PNAS
www.pnas.org

Supplementary Information for

Genome-wide mapping of spontaneous genetic alterations in diploid yeast cells

Yang Sui[1,2,#], Lei Qi[1,2,#], Jian-Kun Wu[1], Xue-Ping Wen[1], Xing-Xing Tang[1], Zhong-Jun Ma[1], Xue-Chang Wu[3], Ke Zhang[3,*], Robert J. Kokoska[4], Dao-Qiong Zheng[1,*], Thomas D. Petes[2,*]

Co-corresponding authors:
Ke Zhang
Email: zhangke726@zju.edu.cn
Dao-Qiong Zheng
Email: zhengdaoqiong@zju.edu.cn
Thomas D. Petes
Email: tom.petes@duke.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

www.pnas.org/cgi/doi/10.1073/pnas.2018633117

Supplementary Information Text

Whole-genome sequencing of diploid strains

Genomic DNA was extracted using the Omega yeast DNA kit (Life Science Products) and sequenced by 150 bp pair-end strategy of the Illumina Nextseq500 platform. To obtain .sam files (1), we mapped the reads to the S288c reference sequence using the "mem" algorithm of the BWA software. Samtools was used to transform the .sam files to the .bam files (2). We identified LOH regions using the same protocols described in Zheng *et al.* (2016) (3). Single-base changes and in/dels less than 10 bp were analyzed using the software FreeBayes with the .bam files as inputs (4). Deletions in the range of 10-100 bp were identified by LUMPY software (5). The effects of point mutations on protein sequence was assessed by SnpEff software (6). We also confirmed 12 of the mutations diagnosed by whole-genome sequencing (7 single-base mutations, 4 in/dels, and one complex mutation) by PCR amplification of sequences (using primers described in Table S1-2), followed by Sanger sequencing. All were confirmed.

Details concerning recombination with the rRNA gene cluster.

As discussed in the main text, we often observed one ribosomal RNA gene cluster that increased in size in the same cell in which the other cluster decreased in size. There are two mechanisms that could produce this pattern. First, as shown in Fig. S8A, unequal sister-strand crossovers would result in one daughter cell with a shorter array and one with a longer array. If the daughter cell with the shorter array grew more slowly that the original cell, there would be a selective advantage for a second unequal crossover to increase the array size. If the second event occurred in the opposite homolog from the first event, the coupled deletion/duplication pattern would be observed. The second mechanism would be a crossover within the rDNA cluster between the two homologs (Fig. 8B). One daughter would have fewer W303-1A rDNA repeats and more YJM789 repeats, and the other daughter would have the reciprocal pattern.

The distinction between these two models is that a crossover between rRNA gene clusters on different homologs would result in LOH for all SNPs located centromere-distal to the cluster on chromosome XII. Of the 93 isolates, 19 had LOH patterns expected for crossovers within the clusters on different homologs (Dataset S4-1). In 16 of these 19 isolates, we observed a loss of one type of repeat and a gain of the other. In 14 of these 16 events, the sequences distal to the cluster are those predicted by the clusters that gained or lost repeats. For example, in Fig. S8B, the D1 daughter cell in which the number of W303-type repeats is increased is homozygous for W303-derived SNPs distal to the rRNA gene cluster. Although the reciprocal crossover events within the cluster may account for some of reciprocal gains and losses in our experiments, the majority of the 76 isolates with this pattern likely involve other mechanisms such as unequal sister-strand crossovers.

A crossover that occurs proximal to the rRNA gene cluster should result in a cell that has all rRNA genes derived from one of the two homologs. However, for several such isolates, although most of the repeats were derived from one homolog, a few were derived from the other. For example, in isolate Spo11-143 (Dataset S3), 202 rRNA genes had SNPs derived from W303-1A and 2 had SNPs derived from YJM789. A simple explanation of this pattern is that, prior to the crossover centromere-proximal to the rRNA gene cluster, there was a gene conversion event in which two repeats were transferred from the YJM789 homolog to the W303-1A-derived homolog.

Calculations of hotspot activity

We calculated hotspot activity separately for I-LOH events and T-LOH events. Dataset S6, we calculated that, among the 93 isolates, 18327 SNPs were included within interstitial LOH tracts. Since there are 45174 SNPs in each isolate, the probability that an individual SNP would be included in an interstitial LOH event per isolate is 18327/45174x93 or 4.36×10^{-3}. The probability that an individual SNP within one isolate would not be included in a conversion event is 1-.00436 or 0.9956. Based on a binary distribution, the probabilities that any specific SNP will be involved 0, 1, 2, 3, 4, and 5 times among 93 isolates are 0.67, 0.27, 0.05, 0.007, 0.0007, and 0.00006, respectively. If we multiply these numbers by the number of SNPs in the genome (45174), the expected numbers of SNPs with 0, 1, 2, 3, 4 and 5 events among the 93 isolates are 30,266, 12,197, 2,258, 316, 31, and 3, respectively. No SNPs were over-represented in conversion events as calculated by this distribution.

A similar calculation was done for the breakpoints of T-LOH events. We calculated the likely breakpoint of T-LOH events in a window of 20 kb flanking the transition between heterozygous and homozygous SNPs (as described in the footnote of Dataset S4-1). Since there are 358 T-LOH events, the probability that a SNP will be within the breakpoint window within one isolate is (358 x 20 kb)/(45174 x 93) or 1.7×10^{-3}. The probabilities that any specific SNP will be involved 0, 1, 2, 3, 4, and 5 times among 93 isolates are 0.85, 0.13, 0.01, 0.0005, 0.00002, and <0.000001, respectively. The expected numbers of SNPs with 0, 1, 2, 3, 4 and 5 events among the 93 isolates are 38,398, 5,873, 452, 23, 1, and <1, respectively. Since we observed multiple SNPs with 5 or more LOH events within the breakpoints (Fig. 5), these regions are significantly elevated for the initiation of crossovers.

Calculation of expected SNP-specific rates of LOH

The expected rates of LOH (R_{LOH}) for individual SNPs are the sum of two rates: the rate expected from I-LOH events ($R_{I\text{-}LOH}$) and the rate expected from T-LOH events ($R_{T\text{-}LOH}$). Since mitotic gene conversions are distributed reasonably uniformly in the genome in our analysis, we will assume that $R_{I\text{-}LOH}$ is the same for all SNPs. There are two related methods for calculating $R_{I\text{-}LOH}$. One method is to multiply the genomic rate of interstitial LOH times the ratio of the average conversion

tract size divided by the total genome size. Excluding the ribosomal RNA genes, we estimate $R_{I\text{-}LOH}$ as ($3.3x10^{-3}$ interstitial events/cell division) x (2.8 kb/$11.6x10^{3}$ kb) or $7.9x10^{-7}$. This rate is similar to that determined for conversion of a point mutation in the *URA3* gene ($1.3x10^{-6}$) by Yim *et al.* (7).

