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

In haploid budding yeast, evolutionary adaptation to constitutive DNA replication stress alters three genome maintenance modules: DNA replication, the DNA damage checkpoint, and sister chromatid cohesion. We asked how these trajectories depend on a population's genome architecture by comparing the adaptation of haploids to that diploids and recombination deficient haploids. In all three genome architectures, adaptation happens within 1000 generations at rates that are linearly correlated with the initial fitness defect of the ancestors. Mutations in individual genes are selected at different frequencies in different architectures, but the benefits these mutations confer are similar in all three architectures, and combinations of these mutations reproduce the fitness gains of evolved populations. Despite the differences in the selected mutations, the evolutionary trajectories target the same three functional modules in all architectures revealing a common evolutionary response to DNA replication stress.

*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 to different rates of adaptation [5–16] 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 a form of constitutive DNA replication stress [17], a perturbation of DNA synthesis that produces lesions in DNA, increases genome instability, and decreases fitness [18–21]. DNA replication stress is often present in cancer cells as a consequence of oncogene activation [22–25] and can also be generated by the unrestrained proliferation of selfish genetic elements such as viruses or transposons [26–29]. 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 [17].

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 [30,31]. 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, [32,33]) and exploited for anticancer therapies [34–36].

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 [17], a protein that physically links different activities at replication forks [37]. Adaptation to DNA replication stress occurs in all three genome

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

RESULTS

All populations adapt to constitutive DNA replication stress.

DNA replication stress is often induced with drugs or by reducing the level of DNA polymerases [38–41]. To avoid evolving drug resistance or increased polymerase expression, which would rapidly overcome DNA replication stress, we deleted the *CTF4* gene, which encodes a non-essential subunit of the DNA replication machinery (the replisome) [42]. Ctf4 is a homo-trimer that functions as a structural hub within the replisome [37,43] by binding to the replicative DNA helicase, primase (the enzyme that makes the RNA primers that initiate DNA replication), and other accessory factors [37,44–46]. In the absence of Ctf4, the Polα-primase and other lagging strand processing factors are poorly recruited to the replisome [37,44,47], causing several characteristic features of DNA replication stress, such as accumulation of single strand DNA (ssDNA) gaps [48,49], reversed and stalled forks [17,49], cell cycle checkpoint activation [47,50] and altered chromosome metabolism [51,52]. As a consequence of these defects, *ctf4Δ* cells have substantially reduced reproductive fitness [17]. We removed *CTF4* from wild-type (WT) haploid and diploid cells, as well as from haploids impaired in homologous recombination due to the deletion of *RAD52* (Figure 1A), which encodes a conserved enzyme required for pairing homologous DNA sequences during recombination [53]. Because Rad52 is involved in different forms of homologous recombination, it's absence produces the most severe recombination defects and thus allows us to achieve the largest recombination defect achievable with a single gene deletion [54]. Attempts to generate homozygous *ctf4Δ/ctf4Δ rad52Δ/rad52Δ* mutants failed due to the lethality of the resulting zygotes. We refer to the strains we analyzed as haploids (*ctf4Δ,* yellow), diploids (*ctf4Δ/ctf4Δ,* blue) and recombination-deficient (*ctf4Δ rad52Δ,* red). Compared to the previously reported *ctf4Δ* haploid strains, which manifested a fitness decrease of ~27% relative to isogenic wild-type cells [17], constitutive DNA replication stress caused by the

absence of Ctf4 was more detrimental in diploid (~35% reduced fitness) and recombination-deficient strains (~55% reduced fitness) (Figure 1B). Diploid cells require twice as many forks as haploids and Ctf4-deficient diploids are thus more likely to have forks that cause severe cell-cycle delays or cell lethality. We speculate that this increased probability explains the more prominent fitness defect displayed by diploid cells. Interestingly, homologs of Ctf4 are absent in prokaryotes, where the primase is physically linked to the replicative helicase [55] and Ctf4 is essential in the cells of eukaryotes with larger genomes such as chickens [48] and humans [56]. Rad52 is likely involved in rescuing stalled replication forks by recombination-dependent mechanisms [49,57]. We speculate that the absence of Rad52 increases the duration of these stalls and leads some of them to become double-stranded breaks resulting in cell lethality and explaining the decreased fitness of *ctf4Δ rad52Δ* haploid double mutants. In diploids *ctf4Δ rad52Δ* cells, which have twice as many chromosomes, the number of irreparably stalled fork may be sufficient to kill most of the cells in a population, thus explaining the unviability of the strain.

We evolved eight parallel diploid and recombination-deficient populations for 1000 generations in the same conditions and at the same time as the haploid populations we previously analyzed ([17] and see Materials and methods). In this regime, spontaneous mutations that increase fitness and survive genetic drift will spread through the population [58,59]. In addition, we evolved eight wild-type diploid populations like the previously evolved wild-type haploids, to control for mutations that are selected by the culture conditions rather than by ameliorating DNA replication stress [17].

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% [17], 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). One diploid line (EVO14) gave rise to a population with a haploid genome content, suggesting a possible haploidization event during evolution. Sequencing revealed no aneuploidies as a potential explanation of this phenomenon. While diploidization has been recurrently observed during experimental evolution with budding yeast [8,60–62], reports of spontaneous haploidization events have been instead scarce. Given the difficulties introduced by the change of ploidy over the 1000 generations, we have excluded EVO14 from all our analyses. Control evolved wild-type populations showed only small fitness increases: 4 ± 1% for haploids and 6.7 ± 2.7% for diploids (Figure 1-S2).

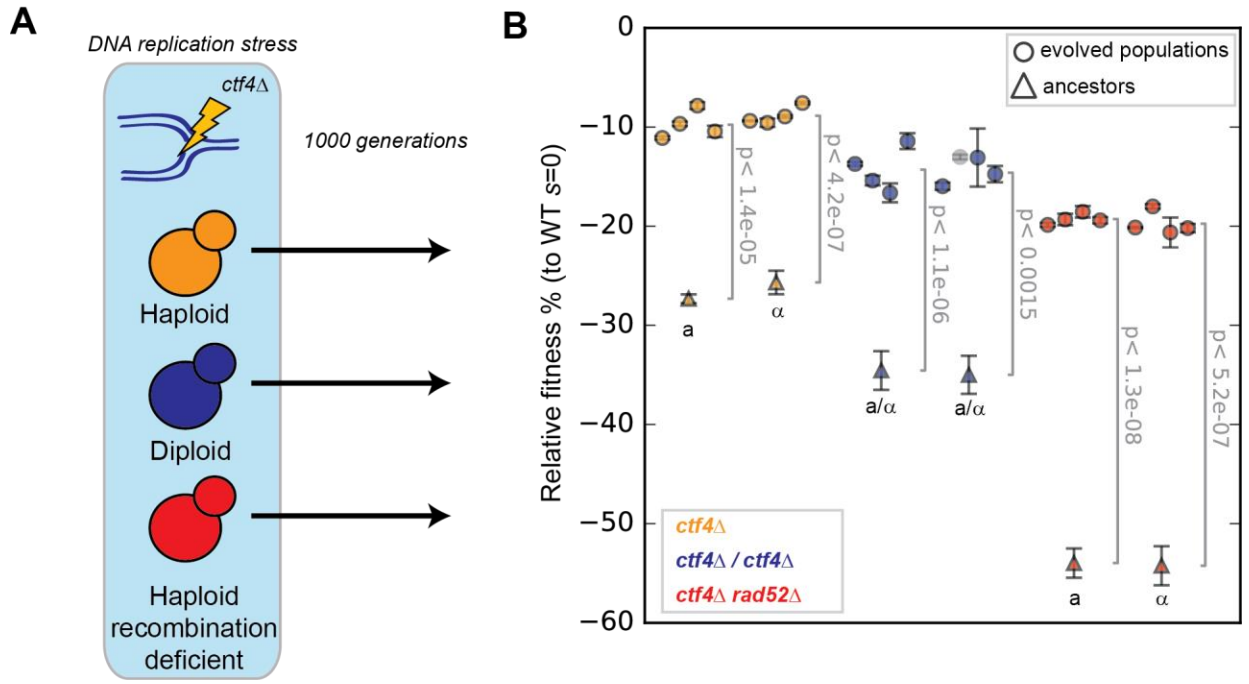
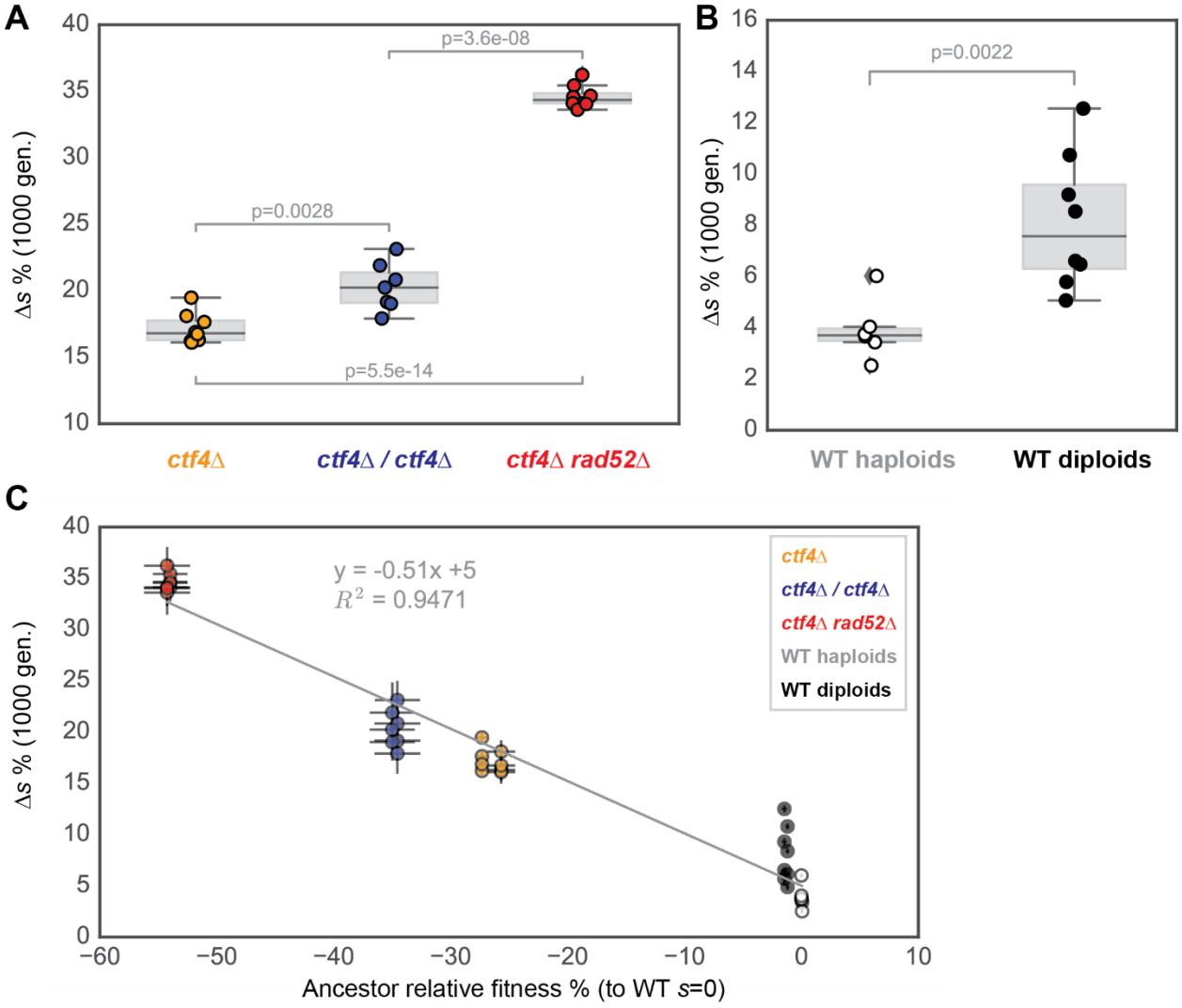


