bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The genetic basis of mutation rate variation in yeast

Liangke Gou[1], Joshua S. Bloom[1,2], Leonid Kruglyak[1,2,3*]

[1]Department of Human Genetics, University of California, Los Angeles, Los Angeles, California 90095, USA

[2]Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, California 90095, USA

[3]Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, California 90095, USA

* Corresponding author

E-mail: lkruglyak@mednet.ucla.edu

Abstract

Mutations are the root source of genetic variation and underlie the process of evolution. Although the rates at which mutations occur vary considerably between species, little is known about differences within species, or the genetic and molecular basis of these differences. Here we leveraged the power of the yeast *Saccharomyces cerevisiae* as a model system to uncover natural genetic variants that underlie variation in mutation rate. We developed a high-throughput fluctuation assay and used it to quantify mutation rates in natural yeast isolates and in 1008 segregant progeny from a cross between BY, a lab strain, and RM, a wine strain. We observed that mutation rate varies among yeast strains and is highly heritable (H^2=0.46). We

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

performed linkage mapping in the segregants and identified four quantitative trait loci (QTLs) underlying mutation rate variation in the cross. We fine-mapped two QTLs to the underlying causal genes, *RAD5* and *MKT1*, that contribute to mutation rate variation. These genes also underlie sensitivity to the DNA damaging agents 4NQO and MMS, suggesting a connection between spontaneous mutation rate and mutagen sensitivity.

Author Summary

Spontaneous mutation rate varies between species, as well as between individuals within species. The genetic basis for mutation rate variation within species is poorly understood. Part of the challenge is accurately measuring mutation rates in many individuals. We addressed this challenge by developing a high-throughput fluctuation assay, and we used this assay to identify and genetically dissect differences in mutation rate between yeast strains. To do so, we measured mutation rates in 1008 segregant progeny from a cross between a laboratory strain and a vineyard strain and used linkage analysis to map four genetic loci underlying the mutation rate variation in this cross. We then identified the genes and variants that underlie the two loci with largest contributions to mutation rate variation. These genes also affect sensitivity to DNA damaging agents, suggesting a connection between spontaneous mutation rate and mutagen sensitivity.

Introduction

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mutations are permanent changes to the genome of an organism that can result from DNA damage that is improperly repaired, from errors in DNA replication, or from the movement of mobile genetic elements. Mutations give rise to genetic variants in populations and are the wellspring of evolution. Mutations also play a major role in both inherited diseases and acquired diseases such as cancer.

The mutation rate can be defined as the number of mutational events per cell division, generation, or unit of time [1]. Mutation rates tends to be approximately 10^{-9} to 10^{-10} mutations per base pair, per cell division, for most microbial species [2], making them difficult to measure and compare across individuals. As a consequence, the effects of genetic background differences on mutation rates have only been investigated on a small scale [3]. Two types of experimental approaches have been used to measure mutation rates in yeast. The first is the fluctuation assay [4]. This method requires a gene with a selectable phenotype such that loss-of-function mutations in the gene enable the mutants to grow in the corresponding selective conditions. Spontaneous mutation rate is then estimated from the distribution of mutant numbers in parallel cultures. Lang and Murray applied the fluctuation assay to *S. cerevisiae* and estimated the per-base-pair mutation rate in yeast [5]. A second method tracks mutation accumulation during experimental evolution and uses whole-genome sequencing to estimate mutation rates [6]. This approach also provides information on the number, locations and types of spontaneous mutations. However, this assay requires growing the mutation accumulation lines over hundreds of generations, as well as sequencing many genomes. Although the fluctuation assay is faster and cheaper, the need for many parallel cultures makes it laborious to extend it to many different strains.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Here we developed a modified version of the fluctuation assay to enable higher-throughput measurements of spontaneous mutation rates. We used the new assay to quantify mutation rates across genetically distinct yeast strains and observed considerable variation. To find the genes underlying the observed variation, we applied the modified fluctuation assay to a large panel of 1,008 segregants from a cross between the laboratory strain BY4724 (hereafter referred to as BY) and the vineyard strain RM11-1a (hereafter referred to as RM). We identified four loci associated with mutation rate variation and narrowed the two loci that contributed the most to mutation rate variation to missense variants in the genes *RAD5* and *MKT1*. We also found interactions between alleles of *RAD5* and *MKT1*.

Results

High-throughput fluctuation assay for measuring mutation rates

The fluctuation assay for measuring mutation rate involves growing many parallel cultures, each starting from a small number of cells, under non-selective conditions, followed by plating to selective medium to identify mutants. The number of mutations that occurs in each culture should follow the Poisson distribution, as mutations arise spontaneously. However, the number of mutant cells that survive on the selective plates can vary greatly, because early mutations are inherited by all offspring of the mutant. This leads to the "jackpot" effect, in which some cultures contain a large number of mutant individuals. The number of observed mutant cells per culture follows the Luria-Delbrück distribution [4], and the Ma-Sandri-Sarkar maximum likelihood method can be used to estimate the expected number of mutations per culture from the observed numbers of mutants [7]. The underlying mutation rate is

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

then calculated by dividing the number of mutations per culture by the average number of cells per culture [1,4]. Here we measured rare spontaneous loss-of-function mutations in the gene *CAN1*, which encodes an arginine permease. Yeast cells carrying loss-of-function mutations in *CAN1* can grow on canavanine, an otherwise toxic arginine analog. Typically, fluctuation assays are labor-intensive and have limited throughput, because a large number of parallel cultures is required for estimating the mutation rate in each assay, and several replicate assays are needed for a robust measurement of the mutation rate in each strain [8]. We modified the fluctuation assay into a high-throughput method for measuring mutation rates in many strains in parallel. We grew cultures in 96-well plates, automated the spotting of cultures, and used high-resolution imaging to rapidly count mutants on many plates (Methods, Fig 1A). The automated spotting process for 96 strains took only approximately twenty minutes, and the imaging process required even less time. These improvements enabled us to measure the spontaneous mutation rates in the hundreds of strains necessary for genetic mapping.

Spontaneous mutation rate varies among yeast isolates

To investigate mutation rate variation among *S. cerevisiae* strains, we measured the spontaneous mutation rate of seven yeast isolates using the high-throughput fluctuation assay (S1 Table). The seven strains span a large range of yeast genetic diversity [9]. We found that the mutation rates of these strains range from 1.1×10^{-7} to 5.8×10^{-7} mutations per gene per generation, with a median of 1.7×10^{-7} (S1 Table, S1 Fig). The median mutation rate was very similar to the previously reported mutation rate at *CAN1* [5]. In particular, the mutation rate we observed for the BY strain

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

125 (1.7×10^{-7}) is very similar to the previously reported rate, which was measured in
126 strain W303 (1.5×10^{-7}) [5], consistent with the fact that W303 shares a large fraction
127 of its genome with BY [10]. An analysis of variance (ANOVA) showed that strain
128 identity explained a significant fraction of the observed variance in mutation rates
129 ($F=69.9$, $df=6$, $p<2\times10^{-16}$) (S1 Fig). The fraction of total variance in mutation rates
130 explained by the repeatability of measurements for each strain, 46%, serves as an
131 upper bound for the estimate of the total contribution of genetic differences between
132 strains to trait variation (broad-sense heritability or H^2). We observed that RM, a
133 vineyard strain, had a mutation rate higher than all other strains (S1 Fig).
134
135 **Four QTLs explain the majority of observed mutation rate variation**
136
137 In order to find the genetic factors underlying the difference in mutation rate between
138 BY and RM, we performed quantitative trait locus (QTL) mapping in 1,008
139 genotyped haploid segregants from a cross between these strains [11]. We measured
140 the mutation rate of each segregant using the high-throughput fluctuation assay
141 (Methods). We estimated the fraction of phenotypic variance explained by the
142 additive effects of all segregating markers (narrow-sense heritability) to be 30%
143 (Methods) [12]. This sets an upper bound for the expectation of the total amount of
144 additive genetic variance that could be explained with a QTL-based model. QTL
145 mapping in the segregant panel identified significant linkage at four distinct loci (Fig
146 1B). At two of the QTLs, on chromosomes XII and V, the RM allele conferred a
147 higher mutation rate, consistent with the higher mutation rate of this strain. At the
148 other two QTLs, on chromosomes XIV and I, the BY allele conferred a higher
149 mutation rate (S2 Fig), showing that a strain with lower trait value can nevertheless

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

harbor trait-increasing alleles. The four detected QTLs explained 20.7% of the phenotypic variance, thus accounting for 69% of the estimated additive heritability. The loci on chromosomes XII, XIV, I and V explained 8.8%, 6.1%, 3.1% and 2.6% of the variance, respectively. We tested the four identified QTLs for pairwise interactions and found a significant interaction between the QTL on chromosome XII and the QTL on chromosome XIV that explained 1% of the phenotypic variance ($F=8.41$, df=1, Bonferroni-corrected $p=0.023$).