We can also calculate $R_{I\text{-}LOH}$ based on the number of SNPs that undergo interstitial LOH in all isolates (Dataset S6). Summed over 93 isolates, we observed 18327 SNPs with interstitial LOH; the total number of SNPs per strain is 45174. The average number of SNPs with interstitial LOH events per isolate is: 18327/93 or 197. The average frequency with which an individual SNP within one isolate will become homozygous by a conversion event following sub-culturing is: 197/45174 or $4.26x10^{-3}$. The rate at which an individual SNP will have an LOH event/cell division is: $4.26x10^{-3}$/2839 (the average number of cell divisions per isolate) or $1.5x10^{-6}$. This rate estimate is similar to the rate calculated above ($7.9x10^{-7}$) based on the conversion tract size and the number of conversions. Thus, using two different methods, we calculate the rate of I-LOH (gene conversion) to be about 10^{-6}/SNP/cell division, equivalent to 10^{-6}/bp/cell division

The second rate ($R_{T\text{-}LOH}$) is a function of the distance of the SNP from the centromere of the chromosome, since a crossover occurring anywhere in this interval will result in LOH for centromere-distal SNPs. Therefore, we can estimate the rate of LOH caused by crossovers/BIR events for individual SNPs by multiplying the rate of terminal LOH events in the genome ($1.4x10^{-3}$/cell division) by the ratio of the distance of the SNP from the centromere (in kb) divided by the total genome length (12.5 Mb); for the genome length, we include the rRNA gene cluster since we can detect terminal LOH events in this region. Thus, $R_{T\text{-}LOH}$ will vary from 0 (for SNPs very close to the centromere) to about $2x10^{-4}$/SNP/cell division for a marker located at the end of the right arm of chromosome XII, the longest chromosome arm in the genome.

For SNPs located near the ends of the chromosome, the expected number of isolates in which the SNP has undergone LOH can be predicted by multiplying the $R_{T\text{-}LOH}$ value by the aggregate number of cell divisions (264,000); alternatively, the expected number can be calculated by multiplying the total number of T-LOH events (358) by the ratio of the distance of the SNP from the centromere divided by the total genome size. For a SNP located near the end of chromosome XII, by either calculation, the predicted number of events is about 52; the observed number is 47. As shown in Dataset S7, in general, there is reasonably good agreement between the calculated and observed numbers of events per chromosome arm.

One exception to the agreement between the numbers of observed and expected events is the number of events involving the terminal SNP on the right end of chromosome IV. We expected 32 events and observed only 11 ($p=0.001$ by chi-square test). A simple explanation of this discrepancy is that the diploid strain WYspol11 used in our study is heterozygous for the *SUP4-o* gene located near the right end of chromosome IV. The *SUP4-o* gene partly suppresses *ade2-1*, an ochre-suppressible mutation that is homozygous in WYspo11. A mitotic crossover located centromere-proximal to *SUP4-o* will result in two daughter cells, one that has lost the suppressor and one that

has duplicated the suppressor. Strains that have an unsuppressed mutation in *ade2* accumulate a red pigment and grow slowly (8). Such strains would be selected against in MA experiments. As expected from these considerations, 10 of 11 strains that had an LOH event at *SUP4-o* involved duplication, rather than loss, of *SUP4-o*.

The rate of T-LOH events can also be used to estimate the number of clonal divisions before a sexual cycle. By analyzing diploid isolates, Magwene *et al.* (2011) (9) found that many strains were heterozygous at many (about 20,000) positions. Within these strains, there were large terminal regions that were homozygous, presumably reflecting mitotic crossovers or BIR. For example, one strain PMY070 had four large T-LOH regions. By dividing 4 by the rate of T-LOH/cell division, one can estimate the number of divisions since the diploid was formed (9). Using our estimate of the rate of crossovers, we calculate that there were about 2857 clonal generations since the previous sexual cycle in PMY070.

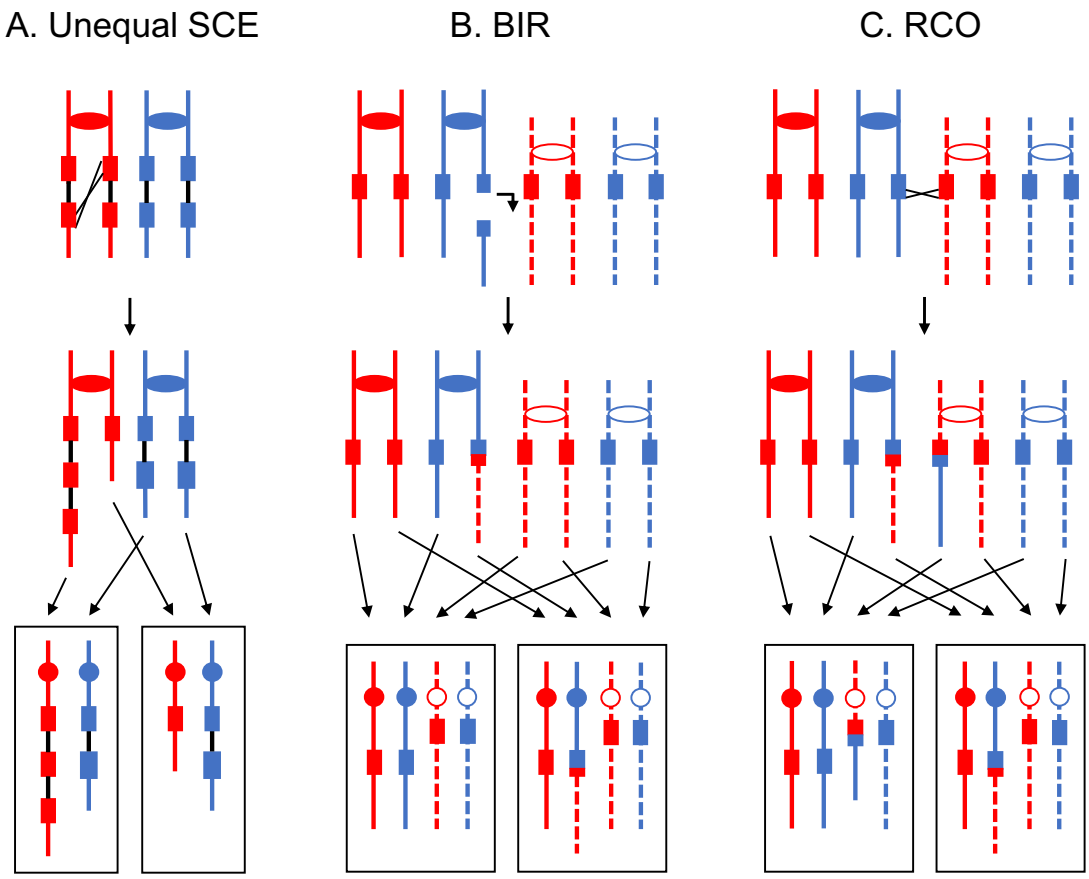


Fig. S1. Large deletions and duplications generated by homologous recombination between dispersed repeats. Red and blue rectangles represent repeated genes, and ovals/circles indicate centromeres. Black rectangles outline daughter cells.

