## Supplemental Figures

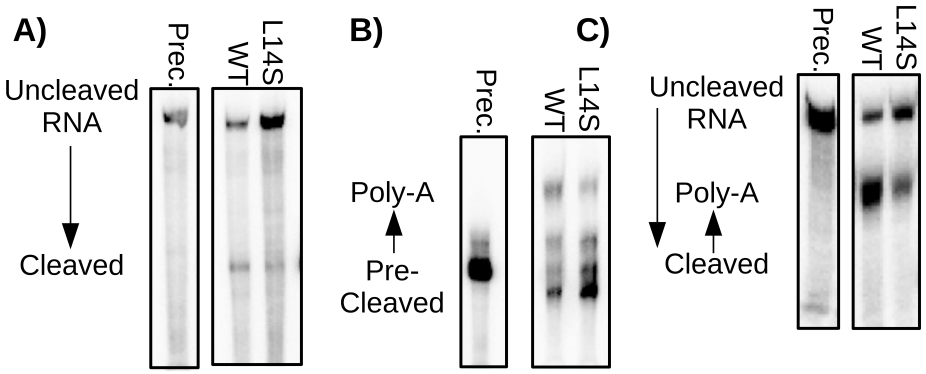


**Figure S1. The *ysh1-L14S* mutant is defective for in-vitro mRNA 3' end-processing. Related to Figure 3.** *In vitro* assay in which a precursor RNA is combined with cell extracts derived from WT or *ysh1-L14S* mutant yeast, which were grown at 30°C. **(A)** The precursor RNA is shortened by the cleavage reaction without poly(A) addition if the reaction is conducted in the presence of 3' dATP. For *ysh1-L14S* mutants, less of the precursor RNA is converted to the cleaved form. **(B)** Pre-cleaved precursor RNA which ends at the *GAL7* poly(A) site is lengthened by the addition of the poly-A tail. For *ysh1-L14S* mutants, less of the precursor RNA is converted to the polyadenylated form. **(C)** Coupled reaction in the presence of ATP, in which an uncleaved precursor RNA is cleaved and subsequently polyadenylated. For *ysh1-L14S* mutants, less of the precursor RNA is converted to the polyadenylated form.

