nature structural & molecular biology

# Evidence that processing of ribonucleotides in DNA by topoisomerase 1 is leading-strand specific

Jessica S Williams[1], Anders R Clausen[1], Scott A Lujan[1], Lisette Marjavaara[2], Alan B Clark[1], Peter M Burgers[3], Andrei Chabes[2,4] & Thomas A Kunkel[1]

**Ribonucleotides incorporated during DNA replication are removed by RNase H2–dependent ribonucleotide excision repair (RER). In RER-defective yeast, topoisomerase 1 (Top1) incises DNA at unrepaired ribonucleotides, initiating their removal, but this is accompanied by RNA-DNA–damage phenotypes. Here we show that these phenotypes are incurred by a high level of ribonucleotides incorporated by a leading strand–replicase variant, DNA polymerase (Pol) ε, but not by orthologous variants of the lagging-strand replicases, Pols α or δ. Moreover, loss of both RNases H1 and H2 is lethal in combination with increased ribonucleotide incorporation by Pol ε but not by Pols α or δ. Several explanations for this asymmetry are considered, including the idea that Top1 incision at ribonucleotides relieves torsional stress in the nascent leading strand but not in the nascent lagging strand, in which preexisting nicks prevent the accumulation of superhelical tension.**





In DNA synthesis reactions containing cellular concentrations of dNTPs and rNTPs, the replicative DNA polymerases of *Saccharomyces cerevisiae*, Pols α, δ and ε, frequently incorporate ribonucleotides into DNA[1]. This fact motivated studies to determine whether ribonucleotides are incorporated by DNA polymerases during nuclear replication *in vivo* and, if so, whether ribonucleotides in the genome have biological effects. To focus on ribonucleotide incorporation by a replicative DNA polymerase (replicase) rather than by the RNA primase, which initiates lagging-strand Okazaki fragments, our first studies examined ribonucleotide incorporation by Pol ε, the primary leading-strand replicase[2,3], encoded by the *S. cerevisiae POL2* gene. We used a variant of Pol ε in which a glycine replaced a conserved methionine (M644) adjacent to the 'steric gate' tyrosine (Y645, in the polymerase active site) that prevents ribonucleotide incorporation by B family polymerases[4,5]. Purified M644G Pol ε incorporates about ten times more ribonucleotides than does wild-type Pol ε (ref. 6). In a *pol2*-M644G yeast strain, deletion of the *RNH201* gene (*rnh201*Δ) encoding the catalytic subunit of RNase H2 results in increased incorporation of ribonucleotides into genomic DNA in excess over those detected in the *pol2*-M644G or *rnh201*Δ single-mutant strains[6]. This observation demonstrates that M644G Pol ε incorporates ribonucleotides into DNA *in vivo* and that RNase H2–dependent repair removes them, as had been predicted earlier[7,8]. This repair has been reconstituted *in vitro*[9] and is designated RER.

We have also examined ribonucleotides incorporated by the major lagging-strand replicase, Pol δ, using a variant in which a conserved leucine adjacent to the steric-gate tyrosine was replaced with methionine. Like its Pol ε counterpart, purified L612M Pol δ also incorporates about ten times more ribonucleotides than its wild-type parent[10]. In yeast encoding the *pol3*-L612M variant, deletion of *RNH201* also resulted in the presence of ribonucleotides in genomic DNA in excess over those detected in the *pol3*-L612M or *rnh201*Δ single-mutant strains. Moreover, strand-specific probing of genomic DNA demonstrated that equivalent Pol ε variants in *Schizosaccharomyces pombe*[11] and *S. cerevisiae*[10,12,13] preferentially incorporate ribonucleotides into the nascent leading-strand DNA, whereas the *S. cerevisiae pol3*-L612M variant preferentially incorporates ribonucleotides into the nascent lagging strand[10]. These preferences strongly support the interpretation, based on mutational signatures[2,3], that Pol ε is primarily a leading-strand replicase and that Pol δ is primarily a lagging-strand replicase. They further demonstrate that RER removes ribonucleotides from both the leading and lagging strands.

In our original study of ribonucleotide incorporation *in vitro*[6], we speculated that the transient presence of ribonucleotides in DNA might serve positive signaling functions, such as directing mismatch repair (MMR) to correct replication errors in the nascent leading strand. Subsequent genetic and biochemical studies have supported this hypothesis[10,14]. Such a signaling role, and evidence that a ribonucleotide imprint in *S. pombe* signals for mating-type switching[15], are examples of beneficial effects of ribonucleotides in DNA; other possibilities have been discussed elsewhere[1,16–18]. Nonetheless, ribonucleotides are noncanonical constituents in DNA (i.e., lesions), and as such they can elicit damage that requires repair via an RNA-DNA–damage response. For example, when RNase H2 cleavage at a ribonucleotide generates a nick in DNA, an attempt to ligate this nick can generate an adenylated 5′-RNA-DNA junction that can elicit

[1]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA. [2]Department of Medical Biochemistry and Biophysics, Umeå University, Umeå, Sweden. [3]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri, USA. [4]Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, Sweden. Correspondence should be addressed to T.A.K. (kunkel@niehs.nih.gov).

Received 26 September 2014; accepted 11 February 2015; published online 9 March 2015; doi:10.1038/nsmb.2989

genome instability if not deadenylated by aprataxin[19]. Moreover, in the absence of RER, Top1 can cleave the DNA backbone where ribonucleotides are present[20,21], thereby initiating removal of ribonucleotides[13]. In doing so, Top1 generates DNA ends with a 5′-OH and a 2′,3′-cyclic phosphate that must be processed to allow eventual ligation. These atypical 'dirty' DNA ends are nonligatable and are problematic for the cell. For example, compared to *pol2*-M644G or *rnh201*Δ single-mutant strains, the *pol2*-M644G *rnh201*Δ double-mutant strain, which contains large numbers of unrepaired ribonucleotides in the genome, grows slowly, is sensitive to the replication inhibitor hydroxyurea (HU) and has a strongly elevated rate of 2- to 5-bp deletions in repetitive DNA sequences[6,21–23]. These negative consequences are largely suppressed by deletion of the *TOP1* gene encoding Top1 (refs. 13,21). In addition, the *pol2*-M644G *rnh201*Δ strain does not survive the additional loss of RNase H1 (ref. 24), which can incise DNA substrates containing four or more consecutive ribonucleotides[25,26].

A particularly interesting feature of the 2- to 5-bp deletion mutagenesis observed in the *pol2*-M644G *rnh201*Δ double-mutant strain is its asymmetry. For example, the rate of deleting a CA dinucleotide from a CACA repeat in the *URA3* reporter gene is 160-fold higher when *URA3* is in one orientation relative to a nearby replication origin as compared to the other orientation[6]. This bias suggests that the mutagenesis results from incorporation of ribonucleotides during Pol ε–mediated leading-strand replication of the 5′-CACA–containing template[21,22]. We therefore set out to probe whether unrepaired ribonucleotides incorporated by the lagging-strand replicases Pols α and δ may also contribute to replicative stress or genome instability. Here we investigate this by using variants of Pols α and δ. We first show that, like M644G Pol ε (ref. 6) and L612M Pol δ (ref. 10), the L868M Pol α variant is much more likely than its wild-type parent enzyme to incorporate ribonucleotides during DNA synthesis *in vitro* and also during replication *in vivo*. We further show that, like L612M Pol δ, L868M Pol α preferentially incorporates ribonucleotides into the nascent lagging strand *in vivo* and that these ribonucleotides are repaired by RER. Surprisingly, however, we observed Top1-dependent removal of ribonucleotides, and the resulting RNA-DNA damage, only for ribonucleotides incorporated by Pol ε, not Pols α or δ. This implies that Top1-dependent ribonucleotide removal and its adverse consequences preferentially result from unrepaired ribonucleotides in the nascent leading strand. We consider three possible models to explain this asymmetry, including the possibility that a difference in the need to relieve torsional stress in the replicated duplexes behind the fork exists during DNA synthesis.

## RESULTS

### Ribonucleotide incorporation by L868M Pol α and L612M Pol δ

We previously showed that substituting glycine for a conserved methionine (Met644) in the Pol ε active site, or substituting methionine for Leu612 at the equivalent location in Pol δ, increases ribonucleotide incorporation during DNA synthesis *in vitro*[6,10]. Here we examined a Pol α variant in which the equivalent conserved leucine (Leu868) was changed to methionine[27]. We measured stable incorporation of ribonucleotides (rNMPs) during extension of a 40-mer primer hybridized to a 70-mer template (**Fig. 1a**). Polymerization reactions contained all four dNTPs and rNTPs at concentrations estimated to be present in yeast[1]. We isolated full-length reaction products and treated them with 0.3 M KOH to hydrolyze the DNA backbone

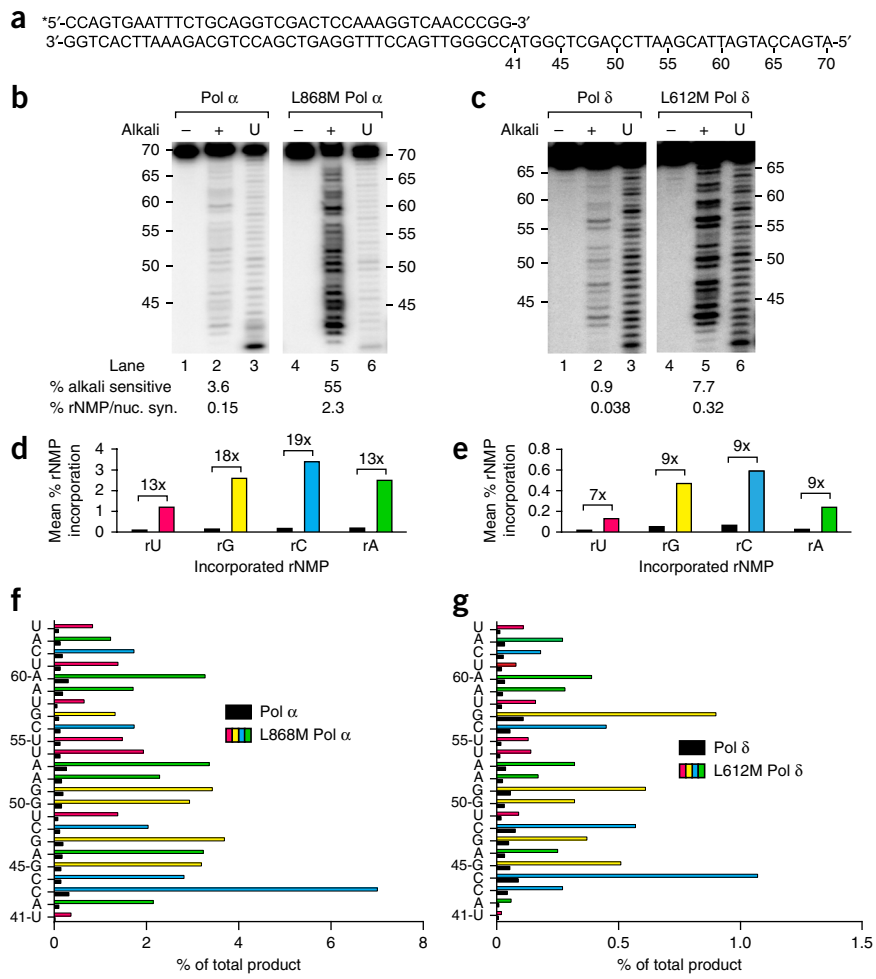



**Figure 1** Ribonucleotide incorporation *in vitro* by variants of Pols α and δ. (**a**) Sequence of primer template used for reactions in **b** and **c**. Asterisk represents the position of the $^{32}$P label. (**b**) Stable rNMP incorporation into DNA. The lane marked U depicts the product generated by Pol α or L868M Pol α before gel purification, as described in ref. 1. Lanes marked with – and + depict gel-purified products treated with 0.3 M KCl (–) or KOH (+). The percentages of alkali-sensitive products and ribonucleotide (rNMP) incorporated per nucleotide synthesized (nuc. syn.) are shown below each lane. The mean and range for duplicate measurements was 3.6 ± 0.3 for Pol α and 55 ± 0.5 for L868M Pol α. (**c**) As in **b**, but for DNA products made by Pol δ (0.9 ± 0.2) and L612M Pol δ (7.7 ± 1.5). (**d**) Average frequency of ribonucleotide incorporation for rU, rA, rC and rG, calculated from **b**. The relative difference in ribonucleotide incorporation between Pol α and L868M Pol α is shown above each base. (**e**) As in **d**, but for Pol δ and L612M Pol δ, with data from **c**. (**f**) Percentage of rNMP incorporation by Pol α or L868M Pol α at each of 24 template positions. The position and identity of each incorporated ribonucleotide is displayed on the *y* axis. (**g**) As in **f**, but for ribonucleotide incorporation by Pol δ or L612M Pol δ. Original gel images are in **Supplementary Data Set 1**.