A. Interstitial deletion and duplication formed by homologous recombination between direct repeats on sister chromatids. The region located between the repeats (black lines) is either duplicated (left daughter cell) or deleted (right daughter cell).

B. Terminal deletion and duplication generated by break-induced replication (BIR) event between repeats on non-homologous chromosomes, forming translocations. Different homologs are indicated by unbroken or dotted lines. The righthand daughter cell has a terminal duplication of sequences derived from the “dotted” red homolog, and a terminal deletion of sequences derived from the “undotted” blue homolog.

C. Terminal deletion and duplication generated by reciprocal crossover (RCO) between repeats on non-homologous chromosomes, forming translocations. Both daughter cells have terminal deletions and duplications.

Fig. S2. Chromosomal location of single-base changes in sub-cultured isolates of WYspo11.

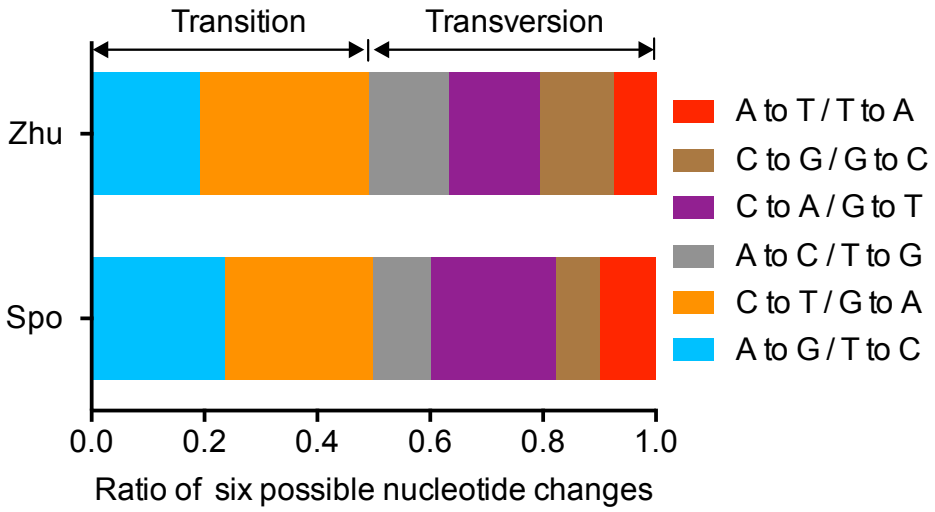


Fig. S3. Different types of base changes observed in sub-cultured isolates of WYspo11 (total of 1276 mutations). In this figure, we compare our distribution of single-base changes to data of Zhu *et al.* (2014) (10).

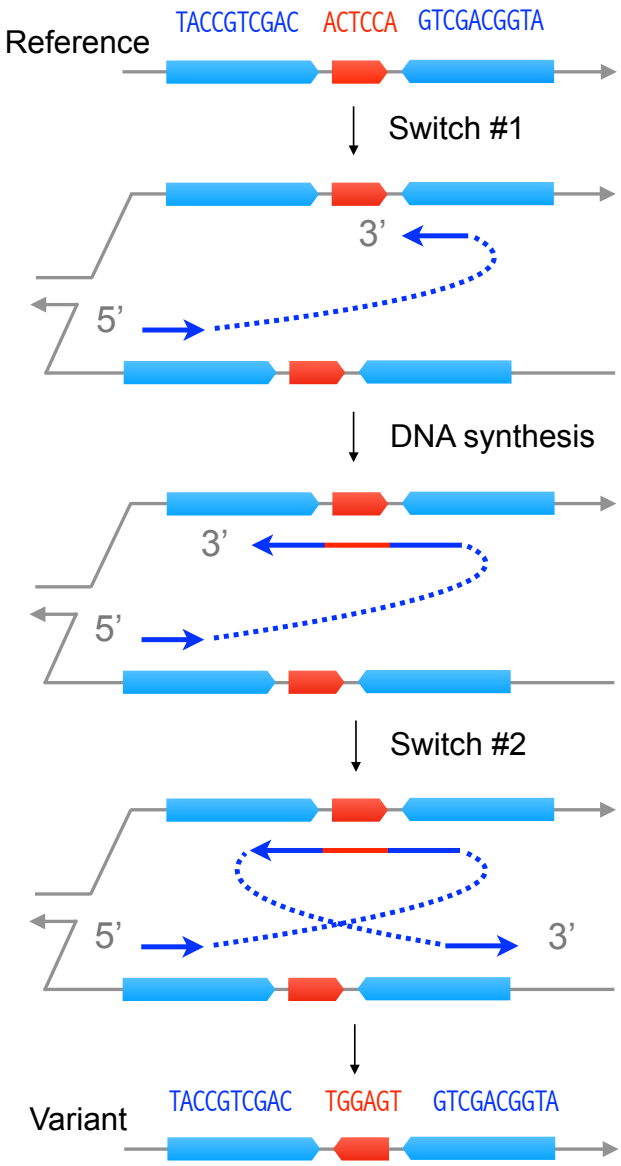


Fig. S4. An example of a complex mutation (inversion between two short non-adjacent palindromes). The palindromes are shown in blue and the inverted region in red. One mechanism by which this class of mutation could be produced is by two template switches between sister chromatids. The mutation is shown in Row 3 of Dataset S1-5.

