The genetic basis of mutation rate variation in yeast

Liangke Gou [*], Joshua S. Bloom [*, †], Leonid Kruglyak [*, †, ‡, 1]

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

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

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

[1] 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 1040 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

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.

Introduction

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 (Baer *et al.* 2007). Mutation rates tends to be approximately 10^{-9} to 10^{-10} mutations per base pair, per cell division, for most microbial species (Drake *et al.* 1998), 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 (Demerec 1937). Two types of experimental approaches have been used to measure mutation rates in yeast. The first is the fluctuation assay (Luria and Delbrück 1943). 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 (Lang and Murray 2008a). A second method tracks mutation accumulation during experimental evolution and uses whole-genome sequencing to estimate mutation rates (Zhu *et al.* 2014). 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.

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 1040 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*.

73 ## Materials and Methods

74 **Yeast strains and media**

75 Seven natural *S. cerevisiae* strains (Table S1) were used in this study. The 1040
76 segregants derived from BY4724 (MATa) and RM11-1a (MATa, *MKT1*-BY,
77 hoΔ::HphMX, flo8Δ::NatMX) were generated, genotyped and described previously
78 (Bloom *et al.* 2013). The RM::*MKT1*-BY strain was made previously by our lab. The
79 BY::*RAD5*-RM strain and the *RAD5* variants substitution strains (Table 1) were from
80 Demogines et al (Demogines *et al.* 2008a). For fluctuation assay, yeast was grown in
81 synthetic complete liquid medium without arginine (SC-Arg) before plating onto
82 selective plates. For DNA damaging agents sensitivity assays, yeast were grown in
83 rich YPD medium (1% yeast extract, 2% peptone and 2% glucose) before plating onto
84 YPD agar plates with DNA damaging agents. SC-Arg and YPD liquid media and agar
85 plates were made according to Amberg et al (Amberg *et al.* 2005).

86

87 **Selection agar plate construction**

88 Selective canavanine plates were made from arginine minus synthetic complete agar
89 medium with 60mg/liter L-canavanine (Sigma C1625). The canavanine plates were
90 dried by incubating at 30ºC overnight. Selective plates for the DNA damaging agents
91 sensitivity assay were made with YPD agar medium containing the respective agents
92 at the concentrations indicated in Table 2. 50ml of the agar medium was poured into
93 each Nunc OmniTray plates (Thermo Scientific 264728) and placed on a flat surface
94 to solidify. Each experiment was performed with the same batch of selection plates.
95 The concentrations for 4NQO (Sigma N8141), MMS (Sigma 64382) and H_2O_2
96 (Sigma 216763) were 0.1μg/ml, 0.01% and 4mM. These concentrations capture the

97 sensitivity difference between the segregants, while maintaining enough colony
98 growth for QTL mapping.
99
100 **Fluctuation assays**
101 To begin the fluctuation assay, yeast were grown in synthetic complete medium
102 without arginine (SC-Arg) in 96-well plates (Costar 3370) for ~48 hours to saturation.
103 Saturated cultures were diluted and pinned into a new 96-well plate with liquid SC-
104 Arg medium. This step ensured a small number of ~1000 yeast cells in the initial
105 inoculum. Plates were sealed with a Breathe-Easy sealing membrane (Sigma Z380059)
106 to prevent evaporation, and incubated at 30ºC with shaking for ~48 hours. 100µl
107 saturated cultures were spot-plated onto canavanine plates in a four by six
108 configuration using a Biomek FX^P automated workstation. Plates with spot-plated
109 yeast culture were dried in the laminar flow hood (Nuair) for half an hour or until dry,
110 and incubated at 30ºC for ~48 hours. We imaged the plates using an imaging robot
111 (S&P Robotics BM3-SC), and the number of colonies in each spot was manually
112 counted from the images. An example of the imaged plate is shown in Figure S1.
113
114 Mutation rate was estimated using the Ma-Sandri-Sarkar Maximum Likelihood
115 Method where the numbers of observed colonies on canavanine plates was fitted into
116 the Luria-Delbrück distribution and a single parameter m was calculated (Sarkar *et al.*
117 1992). The parameter m represents the expected number of mutation events per
118 culture. For the natural isolates and engineered strains, the mutation rate was
119 calculated from the equation $\mu = m/N$, where N is the average number of cells per
120 culture (as a proxy for the number of cell divisions given the starting inoculum is
121 much smaller than N). In the segregant panel, we defined a mutation rate score that

122 was calculated as the residual phenotype after regressing out the effect of average
123 number of cells per strain (N) from the estimate of m per strain across all of the
124 segregants.
125
126 For each of the seven natural isolate strains, we performed ninety-six replicates of the
127 fluctuation assay, which means we had ninety-six estimations of mutation rate. In
128 each replicate three cultures were plated onto canavanine plates, and the number of
129 resistance colonies in these three plates were fitted into the Luria-Delbrück
130 distribution to estimate the mutation events per culture (m). One culture was diluted
131 and plated onto YPD to determine the number of cells per culture (N) in each
132 replicate. Given the number of replicates used for estimate m and N were limited, the
133 mutation rate estimation for the seven natural isolate strains had large variance. For
134 the BYxRM segregants panel, twelve independent replicate cultures were plated onto
135 canavanine plates for every segregant. The number of canavanine resistant colonies in
136 these twelve plates was fitted into the Luria-Delbrück distribution to calculate the
137 number of mutations per culture (m), and one culture was diluted and plated on the
138 YPD plates to determine the number of cells (N). Given only one culture was used to
139 estimate the number of cells (N) for each segregants, which means the mutation rate
140 estimation from the equation $\mu = m/N$ would be largely affected by that one measure
141 of N. To minimize the noise driven by the measure of N, we removed the segregants
142 with relatively less accurate measure of N from the later QTL mapping, only 843
143 segregants with confident measure of m and N were used. Furthermore, we defined a
144 mutation rate score that regressed out the effect of N instead of the division. We used
145 the mutation rate score of segregants for later QTL mapping. For each allele
146 replacement strain (Table 1), ninety-six replicates of fluctuation analysis were