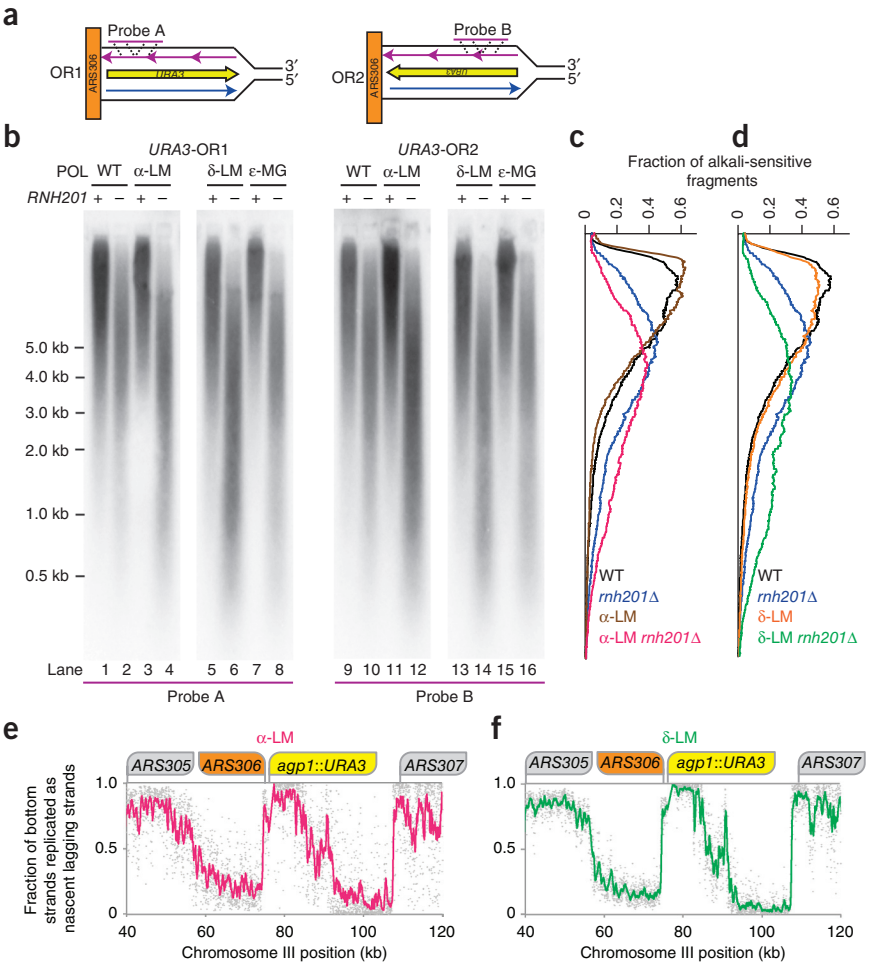



**Figure 2** Strand-specific probing for ribonucleotides in nascent lagging-strand genomic DNA. (**a**) Schematic of the two orientations (OR1 and OR2) of the *URA3* reporter located adjacent to the *ARS306* replication origin. Template strands are black, the nascent leading strand (top strand) synthesized by Pol ε is blue, and the nascent lagging strand (bottom strand) synthesized by Pols α and δ is purple. The probe annealing locations are indicated with dotted lines. (**b**) Alkali-sensitive-site determination and quantitation, performed as previously described[13]. Strains contain one of four versions of the replicative DNA polymerases: wild type (WT), *pol1*-L868M (α-LM), *pol3*-L612M (δ-LM) or *pol2*-M644G (ε-MG). (**c**,**d**) Quantification of data in **b** to determine the average fraction of total alkali-sensitive fragments at each position along the membrane (DNA marker positions in **b**). Curves are plots of the average values from two independent experiments. (**e**) Fraction of replication events in which the bottom strand is replicated as the nascent lagging strand, as estimated from α-LM *rnh201*Δ HydEn-seq data ($n = 1$ cell culture). Noise was reduced via comparison with ε-MG *rnh201*Δ ($n = 4$ independent cell cultures; Online Methods). Gray diamonds represent data for 20-bp bins. The trend line represents a 20-bin (400-bp) moving average. (**f**) As in **e**, but calculated from δ-LM *rnh201*Δ ($n = 2$ independent cell cultures) data. Original blot images are in **Supplementary Data Set 1**.

at positions where rNMPs were present. We then used band intensities to quantify the percentage of rNMPs incorporated at 24 positions. To avoid rNMP incorporation by the RNA primase of the four-subunit Pol α–primase complex, we used the Pol α catalytic subunit only. In order to directly compare ribonucleotide incorporation by both lagging-strand replicases, we extended our initial study of three-subunit L612M Pol δ (ref. 10). The results (**Fig. 1**) for wild-type Pols α and δ are similar to those reported in our initial study[1]. In comparison, average rNMP incorporation by L868M Pol α and L612M Pol δ is 15- and 8-fold higher, respectively (**Fig. 1b**,**c**), corresponding to an average of about one rNMP incorporated per 40 dNMPs by L868M Pol α and about one rNMP incorporated per 300 dNMPs by L612M Pol δ. As for the wild-type enzymes, ribonucleotide incorporation by L868M Pol α and L612M Pol δ varies among the four rNMPs (with rC >rG >rA >rU; **Fig. 1d**,**e**) and at 24 individual nucleotide positions (**Fig. 1f**,**g**). The variation in rNMP incorporation probability *in vitro* by polymerase, by the identity of the base attached to the ribose and by the surrounding sequence may be relevant to any negative and/or positive biological consequences of ribonucleotides in the genome. This possibility is supported by previous studies showing hotspots for ribonucleotide-dependent deletion mutations[6,21] and strand-specific effects on MMR efficiency[10] and by the results below.

### RER removes ribonucleotides from both nascent DNA strands

To test for strand-specific incorporation of ribonucleotides and their removal by RER, we isolated genomic DNA from 16 yeast strains (**Supplementary Table 1**) that differ in the genes encoding the replicases, the catalytic subunit of RNase H2 (Rnh201) and/or Top1. We treated the DNA samples with KOH to hydrolyze molecules containing ribonucleotides and then separated fragments by electrophoresis in an alkaline agarose gel, transferred them to a nylon filter and probed with radiolabeled DNA complementary to either strand of the *URA3* gene located next to the *ARS306* replication origin (**Fig. 2a** and **Supplementary Fig. 1a**). The results (**Fig. 2b** and **Supplementary Fig. 1**) indicate that alkaline hydrolysis of the all *rnh201*Δ genomes, as compared to all *RNH201* genomes, yields shorter DNA fragments, thus demonstrating that RER is occurring and that in its absence unrepaired ribonucleotides are present in the genome. As expected on the basis of earlier studies using a leading strand–specific probe[11–13], nascent leading-strand DNA fragments in the *pol2*-M644G *rnh201*Δ strain (**Supplementary Fig. 1b**, lanes 8 and 16) are shorter than leading-strand fragments in the single-mutant *rnh201*Δ strain (**Supplementary Fig. 1b**, lanes 2 and 10), thus indicating that M644G Pol ε preferentially incorporates ribonucleotides into the nascent leading strand and that they are removed by RER. Similarly, lagging-strand DNA fragments in the *pol3*-L612M *rnh201*Δ strain (**Fig. 2b**, lanes 6 and 14, and **Fig. 2d**, green scan) are shorter than lagging-strand fragments in the *rnh201*Δ strain encoding wild-type replicases (**Fig. 2b**, lanes 2 and 10, and **Fig. 2d**, blue scan). This indicates that RER also removes ribonucleotides incorporated into the nascent lagging strand by L612M Pol δ. The results for L612M Pol δ (**Fig. 2b**) and M644G Pol ε (**Supplementary Fig. 1**) serve as positive controls for the new analysis of ribonucleotides incorporated into the nascent lagging strand by DNA polymerase α. Importantly, lagging-strand DNA fragments in the *pol1*-L868M *rnh201*Δ strain (**Fig. 2b**, lanes 4 and 12, and **Fig. 2c**, red scan) are shorter than lagging-strand

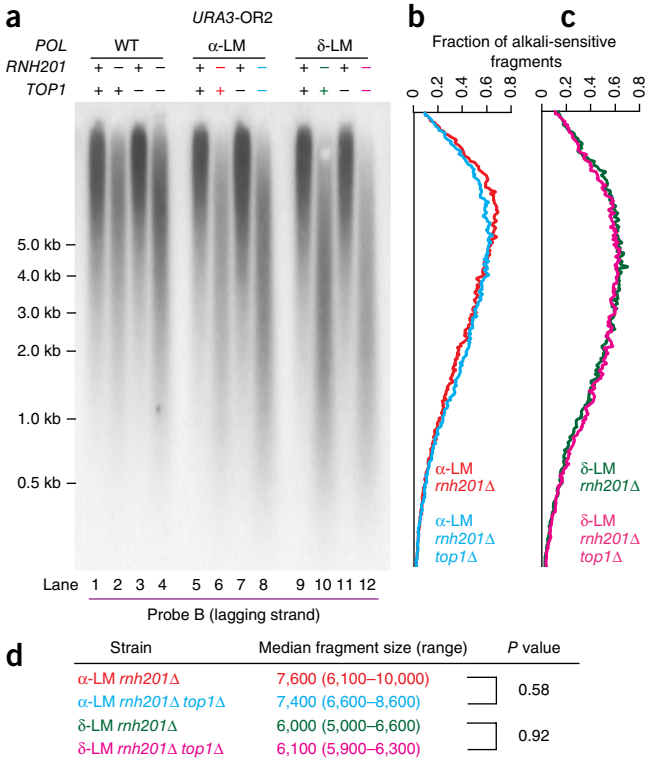



**Figure 3** Lack of Top1-initiated ribonucleotide removal in *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains. (**a**) Detection of alkali-sensitive sites in nascent lagging-strand DNA, performed in strains either proficient or deficient in Top1, with the same approach as in **Figure 2**. All strains contain the *URA3* reporter gene in orientation 2 (OR2, as in **Fig. 2a**). (**b**,**c**) Quantification of data in **a** to determine the fraction of total alkali-sensitive fragments at each position along the membrane. The curves show data from three independent experiments ($n = 3$ independent cell cultures). (**d**) Median DNA fragment sizes (and range) in bases, determined as in ref. 13, with quantitation of the alkali-sensitivity data from five (*pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains) or three (*pol1*-L868M *rnh201*Δ *top1*Δ and *pol3*-L612M *rnh201*Δ *top1*Δ strains) independent experiments. *P* values were calculated by two-tailed Welch's *t* test. Original blot images are in **Supplementary Data Set 1**.