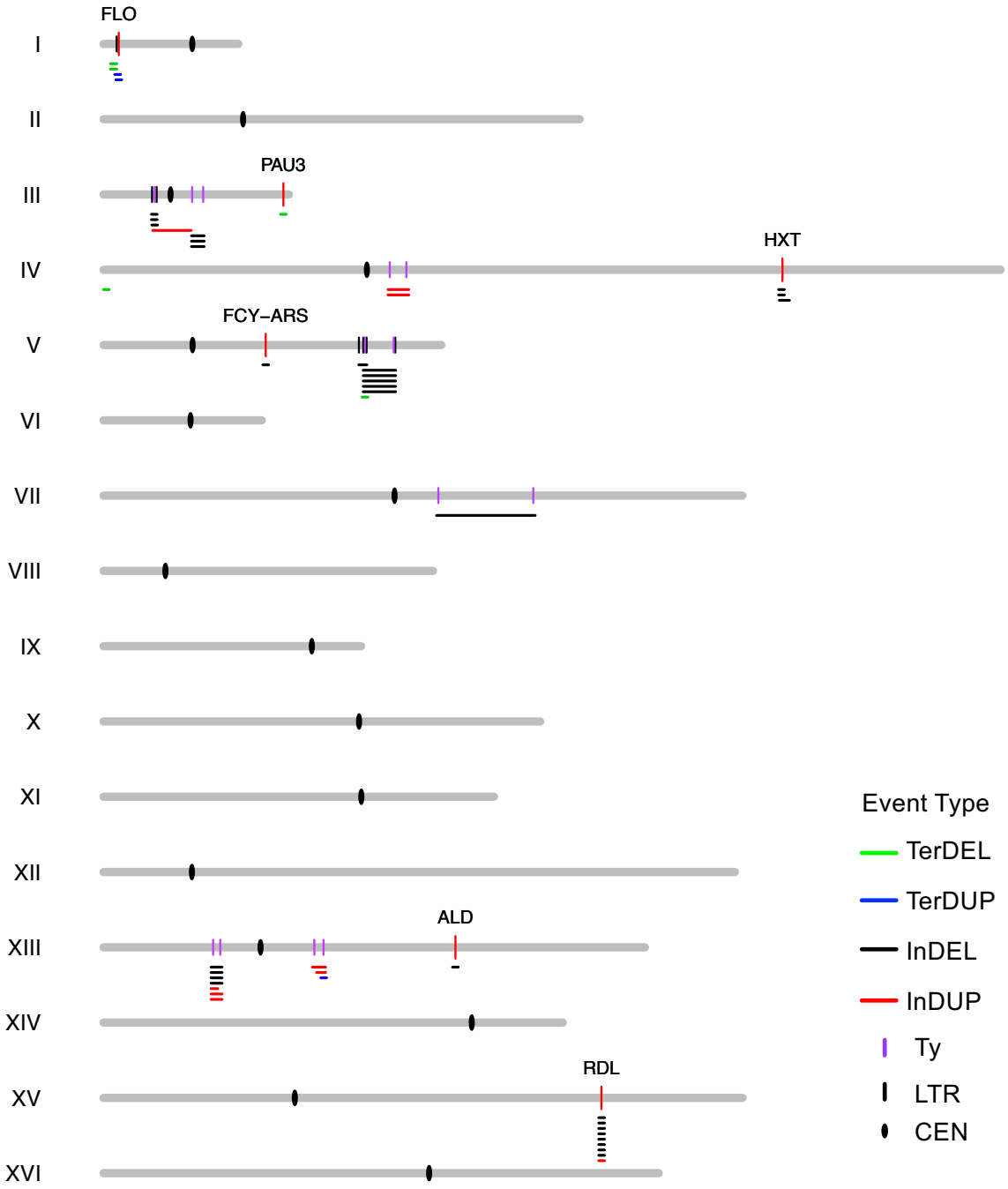


Fig. S5. Distribution of large (>1kb) deletions and duplications in the yeast genome.

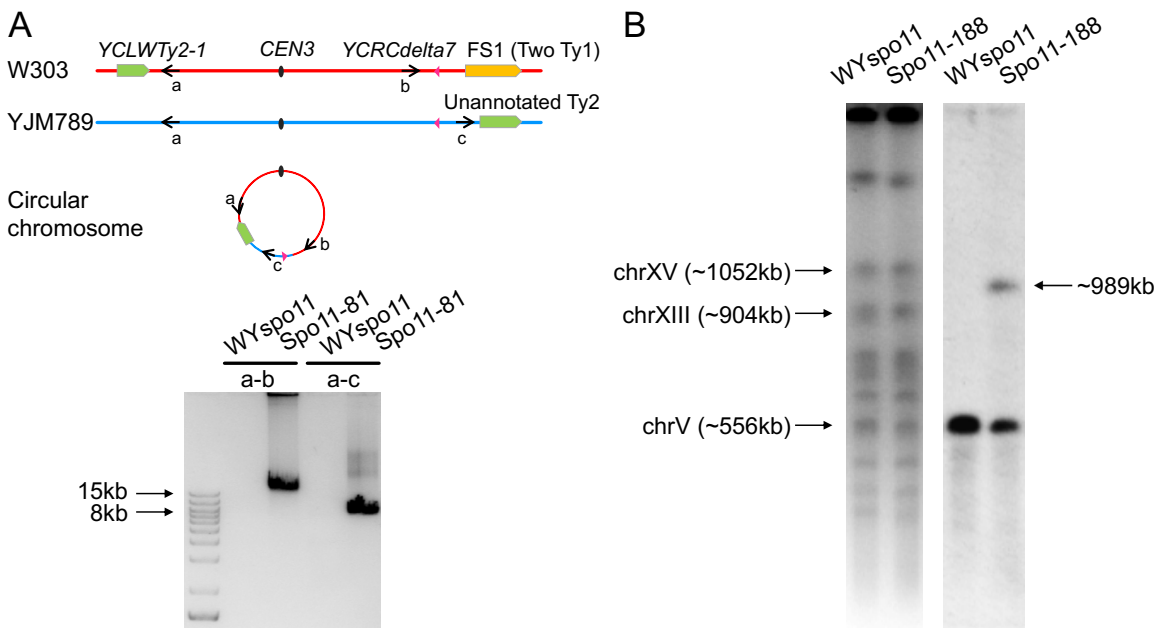
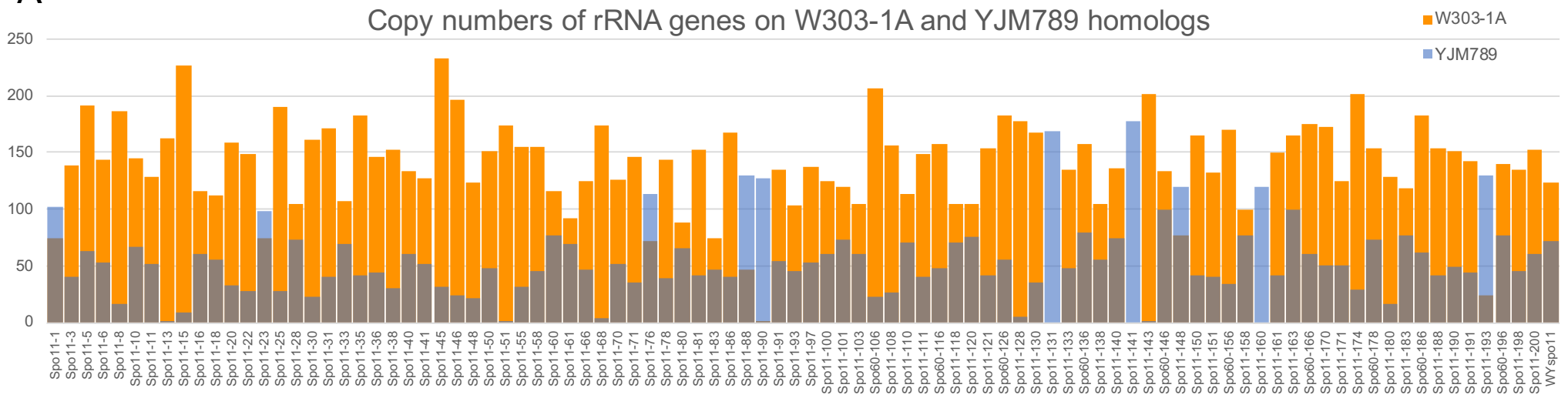


Fig. S6. Chromosome rearrangements formed by recombination between retrotransposons.

A. Duplication caused by circle formation in Spo11-81. By sequencing, the duplication had a complex pattern in which most of the duplication was derived from the W303-1A homolog and a small region on the right side of the duplication was derived from the YJM789 homolog. This pattern could result from a gene conversion between the two homologs, transferring a region flanking the Ty2 on the YJM789 homolog to the W303-1A homolog. Subsequently, a crossover between the unannotated Ty2 and *YCLWTy2-1* could form a circle containing the duplicated region and *CEN3*. This model predicts that PCR fragments generated using primers “a” and “b”, or “a” and “c” would be observed in the strain containing the circle, but not in the control strain, as shown in the gel at the bottom of Fig. S6A as observed.

