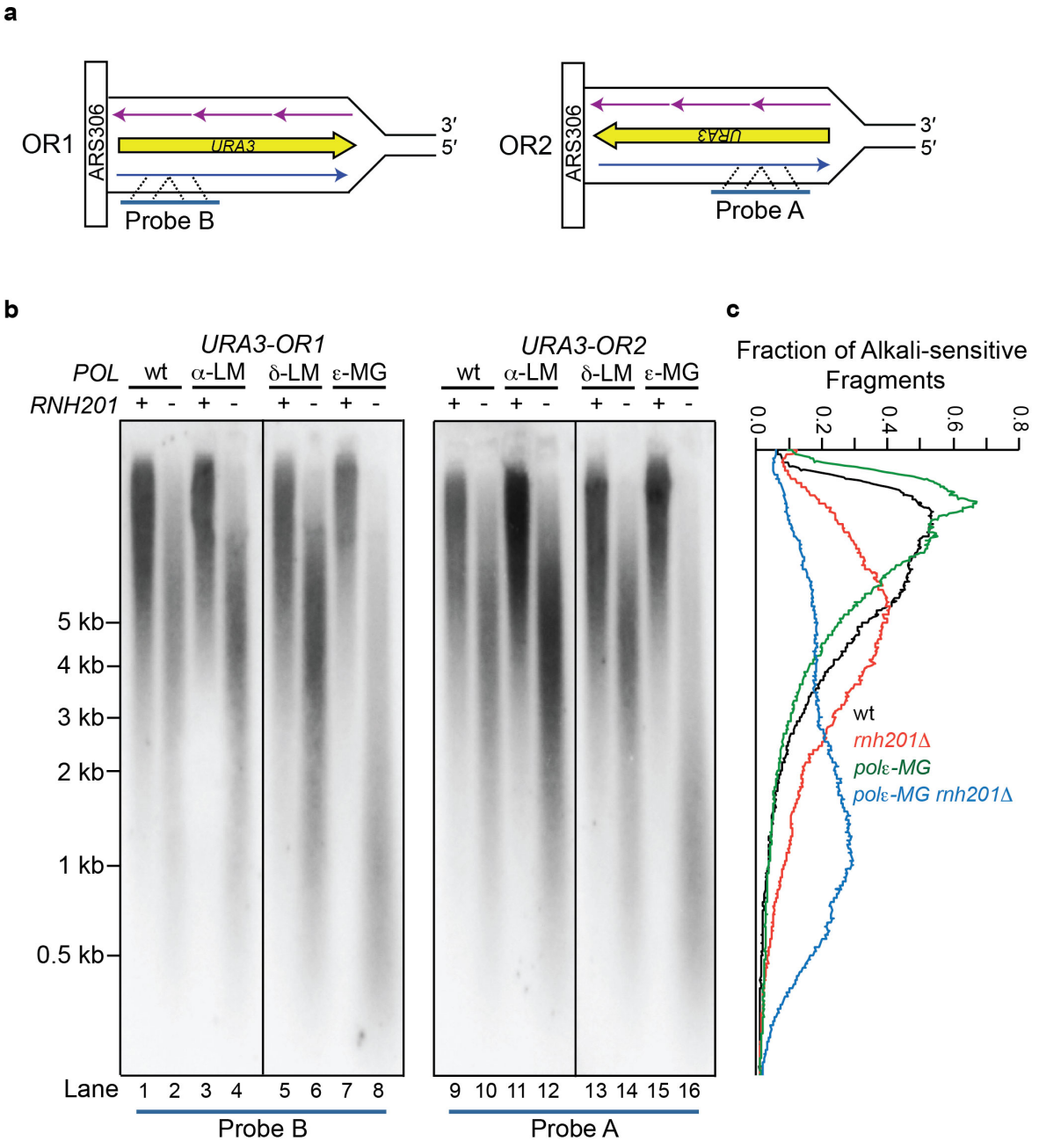