**d**

| Strain | Median fragment size (range) | *P* value |
|---|---|---|
| α-LM *rnh201*Δ | 7,600 (6,100–10,000) | 0.58 |
| α-LM *rnh201*Δ *top1*Δ | 7,400 (6,600–8,600) | |
| δ-LM *rnh201*Δ | 6,000 (5,000–6,600) | 0.92 |
| δ-LM *rnh201*Δ *top1*Δ | 6,100 (5,900–6,300) | |

fragments in the *rnh201*Δ strain encoding wild-type replicases (**Fig. 2b**, lanes 2 and 10, and **Fig. 2c**, blue scan), thus indicating that ribonucleotides are incorporated into the nascent lagging strand by Pol α and that these are subject to RER.

To confirm and extend the observation of the strand specificity of ribonucleotide incorporation during replication *in vivo* seen in the Southern blots, we mapped ribonucleotides by a new procedure, hydrolytic 5′-DNA end–sequencing (HydEn-seq)[28]. In this procedure, DNA ends are generated by alkaline hydrolysis of ribonucleotides in genomic DNA isolated from RNase H2–deficient strains (*rnh201*Δ), fragment libraries are prepared, and the locations of unrepaired ribonucleotides (5′ read ends) are mapped by high-throughput sequencing. We performed HydEn-seq at *ARS306* and the adjacent *URA3* locus (**Fig. 2e**,**f**) and plotted the fraction of replication events in which the bottom strand is replicated as the nascent lagging strand, as estimated by comparison of the HydEn-seq data for the *pol2*-M644G *rnh201*Δ and *pol1*-L868M *rnh201*Δ strains (**Fig. 2e** and **Supplementary Fig. 2**) and the HydEn-seq data for the *pol2*-M644G *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains (**Fig. 2f**). The nascent bottom strand will be the nascent lagging strand immediately to the right of any given bidirectional origin. The abrupt switches in strand specificity clearly identify and correspond to *ARS306* and its neighboring replication origins, *ARS305* and *ARS307*. The bottom strand is replicated as the nascent lagging strand during at least 80% of replication events from *ARS306*, through the *agp1*::*URA3* locus and for more than 8 kb toward *ARS307* (approximately 97.5% of events at *URA3*; **Fig. 2e**,**f**). The high fraction of HydEn-seq ends detected to the right of each origin (**Fig. 2a**) indicate that L868M Pol α and L612M Pol δ synthesize, and insert ribonucleotides into, the nascent lagging strand.

As expected of variations in replication-origin usage and timing in a cell population, the strand preferences are not absolute (hence the scatter in the data points, each depicting 20-nt bins). These data leave open the possibility that forks emanating from *ARS307* may replicate the *URA3* locus adjacent to *ARS306* during perhaps as many as 2.5% of replication events. If ribonucleotide incorporation were a uniform process, then, given that Pol α initiates Okazaki-fragment synthesis and the frequency with which L868M Pol α incorporates ribonucleotides, it would be reasonable to expect Okazaki fragment–sized alkaline-hydrolysis products from *pol1*-L868M *rnh201*Δ genomic DNA. However, our previously published HydEn-seq data have revealed that ribonucleotide incorporation is nonuniform across the genome, with no 5′-DNA ends detected at about 80% of base pairs in these genomes, compared to 5′-DNA-end read counts of 10 or even greater than 100 at other base pairs[28]. This nonuniform distribution means that most ribonucleotides will be concentrated in relatively few Okazaki fragments, thus greatly reducing the likelihood that DNA fragments corresponding to Okazaki fragments of about 200 bases will be observed in Southern blots performed with genomic DNA from the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains. Indeed, DNA fragments of this size are not abundant in the Southern blots that probe the *URA3* locus (**Fig. 2b**) or in visualization of SYBR Gold–stained genomic DNA isolated from the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains that was subjected to alkaline hydrolysis and agarose gel electrophoresis (**Supplementary Fig. 3**).

### Strand specificity of Top1-dependent ribonucleotide removal

A recent study has suggested that yeast Okazaki fragments are about 180 nt long[29]. Pol α is thought to synthesize about 20 nt, with the remainder synthesized by Pol δ (refs. 30–32). Thus, the ability of purified L868M Pol α and L612M Pol δ to incorporate approximately one rNMP per 40 and 300 dNMPs, respectively (**Fig. 1**), predicts that, if RER is the only process that removes them, approximately 50% of Okazaki fragments in *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains might contain an alkali-sensitive site that would result in an Okazaki-size fragment. Although we observed a trace amount of fragments of this size in the scans of genomic DNA from these strains (**Fig. 2b**), the vast majority of fragments were larger. In addition to the possible explanation for this fact mentioned above, it could be that ribonucleotide removal is initiated by Top1 cleavage, which has previously been shown to remove ribonucleotides from the nascent leading strand in the *pol2*-M644G *rnh201*Δ strain[13]. Top1-dependent ribonucleotide removal from the nascent leading strand is evident from the shorter DNA fragments in the *pol2*-M644G *rnh201*Δ *top1*Δ triple-mutant strain as compared to the *pol2*-M644G *rnh201*Δ double-mutant strain[13]. In contrast, we observed little difference in fragment sizes in the equivalent triple- versus double-mutant strains encoding the *pol1*-L868M and *pol3*-L612M variants of the lagging-strand replicases (**Fig. 3**, comparison of lanes 6–8 and 10–12). Thus, Top1 has little or no role in removing ribonucleotides from the nascent lagging strand in the RER-defective *pol1*-L6868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains.

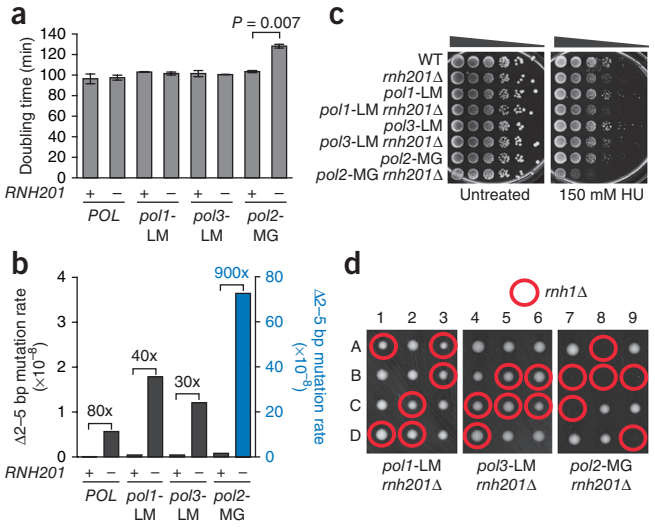



**Figure 4** RNase H2 is dispensable for maintaining genome integrity in strains with increased capacity to incorporate ribonucleotides into lagging-strand DNA. (**a**) Deletion of *RNH201* does not affect growth rate in *pol1*-L868M or *pol3*-L612M mutator strains but does cause an increase in doubling time in the leading-strand mutator variant, *pol2*-M644G. $P = 0.007$ (two-tailed Student's *t* test). Graph shows average values ± s.d. ($n = 3$ independent cell cultures). (**b**) Mutation rates of 2- to 5-bp deletions (Δ2–5 bp), calculated from the data in **Supplementary Table 2** and **Supplementary Figure 5** for the *URA3*-OR2 reporter. The relative fold difference in rate between *RNH201* and *rnh201*Δ is shown above each pair of bars. The rates of 2- to 5-bp deletions corresponding to those displayed are presented in **Supplementary Table 3**. The left *y* axis (black) refers to rates in all strains except *pol2-MG rnh201*Δ. The right *y* axis (blue) corresponds to the rate of 2- to 5-bp deletions in this strain. (**c**) Loss of *RNH201* does not confer sensitivity to 150 mM HU in the *pol1*-LM or pol3-LM strains. The experiment was performed in triplicate; results displayed are representative of $n = 3$ independent cell cultures of this test. (**d**) Tetrad analysis of *RNH1*/*rnh1*Δ diploids in the *pol1*-LM *rnh201*Δ, *pol3*-LM *rnh201*Δ and *pol2*-MG *rnh201*Δ backgrounds. Columns 1–9 are tetrad dissections, and A–D are haploid spore colonies. Plates were photographed after 3 d growth at 30 °C on rich medium.





### Top1-dependent RNA-DNA–damage phenotypes are asymmetric

In *pol2*-M644G *rnh201*Δ strains defective in RER, Top1-dependent removal of ribonucleotides incorporated by M644G Pol ε is associated with several indicators of RNA-DNA damage including slow growth, sensitivity to treatment with the replication inhibitor HU and genome instability in the form of 2- to 5-bp deletions in repetitive sequences[6,13]. The data indicating that Top1 has little or no role in removing ribonucleotides from the nascent lagging strand (**Fig. 3**) suggest that these Top1-dependent RNA-DNA–damage phenotypes may not be observed in the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains. This is indeed the case. In comparison to earlier results with the *pol2*-M644G *rnh201*Δ strain[6,13] and with the single-mutant controls, the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains grow on solid yeast peptone dextrose adenine (YPDA, rich) medium with normal colony sizes (data not shown), and they have doubling times that are equivalent to that of an *RNH201* strain (**Fig. 4a**). This is in contrast to the *pol2*-M644G *rnh201*Δ mutant, which has an increased doubling time (**Fig. 4a**). Also, unlike the *pol2*-M644G *rnh201*Δ mutant, the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains progress normally through the cell cycle (**Supplementary Fig. 4a**). Furthermore, dNTP pools in the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains are not elevated as compared to those in an *rnh201*Δ strain with wild-type replicases (*POL*) (**Supplementary Fig. 4b**). This is distinct from the increased cellular dNTP abundance measured in the *pol2*-M644G *rnh201*Δ mutant compared to an *rnh201*Δ mutant[6]. In the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains, the rate of 2- to 5-bp deletions characteristic of Top1-dependent mutagenesis in RER-defective yeast[6,21,22] is within two- to three-fold of the rate in the *rnh201*Δ single-mutant strain encoding wild-type replicases (**Fig. 4b** and **Supplementary Tables 2** and **3**), as compared to the 130-fold increase in the *pol2*-M644G *rnh201*Δ mutant. In addition, the distribution of 2- to 5-bp deletions within the *URA3* mutational reporter gene is remarkably similar in the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains as compared to their distribution in the *rnh201*Δ single-mutant strain encoding wild-type replicases (**Supplementary Fig. 5** and ref. 22). Moreover, unlike the *pol2*-M644G *rnh201*Δ strain, the *pol1*-L868M *rnh201*Δ and *pol3*-L612M *rnh201*Δ strains are not sensitive to HU (**Fig. 4c**). Thus, although L868M Pol α, L612M Pol δ and M644G Pol ε all incorporate an increased number of ribonucleotides into the nuclear genome, only the *pol2*-M644G variant causes an increase in Top1-dependent RNA-DNA damage in an RER-defective strain.

