buffer [25 mM TRIS pH 7.5, 100 mM NaCl, 1 mM DTT, 10 mM EDTA, 1X protease inhibitor cocktail (Roche, 11836170001) 0.05 % NP-40 (Thermo, 28324)] on ice and ~100 μl glass beads were added before bead beating using the Genie disruptor machine (1 minute shaking, 1 minute rest, 5 times in cold room). The supernatant was carefully removed, centrifuged and the concentrations determined by Bradford reagent (Biorad, 500–0006) and BSA standards (Biorad, 500–0207) and a spectrophotometer. Successively, 5 μg protein extract was loaded on a 12 % TRIS-glycine SDS PAGE gel. After running, the gel was blotted on an Amersham nitrocellulose membrane (GE, 10600003) overnight at 4 °C. The membrane was blocked with 5 % milk and TRIS buffered saline with tween for 1h at room temperature, then either Actin (ab170325) or B8 antibodies (Garfinkel et al., 1991) were added in 1:3000 dilution in 5 % milk + TBST and incubated overnight at 4 °C. The membrane was washed 3 times for 10 minutes with TBST and respective secondary antibodies (Thermo, 31460, 31430) were added in 1:5000 dilution in 5 % milk + TBST and incubated at room temperature for 1h. The membranes were washed again 3 times and developed using ECL Western Blotting solution (Thermo, 32106) according to the manufacturer's instructions.

RNA isolation and Gene Expression measurements—Total RNA was isolated using the Hot Phenol method (Xue et al., 2004). Approximately $3 \times 10^8 – 5 \times 10^8$ cells were collected by centrifugation and washed once with ice cold PBS and stored at −80 TBST °C overnight. Cells were then resuspended in 500 μl TES solution (10 mM TRIS-HCl, pH 7.5, 10 mM EDTA, 0.5 % SDS) and 500 μl acid phenol (VWR, 0981.400ML) was added. Samples were vortexed vigorously and incubated at 65 °C for 1h with brief vortexing every 15 minutes. Cells were chilled on ice for 5 minutes, then centrifuged for 5 minutes at full

speed on a tabletop centrifuge. The upper aqueous supernatants were transferred to clean RNase free tubes (Fisher, AM12450) and extracted twice by 500 µl chloroform. Total RNA was precipitated at −80 °C by sodium acetate (3M, 1/10 volume) and ethanol (2.5 volume), and washed by pre-chilled 70 % ethanol. RNA pellets were resuspended in RNase free water and purified using RNeasy mini kit (Qiagen, 74104) according to the manufacturer's instructions. This was followed by DNase treatment using the Turbo DNA-free kit (Fisher, AM1907) according to the manufacturer's instructions. RNA concentrations were determined using a Nanodrop 1000 machine and 1 µg of total RNA was used to reverse transcribe to cDNA using the iScript cDNA synthesis kit (Biorad, 1708891) according to the manufacturer's instructions. The final product was diluted 10 times to 200 µl and 2 µl of this product was used as template for each qPCR reaction. Quantitative real time measurements were performed according to the manufacturer's recommendations using the 96-well Step One Plus Real Time PCR system (Fisher, 4376600), SYBR select master mix (Fisher, 4472918) and the primers HIS3Q.1, HIS3Q.3 for *his3* and ACT1Q.F, ACT1Q.R for *ACT1* shown in Table S2. Relative fold changes were determined by normalizing to the actin levels of non-induced *pTEF* system samples for each background. At least two duplicate measurements were performed for at least two biological repeats and error bars show the standard deviation for each sample (Banyai et al., 2016).

***In-vitro* D-loop and R-loop assay—**ScRad52 (450 nM) was incubated with a 48-mer 32P-labeled ssDNA (no. 211; 3 µM, nt) or a 48-mer RNA (no. 501; 3 µM, nt) of identical sequence in buffer containing 25 mM TRIS·acetate (pH 7.5), 2 mM DTT, 0.2 mM magnesium acetate, 20 mM KCl (added with the protein stock) and 100 µg/ml BSA for 15 min at 37 °C. D/R-loop formation was initiated by addition of supercoiled pUC19 dsDNA (67.2 µM, nt). Aliquots (10 µl) were withdrawn at indicated time points and deproteinized by incubation in 1 % SDS, 1.6 mg/ml proteinase K, 6 % glycerol and 0.01 % bromophenol blue for 15 min at 37 °C. Samples were analyzed by electrophoresis in 1 % agarose-TAE (40 mM TRIS·acetate, pH 8.0 and 1 mM EDTA) gels. The gels were dried on Amersham Hybond-N+ membranes, and then visualized and quantified using a Typhoon FLA 7000 Phosphor Imager (GE Healthcare). The yield was expressed as a percentage of the total plasmid DNA.

QUANTIFICATION AND STATISTICAL ANALYSIS

Graphs and statistical analysis were made using GraphPad Prism 8.0.2 (Graphpad Software, La Jolla, CA). The results are each expressed as a median and 95 % confidence limits (in parentheses), or alternatively mean with range (in parentheses) when the median is 0. Statistically significant differences between the His^+ frequencies were calculated using the nonparametric two-tailed Mann-Whitney *U*-test (Sokal and Rohlf, 2012). Significance of comparisons is indicated in figures and supplementary data as *, $P < 0.05$; **, $P < 0.01$, and ***, $P < 0.001$. Statistical analysis of *rev3* L979F mutants were performed using a one-tailed Mann-Whitney *U*-test taking the total number of mutations from each isolate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to D. Garfinkel for plasmids pBDG102, pGTyClaI, Ty antibody and for the *S. paradoxus* strain DG-2204, K.D. Koh for construction of KK strains, and Y. Shen for YS strains. We thank E. Graf and D. Sas for their technical assistance. We thank H. Ghalei and S. Marsili for critical reading of the manuscript, and all the members of the Storici lab for assistance and feedback on this research. We acknowledge funding from the National Cancer Institute (NCI) and the National Institute of General Medical Sciences (NIGMS) of the NIH (grant numbers CA188347, P30CA056036 and GM136717 to A.V.M.), Drexel Coulter Program Award (to A.V.M.), the National Institute of General Medical Sciences (NIGMS) of the NIH (grant number GM115927 to F.S.), the National Science Foundation fund (grant number 1615335 to F.S.), the Howard Hughes Medical Institute Faculty Scholar (grant number 55108574 to F.S.), and the grants from the Southeast Center for Mathematics and Biology (NSF, DMS-1764406 and Simons Foundation, 594594 to F.S.) for supporting this work.