147 performed, providing us ninety-six estimations of mutation rate. In each replicate,
148 twelve cultures were plated onto canavanine plate to estimate the number of
149 mutations per culture (m), and three cultures were pooled, diluted and plated on YPD
150 plates to determine the number of cells per culture (N).
151
152 **Yeast growth measurement for DNA damaging agents sensitivity assay**
153 The segregant panel were originally stored in 96-well plates (Costar 3370). During the
154 DNA damaging agents sensitivity assay, individual segregants were inoculated in two
155 plate configurations in 384-well plates (Thermo Scientific 264574) with YPD and
156 grown for ~48 hours in a 30ºC incubator without shaking. Saturated cultures were
157 mixed for 1min at 2,000 r.p.m. using a MixMate (Eppendorf) before pinning. The
158 colony handling robot (S&P Robotics BM3-SC) was used to pin segregants onto
159 selective agar plates with 384 long pins. The plates were incubated at 30ºC for ~48
160 hours and imaged by the colony handling robot (S&P Robotics BM3-SC). Custom R
161 code (Bloom *et al.* 2013) was used to determine the size of each colony and the size
162 was used as a proxy for growth in the presence of the DNA damaging agents.
163
164 **QTL mapping**
165 In order to control for intrinsic growth rate differences and plate position effects, we
166 normalized the traits for growth by fitting a regression for growth of the yeast that
167 were in the same layout configuration on control plate (YPD agar plates for mutagen
168 sensitivity assay). Residuals were used for QTL mapping. We tested for linkage by
169 calculating logarithm likelihood ratio (LOD scores) for each genotypic marker and
170 trait as $-n(\ln(1-r^2)/(2\ln(10))$, where r is the Pearson correlation coefficient
171 between the segregant genotypes and the segregant mutation rate or DNA damaging

172 agents sensitivity. The threshold declaring the significant QTL effect was calculated
173 from the empirical null distribution of the maximum LOD score determined from
174 1,000 permutations (Churchill and Doerge 1994). The estimated 5% family-wise error
175 rate significance thresholds were 3.52, 3.62, 3.61 and 3.64 for mutation rate, mutagen
176 sensitivity for 4NQO, MMS and H_2O_2 respectively. The 95% confidence intervals
177 were determined using a 1.5 LOD score drop. The code and the data for QTL
178 mapping is available at
179 https://github.com/gouliangke/Mutation-rate/tree/master/qtl_mapping
180

181 **Amplicon sequencing of the *CAN1* region in segregants**

182 1040 segregants were assigned into four groups according to their alleles at gene
183 *RAD5* and *MKT1* (Table S2). We collected the cananvanine resistant colonies from
184 the canavanine plates that we used to measure the mutation rate of segregants in the
185 previous fluctuation analysis. A single canavanine resistant colony (if any) was
186 picked from each segregant and the picked colonies from the same group were pooled
187 together for DNA extraction. DNA was extracted from the pool using the Qiagen
188 DNeasy Blood & Tissue Kit. The *CAN1* region was amplified from the DNA of four
189 groups using the Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific)
190 and eight pairs of designed primers (File S3). The amplicon sequencing library was
191 prepared using the Illumina Nextera DNA Library Prep Kit with the adjusted
192 protocols to skip the nextera treatment. The library was then sequenced on the MiSeq
193 platform using the MiSeq Reagent V2 Nano Kit. The sequencing read counts for each
194 group were controlled to be the same (Figure S7). Custom R codes were used to
195 detect the mutation rate spectrum of each group. The code for the mutation rate
196 spectrum analysis is available at

197 https://github.com/gouliangke/Mutation-rate/tree/master/mutation_spectrum

198

199 ## Results

200 **High-throughput fluctuation assay for measuring mutation rates**

201 The fluctuation assay for measuring mutation rate involves growing many parallel
202 cultures, each starting from a small number of cells, under non-selective conditions,
203 followed by plating to selective medium to identify mutants. The number of mutations
204 that occurs in each culture should follow the Poisson distribution, as mutations arise
205 spontaneously. However, the number of mutant cells that survive on the selective
206 plates can vary greatly, because early mutations are inherited by all offspring of the
207 mutant. This leads to the "jackpot" effect, in which some cultures contain a large
208 number of mutant individuals. The number of observed mutant cells per culture
209 follows the Luria-Delbrück distribution (Luria and Delbrück 1943), and the Ma-
210 Sandri-Sarkar maximum likelihood method can be used to estimate the expected
211 number of mutations per culture from the observed numbers of mutants (Sarkar *et al.*
212 1992). The underlying mutation rate is then calculated by dividing the number of
213 mutations per culture by the average number of cells per culture [1,4]. Here we
214 measured rare spontaneous loss-of-function mutations in the gene *CAN1*, which
215 encodes an arginine permease. Yeast cells carrying loss-of-function mutations in
216 *CAN1* can grow on canavanine, an otherwise toxic arginine analog. Typically,
217 fluctuation assays are labor-intensive and have limited throughput, because a large
218 number of parallel cultures is required for estimating the mutation rate in each assay,
219 and several replicate assays are needed for a robust measurement of the mutation rate
220 in each strain (Lang and Murray 2008b). We modified the fluctuation assay into a
221 high-throughput method for measuring mutation rates in many strains in parallel. We

222 grew cultures in 96-well plates, automated the spotting of cultures, and used high-
223 resolution imaging to rapidly count mutants on many plates (Methods, Figure 1A).
224 The automated spotting process for 96 strains took only approximately twenty
225 minutes, and the imaging process required even less time. These improvements
226 enabled us to measure the spontaneous mutation rates in the hundreds of strains
227 necessary for genetic mapping.
228
229 **Spontaneous mutation rate varies among yeast isolates**
230 To investigate mutation rate variation among *S. cerevisiae* strains, we measured the
231 spontaneous mutation rate of seven yeast isolates using the high-throughput
232 fluctuation assay (Table S1). The seven strains span a large range of yeast genetic
233 diversity (Schacherer *et al.* 2009). We found that the mutation rates of these strains
234 range from 1.1×10^{-7} to 5.8×10^{-7} mutations per gene per generation, with a median of
235 1.7×10^{-7} (Table S1, Figure S2). The median mutation rate was very similar to the
236 previously reported mutation rate at *CAN1* (Lang and Murray 2008a). In particular,
237 the mutation rate we observed for the BY strain (1.7×10^{-7}) is very similar to the
238 previously reported rate, which was measured in strain W303 (1.5×10^{-7}) (Lang and
239 Murray 2008a), consistent with the fact that W303 shares a large fraction of its
240 genome with BY (Liti *et al.* 2009). An analysis of variance (ANOVA) showed that
241 strain identity explained a significant fraction of the observed variance in mutation
242 rates ($F=69.9$, df=6, $p<2\times10^{-16}$) (Figure S2). The fraction of total variance in mutation
243 rates explained by the repeatability of measurements for each strain, 46%, serves as
244 an upper bound for the estimate of the total contribution of genetic differences
245 between strains to trait variation (broad-sense heritability or H^2). We observed that
246 RM, a vineyard strain, had a mutation rate higher than all other strains (Figure S2).