Polymorphisms in genes *RAD5* and *MKT1* underlie the major QTLs on chromosomes XII and XIV

Ten genes fell within the confidence interval of the QTL on chromosome XII. A strong candidate was *RAD5*. Previous studies showed that natural variants in *RAD5* contribute to sensitivity to the mutagen 4-nitroquinoline 1-oxide (4NQO) [13]. *RAD5* encodes a DNA repair protein involved in the error-free DNA damage tolerance (DDT) pathway [14,15]. The DDT pathway promotes the bypass of single-stranded DNA lesions encountered by DNA polymerases during DNA replication, thus preventing the stalling of DNA replication [16]. *RAD5* plays a crucial role in one branch of the DDT pathway called template switching (TS), in which the stalled nascent strand switches from the damaged template to the undamaged newly synthesized sister strand for extension past the lesion [16]. Two non-synonymous substitutions exist between BY and RM strains in *RAD5* (Fig 2A), at amino acid positions 783 (glutamic acid in BY and aspartic acid in RM) and 791 (isoleucine in BY and serine in RM). According to Pfam alignments [17], *RAD5* contains a HIRAN domain, an SNF2-related N-terminal domain, a RING-type zinc finger domain, and a

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

helicase C-terminal domain (Fig 2A). Both non-synonymous polymorphisms mapped to the helicase domain of *RAD5* (Fig 2A), and no other sequenced strains of *S. cerevisiae* contain the aspartic acid 783 and serine 791 variants that are private to the RM strain. We used protein variation effect analyzer (PROVEAN) [18] to predict whether the two non-synonymous substitutions have an impact on the biological function of the protein. PROVEAN showed the I791S substitution (score –5.4) might have a strong deleterious effect, while the E783D variant (score –1.8) was not predicted to have a strong effect.

Nineteen genes fell within the confidence interval of the QTL on chromosome XIV. A strong candidate was *MKT1*, which was also reported to affect 4NQO sensitivity [13]. *MKT1* encodes an RNA-binding protein that affects multiple traits and underlies an eQTL hotspot in yeast [19]. The RM allele of *MKT1* increases sporulation rate [20] and improves survival at high temperature [21], in low glucose [22], after exposure to DNA-damaging agents [13], and in high ethanol levels [23]. The coding region of the BY and RM alleles of *MKT1* differs by one synonymous polymorphism and two non-synonymous substitutions. *MKT1* has an XPG domain, which is relevant to DNA repair, and an *MKT1* domain, which is related to the maintenance of K2 killer toxin [24]. One non-synonymous variant is in the XPG domain at amino acid position 30 (aspartic acid in BY and glycine in RM), while the other non-synonymous variant is in the *MKT1* domain at position 453 (lysine in BY and arginine in RM). PROVEAN predicted a large effect of the D30G variant (score 6.7) on the function of *MKT1*, and this variant was previously found to influence sporulation rate [20], mitochondrial genome stability [25] and survival at high temperature [22]. The other variant (K453R) was not predicted to have a strong effect (score 0.6).

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We tested whether *RAD5* and *MKT1* alleles caused differences in mutation rate by using the fluctuation test on allele replacement strains [13,26] (Table 1). The BY strain carrying the RM allele of *RAD5* (BY::*RAD5*-RM) had a higher mutation rate than the BY strain (permutation t-test, mean difference=$2.9x10^{-7}$, $p<1x10^{-4}$), demonstrating that the RM *RAD5* allele increases mutation rate (Fig 3A). This result is consistent with the observed difference between segregants grouped by parental allele at *RAD5* (mean difference=$2.3x10^{-7}$). The RM strain carrying the BY allele of *MKT1* (RM::*MKT1*-BY) had a higher mutation rate than the RM strain (permutation t-test, mean difference=$6.1x10^{-7}$, $p<1x10^{-4}$), showing that the BY *MKT1* allele increases mutation rate (Fig 3A), consistent with the direction of effect observed in the segregants.

Table 1. The allele replacement strains and variant substitution strains

Strain	Background	Relevant Genotype	Source
YLK802	RM	*MATa*, *MKT1*-BY, hoΔ::HphMX, flo8Δ::NatMX	Smith *et al*., 2008
EAY1463	BY	*MATα*, *lys2*Δ, *RAD5*-RM::NatMX	Demogines *et al*., 2008
EAY1471	BY	*MATα*, *lys2*Δ, *RAD5*-I791S::KanMX	Demogines *et al*., 2008
EAY2169	BY	*MATα*, *lys2*Δ, *RAD5*-E783D::KanMX	Demogines *et al*., 2008

To gain a finer-level understanding of the two missense variants between BY and RM in the gene *RAD5*, we tested strains [13] in which these sites in BY were individually replaced with the RM alleles (Table 1) by site-directed mutagenesis. Strains with either variant had a higher mutation rate than BY (permutation t-test, mean difference=$0.9x10^{-7}$, $p<1x10^{-4}$ for BY::*RAD5*-I791S; mean difference=$0.3x10^{-7}$, $p=6x10^{-4}$ for BY::*RAD5*-E783D) (Fig 2B), suggesting that both variants contribute to the higher mutation rate. The BY strain with the I791S substitution had a higher

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mutation rate than the BY strain with the E783D substitution (permutation t-test, mean difference=$0.6x10^{-7}$, $p<1x10^{-4}$) (Fig 2B), consistent with the PROVEAN prediction of a stronger effect for the I791S variant. However, neither variant alone nor the additive effect of the two variants fully recapitulated the increase in mutation rate that we observed when replacing the entire coding region of *RAD5* in BY with the RM allele ($F=67.6$, $df=1$, $p=3.3x10^{-15}$), suggesting an interaction between the two variants.

Mutation rate shares two large effect QTLs with growth on DNA damaging agents 4NQO and MMS

Deficiencies in DNA repair can increase mutation rate [27,28] and increase sensitivity to DNA damaging agents such as alkylating compounds and UV light [29,30]. We hypothesized that genetic variants that cause deficiencies in DNA repair may underlie QTLs for both mutation rate variation and sensitivity to DNA damaging agents. Previously, Demogines *et al*. identified a large-effect QTL on chromosome XII for MMS and 4NQO sensitivity in a panel of 123 segregants from a cross between BY and RM [13]. Additionally, they identified a QTL on chromosome XIV for 4NQO sensitivity by using backcrossing and bulk segregant analysis. These QTLs overlapped with the major QTLs that we identified for mutation rate variation, and the underlying causal genes for 4NQO sensitivity were also *RAD5* and *MKT1*.

To follow up on these results, we measured sensitivity to three different DNA damaging agents in our panel of 1008 segregants (Table 2). The compounds assayed included methyl methanesulfonate (MMS), an alkylating agent that induces DNA

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

double strand breaks and stalls replication forks [31], 4NQO, an ultraviolet light mimetic agent [31] and hydrogen peroxide (H_2O_2), a compound that induces DNA single and double strand breaks [31]. We observed that segregants with higher mutation rate, and presumably less efficient DNA repair systems, were more sensitive to MMS, 4NQO and H_2O_2 (S3 Fig), consistent with our hypothesis that deficiencies in DNA repair increase the rate of spontaneous mutations and increase sensitivity to DNA damaging agents. We identified two large-effect QTLs for 4NQO and MMS sensitivity that overlapped with the major QTLs for mutation rate (Fig 4A and B). An interaction between *RAD5* and *MKT1* was observed for 4NQO sensitivity (F=8.5, df=1, p=0.004) (S4 Fig). The QTLs on chromosome 12 and 14 were still observed in the linkage mapping for H_2O_2, but they had small effects (S5 Fig). The large effect QTLs detected for H_2O_2 sensitivity on other chromosomes likely reflects trait-specific effects of variants acting on sensitivity to H_2O_2 (S5 Fig).