References

Banyai G, Baidi F, Coudreuse D, and Szilagyi Z (2016). Cdk1 activity acts as a quantitative platform for coordinating cell cycle progression with periodic transcription. Nat. Commun 7, 11161. [PubMed: 27045731]

Bonnet A, Grosso AR, Elkaoutari A, Coleno E, Presle A, Sridhara SC, Janbon G, Geli V, de Almeida SF, and Palancade B (2017). Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. Mol Cell 67, 608–621.e606. [PubMed: 28757210]

Burns J, Kukushkin D, Chen X, Landweber LF, Saito M, and Jonoska N (2016). Recurring patterns among scrambled genes in the encrypted genome of the ciliate Oxytricha trifallax. J Theor Biol 410, 171–180. [PubMed: 27593332]

Chakraborty A, Tapryal N, Venkova T, Horikoshi N, Pandita RK, Sarker AH, Sarkar PS, Pandita TK, and Hazra TK (2016). Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes. Nat. Commun 7, 13049. [PubMed: 27703167]

Chan YA, Aristizabal MJ, Lu PYT, Luo Z, Hamza A, Kobor MS, Stirling PC, and Hieter P (2014). Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip. PLoS Genet. 10, e1004288. [PubMed: 24743342]

Chang HHY, Pannunzio NR, Adachi N, and Lieber MR (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat. Rev. Mol 18, 495–506.

Chatterjee N, and Walker GC (2017). Mechanisms of DNA damage, repair, and mutagenesis. Environ Mol Mutagen 58, 235–263. [PubMed: 28485537]

Coic E, Feldman T, Landman AS, and Haber JE (2008). Mechanisms of Rad52-independent spontaneous and UV-induced mitotic recombination in Saccharomyces cerevisiae. J. Genet 179, 199–211.

Derr LK, Strathern JN, and Garfinkel DJ (1991). RNA-mediated recombination in S. cerevisiae. Cell 67, 355–364. [PubMed: 1655280]

DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, and Church GM (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343. [PubMed: 23460208]

Dombroski BA, Feng Q, Mathias SL, Sassaman DM, Scott AF, Kazazian HH, and Boeke JD (1994). An in vivo assay for the reverse transcriptase of human retrotransposon L1 in Saccharomyces cerevisiae. Mol Cell Biol 14, 4485. [PubMed: 7516468]

El Hage A, Webb S, Kerr A, and Tollervey D (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet. 10, e1004716. [PubMed: 25357144]

Esnault C, Maestre J, and Heidmann T (2000). Human LINE retrotransposons generate processed pseudogenes. Nat. Genet 24, 363–367. [PubMed: 10742098]

Feschotte C, and Pritham EJ (2007). DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41, 331–368. [PubMed: 18076328]

Freudenreich CH (2018). R-loops: targets for nuclease cleavage and repeat instability. Curr. Genet 64, 789–794. [PubMed: 29327083]

Gallo D, Kim T, Szakal B, Saayman X, Narula A, Park Y, Branzei D, Zhang Z, and Brown GW (2019). Rad5 Recruits Error-Prone DNA Polymerases for Mutagenic Repair of ssDNA Gaps on Undamaged Templates. Mol Cell 73, 900–914.e909. [PubMed: 30733119]

Garcia-Muse T, and Aguilera A (2019). R Loops: From Physiological to Pathological Roles. Cell 179, 604–618. [PubMed: 31607512]

Garfinkel DJ, Hedge AM, Youngren SD, and Copeland TD (1991). Proteolytic processing of pol-TYB proteins from the yeast retrotransposon Ty1. J Virol 65, 4573–4581. [PubMed: 1714514]

Garfinkel DJ, Mastrangelo MF, Sanders NJ, Shafer BK, and Strathern JN (1988). Transposon tagging using Ty elements in yeast. J. Genet 120, 95–108.

Garfinkel DJ, Nyswaner KM, Stefanisko KM, Chang C, and Moore SP (2005). Ty1 copy number dynamics in Saccharomyces. J. Genet 169, 1845–1857.

Goodier JL (2016). Restricting retrotransposons: a review. Mob DNA 7, 16–16. [PubMed: 27525044]

Haber JE, and Hearn M (1985). Rad52-independent mitotic gene conversion in Saccharomyces cerevisiae frequently results in chromosomal loss. J. Genet 111, 7–22.

Hirano Y, and Sugimoto K (2006). ATR homolog Mec1 controls association of DNA polymerase zeta-Rev1 complex with regions near a double-strand break. Curr Biol. 16, 586–590. [PubMed: 16546083]

Holmes AM, and Haber JE (1999). Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. Cell 96, 415–424. [PubMed: 10025407]

Ivanov EL, and Haber JE (1995). RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in Saccharomyces cerevisiae. Mol Cell Biol 15, 2245–2251. [PubMed: 7891718]

Jasin M, and Rothstein R (2013). Repair of strand breaks by homologous recombination. CSH PERSPECT BIOL 5, a012740.

Jeon J, Park JS, Min B, Chung SK, Kim MK, and Kang YK (2019). Retroelement Insertion in a CRISPR/Cas9 Editing Site in the Early Embryo Intensifies Genetic Mosaicism. Front. Cell Dev. Biol 7, 273. [PubMed: 31781562]

Jinks-Robertson S, and Bhagwat AS (2014). Transcription-Associated Mutagenesis. Annu. Rev. Genet 48, 341–359. [PubMed: 25251854]

Kawasaki Y, and Sugino A (2001). Yeast replicative DNA polymerases and their role at the replication fork. Mol. Cells 12, 277–285. [PubMed: 11804324]

Keskin H, Meers C, and Storici F (2016). Transcript RNA supports precise repair of its own DNA gene. RNA Biol. 13, 157–165. [PubMed: 26637053]

Keskin H, Shen Y, Huang F, Patel M, Yang T, Ashley K, Mazin AV, and Storici F (2014). Transcript-RNA-templated DNA recombination and repair. Nature 515, 436–439. [PubMed: 25186730]

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. [PubMed: 22388286]

Lazzaro F, Novarina D, Amara F, Watt DL, Stone JE, Costanzo V, Burgers PM, Kunkel TA, Plevani P, and Muzi-Falconi M (2012). RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. Mol. Cell 45, 99–110. [PubMed: 22244334]

Lee MH, Siddoway B, Kaeser GE, Segota I, Rivera R, Romanow WJ, Liu CS, Park C, Kennedy G, Long T, et al. (2018). Somatic APP gene recombination in Alzheimer's disease and normal neurons. Nature 563, 639–645. [PubMed: 30464338]

Li X, Stith CM, Burgers PM, and Heyer WD (2009). PCNA is required for initiation of recombination-associated DNA synthesis by DNA polymerase delta. Mol. Cell 36, 704–713. [PubMed: 19941829]

Lyndaker AM, and Alani E (2009). A tale of tails: insights into the coordination of 3' end processing during homologous recombination. BioEssays 31, 315–321. [PubMed: 19260026]

Ma W, Resnick MA, and Gordenin DA (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, 1836–1846. [PubMed: 18267974]

Makarova AV, and Burgers PM (2015). Eukaryotic DNA polymerase zeta. DNA Repair 29, 47–55. [PubMed: 25737057]

Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D, Haft DH, Horvath P, et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol 18, 67–83. [PubMed: 31857715]

Maloisel L, Fabre F, and Gangloff S (2008). DNA polymerase delta is preferentially recruited during homologous recombination to promote heteroduplex DNA extension. Mol Cell Biol 28, 1373–1382. [PubMed: 18086882]

Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 3.