Figure 1: Evolutionary adaptation to constitutive DNA replication stress in different genome architectures. (A) Experimental scheme: Constitutive DNA replication stress was induced by removing Ctf4, a protein that coordinates replisome activities. Haploid (orange), diploids (blue) and cells with a severely decreased ability to perform homologous recombination (red, recombination-deficient), were subjected to 1000 generations of experimental evolution in the presence of constitutive DNA replication stress. **(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). Cells were competed against reference strains with the same ploidy. Fitness data of haploid strains (orange) is from [17]. 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 P-values reported in figures are the result of t-tests assuming unequal variances (Welch's test). 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' [63–66]. 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

154 fitness (R^2 = 0.95). These results agree with previous studies, which observed declining
155 adaptability across a wide range of organisms and selection conditions [11,12,67,68].



156

157 **Figure 2: Fitness increase over 1000 generations. (A)** The fitness increase over 1000 generations for strains
158 experiencing constitutive DNA replication stress. Δ*s* is the difference between the fitness of the evolved populations
159 and the fitness of their ancestors, with both fitnesses measured relative to that of cells that contain Ctf4 (*CTF4*) and
160 have the same ploidy. **(B)**. The fitness increase over 1000 generations for strains that contained wild-type *CTF4*. **(C)**
161 shows the correlation between the fitness increase during evolution (Δ*s*) and the fitness defect of the ancestors
162 relative to cells of the same architecture containing Ctf4 (the ancestor relative fitness). Fitness data of haploid strains
163 (orange and white) is from [17]. Error bars represent standard deviations. Grey line represents linear regression
164 between the datapoints**.** The P-values reported in figures are the result of t-tests assuming unequal variances
165 (Welch's test). The data shown here are reported in or derived from Source data 1.

166 The majority of previous studies have suggested that diploids adapt more slowly than isogenic
167 haploids [15,69,70]. Our diploid *ctf4Δ/ctf4Δ* populations gained more fitness than *ctf4Δ* haploids
168 over the course of 1000 generations (Figure 2A). Wild-type diploids increased more in absolute
169 fitness than wild-type haploids (Figure 2B), suggesting that the discrepancy between our work

170 and previous studies is not due to the presence of replication stress. Diploid populations
171 recovered a slightly smaller fraction of their initial fitness defect compared to haploid and
172 recombination-deficient strains (Figure 2-S2A). Collectively, our results suggest that ploidy is a
173 poor predictor of the speed and extent of adaptation (Figure 2-S1). We conclude that the fitness
174 of the ancestral cells is the major determinant influencing how much populations adapt, adding
175 further support to the concept of 'declining adaptability'.

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

177 We used whole-genome sequencing to understand the genetic basis of the evolutionary
178 adaptation to DNA replication stress. For each different genome architecture, we sequenced all
179 eight evolved populations, as well as four individual clones isolated from each of them. A total of
180 195, 276 and 905 genes or associated regulatory sequences were found mutated in evolved
181 haploid, diploid, and recombination-deficient populations respectively (Figure 3A,
182 Supplementary table 1). When normalized for DNA content, haploid (13 mutations/haploid
183 genome) and diploid clones (9 mutations/haploid genome) had similar but statistically different
184 mutation frequencies ($p = 4x10^{-6}$, Welch's test, Figure 3-S1A) but recombination-deficient clones
185 carried many more mutations, both in total (50/haploid genome, Figure 3-S1A) and as
186 synonymous mutations (7/haploid genome, Figure 3-S1B). Our results agree with the reported
187 increase in spontaneous mutagenesis in *rad52Δ* cells [71–73].

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

193 Only 8 of these genes were mutated in all three genome architectures (Figure 3A). This small
194 number could be due to different adaptive strategies adopted by the different architectures,
195 differences in the rate for the same mutation in different architectures, differences in the
196 selective benefit the mutation conferred, or a combination of these three effects. We found 28
197 genes with putative adaptive mutations that are specific to individual genome architectures, as
198 well as 14 genes with putative adaptive mutations in two architectures (Figure 3A, genes
199 highlighted in boxes and Supplementary table 2). Analyzing gene ontology (GO) terms associated
200 with the putative adaptive mutations revealed an enrichment of biological process terms
201 associated with genome maintenance (Supplementary table 3). The genes that were significantly
202 mutated in all genomic architectures included two that we had previously demonstrated to be
203 adaptive in haploid *ctf4Δ* cells: mutations in *RAD9* inactivate the DNA damage checkpoint and
204 those in *SLD5* alter the function of the replicative helicase [17]. All three architectures also had
205 mutations in *RAD61* and *PDS5*, which are implicated in chromosome segregation [74,75], a
206 cellular module shown to play a role during the adaptation of haploid *ctf4Δ* populations [17].