Table 2. DNA damaging agents used for the sensitivity assay

Agent	Agent characteristic
Hydrogen peroxide (H_2O_2)	Altering DNA structure
Methyl methane sulfonate (MMS)	Altering (alkylating) DNA bases
4-nitroquinoline 1-oxide (4NQO)	UV mimetic

Discussion

We developed and implemented a high-throughput fluctuation assay to directly measure mutation rates in yeast. We used this assay to map four QTLs that influence differences in the spontaneous mutation rate.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We identified *RAD5* as the gene underlying the QTL with the largest effect on mutation rate. *RAD5* encodes a DNA helicase and ubiquitin ligase involved in error-free DNA damage tolerance (DDT), a pathway that facilitates chromosome replication through DNA lesions [32,33]. Previous work showed that Rad5 is a structure-specific DNA helicase that is able to carry out replication fork regression [14], a process of remodeling the replication fork into four-way junctions when replication perturbations are encountered [34]. This process was hypothesized to promote DNA damage tolerance and repair during replication [34]. We showed that two non-synonymous variants between BY and RM in the helicase domain affect mutation rate. The RM allele of *RAD5* increases the sensitivity of yeast to 4NQO and MMS [35], probably due to a defect in replication fork regression. Thus the RM allele of *RAD5* causes both decreased growth in mutagenic conditions and a higher mutation rate in non-stressful normal conditions.

We furthermore showed that polymorphisms in *MKT1* contribute to mutation rate variation. *MKT1* is a highly pleiotropic gene that has been shown to affect levels of transcript and protein abundance for numerous genes [26] [36], as well as numerous cellular phenotypes [13,19–23,37]. The BY and RM alleles of *MKT1* differ by two non-synonymous substitutions that map to amino acid positions 30 (aspartic acid in BY; glycine in RM) and 453 (lysine in BY; arginine in RM). The latter variant (K453R) is located in the MKT1 domain, which is required for activity of the Mkt1 protein in maintaining K2 killer toxin [38]. The former variant (D30G) localizes to the XPG-N (the N-terminus of XPG) domain. Four other yeast proteins contain this domain: Exo1, Din7, Rad27 and Rad2. All of these proteins have functions related to DNA repair and cellular response to DNA damage, including DNA double-strand

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

296 break repair (Exo1) [39], DNA mismatch repair (Exo1, Din7) [40,41], nucleotide
297 excision repair (Rad2) [42], ribonucleotide excision repair (Rad27) [43] and large
298 loop repair (LLR) (Rad27) [44]. The internal XPG (XPG-I) domain, together with
299 XPG-N, forms the catalytic domain of the Xeroderma Pigmentosum
300 Complementation Group G (XPG) protein. The XPG protein has well-established
301 catalytic and structural roles in nucleotide excision repair, a DNA repair process, and
302 acts as a cofactor for a DNA glycosylase that removes oxidized pyrimidines from
303 DNA [45]. In humans, mutations in the XPG protein commonly cause Xeroderma
304 Pigmentosum, which often leads to skin cancer [46]. The aspartic acid at position 30
305 in the XPG domain of Mkt1 is only found in BY and related laboratory strains. We
306 hypothesize that Mkt1 has a previously unknown function in DNA damage repair,
307 mediated through its XPG domain.
308
309 We found that variants in *RAD5* and *MKT1* contribute to both mutation rate variation
310 and mutagen sensitivity. These results suggest that spontaneously occurring mutations
311 may have a similar mutation spectrum to those created by 4NQO and MMS, and are
312 potentially repaired by the same mechanisms. Deficient DNA repair can lead to
313 increased sensitivity to agents such as alkylating compounds and UV light [29,30,47]
314 and to higher mutation rates at sites that are less accessible to the DNA repair system
315 [27]. Because mutation rates can be difficult to measure, sensitivity to mutagens may
316 serve as a useful proxy.
317
318 Recently, Jerison et al. reported heritable differences in adaptability in 230 yeast
319 segregants from the same cross we studied here [48]. They measured adaptability as
320 the difference in fitness between a given segregant (‘founder’) and a descendant of

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

that founder after 500 generations of experimental evolution. Interestingly, *RAD5* fell within one of the QTLs found to influence adaptability. Together with our observation that *RAD5* influences mutation rate, this finding suggests that differences in mutation rate can affect the adaptability of organisms.

Materials and Methods

Yeast strains and media

Seven natural *S. cerevisiae* strains (S1 Table) were used in this study. The 1008 segregants derived from BY4724 (MATa) and RM11-1a (MATa, *MKT1*-BY, hoΔ::HphMX, flo8Δ::NatMX) were generated, genotyped and described previously [11]. The RM::*MKT1*-BY strain was made previously by our lab. The BY::*RAD5*-RM strain and the *RAD5* variants substitution strains (Table 1) were from Demogines et al [13]. For fluctuation assay, yeast was grown in synthetic complete liquid medium without arginine (SC-Arg) before plating onto selective plates. For DNA damaging agents sensitivity assays, yeast were grown in rich YPD medium (1% yeast extract, 2% peptone and 2% glucose) before plating onto YPD agar plates with DNA damaging agents. SC-Arg and YPD liquid media and agar plates were made according to Amberg et al [49].

Selection agar plate construction

Selective canavanine plates were made from arginine minus synthetic complete agar medium with 60mg/liter L-canavanine (Sigma C1625). The canavanine plates were dried by incubating at 30ºC overnight. Selective plates for the DNA damaging agents sensitivity assay were made with YPD agar medium containing the respective agents at the concentrations indicated in Table 2. 50ml of the agar medium was poured into

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

each Nunc OmniTray plates (Thermo Scientific 264728) and placed on a flat surface to solidify. Each experiment was performed with the same batch of selection plates. The concentrations for 4NQO (Sigma N8141), MMS (Sigma 64382) and H_2O_2 (Sigma 216763) were 0.1μg/ml, 0.01% and 4mM. These concentrations capture the sensitivity difference between the segregants, while maintaining enough colony growth for QTL mapping.

Fluctuation assays

To begin the fluctuation assay, yeast were grown in synthetic complete medium without arginine (SC-Arg) in 96-well plates (Costar 3370) for ~48 hours to saturation. Saturated cultures were diluted and pinned into a new 96-well plate with liquid SC-Arg medium. This step ensured a small number of ~1000 yeast cells in the initial inoculum. Plates were sealed with a Breathe-Easy sealing membrane (Sigma Z380059) to prevent evaporation, and incubated at 30ºC with shaking for ~48 hours. 100μl saturated cultures were spot-plated onto canavanine plates in a four by six configuration using a Biomek FX^P automated workstation. Plates with spot-plated yeast culture were dried in the laminar flow hood (Nuair) for half an hour or until dry, and incubated at 30ºC for ~48 hours. We imaged the plates using an imaging robot (S&P Robotics BM3-SC), and the number of colonies in each spot was manually counted from the images.

For each of the seven natural isolate strains, we performed ninety-six replicates of the fluctuation assay. In each replicate three cultures were plated onto canavanine plates to estimate the mutation events per culture. One culture was diluted and plated onto YPD to determine the number of cells per culture in each replicate. For the panel of

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

BYxRM segregants twelve cultures per segregant were plated onto canavanine plates to calculate the number of mutations per culture, and one culture was used to determine the number of cells. For each allele replacement strain (Table 1), ninety-six replicates of fluctuation analysis were performed. In each replicate, twelve cultures were plated onto canavanine plate to estimate the number of mutations per culture, and three cultures were pooled, diluted and plated to determine the number of cells per culture.