### Asymmetric synthetic lethality upon deletion of *RNH1*

Tetrad dissection of a diploid *pol2*-M644G *rnh201*Δ strain previously demonstrated that deleting the gene encoding RNase H1 (*RNH1*) results in lethality[24]. That result is recapitulated here (**Fig. 4d**). This synthetic lethality contrasts with the survival and normal haploid spore colony sizes of triple-mutant *pol1*-L868M *rnh201*Δ *rnh1*Δ and *pol3*-L612M *rnh201*Δ *rnh1*Δ strains (**Fig. 4d**). Thus, RNase H1 is essential only when the ribonucleotide burden is increased in the nascent leading strand but not when the ribonucleotide burden is increased in the nascent lagging strand.

## DISCUSSION

Among many known types of DNA repair, two pathways are renowned for their strand specificity: transcription-coupled nucleotide excision repair (TC-NER) and MMR. As its name implies, TC-NER is coupled to transcription rather than replication, and TC-NER selectively removes DNA lesions from the transcribed DNA strand (reviewed in ref. 33). MMR is also strand specific but differs from TC-NER in that it removes replication errors from the newly replicated DNA strand while leaving the parental strand intact, and, importantly, MMR efficiently corrects replication errors present in both the nascent leading and nascent lagging DNA strands[34]. The current study adds a third type of DNA repair whose strand specificity is distinct from those of both TC-NER and MMR. Top1-dependent repair removes a different type of replication error or lesion that is much more abundant than single-base mismatches[13], and, unlike MMR, it preferentially operates on only one of the two the nascent DNA strands. To our knowledge, our observation of this type of strand-specific repair is unique in the DNA-repair field.

The consequences of increased ribonucleotide incorporation into DNA by the three nuclear replicases (summarized in **Table 1**) allow several new conclusions. The observation of increased ribonucleotides in the genome of the *pol1*-L868M *rnh201*Δ strain demonstrates that L868M Pol α incorporates ribonucleotides as it extends RNA primers to initiate DNA synthesis of Okazaki fragments. This result demonstrates that just as for base-base mismatches made by L868M Pol α (refs. 27,35), some ribonucleotides incorporated by Pol α survive the Okazaki fragment–maturation process catalyzed by Pol δ–dependent

**Table 1 Summary of phenotypes in polymerase-variant yeast strains**

| Phenotype | L868M Pol α | L612M Pol δ | M644G Pol ε |
|---|---|---|---|
| RNase H2–dependent repair | Yes | Yes | Yes |
| Top1-dependent removal | No | No | Yes |
| **Phenotypes of *rnh201*Δ derivatives** | | | |
| Doubling time | Normal | Normal | Increased |
| Rate Δ2–5 bp | Wild type | Wild type | Elevated |
| HU sensitivity | Normal | Normal | Sensitive |
| Viability in *rnh1*Δ | Viable | Viable | Inviable |

strand displacement[36]. Here, we show that, just as for ribonucleotides incorporated into the nascent leading strand by Pol ε and into the nascent lagging strand by Pol δ, ribonucleotides incorporated into the nascent lagging strand by Pol α are subject to RER. Collectively, our results demonstrate that ribonucleotides are incorporated by all three replicases in a strand-specific manner, thereby supporting the current model for the primary roles of the polymerases in leading- and lagging-strand replication. This study clearly shows that, like MMR, RER operates on both nascent DNA strands; moreover, it further reveals that the phenotypic consequences of ribonucleotide incorporation by the three variant replicases are different (**Table 1**). Relevant to this concept of asymmetric genome instability, unrepaired ribonucleotides in nascent leading-strand DNA may be recombinogenic. This possibility is supported by previous observations of increased recombination in the absence of RNase H2 activity[37,38].

We suggest three testable ideas for the observed asymmetries in Top1-dependent RNA-DNA damage by ribonucleotides incorporated into the nascent leading and lagging strands and note that these three ideas are not mutually exclusive. The first idea stems from the observation that in the RER defective–replicase variant strains examined here, the density of ribonucleotides is higher in the nascent leading strand (**Supplementary Fig. 1**) than in the nascent lagging strand (**Fig. 2b–d**). Thus, one possibility (model in **Fig. 5a**) is that at least some Top1-dependent RNA-DNA damage is observable only when ribonucleotide density exceeds a certain threshold. This idea is consistent with several observations: (i) as compared to a *pol2*-M644G *rnh201*Δ strain, an *rnh201*Δ strain encoding all wild-type replicases exhibits relatively normal growth and resistance to HU (**Fig. 4a**,**c** and refs. 6,13,39), and (ii) a wild-type strain survives loss of *HNT3* (aprataxin), whereas deletion of *HNT3* in a *pol2*-M644G strain with elevated ribonucleotide incorporation into the nascent leading strand confers slow growth and genome instability—phenotypes that are substantially improved upon loss of RNase H2 activity[19].

A second possibility (**Fig. 5b**) is that ribonucleotides are removed from the nascent lagging strand by a mechanism that is less available to the leading strand. For example, replication errors generated by Pol α are removed by MMR[27,35], and some ribonucleotides may be removed during Okazaki-fragment maturation[36] by processes that are independent of RNase H2 or Top1. An additional repair process, and/or the distinctly nonuniform distribution of newly incorporated ribonucleotides mentioned above, may explain why Okazaki fragment–sized DNA fragments were not observed for the *pol1*-L868M and *pol3*-L612M strains lacking RNase H2 and/or Top1, despite the biochemical potential of L868M Pol α and L612M Pol δ to incorporate a rNMP into one of every two Okazaki fragments (**Fig. 1**).

A third possibility is suggested by the observation that several of the phenotypes that reflect RNA-DNA damage in the *pol2*-M644G *rnh201*Δ strain, for example, slow growth, HU sensitivity and deletion mutagenesis, result from Top1-dependent incision at ribonucleotides in the nascent leading strand[13]. These asymmetries in RNA-DNA damage might result from preferential Top1 incision at ribonucleotides in the nascent leading strand (**Fig. 5c**). These incisions could relieve torsional stress that may accumulate in the nascent leading strand behind the replication fork (ref. 40). Indeed, the possibility of replication-induced negative supercoiling providing a structural change in DNA that regulates different transactions was discussed in a recent review[41]. Theoretically, nicking at rare lesions such as deoxyuracil incorporated during replication might also be important for relief of torsional stress that accumulates behind DNA replication forks, although published data[42] have suggested that deoxyuracil is present in DNA at less than 1% of the frequency of incorporated ribonucleotides. In principle, Top1 incision at ribonucleotides would not be needed to relieve torsional stress in the discontinuously replicated nascent lagging strand, because frequent DNA ends would prevent torsional stress from accumulating. In this scenario, only the nascent leading strand would contain the dirty, unligatable DNA ends

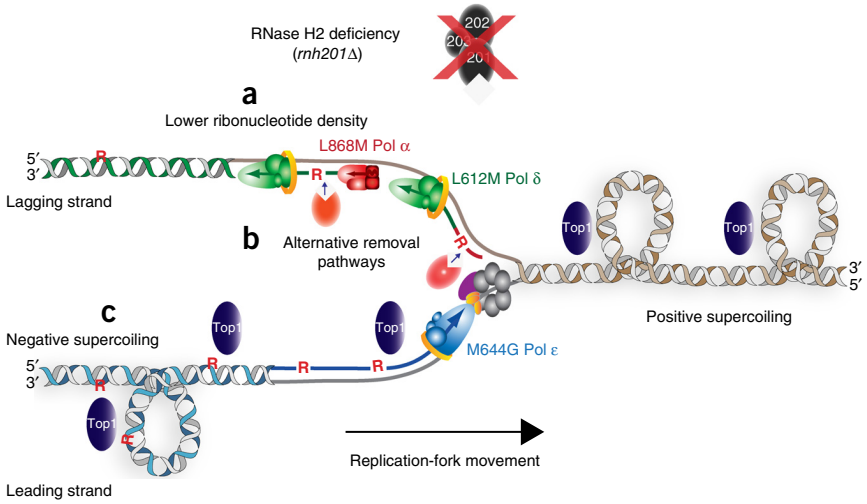



**Figure 5** A model depicting three possibilities for strand-specific consequences of unrepaired ribonucleotides in the genomes of RER-defective yeast strains. (**a**) Failure to observe a Top1-dependent effect in the L868M Pol α or L612M Pol δ strains may be related to these enzymes incorporating fewer ribonucleotides into DNA than does the M644G Pol ε variant. Generation of alternative variants of Pol α and Pol δ that elevate ribonucleotide incorporation into nascent lagging-strand DNA is an approach that could be taken to test this idea. (**b**) A second possibility is the existence of alternative ribonucleotide-repair pathways available on the lagging strand, perhaps involving enzymes required for Okazaki-fragment maturation, such as the Fen1 or Exo1 nucleases or the Dna2 helicase. (**c**) A third idea involves the possibility that negative supercoils accumulate in leading-strand DNA in the wake of the replication fork. This type of superhelical tension may be related to the continuous nature of the leading strand, in contrast to the discontinuity of the lagging strand in the form of preexisting DNA nicks that may allow for DNA rotation and negate the need for relief of such supercoiling by the action of Top1. As a consequence, Top1 incision at ribonucleotides in the nascent leading strand then initiates the RNA-DNA damage observed in a *pol2*-M644G *rnh201*Δ mutant.

generated by Top1 cleavage at ribonucleotides that initiate genome instability. Further support for this model of Top1-dependent effects being leading-strand specific is the demonstration that Top1 localizes to the replisome[43] through an interaction with the CMG complex[44], a helicase that selects Pol ε on the leading strand[45]. In this way, Top1 is well positioned to incise at ribonucleotides incorporated into the nascent leading strand by Pol ε.

## METHODS

Methods and any associated references are available in the online version of the paper.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*


### ACKNOWLEDGMENTS

We thank K. Bebenek, C. Orebaugh and S. Williams for helpful comments on the manuscript and all members of the Kunkel laboratory for thoughtful discussions. We acknowledge the US National Institute of Environmental Health Sciences (NIEHS) Molecular Genetics Core Facility for sequence analysis of 5-FOA–resistant mutants and the NIEHS Flow Cytometry Center for fluorescence-activated cell-sorting analysis. This work was supported by project Z01 ES065070 to T.A.K. from the Division of Intramural Research of the US National Institutes of Health (NIH), NIEHS, by the Swedish Cancer Society to A.C. and by US NIH grant GM032431 to P.M.B.


### AUTHOR CONTRIBUTIONS

J.S.W., A.R.C., S.A.L., L.M., A.B.C. and A.C. designed and performed experiments. P.M.B. provided reagents. All authors were involved in data analysis. J.S.W. and T.A.K. wrote the manuscript, with input from all authors.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.