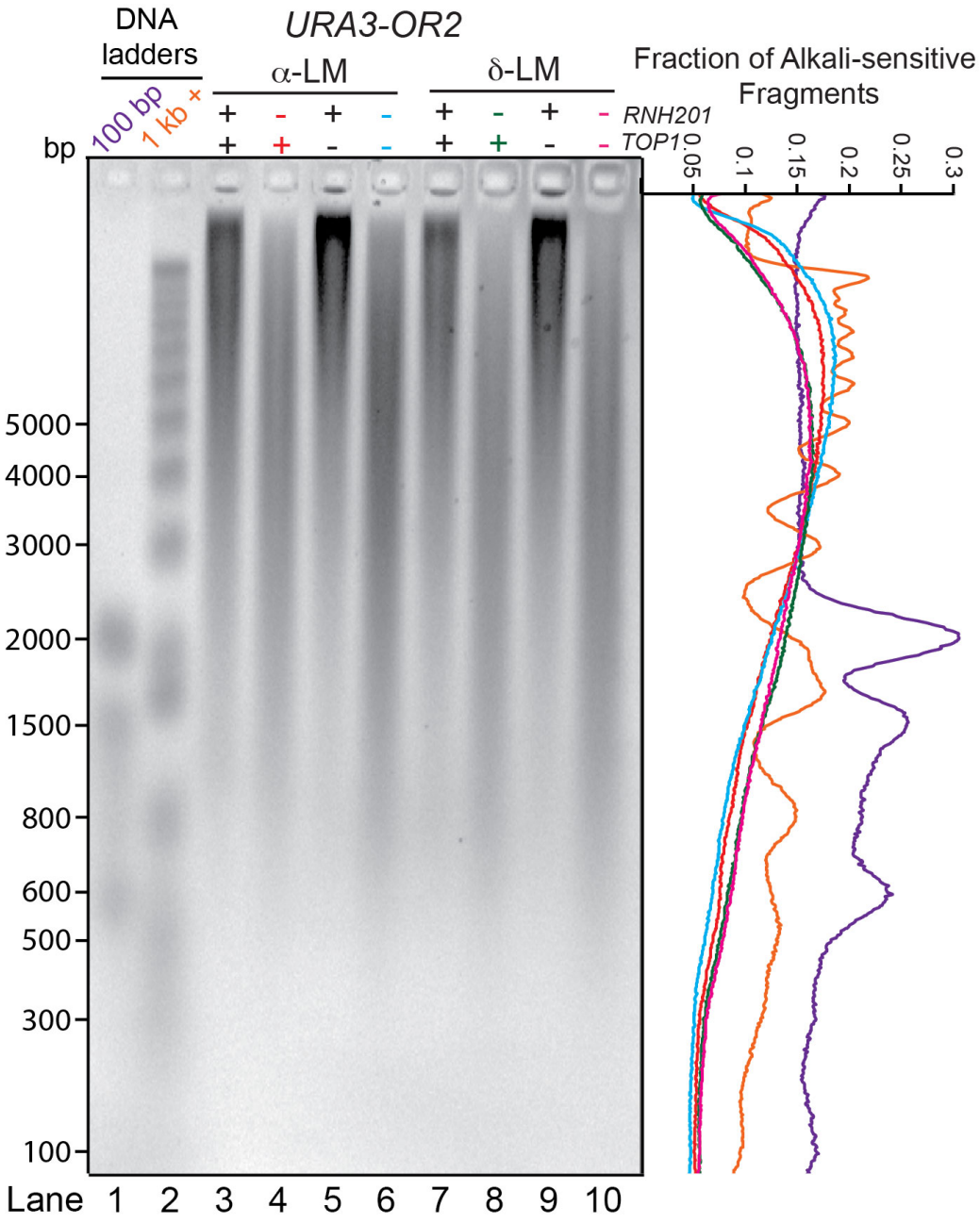
SUPPLEMENTARY INFORMATION