B. Translocation between chromosomes V and XIII. A clamped-homogeneous-electric-field (CHEF) gel was used to analyze chromosomes in a strain (Spo11-188) with a terminal deletion and a terminal duplication (Fig. 2D). The first two lanes show ethidium bromide-stained chromosomes with the positions of chromosomes V, XIII, and XV marked; the first lane has DNA from the control strain WYspo11. Based on sequencing analysis, the expected size of the translocation between V and XIII was about 1000 kb. Although a chromosome of this size is not clearly visualized by staining, when the same gel was hybridized to a probe to chromosome V, in Spo11-188, we observed two bands, one of the expected sizes of V and one at a position of about 1000 kb. The control strain had a single band of the size of V.

A


Copy numbers of rRNA genes on W303-1A and YJM789 homologs
W303-1A
YJM789
250
200
150
100
50
0
Spo11-1
Spo11-3
Spo11-5
Spo11-6
Spo11-8
Spo11-10
Spo11-11
Spo11-13
Spo11-15
Spo11-16
Spo11-18
Spo11-20
Spo11-22
Spo11-23
Spo11-25
Spo11-28
Spo11-30
Spo11-31
Spo11-33
Spo11-35
Spo11-36
Spo11-38
Spo11-40
Spo11-41
Spo11-45
Spo11-46
Spo11-48
Spo11-50
Spo11-51
Spo11-55
Spo11-58
Spo11-60
Spo11-61
Spo11-66
Spo11-68
Spo11-70
Spo11-71
Spo11-76
Spo11-78
Spo11-80
Spo11-81
Spo11-83
Spo11-86
Spo11-88
Spo11-90
Spo11-91
Spo11-93
Spo11-97
Spo11-100
Spo11-101
Spo11-103
Spo60-106
Spo11-108
Spo11-110
Spo11-111
Spo60-116
Spo11-118
Spo11-120
Spo11-121
Spo60-126
Spo11-128
Spo11-130
Spo11-131
Spo11-133
Spo60-136
Spo11-138
Spo11-140
Spo11-141
Spo11-143
Spo60-146
Spo11-148
Spo11-150
Spo11-151
Spo60-156
Spo11-158
Spo11-160
Spo11-161
Spo11-163
Spo60-166
Spo11-170
Spo11-171
Spo11-174
Spo60-178
Spo11-180
Spo11-183
Spo60-186
Spo11-188
Spo11-190
Spo11-191
Spo11-193
Spo60-196
Spo11-198
Spo11-200
WYspo11

B

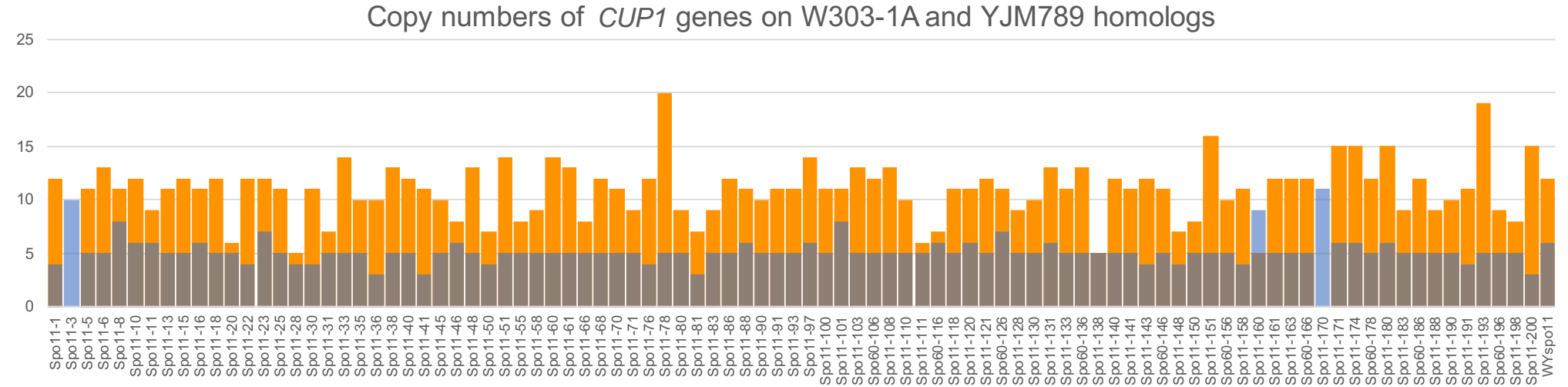

Copy numbers of CUP1 genes on W303-1A and YJM789 homologs
25
20
15
10
5
0
Spo11-1
Spo11-3
Spo11-5
Spo11-6
Spo11-8
Spo11-10
Spo11-11
Spo11-13
Spo11-15
Spo11-16
Spo11-18
Spo11-20
Spo11-22
Spo11-23
Spo11-25
Spo11-28
Spo11-30
Spo11-31
Spo11-33
Spo11-35
Spo11-36
Spo11-38
Spo11-40
Spo11-41
Spo11-45
Spo11-46
Spo11-48
Spo11-50
Spo11-51
Spo11-55
Spo11-58
Spo11-60
Spo11-61
Spo11-66
Spo11-68
Spo11-70
Spo11-71
Spo11-76
Spo11-78
Spo11-80
Spo11-81
Spo11-83
Spo11-86
Spo11-88
Spo11-90
Spo11-91
Spo11-93
Spo11-97
Spo11-100
Spo11-101
Spo11-103
Spo60-106
Spo11-108
Spo11-110
Spo11-111
Spo60-116
Spo11-118
Spo11-120
Spo11-121
Spo60-126
Spo11-128
Spo11-130
Spo11-131
Spo11-133
Spo60-136
Spo11-138
Spo11-140
Spo11-141
Spo11-143
Spo60-146
Spo11-148
Spo11-150
Spo11-151
Spo60-156
Spo11-158
Spo11-160
Spo11-161
Spo11-163
Spo60-166
Spo11-170
Spo11-171
Spo11-174
Spo60-178
Spo11-180
Spo11-183
Spo60-186
Spo11-188
Spo11-190
Spo11-191
Spo11-193
Spo60-196
Spo11-198
Spo11-200
WYspo11

Fig. S7. Changes in the number of ribosomal RNA genes and *CUP1* genes in sub-cultured isolates of WYspo11.

A. In this figure, we show the numbers of ribosomal RNA genes derived from the W303-1A-derived homolog (orange) and the numbers derived from the YJM789-derived homolog (purple/blue). The values for the control strain WYspo11 are shown at the right end of the figure. As described in the text, these numbers are determined by the number of “reads” of ribosomal RNA genes relative to the number of “reads” for single-copy sequences.

