File S1. Expanded Materials and Methods

Strain construction

Most of the details of the strain constructions are in Table S1. As in previous studies (for example, Lee *et al.*, 2009), we used derivatives of the sequence-diverged haploids W303-1A and YJM789. These haploids with various genetic alterations were crossed to generate the diploids used to assay genetic instability. Most haploid strains were constructed by transformation with PCR-generated DNA fragments or by sporulating nearly-isogenic diploids. The genotypes of spores for auxotrophic markers were determined by replica-plating spore-derived colonies to omission media. The replacements of genes with drug-resistance markers were confirmed by PCR analysis as described in Tables S2 and S3. Mating type was determined by PCR with primers MATaF, MATalphaF, and MATR (Table S3). *MAT***a** and *MAT*α loci were associated with 500 and 400 bp fragments, respectively.

Measurements of rates of genetic instability induced by loss of RNase H

We used three methods to examine the rates of instability in strains with mutations affecting RNase H activity. We first measured the frequency of genomic alterations in sub-cultured diploid strains of the following genotypes: wild-type, *rnh1*Δ, *rnh201*Δ, *rnh1*Δ *rnh201*Δ, *pol2-M644L*, and *rnh201*Δ *pol2-M644L*. All diploids were generated by crossing haploids isogenic with W303-1A and YJM789. Two independently-derived isolates from each strain were streaked with a toothpick to single colony density on rich growth medium (YPD) for the first subculture. For the second subculture, five-ten colonies from each isolate were then re-streaked to YPD. One colony derived from each of these ten to twenty colonies was then re-streaked again. For the mutant backgrounds, this procedure was repeated twenty times. For the wild-type, only ten sub-culturings

were performed. All sub-culturing experiments were done at 30^0 C. Following sub-culturing, the passaged strains were examined by whole-genome microarrays as described below.

The second method of measuring genome stability was to monitor the frequency of formation of red/white sectored colonies in strains in which the *ADE2* gene was inserted near the right telomere of one copy of chromosome IV and the *K. lactis URA3* was inserted at the allelic position on the other copy (Fig. 3). The homologs with the *ADE2* and *URA3* genes had wild-type and mutant alleles of the centromere-linked *TRP1* gene, respectively. Experiments were initiated using colonies formed on YPD plates. For each genotype, we used one colony from two independently-constructed diploids. Each colony was suspended in water, and diluted to a concentration that resulted in about 1000 cells per plate on the diagnostic medium (SD-arginine with 10 micrograms/ml adenine). After three days of growth, we scored plates for red/white sectors using a dissecting microscope. Cells from each sector were then re-streaked to YPD plates and, after two days of growth, the resulting colonies were replica-plated to media lacking uracil or tryptophan. If all of the colonies purified from the white sector were Ura^-, the sectored colony was classified as resulting from a reciprocal crossover (Fig. 3A). If all colonies derived from the white sector were Ura^+, the strain was classified as resulting from BIR (Fig. 3B). White sectors that had mixtures of Ura^+ and Ura^- colonies were not used in our analysis; such sectored colonies could reflect events that occurred subsequent to the first division. Sectored colonies could also result from chromosome loss. In such colonies, the red sector would be Trp^-. Of 173 red/white sectored colonies, only one example of chromosome loss was observed.

The assay of genome instability based on red/white colony formation has the unfortunate characteristic of being non-selective. For the third assay, we selected for loss of the heterozygous *URA3* gene located near the right telomere by plating cells on

medium containing 5-fluoro-orotate (5-FOA). For this assay, we suspended colonies of each strain in 1 ml of water, and plated about 100 microliters of each undiluted suspension on medium containing 5-FOA (1 mg/ml), and a dilution of the suspension on non-selective medium (SD-complete medium). Between 15 and 25 colonies were examined for each genotype. From measurements of the number of 5-FOA^R colonies and the total number of cells in each colony/culture, we calculated the rate of 5-FOA^R using the method of the median (Lea and Coulson, 1949). To obtain the 95% confidence intervals for the rate estimates, we used Table B11 in Altman (1990).