1. Nick McElhinny, S.A. *et al.* Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. USA* **107**, 4949–4954 (2010).
2. Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E. & Kunkel, T.A. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**, 127–130 (2007).
3. Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M. & Kunkel, T.A. Division of labor at the eukaryotic replication fork. *Mol. Cell* **30**, 137–144 (2008).
4. Wang, J. *et al.* Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087–1099 (1997).
5. Pavlov, Y.I., Shcherbakova, P.V. & Kunkel, T.A. *In vivo* consequences of putative active site mutations in yeast DNA polymerases alpha, epsilon, delta, and zeta. *Genetics* **159**, 47–64 (2001).
6. Nick McElhinny, S.A. *et al.* Genome instability due to ribonucleotide incorporation into DNA. *Nat. Chem. Biol.* **6**, 774–781 (2010).
7. Eder, P.S., Walder, R.Y. & Walder, J.A. Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. *Biochimie* **75**, 123–126 (1993).
8. Rydberg, B. & Game, J. Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc. Natl. Acad. Sci. USA* **99**, 16654–16659 (2002).
9. Sparks, J.L. *et al.* RNase H2-initiated ribonucleotide excision repair. *Mol. Cell* **47**, 980–986 (2012).
10. Lujan, S.A., Williams, J.S., Clausen, A.R., Clark, A.B. & Kunkel, T.A. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol. Cell* **50**, 437–443 (2013).
11. Miyabe, I., Kunkel, T.A. & Carr, A.M. The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet.* **7**, e1002407 (2011).
12. Lujan, S.A. *et al.* Mismatch repair balances leading and lagging strand DNA replication fidelity. *PLoS Genet.* **8**, e1003016 (2012).
13. Williams, J.S. *et al.* Topoisomerase 1-mediated removal of ribonucleotides from nascent leading-strand DNA. *Mol. Cell* **49**, 1010–1015 (2013).
14. Ghodgaonkar, M.M. *et al.* Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol. Cell* **50**, 323–332 (2013).
15. Vengrova, S. & Dalgaard, J.Z. The wild-type *Schizosaccharomyces pombe* mat1 imprint consists of two ribonucleotides. *EMBO Rep.* **7**, 59–65 (2006).
16. Dalgaard, J.Z. Causes and consequences of ribonucleotide incorporation into nuclear DNA. *Trends Genet.* **28**, 592–597 (2012).
17. Williams, J.S. & Kunkel, T.A. Ribonucleotides in DNA: origins, repair and consequences. *DNA Repair (Amst.)* **19**, 27–37 (2014).
18. Caldecott, K.W. Ribose: an internal threat to DNA. *Science* **343**, 260–261 (2014).
19. Tumbale, P., Williams, J.S., Schellenberg, M.J., Kunkel, T.A. & Williams, R.S. Aprataxin resolves adenylated RNA-DNA junctions to maintain genome integrity. *Nature* **506**, 111–115 (2014).
20. Sekiguchi, J. & Shuman, S. Site-specific ribonuclease activity of eukaryotic DNA topoisomerase I. *Mol. Cell* **1**, 89–97 (1997).
21. Kim, N. *et al.* Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science* **332**, 1561–1564 (2011).
22. Clark, A.B., Lujan, S.A., Kissling, G.E. & Kunkel, T.A. Mismatch repair-independent tandem repeat sequence instability resulting from ribonucleotide incorporation by DNA polymerase epsilon. *DNA Repair (Amst.)* **10**, 476–482 (2011).
23. Cho, J.E., Kim, N., Li, Y.C. & Jinks-Robertson, S. Two distinct mechanisms of Topoisomerase 1-dependent mutagenesis in yeast. *DNA Repair (Amst.)* **12**, 205–211 (2013).
24. Lazzaro, F. *et al.* RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol. Cell* **45**, 99–110 (2012).
25. Cerritelli, S.M. & Crouch, R.J. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* **276**, 1494–1505 (2009).
26. Chon, H. *et al.* RNase H2 roles in genome integrity revealed by unlinking its activities. *Nucleic Acids Res.* **41**, 3130–3143 (2013).
27. Nick McElhinny, S.A., Kissling, G.E. & Kunkel, T.A. Differential correction of lagging-strand replication errors made by DNA polymerases α and δ. *Proc. Natl. Acad. Sci. USA* **107**, 21070–21075 (2010).
28. Clausen, A.R. *et al.* Tracking replication enzymology *in vivo* by genome-wide mapping of ribonucleotide incorporation. *Nat. Struct. Mol. Biol.* **22**, 185–191 (2015).
29. Smith, D.J. & Whitehouse, I. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**, 434–438 (2012).
30. Burgers, P.M. Polymerase dynamics at the eukaryotic DNA replication fork. *J. Biol. Chem.* **284**, 4041–4045 (2009).
31. Zheng, L. & Shen, B. Okazaki fragment maturation: nucleases take centre stage. *J. Mol. Cell Biol.* **3**, 23–30 (2011).
32. Balakrishnan, L. & Bambara, R.A. Okazaki fragment metabolism. *Cold Spring Harb. Perspect. Biol.* **5**, a010173 (2013).
33. Marteijn, J.A., Lans, H., Vermeulen, W. & Hoeijmakers, J.H. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* (2014).
34. Lujan, S.A. *et al.* Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* **24**, 1751–1764 (2014).
35. Niimi, A. *et al.* Palm mutants in DNA polymerases alpha and eta alter DNA replication fidelity and translesion activity. *Mol. Cell. Biol.* **24**, 2734–2746 (2004).
36. Garg, P., Stith, C.M., Sabouri, N., Johansson, E. & Burgers, P.M. Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. *Genes Dev.* **18**, 2764–2773 (2004).
37. Aguilera, A. & Klein, H.L. Genetic control of intrachromosomal recombination in *Saccharomyces cerevisiae*. I. Isolation and genetic characterization of hyper-recombination mutations. *Genetics* **119**, 779–790 (1988).
38. Potenski, C.J., Niu, H., Sung, P. & Klein, H.L. Avoidance of ribonucleotide-induced mutations by RNase H2 and Srs2-Exo1 mechanisms. *Nature* **511**, 251–254 (2014).
39. Williams, J.S. *et al.* Proofreading of ribonucleotides inserted into DNA by yeast DNA polymerase epsilon. *DNA Repair (Amst.)* **11**, 649–656 (2012).
40. Wang, J.C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* **3**, 430–440 (2002).
41. Yu, H. & Droge, P. Replication-induced supercoiling: a neglected DNA transaction regulator? *Trends Biochem. Sci.* **39**, 219–220 (2014).
42. Nilsen, H. *et al.* Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. *Mol. Cell* **5**, 1059–1065 (2000).
43. Bermejo, R. *et al.* Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes Dev.* **21**, 1921–1936 (2007).
44. Gambus, A. *et al.* GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.* **8**, 358–366 (2006).
45. Georgescu, R.E. *et al.* Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat. Struct. Mol. Biol.* **21**, 664–670 (2014).

## ONLINE METHODS

**Materials and reagents.** DNA-modification and restriction enzymes were from New England BioLabs, oligonucleotides were from Integrated DNA Technologies, and dNTPs and rNTPs were from Amersham Biosciences.

**Stable incorporation of rNMPs into DNA.** Oligonucleotide primer templates were prepared as previously described[1]. The catalytic subunit of wild-type and L868M Pol α (ref. 35) and the three-subunit form of wild-type and L612M Pol δ (ref. 46) were purified as previously described. Their abilities to stably incorporate rNMPs into DNA were assessed as previously described[1]. Reactions (20 µL) contained 4.0 pmol (100 nM) of a 70-mer template annealed to a 5′-[γ-$^{32}$P]-labeled 40-mer DNA primer and 10 nM Pol α (wild type or L868M) or 40 nM Pol δ (wild type or L612M) enzyme. Reaction mixtures contained the following physiologically relevant nucleotide concentrations: dATP, 16 µM; dCTP, 14 µM; dGTP, 12 µM; dTTP, 30 µM; rATP, 3,000 µM; rCTP, 500 µM; rGTP, 700 µM; and rUTP, 1,700 µM (ref. 1).

**Yeast strains.** *S. cerevisiae* strains used are isogenic derivatives of strain Δ|(−2)|-7B-YUNI300 (*MATa CAN1 his7-2 leu2-Δ::kanMX ura3-Δ trp1-289 ade2-1 lys2-ΔGG2899-2900*)[5]. Relevant strain genotypes are listed in **Supplementary Table 1**. The *URA3* reporter gene was introduced in either orientation 1 (OR1) or orientation 2 (OR2) at position *AGP1* (ref. 47) by transformation of a PCR product containing *URA3* and its endogenous promoter flanked by sequence targeting the reporter to *AGP1*. *rnh201*Δ variants of *POL*, *pol2*-M644G, *pol1*-L868M and *pol3*-L612M were generated by deletion-replacement of *RNH201* via transformation with a PCR product containing the hygromycin-resistance cassette (HYG-R) flanked by 60 nt of sequence homologous to intergenic regions upstream and downstream of the *RNH201* ORF. Transformants that arose after homologous recombination replacing *RNH201* with HYG-R were verified by PCR analysis. *top1*Δ strain construction was performed as previously described[13]. Construction of diploid strains heterogeneous for *rnh1::natMX4* was performed by deletion-replacement of one copy of *RNH1* via transformation with a PCR product containing the nourseothricin-resistance cassette (natMX4) amplified from pAG25 and flanked by 60 nt of sequence homologous to the intergenic regions upstream and downstream of the *RNH1* open reading frame. Transformants that arose from homologous recombination in the diploid strains (homozygous for the polymerase mutator and *rnh201::hphMX4*) were verified by PCR analysis with multiple primer sets (that annealed in locations either internal and external to the *RNH201* or *RNH1* ORF, respectively, or within the appropriate disruption cassette). This analysis demonstrated the absence of the expected genes in all tested strains.

**Alkaline hydrolysis and Southern blotting of genomic DNA.** Genomic DNA was isolated from asynchronously growing cultures at mid-log phase (in YPDA at 30 °C) with the Epicentre Yeast DNA purification kit (MPY80200). 5 µg of DNA was treated with 0.3 M KOH for 2 h at 55 °C and subjected to alkaline hydrolysis and alkaline agarose electrophoresis as previously described[13]. The gel was neutralized, and the DNA was transferred to a nylon membrane (Hybond N+) by capillary action in alkaline transfer buffer (0.4 N NaOH and 1 M NaCl) overnight. Southern analysis was performed with single-stranded radiolabeled probes prepared from a PCR-amplified fragment of the *URA3* reporter gene integrated at the *AGP1* locus on chromosome III with a previously described procedure[13]. Quantitation of alkali-sensitive fragments was performed with ImageQuant software (GE Healthcare). The fraction was calculated by division of the radioactive intensity (arbitrary units) at 0.1-mm intervals by the total intensity for each lane. Mean fragment-size determination was performed with data from either two (**Fig. 2**) or three (**Fig. 3**) independent experiments, as described in ref. 13. Mean DNA fragment sizes (±s.d.) and *P* values were determined as in ref. 13. SYBR Gold Nucleic Acid Staining (Life Technologies) was for 2 h at room temperature with a 1:10,000 dilution of the stock solution in 1× TBE buffer.

**HydEn-seq protocol and analysis.** 5′ DNA ends were generated by alkaline hydrolysis of genomic DNA and mapped by high-throughput sequencing (HydEn-seq) as described in ref. 28.

*Data analysis: calculation of how often a given locus is replicated as either the nascent leading or lagging strand.* Herein, data sets are defined as either leading-strand biased (lead) or lagging-strand biased (lag), but it is important to note that the method does not rely on an *ab initio* assignment of these characters. So long as the data sets used have opposite strand biases, the fraction of replication events that generate HydEn-seq ends on the bottom strand ($F_{ij,b}$ below) in the lagging strand–biased data set will approach 1 to the right of each bidirectional origin and will approach 0 to the left. The opposite will be true for the leading strand–biased data set.