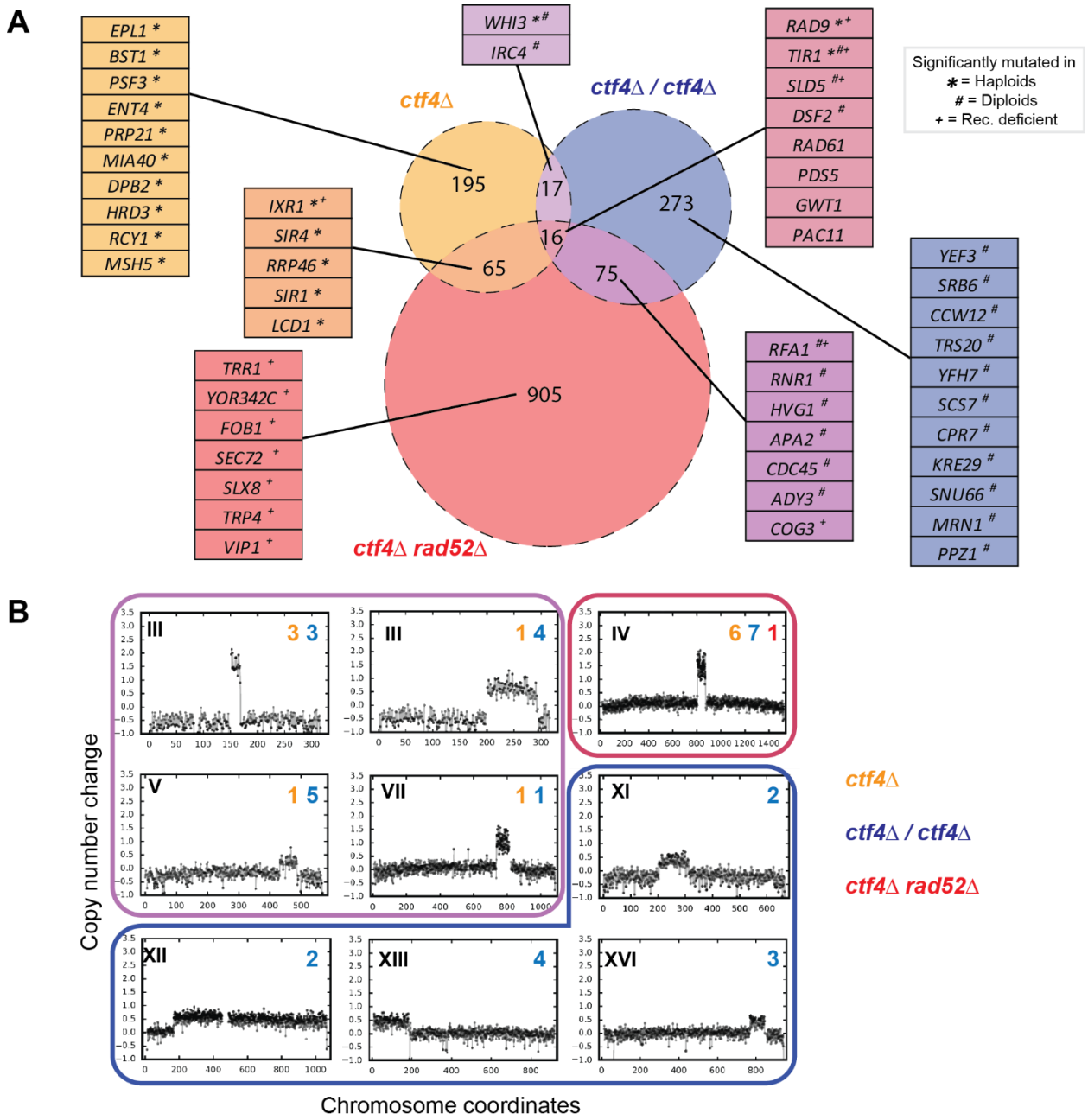


Figure 3: Putative adaptive mutations in the evolved clones. (A) Venn diagram representing the mutations (SNPs and small InDels 1-55bp) found in the evolved clones and populations. The area of the circles is proportional to the list of genes mutated at least once in the set of evolved clones for each 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 the genes in the Venn diagram intersections are reported in Source data 3 **(B)** Copy number variations (CNVs) affecting different chromosomes (roman numbers) which appeared in at least two independent populations. Arabic numbers are the number of independent populations, in each genetic

architecture (highlighted by a different color), in which the CNVs were detected. Copy number change refers to the fragment's gain or loss during the evolution experiment (i.e. +1 means that one copy was gained in haploid cells and that two copies were gained in diploid cells). Frames around the plots have the same color as the sectors in the Venn diagram, and encircle CNVs that have been found in one, two, or three genetic architectures. All the segmental amplifications detected are reported in Supplementary table 4.

We also examined the evolved genomes for changes in gene copy numbers (Copy Number Variations, CNVs) that might have played a role in adaptation. Many clones displayed an increased copy number of defined chromosome segments (segmental amplification) compared to the control populations (Supplementary table 4). Segmental amplifications were common in diploid clones (Figure 3-S1C). We inferred that segmental amplifications were putatively adaptive if the same chromosome region was found amplified in two clones isolated from independent populations (Figure 3B). Among these segmental amplifications, four CNVs were found only in evolved diploids, four were found in at least an evolved haploid and a diploid clone, and only one was detected in populations of all three genome architectures (Figure 3B). This latter fragment is located on chromosome IV and contains *SCC2*, which encodes a subunit of the cohesin loader complex. This essential complex is responsible for loading the proteinaceous cohesin ring that holds sister chromatids together until they segregate from each other during mitosis [76–78]. The gene encoding for the other subunit of the cohesin loader complex, *SCC4*, lies on the segment of chromosome V that was amplified in both haploid and diploid clones. We previously showed that extra copies of these genes alleviate the cohesion defects associated with the absence of Ctf4 in haploid cells and increase their fitness [17].

Overall, genome sequencing reveals that the identity of the genes carrying selected mutations varies depending on genome architecture. Nevertheless, we found evidence of common evolutionary strategies across different populations: Mutations affecting sister chromatid cohesion, DNA replication, and the DNA damage checkpoint, which were previously shown to drive adaptation to a form of DNA replication stress in haploid cells [17], are significantly mutated regardless of the genome architecture of the ancestral cells.

Evolved diploids show adaptive loss of heterozygosity

We analyzed the frequency of mutated reads in evolved diploid clones to estimate whether mutations were homozygous or heterozygous (Figure 4-S1). We detected 36 homozygous mutations across the evolved diploid clones (Figure 4A, Figure 4-S1B, Supplementary table 5). Homozygous mutations were evenly distributed across the evolved clones analyzed and thus were not the product of a single unrepresentative event (Figure 4-S1C). Individual clones often carried multiple homozygous mutations on the same chromosome arm (Supplementary table 5) suggesting that the loss-of-heterozygosity (LOH) was likely the consequence of mitotic recombination events affecting large chromosomal segments [79]. In addition to homozygous 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.

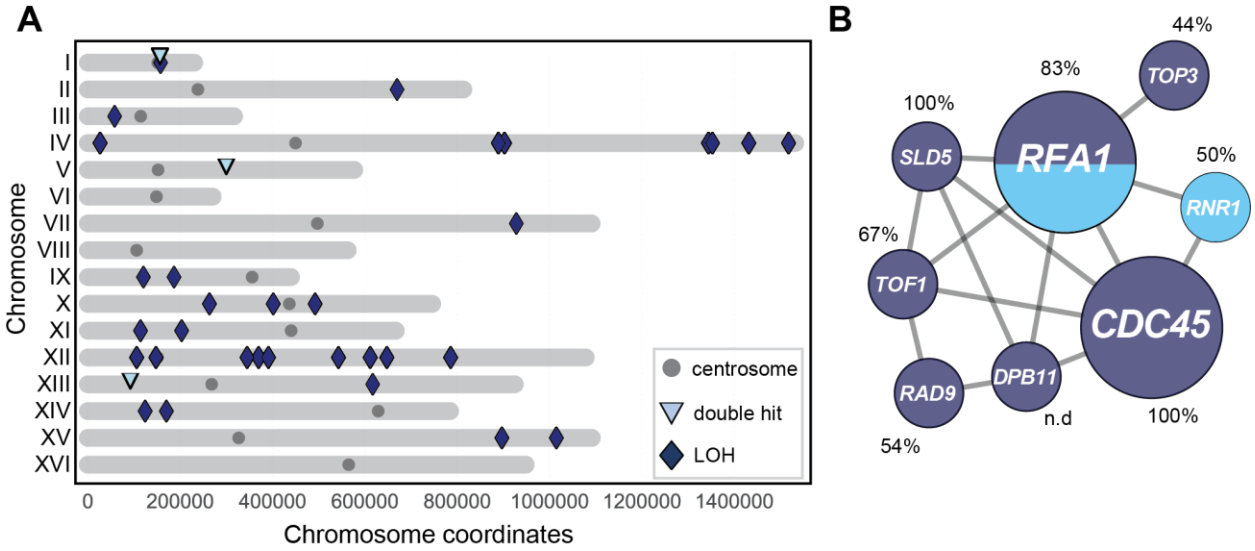


Figure 4: Homozygous mutations in diploids evolved in the presence of constitutive DNA replication stress. (A) Homozygous mutations (SNPs and small InDels of 1-55bp) detected in evolved *ctf4Δ/ctf4Δ* diploids clones distributed by chromosome (roman numbers). Dark blue diamonds represent loss of heterozygosity 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 influences 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 [73,80–84]). Our analysis revealed that all three architectures acquired putative adaptive mutations in three genome maintenance modules: sister chromatid cohesion, DNA replication and the DNA damage checkpoint. Out of the 23 populations analyzed, 18 of them contained at least one clone with mutations affecting all three of these modules, while all of them contain a

286 clone with mutations in two of the three modules (Figure 5A). Despite the ubiquity of mutations
287 in genome maintenance, the frequency of mutations in genes known to affect individual cellular
288 modules varied between genome architectures: mutations affecting sister chromatid cohesion
289 were underrepresented in recombination-deficient strains (p=2.7e-04, χ^2 test) and there were
290 fewer mutations affecting the DNA damage checkpoint in diploid clones (p=0.03, χ^2 test, Figure
291 6A).

292 If amplifying the genes for the cohesin loader, *SCC2* and *SCC4*, depended on homologous
293 recombination between repeated sequences, the rate of amplification should be dramatically
294 decreased in the recombination-deficient strain, which lacks Rad52. Analysis of the segmental
295 amplifications containing *SCC2* or *SCC4* revealed that their boundaries correspond to three
296 classes of repeated sequences: transposons (Ty elements), transposon-associated long terminal
297 repeats (LTRs) and tRNA genes (Figure 5-S1 and Supplementary table 7). Recombination at these
298 repeated elements has been previously shown to produce segmental amplifications selected
299 under limited nutrient availability [85–87]. Amplification of *SCC2* thus occurs as a result of
300 homologous recombination between two different copies of the repeated sequences on
301 chromosome IV, one on either side of *SCC2* and amplification of *SCC4* is due to similar events on
302 chromosome V. This dependence on homologous recombination explains the absence of
303 segmental amplifications in all but one of the recombination-deficient clones (Figure 3B and
304 Supplementary table 4). Because changes affecting sister chromatid cohesion rely predominantly
305 on gene amplifications (Figure 5C), we detect fewer mutations affecting this module in
306 recombination-deficient strains (Figure 5A).