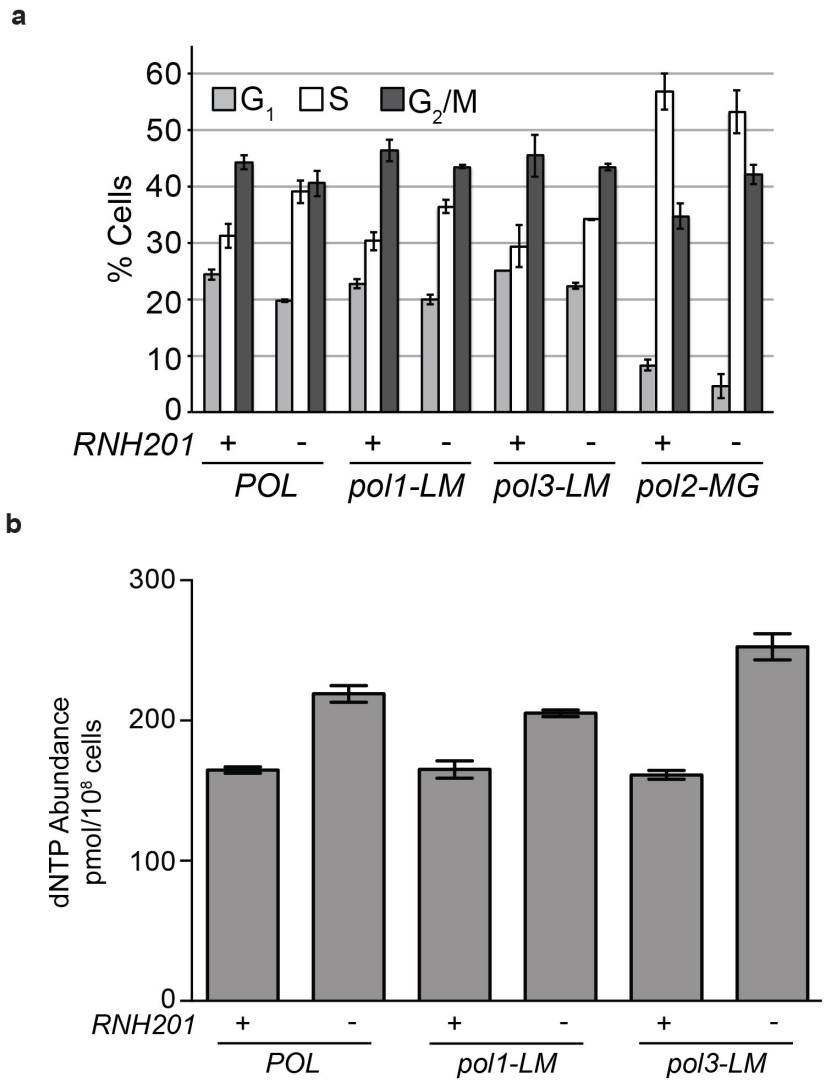


Supplementary Figure 1. Probing for ribonucleotides in nascent leading strand DNA. (**a**) The annealing locations of radiolabeled leading strand-specific probes on the *URA3* reporter in OR1 or OR2. (**b**) Detection of alkali-sensitive sites in nascent leading strand DNA synthesized by Pol ε

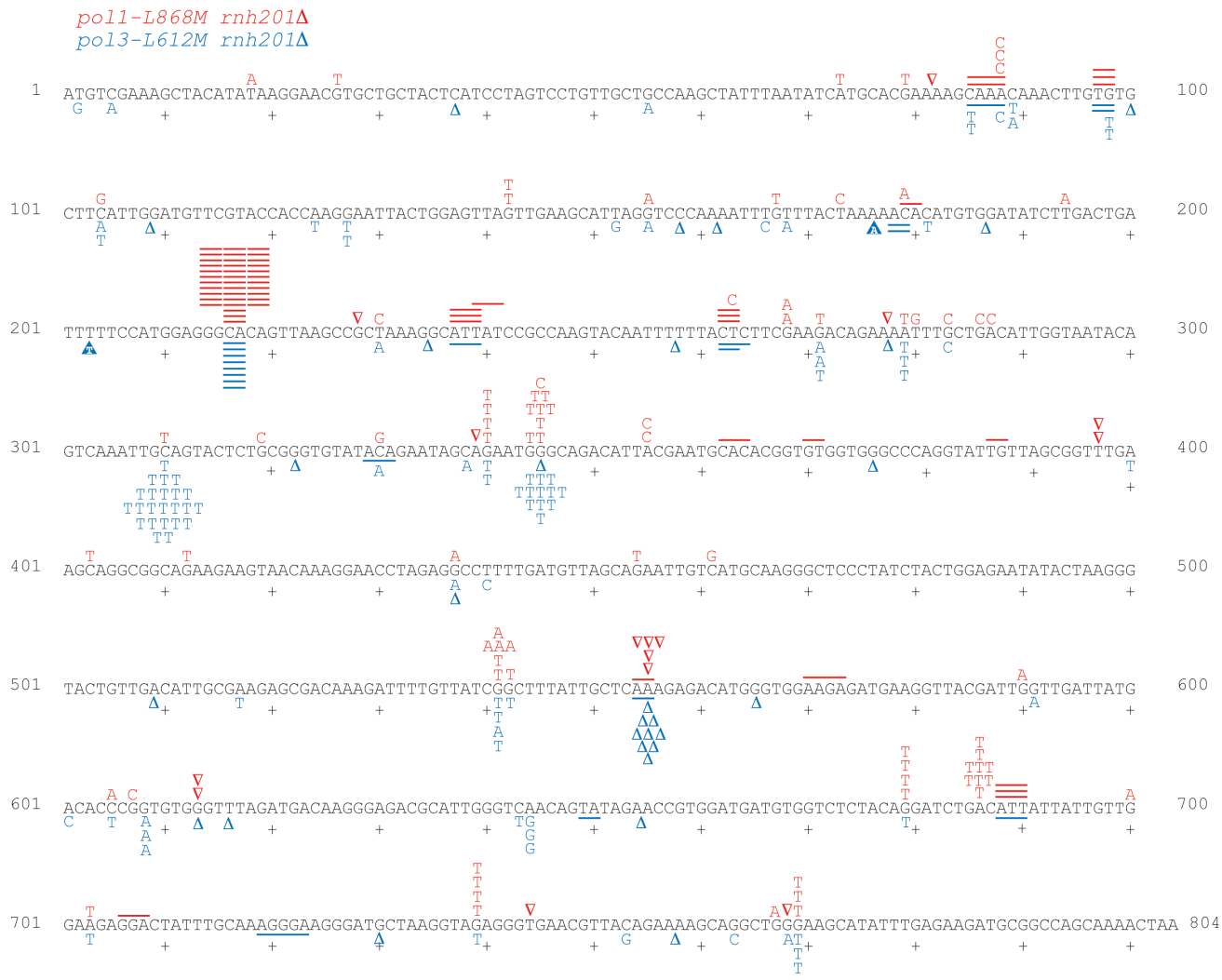
was performed as described [13]. Smaller DNA fragments observed for the α-LM *rnh201*Δ and δ-LM *rnh201*Δ mutants when using probes that anneal to the nascent leading strand (lanes 4, 6, 12 and 14) may be related to the close proximity of *URA3* to *ARS306* (1.6 kb). These alkali-sensitive sites may 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. (**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.