Analysis of fluctuation analysis data

Mutation rate was estimated using the Ma-Sandri-Sarkar Maximum Likelihood Method where the number of observed colonies on canavanine plate was fitted into the Luria-Delbrück distribution on the basis of a single parameter m [7]. The parameter m represents the expected number of mutation events per culture. For the natural isolates and engineered strains, the mutation rate was calculated from the equation $\mu = m/N$, where N is the average number of cells per culture (as a proxy for the number of cell divisions given the starting inoculum is much smaller than N). In the segregant panel, mutation rate was calculated as the residual phenotype after regressing out the effect of average number of cells per strain from the estimate of m per strain across all of the segregants.

Yeast growth measurement for DNA damaging agents sensitivity assay

The segregant panel were originally stored in 96-well plates (Costar 3370). During the DNA damaging agents sensitivity assay, individual segregants were inoculated in two plate configurations in 384-well plates (Thermo Scientific 264574) with YPD and grown for ~48 hours in a 30ºC incubator without shaking. Saturated cultures were

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mixed for 1min at 2,000 r.p.m. using a MixMate (Eppendorf) before pinning. The colony handling robot (S&P Robotics BM3-SC) was used to pin segregants onto selective agar plates with 384 long pins. The plates were incubated at 30ºC for ~48 hours and imaged by the colony handling robot (S&P Robotics BM3-SC). Custom R code [11] was used to determine the size of each colony and the size was used as a proxy for growth in the presence of the DNA damaging agents.

QTL mapping

In order to control for intrinsic growth rate differences and plate position effects, we normalized the traits for growth by fitting a regression for growth of the yeast that were in the same layout configuraton on control plate (YPD agar plates for mutagen sensitivity assay). Residuals were used for QTL mapping. We tested for linkage by calculating logarithm likelihood ratio (LOD scores) for each genotypic marker and trait as $-n(\ln(1-r^2)/(2\ln(10))$, where r is the Pearson correlation coefficient between the segregant genotypes and the segregant mutation rate or DNA damaging agents sensitivity. The threshold declaring the significant QTL effect was calculated from the empirical null distribution of the maximum LOD score determined from 1,000 permutations [50]. The estimated 5% family-wise error rate significance thresholds were 3.52, 3.62, 3.61 and 3.64 for mutation rate, mutagen sensitivity for 4NQO, MMS and H_2O_2 respectively. The 95% confidence intervals were determined using a 1.5 LOD score drop.

Acknowledgments

We are grateful to members of the Kruglyak lab for insightful comments on this manuscript and suggestions for experiments and data analyses. We thank Meru Sadhu

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for helpful discussion. We would like to especially thank the Alani lab in Cornell University for the *RAD5* allele replacement and variants substitution strains.

Author Contributions

Conceived and designed the experiments: LG JSB LK. Performed the experiments: LG. Analyzed the data: LG JSB. Wrote the paper: LG JSB LK.

References

1. Baer CF, Miyamoto MM, Denver DR. Mutation rate variation in multicellular eukaryotes: causes and consequences. Nat Rev Genet. 2007;8: 619–631. doi:10.1038/nrg2158
2. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. Genetics. 1998;148: 1667–1686. doi:citeulike-article-id:610966
3. Demerec M. Frequency of spontaneous mutations in certain stocks of drosophila melanogaster. Genetics. 1937;22: 469–478.
4. Luria S, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance. Genetics. 1943;28: 491–511. doi:10.1038/nature10260
5. Lang GI, Murray AW. Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. Genetics. 2008;178: 67–82. doi:10.1534/genetics.107.071506
6. Zhu YO, Siegal ML, Hall DW, Petrov D a. Precise estimates of mutation rate and spectrum in yeast. Proc Natl Acad Sci U S A. 2014; 1–9. doi:10.1073/pnas.1323011111
7. Sarkar S, Ma WT, Sandri GH. On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. Genetica. 1992;85: 173–179. doi:10.1007/BF00120324
8. Lang GI, Murray AW. Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. Genetics. 2008;178: 67–82. doi:10.1534/genetics.107.071506
9. Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. Comprehensive polymorphism survey elucidates population structure of Saccharomyces cerevisiae. Nature. 2009;458: 342–345. doi:10.1038/nature07670
10. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, et al. Population genomics of domestic and wild yeasts. Nature. 2009;458: 337–341. doi:10.1038/nature07743
11. Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. Finding the sources of missing heritability in a yeast cross. Nature. Nature Publishing Group; 2013;494: 234–237. doi:10.1038/nature11867
12. Lynch M, Walsh B. Genetics and analysis of quantitative traits. Genetics and

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Analysis of Quantitative Traits. 1998. p. 980. doi:10.1086/318209

13. Demogines A, Smith E, Kruglyak L, Alani E. Identification and dissection of a complex DNA repair sensitivity phenotype in baker's yeast. PLoS Genet. 2008;4. doi:10.1371/journal.pgen.1000123

14. Blastyák A, Pintér L, Unk I, Prakash L, Prakash S, Haracska L. Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell. 2007;28: 167–175. doi:10.1016/j.molcel.2007.07.030

15. Torres-Ramos CA, Prakash S, Prakash L. Requirement of RAD5 and MMS2 for postreplication repair of UV-damaged DNA in Saccharomyces cerevisiae. Mol Cell Biol. 2002;22: 2419–2426. doi:10.1128/MCB.2419

16. Bi X. Mechanism of DNA damage tolerance. World J Biol Chem. 2015;6: 48. doi:10.4331/wjbc.v6.i3.48

17. Sonnhammer ELL, Eddy SR, Durbin R. Pfam: A comprehensive database of protein domain families based on seed alignments. Proteins Struct Funct Genet. 1997;28: 405–420. doi:10.1002/(SICI)1097-0134(199707)28:3<405::AID-PROT10>3.0.CO;2-L

18. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. Oxford University Press; 2015;31: 2745–7. doi:10.1093/bioinformatics/btv195

19. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat Rev Genet. 2015;16: 197–212. doi:10.1038/nrg3891

20. Deutschbauer AM, Davis RW. Quantitative trait loci mapped to single-nucleotide resolution in yeast. Nat Genet. Nature Publishing Group; 2005;37: 1333–1340. doi:10.1038/ng1674

21. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ, McCusker JH, et al. Dissecting the architecture of a quantitative trait locus in yeast. Nature. 2002. pp. 326–330. doi:10.1038/416326a

22. Parreiras LS, Kohn LM, Anderson JB. Cellular Effects and Epistasis among Three Determinants of Adaptation in Experimental Populations of Saccharomyces cerevisiae. Eukaryot Cell. 2011;10: 1348–1356. doi:10.1128/EC.05083-11

23. Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann G, Yang Y, et al. Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. Genome Res. Cold Spring Harbor Laboratory Press; 2012;22: 975–84. doi:10.1101/gr.131698.111

24. Wickner RB. Plasmids controlling exclusion of the K2 killer double-stranded RNA plasmid of yeast. Cell. 1980;21: 217–226. doi:10.1016/0092-8674(80)90129-4

25. Dimitrov LN, Brem RB, Kruglyak L, Gottschling DE. Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of Saccharomyces cerevisiae S288C strains. Genetics. 2009;183: 365–383. doi:10.1534/genetics.109.104497

26. Smith EN, Kruglyak L. Gene-environment interaction in yeast gene expression. PLoS Biol. 2008;6: 810–824. doi:10.1371/journal.pbio.0060083

27. Sabarinathan R, Mularoni L, Deu-Pons J, Gonzalez-Perez A, Lopez-Bigas N. Nucleotide excision repair is impaired by binding of transcription factors to DNA. Nature. 2016;532: 264–267. doi:10.1038/nature17661

28. Supek F, Lehner B. Differential DNA mismatch repair underlies mutation rate

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

variation across the human genome. Nature. Nature Research; 2015;521: 81–84. doi:10.1038/nature14173

29. Frankfurt OS. Inhibition of DNA repair and the enhancement of cytotoxicity of alkylating agents. Int J cancer. 1991;48: 916–23. Available: http://www.ncbi.nlm.nih.gov/pubmed/1907257

30. Sun Y, Moses RE. Reactivation of psoralen-reacted plasmid DNA in Fanconi anemia, xeroderma pigmentosum, and normal human fibroblast cells. Somat Cell Mol Genet. 1991;17: 229–238. Available: http://www.ncbi.nlm.nih.gov/pubmed/2047939