307 In haploid *ctf4Δ* cells, inactivating the DNA damage checkpoint increases fitness [17]. The primary
308 target for mutation is *RAD9*, both because of its large size (3927 bp) and the presence of a run of
309 11 As, which repeatedly gain an extra nucleotide causing a frameshift mutation. Because this is a
310 loss-of-function mutation [17], removing Rad9 function in diploids requires two events, thus
311 reducing the probability of adapting by destroying the DNA damage checkpoint. Despite this
312 difficulty, one diploid clone is homozygous for the frameshift mutation in *RAD9*, most likely as a
313 result of a mutation in one copy of the gene followed by mitotic recombination or gene
314 conversion that homozygosed this mutation and fully inactivated the checkpoint. The need to
315 follow a two-hit pathway to fully inactivate the DNA damage checkpoint explains why fewer RAD9
316 mutations were found in diploids.

317 We conclude that although mutations affecting sister chromatid cohesion, DNA replication and
318 DNA damage checkpoint are found in all the three genome architectures, the frequency at which
319 we detect them depends on the genome architecture of an evolving population: gene
320 amplification occurs much less frequently in recombination-deficient populations, making one
321 pathway for improving sister chromatid cohesion less accessible, and recessive null mutations
322 are rarer in diploid populations, blocking a major route by which haploids inactivate the DNA
323 damage checkpoint.

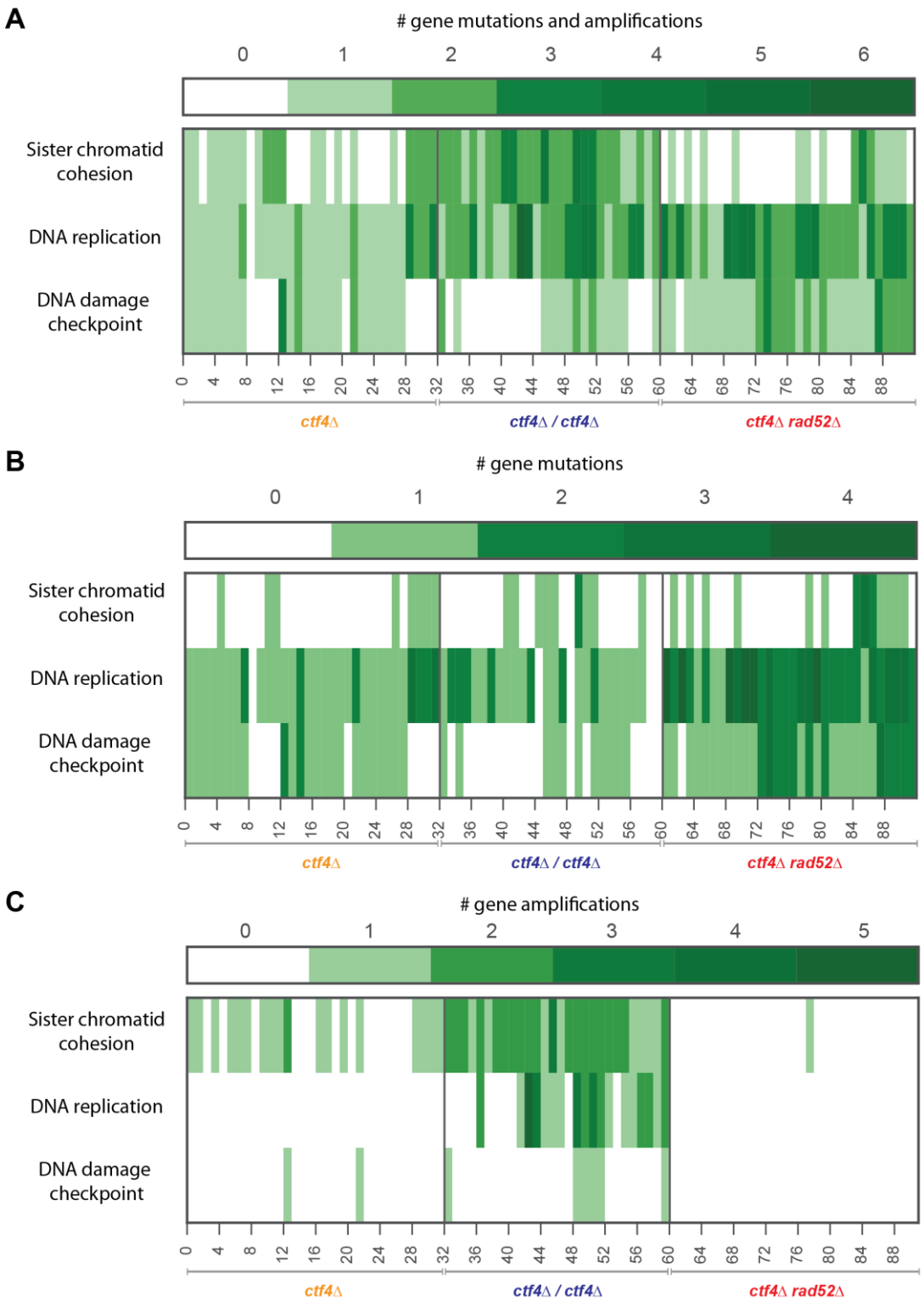


Figure 5: Mutations affecting sister chromatid cohesion, DNA replication and the DNA damage checkpoint. (A) Heatmaps representing all the mutations (SNPs, small InDels of 1-55bp and segmental amplifications) causally implicated in these phenotypes in the 92 sequenced evolved clones. **(B)** Frequency of SNPs and small InDels (1-55bp) affecting genes (Open reading frames and associated regulatory regions). **(C)** Frequency of genes associated with

329 the three processes that are present on segmental amplifications. The number of hits and the identities of the genes
330 which were mutated are reported in Source data 5.

331

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

334 We asked whether genome architecture affected the benefit that mutations conferred as well as
335 the frequency at which they were observed. To test this hypothesis, we chose a group of recurring
336 mutations to validate their causality and measure their fitness benefit in different genome
337 architectures. We mimicked the segmental amplification of chromosome IV (Figure 3B), affecting
338 the cohesin loader *SCC2*, by integrating an extra copy of the gene at the *URA3* locus (referred to
339 as *2xSCC2*). Ixr1 is a transcription factor that indirectly affects DNA replication by positively
340 regulating the production of deoxyribonucleotide triphosphates (dNTPs, [88]). The abundance of
341 mutations causing premature stop codons in *IXR1* suggested that evolution selected for the
342 inactivation of this gene (Supplementary table 8). We mimicked the loss of a functional Ixr1
343 protein by gene deletion (*ixr1⊿*). Another recurrently mutated gene affecting DNA replication
344 was *SLD5*, which encodes the subunit of the main replicative helicase (CMG, [89]) that is bound
345 *in vivo* by Ctf4 [47]. *SLD5* is an essential gene and our earlier analysis suggested that mutations
346 like *sld5-E130K* reduced or altered the helicase's function [17]. Finally, the majority of mutations
347 affecting the DNA damage checkpoint factor *RAD9* resulted in early stop codons (Supplementary
348 table 9), leading us to test the effect of the *rad9Δ* gene deletion. These mutations were all
349 previously demonstrated to be adaptive in haploid cells [17] and were used because they
350 occurred, albeit at different frequencies, in all three (point mutations in *SLD5* and *RAD9* and
351 amplification of *SCC2*) or at least two genome architectures (point mutations in *IXR1* in both
352 haploid populations).

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

362 For the diploid *ctf4Δ/ctf4Δ* strain, we tested each mutation as a heterozygote (e.g. *IXR1/ixr1Δ*)
363 and a homozygote (e.g. *ixr1Δ/ixr1Δ*). The homozygous deletion of *IXR1* gave a substantial fitness
364 benefit (Figure 6A, dark blue), but strains heterozygous for the deletion were slightly less fit than
365 their ancestors (Figure 6A, light blue). For the other three mutations (an extra copy of *SCC2*, *sld5-*
366 *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.