Microarray analysis

DNA samples for microarray analysis were prepared by methods similar to those described in St. Charles *et al.* (2012). In brief, yeast cells were grown to stationary phase in liquid YPD cultures (5-15 ml). Cells were harvested by centrifugation, and the resulting cell pellet was resuspended in about 500 microliters of 42^0 molten agarose (0.5% low-melt agarose in 100 mM EDTA); 20 microliters of a 25 mg/ml solution of Zymolyase was then added. This mixture was distributed among about seven plug molds, each containing about 100 microliters. The samples were allowed to solidify at 4°C for 30 minutes. After solidifying, the plugs were suspended in 1ml of 10mM Tris/500mM EDTA (TE buffer), and incubated in a 15 ml tube at 37°C for at least 10 hours. 100 microliters of a 5% sarcosyl, 5 mg/ml proteinase-K in 500mM of EDTA (pH7.5) solution was then added to each sample, and the samples were incubated at 50° C for at least 12 hours. Each plug was then washed twice with TE buffer. The second incubation was performed with shaking at 4° C for at least 12 hours. After 12 hours, we did a third wash at 4^0 with TE buffer.

DNA was then extracted from four plugs of each sample using methods described in St. Charles *et al.* (2012). The samples were sonicated to yield DNA fragments of about 250 bp. The samples derived from each plug were pooled for labeling with fluorescent dyes. Each sample had about 100 micrograms/ml of DNA, and about 10 microliters was used in the labeling reactions.

For our method of analysis, the hybridization of DNA derived from experimental strains with LOH events was performed in competition with DNA from control strains heterozygous for all SNPs. The experimental strains were labeled with Cy3-dUTP, whereas the control strain was labeled with Cy5-dUTP (details in St. Charles *et al.*, 2012). The labeled nucleotides were provided as part of the Invitrogen Bioprime Array CGH Genome Labeling Module. The control and experimental labeled samples were combined, and hybridized to the microarrays. Two types of custom-made Agilent microarrays were used, one to analyze LOH events throughout the genome (St. Charles *et al.*, 2012) and one to examine LOH events on the right arm of chromosome IV (St. Charles and Petes, 2013). The sequences and locations of oligonucleotides on the whole-genome array are in Table S5 (St. Charles *et al.*, 2012), and the sequences and locations of oligonucleotides on the chromosome IV-specific array are in Table S9 (St. Charles and Petes, 2013).

After hybridization (conditions described in St. Charles *et al.*, 2012), the arrays were scanned using the GenePix scanner and GenePix Pro 6.1 software. A GenePix Results (.gpr) file was generated for each sample using the Batch Analysis feature in Gene Pix Pro 6.1. This file contains a “ratio of medians (635 nm/532 nm)” for each oligonucleotide represented on the microarray. This ratio reflects the fluorescence of the Cy5-labeled experimental sample relative to the Cy3-labeled control. Each SNP analyzed was represented by at least four 25-base oligonucleotides, two identical to the Watson and Crick strands containing the W303-1A allele and two identical to the Watson and Crick

strands of the YJM789 allele. We used programs written in Perl and R to automate the analysis and to plot hybridization levels throughout the genome (programs available on request). The resulting plots were done at two levels of resolution. Low-resolution plots depicted hybridization ratios that are the moving average of ten adjacent SNPs, whereas high-resolution plots show the ratio of medians at each individual SNP. We eliminated from the analysis any oligonucleotides that were "flagged" by the GenePix Pro 6.1 software or that had a level of fluorescence that was in the bottom 5% of intensities in both the 635 nm and 532 nm channels.

Following normalization, the hybridization ratios of heterozygous SNPs for the experimental samples were about 1 to both the W303-1A- and YJM789-related SNPs. For LOH regions in which W303-1A-related SNPs were homozygous, the ratio of hybridization to these SNPs was about 1.5 and the ratio of hybridization to YJM789-related SNPs was about 0.2. For LOH regions in which YJM789-related SNPs became homozygous, these ratios were reversed.

Microarray slides were re-used about six times by stripping the slides of the labeled DNA. Microarray slides containing DNA probes were stripped by placing them in stripping buffer (10 mM potassium phosphate, pH 6.6), and slowly heating to them to the boiling point over about one hour. After rinsing in water, the slides were stored in a nitrogen-containing cabinet. The slides that cover the microarray slides (gasket slides) were boiled in the stripping buffer for 40 minutes, rinsed in water, and dried by centrifugation. Following stripping, we usually allowed the slides to dry for two days before they were used again.