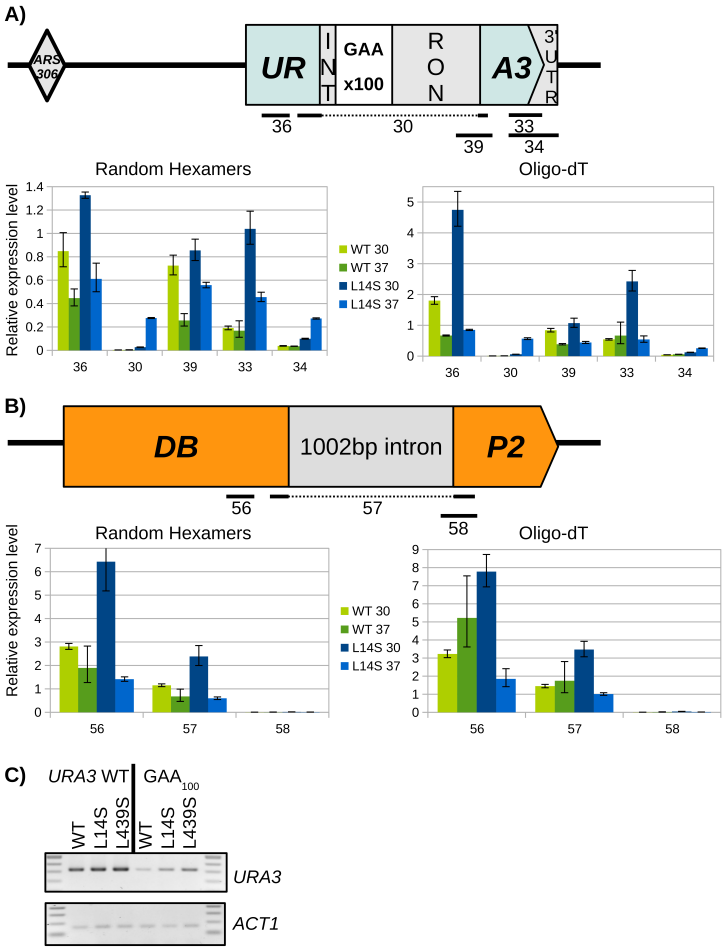


**Figure S2. Further analysis of splicing efficiency. Related to Figure 4. (A)** Diagram of the *UR*-GAA-A3 cassette, indicating the position of primers used for RNA analysis. qRT-PCR results indicate that the repeat-containing intron is poorly spliced in the WT, but the splicing is considerably improved in the *ysh1-L14S* mutant, especially at 37°C. In contrast, the total amount of *URA3* RNA is not greatly affected by the *ysh1-L14S* mutation. **(B)** Diagram of the *DBP2* gene, containing the longest natural yeast intron, which, at 1002bp, is just slightly longer than the 974bp repeat-containing intron in the *URA3* cassette. qRTPCR results indicate that splicing of this intron is very efficient, with almost no unspliced product recovered. The *ysh1-L14S* mutation does not appear to affect splicing efficiency, indicating that it may have a repeat-specific effect on splicing in the *URA3* cassette. qRTPCR results shown for cDNA made from either random hexamers or oligo-dT primers, as indicated. Error bars represent the SD of qPCR technical replicates. **(C)** RTPCR analysis of polysomal RNA, comparing spliced *URA3* mRNA (primer pair #3), and loading control *ACT1* spliced mRNA. Results are shown for both *ysh1-L14S* and *ysh1-L439S* mutants, which show an increase in mRNA relative to the WT for the UR-GAA-A3 cassette (right panel). In contrast, the *URA3* gene lacking an intron is unaffected (left panel). This supports the results in (**A**), suggesting that the increase in splicing efficiency leads to RNA that is stably exported from the nucleus.

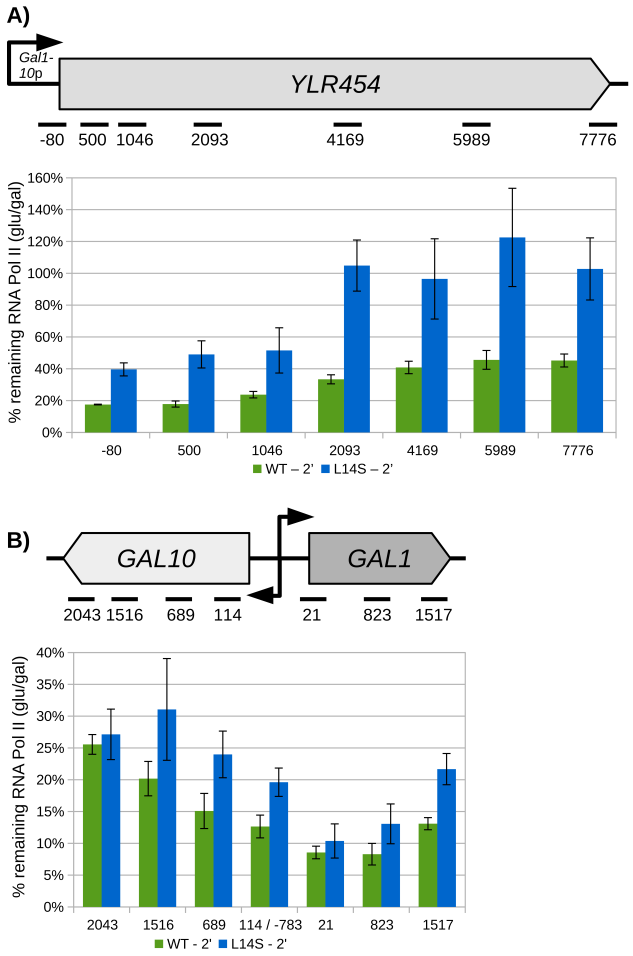


**Fig. S3. Slow transcription elongation at genes not containing GAA repeats. Related to Figure 6. (A)** The 8 kb-long gene *YLR454w* was placed under control of the *GAL1-10* promoter. This modified gene was used to measure transcription elongation speed via RNA polymerase clearance assays. The *ysh1-L14S* mutant displays markedly slower elongation speed, with ~2.5-fold greater occupancy of RNA Pol II observed past 2 kb in the two minutes following glucose repression. In contrast, the WT strain shows that most RNA Pol II has proceeded past the end of the 8 kb gene. **(B)** RNA polymerase clearance assays measuring transcription elongation through the endogenous *GAL1* and *GAL10* genes, performed concurrently with the experiment in Fig. 6A. Slower transcription elongation is evident in the *ysh1-L14S* strains. Error bars represent standard error of two trials. Primer pairs are numbered by the midpoint of the PCR product, with respect to the ORF start position.

**Fig. S4. Location of mutants within Ysh1 protein structure. Related to Figure 3.** A portion of the protein structure of the human homolog of Ysh1, CPSF-73, is shown (Mandel et al., 2006). The *ysh1-L14S* mutant in (magenta spheres) is located near the surface and partially exposed, while the *ysh1-L439S* (yellow spheres) is buried within the protein.