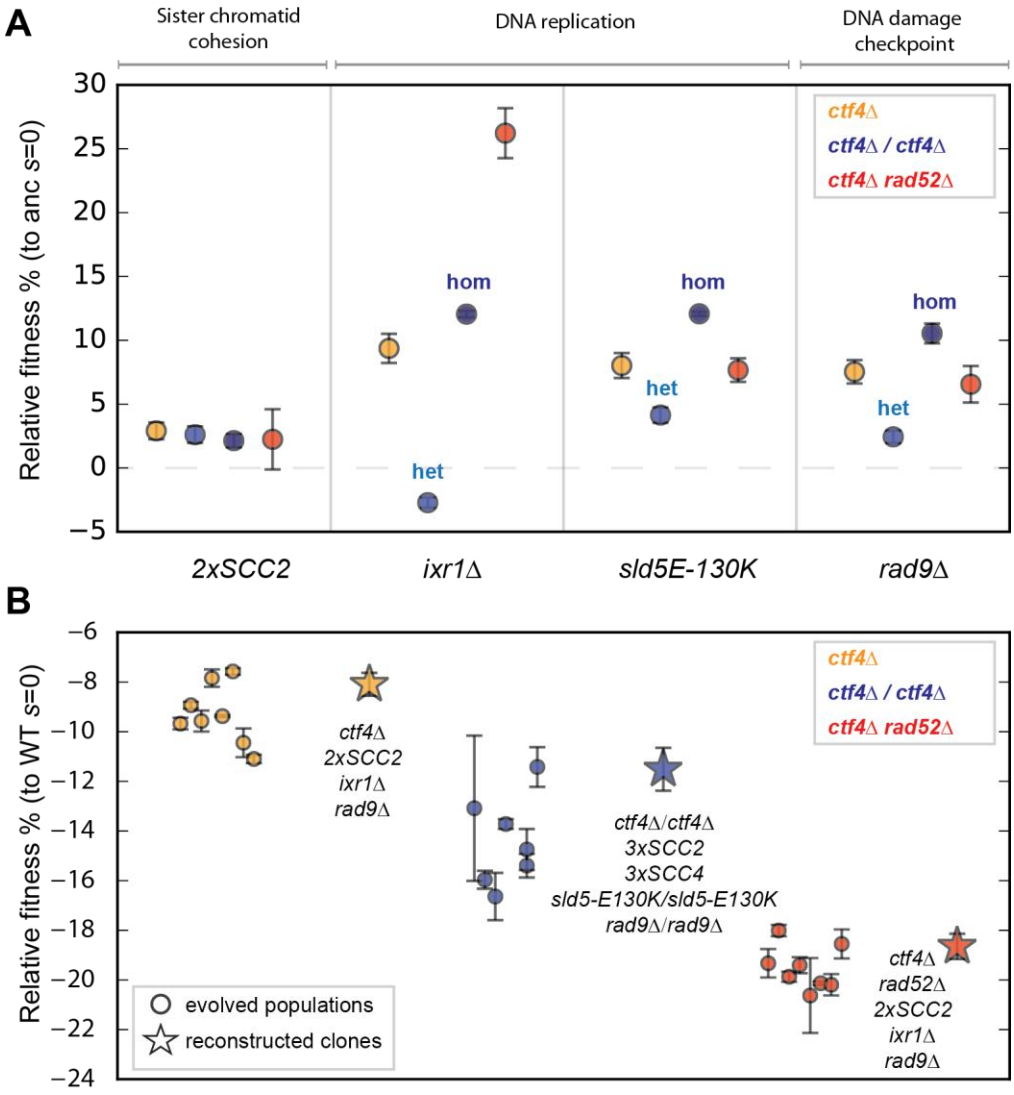


Figure 6: Reconstructed adaptive mutations can produce evolved fitness increases. (A) A subset of mutations affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint were re-constructed in the ancestor cells for all three genomic architectures. The fitness of the re-constructed strains, 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

383 diploids). Fitness data of haploid strains (orange) is from [17]. The P-values reported in figures are the result of t-
384 tests assuming unequal variances (Welch's test). The fitness values shown here are reported in Source data 6b.

385 Are mutations in sister chromatid cohesion, DNA replication and the DNA damage checkpoint
386 enough to recapitulate the evolved fitness increase in all three genomic architectures? We
387 previously showed that a haploid *ctf4Δ* strain carrying a mutation in all three of these modules
388 recapitulated the fitness of the evolved strains (Figure 6B and [17]). We therefore asked if
389 mutations in these three modules were sufficient to recapitulate 1000 generations of evolution
390 in diploid and recombination-deficient strains. We constructed strains that carried mutations
391 affecting sister chromatid cohesion, DNA replication and DNA damage checkpoint (Figure 6B),
392 using mutations that had been found in the same genome architecture (Figure 3,4 and
393 Supplementary table 1,4) and shown to be adaptive in isolation (Figure 6A).

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

399

400 **DISCUSSION**

401 We previously showed that haploid cells adapt to a form of constitutive DNA replication stress
402 by accumulating mutations affecting DNA replication, sister chromatid cohesion, and the DNA
403 damage checkpoint [17]. We repeated this experiment in two other genomic architectures, a
404 diploid and a recombination-deficient haploid strain, to ask which features of this evolutionary
405 trajectory were general and which were specific to a particular genomic architecture. We find
406 that the genes that acquire adaptive mutations, the frequency at which they are mutated, and
407 the frequency at which these mutations are selected all differ between architectures but that
408 mutations that confer strong benefits can occur in all three modules in each architecture.
409 Engineering one mutation in each module into an ancestral strain lacking Ctf4 is enough to
410 produce the evolved fitness increase in all three genomic architectures. Furthermore,
411 reconstruction of a panel of mutations into all three architectures proved they are adaptive even
412 in architectures where the affected genes were not found significantly mutated by the end of the
413 experiment. Altogether our results demonstrate the existence of a common pathway for yeast
414 cells to adapt to a form of constitutive DNA replication stress.

415 **The speed of adaptation to DNA replication stress depends on initial fitness not genome**
416 **architecture**

417 What determines the rate of evolutionary adaptation? One possibility is genome architecture:
418 beneficial, recessive mutations are harder to select in diploids [69,90] and gene amplification is
419 less frequent in recombination-defective cells [91,92], suggesting that recombination proficient

haploids should evolve faster than the other two genome architectures. Another explanation is that the initial fitness of the population is the principal factor in setting the rate of adaptation, with less fit ancestral populations increasing in fitness faster than more fit ones, a phenomenon referred to as declining adaptability [63]. We observed that adaptation rates scale linearly with the fitness difference between a mutant ancestor and a non-mutant wild-type, showing that the recovery from DNA replication stress follows the declining adaptability documented in other experimental conditions.

Previous work has produced a variety of results with the majority reporting that increasing ploidy decreases the rate of evolution [15,69,70] and a minority reaching the opposite conclusion [14,16]. We suggest that this variability arises because genome architecture affects two aspects of the mutations that produce adaptation: the frequency at which they occur and the selective advantage they confer. Thus, we find that defects in recombination reduce the rate of gene amplification and diploidy makes the accumulation of fully recessive mutations require two successive genetic events rather than a single mutation. Despite these barriers, our diploids and recombination-defective haploids evolved faster than recombination-proficient haploids. We can only speculate on why this is so but note that the recombination-proficient strains have a higher mutation rate and diploids may be more resistant to the unbalanced gene dosage caused by segmental amplifications. Our results thus go against the trend of slower adaptation in diploids as compared to haploids reported by the majority other studies [15,69,70]. This effect is not limited to populations experiencing DNA replication stress (Figure 2A) but is also present in control wild-type populations (Figure 2B). Our results support the idea that the details of genotypes, selections, and experimental protocols can determine the effect of ploidy on adaptation.

We engineered the same mutations into all three genome architectures to ask if the benefit of individual mutations depended on either genome architecture or the fitness of the recipient strain. For three of the four mutations we tested (*sld5-E130K*, *rad9Δ*, and 2x*SCC2*) their effect was independent of both these variables. The fourth (*ixr1Δ*) showed diminishing return epistasis, defined as an adaptive mutation producing a smaller benefit when present in fitter genetic backgrounds [65]. This effect has been invoked to explain declining adaptability and is thought to arise from the complex pattern of epistasis between the adaptive mutations and the set of alleles present in the genetic background [68,93–96]. Our results therefore agree with previous reports observing declining adaptability across strains with different initial fitness but largely fail to observe diminishing return epistasis as a potential justification of this phenomenon. Our experiments and two previous evolutionary repair experiments (Hsieh et al., 2020; Laan et al., 2015) both show interactions that are approximately additive between different selected mutations. The reasons for this difference are currently unknown.

General principles of adaptation to constitutive DNA replication stress

Evolutionary change can be examined at multiple levels from the details of adaptive mutations in individual genes, the identity of the functional modules in which the mutations lie, cellular

459 phenotypes, such as the integrity of the DNA damage checkpoint, to overall reproductive fitness.
460 For each of these levels, we examined the effect of genome architecture on the adaptation to a
461 form of constitutive DNA replication stress. At both ends of this scale we see substantial variation:
462 genome architecture influences the frequency at which mutations occur, the fitness benefit they
463 confer, and the extent of overall adaptation. In the middle, we see more conservation, with all
464 three architectures accumulating mutations in the same three functional modules, suggesting
465 that they all follow similar phenotypic paths by increasing the stability of the linkage between
466 sister chromatids, stabilizing replication forks, and inactivating the DNA damage checkpoint. To
467 test this hypothesis, we reconstructed mutations from all three modules into the ancestral, Ctf4-
468 deficient strains of all three genome architectures. This manipulation reproduced fitness
469 observed after 1000 generations of evolution leading us to conclude that the principles that
470 govern evolutionary adaptation to constitutive DNA replication stress are conserved across the
471 genome architectures we examined.

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

473 The phenotype of heterozygous mutations in diploid cells depends on their dominance. Although
474 loss-of-function mutations are often assumed to be fully recessive, a substantial minority have
475 detectable phenotypes as heterozygotes (referred to as semi-dominance or haploinsufficiency),
476 typically much weaker than that of the homozygous loss-of-function mutations [97]. The
477 phenotype of the heterozygote determines whether mutations will be selected for in diploid
478 populations [98,99]. As an example, loss-of-function mutations in *IXR1* were repeatedly found in
479 haploid and recombination-deficient populations but were not present in evolved diploid strains.
480 This observation is consistent with the slight fitness cost of the heterozygous *IXR1* deletion,
481 despite the strong benefit provided by the homozygous deletion: the cost of the heterozygote
482 acts as a barrier that prevents access to the beneficial homozygous mutations. Like *IXR1*, deletion
483 of both copies of *RAD9* has a strong fitness benefit in diploids, but in this case there is a small
484 fitness benefit in the *rad9Δ/RAD9* heterozygotes providing a path to inactivating the DNA damage
485 checkpoint in which both steps increase fitness. This two-step path in diploids is less favorable
486 than the one-step path that inactivates the checkpoint in haploids and likely accounts for the
487 lower frequency of mutations affecting *RAD9* and other known checkpoint genes in evolved
488 diploid clones (Figure 5 and Source data 5).

