

SUPPORTING ONLINE MATERIAL

Materials and Methods

CAN1 forward mutation assays. All data (with the exception of the WT spectrum from Lippert *et al.* 3) were derived using isogenic derivatives of strain $\Delta[(-2)]$ -7B-YUNI300 (*MATa CAN1 his7-2 leu2- Δ ::kanMX ura3- Δ trp1-289 ade2-1 lys2- Δ GG2899-2900*; ref. 18). The *rnh201 Δ* strain was described previously (1). The *TOP1* ORF was deleted via transformation with a PCR product amplified from pAG25 (19). Cells were grown nonselectively at 30°C in YEP medium (1% yeast extract, 2% bacto-peptone; 2% agar for plates) supplemented with 2% dextrose and 250 μ g/ml adenine. Can-R mutants were selected on synthetic complete medium lacking arginine and supplemented with 60 μ g/ml canavanine. Rates of spontaneous Can-R colonies were measured as described previously (20). To generate spectra, genomic DNA isolated from independent Can-R mutants was PCR-amplified and sequenced. Rates of individual mutation types were calculated by multiplying the proportion of the mutation type in the corresponding spectrum by the total mutation rate.

lys2 reversion assays. All *lys2* alleles were located at the *HIS4* locus of strain YPH45 (*MATa ura3-52 ade2-101_{oc} trp1 Δ 1*; ref. 21) and were transcribed in the same direction as the replication fork initiated at *ARS306*. The *lys2 Δ Bgl* mutation (9) was introduced into the *his4 Δ ::pTET-LYS2* allele (22) by two-step allele replacement. All *lys2 Δ A746NR* alleles were under control of the native *LYS2* promoter, and construction of the *lys2 Δ A746NR,(AT)₂* and *lys2 Δ A746NR,(TC)₃* alleles was described previously (3). The *lys2 Δ A746NR,(AG)₄* allele was similarly constructed using oligos 5'-GATCAAGTGAAGAGAGAGCTTAcGCAA and 5'-GATCTTGcGTAAGCTCTCTCTCACTT; runs are underlined and lowercase letters reflect an AT to CG mutation introduced to remove a stop codon. The *RNH201* and *TOP1* genes were

deleted using PCR-amplified fragments containing dominant drug-resistant or nutritional markers as appropriate. Lys⁺ rates and reversion spectra were obtained as described previously (22). For nonselective growth of *pTET-lys2ΔBgl* strains, medium was supplemented with 2 μg/ml doxycycline hyclate (Sigma) to maintain a low-transcription status.

Top1 cleavage assays. Cleavage assays were carried out as described previously (23). Briefly, custom-synthesized oligonucleotides (Midland Certified, Midland, Texas) were end-labeled prior to annealing with the cold, complementary strand at a 1:1 ratio. Reaction mixtures contained 50-200 nM of labeled DNA construct, 10 mM Tris-HCl (pH=7.5), 50 mM KCl, 1 mM EDTA, 1 mM DTT, 15 ug/mL BSA and 10% DMSO. Recombinant human Top1 (24) and/or CPT were added to 70 nM and 10 μM, respectively, as indicated. Reactions were for 1 hr at 25°C and were stopped by addition of 0.5% SDS (final concentration). In reversal assays, samples were incubated for 1 hr at 25°C before adding 0.5 M NaCl (final concentration). A small volume was withdrawn from the sample and the reaction stopped by adding 0.5% SDS (final concentration) at the specified time points (in min). Samples were analyzed on 20% denaturing polyacrylamide gels and labeled fragments detected using a PhosphorImager.