**Table S1. UV Dosage Determination. Related to Experimental Methods – Screening Approach.** Determining a suitable UV dosage by measuring the rate of mutation of the *CAN1* gene under various UV doses. $\mu_g$ represents mutation rates determined as described in (Drake, 1991), using three independent trials. The relative number of mutations is scaled to the 0 UV dose.

| **Dose (x100 µJ):** | **µg:** | **Relative # Mutations:** |
|---|---|---|
| 0 | 0.00192 | 1 |
| 25 | 0.00193 | 1.01 |
| 50 | 0.00218 | 1.14 |
| 75 | 0.00818 | 4.27 |
| 100 | 0.0632 | 33.0 |

**Table S2. Yeast strains in this study.**

| Strain name: | Parent strain: | Cassettes: | Additional genotype info: |
|---|---|---|---|
| SMY706 | CH1585 | none | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| SMY720 | SMY706 | none | ade2::URA3, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| SMY724 | SMY720 | AD-GAA63-E2 | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| SMY732 | SMY724 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG28 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | ysh1 L439S, ysh1 S59F, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa, various UV-induced mutations |
| RMG35 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa, various UV-induced mutations |
| RMG87 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | ysh1 L439S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG89 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG92 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG108 | G4G1C1T150 (Shah et al., 2014) | UAS-GAA100-Galp-CAN1 | ysh1 L14S, CAN1::kanMX4, BAR1::URA3, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG110 | G4G1C1T150 (Shah et al., 2014) | UAS-GAA100-Galp-CAN1 | CAN1::kanMX4, BAR1::URA3, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG385 | SMY706 | none | leu2Δ1, trp1Δ63, ura3::HPH, his3Δ200, MIP1, HAP1, MATa |
| RMG389 | RMG385 | UR-GAA100-A3 at Chr. V non-essential arm | leu2Δ1, trp1Δ63, ura3::HPH, his3Δ200, MIP1, HAP1, MATa |
| RMG391 | RMG389 | UR-GAA100-A3 at Chr. V non-essential arm | ysh1 L14S, leu2Δ1, trp1Δ63, ura3::HPH, his3Δ200, MIP1, HAP1, MATa |
| RMG338 | SMY724 | AD-GAA63-E2 | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| FPY01 | SMY724 | GAL1-10p-YLR454, AD-GAA63-E2 | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| FPY02 | RMG338 | GAL1-10p-YLR454, AD-GAA63-E2 | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG431 | SMY732 | Gal1-10p-UR-GAA100-A3 at ARS306 locus | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG434 | RMG89 | Gal1-10p-UR-GAA100-A3 at ARS306 locus | ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG407 | SMY732 | UR-GAA100-A3 at ARS306 locus | leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG409 | RMG89 | UR-GAA100-A3 at ARS306 locus | ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG167 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RAD52::HPH, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG169 | RMG89 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RAD52::HPH, ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG232 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RAD51::HPH, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG236 | RMG89 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RAD51::HPH, ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG114 | RMG92 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RNH201::LEU2, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |

| | | | |
|---|---|---|---|
| RMG112 | RMG89 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RNH201::LEU2, ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG118 | RMG114 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RNH1::HPH, RNH201::LEU2, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG120 | RMG112 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | RNH1::HPH, RNH201::LEU2, ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG242 | SMY732 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | pRNH1::Nat-pMet25, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |
| RMG247 | RMG89 | AD-GAA63-E2, UR-GAA100-A3 at ARS306 locus | pRNH1::Nat-pMet25, ysh1 L14S, leu2Δ1, trp1Δ63, ura3Δ52, his3Δ200, MIP1, HAP1, MATa |