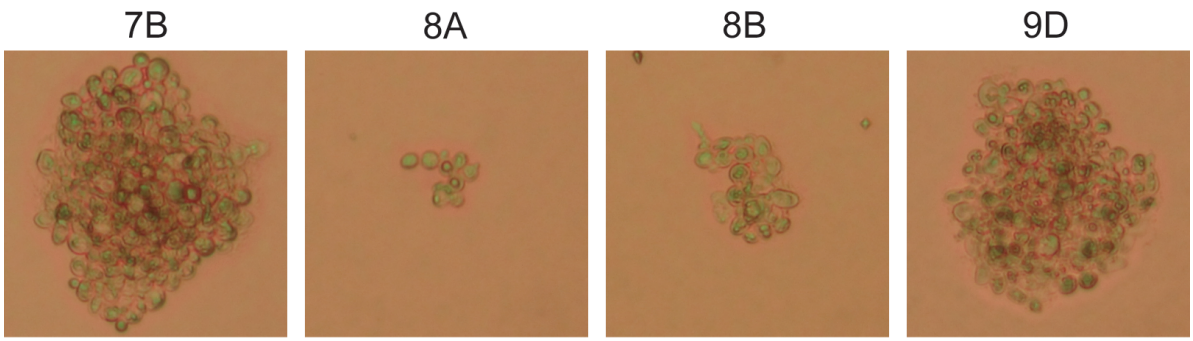
Supplementary Figure 2. Alkali-sensitive Okazaki fragment-sized DNA is not observed in the α-LM *rnh201*Δ or δ-LM *rnh201*Δ strains. Purified genomic DNA was subjected to alkaline hydrolysis, alkaline-agarose electrophoresis and stained with SYBR® Gold Nucleic Acid Gel Stain, a sensitive fluorescent 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.



Supplementary Figure 3. 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, 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 4. Ura3 mutation spectra for *pol1-L868M rnh201Δ* and the *pol3-L612M rnh201*Δ strains. 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.



pol2-M644G rnh201Δ rnh1Δ

Supplementary Figure 5. Microscopic analysis of spore colonies reveals that *pol2-M644G rnh201::hphMX4 rnh1::natMX4* spores germinate but do not proliferate beyond a few cell divisions. Spore colonies were obtained by tetrad analysis of an *RNH1/rnh1::natMX4* diploid. All viable spores are Nat$^-$. Plates were photographed after 5 days growth at 30°C.

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 top1*Δ	YJW289	*pol1-L868M top1*::*natMX4 agp1*::*URA3-OR2*	This study

*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 4). 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 Online Methods

Materials and Reagents. DNA modification and restriction enzymes were from New England Biolabs (Ipswich, MA), oligonucleotides were from Integrated DNA Technologies (Coralville, IA) and dNTPs and rNTPs were from Amersham Biosciences (Piscataway, NJ).

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 [35] and the three-subunit form of wild type and L612M Pol δ [47] 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 physiologically-relevant nucleotide concentrations: dATP, 16 μM; dCTP, 14 μM; dGTP, 12 μM; dTTP, 30 μM; rATP, 3000 μM; rCTP, 500 μM; rGTP, 700 μM; and rUTP, 1700 μM [1].