B. The numbers of two types of *CUP1* genes for each isolate are shown.

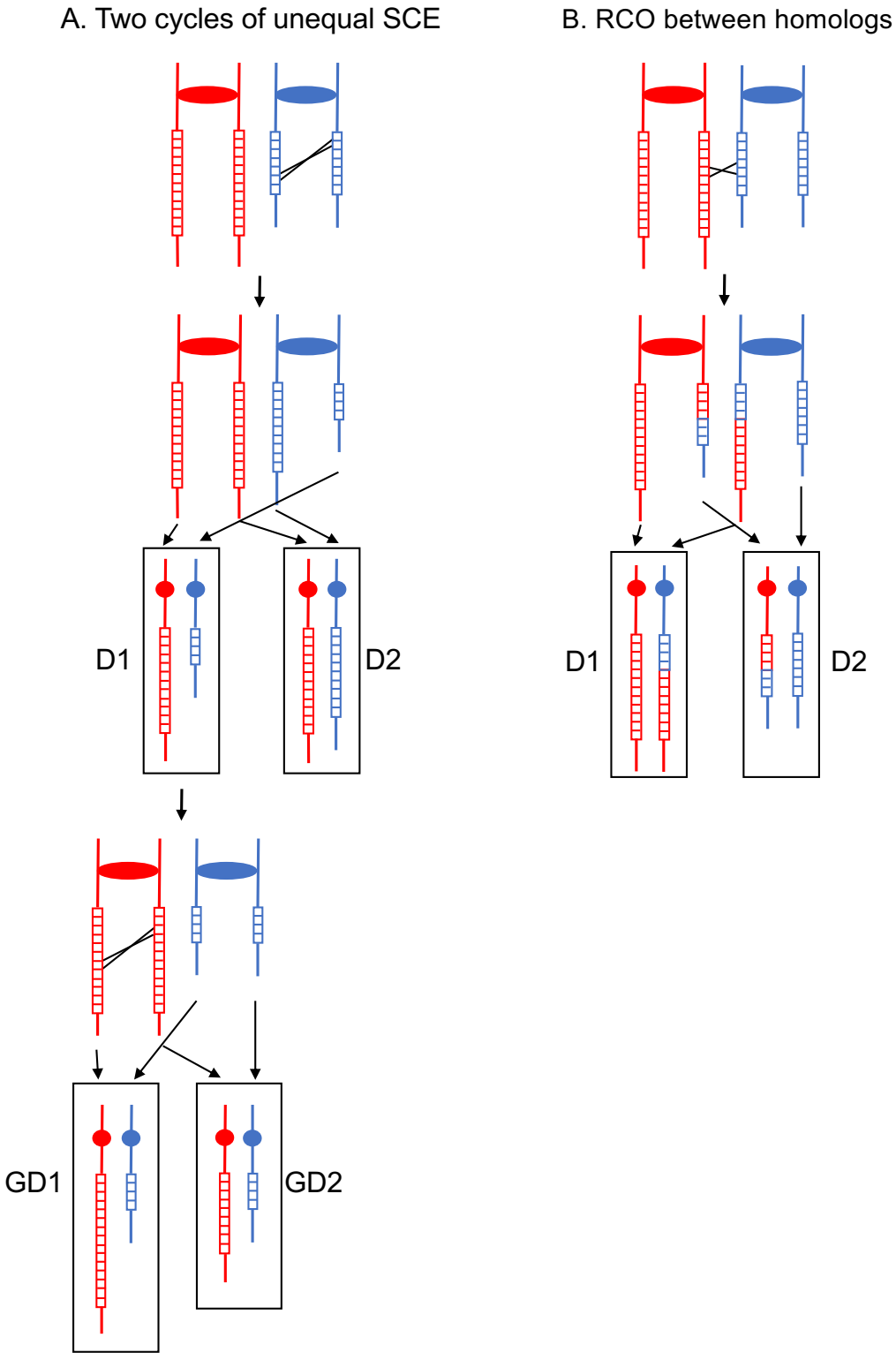

A. Two cycles of unequal SCE
B. RCO between homologs
D1
D2
D1
D2
GD1
GD2

Fig S8. Recombination within the rRNA gene cluster. Each small rectangle indicates 10 rRNA gene repeats. The rRNA gene clusters on the W303- and YJM789-derived clusters have about 120 and 70 repeats, respectively; the W303- and YJM789-derived homologs are shown in red and blue, respectively. As described in the text, many of the sub-cultured isolates had elevated numbers of one type of repeat and reduced numbers of the other. This pattern can be produced by the two mechanisms shown in A and B.

A. Two cycles of unequal sister-chromatid exchange (SCE). Following the first exchange, the number of YJM789-derived repeats is reduced in one daughter cell (D1) and increased in the other (D2). If there is a subsequent unequal SCE in the D1 cell involving the W303-derived cluster, one granddaughter cell (GD1) would have an increased number of W303-derived repeats in addition to the reduced number of YJM789-derived repeats.

B. Reciprocal crossover (RCO) between homologs. A reciprocal crossover between clusters on different homologs would increase the type of one repeat and decrease the other type of repeat. For example, the D1 cell would have 250 copies of the W303-derived repeat and 40 copies of the YJM789-derived repeat.