**Table S3. PCR products in this study**

| Set #: | Purpose: | Template: | Primer sequence (FWD, REV): |
|---|---|---|---|
| 1 | Amplifies GAA repeats in all cassettes | colony PCR | CTCGATGTGCAGAACCTGAAGCTTGATCT |
| | | | GCTCGAGTGCAGACCTCAAATTCGATGA |
| 2 | For replacement of ADE2 gene with URA3 gene | pRS306 (Sikorski and Hieter, 1989) | GAGAAGTGACGCAAGCATCAATGGTATAATGTCCAGAGTTGT GAGGCCTTGACACCGCAGGGTAATAACTGAT |
| | | | CTTTTCCCGGTTGTGGTATATTTGGTGTGGAAATGTTCTATTT AGAAACACGGTCACAGCTTGTCTGTAA |
| 3 | Amplifies repeats in AD-GAA-E2 cassette for sequencing | colony PCR | CCCGGTTGTGGTATATTTGGTG |
| | | | GCATAATGGCGTTCGTTGTAATGG |
| 4 | Amplifies YSH1 flanking sequence for plasmid construction | SMY732 gDNA | TATCGTCTCACGCGCTATTCCTACAGTAACTAACGCAGACAT CA |
| | | | CCAGATTTTGGTTTGGTATTACTTCTATAAAGTAGTCTA |
| 5 | Amplifies YSH1 gene for plasmid construction | RMG28 or RMG35 gDNA | TAGAACTGCCTTTTATTGTTTACCTTAATCACT |
| | | | ATAGCGCGCCTGTATATCGTCATTTAGGGTT |
| 6 | Confirmation of HIS3–YSH1 integration in yeast | colony PCR | GGAGCGAACAAATACAACAAC |
| | | | ATAGCGCGCCTGTATATCGTCATTTAGGGTT |
| 7 | Sequence confirmation of ysh1 mutations | colony PCR | CTAGCGGATCTTTAACAACT |
| | | | GTCTTTGTGTTCCTCTTCAT |
| 8 | Amplify HIS3–YSH1 from yeast for propagation into new strains | RMG89 gDNA | GGAGCGAACAAATACAACAAC |
| | | | ATAGCGCGCCTGTATATCGTCATTTAGGGTT |
| 9 | Replace ura3Δ52 with hphMX4 | pCORE-UH (Storici and Resnick, 2003) | TTTTGACCATCAAAGAAGGTTAATGTGGCTGTGGTTTCAGGG TCCATAAAGTTTAGCTTGCCTCGTCCC |
| | | | TATTGGATAGTTCCTTTTTATAAAGGCCATGAAGCTTTTTCTT TCCAATTCTTATCCTTCACCATAAATATGCCTCGCAAAAAAG GTAATAATACGCAAACCGCCTCTC |
| 10 | Confirmation of ura3Δ52 absence | colony PCR | CATTATCCGCCAAGTACAAT |
| | | | GTCTCCCTTGTCATCTAAAC |
| 11 | Amplifies UR-GAA-A3 cassette along with TRP1 marker, with flanking homology to the non-essential arm of Chr. V | SMY732 gDNA | ACAGAAACCCTAGATTTCTATAGGGCAAATTTCAGGGTTATA CTAAACACTCGCAGACATTGTTATTTCC |
| | | | AATTAAACCTATTTCTTTATCATCATATTTACTTATATCTTTAA CAGATTCCAGGCGATAAGCTAGCAGGCAA |
| 12 | Located outside integration site for UR-GAA-A3 cassette on Chr. V non-essential arm | colony PCR | GATTCCCTGATTCGGTTTACTCT |
| | | | CTCTTGTCCCTTATTAGCCTTGA |
| 13 | Replacement of URA3 promoter with Gal1-10 promoter | pCORE-Hp53 (Storici and Resnick, 2003) | TAATCAAGTAACACTCGCAGACATTGTTATTTCCGCGGATCC GGAGATCTGTTTAGCTTGCCTCGTCCC |
| | | | GCAGCACGTTCCTTATATGTAGCTTTCGACATGATTTATCTTC GTTTCCTTCTTAAAGTTAAACAAAATTATTTCTAGTCGA |
| 14 | Confirmation of GAL1-10 promoter integration | colony PCR | ATGACCCACTCAGGTGTTAAA |
| | | | TAATGCCTTTAGCGGCTTAACT |
| 15 | ADE2 WT restoration | Strain #1 gDNA | TATTAGTGAGAAGCCGAGAA |
| | | | TTGAGCCGCCTTATATGAA |
| 16 | RAD52 KO integration | pAG32 (Goldstein and McCusker, 1999) | CGAATGGCGTTTTTAAGCTATTTTGCCACTGAGAATCAACAA ATGCAAACAAGGAGGTTGCCAGATCTGTTTAGCTTGCCT |
| | | | GGTTTCACGCGGTACTTGATTCCCAGCCCCTTCTAGCATATGA GGCCCCAGTTCTTTATCATCGATGAATTCGAGCTCGTT |
| 17 | RAD52 KO external confirmation | colony PCR | CTAGAGGATTTTGGAGTAATAAATAATGATG |
| | | | ACGTCGCTAAAGATGGTATGGTA |
| 18 | RAD52 KO internal confirmation | colony PCR | GCCAAGAAATCTGCCGTTAC |
| | | | TGAGCTTTCGCTGATTTCATCC |