247

248 **Four QTLs explain the majority of observed mutation rate variation**

249 In order to find the genetic factors underlying the difference in mutation rate between
250 BY and RM, we performed quantitative trait locus (QTL) mapping in 1040 genotyped
251 haploid segregants from a cross between these strains (Bloom *et al.* 2013). We
252 measured the mutation rate of each segregant using the high-throughput fluctuation
253 assay (Methods). We estimated the fraction of phenotypic variance explained by the
254 additive effects of all segregating markers (narrow-sense heritability) to be 30%
255 (Methods) (Lynch and Walsh 1998). This sets an upper bound for the expectation of
256 the total amount of additive genetic variance that could be explained with a QTL-
257 based model. QTL mapping in the segregant panel identified significant linkage at
258 four distinct loci (Figure 1B). At two of the QTLs, on chromosomes XII and I, the
259 RM allele conferred a higher mutation rate, consistent with the higher mutation rate of
260 this strain. At the other two QTLs, on chromosomes XIV and V, the BY allele
261 conferred a higher mutation rate (Figure S3), showing that a strain with lower trait
262 value can nevertheless harbor trait-increasing alleles. The four detected QTLs
263 explained 20.7% of the phenotypic variance, thus accounting for 69% of the estimated
264 additive heritability. The loci on chromosomes XII, XIV, I and V explained 8.8%,
265 6.1%, 3.1% and 2.6% of the variance, respectively. We tested the four identified
266 QTLs for pairwise interactions and found a significant interaction between the QTL
267 on chromosome XII and the QTL on chromosome XIV that explained 1% of the
268 phenotypic variance (F=8.41, df=1, Bonferroni-corrected p=0.023).

269

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

272 Ten genes fell within the confidence interval of the QTL on chromosome XII. A
273 strong candidate was *RAD5*. Previous studies showed that natural variants in *RAD5*
274 contribute to sensitivity to the mutagen 4-nitroquinoline 1-oxide (4NQO) (Demogines
275 *et al.* 2008a). *RAD5* encodes a DNA repair protein involved in the error-free DNA
276 damage tolerance (DDT) pathway (Torres-Ramos *et al.* 2002; Blastyák *et al.* 2007).
277 The DDT pathway promotes the bypass of single-stranded DNA lesions encountered
278 by DNA polymerases during DNA replication, thus preventing the stalling of DNA
279 replication (Bi 2015). *RAD5* plays a crucial role in one branch of the DDT pathway
280 called template switching (TS), in which the stalled nascent strand switches from the
281 damaged template to the undamaged newly synthesized sister strand for extension
282 past the lesion (Bi 2015). Two non-synonymous substitutions exist between BY and
283 RM strains in *RAD5* (Figure 2A), at amino acid positions 783 (glutamic acid in BY
284 and aspartic acid in RM) and 791 (isoleucine in BY and serine in RM). According to
285 Pfam alignments (Sonnhammer *et al.* 1997), *RAD5* contains a HIRAN domain, an
286 SNF2-related N-terminal domain, a RING-type zinc finger domain, and a helicase C-
287 terminal domain (Figure 2A). Both non-synonymous polymorphisms mapped to the
288 helicase domain of *RAD5* (Figure 2A), and no other sequenced strains of *S. cerevisiae*
289 contain the aspartic acid 783 and serine 791 variants that are private to the RM strain.
290 We used protein variation effect analyzer (PROVEAN) (Choi and Chan 2015) to
291 predict whether the two non-synonymous substitutions have an impact on the
292 biological function of the protein. PROVEAN showed the I791S substitution (score –
293 5.4) might have a strong deleterious effect, while the E783D variant (score –1.8) was
294 not predicted to have a strong effect.

295

296 Nineteen genes fell within the confidence interval of the QTL on chromosome XIV.
297 A strong candidate was *MKT1*, which was also reported to affect 4NQO sensitivity
298 (Demogines *et al.* 2008a). *MKT1* encodes an RNA-binding protein that affects
299 multiple traits and underlies an eQTL hotspot in yeast (Albert and Kruglyak 2015).
300 The RM allele of *MKT1* increases sporulation rate (Deutschbauer and Davis 2005)
301 and improves survival at high temperature (Steinmetz *et al.* 2002), in low glucose
302 (Parreiras *et al.* 2011), after exposure to DNA-damaging agents (Demogines *et al.*
303 2008a), and in high ethanol levels (Swinnen *et al.* 2012). The coding region of the BY
304 and RM alleles of *MKT1* differs by one synonymous polymorphism and two non-
305 synonymous substitutions. *MKT1* has an XPG domain, which is relevant to DNA
306 repair, and an *MKT1* domain, which is related to the maintenance of K2 killer toxin
307 (Wickner 1980). One non-synonymous variant is in the XPG domain at amino acid
308 position 30 (aspartic acid in BY and glycine in RM), while the other non-synonymous
309 variant is in the *MKT1* domain at position 453 (lysine in BY and arginine in RM).
310 PROVEAN predicted a large effect of the D30G variant (score 6.7) on the function of
311 *MKT1*, and this variant was previously found to influence sporulation rate
312 (Deutschbauer and Davis 2005), mitochondrial genome stability (Dimitrov *et al.* 2009)
313 and survival at high temperature [22]. The other variant (K453R) was not predicted to
314 have a strong effect (score 0.6).

315

316 We tested whether *RAD5* and *MKT1* alleles caused differences in mutation rate by
317 using the fluctuation test on allele replacement strains (Demogines *et al.* 2008a; Smith
318 and Kruglyak 2008) (Table 1). The BY strain carrying the RM allele of *RAD5*
319 (BY::*RAD5*-RM) had a higher mutation rate than the BY strain (permutation t-test,

320 mean difference=2.9×10^{-7}, $p<1 \times 10^{-4}$), demonstrating that the RM *RAD5* allele
321 increases mutation rate (Figure 3A). This result is consistent with the observed
322 difference between segregants grouped by parental allele at *RAD5* (mean
323 difference=2.3×10^{-7}). The RM strain carrying the BY allele of *MKT1* (RM::*MKT1*-BY)
324 had a higher mutation rate than the RM strain (permutation t-test, mean
325 difference=6.1×10^{-7}, $p<1 \times 10^{-4}$), showing that the BY *MKT1* allele increases mutation
326 rate (Figure 3A), consistent with the direction of effect observed in the segregants.
327
328 To gain a finer-level understanding of the two missense variants between BY and RM
329 in the gene *RAD5,* we tested strains (Demogines *et al.* 2008a) in which these sites in
330 BY were individually replaced with the RM alleles (Table 1) by site-directed
331 mutagenesis. Strains with either variant had a higher mutation rate than BY
332 (permutation t-test, mean difference=0.9×10^{-7}, $p<1 \times 10^{-4}$ for BY::*RAD5*-I791S; mean
333 difference=0.3×10^{-7}, $p=6 \times 10^{-4}$ for BY::*RAD5*-E783D) (Figure 2B), suggesting that
334 both variants contribute to the higher mutation rate. The BY strain with the I791S
335 substitution had a higher mutation rate than the BY strain with the E783D substitution
336 (permutation t-test, mean difference=0.6×10^{-7}, $p<1 \times 10^{-4}$) (Figure 2B), consistent with
337 the PROVEAN prediction of a stronger effect for the I791S variant. However, neither
338 variant alone nor the additive effect of the two variants fully recapitulated the increase
339 in mutation rate that we observed when replacing the entire coding region of *RAD5* in
340 BY with the RM allele ($F=67.6$, $df=1$, $p=3.3 \times 10^{-15}$), suggesting an interaction between
341 the two variants.
342