Martin SK, and Wood RD (2019). DNA polymerase zeta in DNA replication and repair. Nucleic Acids Res. 47, 8348–8361. [PubMed: 31410467]

Mazina OM, Keskin H, Hanamshet K, Storici F, and Mazin AV (2017). Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. Mol. Cell 67, 19–29.e13. [PubMed: 28602639]

McVey M, Khodaverdian VY, Meyer D, Cerqueira PG, and Heyer W-D (2016). Eukaryotic DNA Polymerases in Homologous Recombination. Annu. Rev. Genet 50, 393–421. [PubMed: 27893960]

Meers C, Keskin H, and Storici F (2016). DNA repair by RNA: Templated, or not templated, that is the question. DNA Repair 44, 17–21. [PubMed: 27237587]

Melamed C, Nevo Y, and Kupiec M (1992). Involvement of cDNA in homologous recombination between Ty elements in Saccharomyces cerevisiae. Mol Cell Biol 12, 1613–1620. [PubMed: 1372387]

Michelini F, Jalihal AP, Francia S, Meers C, Neeb ZT, Rossiello F, Gioia U, Aguado J, Jones-Weinert C, Luke B, et al. (2018). From “Cellular” RNA to “Smart” RNA: Multiple Roles of RNA in Genome Stability and Beyond. Chem. Rev 118, 4365–4403. [PubMed: 29600857]

Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, and Moran JV (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat. Genet 31, 159–165. [PubMed: 12006980]

Moseley PL (1989). Augmentation of bleomycin-induced DNA damage in intact cells. Am. J. Physiol. Cell Physiol 257, C882–887.

Mules EH, Uzun O, and Gabriel A (1998). In Vivo Ty1 Reverse Transcription Can Generate Replication Intermediates with Untidy Ends. J VIROL 72, 6490–6503. [PubMed: 9658092]

Nava GM, Grasso L, Sertic S, Pellicioli A, Muzi Falconi M, and Lazzaro F (2020). One, No One, and One Hundred Thousand: The Many Forms of Ribonucleotides in DNA. Int. J. Mol. Sci 21(5).

Nelson JR, Lawrence CW, and Hinkle DC (1996). Deoxycytidyl transferase activity of yeast REV1 protein. Nature 382, 729–731. [PubMed: 8751446]

Nevo-Caspi Y, and Kupiec M (1996). Induction of Ty recombination in yeast by cDNA and transcription: role of the RAD1 and RAD52 genes. J. Genet 144, 947–955.

Nowacki M, Vijayan V, Zhou Y, Schotanus K, Doak TG, and Landweber LF (2008). RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature 451, 153–158. [PubMed: 18046331]

Ono R, Ishii M, Fujihara Y, Kitazawa M, Usami T, Kaneko-Ishino T, Kanno J, Ikawa M, and Ishino F (2015). Double strand break repair by capture of retrotransposon sequences and reverse-transcribed spliced mRNA sequences in mouse zygotes. Sci. Rep 5, 12281. [PubMed: 26216318]

Onozawa M, Zhang Z, Kim YJ, Goldberg L, Varga T, Bergsagel PL, Kuehl WM, and Aplan PD (2014). Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. Proc. Natl. Acad. Sci. U.S.A 111, 7729–7734. [PubMed: 24821809]

Paques F, and Haber JE (1997). Two pathways for removal of nonhomologous DNA ends during double-strand break repair in Saccharomyces cerevisiae. Mol Cell Biol 17, 6765–6771. [PubMed: 9343441]

Piazza A, and Heyer WD (2019). Moving forward one step back at a time: reversibility during homologous recombination. Curr. Genet 65, 1333–1340 [PubMed: 31123771]

Plosky BS, and Woodgate R (2004). Switching from high-fidelity replicases to low-fidelity lesion-bypass polymerases. Curr Opin Genet Dev 14, 113–119. [PubMed: 15196456]

Posse V, Al-Behadili A, Uhler JP, Clausen AR, Reyes A, Zeviani M, Falkenberg M, and Gustafsson CM (2019). RNase H1 directs origin-specific initiation of DNA replication in human mitochondria. PLoS Genet. 15, e1007781. [PubMed: 30605451]

Rattray AJ, Shafer BK, McGill CB, and Strathern JN (2002). The roles of REV3 and RAD57 in double-strand-break-repair-induced mutagenesis of Saccharomyces cerevisiae. J. Genet 162, 1063–1077.

Reik W, Dean W, and Walter J (2001). Epigenetic reprogramming in mammalian development. Science 293, 1089–1093. [PubMed: 11498579]

Rodriguez GP, Romanova NV, Bao G, Rouf NC, Kow YW, and Crouse GF (2012). Mismatch repair-dependent mutagenesis in nondividing cells. Proc. Natl. Acad. Sci. U.S.A 109, 6153–6158. [PubMed: 22474380]

Rondon AG, and Aguilera A (2019). What causes an RNA-DNA hybrid to compromise genome integrity? DNA Repair 81, 102660. [PubMed: 31302006]

Sneeden JL, Grossi SM, Tappin I, Hurwitz J, and Heyer WD (2013). Reconstitution of recombination-associated DNA synthesis with human proteins. Nucleic Acids Res. 41, 4913–4925. [PubMed: 23535143]

Sokal RR, and Rohlf FJ (2012). Biometry : the principles and practice of statistics in biological research (New York: W.H. Freeman).

Sollier J, Stork CT, Garcia-Rubio ML, Paulsen RD, Aguilera A, and Cimprich KA (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. Mol. Cell 56, 777–785. [PubMed: 25435140]

Steele EJ (2017). Reverse Transcriptase Mechanism of Somatic Hypermutation: 60 Years of Clonal Selection Theory. Front Immunol 8, 1611–1611. [PubMed: 29218047]

Stone JE, Kissling GE, Lujan SA, Rogozin IB, Stith CM, Burgers PMJ, and Kunkel TA (2009). Low-fidelity DNA synthesis by the L979F mutator derivative of Saccharomyces cerevisiae DNA polymerase ζ. Nucleic Acids Res. 37, 3774–3787. [PubMed: 19380376]

Storici F, Bebenek K, Kunkel TA, Gordenin DA, and Resnick MA (2007). RNA-templated DNA repair. Nature 447, 338–341. [PubMed: 17429354]

Storici F, Lewis LK, and Resnick MA (2001). In vivo site-directed mutagenesis using oligonucleotides. Nat. Biotechnol 19, 773–776. [PubMed: 11479573]

Stuckey R, García-Rodríguez N, Aguilera A, and Wellinger RE (2015). Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. Proc. Natl. Acad. Sci. U.S.A 112, 5779–5784. [PubMed: 25902524]

Su XA, and Freudenreich CH (2017). Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A 114, E8392–e8401. [PubMed: 28923949]

Su Y, Ghodke PP, Egli M, Li L, Wang Y, and Guengerich FP (2019). Human DNA polymerase eta has reverse transcriptase activity in cellular environments. J. Biol 294, 6073–6081.