31. Hampsey M. A review of phenotypes in Saccharomyces cerevisiae. Yeast. 1997. pp. 1099–1133. doi:10.1002/(SICI)1097-0061(19970930)13:12<1099::AID-YEA177>3.0.CO;2-7

32. Unk I, Hajdú I, Blastyák A, Haracska L. Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair. 2010. pp. 257–267. doi:10.1016/j.dnarep.2009.12.013

33. Hishida T, Kubota Y, Carr AM, Iwasaki H. RAD6-RAD18-RAD5-pathway-dependent tolerance to chronic low-dose ultraviolet light. Nature. 2009;457: 612–615. doi:10.1038/nature07580

34. Neelsen KJ, Lopes M. Replication fork reversal in eukaryotes: from dead end to dynamic response. Nat Rev Mol Cell Biol. 2015;16: 207–220. doi:10.1038/nrm3935

35. Demogines A, Wong A, Aquadro C, Alani E. Incompatibilities involving yeast mismatch repair genes: a role for genetic modifiers and implications for disease penetrance and variation in genomic mutation rates. PLoS Genet. 2008;4: e1000103. doi:10.1371/journal.pgen.1000103

36. Foss EJ, Radulovic D, Shaffer SA, Goodlett DR, Kruglyak L, Bedalov A. Genetic variation shapes protein networks mainly through non-transcriptional mechanisms. PLoS Biol. 2011;9. doi:10.1371/journal.pbio.1001144

37. Wang X, Kruglyak L. Genetic basis of haloperidol resistance in Saccharomyces cerevisiae is complex and dose dependent. PLoS Genet. 2014;10. doi:10.1371/journal.pgen.1004894

38. Vermut M, Widner WR, Dinman JD, Wickner RB. XIV. Yeast sequencing reports. Sequence of MKT1, needed for propagation of M2 satellite dsRNA of the L□A virus of Saccharomyces cerevisiae. Yeast. 1994;10: 1477–1479. doi:10.1002/yea.320101111

39. Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, Jaschob D, et al. Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. Nat Cell Biol. 2012;14: 966–976. doi:10.1038/ncb2549

40. Tran PT, Fey JP, Erdeniz N, Gellon L, Boiteux S, Liskay RM. A mutation in EXO1 defines separable roles in DNA mismatch repair and post-replication repair. DNA Repair (Amst). 2007;6: 1572–1583. doi:10.1016/j.dnarep.2007.05.004

41. Koprowski P, Fikus MU, Dzierzbicki P, Mieczkowski P, Lazowska J, Ciesla Z. Enhanced expression of the DNA damage-inducible gene DIN7 results in increased mutagenesis of mitochondrial DNA in Saccharomyces cerevisiae. Mol Genet Genomics. 2003;269: 632–639. doi:10.1007/s00438-003-0873-8

42. Northam MR, Robinson HA, Kochenova O V., Shcherbakova P V. Participation of DNA polymerase in replication of undamaged DNA in Saccharomyces cerevisiae. Genetics. 2010;184: 27–42.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

doi:10.1534/genetics.109.107482

43. Sparks JL, Chon H, Cerritelli SM, Kunkel TA, Johansson E, Crouch RJ, et al. RNase H2-Initiated ribonucleotide excision repair. Mol Cell. 2012;47: 980–986. doi:10.1016/j.molcel.2012.06.035

44. Sommer D, Stith CM, Burgers PMJ, Lahue RS. Partial reconstitution of DNA large loop repair with purified proteins from Saccharomyces cerevisiae. Nucleic Acids Res. Oxford University Press; 2008;36: 4699–707. doi:10.1093/nar/gkn446

45. Clarkson SG. The XPG story. Biochimie. 2003. pp. 1113–1121. doi:10.1016/j.biochi.2003.10.014

46. O'Donovan A, Scherly D, Clarkson SG, Wood RD. Isolation of active recombinant XPG protein, a human DNA repair endonuclease. J Biol Chem. 1994;269: 15965–15968.

47. O'Driscoll M, Macpherson P, Xu YZ, Karran P. The cytotoxicity of DNA carboxymethylation and methylation by the model carboxymethylating agent azaserine in human cells. Carcinogenesis. 1999;20: 1855–1862. doi:10.1093/carcin/20.9.1855

48. Jerison ER, Kryazhimskiy S, Mitchell J, Bloom JS et al. Genetic variation in adaptability and pleiotropy in budding yeast. Elife. eLife Sciences Publications Limited; 2017;6: 1–38. doi:10.1101/121749

49. Amberg DC, Burke DJ, Strathern JN. Methods in yeast genetics: a cold spring harbor laboratory course manual. 2005. Available: http://www.amazon.com/Methods-Yeast-Genetics-Spring-Laboratory/dp/0879697288

50. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. Genetics. 1994;138: 963–971. doi:10.1534/genetics.107.080101

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

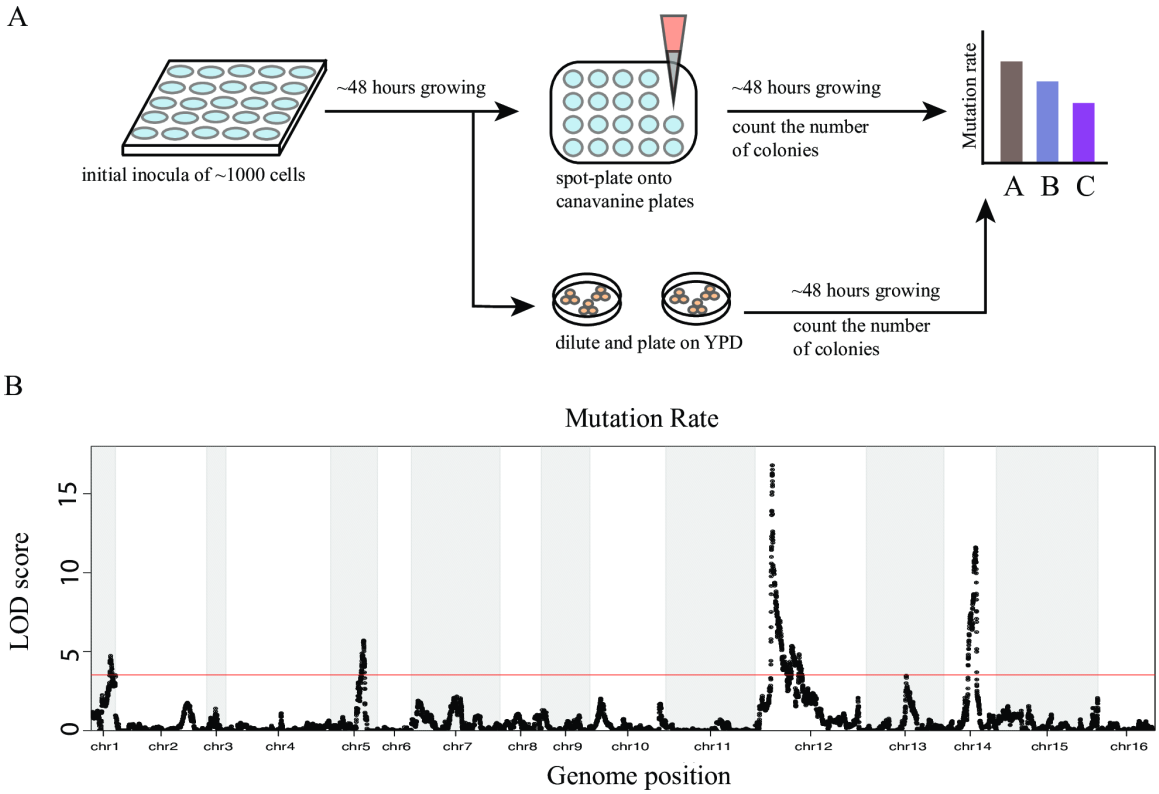


Fig 1. Linkage analysis identified four loci underlying mutation rate variation.