Define $R_{ij,b}$ in bin $i$ of data set $j$ (lead or lag) as the true ratio of replication events that generate HydEn-seq reads that map to the bottom strand (bottom; $b$) versus the top strand (top; $t$). If ribonucleotides inserted by replicative polymerases were the only source of HydEn-seq ends, and if ribonucleotide insertion were uniform, then $R_{ij,b}$ would be a simple ratio of forward-mapping ($n_{ij,t}$) and reverse-mapping ($n_{ij,b}$) end counts:

$$R_{ij,b} = \frac{n_{ij,b}}{n_{ij,t}} \tag{1}$$

$R_{ij,b}$ would then serve as a measure of strand-biased replication, and leading strand–biased (lead) and lagging strand–biased (lag) ratios would be multiplicative inverses of each other:

$$R_{i\ lag,b} = R_{i\ lead,t} = R_{i\ lead,b}{}^{-1} \tag{2}$$

However, ribonucleotide incorporation is neither uniform nor the sole source of DNA ends detected via HydEn-seq. Various factors, such as sequence context, modulate the ratio by a multiplicative factor ($W_i$) that varies from bin to bin but should be largely independent of genetic manipulations to the replicative polymerases. Thus the ratio ($O_{ij,t}$) of observed end counts ($m_{ij}$) is the product of the true ratio of replication events and the multiplicative factor:

$$O_{ij,b} = \frac{m_{ij,b}}{m_{ij,t}} = R_{ij,b}W_i \tag{3}$$

The inverse relationship in equation (2) means that the ratio of observed ratios is equal to the square of the true ratio in the numerator:

$$\frac{o_{i\ lag,b}}{o_{i\ lead,b}} = \frac{R_{i\ lag,b}W}{R_{i\ lead,b}W} = R_{i\ lag,b}{}^2 \tag{4}$$

Thus each true ratio of replication events acting on a given strand may be calculated from observed HydEn-seq end counts from strains engineered to have opposite ribonucleotide-incorporation strand biases, either as a ratio (equations (5) and (6)) or as a fraction of replication events ($F_{ij,b}$; equation (7)):

$$R_{i\ lag,b} = \sqrt{\frac{o_{i\ lag,b}}{o_{i\ lead,b}}} \tag{5}$$

$$R_{i\ lead,b} = \sqrt{\frac{o_{i\ lead,b}}{o_{i\ lag,b}}} \tag{6}$$

$$F_{ij,b} = \frac{n_{ij,b}}{n_{ij,b} + n_{ij,t}} = \frac{1}{1 + \frac{n_{ij,t}}{n_{ij,b}}} = \frac{1}{1 + R_{ij,b}{}^{-1}} \tag{7}$$

In practice, if no ends were counted for a particular strand in a given bin, $m_{ij}$ was set to 0.5, thus making $O_{ij,bottom}$ a boundary estimate. $O_{ij,bottom}$ was not calculated directly from $m_{ij}$ but rather was calculated from the average of normalized counts ($\overline{m_{ij}}$) from $N$ replicate HydEn-seq experiments (equation (8)), with $M_k$ representing the total read count in replicate $k$:

$$\overline{m_{ij}} = \left(\sum_{k=0}^{N} (m_{ik} / M_k)\right) / N \tag{8}$$

**Phenotype analyses.** Strains were grown in YPDA medium (1% yeast extract, 2% bacto-peptone, 250 mg/l adenine, 2% dextrose, and 2% agar for plates). Doubling time ($D_t$) values were calculated from cultures in the logarithmic phase of growth in rich medium at 30 °C. The experiment was performed in triplicate, and data are displayed as the mean $D_t$ ± s.d. Spot dilution assays were performed by plating

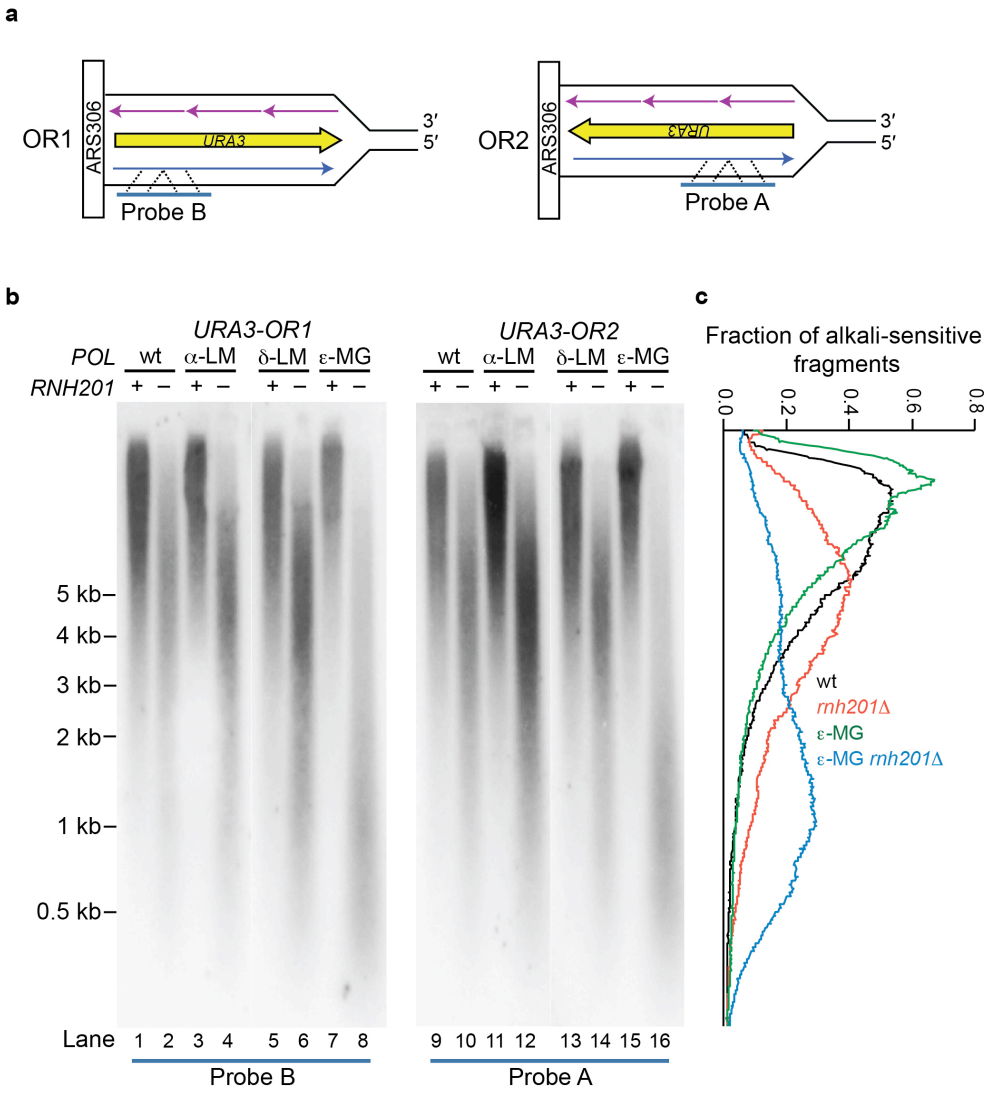



doi:10.1038/nsmb.2989

serial (ten-fold) dilutions of asynchronously growing mid-log-phase cultures from the indicated strains onto YPDA agar plates with or without 150 mM HU (Sigma H8627). Plates were incubated at 30 °C and photographed after 3 d of growth. Three independent biological replicates of this test were performed. Flow cytometry to determine DNA content was performed as previously described[13]. Histograms and plots are representative of at least two independent experiments. Nucleotide pool measurements were analyzed by HPLC as previously described[6,13], and the data displayed are from two independent experiments.

**Measurement of spontaneous mutation rates and sequence analysis.** Spontaneous mutation rates were determined with fluctuation analysis as described previously[48]. Genomic DNA from independent 5-FOA–resistant colonies was isolated, and the *ura3* gene was PCR amplified and sequenced. Rates of individual mutation classes were calculated by multiplication of the proportion of each mutation type by the total mutation rate for each strain.

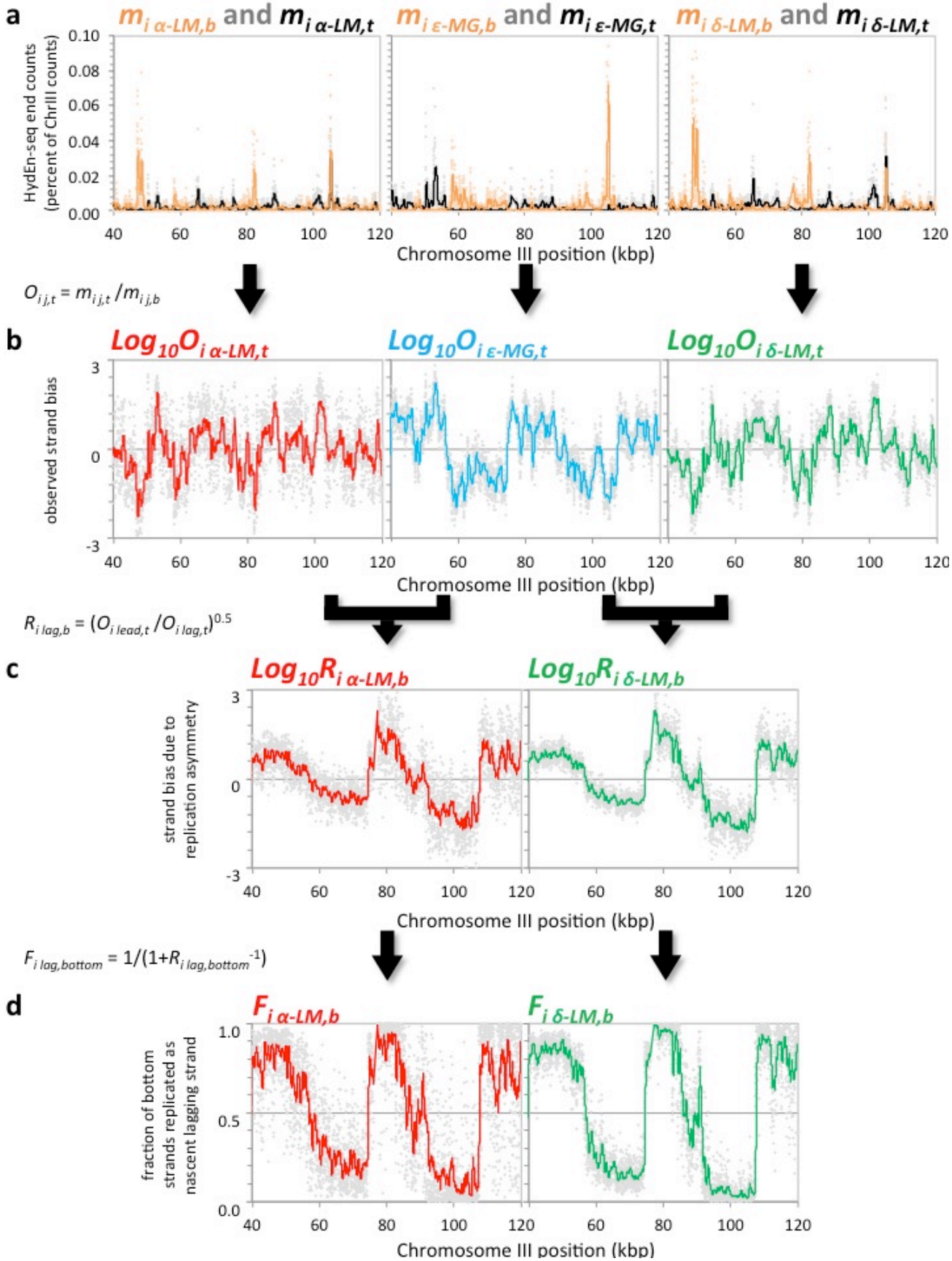

46. Burgers, P.M. & Gerik, K.J. Structure and processivity of two forms of *Saccharomyces cerevisiae* DNA polymerase delta. *J. Biol. Chem.* **273**, 19756–19762 (1998).
47. Shcherbakova, P.V. & Kunkel, T.A. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. *Mol. Cell. Biol.* **19**, 3177–3183 (1999).
48. Pavlov, Y.I., Newlon, C.S. & Kunkel, T.A. Yeast origins establish a strand bias for replicational mutagenesis. *Mol. Cell* **10**, 207–213 (2002).