343 **Mutation rate shares two large effect QTLs with growth on DNA damaging**
344 **agents 4NQO and MMS**
345 Deficiencies in DNA repair can increase mutation rate (Supek and Lehner 2015;
346 Sabarinathan *et al.* 2016) and increase sensitivity to DNA damaging agents such as
347 alkylating compounds and UV light (Sun and Moses 1991; Frankfurt 1991). We
348 hypothesized that genetic variants that cause deficiencies in DNA repair may underlie
349 QTLs for both mutation rate variation and sensitivity to DNA damaging agents.
350 Previously, Demogines *et al.* identified a large-effect QTL on chromosome XII for
351 MMS and 4NQO sensitivity in a panel of 123 segregants from a cross between BY
352 and RM (Demogines *et al.* 2008a). Additionally, they identified a QTL on
353 chromosome XIV for 4NQO sensitivity by using backcrossing and bulk segregant
354 analysis. These QTLs overlapped with the major QTLs that we identified for mutation
355 rate variation, and the underlying causal genes for 4NQO sensitivity were also *RAD5*
356 and *MKT1*.
357
358 To follow up on these results, we measured sensitivity to three different DNA
359 damaging agents in our panel of 1040 segregants (Table 2). The compounds assayed
360 included methyl methanesulfonate (MMS), an alkylating agent that induces DNA
361 double strand breaks and stalls replication forks (Hampsey 1997), 4NQO, an
362 ultraviolet light mimetic agent (Hampsey 1997) and hydrogen peroxide (H_2O_2), a
363 compound that induces DNA single and double strand breaks (Hampsey 1997). We
364 observed that segregants with higher mutation rate, and presumably less efficient
365 DNA repair systems, were more sensitive to MMS, 4NQO and H_2O_2 (Figure S4),
366 consistent with our hypothesis that deficiencies in DNA repair increase the rate of
367 spontaneous mutations and increase sensitivity to DNA damaging agents. We

368 identified two large-effect QTLs for 4NQO and MMS sensitivity that overlapped with
369 the major QTLs for mutation rate (Figure 4A and B). An interaction between *RAD5*
370 and *MKT1* was observed for 4NQO sensitivity (F=8.5, df=1, p=0.004) (Figure S5).
371 The QTLs on chromosome 12 and 14 were still observed in the linkage mapping for
372 H_2O_2, but they had small effects (Figure S6). The large effect QTLs detected for H_2O_2
373 sensitivity on other chromosomes likely reflects trait-specific effects of variants
374 acting on sensitivity to H_2O_2 (Figure S6).
375
376 **Similar mutation spectra in segregants with different *RAD5* and *MKT1***
377 **genotypes**
378 In order to gain a better understanding of how genetic variation in *RAD5* and *MKT1*
379 might influence the DNA damage repair process, we characterized the spontaneous
380 mutation spectrum at *CAN1* in the segregants. We divided 1040 segregants into four
381 groups based on their genotypes at *RAD5* and *MKT1* and sequenced pools of *CAN1*-
382 resistant mutants from each group (Table S2; Figure S7). The mutation spectra of the
383 four groups are shown in Table S3 and Figure S8. C:G > T:A transitions were the
384 most frequently observed mutations. A:T > G:C was the rarest transition, and A:T >
385 T:A was the rarest transversion. The spectra for single base pair substitutions
386 observed here (Figure S9) are similar to previous observations based on whole-
387 genome sequencing of mutation accumulation strains (Zhu *et al.* 2014). While there
388 were some differences in the relative frequencies of specific mutation types (for
389 instance, more C:G > G:C transversions in segregants with the RM *MKT1* allele and
390 more A:T > C:G transversions in segregants with BY *RAD5* allele), these mutation
391 differences were not statistically significant after correction for multiple testing.
392

393 **Discussion**

394 We developed and implemented a high-throughput fluctuation assay to directly
395 measure mutation rates in yeast. We used this assay to map four QTLs that influence
396 differences in the spontaneous mutation rate, and narrowed the two QTLs with the
397 largest effects to causal genes and variants. We attempted to gain insight into how
398 these variants might affect the mutation rate by comparing mutational spectra of
399 segregants grouped by genotype, but the differences we observed did not reach
400 statistical significance.

401

402 We identified *RAD5* as the gene underlying the QTL with the largest effect on
403 mutation rate. *RAD5* encodes a DNA helicase and ubiquitin ligase involved in error-
404 free DNA damage tolerance (DDT), a pathway that facilitates chromosome
405 replication through DNA lesions (Hishida *et al.* 2009; Unk *et al.* 2010). Previous
406 work showed that Rad5 is a structure-specific DNA helicase that is able to carry out
407 replication fork regression (Blastyák *et al.* 2007), a process of remodeling the
408 replication fork into four-way junctions when replication perturbations are
409 encountered (Neelsen and Lopes 2015). This process was hypothesized to promote
410 DNA damage tolerance and repair during replication (Neelsen and Lopes 2015). We
411 showed that two non-synonymous variants between BY and RM in the helicase
412 domain affect mutation rate. The RM allele of *RAD5* increases the sensitivity of yeast
413 to 4NQO and MMS (Demogines *et al.* 2008b), probably due to a defect in replication
414 fork regression. Thus the RM allele of *RAD5* causes both decreased growth in
415 mutagenic conditions and a higher mutation rate in non-stressful normal conditions.