| | | | |
|---|---|---|---|
| 19 | RAD51 KO integration | pAG32 (Goldstein and McCusker, 1999) | AAATGTTGGAAATGCACCACTACCGTTCTTCAACCAATCTAG TTTAGCTATTTAGAACGCGGCTACAATTA |
| | | | AAAGAGGAGAATTGAAAGTAAACCTGTGTAAATAAATAGAG ACAAGAGACCAAATACCTACCCTGATTCTGTGGATAACC |
| 20 | RAD51 KO external confirmation | colony PCR | CAATTCGCAAGAAACGCACT |
| | | | AAGTAGTCATCGGGAAGAAGAGTA |
| 21 | RAD51 KO internal confirmation | colony PCR | AGATCGGAGCTGATTTGTTTGAC |
| | | | CTTCACCGCCACCAATATCC |
| 22 | RNH201 KO integration | pRS305 (Sikorski and Hieter, 1989) | ATGGTACCCCCCACGGTAGAAGCATCCTTGGAGTCTCCTTAC ACTAAGTCGTAGGTGTCGGGGCTGGCTTAA |
| | | | AGGTGATCACCGGTACCAATTATCTAGGGTTCTCAGCCTCTTC CTTCTAACTCCTTACGCATCTGTGCGGTA |
| 23 | RNH201 KO confirmation | colony PCR | ATTTTGCGACGCCTGCCAAT |
| | | | GAAACGGCAAAGCATAGTAGCAGAT |
| 24 | RNH1 KO integration | pAG32 (Goldstein and McCusker, 1999) | AATTATGGCAAGGCAAGGGAACTTCTACGCGGTTAGAAAGG GCAGGGAAACAGCTGAAGCTTCGTACGC |
| | | | CCTGGATCACCATCGTGTCCTTTTACCCATTCAATCTGAAATT TACCATTCATAGGCCACTAGTGGATCTG |
| 25 | RNH1 KO confirmation | colony PCR | GACAGCAGCATCAACAATGAA |
| | | | CACGCTTATAGATAGTTATCGGGTA |
| 26 | RNH1 promoter replacement with pMet | pYM-N35 (Janke et al., 2004) | TCACTCCTTGCTTATCGAAGGAACTATCGATTCCTAATTACGT ACGCTGCAGGTCGAC |
| | | | CCCTGCCCTTTCTAACCGCGTAGAAGTTCCCTTGCCTTGCCAT CGATGAATTCTCTGTCG |
| 27 | pMet-RNH1 confirmation | colony PCR | CAGGGTCGTCAGATACATAGATAC |
| | | | ATAACCTGCCCTTGAAGATGAC |

**Table S4. qPCR primer sets**

| Set #: | Target: | Primer sequence (FWD, REV): |
|---|---|---|
| 28 | SCR1 gene – RNA normalization | GTGGGATGGGATACGTTGAG |
| | | TTTACGACGGAGGAAAGACG |
| 29 | ADE2 cassette (#2) spliced mRNA | TGTAGAGACTATCCACAAGGACA |
| | | TCCAGAGTTGTGAGGCCTT |
| 30 | URA3 cassette (#1) spliced mRNA | CTGCTAACATCAAAAGGCCTCTA |
| | | AGTTGAAGCATTAGGTCCCAAA |
| 31 | ADE2 cassette (#2) – 3' end pre-poly-A site | ATTGTGCAAATGCCTAGAGGT |
| | | TAGATAAGCTTCGTAACCGACAGT |
| 32 | ADE2 cassette (#2) – 3' end post-poly-A site | ATTGTGCAAATGCCTAGAGGT |
| | | ACATTTGATGTAATCATAACAAAGCCT |
| 33 | URA3 cassette (#1) – 3' end pre-poly-A site | ATTTGAGAAGATGCGGCCA |
| | | GTCGACGGGTAATAACTGATATAATTAAA |
| 34 | URA3 cassette (#1) – 3' end post-poly-A site | ATTTGAGAAGATGCGGCCA |
| | | GCCTCGTGATACGCCTATTT |
| 35 | GAL1-10p-URA3 – promoter-exon junction | TACCTCTATACTTTAACGTCAAGGAGAA |
| | | ATCTTCGTTTCCTTCTTAAAGTTAAACA |
| 36 | URA3 cassette – within exon 1 | AGTTGAAGCATTAGGTCCCAAA |
| | | TAATGCCTTTAGCGGCTTAACT |
| 37 | URA3 cassette – exon-intron junction 1 | CCCAGGTATTGTTAGCGGTT |
| | | ATCGAATTTGAGGTCTGCACT |
| 38 | URA3 cassette – within intron | AGATCAAGCTTCAGGTTCT |
| | | GAAGGCTCTCAAGGGCATC |
| 39 | URA3 cassette – intron-exon junction 2 | TTGACTGATCTGTAATAACCACGA |
| | | GAGCCCTTGCATGACAATTC |
| 40 | ChrV intergenic region | CAAGTTTGGTGAGATAGTTTACGC |
| | | CCCTGCATTGAGATCGCATC |
| 41 | YLR454 promoter (-80bp) | CTGGGGTAATTAATCAGCGAAGCGATG |
| | | CACTTGTACAGTAGAACATTAATCGGAAAC |
| 42 | YLR454 (500bp) | CGCAATTAGTCAACAACGATATCACGATTG |
| | | CCTACTTGAAGTCCATCCTTCAGAGG |
| 43 | YLR454 –(1046bp) | CAATACCAACAGGTTCAGAAATGAGATGC |
| | | GAGAGAACAAATTGGTTTCGCCAAATATCG |
| 44 | YLR454 (2093bp) | CATATCATCCACCCTAGGTGCTAGGTCGG |
| | | GAGCTGACCAGACCTAACCATAGTAGCGTG |
| 45 | YLR454 (4169bp) | AGATATTACTCGTTGTTCGTGCCCAG |
| | | AGATATTACTCGTTGTTCGTGCCCAG |
| 46 | YLR454 (5989bp) | CGTACTGTTGAAATGGAACGAGGACGC |
| | | ATCGCTTCCATACTCGTTGTATCATCAGTC |
| 47 | YLR454 (7776bp) | GAGGGTCACAGATCTATTACTTGCCC |
| | | GTTGTGAGTTGCTTCAGTGGTGAAGTG |
| 48 | GAL10 (2043bp) | AATCTGTAGACAATCTTGGACCC |
| | | GCCTGGTAGATACATTGATGCTAT |
| 49 | GAL10 (1516bp) | TTCACCAGCAGTCAATTTACCT |
| | | GCCGAGTACATGCTGATAGATAA |
| 50 | GAL10 (689bp) | ACCATCTCTGGAATCATAATCGTC |
| | | ATATGGCTCAAGTAGCTGTTGGTA |
| 51 | GAL10 (114bp) | TCAAGACCTCTAACCTGGCTA |
| | | TGGTGCTGGATACATTGGTTCA |
| 52 | GAL1 (21bp) | TACCTCTATACTTTAACGTCAAGGAGAA |
| | | TCGGCCAATGGTCTTGGTAA |
| 53 | GAL1 (823bp) | TTATTGCGAACACCCTTGTTGTA |
| | | CACCGTACGTGGCAGCTAA |
| 54 | GAL1 (1517bp) | AAAGAAGCCCTTGCCAATGA |
| | | CTGCCCAATGCTGGTTTAGA |