489 Two-step paths affected the replication machinery or DNA damage checkpoint in 9 of the 28
490 diploid clones; 2 involved independent mutations in the same gene and 8 produced homozygous
491 mutations due to gene conversion or mitotic recombination (Figure 4, one clone contained two
492 homozygous mutations). The high frequency of these events suggests that these mutations
493 conferred some advantage even as heterozygotes. For one mutation, *sld5-E130K*, we verified the
494 fitness benefit of the *sld5-E130K/SLD5* heterozygote, which suggests that the mutant subunit is
495 present in a fraction of the genome-wide replicative helicases, where it helps to stabilize
496 replication forks in the absence of Ctf4. Mitotic recombination can then homozygose the
497 mutation, eliminating the wild-type Sld5 protein and stabilizing all replisomes. The genome

498 architecture of diploid cells thus explains both the reduced frequency of mutations in *RAD9* and
499 dictates which mutations affecting DNA replication are ultimately selected (e.g. *IXR1* versus
500 *SLD5*). We suggest that semi-dominant mutations are an important intermediate in the selection
501 of loss-of-heterozygosity during the evolution of diploid, somatic cells.

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

503 Recombination-deficient clones have much fewer segmental amplifications affecting the genes
504 for the cohesin loader, *SCC2* and *SCC4* (Figure 5C). The lack of the recombinase protein Rad52 in
505 recombination-deficient strains dramatically reduces their ability to achieve segmental
506 amplifications. Only one event, likely mediated by Rad52-independent recombination [100–103],
507 was found in the recombination-deficient evolved clones (Figure 3B and Supplementary table 4).
508 The other mutations in these strains that are likely to affect sister chromatid cohesion were point
509 mutations affecting other proteins involved in pairing and segregation of chromosomes such as
510 Pds5 and Rad61 (Figure 5B). The higher number of point mutations in this and other modules is
511 likely the consequence of increased spontaneous mutagenesis imposed by the lack of homology-
512 mediated repair of spontaneously occurring DNA lesions (Figure 3-S1,2, [104]). The genome
513 architecture of recombination-deficient strains thus explains both the reduced frequency of gene
514 amplification affecting sister chromatid cohesion and the increased number of point mutations
515 detected in evolved clones (Figure 6B).

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

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

520 If the major adaptive mechanisms are shared, how do we explain the limited number of genes
521 that are mutated in all conditions and the genes that are uniquely and significantly mutated in
522 specific genome architectures (Figure 3A)? There are several hypotheses: some of these genes
523 may be false positives, mutations in these genes confer benefits in more than one architecture
524 but were only detected in a single architecture, and mutations in these genes are substantially
525 more, or exclusively, beneficial in one architecture. The fact that our reconstructed strains
526 recapitulate the fitness of the evolved populations with a minimal set of changes in genome
527 maintenance suggests that architecture-specific mutations play a marginal role in adaptation. Of
528 the 22 genes that were mutated in at least two architectures, 10 have known roles in DNA
529 replication, repair, or sister chromatid cohesion, whereas only 5 of the 28 that were uniquely
530 mutated in one architecture have such a role ($p = 0.034$, χ^2 test). The fitness of the reconstructed
531 strains (Figure 6A,B) also implies that the most important mutations for the adaptation to DNA
532 replication stress are those that were found in multiple architectures (Figure 3A). Mutations that
533 were never (*ixr1*), or rarely (*2xSCC2* and *rad9*) seen in a given genome architecture, nevertheless
534 increased fitness in all three architectures. This finding suggests that caution should be used in
535 inferring adaptive strategies from lists of genes that have been mutated more often than

536 expected by chance, especially when comparing populations that were evolved under different
537 conditions or have different ancestral genotypes. We argue that experimental dissection of
538 adaptive trajectories is needed to distinguish the mutations that make major contributions to
539 adaptation from minor players and mutational noise.

540 The distribution of mutations with known phenotypic effects suggests that genome architecture
541 controls how much adaptation in different functional modules contributes to the overall
542 response to selective pressure. Mutations that break the DNA damage checkpoint by inactivating
543 *RAD9* are less frequent in diploid populations than they are in haploids, and amplifications of
544 cohesin loader genes are less frequent in recombination-deficient populations than they are in
545 recombination-proficient ones. Although we suspect that some of the diploid populations have
546 found alternate genetic routes to partially or fully inactivate the DNA damage checkpoint and
547 that some of the recombination-defective populations have selected other mutations that
548 improve sister chromatid cohesion, these changes are likely to be less effective than those that
549 were frequent in other genome architectures.

550 **Implications for cancer and natural evolution**

551 We previously speculated that evolutionary processes similar to the adaptation to constitutive
552 DNA replication stress could assist the evolution of cancer cells and organisms in nature [17].
553 Tumor cells often experience DNA replication stress as a consequence of oncogene activation,
554 promoting genetic instability and increasing the rate of cell death [19,22]. While generated
555 through a different mechanism (unrestrained proliferation, rather than replisome perturbation),
556 oncogene induced DNA replication stress produces cellular consequences [25] which are
557 remarkably similar to the one described in the absence of Ctf4, such as the accumulation of
558 ssDNA gaps, stalled and reversed forks [17,48,49], genetic instability [49,51,52] and DNA damage
559 response activation [47,50]. Based on these similarities we speculate that the evolutionary
560 adaptation to DNA replication stress could reduce its negative effects on cellular fitness and thus
561 assist tumor evolution. In natural populations, interference with DNA replication, through the
562 fixation of deleterious mutations, perturbation by selfish genetic elements or toxins secreted by
563 antagonistic organisms could generate similar evolutionary adaptation. The consequences of
564 these evolutionary processes could rewire genome maintenance modules to alter cancer cells'
565 physiology and generate molecular diversity in nature [105]. For these scenarios to be plausible,
566 the evolutionary adaptation to DNA replication stress needs to happen over short timescales. In
567 cancer, it would need to happen over the limited somatic evolution that accompanies cancer
568 progression. In nature, it would need to occur before the perturbed organisms are outcompeted
569 by unaffected ones or block the source of DNA replication stress. It could be argued that some
570 genome architectures, such as diploidy or reduced recombination, could reduce the likelihood of
571 evolving new strategies for genome replication and maintenance. In our hands, however, these
572 genome architectures do not preclude rapid, evolutionary adaptation to constitutive DNA
573 replication stress: loss of heterozygosity or increased mutation rates, which have both been
574 documented in cancer [106,107] and natural evolution [108,109], allow cells to circumvent these

575 obstacles and adapt quickly to a severe perturbation of the replication machinery. One rapidly
576 evolving clade of yeasts has lost *RAD9* [110], the primary target of checkpoint-inactivating
577 mutations in our experiments. Similarly, many cancers inactivate the DNA damage response
578 during tumor evolution [111,112]. We suggest that the careful examination of the genomic
579 sequences of tumors and fast-evolving species might reveal signs of adaptation that mirror the
580 molecular evolution we have described.

581 **MATERIALS AND METHODS**

582 **Strains**

583 All strains were derivatives of a modified version (Rad5$^+$) of S. cerevisiae strain W303 (*leu2-3,112*
584 *trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, RAD5*). Supplementary table 10 lists each strain's
585 genotype. The ancestors of wild-type, *ctf4Δ* and *ctf4Δ rad52Δ* strains were obtained by
586 sporulating a *CTF4/ctf4Δ RAD52/rad52Δ* heterozygous diploid. This was done to minimize the
587 selection acting on the ancestor strains before the beginning of the experiment. *CTF4/CTF4* and
588 *ctf4Δ/ctf4Δ* diploid strains were obtaining by mating freshly generated *MAT***a** and *MAT*α haploid
589 strains.

590 **Media and growth conditions**

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

593 **Experimental evolution**

594 All populations used in the evolution experiment (EVO1-40) were initially inoculated in glass
595 tubes containing 10 ml of YPD with yeast colonies of the relative genotype. Glass tubes were
596 placed in roller drums at 30˚C and grown for 24 hr. Daily passages were done by diluting 10 μl of
597 the previous culture into 10 ml of fresh YPD (1:1000 dilution, allowing for approximately 10
598 generations/cycle). All populations were passaged for a total of 100 cycles (≈1000 generations).
599 In this regime, the effective population size is calculated as N_0 x g where N_0 is the size of the
600 population bottleneck at transfer and g is the number of generations achieved during a batch
601 growth cycle and corresponds to approximately to 10^7 cells. After 1000 generations four evolved
602 clones were isolated from the populations evolved in the absence of Ctf4 (EVO1-24) by streaking
603 cells on a YPD plate (total of 96 clones isolated). To capture the within-population genetic
604 variability we selected the clones displaying the largest divergence of phenotypes in terms of
605 resistance to genotoxic agents (methyl-methanesulfonate, hydroxyurea and camptothecin). All
606 final populations and clones isolated from them were mixed with 800 ml of 30% v/v glycerol and
607 stored at -70˚C for future analysis. The evolution of haploid, diploid, and recombination defective
608 strains, all lacking *CTF4*, and wild-type haploid and diploid strains were all carried out in parallel.