Associating genomic elements with LOH transitions in sub-cultured strains

One of the main goals of this research was to find out whether the LOH events resulting from loss of RNase H were enriched at the locations of specific genomic

elements. The rationale for our analysis is that the breakpoints associated with LOH events should be located near the site of the recombinogenic lesion. In our analysis, we examined only events in which at least two or more adjacent SNPs underwent LOH. To determine the likely "window" containing the recombination initiation site, we used the same procedure employed in our previous studies (for example, St. Charles and Petes, 2013). For interstitial LOH events, we used an association window that included all sequences located between the heterozygous SNPs that most closely flanked the LOH region. For terminal LOH events, the association window was 10 kb centromere-proximal and 10 kb centromere-distal from the homozygous SNP that was closest to the LOH event. Some events were found in most or all of the sub-cultured strains derived from a single isolate. Since these events (marked "redundant" in Table S5) were likely generated in the isolate before sub-culturing, we included redundant events in each category as single events.

We next determined whether specific genetic elements were over-represented in the association windows of different mutant strains. This analysis involved multiple steps. First, for each genotype, we summed the number of bases in the association windows over all of the individual sub-cultured isolates. For example, for the nineteen sub-cultured isolates of the *rnh201* strain about 1.35 Mb were included in the association windows (Table S5). The yeast nuclear genome as annotated in SGD, which includes only two of the approximately 150 rRNA gene repeats, is about 12.1 Mb. As discussed in the legend to Table S10, our microarrays cover about 11.6 Mb of the genome, since these arrays do not include the repetitive sub-telomeric sequences. Thus, the total amount of DNA represented on the arrays for 19 isolates is about 220.4 Mb. The amount of genomic DNA that is not present in the association windows is, therefore, 220.4 Mb – 1.35 Mb or about 219 Mb. Second, we determined the total number of specific genomic elements represented on the array. For example, there are 352 ARS elements in the

genome and 317 ARS elements represented on the array (Table S10). If these elements are placed randomly with respect to the association windows, we expect 37 ARS elements within the association windows: 317 x (1.35 Mb/220.4 Mb) x 19. The expected number of ARS elements located outside of the association windows is 5986: (317 x 19) – 37. We then counted the number of ARS elements within the association windows, determining that there were 31; the observed number of these elements outside of the association windows was 5992. Finally, we compared the observed and expected numbers by Chi-square analysis (Table S11), finding a p value (0.362) that indicates no significant association between LOH breakpoints and ARS elements. We repeated this analysis with twenty other genomic features (described below). After correction of the p values for multiple comparisons (Benjamini and Hochberg, 1995), none of these values were significant.

The numbers and locations of each genomic element tested were assembled from a variety of sources. Ty elements, solo LTR elements, centromeres, intron-containing genes, ARS elements, and tRNA genes were extracted from the S288c reference genome using the YeastMine tool on SGD (Engel *et al.* 2013, genome version R64-1-1; http://www.yeastgenome.org/help/video-tutorials/yeastmine). We also used YeastMine to determine the locations of the genes that were among the top 5% in length ("long gene" category in Tables S10 and S11). The same tool was used to identify the genes with the highest (top 5%) and lowest (bottom 5%) rates of transcription. From the sequence of the ORFs, we calculated the percentage of G bases on the non-transcribed strand. 41 ORFs with ≥ 29% G were identified and used in the association analysis. In addition, we identified 115 ORFs that had a GC-content ≥ 50%; these genes were also used in our analysis (Table S11).