| | | |
|---|---|---|
| 55 | RNH1 gene | GGTTAGAAAGGGCAGGGAAA |
| | | TGTCAAGGAGGAATAATGAAGTGG |
| 56 | DBP2 5' exon | TACTTAGAAACAGCCTCTCAAGAC |
| | | TCTTTGTCACCATGAATAGCCA |
| 57 | DBP2 spliced | CGCCAGAGGTATCGATGTCAA |
| | | AGATATAGCAGTACCAGTAGCACC |
| 58 | DBP2 unspliced | ATTACTAACAGTATGCTTTGTAAACGT |
| | | AGATATAGCAGTACCAGTAGCACC |

**Supplemental Experimental Methods**

**Strain construction**

The list of our strains is presented in Table S1.

The strain used for the screen (SMY732) was constructed in several steps. First, the *ADE2* gene in the starting strain SMY706 was replaced with the *URA3* gene using PCR product #2 by selecting for $URA^+$ red colonies. The resulting strain SMY720 was then transformed with the AD-GAA-E2 cassette (Fig. 1A) by selecting for $5FOA^r$ white or light-pink colonies. An isolate of this transformation with 63 uninterrupted GAA repeats, verified by Sanger sequencing of PCR product #3, became strain SMY724. This strain was then transformed with cassette UR-GAA-A3, as previously described (Shishkin et al., 2009). The resulting strain SMY732 was used for the mutagenesis screen.

To create *YSH1* point mutant strains, first *YSH1* flanking sequence was amplified from genomic DNA using PCR product #4 (Table S3) and cloned into pRS303 (Sikorski and Hieter, 1989) at the *Bst*Z17I and *Bss*HII sites. Second, the mutant alleles of *YSH1* were created by PCR-amplification of genomic DNA from mutagenized strains RMG28 and RMG35 (PCR product #5, Table S3). They were then cloned into the *Bsm*BI and *Afe*I sites of the aforementioned pRS303 derivative. The resultant plasmids were cut with *Bss*HII. The fragment containing the mutant *YSH1* allele together with the *HIS3* marker was isolated and transformed into strain SMY732 selecting for HIS+ clones. The correct integration of the mutant *YSH1* alleles was confirmed by the presence of PCR product #6 and by Sanger sequencing of PCR product #7, resulting in strains RMG87 (*ysh1-L439S*), RMG89 (*ysh1-L14S*) and RMG92 (*YSH1*).

The *ysh1-L14S* allele from the strain RMG89 was subsequently used as a template for PCR product #8 in order to propagate the *ysh1-L14S* mutation, along with the *HIS3* marker, to further strains. The strain G4G1C1T150 (Shah et al., 2014), which contains the UAS-GAA-CAN1 cassette was transformed in this manner to create the strains RMG108 (*ysh1-L14S*) and RMG110 (*YSH1*). The strain SMY724 was modified in the same manner to create RMG338 (*ysh1-L14S*).

The wild-type strain for the arm loss assay was created by transforming SMY706 with PCR product #9, which directly replaced the *ura3Δ52* allele with the *HPH* selectable marker, in order to discourage intra-chromosomal repair events with the UR-GAA-A3 cassette. The resulting strain, RMG385, was confirmed by the absence of PCR product #10. To create RMG389, RMG385 was then transformed with PCR product #11, consisting of the UR-GAA-A3 cassette together with the *TRP1* selectable marker, along with primer tails directing integration to the non-essential arm of chromosome V, just centromeric to the endogenous *CAN1* gene. Correct integration of the cassette was verified using PCR products #1 and #12. The *ysh1-L14S* mutant strain for the arm-loss assay, RMG391, was then created by transforming RMG389 with the PCR product #8 from RMG89, containing the *ysh1-L14S* allele together with *HIS3* selectable marker.