(A) The fluctuation assay was performed as shown in the workflow. The assay started with a small number of cells growing in 96-well plates in liquid SC-Arg medium for ~48 hours, followed by plating onto selective agar plates with canavanine. A proportion of the cultures were diluted to measure the number of cells per culture (Methods). Plates were imaged two days after spot-plating, and the number of colonies on canavanine plate was counted. (B) LOD score for mutation rate variation is plotted against the genetic map. The 4 significant QTLs explain 20.7% of the phenotypic variance. The red line indicates a 5% FWER significance threshold (LOD =3.52).

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

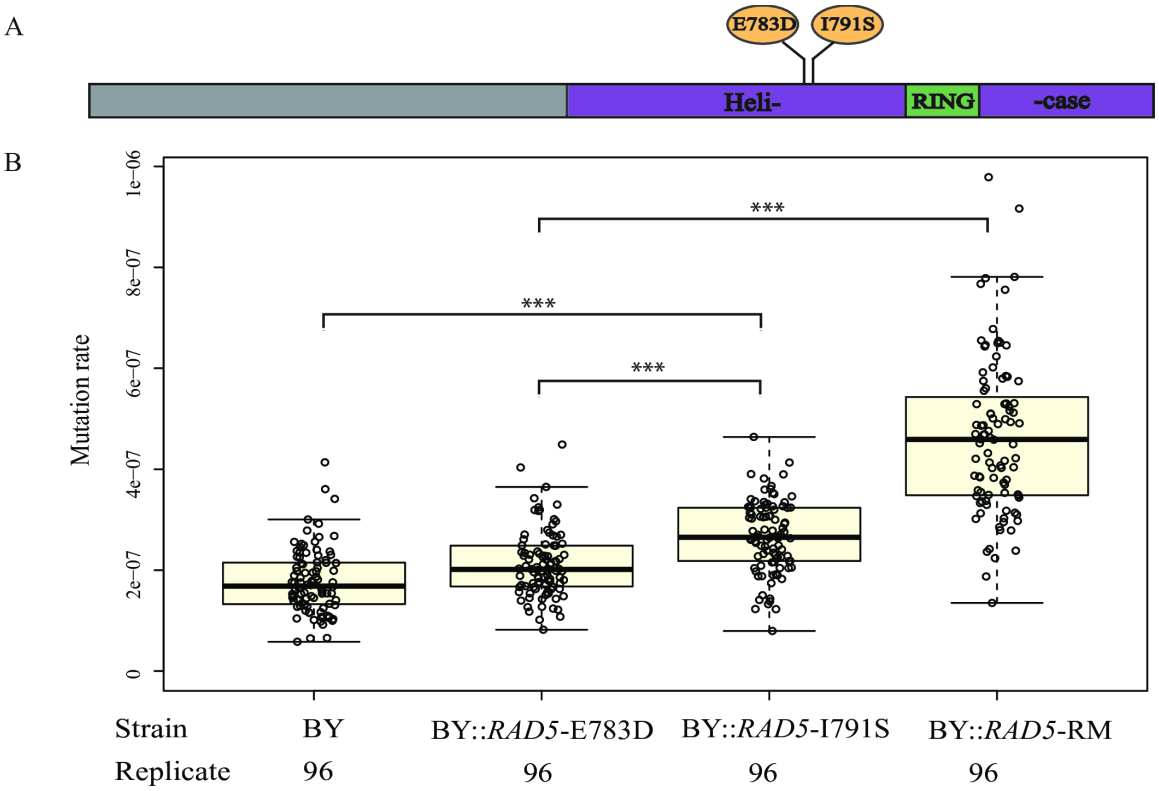


Fig 2. Polymorphisms in *RAD5* underlie mutation rate variation.

(A) *RAD5* polymorphisms between BY and RM are located in the helicase region. The first letter for each polymorphism indicates the BY polymorphisms (E783, I791) and the second letter indicates the RM polymorphisms (D783, S791). (B) The effect of single *RAD5* polymorphism and *RAD5* whole gene replacement was tested in the BY strain background for mutation rate. For each strain, the mutation rates of ninety-six replicates were measured. Bold lines show the mean. Boxes show the interquartile range. Statistical significance was tested using a permutation t-test. Permutation p value < 0.001 is shown as ***.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

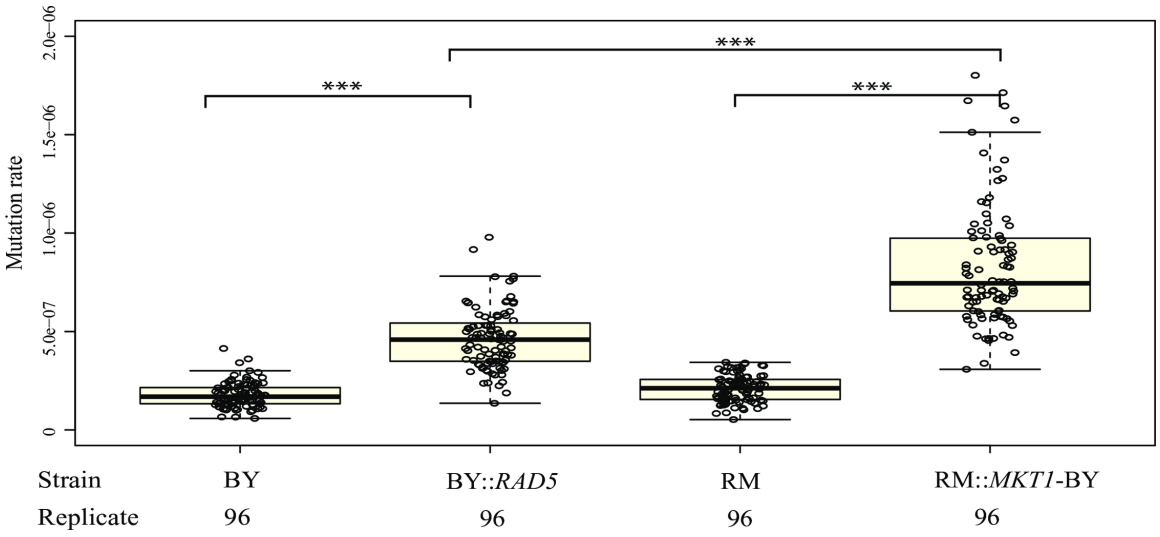


Fig 3. The RM allele of *RAD5* and BY allele of *MKT1* increase mutation rate.

The mutation rate of two allele replacement strains, the BY strain and the RM strain are measured and compared. For each strain, ninety-six replicate measurement for mutation rate was performed. Bold lines show the mean. Boxes show the interquartile range. Statistical significance was tested using permutation t-test. Permutation p value < 0.001 is shown as ***.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