Most of the other references for the locations of various genomic elements are in Table S9. Regions with converging replication forks (TER sites) were described in Table

S2 of Fachinetti *et al*. (2010). In the same paper, binding sites for the Rbp3p subunit of RNA polymerase II were mapped by chromatin immunoprecipitation followed by microarray analysis. We downloaded these data (GSM409326 on GEO, GSM409326_Rpb3_signal.bar.gz) and converted them to a .txt file. All sites with a normalized log2 value less than 0.4 were eliminated from analysis. Adjacent sites less than 1 kb apart were collapsed into single intervals, and the signal was averaged over all collapsed sites. There were 933 such intervals. As sorted by the hybridization values, we used the top 10% (93) of the intervals for our association analysis. 58 of the 71 genomic TER sites were associated with Rpb3p binding (Table S4; Fachinetti *et al.*, 2010). These sites were designated "TER sites related to high transcription" in Tables S9-S12. We also examined the association of LOH breakpoints with sites enriched for the binding of Rrm3p, a helicase involved in promoting replication through certain hard-to-replicate sequences; the map locations of these sites are in Supplemental Table 7 of Azvolinsky *et al.* (2009).

Palindromic sequences greater than 16 bp were examined using data from Lisnic *et al.* (2005), and sequences likely to form G4 quadruplex structures were obtained from Dataset S1 of Capra *et al*. (2010). Hershman *et al.* (2008) examined differential expression of genes by *S. cerevisiae* in response to N-methyl mesoporphyrin IX (NMM), a drug that stabilizes G4 quadruplexes *in vitro*. We examined the association of genes whose transcription was significantly ($p<0.001$) altered by the drug (Supplementary Table 5 of Hershman *et al.*, 2008) with the LOH events. We also examined associations with genomic regions with high levels of Elc1p, a protein involved in resolving conflicts between converging transcripts (Hobson *et al.*, 2012). The top 10% of Elc1p-binding sites (Table S1 of Hobson *et al*., 2012) were used to look for associations. The locations of RNA/DNA hybrids in *rnh1Δ rnh201Δ* strains have been recently mapped (Chan *et al.*, 2014). We examined the association of those sites that were at least ten-fold enriched

(Dataset S7 of Chan *et al.*, 2014) with our LOH data. Lastly, based on observations of a non-random association of R-loops and poly A tracts (Doug Koshland, University of California, Berkeley), we identified all uninterrupted poly A or poly T tracts that were at least 25 bases in length. We examined the association of these 41 tracts with LOH events. As described in the main text, after correction for multiple comparisons, none of the genomic elements that we examined were significantly associated with the LOH events in sub-cultured strains.

In our analysis, any elements that are within the association windows or that span the association windows are included in our analysis. For most of the genomic elements examined, the size of the element was small, less than 10% of the average size of the association window. For four of the elements (Ty elements, TER sites, TER sites associated with high levels of transcription, and "Long Genes"), however, the size of the element was greater than 10% of the size of the association window. For these comparisons, we expanded all association windows by an amount equivalent to the average size of the element. For example, when we examined the associations between Ty elements and LOH events, the association windows were expanded by 6 kb, the size of a Ty element.

Associating genomic elements with LOH transitions in sectored colonies

In sectored colonies (reflecting crossovers on the right arm of chromosome IV), the borders of the association window were the coordinates of the heterozygous SNP closest to the most centromere-proximal LOH transition and the homozygous SNP closest to the most centromere-distal LOH transition (Table S8). All association windows were used in our analysis. The right arm of chromosome IV represents about 1.4 Mb, and the number of genomic elements on the right arm of IV are given in Table S10. Our methods of calculating significant associations between genomic elements and LOH

events are chromosome IV were analogous to those described for the sub-cultured strains. We performed microarray analysis on sectored colonies of only two genotypes: *rnh201*Δ and *rnh1*Δ *rnh201*Δ. No significant associations were found between LOH events in these strains and any of the tested genomic elements (Table S12).

Regions of apparent terminal duplications/deletions at repetitive sub-telomeric regions

By microarray analysis, regions of LOH are unambiguous since the hybridization signals for one set of allelic SNPs increases for the same genomic region in which the other set of allelic SNPs decreases. From previous studies (Y. Yin and T. Petes, unpublished observations), we have found a small number of apparent terminal duplications and deletions that likely reflect LOH events on non-homologous chromosomes with shared sub-telomeric sequences. In the current study, we observed several such events among the sub-cultured strains as described below.