Sugawara N, Pâques F, Colaiácovo M, and Haber JE (1997a). Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination. Proc. Natl. Acad. Sci. U.S.A 94, 9214–9219. [PubMed: 9256462]

Surani MA (2001). Reprogramming of genome function through epigenetic inheritance. Nature 414, 122–128. [PubMed: 11689958]

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43, 11.10.11–11.10.33. [PubMed: 25431634]

Wahba L, Costantino L, Tan FJ, 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. [PubMed: 27298336]

Wimberly H, Shee C, Thornton PC, Sivaramakrishnan P, Rosenberg SM, and Hastings PJ (2013). R-loops and nicks initiate DNA breakage and genome instability in non-growing Escherichia coli. Nat. Commun 4, 2115. [PubMed: 23828459]

Xue Y, Haas SA, Brino L, Gusnanto A, Reimers M, Talibi D, Vingron M, Ekwall K, and Wright AP (2004). A DNA microarray for fission yeast: minimal changes in global gene expression after temperature shift. Yeast 21, 25–39. [PubMed: 14745780]

Yan Z, Xue C, Kumar S, Crickard JB, Yu Y, Wang W, Pham N, Li Y, Niu H, Sung P, et al. (2019). Rad52 Restrains Resection at DNA Double-Strand Break Ends in Yeast. Mol. Cell 76, 699–711.e696. [PubMed: 31542296]

Zhong X, Garg P, Stith CM, Nick McElhinny SA, Kissling GE, Burgers PMJ, and Kunkel TA (2006). The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. Nucleic Acids Res. 34, 4731–4742. [PubMed: 16971464]

Highlights

- DNA polymerase ζ promotes RNA-templated DNA repair and modification
- cDNA, unlike RNA-templated DNA repair, requires end-clipping function
- RNA-mediated DNA modification proceeds in the absence of recombination genes
- Mismatch repair facilitates accurate repair with template RNA

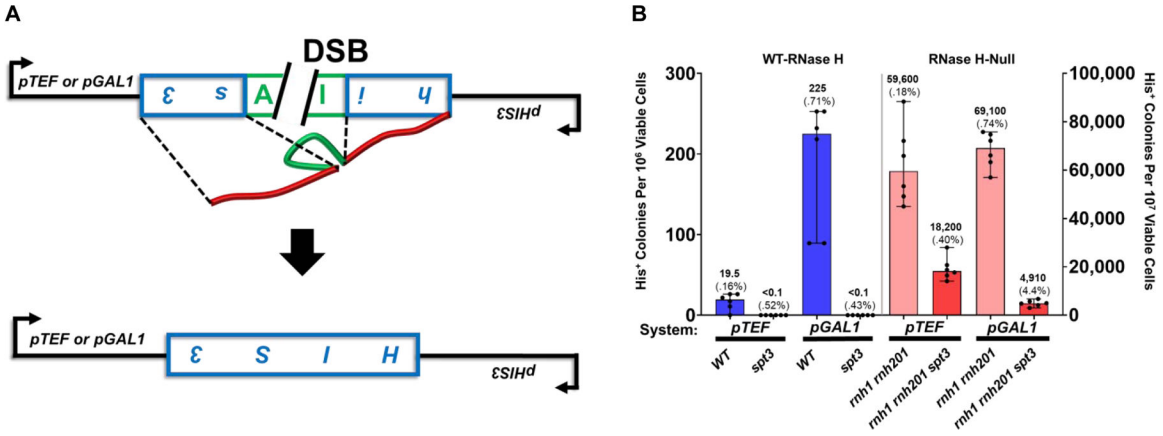


Figure 1. DSB Repair by RNA in *cis* is Facilitated by Constitutive Expression of the Template RNA

(A) Scheme of the genetic system used to detect R-TDR in yeast cells. The system contains a *his3* gene interrupted by an artificial intron (AI, in green) in the antisense orientation under either the galactose inducible promoter (*pGAL1*) or the constitutive translation elongation factor promoter (*pTEF*). Following transcription and splicing in the antisense orientation, the antisense transcript-RNA (in red) is used to guide removal of intronic sequence coded in DNA following a DSB inside the intronic sequence. This results in functional *HIS3* gene and yeast cell growth on medium lacking histidine. The DSB is generated by a galactose inducible HO endonuclease present on chromosome III.

(B) Fluctuation assay showing frequency of His^+ colonies per 10^6 or 10^7 viable cells following DSB induction in wild-type RNase H or RNase H-null cells. The donor antisense RNA is either expressed from the constitutive *pTEF* or the inducible *pGAL1* promoter. Cell genotypes are indicated. Colors of bars indicate interpreted pathway of repair cDNA (blue), cDNA/RNA (pink) and RNA (red). Individual frequencies are plotted. Bars represent median with 95 % confidence interval. The median is shown above each bar and survival following DSB induction is shown in parentheses; N=6. *P*-values are shown in Table S3.