609 **Whole genome sequencing**

610 Genomic DNA library preparation was performed with an Illumina (RRID:SCR_010233, San Diego,
611 CA, US) Nextera DNA Library Prep Kit. Libraries were then pooled and sequenced either with an
612 Illumina HiSeq 2500 (125bp paired end reads) or an Illumina NovaSeq (150 bp paired end reads).
613 The SAMtools software package (RRID:SCR_002105, samtools.source- forge.net) was then used
614 to sort and index the mapped reads into a BAM file. GATK (RRID:SCR_ 001876,
615 www.broadinstitute.org/gatk; McKenna et al., 2010) was used to realign local indels, and VarScan
616 (RRID:SCR_006849, varscan.sourceforge.net) was used to call variants. Clones and populations
617 were sequenced at approximately the following depths: 25-30X for haploid clones, 50-60X for
618 diploid clones, 50-60X for haploid populations and 120-130X for diploid populations. Mutations
619 were found using a custom pipeline written in Python (RRID:SCR_008394, www.python.org). The
620 pipeline (github.com/koschwanez/mutantanalysis) compares variants between the reference
621 strain, the ancestor strain, and the evolved strains. Variants found in less than 25% and 35% of
622 the reads in haploid and diploid populations respectively were discarded, since many of these
623 corresponded to misalignment of repeated regions. For clone sequencing, only variants found in
624 more than 75% of the reads in haploids and 35% of the reads in diploids (to account for
625 heterozygosity) were considered mutations. The frequency of the reads associated with all the
626 variants detected are reported in Supplementary table 1. A variant that occurs between the
627 ancestor and an evolved strain is labeled as a mutation if it either (1) causes a substitution in a
628 coding sequence or (2) occurs in a regulatory region, defined as the 500 bp upstream and
629 downstream of the coding sequence. Gapped alignments of the 150 paired-end reads in our data
630 set permit the identification of small indels ranging in size from 1–55 bp using VarScan
631 pileup2indel tool [113]. All small indels (and the respective sequence affected) are listed
632 together with SNPs in Supplementary table 1.

633 **Identification of putative adaptive mutations**

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

635 *Parallel evolution of genes:* This method relies on the assumption that those genes that have
636 been mutated significantly more than expected by chance alone, represent cases of parallel
637 evolution among independent lines. The mutations affecting those genes are therefore
638 considered putatively adaptive. The same procedure was used independently on the mutations
639 found in wild-type and *ctf4Δ* evolved lines:

640 We first calculated per-base mutation rates as the total number of mutations in coding regions
641 occurring in a given background, divided by the size of the coding yeast genome in bp (including
642 1000bp per ORF to account for regulatory regions)

643 $$\lambda = \frac{SNPs + indels}{bp\text{coding}}$$

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

646 $$\mathrm{P}\ (n\ mutations|gene\ of\ length\ N) = \frac{(\lambda N)^n e^{-\lambda N}}{n!}$$

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

649 $$P(\neq\geq n\ mutations|gene\ of\ length\ N) = \sum_{k=n}^{\infty} \frac{(\lambda N)^n e^{-\lambda N}}{k!} = 1 - \frac{\Gamma\ (\ n+1, \lambda N)}{n!}$$

650 (Where Γ is the upper incomplete gamma function) which gives us the p-value for the comparison
651 of the observed mutations with the null, Poisson model. In order to decrease the number of false
652 positives, we then performed multiple-comparison corrections. The more stringent Bonferroni
653 correction (α=0.05) was applied on the wild-type evolved mutations dataset, while Benjamini-
654 Hochberg correction (α =0.05) was used for the *ctf4Δ* mutation dataset. Genes that were found
655 significantly selected in the evolved wild-type populations (after Bonferroni correction) were
656 removed from the list of evolved *ctf4Δ* strains (genes mutated in wild-type evolved populations
657 were removed from the *ctf4Δ* and *ctf4Δ rad52Δ* datasets and genes mutated in wild-type diploid
658 populations were removed from the *ctf4Δ/ctf4Δ* dataset). We argue that since these genes were
659 targets of selection even in wild-type cells, they are likely involved in processes that are unrelated
660 to DNA replication and are instead associated with adaptation to sustained growth by serial
661 dilutions. Supplementary table 2 lists the mutations detected in evolved clones and populations,
662 after filtering out those that occurred in genes that were significantly mutated in the respective
663 wild-type populations. Genes significantly selected are shown in dark grey (after Benjamini-
664 Hochberg correction with α=0.05).

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

672 **Gene ontology (GO) enrichment analysis**

673 The list of genes with putatively selected mutations (Figure 3A) or homozygous mutations in
674 diploids (Figure 4) were input as 'multiple proteins' in the STRING database, which reports on
675 the network of interactions between the input genes (https://string-db.org). The GO term
676 enrichment analysis provided by STRING are reported in Supplementary Table 3 and
677 Supplementary Table 6 respectively. Briefly, the strength of the enrichment is calculated as
678 Log_{10}(O/E), where O is the number of 'observed' genes in the provided list (of length N) which
679 belong to the GO-term, and E is the number of 'expected' genes we would expect to find
680 matching the GO-term providing a list of the same length N made of randomly picked genes. P-
681 values are computed using a Hypergeometric test and corrected for multiple testing using the

682 Benjamini-Hochberg procedure. The resulting P-values are represented as 'False discovery rate'
683 in the supplementary tables and describe the significance of the GO terms enrichment [114]."

684 **Fitness assays**

685 To measure relative fitness, we competed the ancestors and evolved strains against reference
686 strains. Cells were competed against reference strains with the same ploidy. Wild-type refence
687 strains were used for the competition assays in all the figures, except for those in Figure 6A where
688 *ctf4Δ* reference strains were used instead. A *pFA6a-prACT1-yCerulean-HphMX4* plasmid was
689 digested with AgeI and integrated at one of the *ACT1* loci of the original heterozygous diploid
690 (*CTF4/ctf4Δ RAD52/rad52Δ*) strain. This allows for the expression of fluorescent protein
691 yCerulean under the strong actin promoter. The heterozygous diploid was then sporulated and
692 dissected to obtain fluorescent wild-type, *ctf4Δ* or *ctf4Δ rad52Δ* reference haploid strains.
693 *ctf4Δ/ctf4Δ* diploid reference strains were obtained by mating two *ctf4Δ* haploid strains, one of
694 which expressed yCerulean under the *ACT1* promoter. For measuring a strain's relative fitness,
695 10 ml of YPD were inoculated in individual glass tubes with either the frozen reference or test
696 strains. After 24 hrs. the strains were mixed in fresh 10 ml YPD tubes at a ratio dependent on the
697 expected fitness of the test strain compared to the reference (i.e. 1:1 if believed to be nearly
698 equally fit) and allowed to proliferate at 30°C for 24 hrs. 10 μl samples were taken from this
699 mixed culture (day 0) and the ratio of the two starting strains was immediately measured. Tubes
700 were then cultured following in the same conditions as the evolution experiment by diluting them
701 1:1000 into fresh medium every 24hrs for 2-4 days, monitoring the strain ratio at every passage.
702 Samples were analyzed by flow cytometry (Fortessa, BD Bioscience, RRID:SCR_013311, Franklin
703 Lakes, NJ, US). Strain ratios and the number of generations occurred between samples were
704 calculated using a custom Python script utilizing the FlowCytometryTools package
705 (https://pypi.org/project/FlowCytometryTools/). Ratios, *r*, were calculated based on the number of
706 fluorescent and non-fluorescent events detected by the flow cytometer:

707 $$r = \frac{NonFluorescent_{events}}{Fluorescent_{events}}$$

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

710 $$g = \frac{\log_{10}(Events_{t24}/events_{t0})}{\log_{10} 2}$$

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

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

717 Whole genome sequencing and read mapping was done as previously described. The read-depths
718 for every unique 100 bp region in the genome were then obtained by using the VarScan

719 copynumber tool. A custom pipeline written in python was used to visualize the genome-wide
720 CNVs. First, the read-depths of individual 100 bp windows were normalized to the genome-wide
721 median read-depth to control for differences in sequencing depths between samples. The
722 coverage of the ancestor strains was then subtracted from the one of the evolved lines to reduce
723 the noise in read depth visualization due to the repeated sequences across the genome. The
724 resulting CNVs were smoothed across five 100 bp windows for a simpler visualization. Final CNVs
725 were then plotted relative to their genomic coordinate at the center of the smoothed window.
726 Since the wild-type CNVs were subtracted from the evolved CNVs, the y axis refers to the copy
727 number change occurred during evolution (i.e. +1 means that one an extra copy of a chromosome
728 fragment has been gained in haploid cells, and that two extra copies have been gained in
729 diploids). Segmental amplifications were identified as continuous chromosomal regions affected
730 by the same change in copy number (Figure 3B and Supplementary table 4). Amplified genes
731 (Figure 5A,C) were obtained from the lists of ORFs included within the boundaries of the
732 segmental amplifications.