Yeast strains. *Saccharomyces 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*-Δ*GG2899-2900*) [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* [48] 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 nucleotides of sequence homologous to intergenic regions upstream and downstream of the *RNH201* ORF. Transformants that arose following homologous recombination replacing *RNH201* with HYG-R were verified by PCR analysis. *top1*Δ strain construction was performed as described [13].

Construction of diploid strains heterogeneous for *rnh1::natMX4* was performed by deletion-replacement of 1 copy of *RNH1* via transformation with a PCR product containing the nourseothricin-resistance cassette (natMX4) amplified from pAG25 and flanked by 60 nucleotides 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 using 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) using the Epicentre Yeast DNA purification kit (MPY80200). Five 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 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, 1 M NaCl) overnight. Southern analysis was performed using single-strand radiolabeled probes prepared from a PCR-amplified fragment of the *URA3* reporter gene integrated at the *AGP1* locus on chromosome III using a previously described procedure [13]. Quantitation of alkali-sensitive fragments was performed using ImageQuant software (GE Healthcare). The fraction was calculated by dividing the radioactive intensity (arbitrary units) at 0.1 mm intervals by the total intensity for each lane. Mean fragment size determination was performed using data from either two (Figure 2) or three (Figure 3) independent experiments as described in [13]. Mean DNA fragment sizes (± standard deviation) and P-values were determined as in [13]. SYBR® Gold Nucleic Acid Staining (Life Technologies) was for 2 hours at room temperature using a 1:10,000 dilution of the stock solution in 1X 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 [28].

Data analysis

Calculating 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 then fraction of replication events that generate HydEn-seq ends on the bottom strand ($F_{ij,bottom}$ below) in the lagging-strand-biased data set will approach 1 to the right of each bi-directional origin and will approach 0 to the left. The opposite will be true for the leading-strand-biased data set.

Define $R_{ij,bottom}$ in bin *i* of data set *j* (*lead* or *lag*) as the true ratio of replication events which generate HydEn-seq reads that map to the top strand (forward; *f*) versus the bottom strand (reverse; *r*). If ribonucleotides inserted by replicative polymerases were the only source of HydEn-seq ends, and if ribonucleotide insertion was uniform, then $R_{ij,top}$ would be a simple ratio of forward mapping ($n_{ij,f}$) and reverse mapping ($n_{ij,r}$) end counts (Equation 1).

Equation 1 $$R_{ij,bottom} = \frac{n_{ij,r}}{n_{ij,f}}$$

$R_{ij,bottom}$ would then serve as a measure of strand-biased replication and leading-strand-biased (*lead*) and lagging-strand-biased (*lag*) ratios will be multiplicative inverses of one another (Equation 2).

Equation 2 $$R_{i\ lag,bottom} = R_{i\ lead,top} = {R_{i\ lead,bottom}}^{-1}$$

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 ($O_{ij,top}$) of observed end counts (m_{ij}) is the product of the true ratio of replication events and the multiplicative factor (Equation 3).

Equation 3 $$O_{ij,bottom} = \frac{m_{ij,r}}{m_{ij,f}} = R_{ij,bottom}W_i$$

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 (Equation 4).

Equation 4 $$\frac{O_{i\,lag,bottom}}{O_{i\,lead,bottom}} = \frac{R_{i\,lag,bottom}W}{R_{i\,lead,bottom}W} = {R_{i\,lag,bottom}}^2$$

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 ($F_{ij,bottom}$; Equation 7).

Equation 5 $$R_{i\,lag,bottom} = \sqrt{\frac{O_{i\,lag,bottom}}{O_{i\,lead,bottom}}}$$

Equation 6 $$R_{i\,lead,bottom} = \sqrt{\frac{O_{i\,lead,bottom}}{O_{i\,lag,bottom}}}$$

Equation 7 $$F_{ij,bottom} = \frac{n_{ij,r}}{n_{ij,r}+n_{ij,f}} = \frac{1}{1+\frac{n_{ij,f}}{n_{ij,r}}} = \frac{1}{1+{R_{ij,bottom}}^{-1}}$$

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

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

Phenotype Analyses. Strains were grown in YPDA media (1% yeast extract, 2% bacto-peptone, 250 mg/l adenine, 2% dextrose, 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, data are displayed as the mean D_t ± standard deviation. Spot dilution assays were performed by plating serial (10-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 ºC and photographed after 3 days 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 is from two independent experiments.

Measurement of spontaneous mutation rates and sequence analysis. Spontaneous mutation rates were determined using fluctuation analysis as described previously [46]. 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 multiplying the proportion of each mutation type by the total mutation rate for each strain.

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).