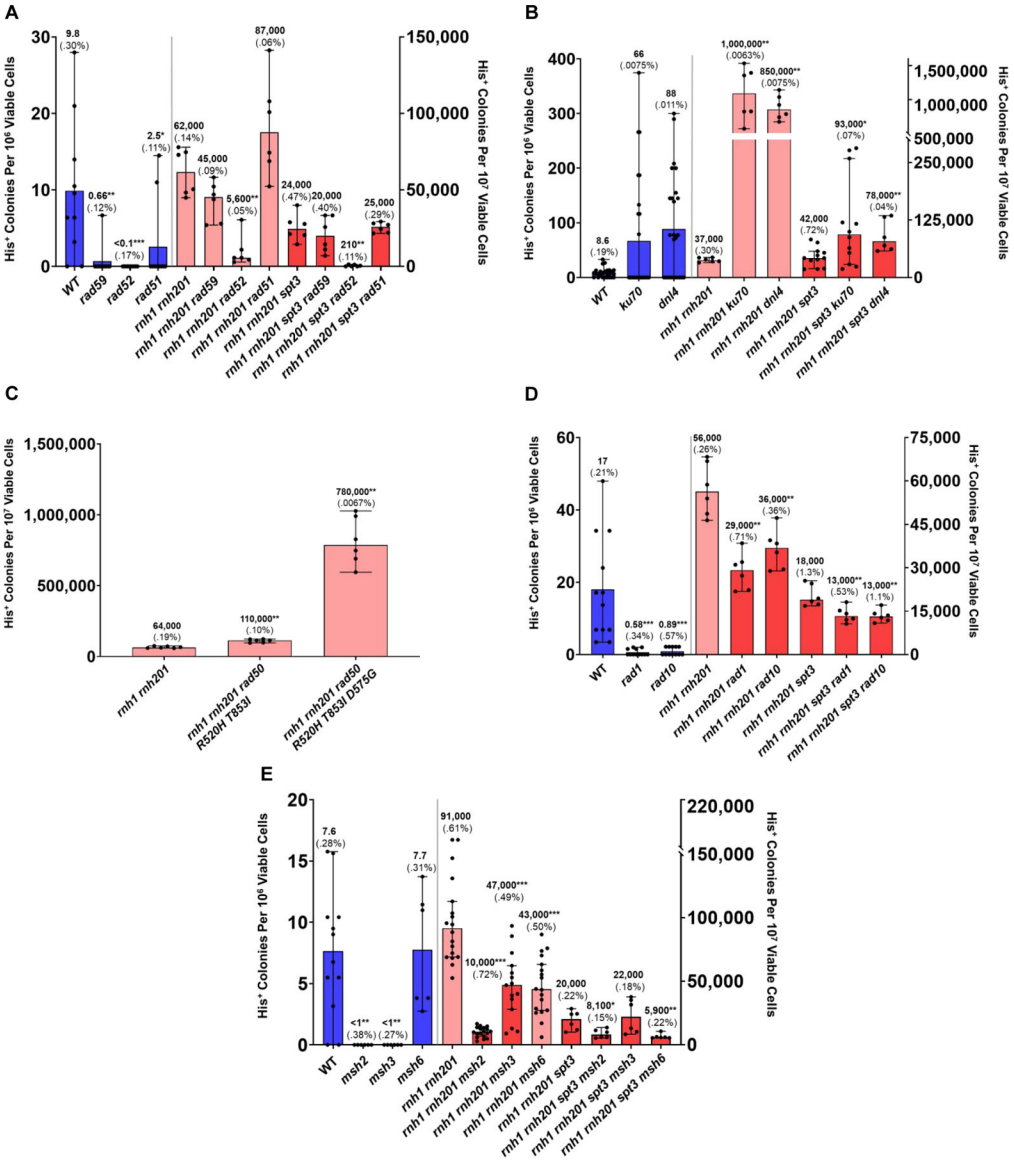


Figure 2. R-TDR requires Rad52 but not NHEJ proteins

Bar graphs of His$^+$ frequencies. Colors of bars indicate interpreted pathway of repair, cDNA (blue), cDNA/RNA (pink) and RNA (red). The genotype of the samples is indicated under each bar. Median or mean for each genotype is shown above each bar and survival following DSB is shown in parentheses. *P*-values are shown in Table S4. (A) Fluctuation assay showing frequency of His$^+$ colonies per 10^6 or 10^7 viable cells following DSB induction in different HR mutants. Individual frequencies are plotted. Bars represent mean with range (WT, *rad59*, *rad52*, *rad51*) or median with 95 % confidence interval (*rnh1 rnh201, rnh1 rnh201 rad59, rnh1 rnh201 rad52, rnh1 rnh201 rad51, rnh1 rnh201 spt3, rnh1 rnh201 spt3 rad59, rnh1 rnh201 spt3 rad52, rnh1 rnh201 spt3 rad51*). N=6–10.

(B) Fluctuation assay showing frequency of His$^+$ colonies per 10^6 or 10^7 viable cells following DSB induction in NHEJ mutants. Individual frequencies are plotted. Bars represent mean with range (WT, *ku70*, *dnl4*) *or* median with 95 % confidence interval (*rnh1*

rnh201, rnh1 rnh201 ku70, rnh1 rnh201 dnl4, rnh1 rnh201 spt3, rnh1 rnh201 spt3 dnl4, rnh1 rnh201 spt3 ku70). N=6–26.

(C) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells following DSB induction in *rad50* mutants. Individual frequencies are plotted. Bars represent median with 95 % confidence interval. N=6. *P*-values *rnh1 rnh201 - rnh1 rnh201 rad50 R520H T853I* (0.0022**), *rnh1 rnh201 - rnh1 rnh201 rad50 R520H T853I D575G* (0.0022**).

(D) Fluctuation assay showing frequency of His^+ colonies per 10^6 or 10^7 viable cells following DSB induction in clippase mutants. Individual frequencies are plotted. Bars represent mean with range (WT, *rad1*, *rad10*) or median with 95 % confidence interval (*rnh1 rnh201, rnh1 rnh201 rad1, rnh1 rnh201 rad10, rnh1 rnh201 spt3, rnh1 rnh201 spt3 rad1, rnh1 rnh201 spt3 rad10*). N=6–12.

(E) Fluctuation assay showing frequency of His^+ colonies per 10^6 or 10^7 viable cells following DSB induction in mismatch repair mutants. Individual frequencies are plotted. Bars represent mean with range (WT, *msh2*, *msh3, msh6) or* median with 95 % confidence interval (*rnh1 rnh201, rnh1 rnh201 msh2, rnh1 rnh201 msh3, rnh1 rnh201 msh6, rnh1 rnh201 spt3, rnh1 rnh201 spt3 msh2, rnh1 rnh201 msh3, rnh1 rnh201 msh6*). N=6–18.

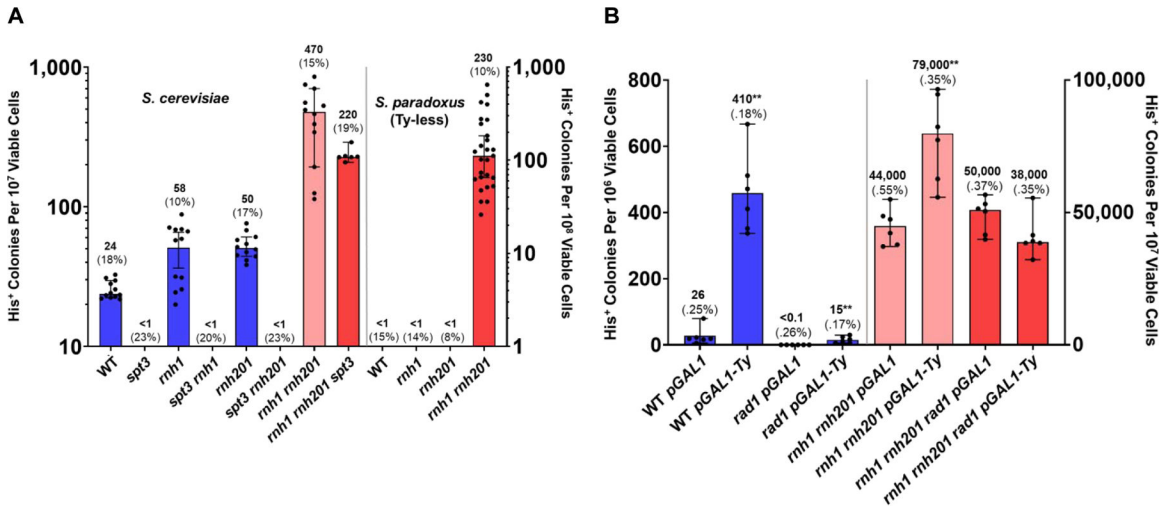


Figure 3. c-TDR is driven by the Ty RT

Bar graphs of His$^+$ frequencies. The genotype of the samples is indicated under each bar. Colors of bars indicate interpreted pathway of repair, cDNA (blue), cDNA/RNA (pink) and RNA (red). Median or mean of each genotype described above is shown above each bar and survival following DSB is shown in parentheses. *P*-values are shown in Table S5.

(A) Fluctuation assay of His$^+$ colonies per 10^7 or 10^8 viable cells following DSB induction comparing frequencies in *S. cerevisiae* and *S. paradoxus* (Ty-less) cells. Individual frequencies are plotted. Bars represent median with 95 % confidence interval on a log scale. N=6–36.

