Genome architecture shapes evolutionary adaptation to DNA replication stress

Marco Fumasoni*, Andrew W. Murray*

Department of Molecular and Cellular Biology, Harvard University, Cambridge, United States

ABSTRACT

Evolutionary adaptation to perturbations in DNA replication follows reproducible trajectories that lead to changes in three important aspects of genome maintenance: DNA replication, the DNA damage checkpoint, and sister chromatid cohesion. We asked how these trajectories depend on a population's genome architecture by testing whether ploidy or the ability to perform homologous recombination influence the evolutionary fate of the budding yeast, *Saccharomyces cerevisiae*, as it adapts to constitutive DNA replication stress, a condition that characterizes many cancer cells. In all three genome architectures, adaptation happens within 1000 generations at rates that are linearly correlated with the initial fitness defect of the ancestors. Which genes are mutated depends on the frequency at which mutations occur and the selective advantage they confer. The recombination-deficient strain amplifies adaptive chromosomal regions less often, whereas the selective advantage of loss-of-function mutations, such as those that inactivate the DNA damage checkpoint, is reduced in diploids because of the presence of a second, wild-type copy of the gene. Despite these differences, selection targets the same three functional modules in all three architectures, suggesting that genome architecture controls which genes are mutated but not which modules are modified.

*For correspondence: marcofumasoni@fas.harvard.edu, awm@mcb.harvard.edu

INTRODUCTION

Organisms and cells adapt to selective pressures through the acquisition of beneficial mutations. The interactions between beneficial mutations that produce evolved phenotypes are complex and difficult to determine both in laboratory and natural evolution. These interactions are also likely to depend on two factors. The first, genetic background, refers to the collection of individual sequence variants in a strain or population [1–4]. The second is genome architecture, which we use to refer to functional features of the genome including its ploidy, rates of recombination, and the activity of mobile genetic elements. Previous work has revealed that different genetic backgrounds can select different spectra of mutations and different genome architectures lead do different rates of adaptation [5–14] but we lack a general, mechanistic explanation of these differences. Here, we take advantage of our earlier work on adaptation to DNA replication stress to investigate how different genome architectures control adaptive evolutionary trajectories.

We previously characterized the evolutionary adaptation of haploid cells of the budding yeast, *Saccharomyces cerevisiae*, experiencing constitutive DNA replication stress [15], a perturbation of DNA synthesis that produces lesions in DNA, increases genome instability, and decreases fitness [16–19]. DNA replication stress is often present in cancer cells as a consequence of oncogene activation [20–23] and can also be generated by the unrestrained proliferation of selfish genetic elements such as viruses or transposons [24–27]. In haploid *S. cerevisiae*, evolutionary adaptation to constitutive DNA replication stress is driven by mutations affecting three different functional modules that contribute to genome maintenance: gain-of-function mutations that stabilize the linkage between sister chromatids, mutations that improve the completion of DNA replication, most likely by stabilizing the progression of replication forks, and mutations that inactivate the DNA damage checkpoint, thus accelerating the cell division cycle. Full adaptation to the defects imposed by DNA replication stress requires a combination of mutations in all three modules, which are acquired in an order dictated by the epistatic interactions between the mutations [15].

We asked how genome architecture affected evolutionary adaptation to constitutive DNA replication stress by testing two additional architectures: i) diploid cells, which carry two copies of every chromosome and thus make fully recessive mutations phenotypically silent, ii) haploid, recombination-deficient cells which cannot produce gene amplifications or deletions that depend on homologous recombination between repeated sequences [28,29]. These architectures are relevant for cancer biology: cancer progression begins in diploid cells and defects in homologous recombination have been implicated in predisposition to tumorigenesis (i.e. BRCA2 mutations, [30,31]) and exploited for anticancer therapies [32–34].

We evolved cells with these two genome architectures and compared their responses with those of haploid, recombination-proficient cells. All three architectures experienced constitutive DNA replication stress caused by removing Ctf4 [15], a protein that physically links different activities at replication forks [35]. Adaptation to DNA replication stress occurs in all three genome

62 architectures at rates that are correlated with their initial fitness defects. Whole genome
63 sequencing of the evolved populations revealed diverse mutational spectra with only modest
64 convergence among the three different architectures. Despite the differences in which genes
65 were mutated, all three architectures accumulated mutations in the DNA replication machinery,
66 the DNA damage checkpoint, and sister chromatid cohesion. Individually engineering a subset of
67 these mutations into all three architectures demonstrated their ability to reduce the fitness cost
68 of DNA replication stress, even in architectures where they didn't appear. Combining mutations
69 in all three modules produced fitness increases that matched those of the evolved strains in all
70 three architectures. We conclude that although differences in the observed frequency of
71 mutations in different genes depend on genome architecture, the principles that govern
72 adaptation to constitutive DNA replications stress are largely architecture-independent. We
73 discuss the implication of our results for tumor evolution and for generating inter-species
74 molecular diversity in genome maintenance.

75

76 **RESULTS**

77 **All populations adapt to constitutive DNA replication stress.**

78 To induce constitutive DNA replication stress in *S. cerevisiae*, we deleted the *CTF4* gene, which
79 encodes a non-essential subunit of the DNA replication machinery (the replisome) [36]. Ctf4 is a
80 homo-trimer that functions as a structural hub within the replisome [35,37] by binding to the
81 replicative DNA helicase, primase (the enzyme that makes the RNA primers that initiate DNA
82 replication), and other accessory factors [35,38–40]. In the absence of Ctf4, cells experience DNA
83 replication stress [41–43] and have substantially reduced reproductive fitness [15]. We removed
84 *CTF4* from wild-type (WT) haploid and diploid cells, as well as from haploids impaired in
85 homologous recombination due to the deletion of *RAD52* (Figure 1A), which encodes a conserved
86 enzyme required for pairing homologous DNA sequences during recombination [44]. Attempts
87 to generate homozygous *ctf4Δ/ctf4Δ rad52Δ/rad52Δ* mutants failed due to the lethality of the
88 resulting zygotes. We refer to the stains we analyzed as haploids (*ctf4Δ,* yellow), diploids
89 (*ctf4Δ/ctf4Δ,* blue) and recombination-deficient (*ctf4Δ rad52Δ,* red). Compared to the previously
90 reported *ctf4Δ* haploid strains, which manifested a fitness decrease of ~27% relative to isogenic
91 wild-type cells [15], constitutive DNA replication stress caused by the absence of Ctf4 was more
92 detrimental in diploid (~35% reduced fitness) and recombination-deficient strains (~55% reduced
93 fitness) (Figure 1B). We evolved eight parallel diploid and recombination-deficient populations
94 for 1000 generations in the same conditions and at the same time as the haploid populations we
95 previously analyze ([15] and see Materials and methods). In this regime, spontaneous mutations
96 that increase fitness and survive genetic drift will spread through the population [45,46]. In
97 addition, we evolved eight wild-type diploid populations like the previously evolved wild-type
98 haploids, to control for mutations that are selected by the culture conditions rather than by
99 ameliorating DNA replication stress [15].

After 1000 generations experiencing constitutive DNA replication stress, we asked how the fitness of the evolved diploid and recombination-deficient populations compared to those of evolved haploids. Regardless of their genome architecture, all populations lacking Ctf4 increased in fitness compared to their respective ancestors: haploids from -27% to -9.6 ± 1.2% [15], diploids from -35% to -14.4 ± 1.8%, and recombination-deficient haploids from -55% to -19 ± 1% (Figure 1B). The fitness of four clones isolated from each evolved population were similar to the fitness of the population they were derived from (Figure 1-S1). Control evolved wild-type populations showed only small fitness increases: 4 ± 1% for haploids and 6.7 ± 2.7% for diploids (Figure 1-S2).

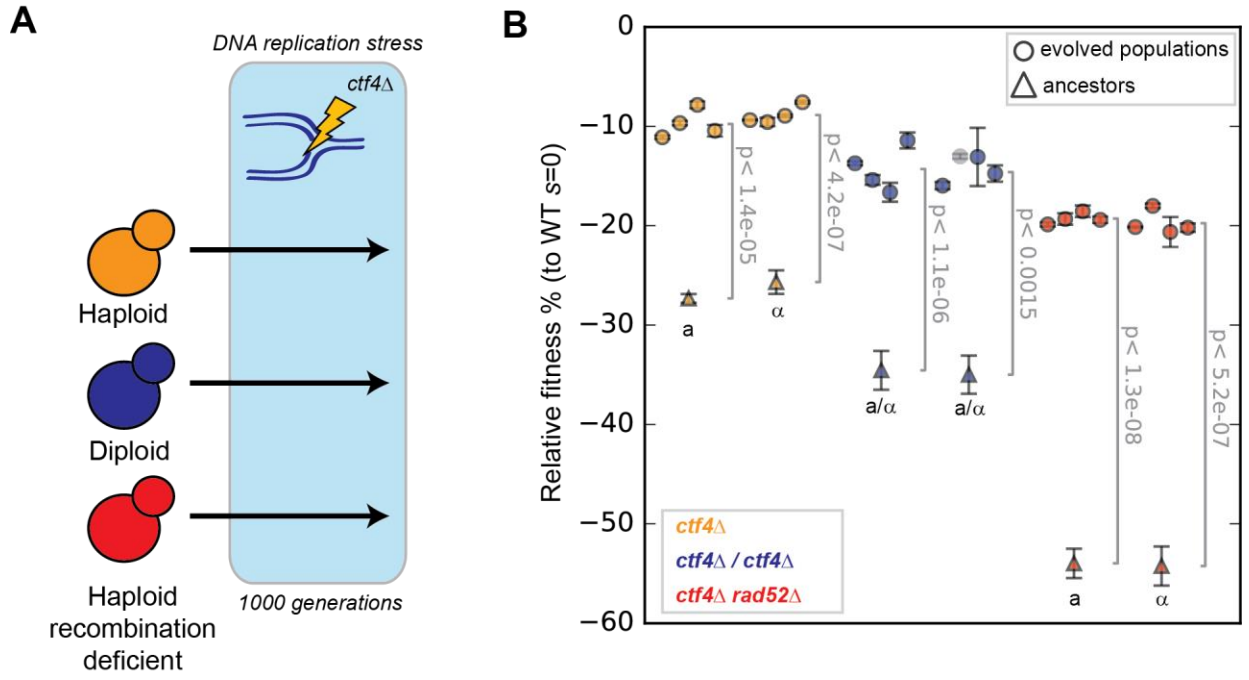


Figure 1: Evolutionary adaptation to constitutive DNA replication stress in different genome architectures. (A) Experimental scheme: Constitutive DNA replication stress was induced by removing Ctf4, a protein that coordinates replisome activities. Haploid (orange), diploids (blue) and cells with a severely decreased ability to perform homologous recombination (red, recombination-deficient), were subjected to 1000 generations of experimental evolution. **(B)** The fitness of the 6 ancestral strains (2 for each genome architecture) and of 24 evolved populations derived from them, relative to WT cells (s = 0). Fitness data of haploid strains (orange) is from [15]. The semi-transparent gray dot represents a diploid strain that became haploid during the experiment and is thus excluded from subsequent analysis. Error bars represent standard deviations. a and α refer to the strains' mating type (*MAT* locus). a/α indicates diploid strains. The fitness values shown here are reported in Source data 1.

The rate of adaptation to DNA replication stress depends on the ancestor's fitness

Although all populations increased their fitness in response to constitutive DNA replication stress, strains with different genome architectures increased in fitness to different extents (measured as the absolute fitness gain of populations over 1000 generations). Recombination-deficient populations showed the largest fitness gain (+34.6 ± 0.8%), followed by diploid (+20.3 ± 1.6%) and haploid populations (+17.2 ± 1.0%, Figure 2A). We asked what could have generated different

adaptation rates in response to the same selective pressures. Less fit ancestral strains have been proposed to increase in fitness faster than more fit ancestors, a phenomenon referred as 'declining adaptability' [47–50]. To test for this possibility, we plotted the final fitness of all evolved populations against the fitness of their respective ancestors (Figure 2C). We found that the absolute fitness gain over 1000 generations depended linearly on the ancestors' relative fitness (R^2 = 0.95). These results agree with previous studies, which observed declining adaptability across a wide range of organisms and selection conditions [9,10,51,52].

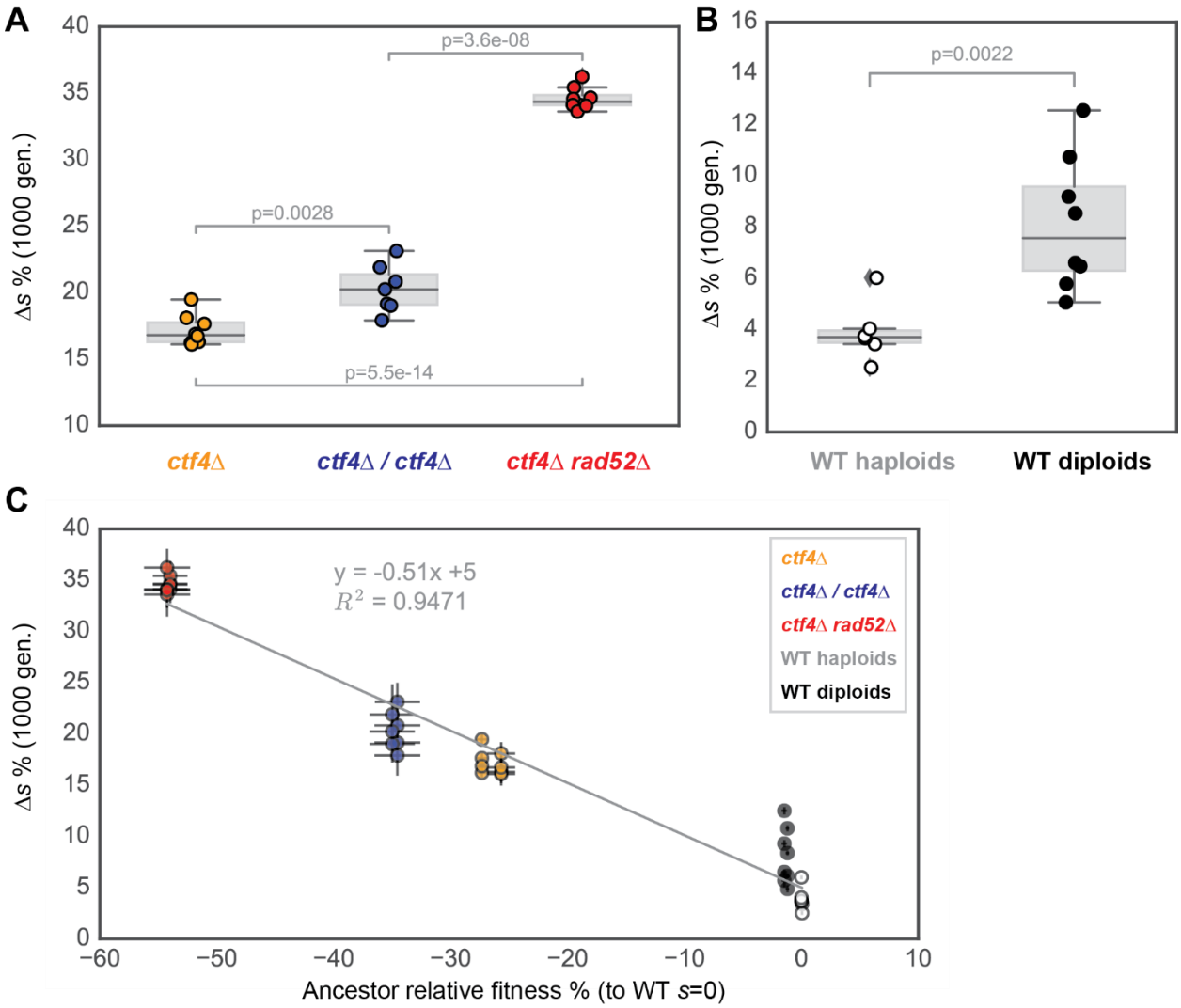


Figure 2: Fitness increase over 1000 generations. (A) The fitness increase over 1000 generations for strains experiencing constitutive DNA replication stress. Δ*s* is the difference between the fitness of the evolved populations and the fitness of their ancestors, with both fitnesses measured relative to that of cells that contain Ctf4 (*CTF4*) and have the same ploidy. **(B)**. The fitness increase over 1000 generations for strains that contained wild-type *CTF4*. **(C)** shows the correlation between the fitness increase during evolution (Δ*s*) and the fitness defect of the ancestors relative to cells of the same architecture containing Ctf4 (the ancestor relative fitness). Fitness data of haploid strains (orange and white) is from [15]. Error bars represent standard deviations. Grey line represents linear regression between the datapoints**.** The data shown here are reported in or derived from Source data 1.

141 The majority of previous studies have suggested that diploids adapt more slowly than isogenic
142 haploids [13,53,54]. Our diploid *ctf4Δ/ctf4Δ* populations gained more fitness than *ctf4Δ* haploids
143 over the course of 1000 generations (Figure 2A). Wild-type diploids increased more in absolute
144 fitness than wild-type haploids (Figure 2B), suggesting that the discrepancy between our work
145 and previous studies is not due to the presence of replication stress. Diploid populations
146 recovered a slightly smaller fraction of their initial fitness defect compared to haploid and
147 recombination-deficient strains (Figure 2-S2A). Collectively, our results suggest that ploidy is a
148 poor predictor of the speed and extent of adaptation (Figure 2-S1). We conclude that the fitness
149 of the ancestral cells is the major determinant influencing how much populations adapt, adding
150 further support to the concept of 'declining adaptability'.

151 **Limited parallelism in adaptively mutated genes across genome architectures.**

152 We used whole-genome sequencing to understand the genetic basis of the evolutionary
153 adaptation to DNA replication stress. For each different genome architecture, we sequenced all
154 eight evolved populations, as well as four individual clones isolated from each of them. A total of
155 195, 276 and 905 genes or associated regulatory sequences were found mutated in evolved
156 haploid, diploid, and recombination-deficient populations respectively (Figure 3A,
157 Supplementary table 1). When normalized for DNA content, haploid (13 mutations/haploid
158 genome) and diploid clones (9 mutations/haploid genome) had similar but statistically different
159 mutation frequencies (Figure 3-S1) but recombination-deficient clones carried many more
160 mutations, both in total (50/haploid genome, Figure 3-S1A) and as synonymous mutations
161 (7/haploid genome, Figure 3-S1B). Our results agree with the reported increase in spontaneous
162 mutagenesis in *rad52Δ* cells [55–57].

163 We used statistical analysis to identify candidate adaptive mutations, by identifying genes that
164 were mutated more often than expected by chance in independent populations (see Materials
165 and methods for details). We performed this analysis both within individual genome
166 architectures and across all three architectures. Using a false discovery rate of 5% we found a
167 total 50 genes that met one or both criteria.