Table S1. Mutation types and rates in *lys2* frameshift-reversion assays

<i>lys2</i> frameshift allele	Relevant genotype	Rate (x 10 ⁻⁹)				
		Hotspot	1-bp indels	Other	Total (CI)	Total/WT
$\Delta A746NR$	WT*	NA	0.14	0.81	0.95	
	(N=90)				(0.86-1.04)	
$\Delta A746NR,(AG)_4$	WT	0.65	0.47	1.92	3.05	1.0
	(N=84)				(2.58-4.22)	
	<i>rnh201</i> Δ	158	<2.0	9.9	189	62
	(N=95)				(178-200)	
	<i>rnh201</i> Δ <i>top1</i> Δ	0.16	0.59	2.11	2.86	0.9
	(N=88)				(1.68-3.61)	
$\Delta A746NR,(TC)_3$	WT*	1.43	0.19	1.17	2.79	1.0
	(N=74)				(2.46-3.10)	
	<i>rnh201</i> Δ	17.7	<0.19	1.92	19.9	7.1
	(N=92)				(15.7-22.7)	
	<i>rnh201</i> Δ <i>top1</i> Δ	<0.02	0.41	1.12	1.53	0.5
	(N=74)				(1.34-1.89)	
$\Delta A746NR,(AT)_2$	WT*	6.21	0.56	0.65	7.42	1.0
	(N=80)				(6.96-7.57)	
	<i>rnh201</i> Δ	8.15	0.44	2.62	11.2	1.5
	(N=77)				(10.3-11.9)	
	<i>rnh201</i> Δ <i>top1</i> Δ	<0.03	0.73	1.21	1.94	0.3
	(N=74)				(1.38-5.06)	
ΔBgl	WT	0.22	1.66	1.31	3.19	1.0
	(N=73)				(2.63-4.73)	
	<i>rnh201</i> Δ	5.65	0.57	0.57	6.88	2.2
	(N=84)				(4.43-12.0)	
	<i>rnh201</i> Δ <i>top1</i> Δ	<0.05	1.05	3.74	4.80	1.5
	(N=91)				(4.09-6.14)	

Table legend

Rates of individual mutation types were calculated by multiplying the total reversion rate by the proportion of the mutation type in the corresponding spectrum. When no events of a given type were observed, the rate is designated as “<” and was calculated assuming a single event.

NA, not applicable; CI, 95% confidence interval.

*Data are from Lippert *et al.* (3).

Supplemental Figure legends

Figure S1. Mutation spectra of Can-R mutants. Nucleotides are numbered beginning with the ATG start codon. Base substitutions and indels are in red below and above the sequence, respectively; the lengths of the red bars correspond to deletion sizes.

Figure S2. Reversion spectra for *lys2* alleles. In **A-C** hotspot-containing sequences transplanted from the *CAN1* locus are highlighted in yellow and sequences associated with the flanking *Bgl*II site are highlighted in gray. + or Δ correspond to the insertion or deletion, respectively, of a single bp; cins or cdel, insertions or deletions associated with one or more base substitutions; del, deletion; dup, duplication; N, number of independent revertants sequenced.

Figure S3. Human Top1 cleavage assays. **A** shows fragments used as substrates, with relevant dinucleotide repeats highlighted in gray. Transcribed and nontranscribed strands are designated TS and NTS, respectively; arrows indicate positions of Top1 cleavage. In **B**, the indicated strand was labeled on the 3' end.