(B) Fluctuation assay showing frequency of His$^+$ colonies per 10^6 or 10^7 viable cells following DSB induction in cells containing integrated *pGAL1-Ty*. Individual frequencies are plotted. Bars represent mean with range (WT *pGAL1*, WT *pGAL1-Ty, rad1 pGAL1, rad1 pGAL1-Ty*) or median with 95 % confidence interval (*rnh1 rnh201 pGAL1, rnh1 rnh201 pGAL1-Ty, rnh1 rnh201 rad1 pGAL1, rnh1 rnh201 rad1 pGAL1-Ty*). N=6.

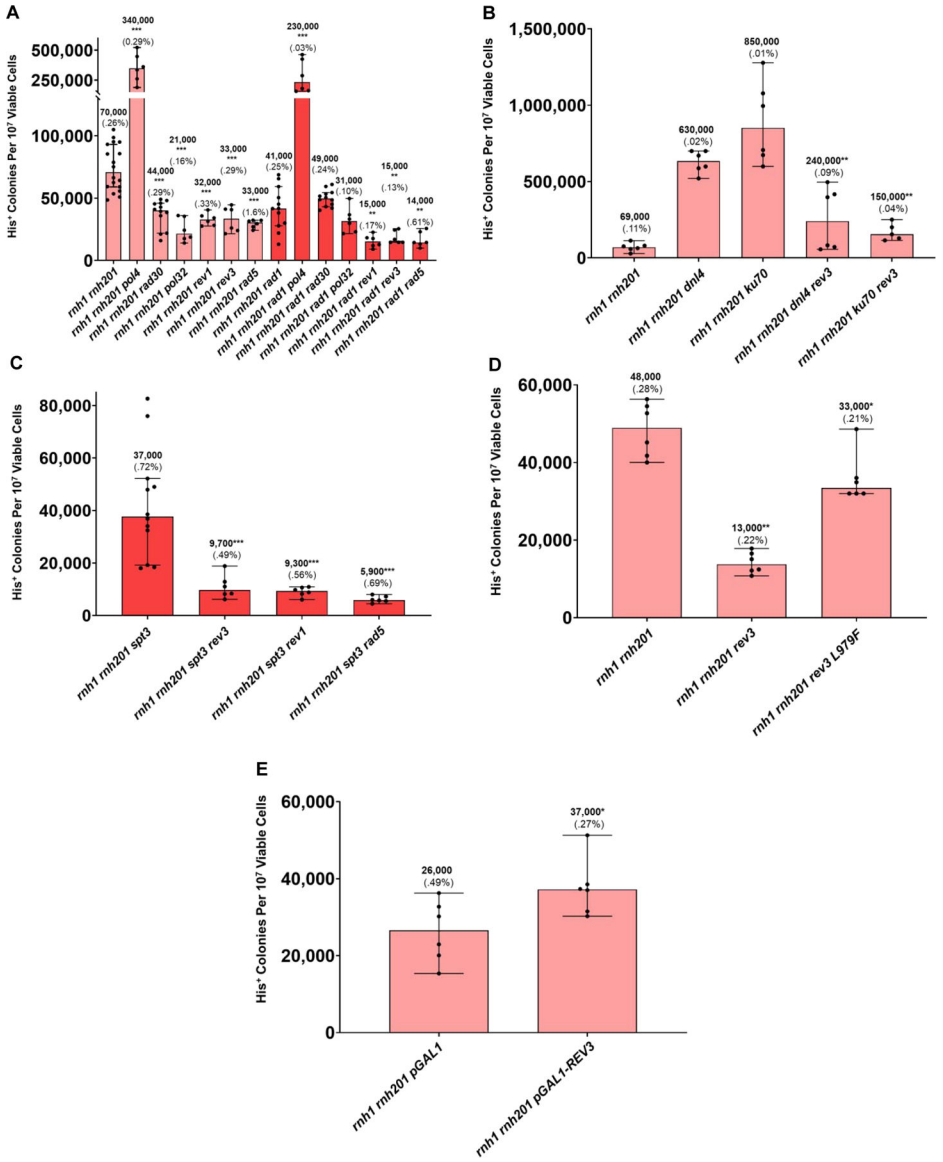


Figure 4. DNA polymerase ζ promotes RNA-DNA recombination triggered by a DSB

Bar graphs of His$^+$ frequencies. The genotype of the samples is indicated under each bar. Colors of bars indicate interpreted pathway of repair, cDNA/RNA (pink) and RNA (red). Median or mean of each genotype described above is shown above each bar and survival following DSB is shown in parentheses. *P*-values are shown in Table S6.

(A) Fluctuation assay showing frequency of His$^+$ colonies per 10^7 viable cells following DSB induction in yeast non-essential DNA polymerase mutants. Individual frequencies are plotted. Bars represent median with 95 % confidence interval. N=6–18.

(B) Fluctuation assay showing frequency of His$^+$ colonies per 10^7 viable cells following DSB induction in NHEJ and Pol Zeta mutants. Individual frequencies are plotted. Bars represent median with 95 % confidence interval. N=6. *P*-values *rnh1 rnh201 dnl4* – rnh1 *rnh201 dnl4 rev3* (0.0022**), *rnh1 rnh201 ku70* – *rnh1 rnh201 ku70 rev3* (0.0043**).

(C) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells following DSB induction in Pol Zeta mutants. Individual frequencies are plotted. Bars represent median with 95 % confidence interval. N=6–12.
(D) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells following DSB induction in *rev3* L979F low fidelity mutant. Individual frequencies are plotted. Bars represent median with 95 % confidence interval. N=6. *P*-values *rnh1 rnh201 - rnh1 rnh201 rev3* (0.0022), *rnh1 rnh201 - rnh1 rnh201 rev3 L979F* (0.0152).
(E) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells following DSB induction in cells overexpressing *REV3*. Individual frequencies are plotted. Bars represent median with 95% confidence interval. N=6. *P*-values *rnh1 rnh201 pGAL1 - rnh1 rnh201 pGAL1-REV3* (0.0260).