168 Only 8 of these genes were mutated in all three genome architectures (Figure 3A). This small
169 number could be due to different adaptive strategies adopted by the different architectures,
170 differences in the rate for the same mutation in different architectures, differences in the
171 selective benefit the mutation conferred, or a combination of these three effects. We found 28
172 genes with putative adaptive mutations that are specific to individual genome architectures, as
173 well as 14 genes with putative adaptive mutations in two architectures (Figure 3A, genes
174 highlighted in boxes and Supplementary table 2). Analyzing gene ontology (GO) terms associated
175 with the putative adaptive mutations revealed an enrichment of biological process terms
176 associated with genome maintenance (Supplementary table 3). The genes that were significantly
177 mutated in all genomic architectures included two that we had previously demonstrated to be
178 adaptive in haploid *ctf4Δ* cells: mutations in *RAD9* inactivate the DNA damage checkpoint and
179 those in *SLD5* alter the function of the replicative helicase [15]. All three architectures also had

mutations in *RAD61* and *PDS5*, which are implicated in chromosome segregation [58,59], a cellular module shown to play a role during the adaptation of haploid *ctf4Δ* populations [15].

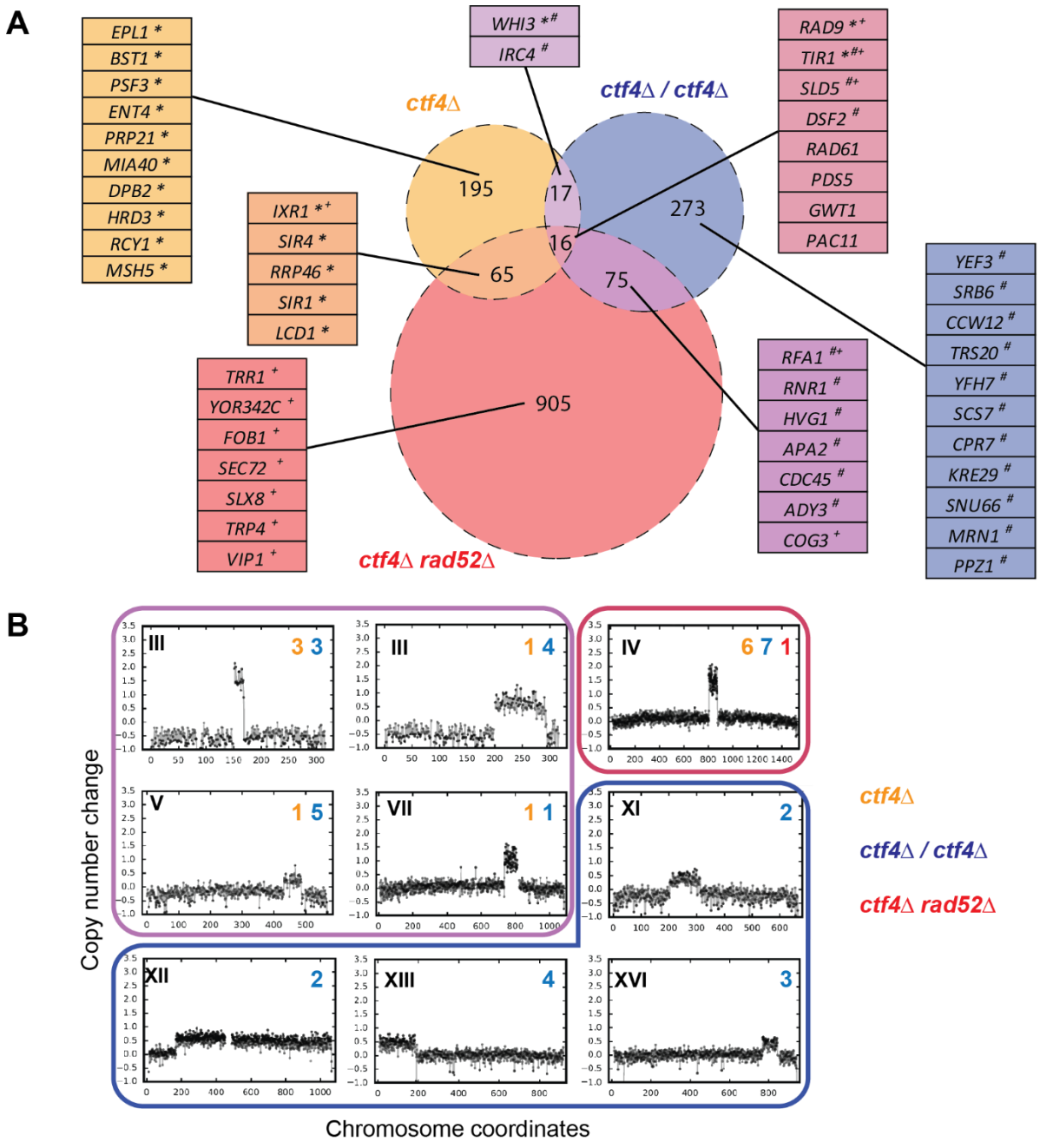


Figure 3: Putative adaptive mutations in the evolved clones. (A) Venn diagram representing the mutations found in the evolved clones and populations. The area of the circles is proportional to the list of genes mutated at least once in the set of evolved clones for each genome architecture. The number within each sector represents number of genes which were mutated in one, two, or three genetic architectures. Tables include the genes whose mutations were significantly mutated in either global or architecture-specific analysis. Symbols next to the genes highlight that they were statistically significantly mutated in individual architectures (* refers to haploid, # to diploid and + to recombination-deficient architectures, e.g. *SLD5* $^{+\#}$ means that the gene is found mutated in all three architectures, but is statistically significantly mutated only in diploid and recombination-deficient populations). The identities of all

191 the genes in the Venn diagram intersections are reported in Source data 3 **(B)** Copy number variations (CNVs)
192 affecting different chromosomes (roman numbers) which appeared in at least two independent populations. Arabic
193 numbers are the number of independent populations, in each genetic architecture (highlighted by a different color),
194 in which the CNVs were detected. Copy number change refers to the fragment's gain or loss during the evolution
195 experiment (i.e. +1 means that one copy was gained in haploid cells and that two copies were gained in diploid cells).
196 Frames around the plots have the same color as the sectors in the Venn diagram, and encircle CNVs that have been
197 found in one, two, or three genetic architectures. All the segmental amplifications detected are reported in
198 Supplementary table 4.

199 We also examined the evolved genomes for changes in gene copy numbers (Copy Number
200 Variations, CNVs) that might have played a role in adaptation. Many clones displayed an
201 increased copy number of defined chromosome segments (segmental amplification) compared
202 to the control populations (Supplementary table 4). Segmental amplifications were common in
203 diploid clones (Figure 3-S1C). We inferred that segmental amplifications were putatively adaptive
204 if the same chromosome region was found amplified in two clones isolated from independent
205 populations (Figure 3B). Among these segmental amplifications, four CNVs were found only in
206 evolved diploids, four were found in at least an evolved haploid and a diploid clone, and only one
207 was detected in populations of all three genome architectures (Figure 3B). This latter fragment is
208 located on chromosome IV and contains *SCC2*, which encodes a subunit of the cohesin loader
209 complex. This essential complex is responsible for loading the proteinaceous cohesin ring that
210 holds sister chromatids together until they segregate from each other during mitosis [60–62].
211 The gene encoding for the other subunit of the cohesin loader complex, *SCC4*, lies on the segment
212 of chromosome V that was amplified in both haploid and diploid clones. We previously showed
213 that extra copies of these genes alleviate the cohesion defects associated with the absence of
214 Ctf4 in haploid cells and increase their fitness [15].

215 Overall, genome sequencing reveals that the identity of the genes carrying selected mutations
216 varies depending on genome architecture. Nevertheless, we found evidence of common
217 evolutionary strategies across different populations: Mutations affecting sister chromatid
218 cohesion, DNA replication, and the DNA damage checkpoint, which were previously shown to
219 drive adaptation to DNA replication stress in haploid cells, are significantly mutated regardless of
220 the genome architecture of the ancestral cells.

221 **Evolved diploids show adaptive loss of heterozygosity**

222 We analyzed the frequency of mutated reads in evolved diploid clones to estimate whether
223 mutations were homozygous or heterozygous (Figure 4-S1). We detected 36 homozygous
224 mutations across the evolved diploid clones (Figure 4A, Figure 4-S1B, Supplementary table 5).
225 Homozygous mutations were evenly distributed across the evolved clones analyzed and thus
226 were not the product of a single unrepresentative event (Figure 4-S1C). Individual clones often
227 carried multiple homozygous mutations on the same chromosome arm (Supplementary table 5)
228 suggesting that the loss-of-heterozygosity (LOH) was likely the consequence of mitotic
229 recombination events affecting large chromosomal segments [63]. In addition to homozygous
230 mutations, three genes were found to carry two different mutations, most likely one in each of

the two copies of the gene (double hits, Figure 4A, Supplementary table 5). Ten of the homozygous or double hit mutations were abundant within their respective populations (mutation penetrance > 70%, Figure 4B and 4-S2). Gene ontology analysis of the 39 genes with homozygous mutations or double hit alleles revealed the enrichment of many terms associated with genome maintenance (Supplementary table 6). One cluster of genes, linked by genetic and physical interactions, affected components of the replisome and factors involved in the DNA damage checkpoint (Figure 4B). Because these genes had been changed by two events (either two mutations, or a mutation followed by LOH) and the altered forms had risen to high frequency in the evolved populations (Figure 4B), we argue that the inactivation or modification of many of these genes was selected for by constitutive DNA replication stress.

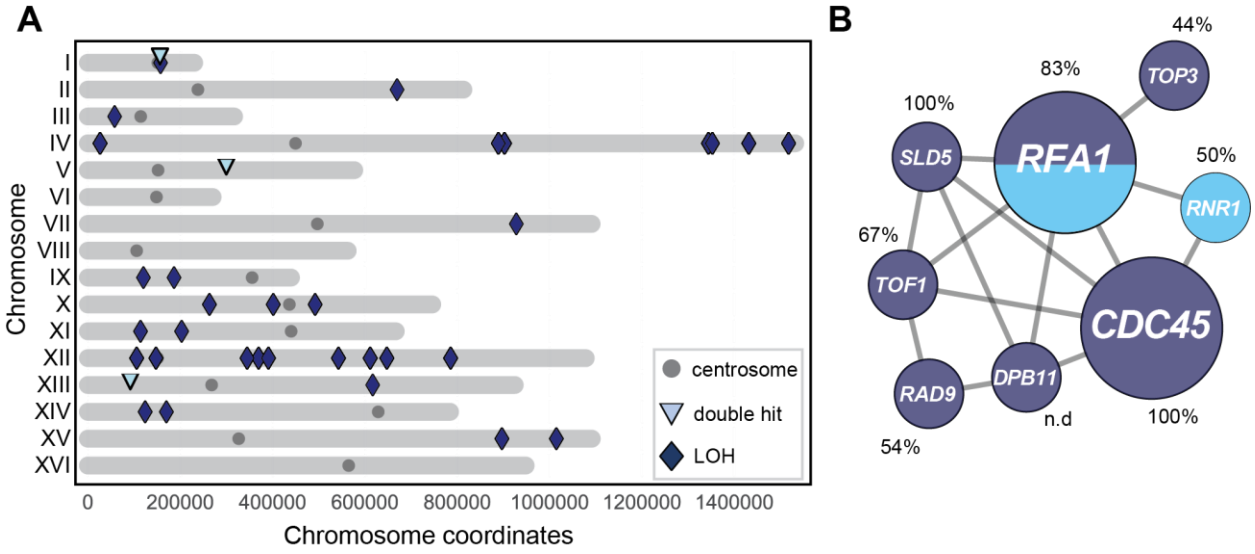


Figure 4: Homozygous mutations in diploids evolved in the presence of constitutive DNA replication stress. (A) Homozygous mutations detected in evolved *ctf4Δ/ctf4Δ* diploids clones distributed by chromosome (roman numbers). Dark blue diamonds represent loss of heterozygosity events (LOH). Light-blue triangles represent genes which contain two independent mutations (double hits) and the gray circles are the centromeres. The identities of the mutations shown here are reported in Supplementary table 5. **(B)** DNA replication and DNA damage checkpoint genes with mutations in both copies in evolved *ctf4Δ/ctf4Δ* diploid clones. Gray lines represent evidence of genetic and physical interactions from the literature (https://string-db.org). Node diameter is proportional to the number of populations in which the gene was mutated. Dark blue represents LOH events, light blue represents double hits . Percentages refer to the average estimated frequency within populations of the clones carrying the respective mutation in homozygosity (LOH) or carrying both mutations (double hits).

Genome architecture controls the distribution of adaptive mutants

In addition to altering which genes get adaptive mutations, we hypothesize that genome architecture could affect evolutionary trajectories in two ways: changing the frequency at which different modules are altered and changing the distribution of the types of mutations that are selected (for example by altering the ratio of adaptive point mutations to copy number variation [57,64–68]). Our analysis revealed that all three architectures acquired putative adaptive mutations in three genome maintenance modules: sister chromatid cohesion, DNA replication

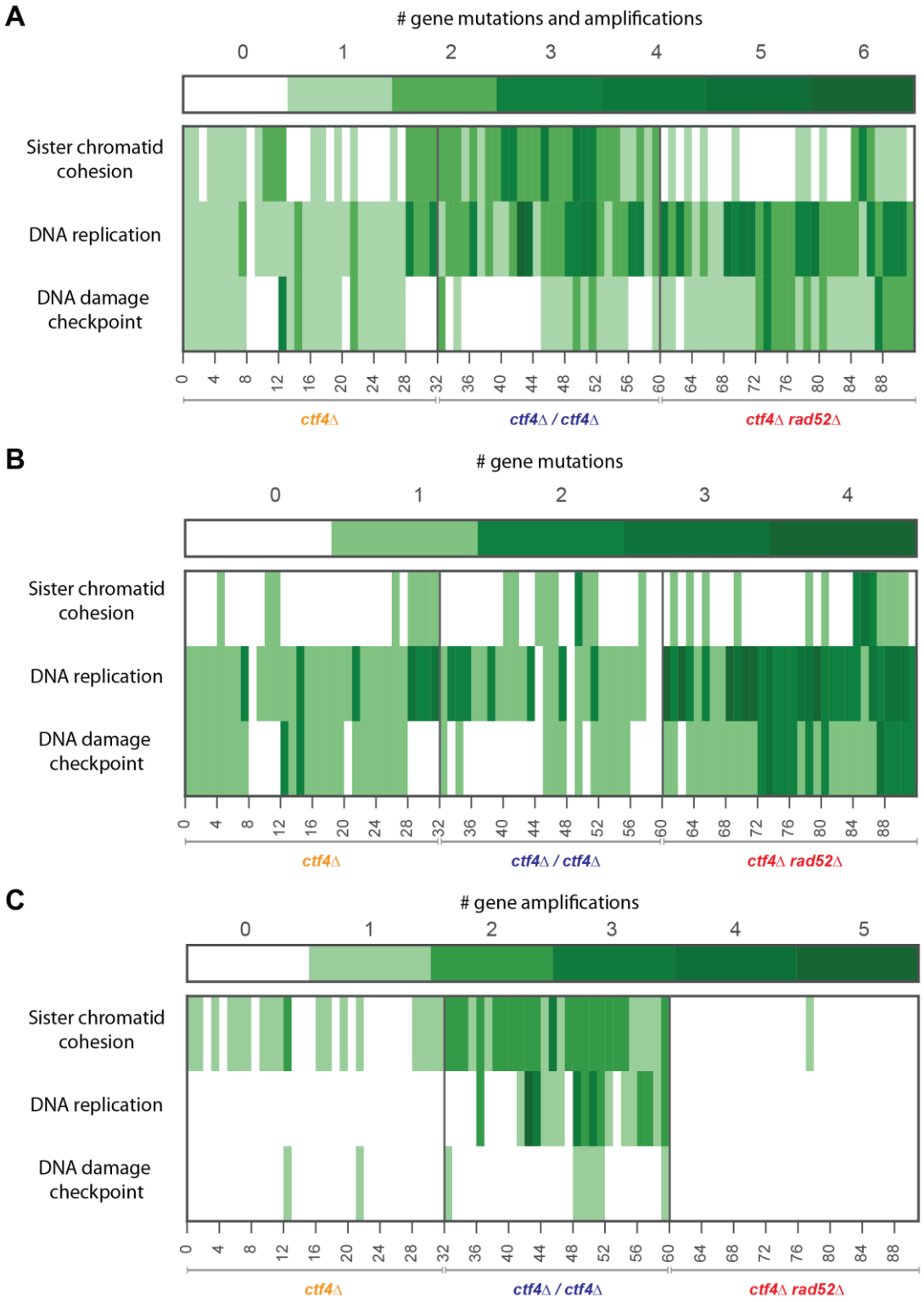
259 and the DNA damage checkpoint. Out of the 23 populations analyzed, 18 of them contained at
260 least one clone with mutations affecting all three of these modules, while all of them contain a
261 clone with mutations in two of the three modules (Figure 5A). Despite the ubiquity of mutations
262 in genome maintenance, the frequency of mutations in genes known to affect individual cellular
263 modules varied between genome architectures: mutations affecting sister chromatid cohesion
264 were underrepresented in recombination-deficient strains (p=2.7e-04, χ^2 test) and there were
265 fewer mutations affecting the DNA damage checkpoint in diploid clones (p=0.03, χ^2 test, Figure
266 6A).

267 If amplifying the genes for the cohesin loader, *SCC2* and *SCC4*, depended on homologous
268 recombination between repeated sequences, the rate of amplification should be dramatically
269 decreased in the recombination-deficient strain, which lacks Rad52. Analysis of the segmental
270 amplifications containing *SCC2* or *SCC4* revealed that their boundaries correspond to three
271 classes of repeated sequences: transposons (Ty elements), transposon-associated long terminal
272 repeats (LTRs) and tRNA genes (Figure 5-S1 and Supplementary table 7). Recombination at these
273 repeated elements has been previously shown to produce segmental amplifications selected
274 under limited nutrient availability [69–71]. Amplification of *SCC2* thus occurs as a result of
275 homologous recombination between two different copies of the repeated sequences on
276 chromosome IV, one on either side of *SCC2* and amplification of *SCC4* is due to similar events on
277 chromosome V. This dependence on homologous recombination explains the absence of
278 segmental amplifications in all but one of the recombination-deficient clones (Figure 3B and
279 Supplementary table 4). Because changes affecting sister chromatid cohesion rely predominantly
280 on gene amplifications (Figure 5C), we detect fewer mutations affecting this module in
281 recombination-deficient strains (Figure 5A).

282 In haploid *ctf4Δ* cells, inactivating the DNA damage checkpoint increases fitness [15]. The primary
283 target for mutation is *RAD9*, both because of its large size (3927 bp) and the presence of a run of
284 11 As, which repeatedly gain an extra nucleotide causing a frameshift mutation. Because this is a
285 loss-of-function mutation [15], removing Rad9 function in diploids requires two events, thus
286 reducing the probability of adapting by destroying the DNA damage checkpoint. Despite this
287 difficulty, one diploid clone is homozygous for the frameshift mutation in *RAD9*, most likely as a
288 result of a mutation in one copy of the gene followed by mitotic recombination or gene
289 conversion that homozygosed this mutation and fully inactivated the checkpoint. The need to
290 follow a two-hit pathway to fully inactivate the DNA damage checkpoint explains why fewer RAD9
291 mutations were found in diploids.