To replace the *URA3* promoter in the UR-GAA-A3 cassette with the Gal1-10 promoter, strains SMY732 and RMG89 were transformed with the PCR product #13, containing the GAL1-10 promoter together with the *HPH* selectable marker. Correct integration of the GAL1-10 promoter was confirmed by Sanger sequencing of PCR product #14. The resultant strains RMG431 and RMG434 contained the GAL-UR-GAA-A3 hybrid in the *YSH1* or *ysh-L14S* genetic background, respectively.

To create strains with the single UR-GAA-A3 cassette, the AD-GAA-E2 cassette in the SMY732 and RMG89 strains was replaced with the wild-type *ADE2* gene, which was PCR amplified from SMY706 (PCR product #15), resulting in strains RMG407 and RMG409, respectively. Restoration of the WT gene was confirmed by the ADE+ phenotype combined with the absence of the 240bp band in PCR product #1.

To assay elongation on the *YLR454* gene, two strains FPY01 and FPY02 were created from the SMY724 strain and RMG338, respectively. by single-step integration of a *TRP1* plasmid containing the *GAL1* promoter fused to the 5'-most 300 bp of the *YLR454w* open-reading frame at the *YLR454w* locus (Mason and Struhl, 2005).

*RAD52* knockouts in strains SMY732 and RMG89 were obtained by direct replacement with the *HPH* gene obtained by PCR amplification from pAG32 (Goldstein and McCusker, 1999) (PCR product #16), resulting in strains RMG167 and RMG169, respectively. The knockouts were confirmed by the presence of the insertion and the absence of the original gene, via PCR primer pairs #17 and #18, respectively. *RAD51* knockouts were obtained in the same manner, using PCR product #19 for transformation and PCR primer pairs #20 and #21 for confirmation, resulting in strains RMG232 and RMG236.

RNH1/201 knockouts were made in two steps. First, RNH201 knockouts in strains RMG92 (wild type) and RMG89 (*ysh1*-L14S) were obtained by direct replacement with the LEU2 gene obtained by PCR amplification from pRS305 (Sikorski and Hieter, 1989) (PCR product #22), resulting in strains RMG114 and RMG112, respectively. RNH201 replacement was confirmed via PCR using primer pair #23. Next, RNH1 was knocked out in strains RMG114 and RMG112 by direct replacement with the HPH gene obtained by PCR amplification from pAG32 (Goldstein and McCusker, 1999) (PCR product #24), resulting in strains RMG118 and RMG120, respectively. RNH1 replacement was confirmed via PCR using primer pair #25.

Strains for RNH1 overexpression were generated by replacing the RNH1 promoter with the inducible MET25 promoter. pMET was amplified from pYM-N35 (Janke et al., 2004), along with the nourseothricin-resistance marker gene, using primer pair #26 containing homology to the pRNH1 region. This PCR product was used to transform strains SMY732

and RMG89, resulting in strains RMG242 and RMG247, respectively. Promoter replacement was confirmed via PCR primer pair #27.

**Screening approach**

First, a suitable UV dosage was chosen by measuring the rate of mutation of the *CAN1* gene under various UV doses. Strain SMY732 was resuspended in water and subjected to 0J, 0.025J, 0.05J, 0.075J and 0.1J UV radiation using a UV Stratalinker 1800 (Stratagene). Appropriate dilutions were grown on YPD and canavanine-containing media for 3 days at 30C. The resulting colonies were counted and used to estimate mutation rate as described in (Drake, 1991) (Table S1).

For the screen, strain SMY732 was spread on YPD plates and subjected to 0.1J UV radiation followed by growth for 3 days at 30°C. Dark red colonies were isolated and spread on synthetic complete media lacking uracil and containing 0.1% 5-FOA. 5-FOA-resistant colonies were checked for large-scale repeat expansions by PCR using primer set #1, as previously described (Shishkin et al., 2009). Genomic DNA was isolated from strains with high frequencies of repeat expansions, and then used for whole genome sequencing (Illumina Genome Analyzer II, 100bp paired-end reads, barcoded libraries), resulting in an average of ~80X coverage per strain. Reads were aligned to the S288C reference genome using Bowtie (Langmead et al., 2009), and variants were called using Samtools (Li et al., 2009). Variants were further analyzed for potential deleterious effects using snpEFF (Cingolani et al., 2012) and PolyPhen2 (Adzhubei et al., 2013). Strains RMG28 and RMG35, containing *ysh1-L439S* and *ysh1-L14S* variants, respectively, were subjected to further analysis.