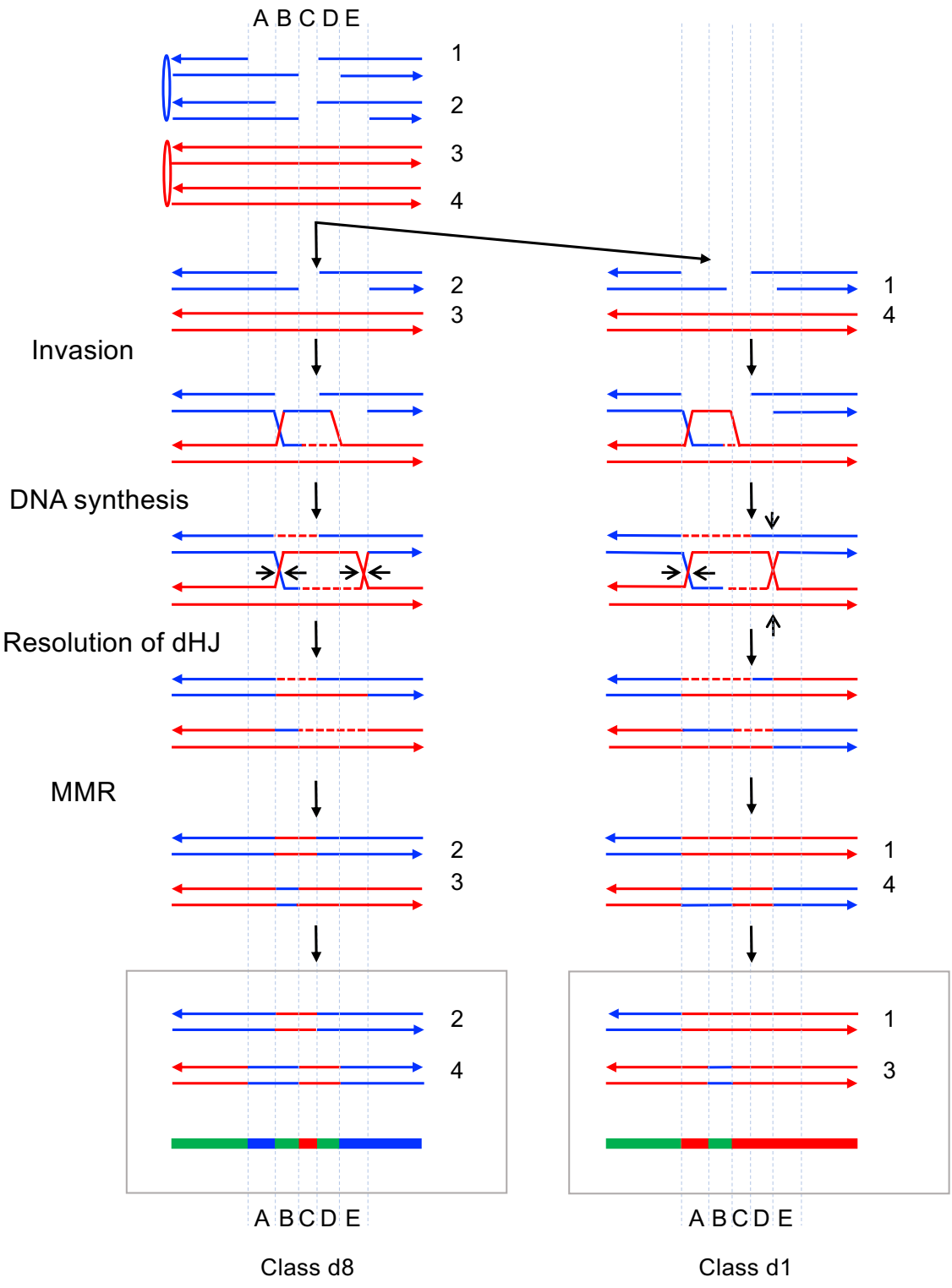

A B C D E
1
2
3
4
Invasion
2
3
1
4
DNA synthesis
Resolution of dHJ
MMR
2
3
1
4
2
4
1
3
A B C D E
A B C D E
Class d8
Class d1

Fig. S9. Possible mechanisms giving rise to two complex conversion events associated with crossovers (Classes d8 and d1 in Dataset S4-2). In this figure, each chromosome is depicted as a double-stranded structure with the 3' ends indicated by arrows. We show both blue chromatids broken as a consequence of a G1-induced break, and subsequent replication of the broken chromosome. Each broken end is associated with a small double-stranded gap that is then processed 5' to 3'. Following strand invasion, double Holliday junctions are formed. The junction shown on the left side of the figure is processed to yield a non-crossover whereas the junction shown on the right side is processed to generate a crossover. Disjunction of chromatids 2 and 4 would form the LOH pattern shown as Class d8 in Dataset 4-2; the disjunction of chromatids 1 and 3 would produce the pattern of LOH shown as Class d1 in Dataset 4-2. As in Dataset 4-2, the green line at the bottom of the figure indicates heterozygosity, and the blue and red lines show LOH for the YJM789- and W303-1A-derived SNPs, respectively.

Table S1. Strains and primers used in the study.

Table S1-1. Strains used in this study

Name	Background	Genotype	Reference or Construction
W1588-4C	W303-1A	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5*	St. Charles and Petes, 2013 (11).
JSC20-1	YJM789	*MATα ade2-1 ura3 gal2 ho::hisG IV1510386::SUP4-o*	St. Charles and Petes, 2013 (11).
Wspo11	W303-1A	*MATa leu2-3,112 his3-11,15 ura3-1 ade2-1 trp1-1 can1-100 RAD5 spo11::KANMX*	Replacement of *SPO11* with PCR fragment amplified from pUG6 (12).
Yspo11	YJM789	*MATα ade2-1 ura3 gal2 ho::hisG IV1510386::SUP4-o spo11::KANMX*	Replacement of *SPO11* with PCR fragment amplified from pUG6 (12).
WYspo11	W303-1AxYJM789	*MATa/α leu2-3,112/LEU his3-11,15/HIS ura3-1/ura3 ade2-1/ade2-1 trp1- 1/TRP can1-100/CAN1 RAD5/RAD5 GAL2/gal2 ho::hisG IV1510386/IV1510386::SUP4-o spo11::KANMX/spo11::KANMX*	Cross of Wspo11 and Yspo11.
Spo11-*	WYspo11	*MATa/α leu2-3,112/LEU his3-11,15/HIS ura3-1/ura3 ade2-1/ade2-1 trp1- 1/TRP can1-100/CAN1 RAD5/RAD5 GAL2/gal2 ho::hisG IV1510386/IV1510386::SUP4-o spo11::KANMX/spo11::KANMX*	Strains of WYspo11 subcultured 120 passages.
Spo60-*	WYspo11	*MATa/α leu2-3,112/LEU his3-11,15/HIS ura3-1/ura3 ade2-1/ade2-1 trp1- 1/TRP can1-100/CAN1 RAD5/RAD5 GAL2/gal2 ho::hisG IV1510386/IV1510386::SUP4-o spo11::KANMX/spo11::KANMX*	Strains of WYspo11 subcultured 60 passages.

Table S1-2. Primers used in this study

Order	Primer	Sequence (5' to 3')	Purpose
1	dspo11 S	ACGATTTACTAAGTTCACCTTCTCATGGCTTTGGA GGGATTGCGGCATCAGAGCAGATT	Amplification of *spo11::KANMX* cassette from pUG6 (12).
	dspo11 A	CGCTGACATTCTATGATGACGTTCTTTAATGGTG AAGTTGCAATACCGCTCGCCGCAG	
2	vspo11S	CTTCATTACAGAAACGGC	Verification of deletion of *SPO11* with *spo11::KANMX* cassette from pUG6 (12).
	vspo11A	GGTTCGATTCTAGGAGATG	
3	**vXIII652256S**	CATCATTGCTGCTTCTCG	Verification of the deletion at chrXIII 652256 in Spo11-1 strain. Bolded primer was used for Sanger sequencing.
	vXIII652256A	CGAGCACGACAGTGAAAT	
4	vVII996350S	GTGAGTTTATTGAGGGTGAT	Verification of the complex variation at chrVII 996350 in Spo11-6 strain. Bolded primer was used for Sanger sequencing.
	vVII996350A	TAATCAAACATTTACGCAGT	
5	**vXII832573S**	TGACTCGTGGCAACAGATT	Verification of the deletion at chrXII 832573 in Spo11-50 strain. Bolded primer was used for Sanger sequencing.
	vXII832573A	TGATTCTCCACCCACCAC	
6	**vXV273942S**	TGACCATAAGCGTCTCCT	Verification of the SNV at chrXV 273942 in Spo11-50 strain. Bolded primer was used for Sanger sequencing.
	vXV273942A	CCCTTCTTACGAAACTGC	
7	**vXV258123S**	AGGAGTAGGAAGAATACGGTG	Verification of the SNV at chrXV 258123 in Spo11-80 strain. Bolded primer was used for Sanger sequencing.
	vXV258123A	CACGCCATCTATTGAAACTC	
8	**vXIV187636S**	ATTCCTCTGGGACACCTTG	Verification of the SNV at chrXIV 187636 in Spo11-90 strain. Bolded primer was used for Sanger sequencing.
	vXIV187636A	TCTCCTCAACGGAACACG	