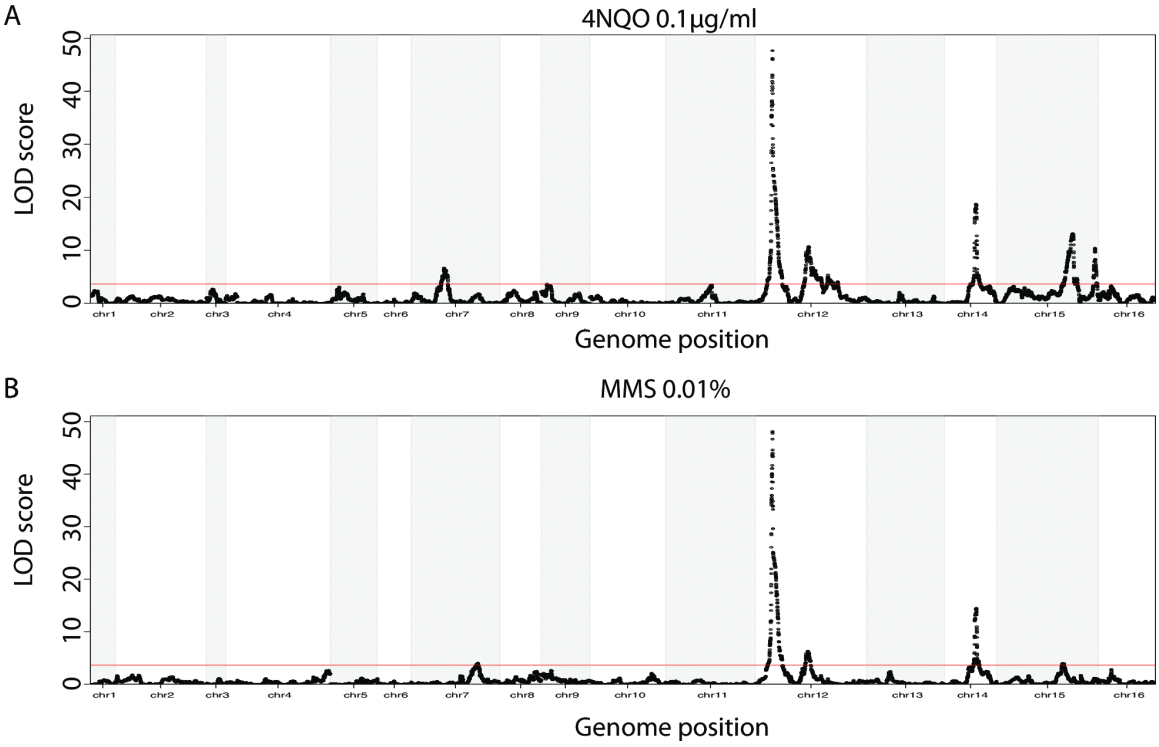
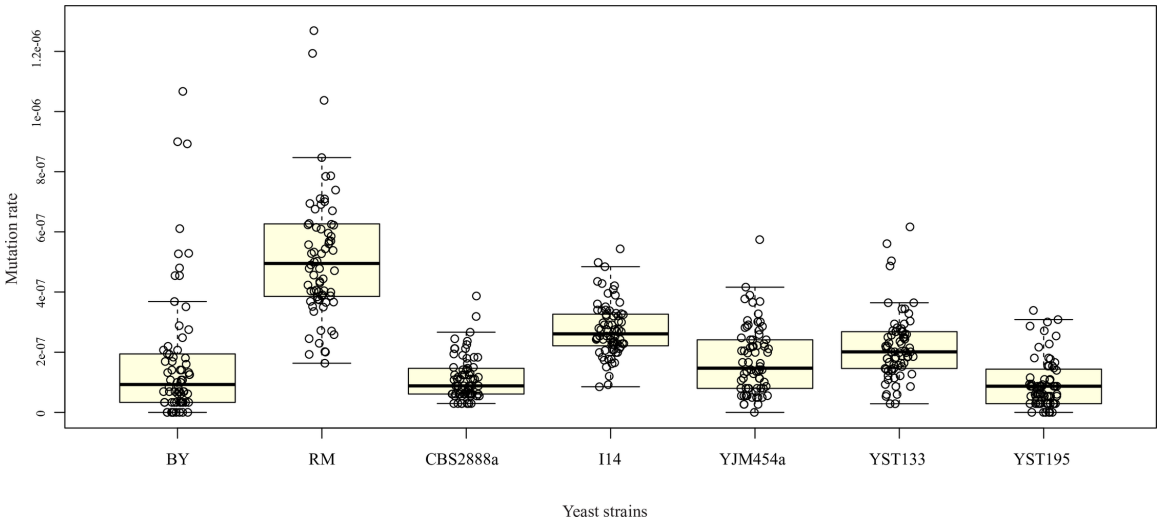


Fig 4. Loci underlying mutation rate variation, 4NQO sensitivity and MMS sensitivity are overlapped.

(A-B) The LOD scores for 4NQO (0.1 µg/ml) sensitivity and MMS (0.01%) sensitivity are plotted against the genetic map. The red line indicates a 5% FWER significance threshold (LOD=3.62 for 4NQO and LOD=3.61 for MMS).

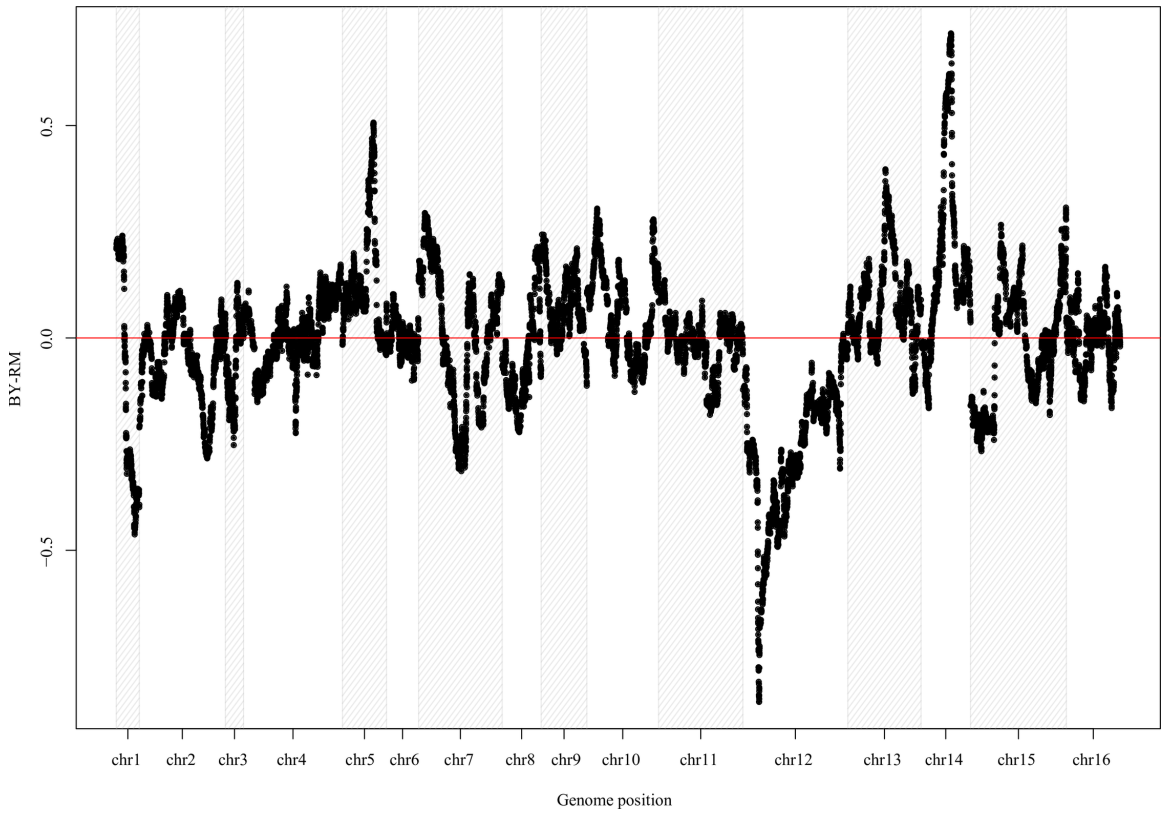
bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



S1 Fig. Mutation rate differs between seven natural yeast strains.

Ninety-six measure of mutation rate was performed for each strain. Means of the mutation rate are plotted as the line. Boxes show the 25%-75% percentile.