**Fluctuation Assays**

Fluctuation assays were performed as previously described (Shah et al., 2014; Shishkin et al., 2009). Briefly, frozen stocks are spread for single colonies on YPD media supplemented with uracil, to ensure that cells with an inactivated *URA3* cassette will be able to grow. At least eight colonies per strain isolate are picked after ~40 hours and diluted in 200ul $dH_2O$. Five steps of 10-fold serial dilutions are performed. For the wild-type strain, YPD plates receive 50ul of suspended cells from the $4^h$ dilution and 5FOA plates receive 50ul of undiluted cells. Volumes and concentrations may be adjusted for the various mutant strains in order to plate an appropriate density of cells. Plates are grown at $30^oC$ for three days, and colonies are counted. For 5FOA plates, colonies are examined for repeat length via PCR using primer set #1. Numbers of colonies on 5FOA are thus adjusted to reflect only expansions, and the expansion rates are calculated (Drake, 1991). Slow growing strains were given extra time to grow at all steps, as compared with the wild type strain. For the arm loss assays, selection for canavanine-resistance preceded selection for 5-FOA resistance. Absence of PCR product #1 was used to distinguish chromosomal arm loss events from *can1ura3* double point mutants.

***RNH1* Overexpression**

Strains overexpressing *RNH1* were obtained via replacement of the endogenous *RNH1* promoter with the inducible *MET25* promoter. Fluctuation assays were carried out as above, with the following modification: strains RMG242 (wild type) and RMG247 (*ysh1*-L14S) were pre-grown on solid media lacking the amino acids methionine and cysteine and supplemented with uracil (high expression), as well as on synthetic complete media supplemented with uracil (low expression). Concurrently, *RNH1* expression was measured under these conditions by extracting RNA and gDNA from cells on the same solid media plates and performing qPCR on each (see "Quantitative RNA Analysis" below for further details) using PCR primer pair #55. *RNH1* expression analysis was performed as described below, using primer pair #55, and was found to increase ~3-9 fold in media lacking methionine (data not shown), which is in line with published expectations (Mumberg et al., 1994).

**Quantitative RNA Analysis**

RNA levels were determined using the strategy described in the main text. Strains SMY732, RMG89, RMG407 and RMG409 were grown overnight in 2 ml of YPD media at 30°C, split into two tubes each of 1 ml culture and 9 ml YPD, then grown for 4 hours at either 30°C or 37°C. 1.5 ml aliquots of each were used to extract RNA and genomic DNA using the Zymo YeaStar RNA and Genomic DNA kits, respectively. RNA was then DNase treated (Turbo DNA-free kit – Thermo-Fisher), followed by cDNA generation (Superscript IV First Strand Synthesis system – Thermo-Fisher) using either random hexamers, or poly-dT primers. qPCR was then performed using SYBR Select Master Mix (Thermo-Fisher) and QuantStudio 6 Flex RTPCR system (Applied Biosystems) for both cDNA and gDNA samples with specific primer sets as noted. qPCR values from gDNA for PCR product #28 were then used for normalization of all cDNA results, on the basis that each haploid cell contains a single copy of the genome, and that the number of cells was kept consistent between the RNA and gDNA preparations. This normalization method is therefore robust to strains with RNA processing defects, or any other conditions that might alter the levels of various RNA species typically used for normalization.

**RNA Pol II elongation assays**

The RNA Pol II transcription elongation assay was performed as previously described (Mason and Struhl, 2005). Details particular to this experiment are described here: Briefly, strains RMG431 and RMG434 were cultured to OD 0.4 in

raffinose-containing minimal medium at 30°C. Expression from the Gal1-10 promoter was induced with the addition of 2% galactose for 2.5 hours, and cultures were subsequently shifted to 37°C for 1 hour. Cells were formaldehyde-fixed two minutes after the addition of 2% glucose, or directly from the galactose culture. Fixed cells were lysed by cryo-grinding. Cross-linked chromatin was sheared to an average length of 200 bp. RNA Pol II-bound chromatin was then incubated with anti-Rpb1 CTD antibody 4H8 (Santa Cruz) for 5 hours at 4°C. Immunoprecipitated chromatin was then de-crosslinked, purified and analyzed by qPCR as described above. Pol II occupancy was normalized via PCR product #40, representing an intergenic region on chromosome V that should contain minimal bound RNA Pol II. Strains FPY01 and FPY02 were monitored at the pGAL-YLR454 locus using primer sets #41-47, while strains RMG431 and RMG434 were monitored at the Gal1-10p–UR-GAA-A3 cassette using primer sets #33-39, as well as the endogenous Gal1-10 locus, using primer sets #48-54. Each PCR locus was then normalized to itself by dividing the glucose over the galactose values.

**Supplemental References**

Drake, J.W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A *88*, 7160-7164.
Goldstein, A.L., and McCusker, J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast *15*, 1541-1553.
Mumberg, D., Muller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Res *22*, 5767-5768.
Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.
Storici, F., and Resnick, M.A. (2003). Delitto perfetto targeted mutagenesis in yeast with oligonucleotides. Genet Eng (N Y) *25*, 189-207.