416

417 We furthermore showed that polymorphisms in *MKT1* contribute to mutation rate
418 variation. *MKT1* is a highly pleiotropic gene that has been shown to affect levels of
419 transcript and protein abundance for numerous genes (Smith and Kruglyak 2008)
420 (Foss *et al.* 2011), as well as numerous cellular phenotypes (Steinmetz *et al.* 2002;
421 Deutschbauer and Davis 2005; Demogines *et al.* 2008a; Parreiras *et al.* 2011;
422 Swinnen *et al.* 2012; Wang and Kruglyak 2014; Albert and Kruglyak 2015). The BY
423 and RM alleles of *MKT1* differ by two non-synonymous substitutions that map to
424 amino acid positions 30 (aspartic acid in BY; glycine in RM) and 453 (lysine in BY;
425 arginine in RM). The latter variant (K453R) is located in the MKT1 domain, which is
426 required for activity of the Mkt1 protein in maintaining K2 killer toxin (Vermut *et al.*
427 1994). The former variant (D30G) localizes to the XPG-N (the N-terminus of XPG)
428 domain. Four other yeast proteins contain this domain: Exo1, Din7, Rad27 and Rad2.
429 All of these proteins have functions related to DNA repair and cellular response to
430 DNA damage, including DNA double-strand break repair (Exo1) (Tkach *et al.* 2012),
431 DNA mismatch repair (Exo1, Din7) (Koprowski *et al.* 2003; Tran *et al.* 2007),
432 nucleotide excision repair (Rad2) (Northam *et al.* 2010), ribonucleotide excision
433 repair (Rad27) (Sparks *et al.* 2012) and large loop repair (LLR) (Rad27) (Sommer *et*
434 *al.* 2008). The internal XPG (XPG-I) domain, together with XPG-N, forms the
435 catalytic domain of the Xeroderma Pigmentosum Complementation Group G (XPG)
436 protein. The XPG protein has well-established catalytic and structural roles in
437 nucleotide excision repair, a DNA repair process, and acts as a cofactor for a DNA
438 glycosylase that removes oxidized pyrimidines from DNA (Clarkson 2003). In
439 humans, mutations in the XPG protein commonly cause Xeroderma Pigmentosum,
440 which often leads to skin cancer (O'Donovan *et al.* 1994). The aspartic acid at
441 position 30 in the XPG domain of Mkt1 is only found in BY and related laboratory

442 strains. We hypothesize that Mkt1 has a previously unknown function in DNA
443 damage repair, mediated through its XPG domain.
444
445 We found that variants in *RAD5* and *MKT1* contribute to both mutation rate variation
446 and mutagen sensitivity. These results suggest that spontaneously occurring mutations
447 may have a similar mutation spectrum to those created by 4NQO and MMS, and are
448 potentially repaired by the same mechanisms. Deficient DNA repair can lead to
449 increased sensitivity to agents such as alkylating compounds and UV light (Sun and
450 Moses 1991; Frankfurt 1991; O'Driscoll *et al.* 1999) and to higher mutation rates at
451 sites that are less accessible to the DNA repair system (Sabarinathan *et al.* 2016).
452 Because mutation rates can be difficult to measure, sensitivity to mutagens may serve
453 as a useful proxy.
454
455 Recently, Jerison et al. reported heritable differences in adaptability in 230 yeast
456 segregants from the same cross we studied here (Jerison *et al.* 2017). They measured
457 adaptability as the difference in fitness between a given segregant ('founder') and a
458 descendant of that founder after 500 generations of experimental evolution.
459 Interestingly, *RAD5* fell within one of the QTLs found to influence adaptability.
460 Together with our observation that *RAD5* influences mutation rate, this finding
461 suggests that differences in mutation rate can affect the adaptability of organisms.
462

463 **Acknowledgments**

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

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

468

469 Author Contributions

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

472

473 References

474

475 Albert F. W., and L. Kruglyak, 2015 The role of regulatory variation in complex traits
476 and disease. Nat. Rev. Genet. 16: 197–212. https://doi.org/10.1038/nrg3891
477 Amberg D. C., D. J. Burke, and J. N. Strathern, 2005 *Methods in yeast genetics: a*
478 *cold spring harbor laboratory course manual*.
479 Baer C. F., M. M. Miyamoto, and D. R. Denver, 2007 Mutation rate variation in
480 multicellular eukaryotes: causes and consequences. Nat. Rev. Genet. 8: 619–631.
481 https://doi.org/10.1038/nrg2158
482 Bi X., 2015 Mechanism of DNA damage tolerance. World J. Biol. Chem. 6: 48.
483 https://doi.org/10.4331/wjbc.v6.i3.48
484 Blastyák A., L. Pintér, I. Unk, L. Prakash, S. Prakash, *et al.*, 2007 Yeast rad5 protein
485 required for postreplication repair has a DNA helicase activity specific for
486 replication fork regression. Mol. Cell 28: 167–175.
487 https://doi.org/10.1016/j.molcel.2007.07.030
488 Bloom J. S., I. M. Ehrenreich, W. T. Loo, T.-L. V. Lite, and L. Kruglyak, 2013
489 Finding the sources of missing heritability in a yeast cross. Nature 494: 234–237.
490 https://doi.org/10.1038/nature11867
491 Choi Y., and A. P. Chan, 2015 PROVEAN web server: a tool to predict the functional
492 effect of amino acid substitutions and indels. Bioinformatics 31: 2745–7.
493 https://doi.org/10.1093/bioinformatics/btv195
494 Churchill G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative
495 trait mapping. Genetics 138: 963–971.
496 https://doi.org/10.1534/genetics.107.080101
497 Clarkson S. G., 2003 The XPG story. Biochimie 85: 1113–1121.
498 Demerec M., 1937 Frequency of spontaneous mutations in certain stocks of
499 drosophila melanogaster. Genetics 22: 469–478.
500 Demogines A., E. Smith, L. Kruglyak, and E. Alani, 2008a Identification and
501 dissection of a complex DNA repair sensitivity phenotype in baker's yeast. PLoS
502 Genet. 4. https://doi.org/10.1371/journal.pgen.1000123
503 Demogines A., A. Wong, C. Aquadro, and E. Alani, 2008b Incompatibilities
504 involving yeast mismatch repair genes: a role for genetic modifiers and
505 implications for disease penetrance and variation in genomic mutation rates.