292 We conclude that although mutations affecting sister chromatid cohesion, DNA replication and
293 DNA damage checkpoint are found in all the three genome architectures, the frequency at which
294 we detect them depends on the genome architecture of an evolving population: gene
295 amplification occurs much less frequently in recombination-deficient populations, making one
296 pathway for improving sister chromatid cohesion less accessible, and recessive null mutations

297 are rarer in diploid populations, blocking a major route by which haploids inactivate the DNA
298 damage checkpoint.



299

300 **Figure 5: Mutations affecting sister chromatid cohesion, DNA replication and the DNA damage checkpoint. (A)**
301 Heatmaps representing all the mutations (SNPs, InDels and segmental amplifications) causally implicated in these

302 phenotypes in the 92 sequenced evolved clones **(B)** Frequency of SNPs and small InDels. **(C)** Frequency of genes
303 associated with the three processes that are present on segmental amplifications. The number of hits and the
304 identities of the genes which were mutated are reported in Source data 5.

305

306 **Mutations affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint**
307 **recapitulate evolution in all genome architectures.**

308 We asked whether genome architecture affected the benefit that mutations conferred as well as
309 the frequency at which they were observed. To test this hypothesis, we chose a group of recurring
310 mutations to validate their causality and measure their fitness benefit in different genome
311 architectures. We mimicked the segmental amplification of chromosome IV (Figure 3B), affecting
312 the cohesin loader *SCC2*, by integrating an extra copy of the gene at the *URA3* locus (referred to
313 as *2xSCC2*). Ixr1 is a transcription factor that indirectly affects DNA replication by positively
314 regulating the production of deoxyribonucleotide triphosphates (dNTPs, [72]). The abundance of
315 mutations causing premature stop codons in *IXR1* suggested that evolution selected for the
316 inactivation of this gene (Supplementary table 8). We mimicked the loss of a functional Ixr1
317 protein by gene deletion (*ixr1Δ*). Another recurrently mutated gene affecting DNA replication
318 was *SLD5*, which encodes the subunit of the main replicative helicase (CMG, [73]) that is bound
319 *in vivo* by Ctf4 [74]. *SLD5* is an essential gene and our earlier analysis suggested that mutations
320 like *sld5-E130K* reduced or altered the helicase's function [15]. Finally, the majority of mutations
321 affecting the DNA damage checkpoint factor *RAD9* resulted in early stop codons (Supplementary
322 table 9), leading us to test the effect of the *rad9Δ* gene deletion. These mutations were all
323 previously demonstrated to be adaptive in haploid cells [15] and were used because they
324 occurred, albeit at different frequencies, in all three (point mutations in *SLD5* and *RAD9* and
325 amplification of *SCC2*) or at least two genome architectures (point mutations in *IXR1* in both
326 haploid populations).

327 We found that all four mutations were adaptive in all three genomic architectures. Remarkably,
328 mutations affecting individual genes were adaptive even in architectures where they were never
329 (*IXR1* in diploids) or rarely detected (*SLD5* in haploids, *RAD9* in diploids and *SCC2* in
330 recombination-deficient strains). Three of the mutations we tested (*SCC2* amplification, *sld5-*
331 *E130K*, and *rad9Δ*) conferred similar benefits in all three architectures, but the fourth, *ixr1Δ*, gave
332 a much larger benefit in the recombination-deficient strain. One possibility is that the absence of
333 Ixr1 reduces the frequency of double-stranded breaks associated with DNA replication in the
334 absence of Ctf4. These breaks are lethal in the absence of Rad52 but can be repaired in the other
335 two architectures.

336 For the diploid *ctf4Δ/ctf4Δ* strain, we tested each mutation as a heterozygote (e.g. *IXR1/ixr1Δ*)
337 and a homozygote (e.g. *ixr1Δ/ixr1Δ*). The homozygous deletion of *IXR1* gave a substantial fitness
338 benefit (Figure 6A, dark blue), but strains heterozygous for the deletion were slightly less fit than
339 their ancestors (Figure 6A, light blue). For the other three mutations (an extra copy of *SCC2*, *sld5-*
340 *E130K*, and *rad9*), the heterozygotes conferred a fitness benefit that was significant, albeit

weaker than that of the homozygote. These results suggest that homozygous mutations can be selected by a two-step process in which both the first step, mutation in one copy of the gene, and the second, mutation in the other copy or loss of heterozygosity, confer a selective advantage. We conclude that, mutations affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint are beneficial in the presence of constitutive DNA replication stress independently of the genome architecture of the strains.

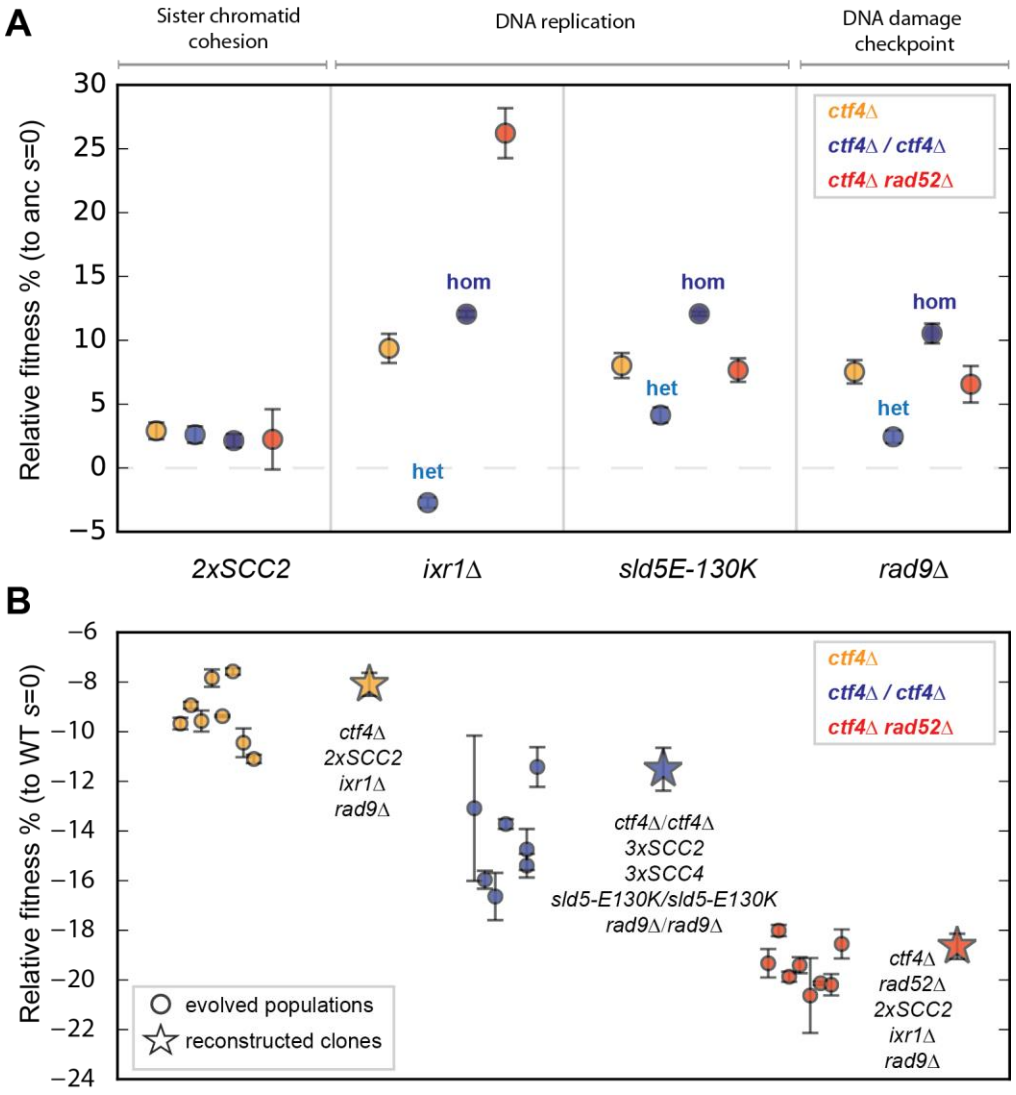


Figure 6: Reconstructed adaptive mutations can produce evolved fitness increases. (A) A subset of mutations affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint were re-constructed in the ancestor cells for all three genomic architectures. The fitness of the re-constructed strains, relative to their respective *ctf4Δ* ancestor cells (*s* = 0) is depicted. For diploid strains, we replaced one (het) or two (hom) copies of the wild-type gene with the mutant allele. The fitness values shown here are reported in Source data 6a. **(B)** Reconstructed strains carrying combinations of adaptive mutations affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint recapitulate the fitness of their respective evolved populations after 1000 generations. The most frequently mutated gene affecting DNA replication was chosen for the reconstruction in individual architectures (e.g. *IXR1* was mutated more often that *SLD5* in haploids whereas the reverse was true in

357 diploids). Fitness data of haploid strains (orange) is from [15]. The fitness values shown here are reported in Source
358 data 6b.

359 Are mutations in sister chromatid cohesion, DNA replication and the DNA damage checkpoint
360 enough to recapitulate the evolved fitness increase in all three genomic architectures? We
361 previously showed that a haploid *ctf4Δ* strain carrying a mutation in all three of these modules
362 recapitulated the fitness of the evolved strains (Figure 6B and [15]). We therefore asked if
363 mutations in these three modules were sufficient to recapitulate 1000 generations of evolution
364 in diploid and recombination-deficient strains. We constructed strains that carried mutations
365 affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint (Figure 6B),
366 using mutations that had been found in the same genome architecture (Figure 3,4 and
367 Supplementary table 1,4) and shown to be adaptive in isolation (Figure 6A).

368 For all three genome architectures, the reconstructed strains had a fitness that was similar to the
369 evolved populations with the same architecture (Figure 6B). This result demonstrates that for all
370 three genome architectures, acquiring mutations in all three modules (sister chromatid cohesion,
371 DNA replication and DNA damage checkpoint) can recapitulate the fitness increase seen in
372 response to evolution in the presence of DNA replication stress.

373

374 **DISCUSSION**

375 We previously showed that haploid cells adapt to constitutive DNA replication stress by
376 accumulating mutations affecting DNA replication, sister chromatid cohesion, and the DNA
377 damage checkpoint [15]. We repeated this experiment in two other genomic architectures, a
378 diploid and a recombination-deficient haploid strain, to ask which features of this evolutionary
379 trajectory were general and which were specific to a particular genomic architecture. We find
380 that the genes that acquire adaptive mutations, the frequency at which they are mutated, and
381 the frequency at which these mutations are selected all differ between architectures but that
382 mutations that confer strong benefits always lie in the same three modules. Engineering one
383 mutation in each module into an ancestral strain lacking Ctf4 is enough to produce the evolved
384 fitness increase in all three genomic architectures, demonstrating the existence of a common
385 pathway for yeast cells to adapt to constitutive DNA replication stress.

386 **The speed of adaptation depends on initial fitness not genome architecture**

387 What determines the rate of evolutionary adaptation? One possibility is genome architecture:
388 beneficial, recessive mutations are harder to select in diploids [53,75] and gene amplification is
389 less frequent in recombination-defective cells [76,77], suggesting that recombination proficient
390 haploids should evolve faster than the other two genome architectures. Another explanation is
391 that the initial fitness of the population is the principal factor in setting the rate of adaptation,
392 with less fit ancestral populations increasing in fitness faster than more fit ones, a phenomenon
393 referred to as declining adaptability [47]. We observed that adaptation rates scale linearly with

394 the fitness difference between a mutant ancestor and a non-mutant wild-type, showing that the
395 recovery from DNA replication stress follows the declining adaptability documented in other
396 experimental conditions.

397 Previous work has produced a variety of results with the majority reporting that increasing ploidy
398 decreases the rate of evolution [13,53,54] and a minority reaching the opposite conclusion
399 [12,14]. We suggest that this variability arises because genome architecture affects two aspects
400 of the mutations that produce adaptation: the frequency at which they occur and the selective
401 advantage they confer. Thus, we find that defects in recombination reduce the rate of gene
402 amplification and diploidy makes the accumulation of fully recessive mutations require two
403 successive genetic events rather than a single mutation. Despite these barriers, our diploids and
404 recombination-defective haploids evolved faster than recombination-proficient haploids. We can
405 only speculate on why this is so but note that the recombination-proficient strains have a higher
406 mutation rate and diploids may be more resistant to the unbalanced gene dosage caused by
407 segmental amplifications.

408 We engineered the same mutations into all three genome architectures to ask if the benefit of
409 individual mutations depended on either genome architecture or the fitness of the recipient
410 strain. For three of the four mutations we tested (*sld5-E130K*, *rad9Δ*, and 2x*SCC2*) their effect
411 was independent of both these variables. The fourth (*ixr1Δ*) showed diminishing return epistasis,
412 defined as an adaptive mutation producing a smaller benefit when present in fitter genetic
413 backgrounds [49]. This effect has been invoked to explain declining adaptability and is thought
414 to arise from the complex pattern of epistasis between the adaptive mutations and the set of
415 alleles present in the genetic background [52,78–80]. Because most of the mutations did not
416 show diminishing returns epistasis, we speculate that the mutations that are selected during
417 evolutionary repair of genetic damage are likely to have fewer and less complex interactions than
418 those that accumulate when wild type strains are selected to grow faster in variants of standard
419 laboratory environments [8,52].

420 **General principles of adaptation to constitutive DNA replication stress**

421 Evolutionary change can be examined at multiple levels from the details of adaptive mutations
422 in individual genes, the identity of the functional modules in which the mutations lie, cellular
423 phenotypes, such as the integrity of the DNA damage checkpoint, to overall reproductive fitness.
424 For each of these levels, we examined the effect of genome architecture on the adaptation to
425 constitutive DNA replication stress. At both ends of this scale we see substantial variation:
426 genome architecture controls the frequency at which mutations occur, the fitness benefit they
427 confer, and the extent of overall adaptation. In the middle, we see more conservation, with all
428 three architectures accumulating mutations in the same three functional modules, suggesting
429 that they all follow similar phenotypic paths by increasing the stability of the linkage between
430 sister chromatids, stabilizing replication forks, and inactivating the DNA damage checkpoint. To
431 test this hypothesis, we reconstructed mutations from all three modules into the ancestral, Ctf4-
432 deficient strains of all three genome architectures. This manipulation reproduced fitness

433 observed after 1000 generations of evolution leading us to conclude that the principles that
434 govern evolutionary adaptation to constitutive DNA replication stress are conserved across the
435 genome architectures we examined.

436 **Molecular mechanisms affecting adaptation in diploid cells.**

437 The phenotype of heterozygous mutations in diploid cells depends on their dominance. Although
438 loss-of-function mutations are often assumed to be fully recessive, a substantial minority have
439 detectable phenotypes as heterozygotes (referred to as semi-dominance or haploinsufficiency),
440 typically much weaker than that of the homozygous loss-of-function mutations [81]. The
441 phenotype of the heterozygote determines whether mutations will be selected for in diploid
442 populations [82,83]. As an example, loss-of-function mutations in *IXR1* were repeatedly found in
443 haploid and recombination-deficient populations but were not present in evolved diploid strains.
444 This observation is consistent with the slight fitness cost of the heterozygous *IXR1* deletion,
445 despite the strong benefit provided by the homozygous deletion: the cost of the heterozygote
446 acts as a barrier that prevents access to the beneficial homozygous mutations. Like *IXR1*, deletion
447 of both copies of *RAD9* has a strong fitness benefit in diploids, but in this case there is a small
448 fitness benefit in the *rad9Δ/RAD9* heterozygotes providing a path to inactivating the DNA damage
449 checkpoint in which both steps increase fitness. This two-step path in diploids is less favorable
450 than the one-step path that inactivates the checkpoint in haploids and likely accounts for the
451 lower frequency of mutations affecting *RAD9* and other known checkpoint genes in evolved
452 diploid clones (Figure 5 and Source data 5).

453 Two-step paths affected the replication machinery or DNA damage checkpoint in 9 of the 28
454 diploid clones; 2 involved independent mutations in the same gene and 8 produced homozygous
455 mutations due to gene conversion or mitotic recombination (Figure 4, one clone contained two
456 homozygous mutations). The high frequency of these events suggests that these mutations
457 conferred some advantage even as heterozygotes. For one mutation, *sld5-E130K*, we verified the
458 fitness benefit of the *sld5-E130K/SLD5* heterozygote, which suggests that the mutant subunit is
459 present in a fraction of the genome-wide replicative helicases, where it helps to stabilize
460 replication forks in the absence of Ctf4. Mitotic recombination can then homozygose the
461 mutation, eliminating the wild-type Sld5 protein and stabilizing all replisomes. The genome
462 architecture of diploid cells thus explains both the reduced frequency of mutations in *RAD9* and
463 dictates which mutations affecting DNA replication are ultimately selected (e.g. *IXR1* versus
464 *SLD5*). We suggest that semi-dominant mutations are an important intermediate in the selection
465 of loss-of-heterozygosity during the evolution of diploid, somatic cells.

466 **Molecular mechanisms affecting adaptation in recombination-deficient cells.**

467 Recombination-deficient clones have much fewer segmental amplifications affecting the genes
468 for the cohesin loader, *SCC2* and *SCC4* (Figure 5C). The lack of the recombinase protein Rad52 in
469 recombination-deficient strains dramatically reduces their ability to achieve segmental
470 amplifications. Only one event, likely mediated by Rad52-independent recombination [84–87],

471 was found in the recombination-deficient evolved clones (Figure 3B and Supplementary table 4).
472 The other mutations in these strains that are likely to affect sister chromatid cohesion were point
473 mutations affecting other proteins involved in pairing and segregation of chromosomes such as
474 Pds5 and Rad61 (Figure 5B). The higher number of point mutations in this and other modules is
475 likely the consequence of increased spontaneous mutagenesis imposed by the lack of homology-
476 mediated repair of spontaneously occurring DNA lesions (Figure 3-S1,2, [88]). The genome
477 architecture of recombination-deficient strains thus explains both the reduced frequency of gene
478 amplification affecting sister chromatid cohesion and the increased number of point mutations
479 detected in evolved clones (Figure 6B).

480 **Impact of genome architecture on evolutionary adaptation to DNA replication stress**

481 Our analysis argues that although genome architecture doesn't affect the ability of cells to adapt
482 to DNA replication stress or the modules that alter to produce adaptation, it does affect the
483 mutational pathways by which different architectures adapt.

484 If the major adaptive mechanisms are shared, how do we explain the limited number of genes
485 that are mutated in all conditions and the genes that are uniquely and significantly mutated in
486 specific genome architectures (Figure 3A)? There are several hypotheses: some of these genes
487 may be false positives, mutations in these genes confer benefits in more than one architecture
488 but were only detected in a single architecture, and mutations in these genes are substantially
489 more, or exclusively, beneficial in one architecture. The fact that our reconstructed strains
490 recapitulate the fitness of the evolved populations with a minimal set of changes in genome
491 maintenance suggests that architecture-specific mutations play a marginal role in adaptation. Of
492 the 22 genes that were mutated in at least two architectures, 10 have known roles in DNA
493 replication, repair, or sister chromatid cohesion, whereas only 5 of the 28 that were uniquely
494 mutated in one architecture have such a role ($p = 0.034$, χ^2 test). The fitness of the reconstructed
495 strains (Figure 6A,B) also implies that the most important mutations for the adaptation to DNA
496 replication stress are those that were found in multiple architectures (Figure 3A). Mutations that
497 were never (*ixr1*), or rarely (*2xSCC2* and *rad9*) seen in a given genome architecture, nevertheless
498 increased fitness in all three architectures. This finding suggests that caution should be used in
499 inferring adaptive strategies from lists of genes that have been mutated more often than
500 expected by chance, especially when comparing populations that were evolved under different
501 conditions or have different ancestral genotypes. We argue that experimental dissection of
502 adaptive trajectories is needed to distinguish the mutations that make major contributions to
503 adaptation from minor players and mutational noise.