KO_244_1_XX_D (*rnh201*Δ *pol2-M644L*). In this isolate, we observe a terminal LOH event (YJM789-derived SNPs becoming homozygous) on chromosome X (transition coordinates 708414-728414). This strain also has a terminal deletion on the left arm of chromosome IV (transition coordinates 15561-18870), resulting in loss of W303-1A-derived sequences. In the sequence of S288c (nearly isogenic with W303-1A), we found that the chromosome X sequences between 730-742 kb are almost identical to the region 3-15 kb on chromosome IV. For example, the oligonucleotide 5211 (Table S3 in St. Charles *et al.,* 2012), near the right telomere of IV is repeated near the right telomere of X. Thus, an LOH event causing loss of W303-1A-derived SNPs from chromosome X will appear as a reduced signal of hybridization to W303-1A-specific SNPs near the right telomere of chromosome IV. It is unclear whether the YJM789-derived copy of chromosome X has the same duplication as the W303-1A-derived homolog. When sequences from a portion of the repeated region on chromosome IV (10 kb to 12 kb) are

used in a BLAST search of the YJM789 database in SGD, the sequences match to chromosome IV contigs without matching to chromosome X contigs. Finally, we note that LOH events that involve the left end of IV will not have a detectable effect on the microarray pattern observed on the right end of chromosome X because the most centromere-distal oligonucleotide on the array is at position 727 kb which is centromere-proximal to the repeated sequences.

KO_5_6_E (*rnh1*Δ *rnh201*Δ); KO_5_9_H (*rnh1*Δ *rnh201*Δ); KO_75_2_XX_H (*rnh201*Δ). These three strains had an apparent terminal deletion on chromosome VI, resulting in loss of YJM789-derived sequences. The coordinates for the deletions were similar in all three isolates beginning near coordinate 30,000 kb and proceeding to the telomere. All three strains also had terminal LOH events on chromosome X, resulting in loss of YJM789-derived SNPs and duplication of W303-1A-derived SNPs. The breakpoints of these LOH events were different in the three strains: KO_5_6_E (about 135 kb); KO_5_9_H (about 127 kb); KO_75_2_XX_H (about 36 kb). According to Wei *et al.* (2007), in YJM789, an approximately 30 kb segment derived from the left end of chromosome VI is translocated to the left end of chromosome X. From these data, the W303-1A SNPs are at the right end of chromosome VI, whereas the YJM789 SNPs are at the right end of chromosome X. Since sub-telomeric repeats are difficult to assemble, it is also possible that YJM789 has two copies of the 30 kb segment, one on VI and one on X. An LOH event on the left arm of chromosome X in which YJM789-derived sequences are lost will result in an apparent deletion of YJM789-specific sequences near the left telomere of VI. It should also be noted that there are only eight SNPs on the microarray from the 30 kb segment.

KO_75_2_XX_I (*rnh201*Δ); KO_5_9_I (*rnh1*Δ *rnh201*Δ). These strains both contain apparent terminal duplications of YJM789-derived sequences on the left arm of

chromosome XV with the starting point of the duplication near coordinate 22 kb. Both strains have terminal LOH events on the left end of chromosome IX, resulting in duplication of YJM789-derived SNPs and loss of W303-1A-derived SNPs. The breakpoints for the LOH event are near coordinate 105 kb for KO_75_2_XX_I and near coordinate 336 for KO_5_9_I. The sub-telomeric regions of chromosomes IX and XV share considerable homology in both the YJM789 and W303-1A/S288c genomes. In the S288c genomes, sequences from chromosome XV with coordinates about 22-31 kb share extensive homology with sequences from chromosome IX located between coordinates 17-26 kb. However, according the SGD database and our analysis, the oligonucleotides that have an elevated signal on chromosome XV are not in the region of the genome that is repeated on chromosome IX in either the W303-1A or the YJM789 genomes. For example, in isolate KO_75_2_XX_I, the YJM789-specific oligonucleotide at position 22005 clearly has an elevated level of hybridization. The sequence of this oligonucleotide, however, is not present on chromosome IX. Similarly, in isolate KO_5_9_I, the level of hybridization to the YJM789-specific oligonucleotide 19925 is elevated, although the sequence of this oligonucleotide is not on chromosome IX.