506 PLoS Genet. 4: e1000103. https://doi.org/10.1371/journal.pgen.1000103
507 Deutschbauer A. M., and R. W. Davis, 2005 Quantitative trait loci mapped to single-
508 nucleotide resolution in yeast. Nat. Genet. 37: 1333–1340.
509 https://doi.org/10.1038/ng1674
510 Dimitrov L. N., R. B. Brem, L. Kruglyak, and D. E. Gottschling, 2009
511 Polymorphisms in multiple genes contribute to the spontaneous mitochondrial
512 genome instability of Saccharomyces cerevisiae S288C strains. Genetics 183:
513 365–383. https://doi.org/10.1534/genetics.109.104497
514 Drake J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow, 1998 Rates of
515 spontaneous mutation. Genetics 148: 1667–1686. https://doi.org/citeulike-article-
516 id:610966
517 Foss E. J., D. Radulovic, S. A. Shaffer, D. R. Goodlett, L. Kruglyak, *et al.*, 2011
518 Genetic variation shapes protein networks mainly through non-transcriptional
519 mechanisms. PLoS Biol. 9. https://doi.org/10.1371/journal.pbio.1001144
520 Frankfurt O. S., 1991 Inhibition of DNA repair and the enhancement of cytotoxicity
521 of alkylating agents. Int. J. cancer 48: 916–23.
522 Hampsey M., 1997 A review of phenotypes in Saccharomyces cerevisiae. Yeast 13:
523 1099–1133.
524 Hishida T., Y. Kubota, A. M. Carr, and H. Iwasaki, 2009 RAD6-RAD18-RAD5-
525 pathway-dependent tolerance to chronic low-dose ultraviolet light. Nature 457:
526 612–615. https://doi.org/10.1038/nature07580
527 Jerison E. R., S. Kryazhimskiy, J. Mitchell, and J. S. et al. Bloom, 2017 Genetic
528 variation in adaptability and pleiotropy in budding yeast. Elife 6: 1–38.
529 https://doi.org/10.1101/121749
530 Koprowski P., M. U. Fikus, P. Dzierzbicki, P. Mieczkowski, J. Lazowska, *et al.*, 2003
531 Enhanced expression of the DNA damage-inducible gene DIN7 results in
532 increased mutagenesis of mitochondrial DNA in Saccharomyces cerevisiae. Mol.
533 Genet. Genomics 269: 632–639. https://doi.org/10.1007/s00438-003-0873-8
534 Lang G. I., and A. W. Murray, 2008a Estimating the per-base-pair mutation rate in the
535 yeast Saccharomyces cerevisiae. Genetics 178: 67–82.
536 https://doi.org/10.1534/genetics.107.071506
537 Lang G. I., and A. W. Murray, 2008b Estimating the per-base-pair mutation rate in
538 the yeast Saccharomyces cerevisiae. Genetics 178: 67–82.
539 https://doi.org/10.1534/genetics.107.071506
540 Liti G., D. M. Carter, A. M. Moses, J. Warringer, L. Parts, *et al.*, 2009 Population
541 genomics of domestic and wild yeasts. Nature 458: 337–341.
542 https://doi.org/10.1038/nature07743
543 Luria S., and M. Delbrück, 1943 Mutations of bacteria from virus sensitivity to virus
544 resistance. Genetics 28: 491–511. https://doi.org/10.1038/nature10260
545 Lynch M., and B. Walsh, 1998 Genetics and analysis of quantitative traits, p. 980 in
546 *Genetics and Analysis of Quantitative Traits*,.
547 Neelsen K. J., and M. Lopes, 2015 Replication fork reversal in eukaryotes: from dead
548 end to dynamic response. Nat. Rev. Mol. Cell Biol. 16: 207–220.
549 https://doi.org/10.1038/nrm3935
550 Northam M. R., H. A. Robinson, O. V. Kochenova, and P. V. Shcherbakova, 2010
551 Participation of DNA polymerase ?? in replication of undamaged DNA in
552 Saccharomyces cerevisiae. Genetics 184: 27–42.
553 https://doi.org/10.1534/genetics.109.107482
554 O'Donovan A., D. Scherly, S. G. Clarkson, and R. D. Wood, 1994 Isolation of active
555 recombinant XPG protein, a human DNA repair endonuclease. J. Biol. Chem.

556 269: 15965–15968.
557 O'Driscoll M., P. Macpherson, Y. Z. Xu, and P. Karran, 1999 The cytotoxicity of
558 DNA carboxymethylation and methylation by the model carboxymethylating
559 agent azaserine in human cells. Carcinogenesis 20: 1855–1862.
560 https://doi.org/10.1093/carcin/20.9.1855
561 Parreiras L. S., L. M. Kohn, and J. B. Anderson, 2011 Cellular Effects and Epistasis
562 among Three Determinants of Adaptation in Experimental Populations of
563 Saccharomyces cerevisiae ??? Eukaryot. Cell 10: 1348–1356.
564 https://doi.org/10.1128/EC.05083-11
565 Sabarinathan R., L. Mularoni, J. Deu-Pons, A. Gonzalez-Perez, and N. Lopez-Bigas,
566 2016 Nucleotide excision repair is impaired by binding of transcription factors to
567 DNA. Nature 532: 264–267. https://doi.org/10.1038/nature17661
568 Sarkar S., W. T. Ma, and G. H. Sandri, 1992 On fluctuation analysis: a new, simple
569 and efficient method for computing the expected number of mutants. Genetica
570 85: 173–179. https://doi.org/10.1007/BF00120324
571 Schacherer J., J. A. Shapiro, D. M. Ruderfer, and L. Kruglyak, 2009 Comprehensive
572 polymorphism survey elucidates population structure of Saccharomyces
573 cerevisiae. Nature 458: 342–345. https://doi.org/10.1038/nature07670
574 Smith E. N., and L. Kruglyak, 2008 Gene-environment interaction in yeast gene
575 expression. PLoS Biol. 6: 810–824. https://doi.org/10.1371/journal.pbio.0060083
576 Sommer D., C. M. Stith, P. M. J. Burgers, and R. S. Lahue, 2008 Partial reconstitution
577 of DNA large loop repair with purified proteins from Saccharomyces cerevisiae.
578 Nucleic Acids Res. 36: 4699–707. https://doi.org/10.1093/nar/gkn446
579 Sonnhammer E. L. L., S. R. Eddy, and R. Durbin, 1997 Pfam: A comprehensive
580 database of protein domain families based on seed alignments. Proteins Struct.
581 Funct. Genet. 28: 405–420. https://doi.org/10.1002/(SICI)1097-
582 0134(199707)28:3<405::AID-PROT10>3.0.CO;2-L
583 Sparks J. L., H. Chon, S. M. Cerritelli, T. A. Kunkel, E. Johansson, *et al.*, 2012 RNase
584 H2-Initiated ribonucleotide excision repair. Mol. Cell 47: 980–986.
585 https://doi.org/10.1016/j.molcel.2012.06.035
586 Steinmetz L. M., H. Sinha, D. R. Richards, J. I. Spiegelman, P. J. Oefner, *et al.*, 2002
587 Dissecting the architecture of a quantitative trait locus in yeast. Nature 416: 326–
588 330.
589 Sun Y., and R. E. Moses, 1991 Reactivation of psoralen-reacted plasmid DNA in
590 Fanconi anemia, xeroderma pigmentosum, and normal human fibroblast cells.
591 Somat Cell Mol Genet 17: 229–238.
592 Supek F., and B. Lehner, 2015 Differential DNA mismatch repair underlies mutation
593 rate variation across the human genome. Nature 521: 81–84.
594 https://doi.org/10.1038/nature14173
595 Swinnen S., K. Schaerlaekens, T. Pais, J. Claesen, G. Hubmann, *et al.*, 2012
596 Identification of novel causative genes determining the complex trait of high
597 ethanol tolerance in yeast using pooled-segregant whole-genome sequence
598 analysis. Genome Res. 22: 975–84. https://doi.org/10.1101/gr.131698.111
599 Tkach J. M., A. Yimit, A. Y. Lee, M. Riffle, M. Costanzo, *et al.*, 2012 Dissecting
600 DNA damage response pathways by analysing protein localization and
601 abundance changes during DNA replication stress. Nat. Cell Biol. 14: 966–976.
602 https://doi.org/10.1038/ncb2549
603 Torres-Ramos C. A., S. Prakash, and L. Prakash, 2002 Requirement of RAD5 and
604 MMS2 for postreplication repair of UV-damaged DNA in Saccharomyces
605 cerevisiae. Mol. Cell. Biol. 22: 2419–2426. https://doi.org/10.1128/MCB.2419