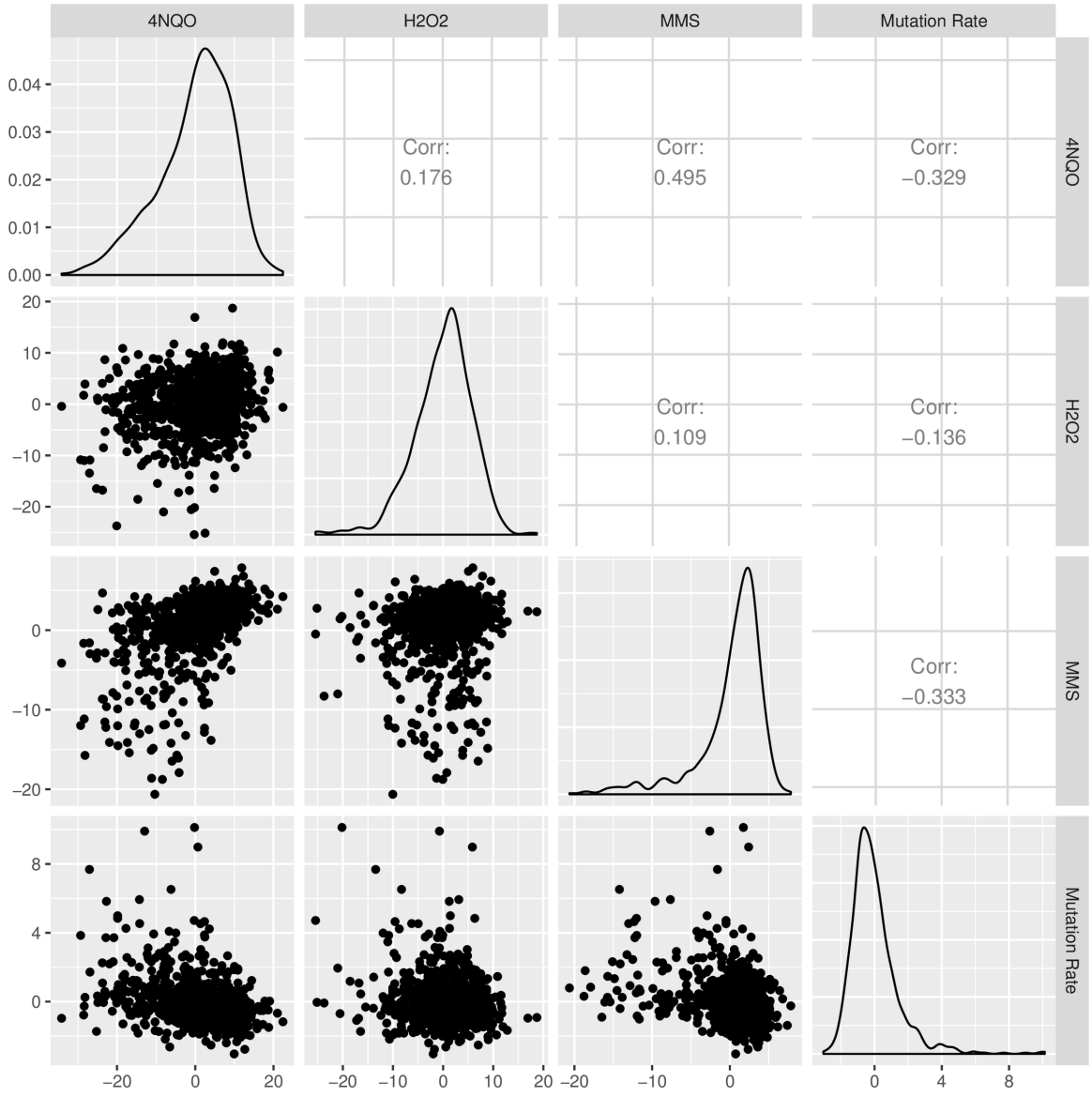
bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



S2 Fig. Loci on chromosome XII and XIV have large effects on mutation rate.

Effect size of genetic markers along the genome shows the BY alleles on chromosome XIV and V increase the mutation rate, while the RM alleles on chromosome XII and I increase the mutation rate.

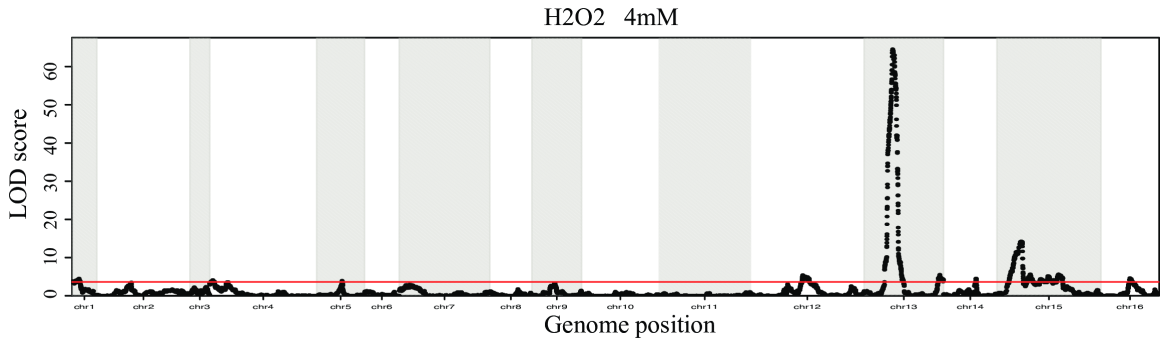
bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



S3 Fig. Mutation rate is positively correlated with 4NQO, MMS and H_2O_2 sensitivity in the segregant panel.

Trait values for mutagens are proxy measurements for mutagen resistance. As shown in the first row, mutation rate is negatively correlated with 4NQO, MMS and H_2O_2 resistances, meaning mutation rate is positively correlated with the sensitivity of these mutagens. Data are displayed in the lower triangle and the linear Pearson correlation values are shown in the upper triangle.

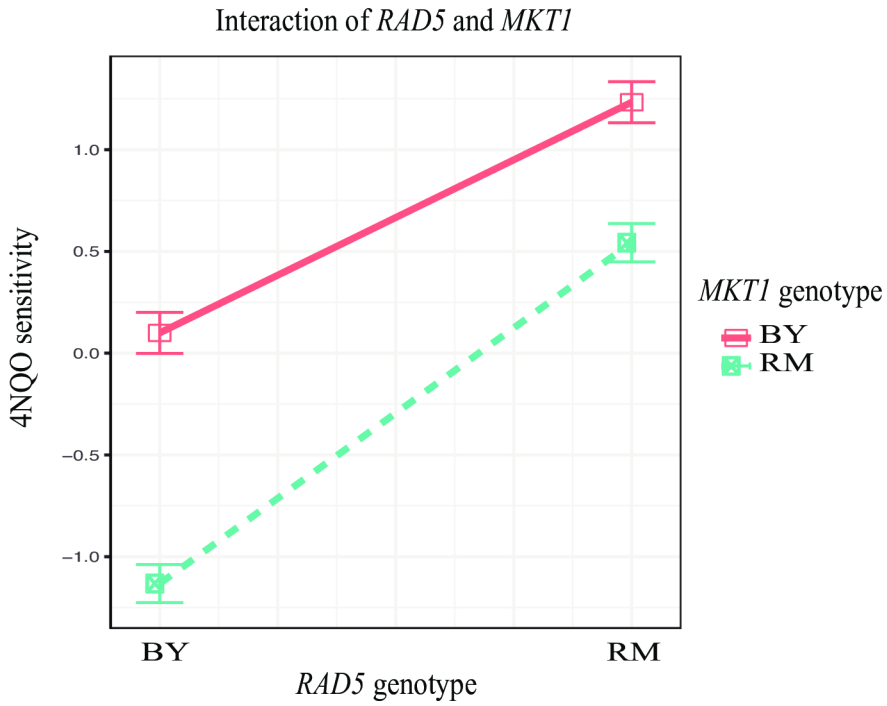
bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



S4 Fig. Loci underlie the H_2O_2 sensitivity.

LOD scores of sensitivity for H_2O_2 (4mM) are plotted against the genetic map. The red line indicates the significant threshold (3.64) from 1000 permutations.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



S5 Fig. 4NQO sensitivity in 1008 segregants shows interaction between *RAD5* and *MKT1*.

The 4NQO sensitivity of segregants is plotted with respect to their *RAD5* and *MKT1* genotypes. The x-axis corresponds to different parental alleles at *RAD5*. The colors correspond to different parental alleles at *MKT* (Red corresponds to the BY allele, and green corresponds to the RM allele). The y-axis corresponds to the sensitivity to 4NQO.

bioRxiv preprint doi: https://doi.org/10.1101/338723; this version posted June 4, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Strain	Mutation rate	Std. Error
BY4724	1.7×10^{-7}	2.2×10^{-7}
RM11-1a	5.8×10^{-7}	4.0×10^{-7}
CBS2888a	1.1×10^{-7}	7.2×10^{-8}
I14	2.8×10^{-7}	9.0×10^{-8}
YJM454a	1.7×10^{-7}	1.1×10^{-7}
YST133	2.2×10^{-7}	1.1×10^{-7}
YST195	1.7×10^{-7}	5.3×10^{-7}

S1 Table. The mutation rate of seven natural yeast strains.

Mutation rate shown in the table is the mean of ninety-six replicates.