9	**vVIII490638S**	TTGATGAGTCTCGGTGCT	Verification of the SNV at chrVIII 490638 in Spo11-91 strain. Bolded primer was used for Sanger sequencing.
	vVIII490638A	ATCCATCCCTGTCTTTCC	
10	**vVI147695S**	ATTTACGCAAACTGGTCC	Verification of the SNV at chrVI 147695 in Spo11-103 strain. Bolded primer was used for Sanger sequencing.
	vVI147695A	TTGCGTCAAAGTTAGGTG	
11	**vXVI686686S**	AGTCCGATTCTACCTCCC	Verification of the SNV at chrXVI 686686 in Spo11-111 strain. Bolded primer was used for Sanger sequencing.
	vXVI686686A	TTGGCTCAGTGTTAGATGTT	
12	vIX35185S	CTTTATGAGTCACGGAGGC	Verification of the insertion at chrIX 35185 in Spo11-151 strain. Bolded primer was used for Sanger sequencing.
	vIX35185A	GTGAGGTAGGGCGACAAG	
13	**vXII317188S**	GAATAAGGGATTGGGCGT	Verification of the insertion at chrXII 317188 in Spo11-160 strain. Bolded primer was used for Sanger sequencing.
	vXII317188A	TTGGGTTAGAGGATTGAGTTG	
14	vXII974909S	CGCTATCACCCTATTACTG	Verification of the SNV at chrXII 974909 in Spo11-188 strain. Bolded primer was used for Sanger sequencing.
	vXII974909A	AAGAGTGCTCAAATACCGT	
15	vIII138504S	TGGAGAACCCAACCAGACC	The circular chromosome III in Spo11-81 was verified by inverse PCR using this pair of primers (a-b) and TaKaRa LA Taq® Hot Start Version DNA polymerase (TaKaRa, Cat#: RR042A).
	vIII91002A	TCGGCTGTGATTTCTTGACC	
16	vIII149313S	CGCTTATCCGATTCTGTTGG	The circular chromosome III in Spo11-81 was verified by inverse PCR using this pair of primers (a-c) and TaKaRa LA Taq® Hot Start Version DNA polymerase (TaKaRa, Cat#: RR042A).
	vIII91002A	TCGGCTGTGATTTCTTGACC	

Table S2. LOH events in relation to telomeres and centromeres.

Chromosome element	Window size	Type of LOH event	Obs. events in window	Exp. events in window	Obs. events outside of window	Exp. events outside of window	p value
Centromere	30 kb	CO/BIR	7	15	349	341	0.048
Centromere	30 kb	CON	50	36	809	823	0.022
Telomere	20 kb*	CO/BIR	60	20	296	336	<.0001
Telomere	20 kb*	CON	51	47	808	812	0.5967

*For the centromeres, the region included 15 kb to each side of the centromere. For the telomeres, the region included 20 kb from the SNP that was located closest to the telomere. On the average, there were about 15 kb from the end of the chromosome to the first SNP. For each I-LOH event or T-LOH event, we determined the center of the window from Dataset S4-1, and looked for its inclusion in the centromeric or telomeric regions. The expected numbers of events were calculated by the following equation: (# of kb in the genome occupied by the windows of each element/total genome size) x (the total # of events for each class). The p values were determined by chi-square analysis.

Dataset S1 (separate file). Spontaneous mutations in the wild-type sub-cultured diploid WYspo11.

Dataset S1-1. Single-base mutations: genomic locations and base pair changes.

Dataset S1-2. Numbers and rates of single-base mutations in each isolate.

Dataset S1-3. Small In/Dels (<10 bp) observed in sub-cultured strains.

Dataset S1-4. In/Dels (≥10 bp but <100 bp) between direct repeats.

Dataset S1-5. Complex mutations (multiple changes within 10 bp).

Dataset S2 (separate file). Chromosome rearrangements: large deletions and duplications.

Dataset S2-1. Genomic location of large (>100 bp) interstitial deletions and duplications in sub-cultured strains.

Dataset S2-2. SGD coordinates for transitions for termination deletion and duplication events in sub-cultured strains.

Dataset S2-3. Depictions of various classes of deletion and duplication events.

Dataset S3 (separate file). Copy-number changes in rRNA and *CUP1* genes.

Dataset S4 (separate file). Analysis of LOH events in sub-cultured WYspo11 isolates.

Dataset S4-1. Location of LOH events.

Dataset S4-2. Depiction of various classes of LOH events.

Dataset S4-3. Number of LOH events per isolate.

Dataset S4-4. Associations of LOH events with various elements of chromosome structure.

Dataset S4-5. Numbers of LOH events per chromosome.

Dataset S4-6. Distribution of LOH events among isolates.

Dataset S5 (separate file). Aneuploidy in sub-cultured WYspo11 isolates.

Dataset S5-1. Aneuploid events in sequenced strains.

Dataset S5-2. Depiction of various classes of aneuploids.

Dataset S5-3. The effect of aneuploidy on LOH and single-base mutations.

Dataset S6 (separate file). Frequency of LOH events per SNP among 93 sub-cultured isolates.

Dataset S7 (separate file). LOH events per chromosome arm.

SI References

1. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
2. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
3. D.-Q. Zheng, K. Zhang, X.-C. Wu, P. A. Mieczkowski, T. D. Petes, Global analysis of genomic instability caused by DNA replication stress in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E8114-E8121 (2016).
4. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. *arXiv* 1207.3907 (2012).
5. R. M. Layer, C. Chiang, A. R. Quinlan, I. M. Hall, LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol.* **15**, R84 (2014).
6 P. Cingolani *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80-92 (2012).
7. E. Yim, K. E. O'Connell, J. St Charles, T. D. Petes, High-resolution mapping of two types of spontaneous mitotic gene conversion events in *Saccharomyces cerevisiae*. *Genetics* **198**, 181-192 (2014).
8. A. Kokina, J. Kibilds, J. Liepins, Adenine auxotrophy–be aware: some effects of adenine auxotrophy in *Saccharomyces cerevisiae* strain W303-1A. *FEMS Yeast Res.* **14**, 697-707 (2014).
9. P. M. Magwene *et al.*, Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1987-1992 (2011).
10. Y. O. Zhu, M. L. Siegal, D. W. Hall, D. A. Petrov, Precise estimates of mutation rate and spectrum in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E2310-E2318 (2014).
11. J. S. Charles, T. D. Petes, High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1 Mb arm of yeast chromosome IV. *PLoS Genet.* **9(4)**, e1003434 (2013).
12. U. Gueldener, J. Heinisch, G. Koehler, D. Voss, J. Hegemann, A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**, e23-e23 (2002).