606 Tran P. T., J. P. Fey, N. Erdeniz, L. Gellon, S. Boiteux, *et al.*, 2007 A mutation in
607 EXO1 defines separable roles in DNA mismatch repair and post-replication
608 repair. DNA Repair (Amst). 6: 1572–1583.
609 https://doi.org/10.1016/j.dnarep.2007.05.004
610 Unk I., I. Hajdú, A. Blastyák, and L. Haracska, 2010 Role of yeast Rad5 and its
611 human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair
612 (Amst). 9: 257–267.
613 Vermut M., W. R. Widner, J. D. Dinman, and R. B. Wickner, 1994 XIV. Yeast
614 sequencing reports. Sequence of MKT1, needed for propagation of M2 satellite
615 dsRNA of the L-A virus of Saccharomyces cerevisiae. Yeast 10: 1477–1479.
616 https://doi.org/10.1002/yea.320101111
617 Wang X., and L. Kruglyak, 2014 Genetic basis of haloperidol resistance in
618 Saccharomyces cerevisiae is complex and dose dependent. PLoS Genet. 10.
619 https://doi.org/10.1371/journal.pgen.1004894
620 Wickner R. B., 1980 Plasmids controlling exclusion of the K2 killer double-stranded
621 RNA plasmid of yeast. Cell 21: 217–226. https://doi.org/10.1016/0092-
622 8674(80)90129-4
623 Zhu Y. O., M. L. Siegal, D. W. Hall, and D. a Petrov, 2014 Precise estimates of
624 mutation rate and spectrum in yeast. Proc. Natl. Acad. Sci. U. S. A. 1–9.
625 https://doi.org/10.1073/pnas.1323011111

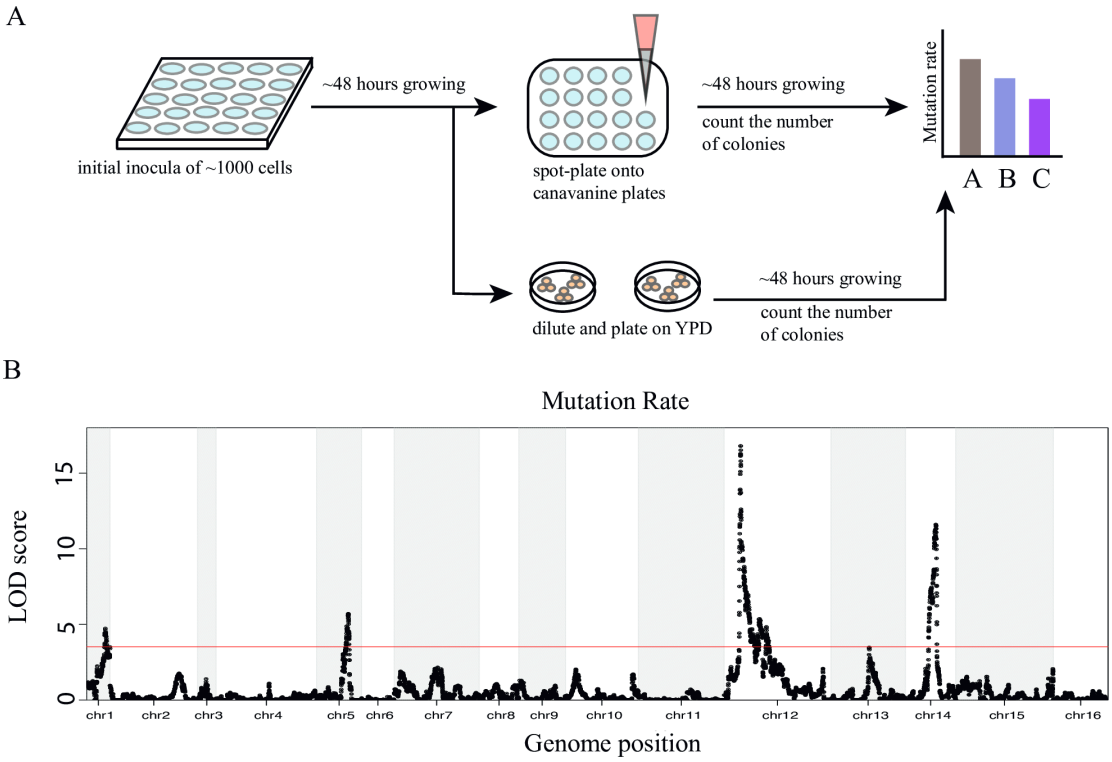

Figure 1. Linkage analysis identified four loci underlying mutation rate variation.

(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).