A *rnh201* (N = 185, rate = 26e-8)

```
1 ATGACAAATT CAAAAGAAGA CGCCGACATA GAGGAGAAGC ATATGTACAA TGAGCCGGTC ACAACCCTCT TTCACGACGT TGAAGCTTCA CAAACACACC
101 ACAGACGTGG GTC AATACCA TTGAAAGATG AGAAAAGTAA AGAATTGTAT CCATTGCGCT CTTTCCCGAC GAGAGTAAAT GCGCAGGATA CGTTCTCTAT
201 GGAGGATGGC ATAGGTGATG AAGATGAAGG AGAAGTACAG AACGCTGAAG TGAAGAGAGA GCTTAAGCAA AGACATATTG GTATGATTGC CTTTGGTGGT
301 ACTATTGGTA CAGGTCTTTT CATTGGTTTA TCCACACCTC TGACCAACGC CGGCCAGTGC GCGCTCTTA TATCATATTT ATTTATGGGT TCTTTGGCAT
401 ATTCTGTAC GCAGTCCCTG GGTGAAATGG CTACATTCAT CCCTGTTACA TCCTCTTTCA CAGTTTCTC ACAAGATTG CTTTCTCCAG CATTGGTGC
501 GGCCAATGGT TACATGTATT GGTTTTCTTG GGCAATCACT TTTGCCCTGG AACTTAGTGT AGTTGGCCAA GTCATTCAAT TTTGGACGTA CAAAGTTCCA
601 CTGGCGGCAT GGATTAGTAT TTTTGGGTA ATTATCACAA TAATGAACCT GTTCCCTGTC AAATATTACG GTGAATTGCA GTTCTGGGTC GCTTCCATCA
701 AAGTTTTAGC CATTATCGGG TTTCTAATAT ACTGTTTTTG TATGGTTTGT GGTGCTGGGG TTACCGGCCC AGTTGGATTG CGTTATTGGA GAAACCCAGG
801 TGCCTGGGGT CCAGGTATAA TATCTAAGGA TAAAAACGAA GGGAGGTTC TTAGTTGGGT TTCCTCTTTG ATTAACGCTG CCTTCACATT TCAAGGTAAT
901 GAACTAGTTG GTATCACTGC TGGTGAAGCT GCAAACCCCA GAAAATCCGT TCCAAGAGCC ATCAA AAAAG TTGTTTCCG TATCTTAACC TTCTACATTG
1001 GCTCTCTATT ATTCATTGGA CTTTGTAGTTC CATAACAATGA CCCTAAACTA ACACAATCTA CTTCTACGT TTCTACTTCT CCCTTTATTA TTGCTATTGA
1101 GAACTCTGGT ACAAAAGTTT TGCCACATAT CTTCAACGCT GTTATCTTAA CAACCAATTAT TTCTGCCGCA AATTCAAATA TTTACGTTGG TTCCCGTATT
1201 TTATTGGTGC TATCAAAGAA CAAGTTGGCT CCTAAATCC TGTC AAGGAC CACCAAAGGT GGTGTTCCAT ACATTGCAGT TTTCTGTACT GCTGCATTTG
1301 GCGCTTTGGC TTACATGGAG ACATCTACTG GTGGTGACAA AGTTTTCGAA TGCGTATTAA ATATCACTGG TGTGCGAGC TTTTGTGCAT GGTATTATTAT
1401 CTC AATCTCG CACATCAGAT TTATGCAAGC TTTGAAATAC CGTGGCATCT CTCGTGACGA GTTACCATTT AAAGCTAAAT TAATGCCCGG CTTGGCTTAT
1501 TATGCGGCCA CATTATGAC GATCATTATC ATTATTCAAG GTTTCACGGC TTTTGCACCA AAATTC AATG GTGTTAGCTT TGCTGCCGCC TATATCTCTA
1601 TTTTCTGTT CTTAGCTGTT TGGATCTTAT TTCAATGCAT ATTCAGATGC AGATTTATTT GGAAGATTGG AGATGTCGAC ATCGATTCCG ATAGAAGAGA
1701 CATTGAGGCA ATTGTATGGG AAGATCATGA ACCAAAGACT TTTTGGGACA AATTTTGGAA TGTTGTAGCA TAG
```