**Supplementary Figure 1**

Probing for ribonucleotides in nascent leading-strand DNA.

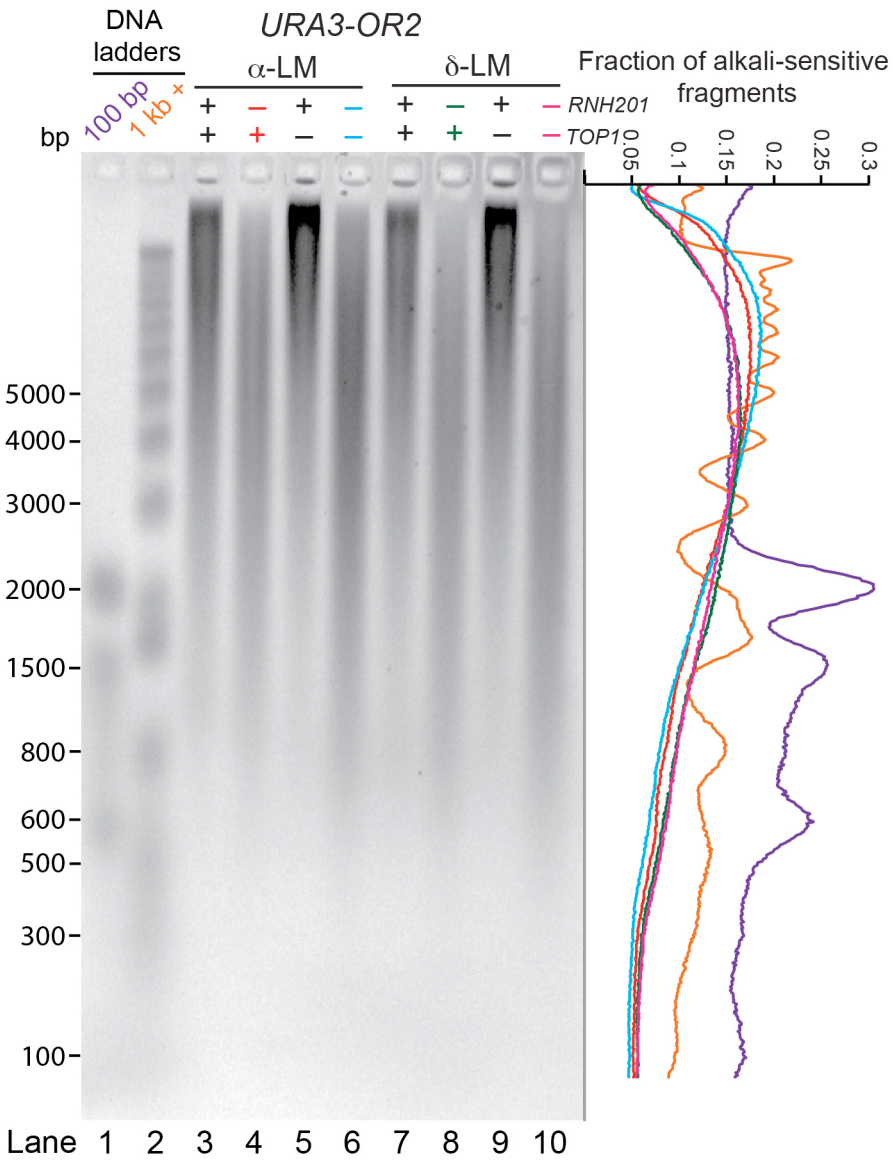
(**a**) The annealing locations of radiolabeled leading strand-specific probes on the *URA3* reporter in OR1 or OR2. (**b**) Detection of alkali-sensitive sites in nascent leading strand DNA synthesized by Pol ε was performed as described [13]. Smaller DNA fragments observed for the α-LM *rnh201*Δ and δ-LM *rnh201*Δ mutants when using probes that anneal to the nascent leading strand (lanes 4, 6, 12 and 14) may be related to the close proximity of *URA3* to *ARS306* (1.6 kb). These alkali-sensitive sites may arise during ribonucleotide incorporation by Pols α or δ into the nascent lagging strand during bidirectional synthesis proceeding from this origin in the opposite direction (to the left of the origin in panel a). In addition, these small fragments that hybridize to the 'leading strand' probe may be generated by synthesis performed by L868M Pol α or L612M Pol δ as they replicate from the adjacent *ARS307* origin. The same explanation applies for small DNA fragments observed for the ε-MG *rnh201*Δ mutant in Figure 2 (main text) when using probes that anneal to the nascent lagging strand (lanes 8 and 16). (**c**) The average fraction of alkali-sensitive fragments along the membrane was determined by quantifying the radioactive signal using data from in panel b (for both Probes A and B) from two independent experiments.


a
$m_{i\,\alpha\text{-}LM,b}$ and $m_{i\,\alpha\text{-}LM,t}$
$m_{i\,\varepsilon\text{-}MG,b}$ and $m_{i\,\varepsilon\text{-}MG,t}$
$m_{i\,\delta\text{-}LM,b}$ and $m_{i\,\delta\text{-}LM,t}$
HydEn-seq end counts (percent of ChrIII counts)
0.10
0.08
0.06
0.04
0.02
0.00
40 60 80 100 120 60 80 100 120 60 80 100 120
Chromosome III position (kbp)
$O_{i\,j,t} = m_{i\,j,t} / m_{i\,j,b}$
b
$Log_{10}O_{i\,\alpha\text{-}LM,t}$
$Log_{10}O_{i\,\varepsilon\text{-}MG,t}$
$Log_{10}O_{i\,\delta\text{-}LM,t}$
observed strand bias
3
0
-3
40 60 80 100 120 60 80 100 120 60 80 100 120
Chromosome III position (kbp)
$R_{i\,lag,b} = (O_{i\,lead,t} / O_{i\,lag,t})^{0.5}$
c
$Log_{10}R_{i\,\alpha\text{-}LM,b}$
$Log_{10}R_{i\,\delta\text{-}LM,b}$
strand bias due to replication asymmetry
3
0
-3
40 60 80 100 120 60 80 100 120
Chromosome III position (kbp)
$F_{i\,lag,bottom} = 1/(1+R_{i\,lag,bottom}^{-1})$
d
$F_{i\,\alpha\text{-}LM,b}$
$F_{i\,\delta\text{-}LM,b}$
fraction of bottom strands replicated as nascent lagging strand
1.0
0.5
0.0
40 60 80 100 120 60 80 100 120
Chromosome III position (kbp)

**Supplementary Figure 2**

Workflow for calculating fractional replication-strand bias from HydEn-seq end counts without background subtraction or internal standards.

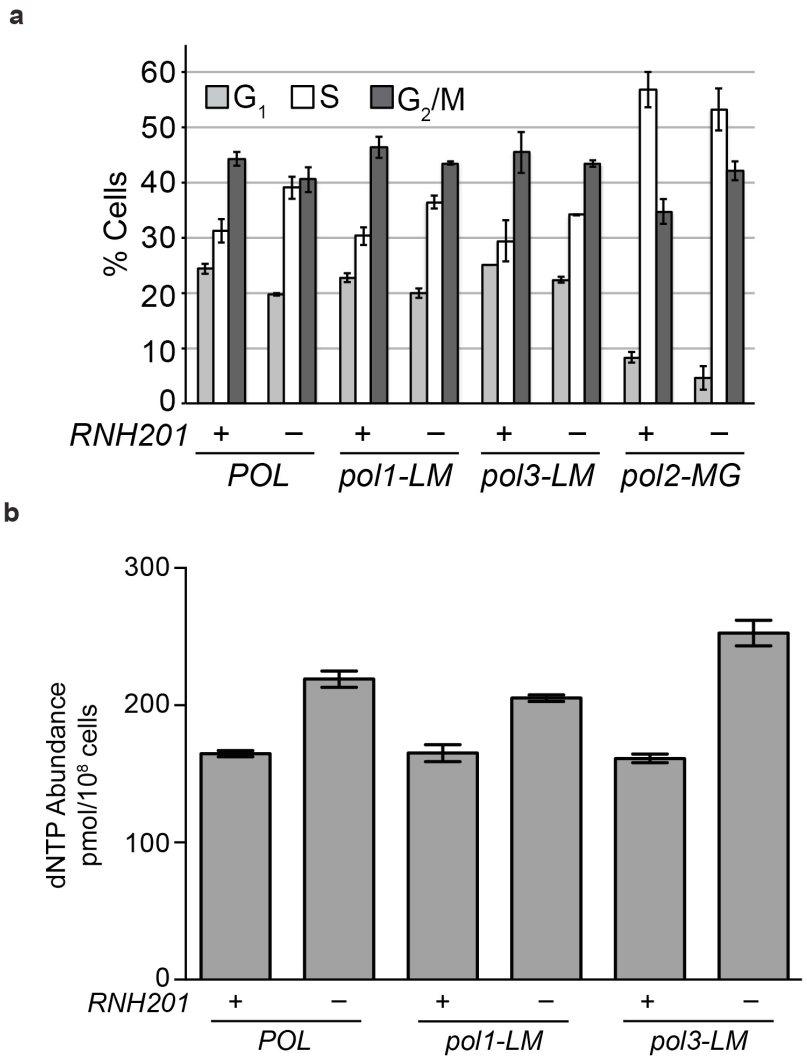
Each set of arrows indicates the application of the equation shown to the left. Gray data points represent bins of 20 base pairs (bp). Labels above each graph are relate the variables in the equations to the values for which they are color-coded. Solid lines are moving averages over 25 bins (500 bp). (**a**) HydEn-seq end counts for reads mapping to the forward (black) and reverse (orange) strands from RNase H2-deficient *pol1-L868M* (α-LM; left), *pol2-M644G* (ε-MG; middle), and *pol3-L612M* (δ-LM; right) strains. (**b**) The observed strand bias, expressed as a log-ratio between forward- and reverse-strand end counts. (**c**) An approximation of the true replication strand bias log-ratio. Extra-replicative contributions were removed by comparing strains with opposite biases. The greater the biases in the chosen strains, the closer the approximation. (**d**) The fraction of replication events in which the bottom strand originates as the nascent lagging strand.




**Supplementary Figure 3**

Alkali-sensitive Okazaki fragment–sized DNA is not observed in the α-LM *rnh201*Δ or δ-LM *rnh201*Δ strains.