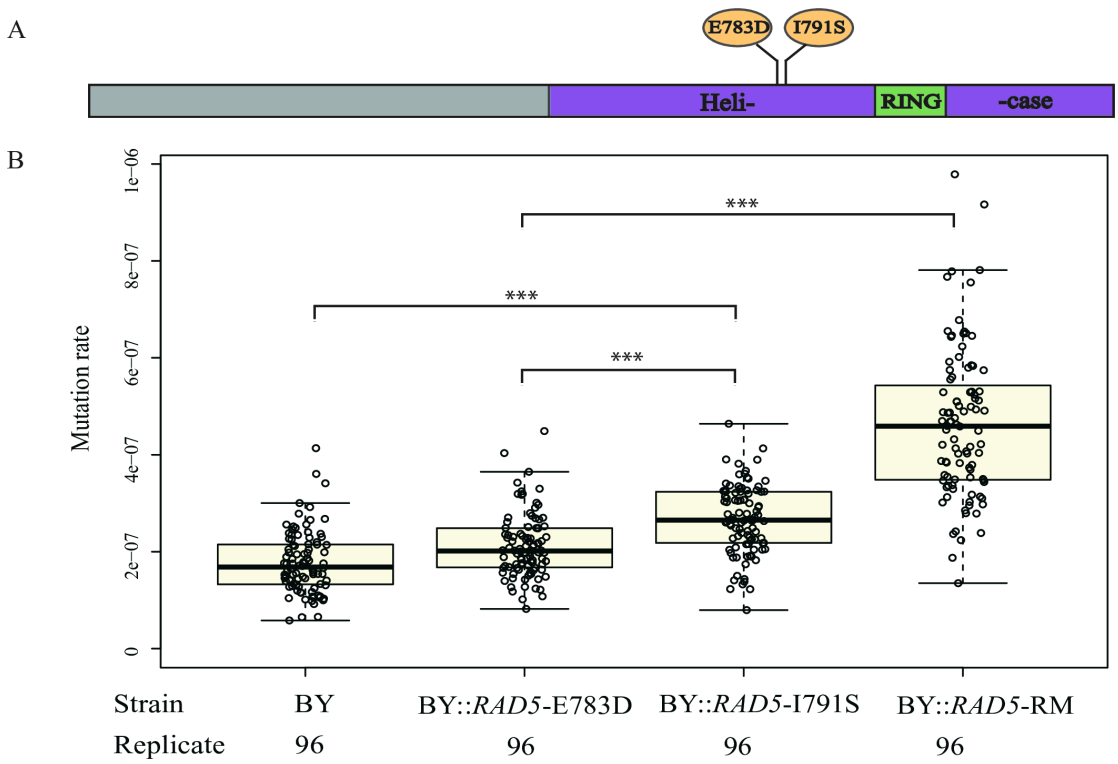

Figure 2. Polymorphisms in *RAD5* underlie mutation rate variation.

(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 ***.

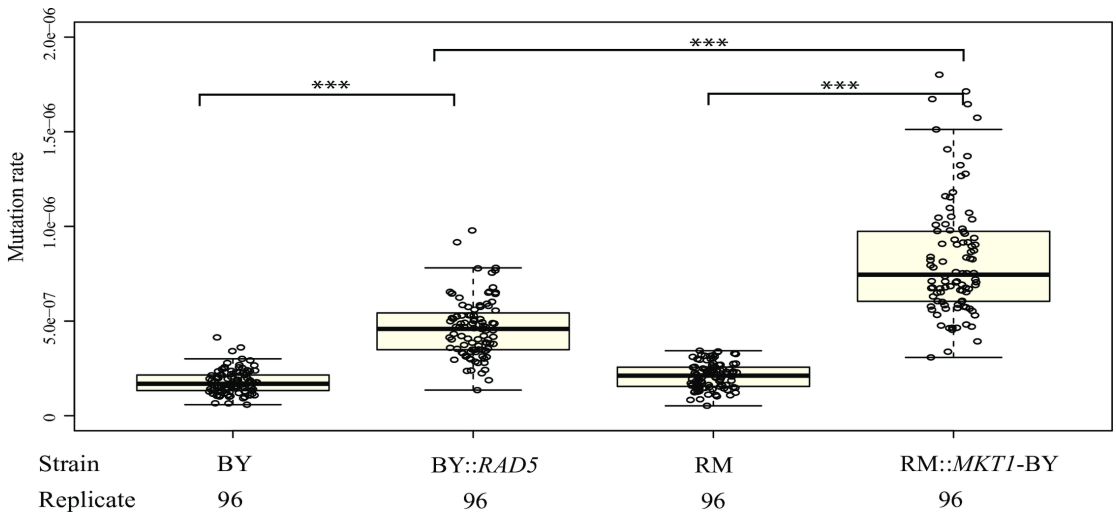

Figure 3. The RM allele of *RAD5* and BY allele of *MKT1* increase mutation rate. The mutation rate of two allele replacement strains, the BY strain and the RM strain are measured and compared. For each strain, ninety-six replicate measurement for mutation rate was performed. Bold lines show the mean. Boxes show the interquartile range. Statistical significance was tested using permutation t-test. Permutation p value < 0.001 is shown as ***.

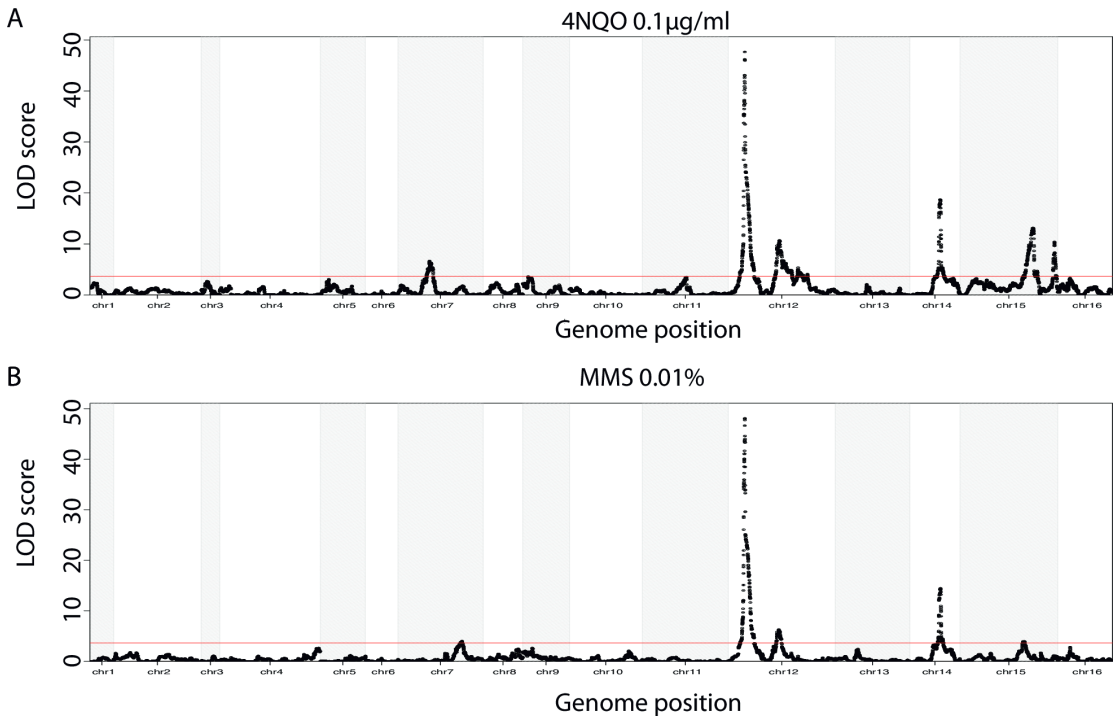

Figure 4. Loci underlying mutation rate variation, 4NQO sensitivity and MMS sensitivity are overlapped.

(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).

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

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