B *rnh201 top1* (N = 165, rate = 18e-8)

```
1 ATGACAAATT CAAAAGAAGA CGCCGACATA GAGGAGAAGC ATATGTACAA TGAGCCGGTC ACAACCCTCT TTCACGACGT TGAAGCTTCA CAAACACACC
101 ACAGACGTGG GTC AATACCA TTGAAAGATG AGAAAAGTAA AGAATTGTAT CCATTGCGCT CTTTCCCGAC GAGAGTAAAT GCGCAGGATA CGTTCTCTAT
201 GGAGGATGGC ATAGGTGATG AAGATGAAGG AGAAGTACAG AACGCTGAAG TGAAGAGAGA GCTTAAGCAA AGACATATTG GTATGATTGC CTTTGGTGGT
301 ACTATTGGTA CAGGTCTTTT CATTGGTTTA TCCACACCTC TGACCAACGC CGGCCAGTGC GCGCTCTTA TATCATATTT ATTTATGGGT TCTTTGGCAT
401 ATTCTGTAC GCAGTCCCTG GGTGAAATGG CTACATTCAT CCCTGTTACA TCCTCTTTCA CAGTTTCTC ACAAGATTG CTTTCTCCAG CATTGGTGGC
501 GGCCAATGGT TACATGTATT GGTTTTCTTG GGCAATCACT TTTGCCCTGG AACTTAGTGT AGTTGGCCAA GTCATTCAAT TTTGGACGTA CAAAGTTCCA
601 CTGGCGGCAT GGATTAGTAT TTTTGGGTA ATTATCACAA TAATGAACCT GTTCCCTGTC AAATATTACG GTGAATTGCA GTTCTGGGTC GCTTCCATCA
701 AAGTTTTAGC CATTATCGGG TTTCTAATAT ACTGTTTTTG TATGGTTTGT GGTGCTGGGG TTACCGGCCC AGTTGGATTG CGTTATTGGA GAAACCCAGG
801 TGCCTGGGGT CCAGGTATAA TATCTAAGGA TAAAAACGAA GGGAGGTTC TTAGTTGGGT TTCCTCTTTG ATTAACGCTG CCTTCACATT TCAAGGTAAT
901 GAACTAGTTG GTATCACTGC TGGTGAAGCT GCAAACCCCA GAAAATCCGT TCCAAGAGCC ATCAA AAAAG TTGTTTCCG TATCTTAACC TTCTACATTG
1001 GCTCTCTATT ATTCATTGGA CTTTGTAGTTC CATAACAATGA CCCTAAACTA ACACAATCTA CTTCTACGT TTCTACTTCT CCCTTTATTA TTGCTATTGA
1101 GAACTCTGGT ACAAAAGTTT TGCCACATAT CTTCAACGCT GTTATCTTAA CAACCAATTAT TTCTGCCGCA AATTCAAATA TTTACGTTGG TTCCCGTATT
1201 TTATTGGTGC TATCAAAGAA CAAGTTGGCT CCTAAATCC TGTC AAGGAC CACCAAAGGT GGTGTTCCAT ACATTGCAGT TTTCTGTACT GCTGCATTTG
1301 GCGCTTTGGC TTACATGGAG ACATCTACTG GTGGTGACAA AGTTTTCGAA TGCGTATTAA ATATCACTGG TGTGCGAGC TTTTGTGCAT GGTATTATTAT
1401 CTC AATCTCG CACATCAGAT TTATGCAAGC TTTGAAATAC CGTGGCATCT CTCGTGACGA GTTACCATTT AAAGCTAAAT TAATGCCCGG CTTGGCTTAT
1501 TATGCGGCCA CATTATGAC GATCATTATC ATTATTCAAG GTTTCACGGC TTTTGCACCA AAATTC AATG GTGTTAGCTT TGCTGCCGCC TATATCTCTA
1601 TTTTCTGTT CTTAGCTGTT TGGATCTTAT TTCAATGCAT ATTCAGATGC AGATTTATTT GGAAGATTGG AGATGTCGAC ATCGATTCCG ATAGAAGAGA
1701 CATTGAGGCA ATTGTATGGG AAGATCATGA ACCAAAGACT TTTTGGGACA AATTTTGGAA TGTTGTAGCA TAG
```


A

(AG)₄	NTS 5' -AACGCTGAAGTGA AGAGAGAG CTTAAGCAA
	TS TTGCGACTTCACT TCTCTCTC GAATTCGTT-5'
(AT)₂	NTS 5' -ACAAAGGTTTTGCCAC ATA T CTTCAACGCT
	TS TGTTTCCAAAACGGTG TA TA GAAGTTGCGA-5'
(TC)₃	NTS 5' -TTTGAATACCGTGGCA TCTCTC GTGACGAGT
	TS AAACTTTATGGCACCGT AGAGAG CACTGCTCA-5'

B