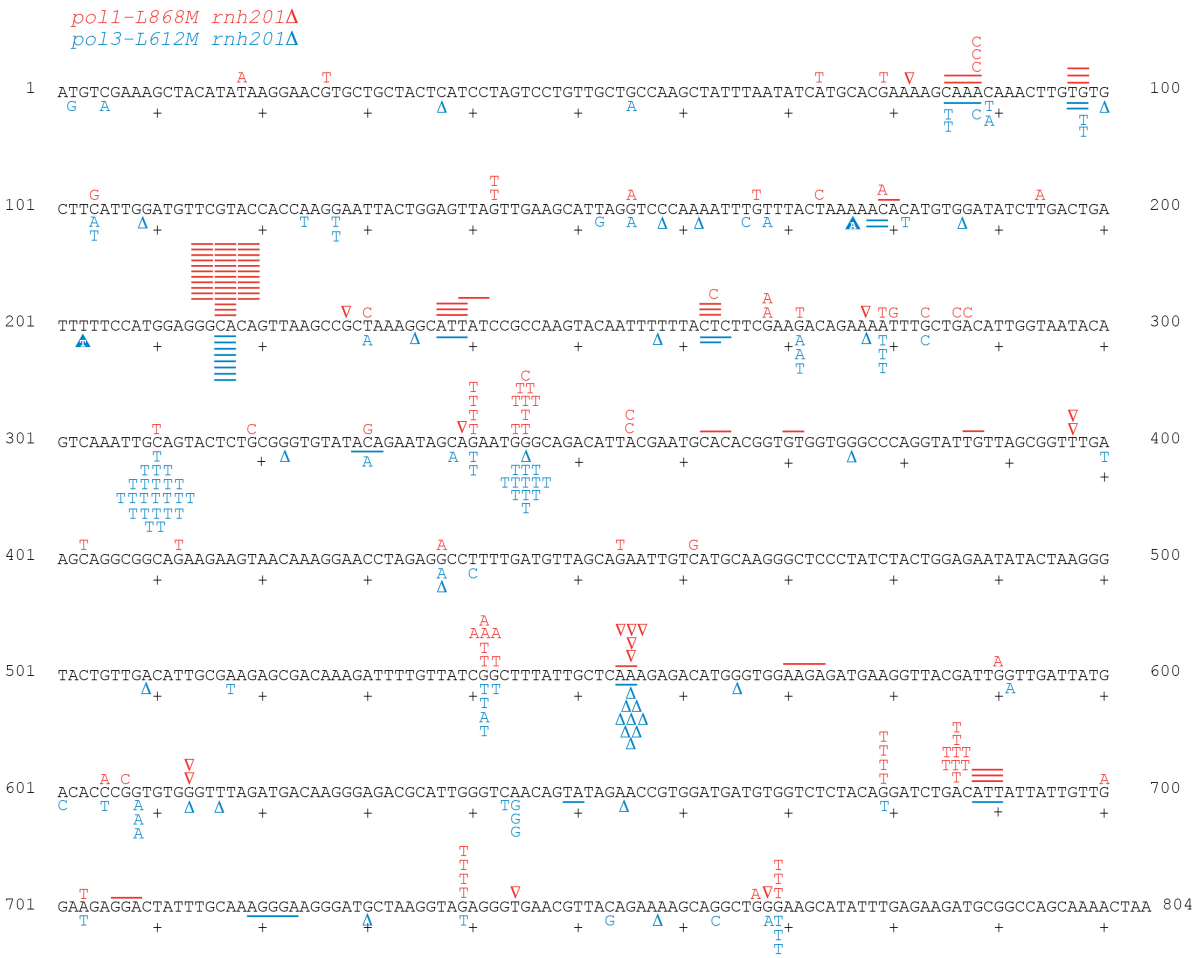
Purified genomic DNA was subjected to alkaline hydrolysis, alkaline-agarose electrophoresis and stained with SYBR® Gold Nucleic Acid Gel Stain, a sensitive fluorescent stain that can be used for detection of single strand DNA. The fraction of alkali-sensitive fragments was calculated by dividing the fluorescence intensity (arbitrary units) at each position along the gel by the total intensity for each lane. The experiment was performed in duplicate, and a representative gel image and quantitation is displayed. We note that Okazaki fragment-sized DNA (e.g. 150-300 base pair) fragments were not observed among the products of the alkaline hydrolysis of genomic DNA from the α-LM *rnh201*Δ or δ-LM *rnh201*Δ strains.




**Supplementary Figure 4**

Asymmetric phenotypes are associated with unrepaired ribonucleotides incorporated into nascent leading- versus lagging-strand DNA.

**(a)** Cell cycle progression is not affected by loss of *RNH201* in the *pol1-LM* or the *pol3-LM* mutator strains with an increased number of unrepaired ribonucleotides in the nascent lagging strand. Cells were grown to mid-log phase at 30°C and processed for flow cytometry as described in [13]. The experiment was performed in duplicate and data are displayed as the mean % ± standard error. **(b)** The increase in total dNTP abundance conferred by loss of *RNH201* is not significantly enhanced in the *pol1-LM* or the *pol3-LM* lagging strand polymerase mutator strains. Data for total dNTP abundance (dCTP, dATP, dTTP and dGTP) is displayed as the mean ± standard error. Each strain genotype was independently analyzed twice.




**Supplementary Figure 5**

*URA3* mutation spectra for *pol1-L868M rnh201*Δ and *pol3-L612M rnh201*Δ.

The coding strand of the 804 base pair *URA3* open reading frame is shown. The sequence changes observed in independent *ura3* mutants are depicted above the coding sequence for *pol1-L868M rnh201*Δ in red and below the coding sequence for *pol3-L612M rnh201*Δ in blue. Letters indicate single base substitutions, open triangles indicate single base deletions, and short lines above or below the coding sequence indicate multibase base deletions of between 2 and 5 base pairs. All strains had the *URA3* reporter in OR2.

## SUPPLEMENTARY INFORMATION

**Supplementary Table 1.** *S. cerevisiae* strains

| Name | Strain | Relevant Genotype | Source |
|---|---|---|---|
| *wt (POL)* | SNM8 | *agp1*::*URA3-OR1* | [6] |
| *wt (POL)* | SNM18 | *agp1*::*URA3-OR2* | [6] |
| *rnh201*Δ | SNM106 | *rnh201*::*hphMX4 agp1*::*URA3-OR1* | [6] |
| *rnh201*Δ | SNM114 | *rnh201*::*hphMX4 agp1*::*URA3-OR2* | [6] |
| *top1*Δ *rnh201*Δ | YJW77 | *top1*::*natMX4 rnh201*::*hphMX4 agp1*::*URA3-OR2* | [13] |
| *top1*Δ | YJW81 | *top1*::*natMX4 agp1*::*URA3-OR2* | [13] |
| *pol2-M644G* | SNM70 | *pol2-M644G agp1*::*URA3-OR1* | [6] |
| *pol2-M644G* | SNM77 | *pol2-M644G agp1*::*URA3-OR2* | [6] |
| *pol2-M644G rnh201*Δ | SNM120 | *pol2-M644G rnh201*::*hphMX4 agp1*::*URA3-OR1* | [6] |
| *pol2-M644G rnh201*Δ | SNM127 | *pol2-M644G rnh201*::*hphMX4 agp1*::*URA3-OR2* | [6] |
| *pol1-L868M* | *SNM15* | *pol1-L868M agp1*::*URA3-OR1* | [27] |
| *pol1-L868M* | *SNM27* | *pol1-L868M agp1*::*URA3-OR2* | [27] |
| *pol1-L868M rnh201*Δ | YJW13 | *pol1-L868M rnh201*::*hphMX4 agp1*::*URA3-OR1* | This study |
| *pol1-L868M rnh201*Δ | YJW17 | *pol1-L868M rnh201*::*hphMX4 agp1*::*URA3-OR2* | This study |
| *pol1-L868M* | YJW289 | *pol1-L868M top1*::*natMX4* | This study |

| | | | |
|---|---|---|---|
| *top1*Δ | | *agp1*::*URA3-OR2* | |
| *pol1-L868M rnh201*Δ *top1*Δ | YJW293 | *pol1-L868M rnh201*::*hphMX4 top1*::*natMX4 agp1*::*URA3-OR2* | This study |
| *pol3-L612M* | SNM11 | *pol3-L612M agp1*::*URA3-OR1* | [27] |
| *pol3-L612M* | SNM24 | *pol3-L612M agp1*::*URA3-OR2* | [27] |
| *pol3-L612M rnh201*Δ | YJW11 | *pol3-L612M rnh201*::*hphMX4 agp1*::*URA3-OR1* | This study |
| *pol3-L612M rnh201*Δ | YJW15 | *pol3-L612M rnh201*::*hphMX4 agp1*::*URA3-OR2* | This study |
| *pol3-L612M top1*Δ | YJW97 | *pol3-L612M top1*::*natMX4 agp1*::*URA3-OR2* | This study |
| *pol3-L612M rnh201*Δ *top1*Δ | YJW99 | *pol3-L612M rnh201*::*hphMX4 top1*::*natMX4 agp1*::*URA3-OR2* | This study |

**Supplementary Table 2.** Mutation rates and sequencing data for the *pol1-L868M* and *pol3-L612M* ± *RNH201* strains.

| *RNH201* | + | Δ | + | Δ |
|---|---|---|---|---|
| Strain | *pol1-L868M* | | *pol3-L612M* | |
| Mutation rate (x $10^{-8}$) | 7.6 | 8.6 | 9.1 | 11 |
| 95% CI | (5.8-11) | (4.9-13) | (6.8-12) | (8.2-13) |
| *ura3* mutants sequenced | 185 | 285 | 196 | 188 |
| Single base substitutions | 115 | 81 | 139 | 95 |
| Single base deletions | 6 | 15 | 24 | 29 |
| Total 2-5 base deletions* | 0 | 59 | 0 | 21 |
| Other[†] | 8 | 6 | 2 | 0 |

*Refers to 2-5 base deletions that occur in repeat sequences. [†]Other mutations include mutations involving multiple bases (deletions of >5 bases, insertions of ≥1 base, duplications and complex mutations). 95% confidence limits (CI) are shown in parentheses and were calculated as described [46]. All strains have the *URA3* reporter in OR2. The sequencing data for the *RNH201* (+) strains is from [27]. The total number of $FOA^r$ mutants sequenced exceeds the number of mutations because some mutants had no sequence change in the *URA3* open reading frame. These mutants may have sequence changes in the *URA3* promoter or another gene that affects resistance to 5-FOA. Because these mutants contribute to the overall mutation rate, they are included in the calculation of rates for individual mutation classes.

**Supplementary Table 3.** Specific mutation rates ($x10^{-8}$) for individual mutation classes.

| Strain | Overall Rate | Δ2–5 bp | BPS | Single base Δ |
|---|---|---|---|---|
| *POL* | 0.91 | 0.0076 | 0.52 | 0.069 |
| *POL rnh201*Δ | 1.6 | 0.56 | 0.52 | 0.082 |
| *pol1-LM* | 7.6 | ≤0.041 | 4.7 | 0.25 |
| *pol1-LM rnh201*Δ | 8.6 | 1.78 | 2.4 | 0.45 |
| *pol3-LM* | 9.1 | ≤0.046 | 6.5 | 1.1 |
| *pol3-LM rnh201*Δ | 11 | 1.2 | 5.6 | 1.7 |
| *pol2-MG* | 6.8 | 0.083 | 2.7 | 0.25 |
| *pol2-MG rnh201*Δ | 82 | 73 | 4.8 | 4.2 |

BPS; base pair substitutions. Specific mutation rates were calculated as the proportion of each type of event among the total mutants sequenced, multiplied by the mutation rate for each strain (using the data in Supplementary Table 2 and Supplementary Figure 5). The sequencing data for the *POL* (wt), *pol1-LM* and *pol3-LM* strains are from [27]. The sequencing data for the *pol2-MG* strain is from [10]. The sequencing data for the *POL rnh201*Δ strain is from [22]. The sequencing data for the *pol2-MG rnh201Δ* strain is from [6]. All strains have the *URA3* reporter in OR2.

## SUPPLEMENTARY REFERENCES


46. Shcherbakova, P.V. & Kunkel, T.A. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. *Mol Cell Biol* **19**, 3177-83 (1999).

47. Burgers, P.M. & Gerik, K.J. Structure and processivity of two forms of Saccharomyces cerevisiae DNA polymerase delta. *J Biol Chem* **273**, 19756-62 (1998).

48. Pavlov, Y.I., Newlon, C.S. & Kunkel, T.A. Yeast origins establish a strand bias for replicational mutagenesis. *Mol Cell* **10**, 207-13 (2002).