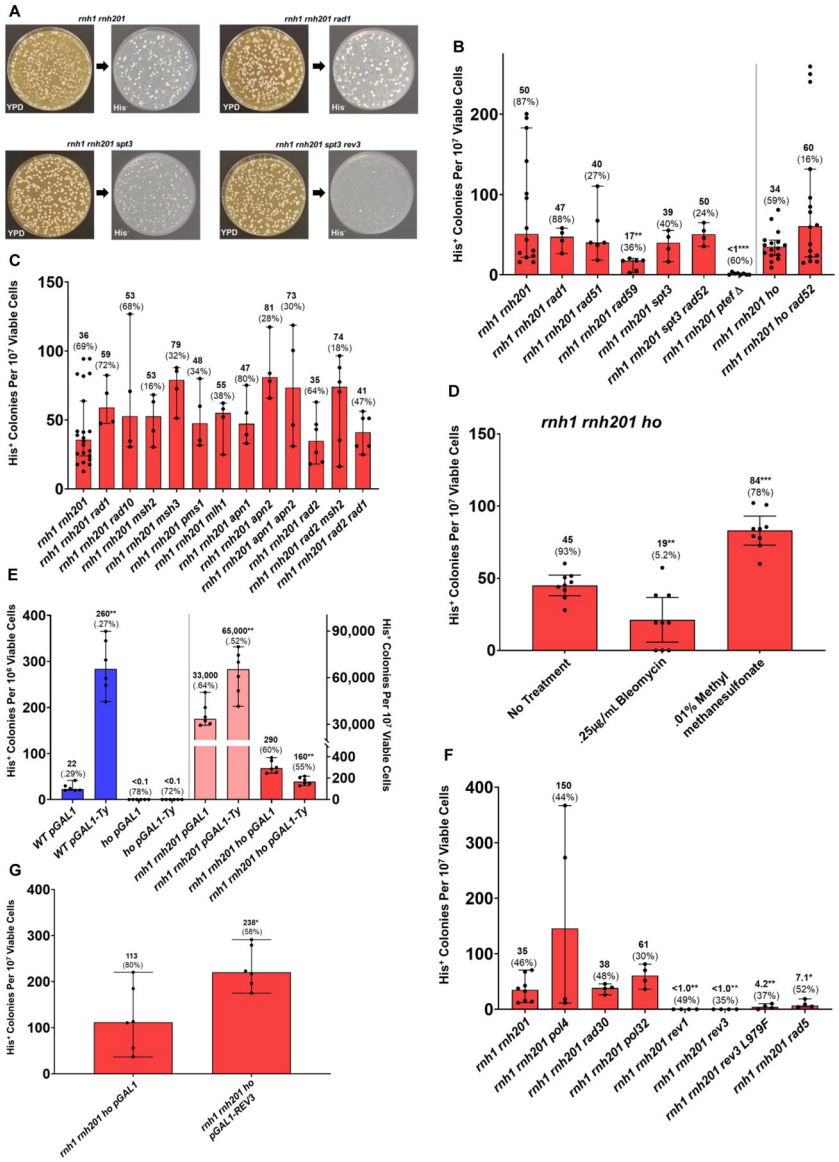


Figure 5. R-TDM requires DNA polymerase ζ

(A) Representative plates showing single colony isolates of *rnh1 rnh201*, *rnh1 rnh201 rad1*, *rnh1 rnh201 spt3* and *rnh1 rnh201 spt3 rev3* mutant strains on YPD medium (no DSB induction) that were replica-plated onto His^- medium.

(B-G) Bar graphs of His^+ frequencies. Bars represent median with 95 % confidence interval. The genotype of the samples is indicated under each bar. Colors of bars indicate interpreted pathway of repair cDNA (blue), cDNA/RNA (pink) and RNA (red). The median is shown above each bar and survival shown in parentheses. *P*-values are shown in Table S7.

(B) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells with no DSB induction for cells of different recombination mutants. Individual frequencies are plotted. N=4–16.

(C) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells with no DSB induction for cells of NER, BER and mismatch repair mutants. Individual frequencies are plotted. N=4–16.

(D) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells without DSB induction following mutagen treatment. Individual frequencies are plotted. N=9 *P*-values No treatment – bleomycin (0.0096**), No treatment – MMS (<0.0001***).

(E) Fluctuation assay showing frequency of His^+ colonies per 10^6 or 10^7 viable cells with and without DSB induction in strains overexpressing the Ty transposon. Individual frequencies are plotted. N=6.

(F) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells with no DSB induction for cells of non-essential DNA polymerase mutants. Individual frequencies are plotted. N=4–8. Mutations generated in L979F mutation is shown in Table S8.

(G) Fluctuation assay showing frequency of His^+ colonies per 10^7 viable cells with no DSB induction for cells overexpressing *REV3*. Individual frequencies are plotted. N=6 *P*-value *rnh1 rnh201 ho pGAL1-rnh1 rnh201 ho pGAL1-REV3* (0.0260).

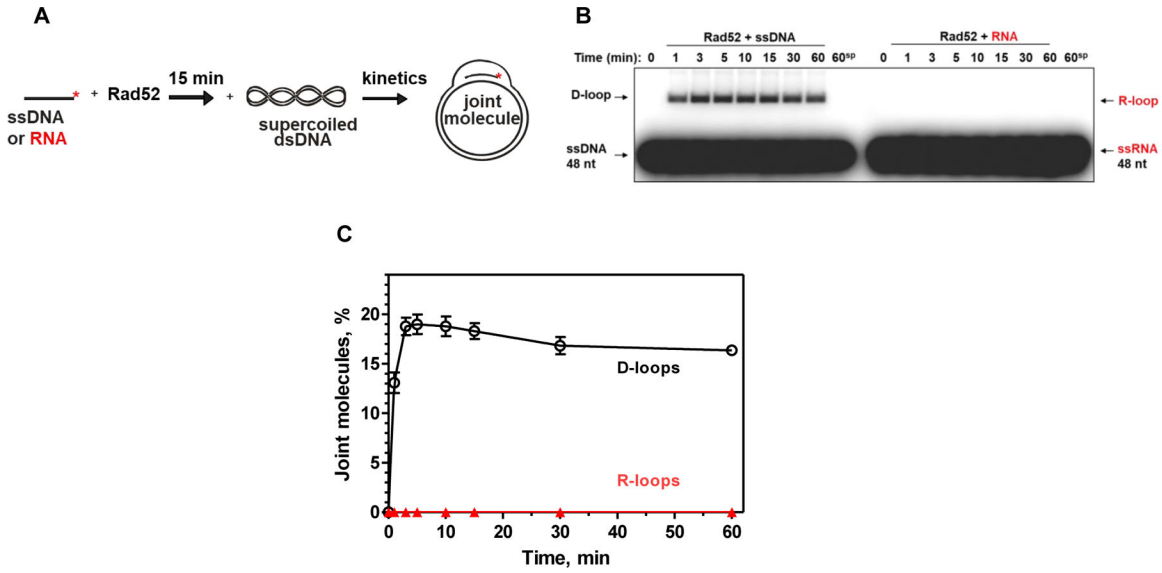


Figure 6. ScRad52 promotes D-loop but not R-loop formation.

(A) The reaction schemes. A red asterisk denotes the ^{32}P label.

(B) The kinetics of D- and R-loop formation. An example of three repeats is shown. ScRad52 protein (450 nM) was preincubated with a 48-mer ^{32}P- labeled ssDNA (no. 211; 3 μM, nt) or with a 48-mer RNA of identical sequence (no. 501; 3 μM, nt) for 15 min at 37 °C, and D-loop formation was initiated by addition of homologous pUC19 dsDNA (67.2 μM, nt). The reaction products at indicated time points were analyzed by electrophoresis in a 1 % agarose gel. In the reaction marked by "60^{sp}", ScRad52 was replaced with storage buffer, and the reaction was carried out for 60 min at 37 °C.