Although we do not have a definitive explanation of these observations, one possibility is that the YJM789 isolate used in our studies has a derivative of chromosome IX in which the terminal 22 kb of chromosome X replaces the terminal 16 kb of chromosome IX. Such a derivative could be formed as a consequence of a break-induced replication event in which a broken end of the YJM789-derived copy of chromosome IX duplicates a portion of the YJM789-derived chromosome XV homolog. The initiation point of this invasion would be in the region of shared homology. In diploid strains with this derivative, an LOH event on chromosome IX, occurring centromere-proximal to the duplicated region, would duplicate both YJM789-related SNPs on chromosome XV and cause a duplication of YJM789-derived sequences from the

terminal region of chromosome XV. It should be noted that an LOH event on chromosome XV that results in LOH for the terminal repeated sequences would not be annotated as an LOH event on chromosome IX since the first oligonucleotide used to diagnose LOH for chromosome IX is located at coordinate 25 kb. Although this model is consistent with our observations, our observations could also reflect an assembly error of the genomic sequences.

KO_5_6_K (*rnh1*Δ *rnh201*Δ). In this isolate, there is an apparent duplication of YJM789-derived sequences on the left arm of chromosome XVI with a transition point between coordinates 23222 and 26225. There were also terminal LOH events on several chromosome arms including the right arm of chromosome XIII, duplicating YJM789-derived SNPs; the LOH event on XV has a transition between coordinates 889 and 892 kb. The left arm of chromosome XVI (25.8-26.4 kb) shares homology with the right arm of chromosome XIII (coordinates 917.5 kb to 917.8 kb), although the oligonucleotides that have increased levels of hybridization on chromosome XVI in KO_5_6_K are not annotated as duplicated on any other homolog. One explanation of the data is that a break within the shared homology occurred on chromosome XIII that was repaired by a BIR event involving chromosome XVI. A subsequent LOH event on XIII could result in the observed apparent duplication of XVI sequences as well as the terminal LOH event duplicating YJM789-derived SNPs. Alternatively, there may be an incorrect assembly or annotation of the sub-telomeric sequences in the databases.

KO_5_9_D (*rnh1*Δ *rnh201*Δ). In this isolate, there is an apparent deletion of W303-1A-derived sequences on the right arm of chromosome I with a transition point between coordinates 195120 and 203572. There is also a terminal LOH event on the right arm of chromosome VIII, resulting in loss of W303-1A-derived sequences, with a transition point between 199775 and 207066. In the S288c genome, there is a large region of conserved

homology that includes the coordinates 207-227 kb on the right end of chromosome I and coordinates 528-556 kb on the right end of VIII. Several of the SNPs located near the right end of chromosome I (for example, oligonucleotide 208214; Table S3, St. Charles *et al.*, 2012) are duplicated on the right end of VIII. Therefore, an LOH event that causes loss of W303-1A-derived SNPs and duplication of YJM789-derived SNPs will result in an apparent deletion of W303-1A sequences from the right end of chromosome I.

Supplemental Figure Legends

Figure S1. Patterns of LOH in sub-cultured strains. Each line represents markers in a diploid isolate. Green indicates heterozygous SNPs; red, homozygous W303-1A-derived SNPs; black, homozygous YJM789-derived SNPs. The yellow circle shows the centromere. Each transition between heterozygous and homozygous SNPs or between two regions with different homozygous SNPs is labeled with a lower case letter. Classes a1-a4 are simple terminal LOH events. In Classes a6-a8, the two transitions (one marked with an asterisk) are separated by distances that are two standard deviations longer than the median length of a mitotic conversion tract. The two transitions are, therefore, likely to reflect two different recombination events. Classes b1 and b2 represent simple interstitial LOH events (gene conversions), whereas in Classes b3-b5, the conversion event is interrupted by a region of heterozygosity. Classes f1-f12 represent terminal LOH events with complex patterns of associated LOH events. Only Classes a1-a4, b1, and b2 were used for our association studies.