504 The distribution of mutations with known phenotypic effects suggests that genome architecture
505 controls how much adaptation in different functional modules contributes to the overall
506 response to selective pressure. Mutations that break the DNA damage checkpoint by inactivating
507 *RAD9* are less frequent in diploid populations than they are in haploids, and amplifications of
508 cohesin loader genes are less frequent in recombination-deficient populations than they are in

509 recombination-proficient ones. Although we suspect that some of the diploid populations have
510 found alternate genetic routes to partially or fully inactivate the DNA damage checkpoint and
511 that some of the recombination-defective populations have selected other mutations that
512 improve sister chromatid cohesion, these changes are likely to be less effective than those that
513 were frequent in other genome architectures.

514 **Implications for cancer and natural evolution**

515 We previously speculated that evolutionary processes similar to the adaptation to constitutive
516 DNA replication stress could assist the evolution of cancer cells and organisms in nature [15].
517 Tumor cells often experience DNA replication stress as a consequence of oncogene activation,
518 promoting genetic instability and increasing the rate of cell death [17,20]. Evolutionary
519 adaptation to DNA replication stress could reduce its negative effects on cellular fitness and thus
520 assist tumor evolution. In natural populations, interference with DNA replication, through the
521 fixation of deleterious mutations, perturbation by selfish genetic elements or toxins secreted by
522 antagonistic organisms could generate similar evolutionary adaptation. The consequences of
523 these evolutionary processes could rewire genome maintenance modules to alter cancer cells'
524 physiology and generate molecular diversity in nature [89]. For these scenarios to be plausible,
525 the evolutionary adaptation to DNA replication stress needs to happen over short timescales. In
526 cancer, it would need to happen over the limited somatic evolution that accompanies cancer
527 progression. In nature, it would need to occur before the perturbed organisms are outcompeted
528 by unaffected ones or block the source of DNA replication stress. It could be argued that some
529 genome architectures, such as diploidy or reduced recombination, could reduce the likelihood of
530 evolving new strategies for genome replication and maintenance. In our hands, however, these
531 genome architectures do not preclude rapid, evolutionary adaptation to constitutive DNA
532 replication stress: loss of heterozygosity or increased mutation rates, which have both been
533 documented in cancer [90,91] and natural evolution [92,93], allow cells to circumvent these
534 obstacles and adapt quickly to a severe perturbation of the replication machinery. One rapidly
535 evolving clade of yeasts has lost *RAD9* [94], the primary target of checkpoint-inactivating
536 mutations in our experiments. Similarly, many cancers inactivate the DNA damage response
537 during tumor evolution [95,96]. We suggest that the careful examination of the genomic
538 sequences of tumors and fast-evolving species might reveal signs of adaptation that mirror the
539 molecular evolution we have described.

540 **MATERIALS AND METHODS**

541 **Strains**

542 All strains were derivatives of a modified version (Rad5^{+}) of S. cerevisiae strain W303 (*leu2-3,112*
543 *trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, RAD5*). Supplementary table 10 lists each strain's
544 genotype. The ancestors of wild-type, *ctf4Δ* and *ctf4Δ rad52Δ* strains were obtained by
545 sporulating a *CTF4/ctf4Δ RAD52/rad52Δ* heterozygous diploid. This was done to minimize the
546 selection acting on the ancestor strains before the beginning of the experiment. *CTF4/CTF4* and

547 *ctf4Δ/ctf4Δ* diploid strains were obtaining by mating freshly generated *MAT***a** and *MAT*α haploid
548 strains.

549 **Media and growth conditions**

550 Standard rich medium, YPD (1% Yeast-Extract, 2% Peptone, 2% D-Glucose) was used for the
551 evolution experiment and all fitness measurements.

552 **Experimental evolution**

553 All populations used in the evolution experiment (EVO1-40) were initially inoculated in glass
554 tubes containing 10 ml of YPD with yeast colonies of the relative genotype. Glass tubes were
555 placed in roller drums at 30˚C and grown for 24 hr. Daily passages were done by diluting 10 μl of
556 the previous culture into 10 ml of fresh YPD (1:1000 dilution, allowing for approximately 10
557 generations/cycle). All populations were passaged for a total of 100 cycles (≈1000 generations).
558 After 1000 generations four evolved clones were isolated from the populations evolved in the
559 absence of Ctf4 (EVO1-24) by streaking cells on a YPD plate (total of 96 clones isolated). All final
560 populations and clones isolated from them were mixed with 800 μl of 30% v/v glycerol and stored
561 at -70˚C for future analysis. The evolution of haploid, diploid, and recombination defective
562 strains, all lacking *CTF4*, and wild-type haploid and diploid strains were all carried out in parallel.

563 **Whole genome sequencing**

564 Genomic DNA library preparation was performed with an Illumina (RRID:SCR_010233, San Diego,
565 CA, US) Nextera DNA Library Prep Kit. Libraries were then pooled and sequenced either with an
566 Illumina HiSeq 2500 (125bp paired end reads) or an Illumina NovaSeq (150 bp paired end reads).
567 The SAMtools software package (RRID:SCR_002105, samtools.source- forge.net) was then used
568 to sort and index the mapped reads into a BAM file. GATK (RRID:SCR_ 001876,
569 www.broadinstitute.org/gatk; McKenna et al., 2010) was used to realign local indels, and VarScan
570 (RRID:SCR_006849, varscan.sourceforge.net) was used to call variants. Mutations were found
571 using a custom pipeline written in Python (RRID:SCR_008394, www.python.org). The pipeline
572 (github.com/koschwanez/mutantanalysis) compares variants between the reference strain, the
573 ancestor strain, and the evolved strains. A variant that occurs between the ancestor and an
574 evolved strain is labeled as a mutation if it either (1) causes a non-synonymous substitution in a
575 coding sequence or (2) occurs in a regulatory region, defined as the 500 bp upstream and
576 downstream of the coding sequence.

577 **Identification of putative adaptive mutations**

578 To identify genes and cellular modules targeted by selection we used the following methods:

579 *Convergent evolution on genes:* This method relies on the assumption that those genes that have
580 been mutated significantly more than expected by chance alone, represent cases of convergent
581 evolution among independent lines. The mutations affecting those genes are therefore

582 considered putatively adaptive. The same procedure was used independently on the mutations
583 found in wild-type and *ctf4Δ* evolved lines:

584 We first calculated per-base mutation rates as the total number of mutations in coding regions
585 occurring in a given background, divided by the size of the coding yeast genome in bp (including
586 1000bp per ORF to account for regulatory regions)

587 $$\lambda = \frac{SNPs + indels}{bp\text{coding}}$$

588 If the mutations were distributed randomly in the genome at a rate λ, the probability of finding
589 n mutations in a given gene of length N is given by the Poisson distribution:

590 $$\mathrm{P}\ (n\ mutations | gene\ of\ length\ N) = \frac{(\lambda N)^n e^{-\lambda N}}{n!}$$

591 For each gene of length N, we then calculated the probability of finding ≥ n mutations if these
592 were occurring randomly.

593 $$P(\neq\geq n\ mutations | gene\ of\ length\ N) = \sum_{k=n}^{\infty} \frac{(\lambda N)^n e^{-\lambda N}}{k!} = 1 - \frac{\Gamma\ (\ n+1, \lambda N)}{n!}$$

594 (Where Γ is the upper incomplete gamma function) which gives us the p-value for the comparison
595 of the observed mutations with the null, Poisson model. In order to decrease the number of false
596 positives, we then performed multiple-comparison corrections. The more stringent Bonferroni
597 correction (α=0.05) was applied on the wild-type evolved mutations dataset, while Benjamini-
598 Hochberg correction (α =0.05) was used for the *ctf4Δ* mutation dataset. Genes that were found
599 significantly selected in the evolved wild-type populations (after Bonferroni correction) were
600 removed from the list of evolved *ctf4Δ* strains (genes mutated in wild-type evolved populations
601 were removed from the *ctf4Δ* and *ctf4Δ rad52Δ* datasets and genes mutated in wild-type diploid
602 populations were removed from the *ctf4Δ/ctf4Δ* dataset). We argue that since these genes were
603 targets of selection even in wild-type cells, they are likely involved in processes that are unrelated
604 to DNA replication and are instead associated with adaptation to sustained growth by serial
605 dilutions. Supplementary table 2 lists the mutations detected in evolved clones and populations,
606 after filtering out those that occurred in genes that were significantly mutated in the respective
607 wild-type populations. Genes significantly selected are shown in dark grey (after Benjamini-
608 Hochberg correction with α=0.05).

609 *Convergent evolution on cellular modules:* Lists of genes that had been mutated more frequently
610 than expected by chance (Supplementary table 2) or found homozygously mutated in diploid
611 strains (Supplementary table 5) were further analyzed for gene ontology (GO) terms enrichment.
612 Lists of mutations were input as 'multiple proteins' in the STRING database, which reports on the
613 network of interactions between the input genes (https://string-db.org). The 'biological process'
614 terms enriched in the significantly or homozygously mutated genes are listed in Supplementary
615 tables 3 and 6 respectively.

616

617 **Fitness assays**

618 To measure relative fitness, we competed the ancestors and evolved strains against reference
619 strains. Wild-type refence strains were used for the competition assays in all the figures, except
620 for those in Figure 6A where *ctf4Δ* reference strains were used instead. A *pFA6a-prACT1-*
621 *yCerulean-HphMX4* plasmid was digested with AgeI and integrated at one of the *ACT1* loci of the
622 original heterozygous diploid (*CTF4/ctf4Δ RAD52/rad52Δ*) strain. This allows for the expression
623 of fluorescent protein yCerulean under the strong actin promoter. The heterozygous diploid was
624 then sporulated and dissected to obtain fluorescent wild-type, *ctf4Δ* or *ctf4Δ rad52Δ* reference
625 haploid strains. *ctf4Δ/ctf4Δ* diploid reference strains were obtained by mating two *ctf4Δ* haploid
626 strains, one of which expressed yCerulean under the *ACT1* promoter. For measuring a strain's
627 relative fitness, 10 ml of YPD were inoculated in individual glass tubes with either the frozen
628 reference or test strains. After 24 hrs. the strains were mixed in fresh 10 ml YPD tubes at a ratio
629 dependent on the expected fitness of the test strain compared to the reference (i.e. 1:1 if
630 believed to be nearly equally fit) and allowed to proliferate at 30°C for 24 hrs. 10 µl samples were
631 taken from this mixed culture (day 0) and the ratio of the two starting strains was immediately
632 measured. Tubes were then cultured following in the same conditions as the evolution
633 experiment by diluting them 1:1000 into fresh medium every 24hrs for 2-4 days, monitoring the
634 strain ratio at every passage. Samples were analyzed by flow cytometry (Fortessa, BD Bioscience,
635 RRID:SCR_013311, Franklin Lakes, NJ, US). Strain ratios and the number of generations occurred
636 between samples were calculated using a custom Python script utilizing the FlowCytometryTools
637 package (https://pypi.org/project/FlowCytometryTools/). Ratios, *r*, were calculated based on the
638 number of fluorescent and non-fluorescent events detected by the flow cytometer:

639 $$r = \frac{NonFluorescent_{events}}{Fluorescent_{events}}$$

640 Generations between time points, *g*, were calculated based on total events measured at time 0
641 hrs and time 24 hrs:

642 $$g = \frac{\log_{10}(Events_{t24}/events_{t0})}{\log_{10} 2}$$

643 Linear regression was performed between the $(g, \log_e r)$ points for each sample. Relative fitness
644 was calculated as the slope of the resulting line. The mean relative fitness, *s*, was calculated from
645 measurements obtained from at least three independent biological replicates. Error bars
646 represent standard deviations. The P-values reported in figures are the result of t-tests assuming
647 unequal variances (Welch's test).

648 **Copy number variations (CNVs) detection by sequencing**

649 Whole genome sequencing and read mapping was done as previously described. The read-depths
650 for every unique 100 bp region in the genome were then obtained by using the VarScan
651 copynumber tool. A custom pipeline written in python was used to visualize the genome-wide
652 CNVs. First, the read-depths of individual 100 bp windows were normalized to the genome-wide
653 median read-depth to control for differences in sequencing depths between samples. The
654 coverage of the ancestor strains was then subtracted from the one of the evolved lines to reduce

655 the noise in read depth visualization due to the repeated sequences across the genome. The
656 resulting CNVs were smoothed across five 100 bp windows for a simpler visualization. Final CNVs
657 were then plotted relative to their genomic coordinate at the center of the smoothed window.
658 Since the wild-type CNVs were subtracted from the evolved CNVs, the y axis refers to the copy
659 number change occurred during evolution (i.e. +1 means that one an extra copy of a chromosome
660 fragment has been gained in haploid cells, and that two extra copies have been gained in
661 diploids). Segmental amplifications were identified as continuous chromosomal regions affected
662 by the same change in copy number (Figure 3B and Supplementary table 4). Amplified genes
663 (Figure 5A,C) were obtained from the lists of ORFs included within the boundaries of the
664 segmental amplifications.

665 **The correlation of chromosomal features with segmental amplification boundaries**

666 We first identified the approximate boundaries (± 500bp) of all segmental amplification by
667 manually inspecting the BAM files with the Integrative Genomic Viewer
668 (https://software.broadinstitute.org/software/igv/). The coordinates of the sequence features that
669 define the boundary between lower and higher read-depths are reported in Supplementary table
670 4. We then examined the sequences within these approximate boundaries to determine whether
671 various chromosomal features were over- or underrepresented. We considered features that
672 previous studies have found to be associated with hotspots for lesions and sources of genetic
673 instability. We first counted how many times a given feature fell in the boundary zones of
674 segmental amplifications. Then we calculated the expected number of features in these zones
675 based on the total number of features in the genome and the percentage of the genome
676 represented by segmental amplification boundaries zones. We compared these numbers by $\chi 2$
677 analysis and reported the associated p-values (Supplementary table 7). The number of tRNA
678 genes, transposable elements, LTRs, ARS elements, snRNA and snoRNA genes and centromeres
679 in the genome were determined using YeastMine (https://yeastmine.yeastgenome.org/).
680 Coordinates of fragile sites were obtained from [97]. G4 sequences were obtained from [98].
681 Highly-(top 5%) and weakly-(least 5%) transcribed genes were identified from the data in [99].
682 Rrm3 binding sites and regions with high levels of γH2AX were derived from [100] and [101],
683 respectively. Sites of DNA replication termination were derived from [15]. The tandemly repeated
684 sequences, with a minimal repeat tract of twenty-four bases, were obtained from the tandem-
685 repeat-database (TRDB; https://tandem.bu.edu/cgi-bin/trdb/trdb.exe).

686 **DATA AND CODE AVAILABILITY**

687 A major dataset, containing the sequencing data used in the manuscript has been made publicly
688 available at the EBI European Nucleotide Archive (Accession no: PRJEB42058). The custom
689 pipeline used for the data analysis present in figures and supplementary tables are available on
690 GitHub: https://github.com/marcofumasoni/

691 **ACKNOWLEDGEMENTS**

692 We thank Michael Laub, María Angélica Bravo Núñez, Sriram Srikant, Thomas LaBar and Andrea
693 Giometto for critical reading of the manuscript; Andrea Giometto for assistance in data analysis;
694 Claire Hartman from the Harvard Bauer Core Facility for technical assistance. We thank the

695 members of the Murray and Nelson labs for helpful discussions. MF acknowledges fellowship
696 support from the Human Frontiers Science Program, EMBO and AIRC.

697 **AUTHOR CONTRIBUTIONS**

698 MF designed and performed the research, analyzed and interpreted the data, and wrote the
699 paper. AWM designed the research, interpreted the data, and wrote the paper.

700 **BIBLIOGRAPHY**

701 1. Chandler CH, Chari S, Dworkin I. Does your gene need a background check? How genetic
702 background impacts the analysis of mutations, genes, and evolution. Trends in Genetics. Elsevier
703 Current Trends; 2013. pp. 358–366. doi:10.1016/j.tig.2013.01.009

704 2. Dowell RD, Ryan O, Jansen A, Cheung D, Agarwala S, Danford T, et al. Genotype to phenotype: A
705 Complex problem. Science. 2010. p. 469. doi:10.1126/science.1189015

706 3. Weinreich DM, Lan Y, Jaffe J, Heckendorn RB. The Influence of Higher-Order Epistasis on
707 Biological Fitness Landscape Topography. J Stat Phys. 2018;172: 208–225. doi:10.1007/s10955-
708 018-1975-3

709 4. Fournier T, Schacherer J. Genetic backgrounds and hidden trait complexity in natural populations.
710 Current Opinion in Genetics and Development. 2017. pp. 48–53. doi:10.1016/j.gde.2017.08.009

711 5. Filteau M, Hamel V, Pouliot M, Gagnon-Arsenault I, Dubé AK, Landry CR. Evolutionary rescue by
712 compensatory mutations is constrained by genomic and environmental backgrounds. Mol Syst
713 Biol. 2015;11: 832. doi:10.15252/msb.20156444

714 6. Kempher ML, Tao X, Song R, Wu B, Stahl DA, Wall JD, et al. Effects of Genetic and Physiological
715 Divergence on the Evolution of a Sulfate-Reducing Bacterium under Conditions of Elevated
716 Temperature. MBio. 2020;11. doi:10.1128/mBio.00569-20

717 7. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in
718 an experimental population of Escherichia coli. Proc Natl Acad Sci USA. 2008;105: 7899–7906.
719 doi:10.1073/pnas.0803151105

720 8. Jerison ER, Kryazhimskiy S, Mitchell JK, Bloom JS, Kruglyak L, Desai MM. Genetic variation in
721 adaptability and pleiotropy in budding yeast. Elife. 2017;6. doi:10.7554/eLife.27167

722 9. Rojas Echenique JI, Kryazhimskiy S, Nguyen Ba AN, Desai MM. Modular epistasis and the
723 compensatory evolution of gene deletion mutants. PLOS Genet. 2019;15: e1007958.
724 doi:10.1371/journal.pgen.1007958

725 10. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, et al. The Genomic Landscape of
726 Compensatory Evolution. PLoS Biol. 2014;12: e1001935. doi:10.1371/journal.pbio.1001935

727 11. Fumasoni M. Tell Me Where You've Been and I'll Tell You How You'll Evolve. MBio. 2020;11:
728 2043–2063. doi:10.1128/mBio.02043-20

729 12. Selmecki AM, Maruvka YE, Richmond PA, Guillet M, Shoresh N, Sorenson AL, et al. Polyploidy can
730 drive rapid adaptation in yeast. Nature. 2015;519: 349–352. doi:10.1038/nature14187

731 13. Gerstein AC, Cleathero LA, Mandegar MA, Otto SP. Haploids adapt faster than diploids across a
732 range of environments. J Evol Biol. 2011;24: 531–540. doi:10.1111/j.1420-9101.2010.02188.x