(C) Data from B represented as a graph. Error bars indicate the standard error of the mean (SEM). N=3

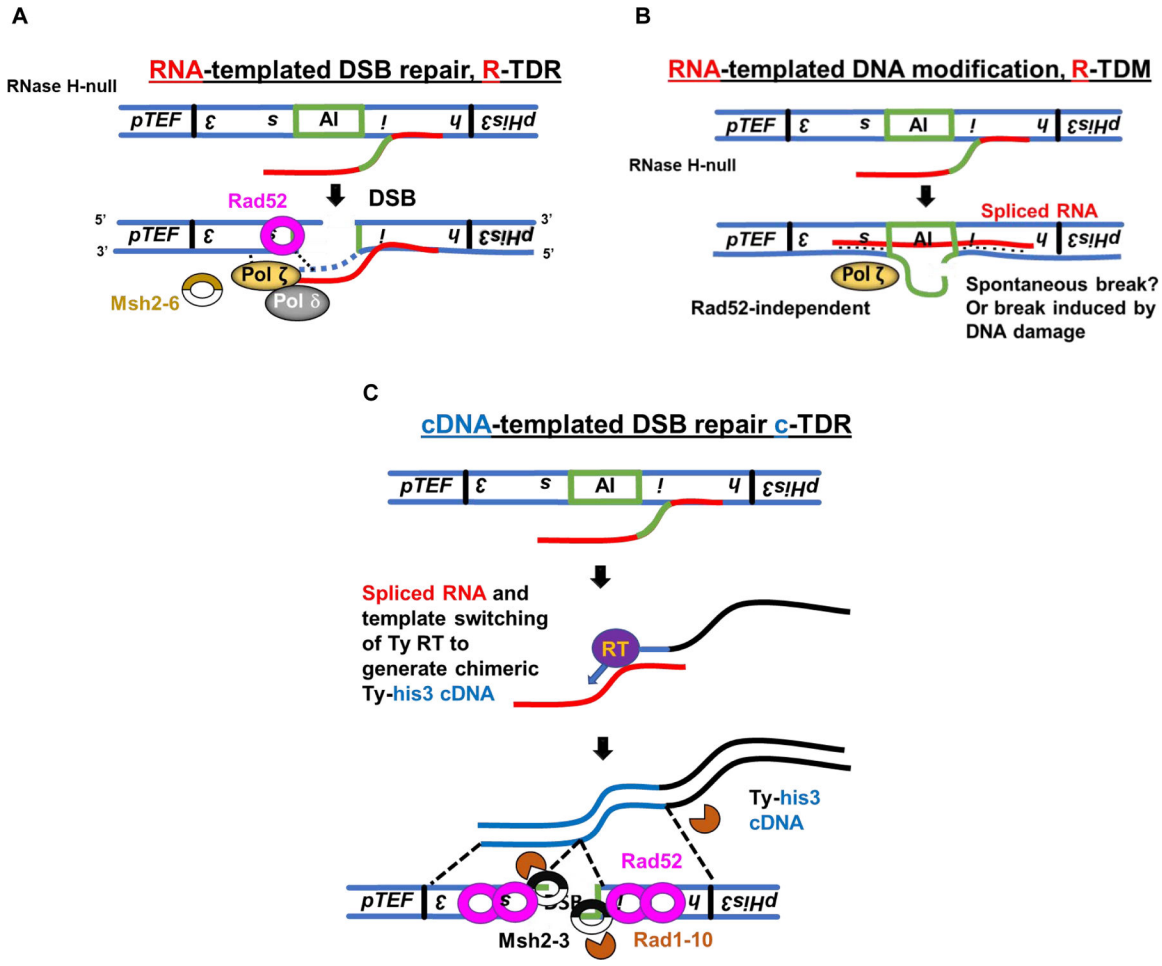


Figure 7. Model of RNA-templated DSB repair (R-TDR), RNA-templated DNA modification (R-TDM) and cDNA-templated DSB repair (c-TDR)

(A) In R-TDR, transcript-RNA anneals back to DNA following DSB aided by Rad52. Following removal of intronic 3′ tail via exonuclease or clipping function, RNA is then used as a template for DSB repair synthesis, aided by DNA Pol ζ. Msh2–6 are required for R-TDR, possibly to remove mismatches introduced by Pol ζ.

(B) In R-TDM, transcript-RNA forms an R-loop structure without the need of Rad52. Likely a spontaneous break, or a break induced by DNA damage provides a 3′ tail for DNA synthesis. RNA is used as a template by DNA Pol ζ to guide removal of the intronic sequence in DNA.

(C) In c-TDR, spliced antisense *his3* RNA is reverse transcribed by Ty RT and used as a templated to facilitate repair of DSB inside of the *his3* locus from which the RNA was generated. This process requires Rad52 for annealing, and Rad1–10 along with Msh2–3 in the removal of non-homologous DNA tails.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Beta-Actin	Abcam	ab170325
B8	Garfinkel et al., 1991	NA
Goat anti-Mouse IgG	Thermo Fisher	31430
Goat anti-Rabbit IgG	Thermo Fisher	31460
Bacterial and Virus Strains		
XL-1 Blue Supercompetent cells	Agilent	50125053
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
PBS	Corning	21-040-CV
1x protease inhibitor cocktail	Roche	11836170001
0.05% NP-40	Thermo Fisher	28324
Bradford reagent	Biorad	500–0006
BSA standards	Biorad	500–0207
Amersham nitrocellulose membrane	GE	10600003
ECL Western Blotting solution	Thermo Fisher	32106
Acid Phenol	VWR	0981.400ML
Qiagen DNA Buffer Set	Qiagen	19060
Nextera XT DNA Library Preparation Kit	Illumina	FC-131-1024
RNeasy mini kit	Qiagen	74104
Turbo DNA-free kit	Thermo Fisher	AM1907
iScript cDNA synthesis kit	Biorad	1708891
Methyl-methanesulfonate	Alfa Aesar	66-27-3
Bleomycin	Alfa Aesar	10814–234
Critical Commercial Assays		
Deposited Data		
DNA-seq reads for sample CM10	NCBI SRA	BioProject PRJNA656525
Experimental Models: Cell Lines		
Experimental Models: Organisms/Strains		
Saccharomyces cerevisiae; *Saccharomyces paradoxus*; Individual genotypes see Table S1		
Oligonucleotides		
Oligonucleotides, see Table S2		
Recombinant DNA		
p414-TEF-Cas9	DiCarlo et al., 2013	43802
pBDG102	Garfinkel et al., 1988	NA
pGTyClaI	Garfinkel et al., 1988	NA
Software and Algorithms		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GraphPad Prism 8.0.2	GraphPad	https://www.graphpad.com/scientific-software/prism/
StepOne Software v2.1	Thermo Fisher	https://www.thermofisher.com/us/en/home/technical-resources/software-downloads/StepOne-and-StepOnePlus-Real-Time-PCR-System.html
Trim Galore 0.3.7	Babraham Institute	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
Bowtie2 2.3.2	Johns Hopkins University	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
GATK 2.7 java -jar GenomeAnalysisTK.jar -T HaplotypeCaller -ploidy 1 -R sacCer3.fa -I BAM -ERC GVCF -o VCF https://github.com/agombolay/Variant-Calling	Broad Institute	https://gatk.broadinstitute.org/hc/en-us
Other		
96-well Step One Plus Real Time PCR system	Thermo Fisher	4376600