Figure S2. Deletions and duplications in sub-cultured strains. This diagram shows the patterns of deletions and duplications in diploid isolates. As in Fig. S1, the green line indicates heterozygous SNPs, and yellow circles show the centromere. The deletion or duplication is shown as a line that is half as wide as the green lines. The Classes dd9 and dd10 show interstitial duplications of W303-1A-derived and YJM789-derived SNPs, respectively. The Class dd12 shows an interstitial deletion in which W303-1A-derived SNPs were removed. The coordinates for the transitions are in Table S6.

Figure S3. Aneuploidy events in sub-cultured strains. Trisomic, but not monosomic, aneuploid events were observed in our studies. For each chromosome, we indicate whether the strain has W303-1A-derived SNPs (red) or YJM789-derived SNPs (black). Note that many of these aneuploid events are associated with recombination on one or more chromosomes.

Figure S4. Locations of LOH events in the sub-cultured *rnh201Δ* strain. Each of the sixteen chromosomes is shown as a thin black horizontal line with SNPs shown as very short vertical yellow lines. The centromeres are represented by black ovals. Red and blue bars show regions of interstitial LOH in which the W303-1A-derived SNPs became homozygous and the YJM789-derived SNPs became homozygous, respectively. Black arrows indicate the positions of terminal LOH events that were unassociated with a gene conversion event. Red arrows and blue arrows indicate terminal LOH events that were associated with a conversion that made W303-1A-derived SNPs and YJM789-derived SNPs homozygous, respectively. Triangles indicate deletions (red for a deletion of W303-1A-derived sequences and blue for a deletion of YJM789-derived sequences) and inverted triangles indicate duplications (same color code as for deletions).

Figure S5. Location of LOH events in the sub-cultured *rnh1Δ rnh201Δ* strain. The mapped events are shown with the same code as in Fig. S4.

Figure S6. Location of LOH events in the sub-cultured *rnh201Δ pol2-M644L* strain. The mapped events are shown with the same code as in Fig. S4.

Figure S7. Patterns of LOH in sectored colonies. In this depiction, each sectored colony is represented by a pair of lines with the red sector shown as the top line. We use the same color code for heterozygous and homozygous regions as in Fig. S1. Other features of these patterns are described in the main text.

Supplemental Tables.

Table S1. Strain list.

Table S2. Plasmid list.

Table S3. Primers list.

Table S4. Strains used in different assays of LOH.

Table S5. LOH events in sub-cultured strains.

Table S6. Deletion/duplication events in sub-cultured strains.

Table S7. Trisomy events in sub-cultured strains.

Table S8. LOH events on chromosome IV in sectored colonies.

Table S9. References used to determine the locations of genomic elements.

Table S10. Number of genomic elements represented on the microarrays.

Table S11. Association of LOH events in sub-cultured strains with various genomic elements.

Table S12. Association of LOH events in sectored colonies with various genomic elements on the right arm of chromosome IV.

References

Altman, D. G. (1990) *Practical Statistics for Medical Research*. (Chapman and Hall/CRC Texts in Statistical Science).

Aksenova, A.Y, P.W. Greenwell, M. Dominska, A.A. Shishkin, J.C. Kim, T.D. Petes, and S.M. Mirkin (2013) Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proc Natl Acad Sci USA* 110(49): 19866-19871.

Azvolinsky, A., P.G. Giresi, J.D. Lieb, and V.A. Zakian (2009) Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. Mol Cell 34: 722-34.

Benjamini Y. and Y. Hochberg (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc A Sta* 57(1): 289-300.

Capra, J.A., K. Paeschke, M. Singh, V.A. Zakian (2010) G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in*Saccharomyces cerevisiae*. *PLoS Comput Biol* 6: e1000861

Chan, Y. A., M.J. Aristizabal, Y.T. Lu Phoebe, Z. Luo, A. Hamza, M.S. Kobor, P.C. Stirling, P. Hieter (2014) Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-Chip. *PLoS Genet* 10(4): e1004288.

Engel S.R., F.S. Dietrich, D.G. Fisk, G. Binkley, R. Balakrishnan, M.C. Costanzo, S.S. Dwight, B.C. Hitz, K. Karra, R.S. Nash, S. Weng, E.D. Wong, P. Lloyd, M.S. Skrzypek, S.R. Miyasato, M. Simison, J.M. Cherry (2013) The Reference Genome Sequence of Saccharomyces cerevisiae: Then and Now. *G3* pii: g3.113.008995v1. doi: 10.1534/g3.113.008995.