733 14. Paquin C, Adams J. Frequency of fixation of adaptive mutations is higher in evolving diploid than
734 haploid yeast populations. Nature. 1983;302: 495–500. doi:10.1038/302495a0

735 15. Fumasoni M, Murray AW. The evolutionary plasticity of chromosome metabolism allows
736 adaptation to constitutive DNA replication stress. Elife. 2020;9. doi:10.7554/eLife.51963

737 16. Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nat Cell Biol. 2014;16:
738 2–9. doi:10.1038/ncb2897

739 17. Muñoz S, Méndez J. DNA replication stress: from molecular mechanisms to human disease.
740 Chromosoma. 2016;126: 1–15. doi:10.1007/s00412-016-0573-x

741 18. Branzei D, Foiani M. Maintaining genome stability at the replication fork. Nat Rev Mol Cell Biol.
742 2010;11: 208–219. doi:10.1038/nrm2852

743 19. Bermejo R, Lai MS, Foiani M. Preventing Replication Stress to Maintain Genome Stability:
744 Resolving Conflicts between Replication and Transcription. Mol Cell. 2012;45: 710–718.
745 doi:10.1016/J.MOLCEL.2012.03.001

746 20. Macheret M, Halazonetis TD. DNA replication stress as a hallmark of cancer. Annu Rev Pathol
747 Mech Dis. 2015;10: 425–48. doi:10.1146/annurev-pathol-012414-040424

748 21. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C, et al. Oncogene-induced
749 senescence is a DNA damage response triggered by DNA hyper-replication. Nature. 2006;444:
750 638–642. doi:10.1038/nature05327

751 22. Neelsen KJ, Zanini IMY, Herrador R, Lopes M. Oncogenes induce genotoxic stress by mitotic
752 processing of unusual replication intermediates. J Cell Biol. 2013;200: 699–708.
753 doi:10.1083/jcb.201212058

754 23. Kotsantis P, Petermann E, Boulton SJ. Mechanisms of Oncogene-Induced Replication Stress:
755 Jigsaw Falling into Place. Cancer Discov. 2018;8: 537–555. doi:10.1158/2159-8290.CD-17-1461

756 24. Weitzman MD, Fradet-Turcotte A. Virus DNA Replication and the Host DNA Damage Response.
757 Annu Rev Virol. 2018;5: 141–164. doi:10.1146/annurev-virology-092917-043534

758 25. Moody CA. Impact of Replication Stress in Human Papillomavirus Pathogenesis. J Virol. 2018;93.
759 doi:10.1128/jvi.01012-17

760 26. Lilley CE, Schwartz RA, Weitzman MD. Using or abusing: viruses and the cellular DNA damage
761 response. Trends in Microbiology. 2007. pp. 119–126. doi:10.1016/j.tim.2007.01.003

762 27. Scheifele LZ, Cost GJ, Zupancic ML, Caputo EM, Boeke JD. Retrotransposon overdose and genome
763 integrity. Proc Natl Acad Sci USA. 2009;106: 13927–13932. doi:10.1073/pnas.0906552106

764 28. San Filippo J, Sung P, Klein H. Mechanism of Eukaryotic Homologous Recombination. Annu Rev
765 Biochem. 2008;77: 229–257. doi:10.1146/annurev.biochem.77.061306.125255

766 29. Symington LS, Rothstein R, Lisby M. Mechanisms and regulation of mitotic recombination in
767 saccharomyces cerevisiae. Genetics. 2014;198: 795–835. doi:10.1534/genetics.114.166140

768 30. Powell SN, Kachnic LA. Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication
769 fidelity and the cellular response to ionizing radiation. Oncogene. 2003;22: 5784–5791.
770 doi:10.1038/sj.onc.1206678

771 31. Walsh CS. Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and
772 a target for ovarian cancer therapy? Gynecologic Oncology. 2015. pp. 343–350.
773 doi:10.1016/j.ygyno.2015.02.017

774 32. Chernikova SB, Game JC, Brown JM. Inhibiting homologous recombination for cancer therapy.
775 Cancer Biol Ther. 2012;13: 61. doi:10.4161/CBT.13.2.18872

776 33. Helleday T. Homologous recombination in cancer development, treatment and development of
777 drug resistance. Carcinogenesis. 2010;31: 955–960. doi:10.1093/carcin/bgq064

778 34. Krajewska M, Fehrmann RSN, De Vries EGE, van Vugt MATM. Regulators of homologous
779 recombination repair as novel targets for cancer treatment. Frontiers in Genetics. 2015. p. 96.
780 doi:10.3389/fgene.2015.00096

781 35. Villa F, Simon AC, Ortiz Bazan MA, Kilkenny ML, Wirthensohn D, Wightman M, et al. Ctf4 Is a Hub
782 in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. Mol Cell.
783 2016;63: 385–96. doi:10.1016/j.molcel.2016.06.009

784 36. Kouprina NYu, Pashina OB, Nikolaishwili NT, Tsouladze AM, Larionov VL. Genetic control of
785 chromosome stability in the yeast Saccharomyces cerevisiae. Yeast. 1988;4: 257–69.
786 doi:10.1002/yea.320040404

787 37. Yuan Z, Georgescu R, Santos R de LA, Zhang D, Bai L, Yao NY, et al. Ctf4 organizes sister
788 replisomes and Pol α into a replication factory. Elife. 2019;8: e47405. doi:10.7554/eLife.47405

789 38. Samora CP, Saksouk J, Goswami P, Wade BO, Singleton MR, Bates PA, et al. Ctf4 Links DNA
790 Replication with Sister Chromatid Cohesion Establishment by Recruiting the Chl1 Helicase to the
791 Replisome. Mol Cell. 2016;63: 371–84. doi:10.1016/j.molcel.2016.05.036

792 39. Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, et al. A Ctf4 trimer couples the
793 CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature. 2014;510: 293–297.
794 doi:10.1038/nature13234

795 40. Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, et al. A key
796 role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic
797 replisome. EMBO J. 2009;28: 2992–3004. doi:10.1038/emboj.2009.226

798 41. Abe T, Kawasumi R, Giannattasio M, Dusi S, Yoshimoto Y, Miyata K, et al. AND-1 fork protection
799 function prevents fork resection and is essential for proliferation. Nat Commun. 2018;9: 3091.
800 doi:10.1038/s41467-018-05586-7

801 42. Fumasoni M, Zwicky K, Vanoli F, Lopes M, Branzei D. Error-Free DNA Damage Tolerance and
802 Sister Chromatid Proximity during DNA Replication Rely on the Polα/Primase/Ctf4 Complex. Mol
803 Cell. 2015;57: 812–823. doi:10.1016/j.molcel.2014.12.038

804 43. Kouprina N, Kroll E, Bannikov V, Bliskovsky V, Gizatullin R, Kirillov A, et al. CTF4 (CHL15) mutants
805 exhibit defective DNA metabolism in the yeast Saccharomyces cerevisiae. Mol Cell Biol. 1992;12:
806 5736–5747. Available: http://www.ncbi.nlm.nih.gov/pubmed/1341195

807 44. Pâques F, Haber JE. Multiple pathways of recombination induced by double-strand breaks in
808 Saccharomyces cerevisiae. Microbiol Mol Biol Rev. 1999;63: 349–404.

809 45. Gerrish PJ, Lenski RE. The fate of competing beneficial mutations in an asexual population.

810 Genetica. 1998;102/103: 127–144. doi:10.1023/A:1017067816551

811 46. Lang GI, Botstein D, Desai MM. Genetic variation and the fate of beneficial mutations in asexual
812 populations. Genetics. 2011;188: 647–661. doi:10.1534/genetics.111.128942

813 47. Couce A, Tenaillon OA. The rule of declining adaptability in microbial evolution experiments.
814 Front Genet. 2015;6: 99. doi:10.3389/fgene.2015.00099

815 48. Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Long-Term Dynamics of
816 Adaptation in Asexual Populations. Science. 2013;342: 1364–1367. doi:10.1126/science.1198914

817 49. Chou H, Chiu H, Delaney NF, Segrè D, Marx CJ. Diminishing Returns Epistasis Among Beneficial
818 Mutations Decelerates Adaptation. Science. 2011;332: 1190–1193. doi:10.1126/science.1203799

819 50. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial
820 mutations in an evolving bacterial population. Science. 2011;332: 1193–1196.
821 doi:10.1126/science.1203801

822 51. Barrick JE, Kauth MR, Strelioff CC, Lenski RE. Escherichia coli rpoB mutants have increased
823 evolvability in proportion to their fitness defects. Mol Biol Evol. 2010;27: 1338–1347.
824 doi:10.1093/molbev/msq024

825 52. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Microbial evolution. Global epistasis makes
826 adaptation predictable despite sequence-level stochasticity. Science. 2014;344: 1519–1522.
827 doi:10.1126/science.1250939

828 53. Marad DA, Buskirk SW, Lang GI. Altered access to beneficial mutations slows adaptation and
829 biases fixed mutations in diploids. Nat Ecol Evol. 2018;2: 882–889. doi:10.1038/s41559-018-
830 0503-9

831 54. Zeyl C, Vanderford T, Carter M. An Evolutionary Advantage of Haploidy in Large Yeast
832 Populations. Science. 2003;299: 555–558. doi:10.1126/SCIENCE.1078417

833 55. Von Borstel RC, Cain KT, Steinberg CM. Inheritance of spontaneous mutability in yeast. Genetics.
834 1971;69: 17–27.

835 56. Kunz BA, Peters MG, Kohalmi SE, Armstrong JD, Glattke M, Badiani K. Disruption of the RAD52
836 gene alters the spectrum of spontaneous SUP4-o mutations in Saccharomyces cerevisiae.
837 Genetics. 1989;122: 535–542.

838 57. Puddu F, Herzog M, Selivanova A, Wang S, Zhu J, Klein-Lavi S, et al. Genome architecture and
839 stability in the Saccharomyces cerevisiae knockout collection. Nature. 2019; 1–5.
840 doi:10.1038/s41586-019-1549-9

841 58. Rolef Ben-Shahar T, Heeger S, Lehane C, East P, Flynn H, Skehel M, et al. Eco1-dependent cohesin
842 acetylation during establishment of sister chromatid cohesion. Science. 2008/07/26. 2008;321:
843 563–566. doi:10.1126/science.1157774

844 59. Panizza S, Tanaka T, Hochwagen A, Eisenhaber F, Nasmyth K. Pds5 cooperates with cohesion in
845 maintaining sister chromatid cohesion. Curr Biol. 2000;10: 1557–1564. doi:10.1016/S0960-
846 9822(00)00854-X

847 60. Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature
848 separation of sister chromatids. Cell. 1997;91: 35–45. doi:doi: 10.1016/s0092-8674(01)80007-6.

849 61. Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Nasmyth K. Cohesin's binding to
850 chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Mol Cell.
851 2000;5: 243–254. doi:10.1016/S1097-2765(00)80420-7

852 62. Srinivasan M, Fumasoni M, Petela NJ, Murray A, Nasmyth KA. Cohesion is established during DNA
853 replication utilising chromosome associated cohesin rings as well as those loaded de novo onto
854 nascent DNAs. Elife. 2020;9. doi:10.7554/eLife.56611

855 63. St. Charles J, Petes TD. High-Resolution Mapping of Spontaneous Mitotic Recombination
856 Hotspots on the 1.1 Mb Arm of Yeast Chromosome IV. Klein H, editor. PLoS Genet. 2013;9:
857 e1003434. doi:10.1371/journal.pgen.1003434

858 64. Sharp NP, Sandell L, James CG, Otto SP. The genome-wide rate and spectrum of spontaneous
859 mutations differ between haploid and diploid yeast. Proc Natl Acad Sci USA. 2018;115: E5046–
860 E5055. doi:10.1073/pnas.1801040115

861 65. Serero A, Jubin C, Loeillet S, Legoix-Né P, Nicolas AG. Mutational landscape of yeast mutator
862 strains. Proc Natl Acad Sci USA. 2014;111: 1897–1902. doi:10.1073/pnas.1314423111

863 66. Loeillet S, Herzog M, Puddu F, Legoix P, Baulande S, Jackson SP, et al. Trajectory and uniqueness
864 of mutational signatures in yeast mutators. Proc Natl Acad Sci USA. 2020;117: 24947–24956.
865 doi:10.1073/pnas.2011332117

866 67. Fisher KJ, Buskirk SW, Vignogna RC, Marad DA, Lang GI. Adaptive genome duplication affects
867 patterns of molecular evolution in Saccharomyces cerevisiae. PLOS Genet. 2018;14: e1007396.
868 doi:10.1371/journal.pgen.1007396

869 68. Scott AL, Richmond PA, Dowell RD, Selmecki AM. The Influence of Polyploidy on the Evolution of
870 Yeast Grown in a Sub-Optimal Carbon Source. Mol Biol Evol. 2017;34: 2690–2703.
871 doi:10.1093/molbev/msx205

872 69. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, et al. Characteristic genome
873 rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc Natl Acad Sci USA.
874 2002;99: 16144–16149. doi:10.1073/pnas.242624799

875 70. Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, et al. The Repertoire and Dynamics
876 of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. PLoS Genet.
877 2008;4: e1000303. doi:10.1371/journal.pgen.1000303

878 71. Lauer S, Avecilla G, Spealman P, Sethia G, Brandt N, Levy SF, et al. Single-cell copy number variant
879 detection reveals the dynamics and diversity of adaptation. PLOS Biol. 2018;16: e3000069.
880 doi:10.1371/journal.pbio.3000069

881 72. Tsaponina O, Barsoum E, Åström SU, Chabes A. Ixr1 is required for the expression of the
882 ribonucleotide reductase Rnr1 and maintenance of dNTP pools. PLoS Genet. 2011;7.
883 doi:10.1371/journal.pgen.1002061

884 73. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H. GINS, a novel multiprotein
885 complex required for chromosomal DNA replication in budding yeast. Genes Dev. 2003;17: 1153–
886 1165. doi:10.1101/gad.1065903

887 74. Tanaka H, Katou Y, Yagura M, Saitoh K, Itoh T, Araki H, et al. Ctf4 coordinates the progression of
888 helicase and DNA polymerase alpha. Genes to cells. 2009;14: 807–820. doi:10.1111/j.1365-

889 2443.2009.01310.x

890 75. Orr HA, Otto SP. Does diploidy increase the rate of adaptation? Genetics. 1994;136: 1475–1480.

891 76. Tlsty TD, Albertini AM, Miller JH. Gene amplification in the lac region of E. coli. Cell. 1984;37:
892 217–224. doi:10.1016/0092-8674(84)90317-9

893 77. Slack A, Thornton PC, Magner DB, Rosenberg SM, Hastings PJ. On the Mechanism of Gene
894 Amplification Induced under Stress in Escherichia coli. PLoS Genet. 2006;2: e48.
895 doi:10.1371/journal.pgen.0020048

896 78. Reddy G, Desai MM. Global epistasis emerges from a generic model of a complex trait. bioRxiv.
897 2020; 2020.06.14.150946. doi:10.1101/2020.06.14.150946

898 79. Wünsche A, Dinh DM, Satterwhite RS, Arenas CD, Stoebel DM, Cooper TF. Diminishing-returns
899 epistasis decreases adaptability along an evolutionary trajectory. Nat Ecol Evol. 2017;1: 0061.
900 doi:10.1038/s41559-016-0061

901 80. Schoustra S, Hwang S, Krug J, de Visser JAGM. Diminishing-returns epistasis among random
902 beneficial mutations in a multicellular fungus. Proc R Soc B Biol Sci. 2016;283: 20161376.
903 doi:10.1098/rspb.2016.1376

904 81. Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, et al.
905 Mechanisms of Haploinsufficiency Revealed by Genome-Wide Profiling in Yeast. Genetics.
906 2005;169: 1915–1925. doi:10.1534/GENETICS.104.036871

907 82. Gerstein AC, Otto SP. Ploidy and the Causes of Genomic Evolution. J Hered. 2009;100: 571–581.
908 doi:10.1093/jhered/esp057

909 83. Otto SP, Gerstein AC. The evolution of haploidy and diploidy. Current Biology. 2008. pp. R1121–
910 R1124. doi:10.1016/j.cub.2008.09.039

911 84. Schacherer J, de Montigny J, Welcker A, Souciet JL, Potier S. Duplication processes in
912 Saccharomyces cerevisiae haploid strains. Nucleic Acids Res. 2005;33: 6319–6326.
913 doi:10.1093/nar/gki941

914 85. Coïc E, Feldman T, Landman AS, Haber JE. Mechanisms of Rad52-independent spontaneous and
915 UV-induced mitotic recombination in Saccharomyces cerevisiae. Genetics. 2008;179: 199–211.
916 doi:10.1534/genetics.108.087189

917 86. Prado F, González-Barrera S, Aguilera A. RAD52-dependent and -independent homologous
918 recombination initiated by Flp recombinase at a single FRT site flanked by direct repeats. Mol
919 Gen Genet. 2000;263: 73–80. doi:10.1007/PL00008677

920 87. Payen C, Koszul R, Dujon B, Fischer G. Segmental Duplications Arise from Pol32-Dependent
921 Repair of Broken Forks through Two Alternative Replication-Based Mechanisms. PLoS Genet.
922 2008;4: e1000175. doi:10.1371/journal.pgen.1000175

923 88. Naiman K, Pagès V, Fuchs RP. A defect in homologous recombination leads to increased
924 translesion synthesis in E. coli. Nucleic Acids Res. 2016;44: 7691–7699. doi:10.1093/nar/gkw488

925 89. LaBar T, Phoebe Hsieh Y-Y, Fumasoni M, Murray AW. Evolutionary Repair Experiments as a
926 Window to the Molecular Diversity of Life. Curr Biol. 2020;30: R565–R574.
927 doi:10.1016/J.CUB.2020.03.046

928 90. Russo M, Crisafulli G, Sogari A, Reilly NM, Arena S, Lamba S, et al. Adaptive mutability of
929 colorectal cancers in response to targeted therapies. Science. 2019;366: 1473–1480.
930 doi:10.1126/science.aav4474

931 91. Thiagalingam S, Laken S, Willson JKV, Markowitz SD, Kinzler KW, Vogelstein B, et al. Mechanisms
932 underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci USA. 2001;98:
933 2698–2702. doi:10.1073/pnas.051625398

934 92. D'Angiolo M, Chiara M De, Yue J-X, Irizar A, Stenberg S, Persson K, et al. A yeast living ancestor
935 reveals the origin of genomic introgressions. Nature. 2020; 1–6. doi:10.1038/s41586-020-2889-1

936 93. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, et al. Genome evolution across
937 1,011 Saccharomyces cerevisiae isolates. Nature. 2018;556: 339–344. doi:10.1038/s41586-018-
938 0030-5

939 94. Steenwyk JL, Opulente DA, Kominek J, Shen X-X, Zhou X, Labella AL, et al. Extensive loss of cell-
940 cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. PLOS Biol. 2019;17:
941 e3000255. doi:10.1371/journal.pbio.3000255

942 95. Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome.
943 Translational Cancer Research. 2013. pp. 107–129. doi:10.3978/j.issn.2218-676X.2013.04.01

944 96. Hohl M, Mojumdar A, Hailemariam S, Kuryavyi V, Ghisays F, Sorenson K, et al. Modeling cancer
945 genomic data in yeast reveals selection against ATM function during tumorigenesis. PLOS Genet.
946 2020;16: e1008422. doi:10.1371/journal.pgen.1008422

947 97. Song W, Dominska M, Greenwell PW, Petes TD. Genome-wide high-resolution mapping of
948 chromosome fragile sites in Saccharomyces cerevisiae. Proc Natl Acad Sci USA. 2014;111: E2210-
949 8. doi:10.1073/pnas.1406847111

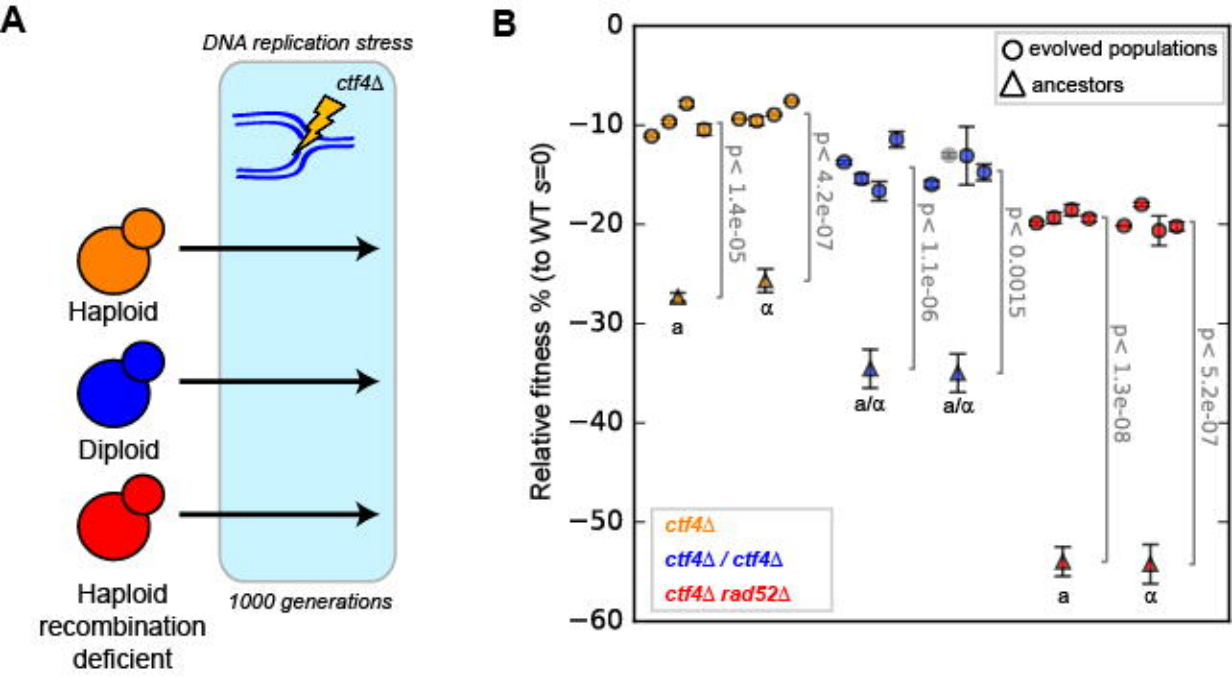
950 98. Capra JA, Paeschke K, Singh M, Zakian VA. G-Quadruplex DNA Sequences Are Evolutionarily
951 Conserved and Associated with Distinct Genomic Features in Saccharomyces cerevisiae. PLoS
952 Comput Biol. 2010;6: e1000861. doi:10.1371/journal.pcbi.1000861

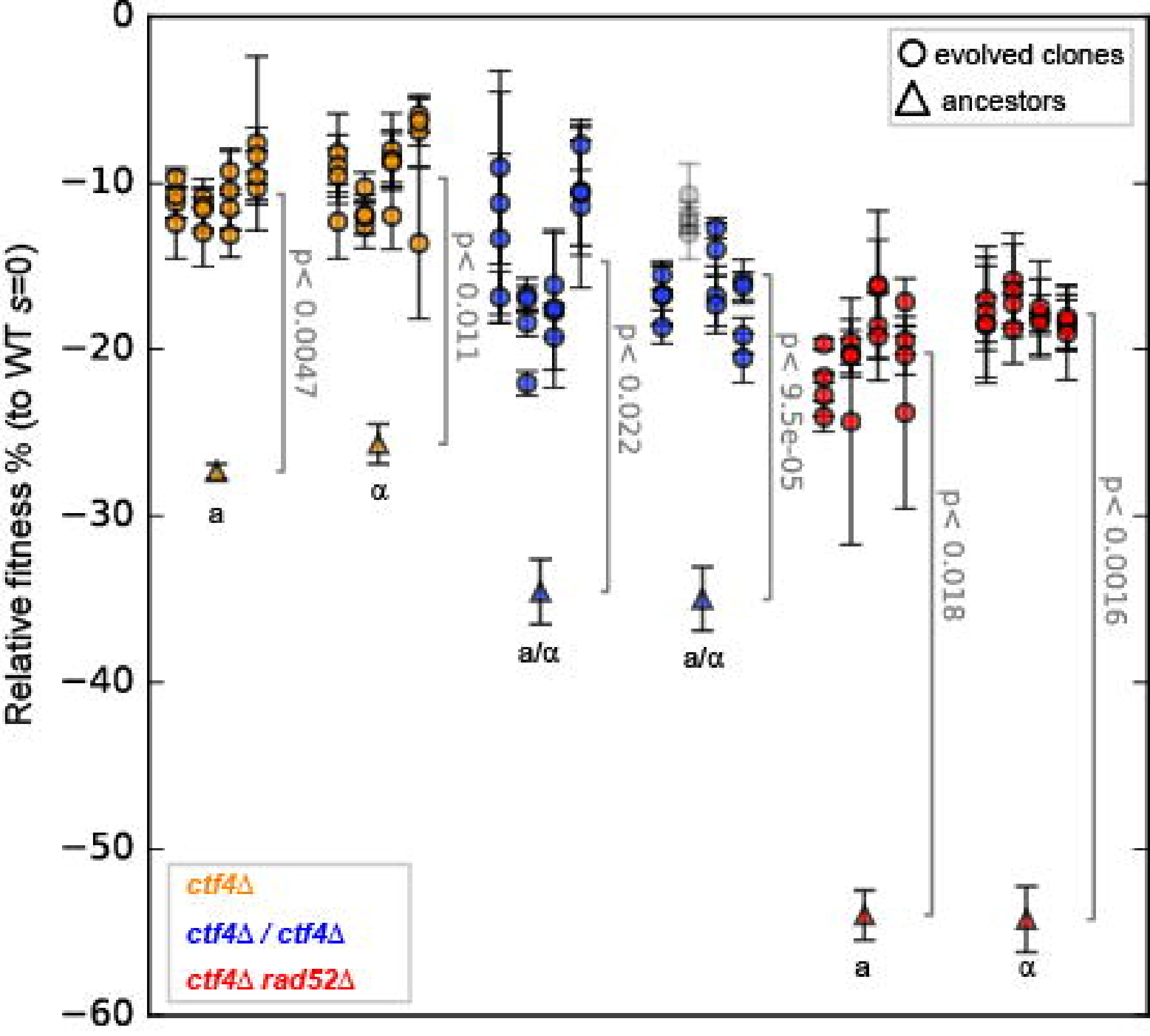
953 99. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, et al. The Transcriptional
954 Landscape of the Yeast Genome Defined by RNA Sequencing. Science. 2008;320: 1344–1350.

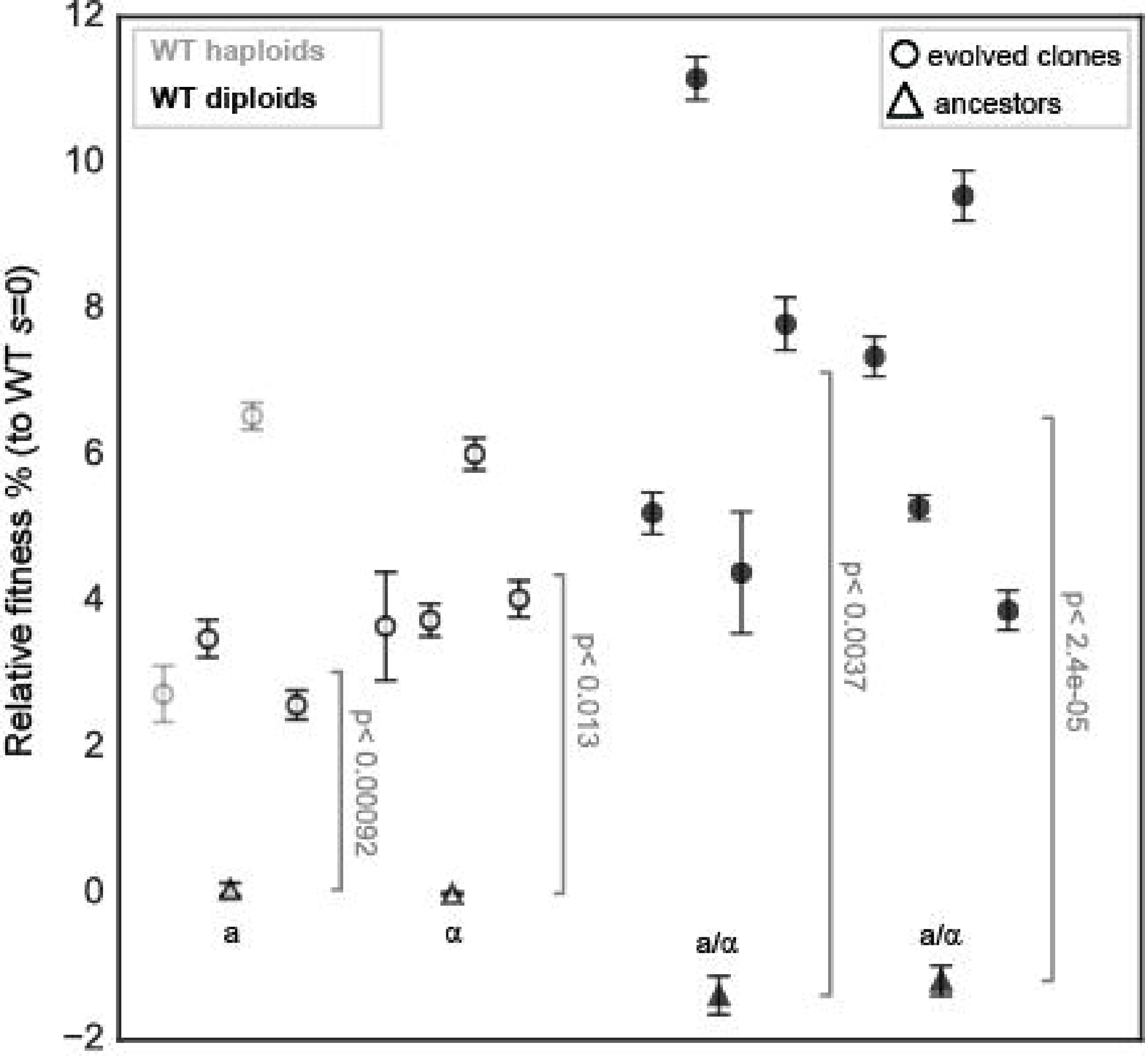
955 100. Azvolinsky A, Dunaway S, Torres JZ, Bessler JB, Zakian VA. The S. cerevisiae Rrm3p DNA helicase
956 moves with the replication fork and affects replication of all yeast chromosomes. Genes Dev.
957 2006;20: 3104–16. doi:10.1101/gad.1478906

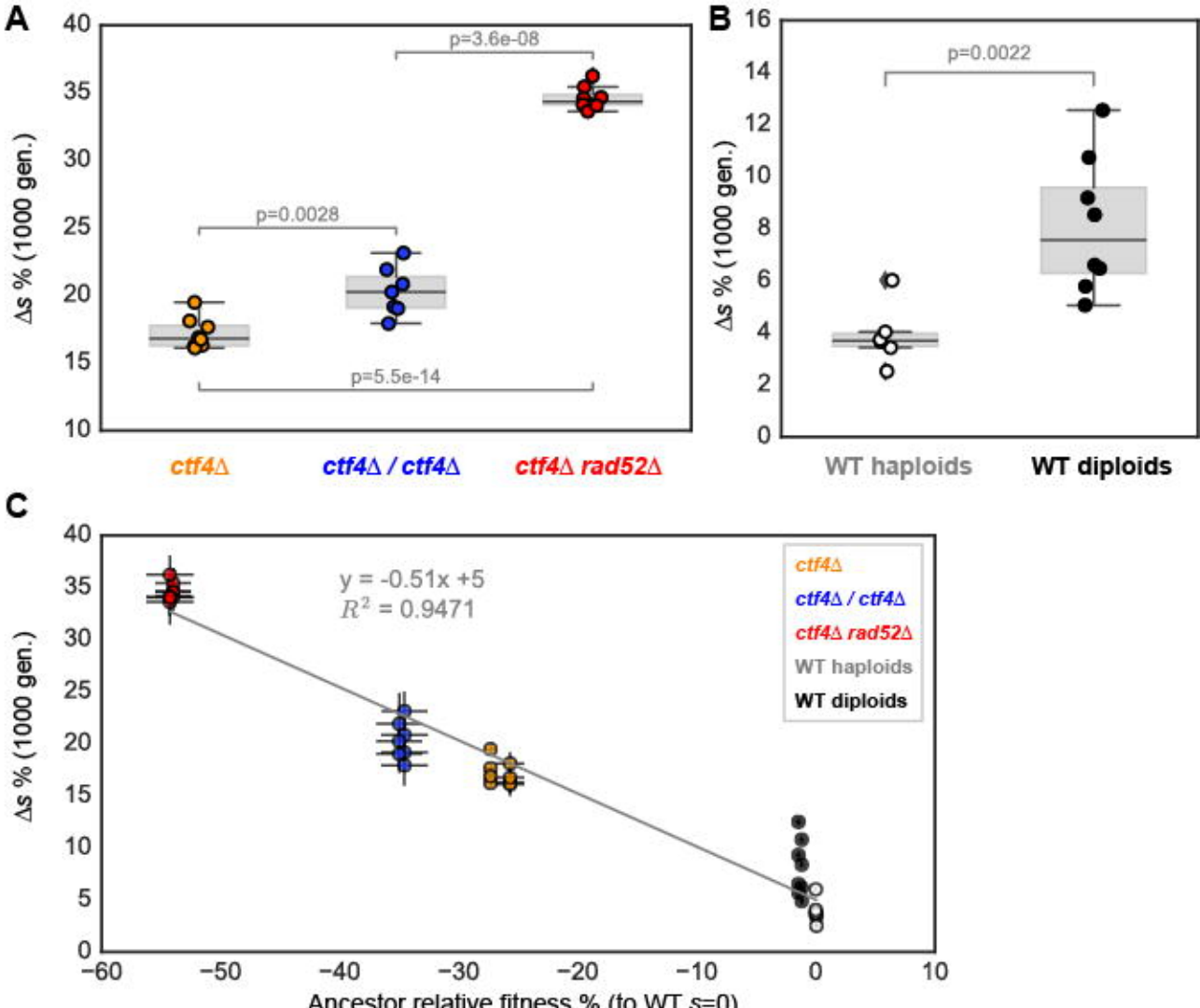
958 101. Szilard RK, Jacques P-E, Laramée L, Cheng B, Galicia S, Bataille AR, et al. Systematic identification
959 of fragile sites via genome-wide location analysis of gamma-H2AX. Nat Struct Mol Biol. 2010;17:
960 299–305. doi:10.1038/nsmb.1754

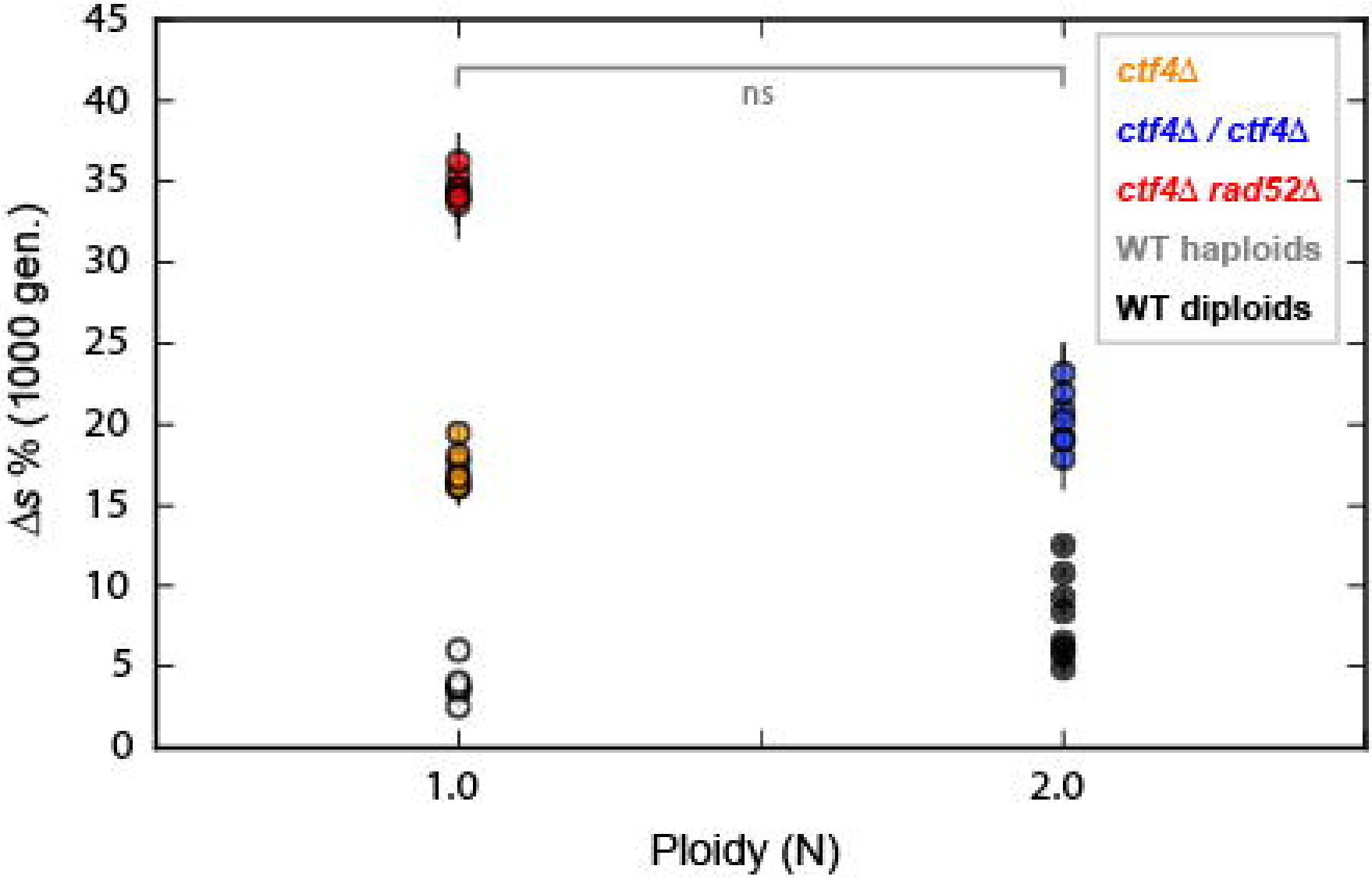
961


A
DNA replication stress
ctf4Δ
Haploid
Diploid
Haploid recombination deficient
1000 generations
B
evolved populations
ancestors
Relative fitness % (to WT s=0)
0
−10
−20
−30
−40
−50
−60
p< 1.4e-05
p< 4.2e-07
p< 1.1e-06
p< 0.0015
p< 1.3e-08
p< 5.2e-07
a
α
a/α
a/α
a
α
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ

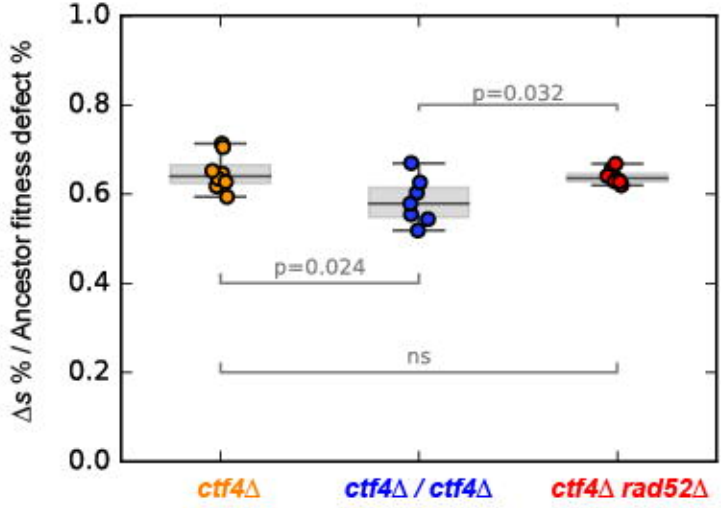

evolved clones
ancestors
Relative fitness % (to WT s=0)
0
−10
−20
−30
−40
−50
−60
p< 0.0047
p< 0.011
p< 0.022
p< 9.5e-05
p< 0.018
p< 0.0016
a
α
a/α
a/α
a
α
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ


WT haploids
WT diploids
evolved clones
ancestors
Relative fitness % (to WT s=0)
12
10
8
6
4
2
0
−2
a
α
a/α
a/α
p< 0.00092
p< 0.013
p< 0.0037
p< 2.4e−05

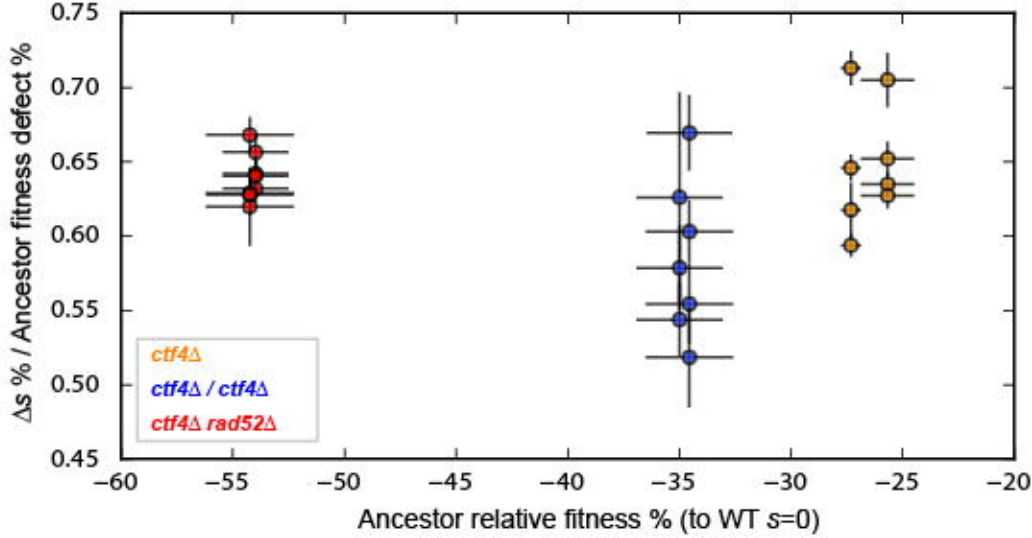

A
Δs % (1000 gen.)
p=3.6e-08
p=0.0028
p=5.5e-14
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
B
Δs % (1000 gen.)
p=0.0022
WT haploids
WT diploids
C
Δs % (1000 gen.)
y = -0.51x +5
R^2 = 0.9471
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
WT haploids
WT diploids
Ancestor relative fitness % (to WT s=0)

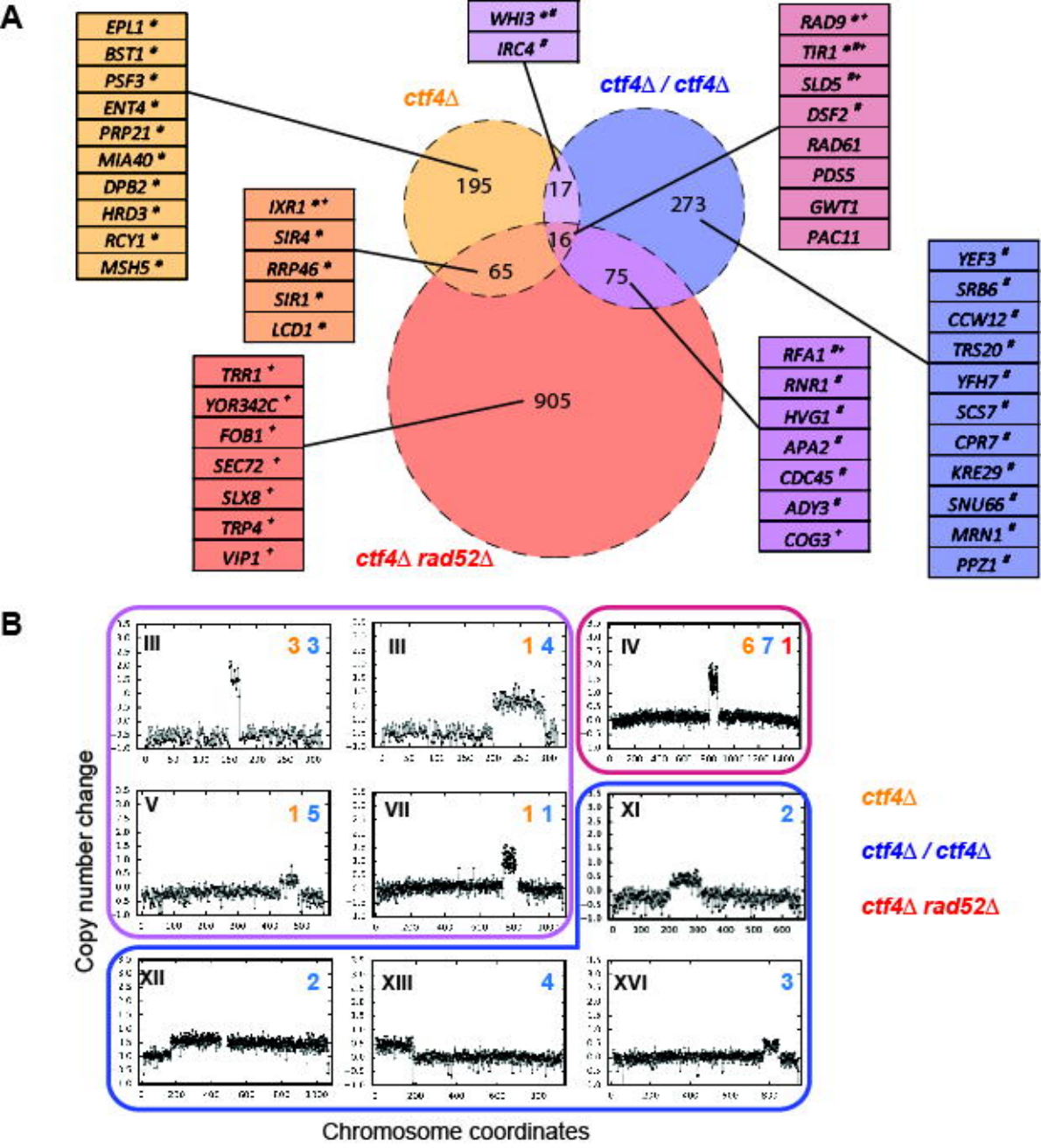

45
40
35
30
25
20
15
10
5
0
Δs % (1000 gen.)
ns
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
WT haploids
WT diploids
1.0
2.0
Ploidy (N)

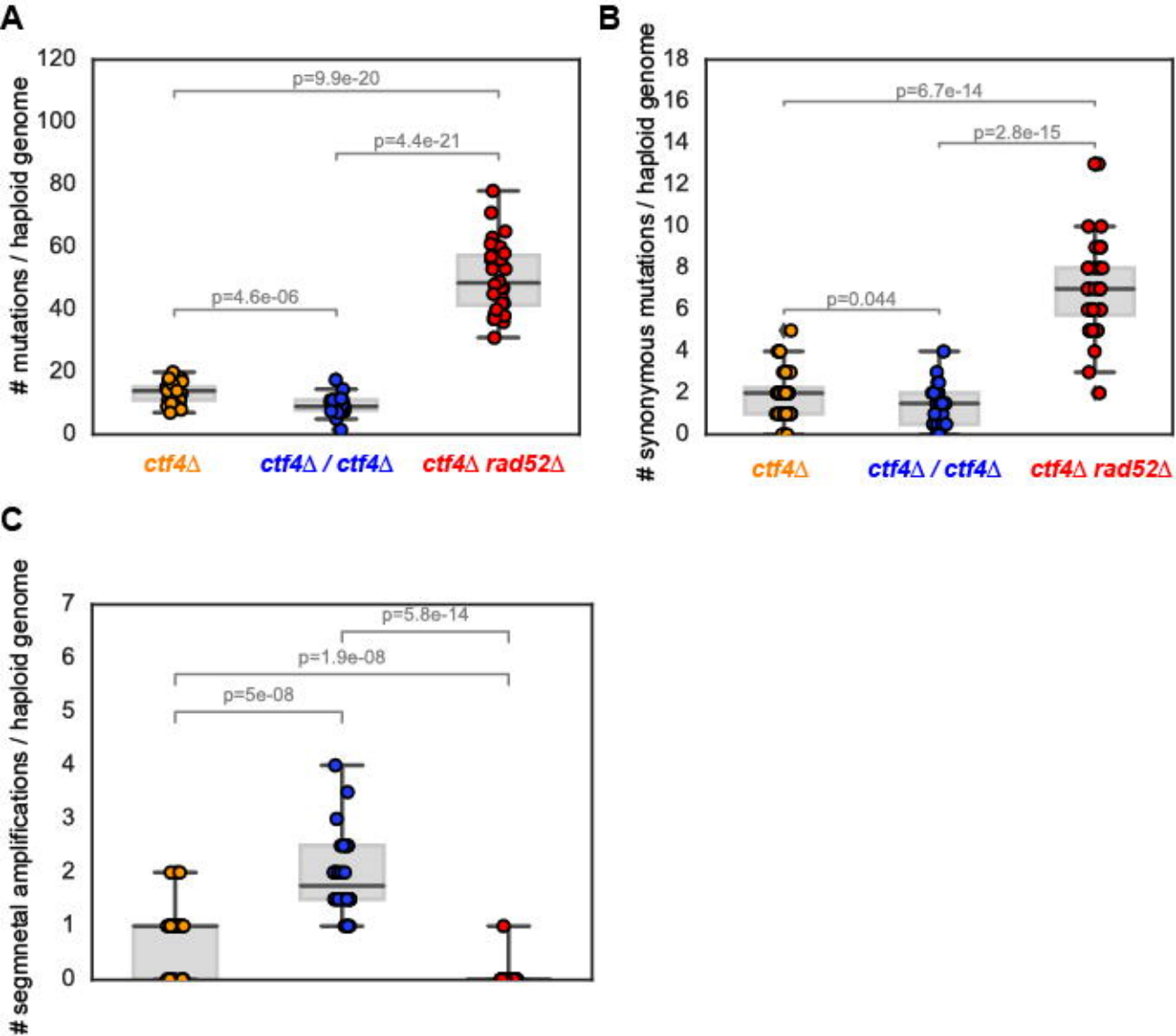
A

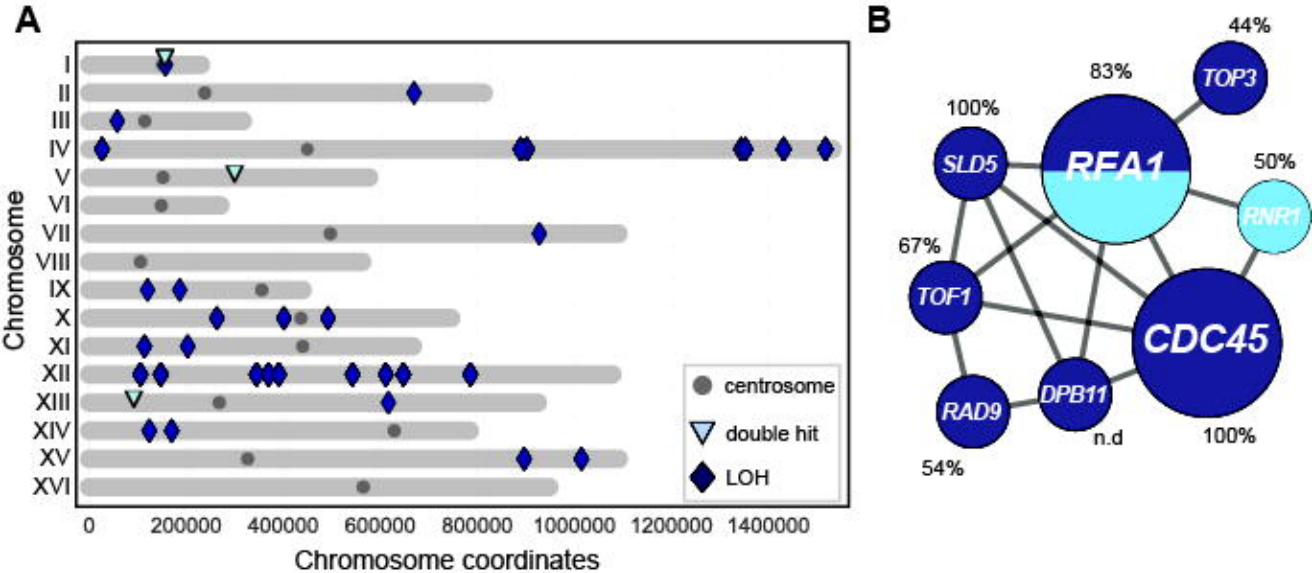

1.0
0.8
0.6
0.4
0.2
0.0
Δs % / Ancestor fitness defect %
p=0.032
p=0.024
ns
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ

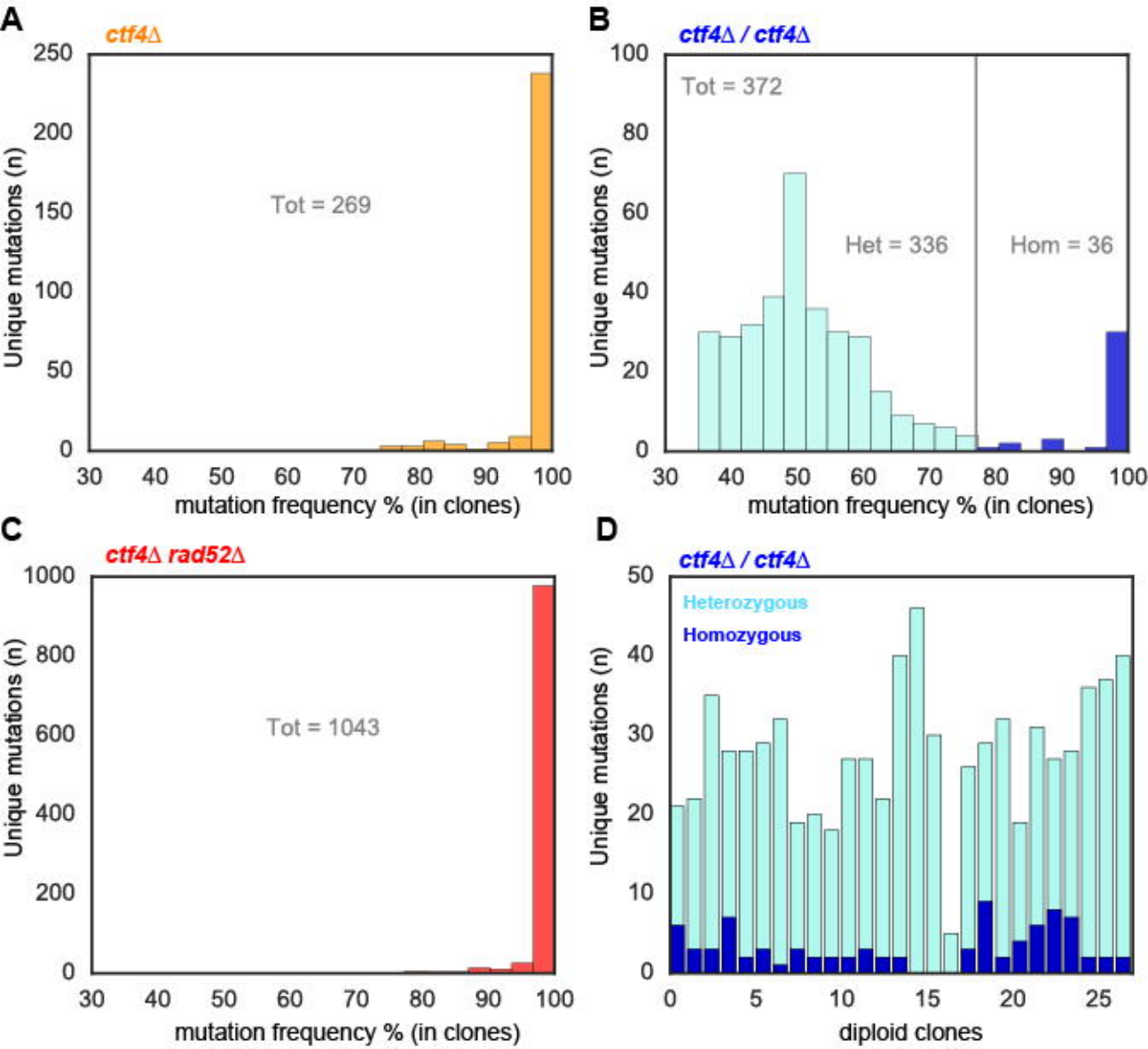
B

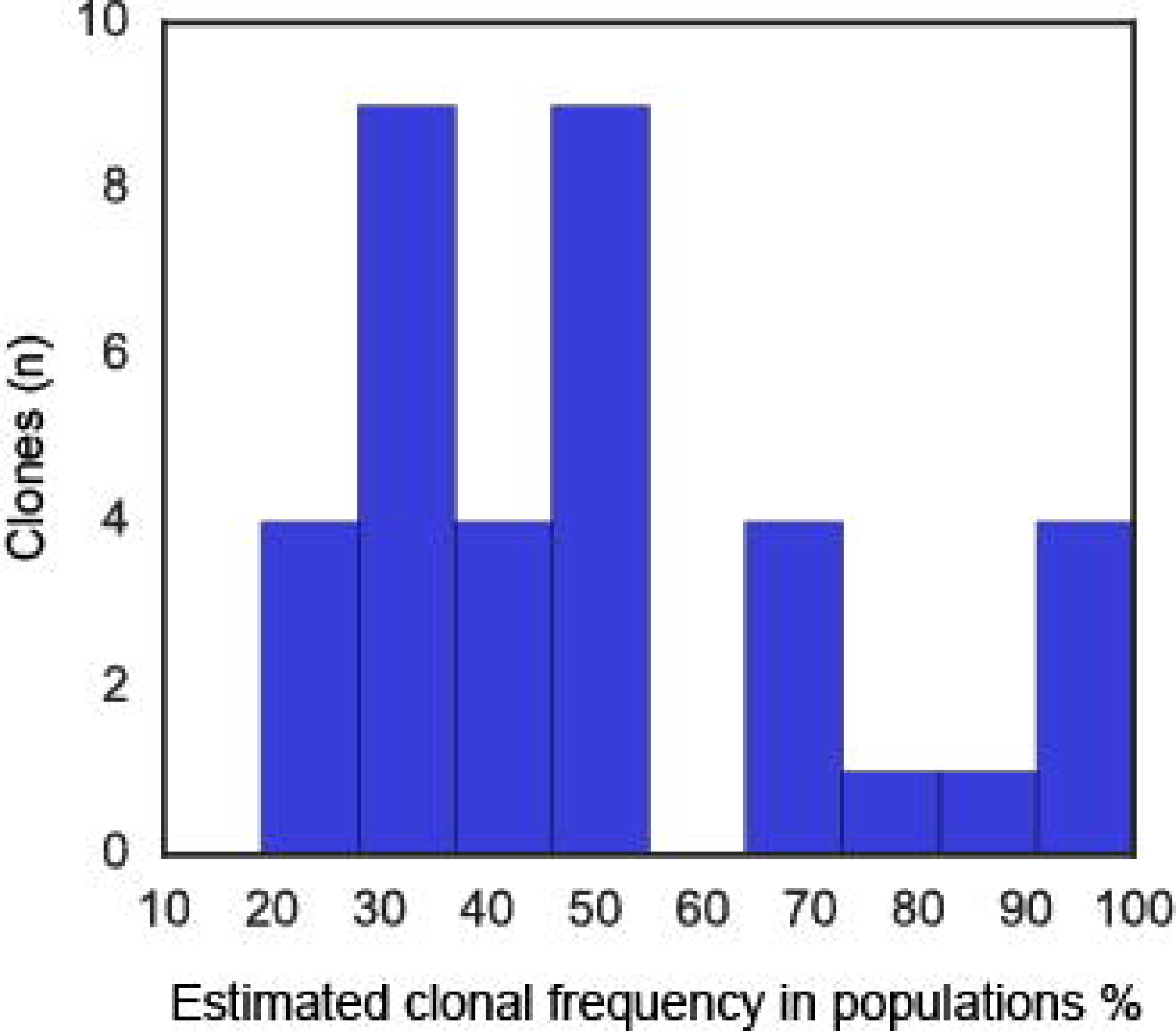

0.75
0.70
0.65
0.60
0.55
0.50
0.45
Δs % / Ancestor fitness defect %
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
−60
−55
−50
−45
−40
−35
−30
−25
−20
Ancestor relative fitness % (to WT s=0)

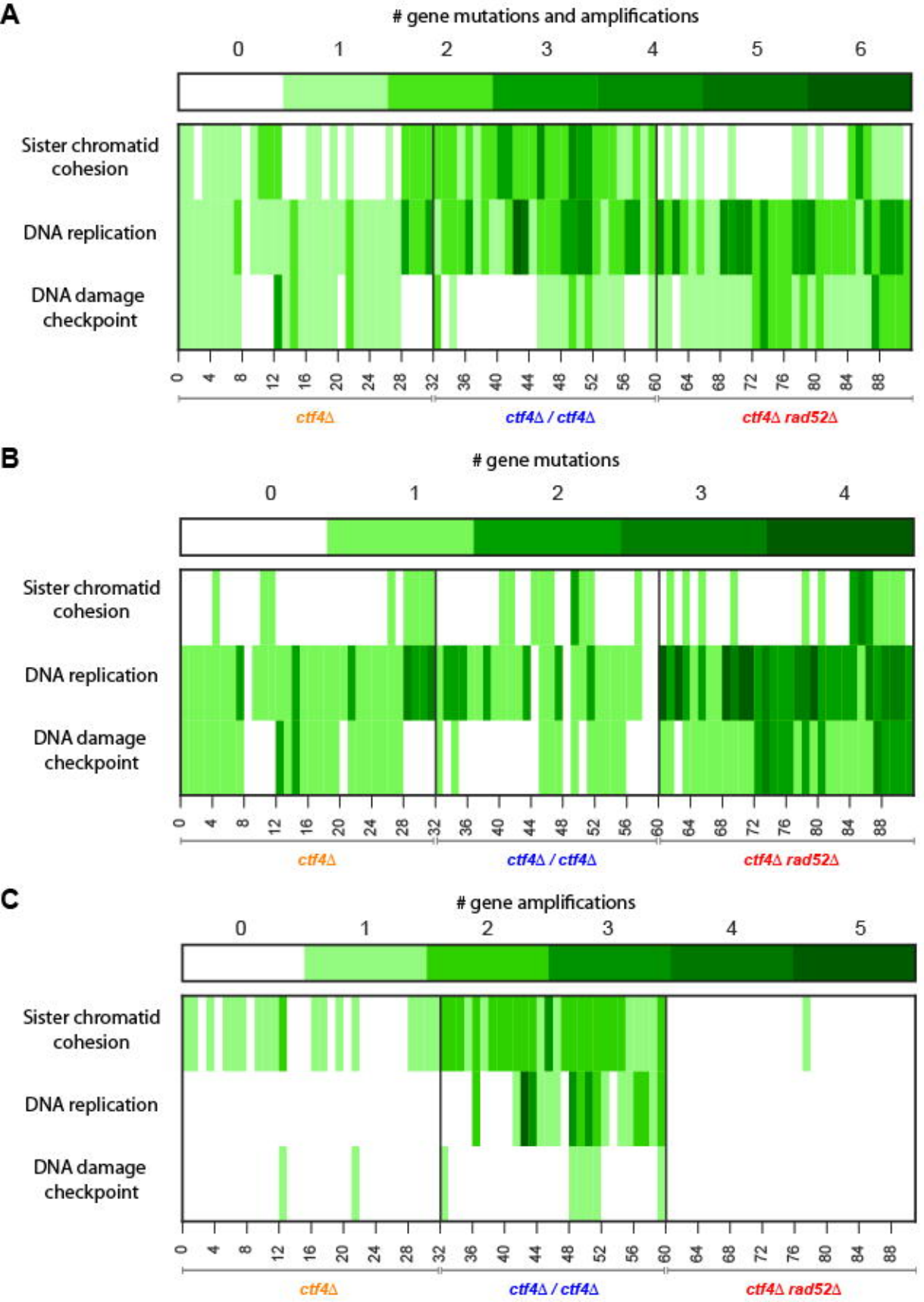

A
EPL1 *
BST1 *
PSF3 *
ENT4 *
PRP21 *
MIA40 *
DPB2 *
HRD3 *
RCY1 *
MSH5 *
WHI3 *#
IRC4 #
RAD9 *+
TIR1 *#+
SLD5 #+
DSF2 #
RAD61
PDS5
GWT1
PAC11
ctf4Δ
ctf4Δ / ctf4Δ
195
17
273
16
65
75
905
IXR1 *+
SIR4 *
RRP46 *
SIR1 *
LCD1 *
YEF3 #
SRB6 #
CCW12 #
TRS20 #
YFH7 #
SCS7 #
CPR7 #
KRE29 #
SNU66 #
MRN1 #
PPZ1 #
TRR1 +
YOR342C +
FOB1 +
SEC72 +
SLX8 +
TRP4 +
VIP1 +
RFA1 #+
RNR1 #
HVG1 #
APA2 #
CDC45 #
ADY3 #
COG3 +
ctf4Δ rad52Δ
B
III 3 3
III 1 4
IV 6 7 1
V 1 5
VII 1 1
XI 2
XII 2
XIII 4
XVI 3
Copy number change
Chromosome coordinates
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ

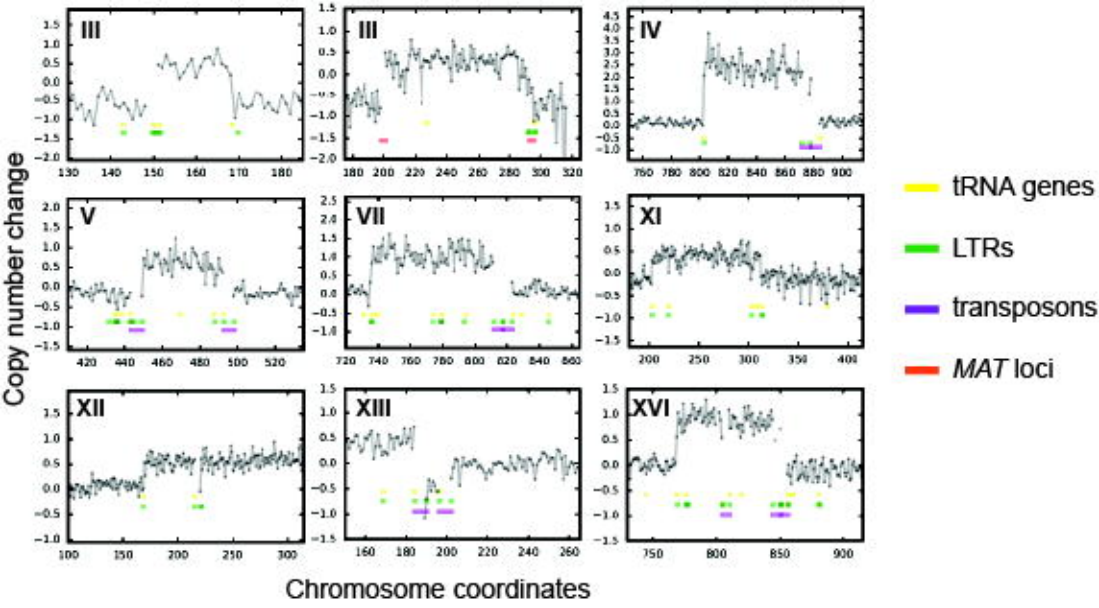

A
mutations / haploid genome
120
100
80
60
40
20
0
p=9.9e-20
p=4.4e-21
p=4.6e-06
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
B
synonymous mutations / haploid genome
18
16
14
12
10
8
6
4
2
0
p=6.7e-14
p=2.8e-15
p=0.044
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
C
segmnetal amplifications / haploid genome
7
6
5
4
3
2
1
0
p=5.8e-14
p=1.9e-08
p=5e-08

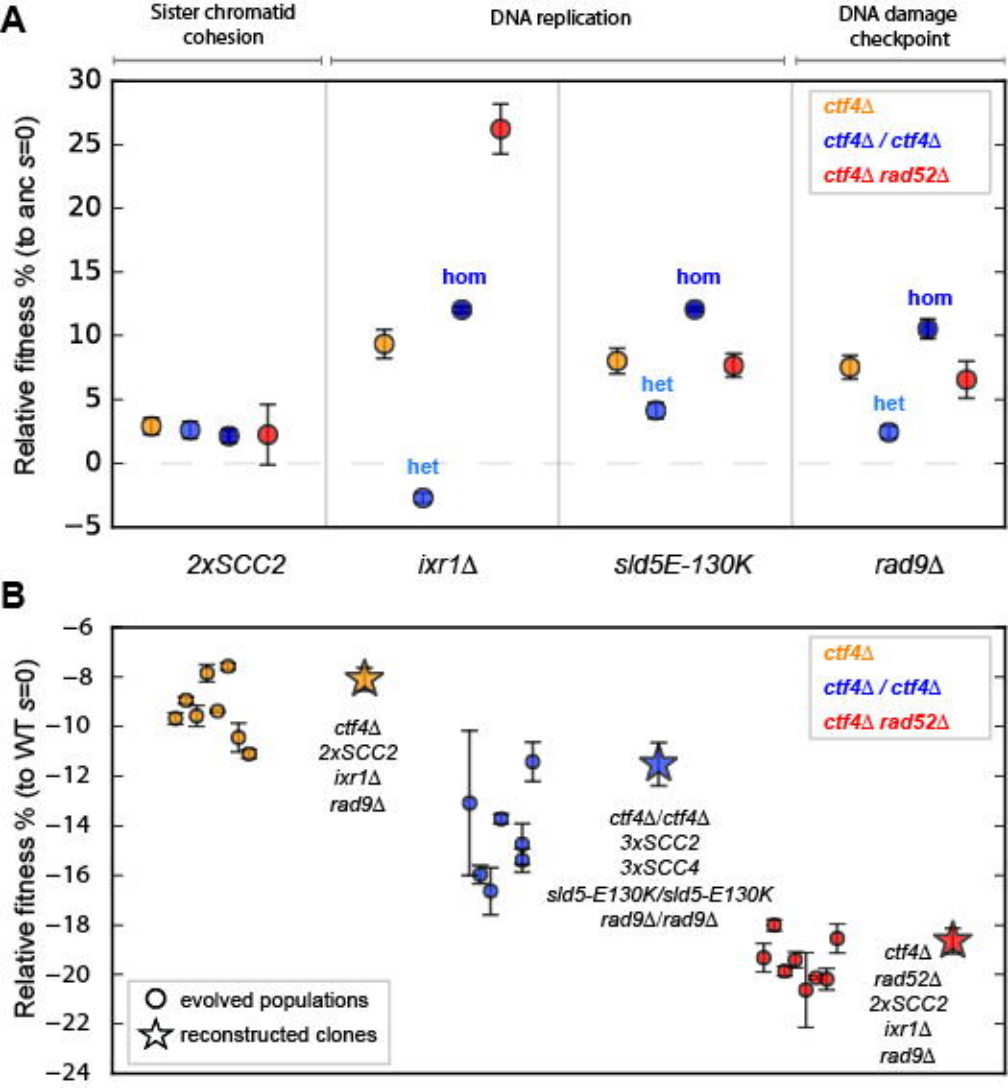

A
Chromosome
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
centrosome
double hit
LOH
0
200000
400000
600000
800000
1000000
1200000
1400000
Chromosome coordinates
B
44%
TOP3
83%
RFA1
100%
SLD5
50%
RNR1
67%
TOF1
CDC45
RAD9
DPB11
n.d
100%
54%

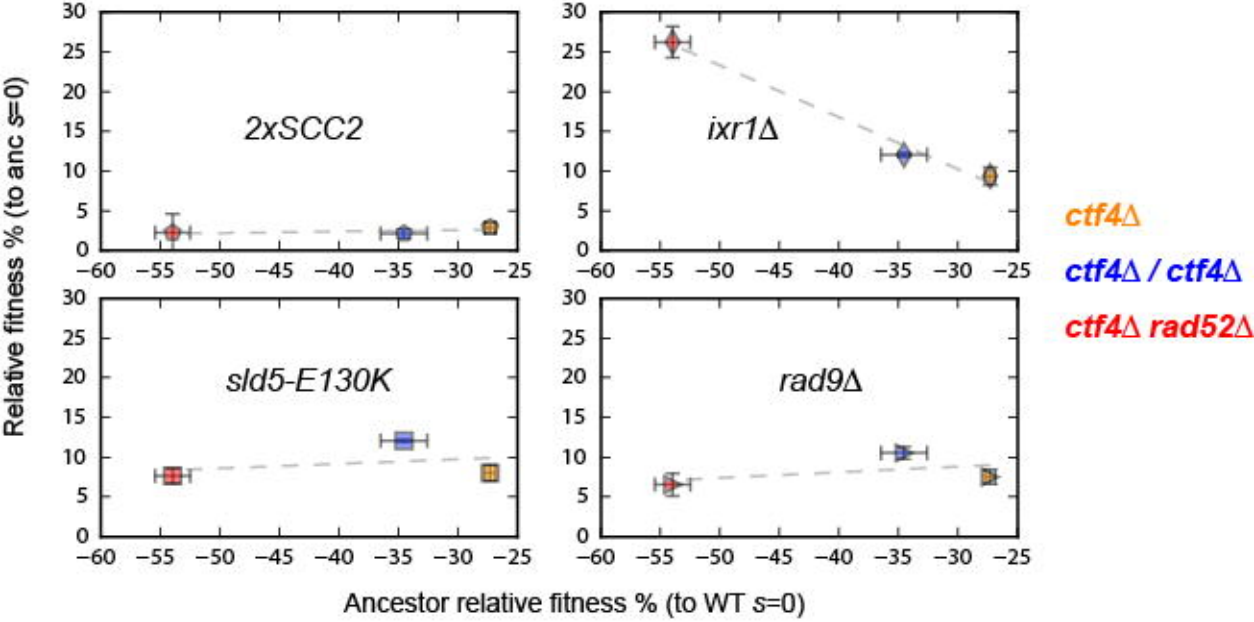

A
ctf4Δ
Tot = 269
Unique mutations (n)
mutation frequency % (in clones)
B
ctf4Δ / ctf4Δ
Tot = 372
Het = 336
Hom = 36
Unique mutations (n)
mutation frequency % (in clones)
C
ctf4Δ rad52Δ
Tot = 1043
Unique mutations (n)
mutation frequency % (in clones)
D
ctf4Δ / ctf4Δ
Heterozygous
Homozygous
Unique mutations (n)
diploid clones


Clones (n)
0
2
4
6
8
10
10
20
30
40
50
60
70
80
90
100
Estimated clonal frequency in populations %


A
gene mutations and amplifications
0
1
2
3
4
5
6
Sister chromatid cohesion
DNA replication
DNA damage checkpoint
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
B
gene mutations
0
1
2
3
4
Sister chromatid cohesion
DNA replication
DNA damage checkpoint
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
C
gene amplifications
0
1
2
3
4
5
Sister chromatid cohesion
DNA replication
DNA damage checkpoint
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ


III
III
IV
V
VII
XI
XII
XIII
XVI
Copy number change
Chromosome coordinates
tRNA genes
LTRs
transposons
MAT loci


A
Sister chromatid cohesion
DNA replication
DNA damage checkpoint
Relative fitness % (to anc s=0)
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ
hom
het
2xSCC2
ixr1Δ
sld5E-130K
rad9Δ
B
Relative fitness % (to WT s=0)
ctf4Δ
2xSCC2
ixr1Δ
rad9Δ
ctf4Δ/ctf4Δ
3xSCC2
3xSCC4
sld5-E130K/sld5-E130K
rad9Δ/rad9Δ
ctf4Δ
rad52Δ
2xSCC2
ixr1Δ
rad9Δ
evolved populations
reconstructed clones


2xSCC2
ixr1Δ
sld5-E130K
rad9Δ
Relative fitness % (to anc s=0)
Ancestor relative fitness % (to WT s=0)
ctf4Δ
ctf4Δ / ctf4Δ
ctf4Δ rad52Δ