Fachinetti, D., R. Bermejo, A. Cocito, S. Minardi, Y. Katou, Y. Kanoh, K. Shirahige, A. Azvolinsky, V.A. Zakian, M. Foiani (2010) Replication termination at eukaryotic

chromosomes is mediated by Top2 and occurs at genomic loci containing pause elements. *Molecular Cell*. 39(4): 595-605.

Goldstein, A.L., and J.H. McCusker (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* 15(14): 1541-1553.

Gueldener, U., J. Heinisch, G.J. Koehler, D. Voss, and J.H. Hegemann (2002) A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 30(6): e23.

Hershman, S.G., Q. Chen, J.Y. Lee, M.L. Kozak, P. Yue, L.S. Wang, and F.B. Johnson (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 36(1): 144-156.

Hobson, D. J., W. Wei, L.M. Steinmetz, J.Q. Svejstrup (2012) RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48(3): 365-374.

Kim, N., J. Cho, Y.C. Li, and S. Jinks-Robertson (2013) RNA:DNA hybrids initiate quasi-palindrome associated mutations in highly transcribed yeast DNA. *PLoS Genet* 9(11): e1003924.

Lea, D.E. and C.A. Coulson (1949) The distribution of number of mutants in a bacterial population. *J Genet* 49: 264-285.

Lee, P.S., P.W. Greenwell, M. Dominska, M. Gawel, M. Hamilton, and T.D. Petes (2009) A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. *PLoS Genet* 5(3): e1000410.

Lisnic, B., I.K. Svetec, H. Saric, I. Nikolic, and Z. Zgaga (2005) Palindrome content of the yeast *Saccharomyces cerevisiae* genome. *Curr Genet* 47: 289-297.

Longtine, M.S., A. McKenzie III, D.J. Demarini, N.G. Shah, A. Wach, A. Brachat, P. Philippsen, and J.R. Pringle (1998) Additional modules for versatile and economical

PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14: 953-961.

Mortimer, R.K. and J.R. Johnston (1986) Genealogy of principal strains of the yeast genetic stock center. *Genetics* 113(1): 35-43.

Nick McElhinny, S.A., D. Kumar, A.B. Clark, D.L. Watt, B.E. Watts, E.B. Lundström, E. Johansson, A. Chabes, and T.A. Kunkel (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol* 6(10): 774-781.

Song, W., M. Dominska, P.W. Greenwell, and T.D. Petes (2014) Genome-wide high-resolution mapping of chromosome fragile sites in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 111(21): e2210-e2218.

St. Charles, J., E. Hazkani-Covo, Y. Yin, S.L. Andersen, F.S. Dietrich, P.W. Greenwell, E. Malc, P. Mieczkowski, and T.D. Petes (2012) High-resolution genome-wide analysis of irradiated (UV and γ-rays) diploid yeast cells reveals a high frequency of genomic loss of heterozygosity (LOH) events. *Genetics* 190(4): 1267-1284.

St. Charles, J. and T.D. Petes (2013) High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1Mb arm of yeast chromosome IV. *PLoS Genet* 9(4): e1003434.

Thomas, B.J. and R. Rothstein (1989) Elevated recombination rates in transcriptionally active DNA. *Cell* 56(4): 619-630.

Vernon, M., K. Lobachev, and T.D. Petes (2008) High rates of “unselected” aneuploidy and chromosome rearrangements in *tel1 mec1* haploid yeast strains. *Genetics* 179(1): 237-247.

Wei, W., J.H. McCusker, R.W. Hyman, T. Jones, Y. Ning, Z. Cao, Z. Gu, D. Bruno, M. Miranda, M. Nguyen, J. Wilhelmy, C. Komp, R. Tamse, X. Wang, P. Jia, P. Luedi, P.J. Oefner, L. David, F.S. Dietrich, Y. Li, R.W. Davis, and L.M. Steinmetz (2007)

Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc Natl Acad Sci USA* 104(31): 12825-12830.

Zhao, X., E.G. Muller, and R. Rothstein (1998) A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Mol Cell* 2: 329-340.