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

734 We first identified the approximate boundaries (± 500bp) of all segmental amplification by
735 manually inspecting the BAM files with the Integrative Genomic Viewer
736 (https://software.broadinstitute.org/software/igv/). The coordinates of the sequence features that
737 define the boundary between lower and higher read-depths are reported in Supplementary table
738 4. We then examined the sequences within these approximate boundaries to determine whether
739 various chromosomal features were over- or underrepresented. We considered features that
740 previous studies have found to be associated with hotspots for lesions and sources of genetic
741 instability. We first counted how many times a given feature fell in the boundary zones of
742 segmental amplifications. Then we calculated the expected number of features in these zones
743 based on the total number of features in the genome and the percentage of the genome
744 represented by segmental amplification boundaries zones. We compared these numbers by $\chi 2$
745 analysis and reported the associated p-values (Supplementary table 7). The number of tRNA
746 genes, transposable elements, LTRs, ARS elements, snRNA and snoRNA genes and centromeres
747 in the genome were determined using YeastMine (https://yeastmine.yeastgenome.org/).
748 Coordinates of fragile sites were obtained from [115]. G4 sequences were obtained from [116].
749 Highly-(top 5%) and weakly-(least 5%) transcribed genes were identified from the data in [117].
750 Rrm3 binding sites and regions with high levels of γH2AX were derived from [118] and [119],
751 respectively. Sites of DNA replication termination were derived from [17]. The tandemly repeated
752 sequences, with a minimal repeat tract of twenty-four bases, were obtained from the tandem-
753 repeat-database (TRDB; https://tandem.bu.edu/cgi-bin/trdb/trdb.exe).

754 **DATA AND CODE AVAILABILITY**

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

759 **ACKNOWLEDGEMENTS**

We thank Michael Laub, María Angélica Bravo Núñez, Sriram Srikant, Thomas LaBar and Andrea Giometto for critical reading of the manuscript; Andrea Giometto for assistance in data analysis; Claire Hartman from the Harvard Bauer Core Facility for technical assistance. We thank the members of the Murray and Nelson labs for helpful discussions. MF acknowledges fellowship support from the Human Frontiers Science Program, EMBO and AIRC.

AUTHOR CONTRIBUTIONS

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

BIBLIOGRAPHY

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

2. Dowell RD, Ryan O, Jansen A, Cheung D, Agarwala S, Danford T, et al. Genotype to phenotype: A Complex problem. Science. American Association for the Advancement of Science; 2010. p. 469. doi:10.1126/science.1189015

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

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

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

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

7. Helsen J, Voordeckers K, Vanderwaeren L, Santermans T, Tsontaki M, Verstrepen KJ, et al. Gene loss predictably drives evolutionary adaptation. Mol Biol Evol. 2020. doi:10.1093/molbev/msaa172

8. Harari Y, Ram Y, Rappoport N, Hadany L, Kupiec M. Spontaneous Changes in Ploidy Are Common in Yeast. Curr Biol. 2018;28: 825-835.e4. doi:10.1016/j.cub.2018.01.062

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

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

11. Rojas Echenique JI, Kryazhimskiy S, Nguyen Ba AN, Desai MM. Modular epistasis and the compensatory evolution of gene deletion mutants. Butler G, editor. PLOS Genet. 2019;15:

799 e1007958. doi:10.1371/journal.pgen.1007958

800 12. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, et al. The Genomic Landscape of
801 Compensatory Evolution. Barton NH, editor. PLoS Biol. 2014;12: e1001935.
802 doi:10.1371/journal.pbio.1001935

803 13. Fumasoni M. Tell Me Where You've Been and I'll Tell You How You'll Evolve. MBio. 2020;11:
804 2043–2063. doi:10.1128/mBio.02043-20

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

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

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

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

813 18. Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nat Cell Biol. 2014;16:
814 2–9. doi:10.1038/ncb2897

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

817 20. Branzei D, Foiani M. Maintaining genome stability at the replication fork. Nat Rev Mol Cell Biol.
818 2010;11: 208–219. doi:10.1038/nrm2852

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

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

824 23. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C, et al. Oncogene-induced
825 senescence is a DNA damage response triggered by DNA hyper-replication. Nature. 2006;444:
826 638–642. doi:10.1038/nature05327

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

830 25. Kotsantis P, Petermann E, Boulton SJ. Mechanisms of oncogene-induced replication stress: Jigsaw
831 falling into place. Cancer Discovery. American Association for Cancer Research Inc.; 2018. pp.
832 537–555. doi:10.1158/2159-8290.CD-17-1461

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

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

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

839 29. Scheifele LZ, Cost GJ, Zupancic ML, Caputo EM, Boeke JD. Retrotransposon overdose and genome
840 integrity. Proc Natl Acad Sci U S A. 2009;106: 13927–13932. doi:10.1073/pnas.0906552106

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

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

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

848 33. Walsh CS. Two decades beyond BRCA1/2: Homologous recombination, hereditary cancer risk and
849 a target for ovarian cancer therapy? Gynecologic Oncology. Academic Press Inc.; 2015. pp. 343–
850 350. doi:10.1016/j.ygyno.2015.02.017

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

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

855 36. Krajewska M, Fehrmann RSN, De Vries EGE, van Vugt MATM. Regulators of homologous
856 recombination repair as novel targets for cancer treatment. Frontiers in Genetics. Frontiers
857 Media S.A.; 2015. p. 96. doi:10.3389/fgene.2015.00096

858 37. Villa F, Simon AC, Ortiz Bazan MA, Kilkenny ML, Wirthensohn D, Wightman M, et al. Ctf4 Is a Hub
859 in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. Mol Cell.
860 2016;63: 385–96. doi:10.1016/j.molcel.2016.06.009

861 38. Crabbé L, Thomas A, Pantesco V, De Vos J, Pasero P, Lengronne A. Analysis of replication profiles
862 reveals key role of RFC-Ctf18 in yeast replication stress response. Nat Struct Mol Biol. 2010;17:
863 1391–1397. doi:10.1038/nsmb.1932

864 39. Wilhelm T, Olziersky AM, Harry D, De Sousa F, Vassal H, Eskat A, et al. Mild replication stress
865 causes chromosome mis-segregation via premature centriole disengagement. Nat Commun.
866 2019;10: 1–14. doi:10.1038/s41467-019-11584-0

867 40. Zheng D-Q, Zhang K, Wu X-C, Mieczkowski PA, Petes TD. Global analysis of genomic instability
868 caused by DNA replication stress in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A.
869 2016;113: E8114–E8121. doi:10.1073/pnas.1618129113

870 41. Vesela E, Chroma K, Turi Z, Mistrik M.
871 Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. Biomolecules.
872 2017;7: 19. doi:10.3390/biom7010019

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

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

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

881 45. Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, et al. A Ctf4 trimer couples the
882 CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature. 2014/05/09.
883 2014;510: 293–297. doi:10.1038/nature13234

884 46. Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, et al. A key
885 role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic
886 replisome. EMBO J. 2009/08/08. 2009;28: 2992–3004. doi:10.1038/emboj.2009.226

887 47. 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. 2009/06/06. 2009;14: 807–820.
889 doi:10.1111/j.1365-2443.2009.01310.x

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

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

896 50. Poli J, Tsaponina O, Crabbé L, Keszthelyi A, Pantesco V, Chabes A, et al. dNTP pools determine
897 fork progression and origin usage under replication stress. EMBO J. 2012;31: 883–894.
898 doi:10.1038/emboj.2011.470

899 51. Kouprina N, Kroll E, Bannikov V, Bliskovsky V, Gizatullin R, Kirillov A, et al. CTF4 (CHL15) mutants
900 exhibit defective DNA metabolism in the yeast Saccharomyces cerevisiae. Mol Cell Biol.
901 1992/12/01. 1992;12: 5736–5747. Available: http://www.ncbi.nlm.nih.gov/pubmed/1341195

902 52. Hanna JS, Kroll ES, Lundblad V, Spencer F a. Saccharomyces cerevisiae CTF18 and CTF4 are
903 required for sister chromatid cohesion. Mol Cell Biol. 2001/04/05. 2001;21: 3144–3158.
904 doi:10.1128/MCB.21.9.3144-3158.2001

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

907 54. Symington LS. Role of RAD52 Epistasis Group Genes in Homologous Recombination and Double-
908 Strand Break Repair. Microbiol Mol Biol Rev. 2002;66: 630–670. doi:10.1128/mmbr.66.4.630-
909 670.2002

910 55. Lu YB, Ratnakar P V, Mohanty BK, Bastia D. Direct physical interaction between DnaG primase
911 and DnaB helicase of Escherichia coli is necessary for optimal synthesis of primer RNA. Proc Natl
912 Acad Sci USA. 1996;93: 12902–7. doi:10.1073/pnas.93.23.12902

913 56. Yoshizawa-Sugata N, Masai H. Roles of human AND-1 in chromosome transactions in S phase. J
914 Biol Chem. 2009/05/15. 2009;284: 20718–20728. doi:10.1074/jbc.M806711200

915 57. Yeeles JT, Poli J, Marians KJ, Pasero P. Rescuing stalled or damaged replication forks. Cold Spring
916 Harb Perspect Biol. 2013/05/03. 2013;5: a012815. doi:10.1101/cshperspect.a012815

917 58. Gerrish PJ, Lenski RE. The fate of competing beneficial mutations in an asexual population.
918 Genetica. 1998;102/103: 127–144. doi:10.1023/A:1017067816551

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

921 60. Venkataram S, Dunn B, Li Y, Agarwala A, Chang J, Ebel ER, et al. Development of a Comprehensive
922 Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast. Cell. 2016;166: 1585-
923 1596.e22. doi:10.1016/J.CELL.2016.08.002

924 61. Gerstein AC, Chun H-JE, Grant A, Otto SP. Genomic Convergence toward Diploidy in
925 Saccharomyces cerevisiae. Dutcher S, editor. PLoS Genet. 2006;2: e145.
926 doi:10.1371/journal.pgen.0020145

927 62. Gerstein AC, Otto SP. Cryptic fitness advantage: Diploids invade haploid populations despite
928 lacking any apparent advantage as measured by standard fitness assays. PLoS One. 2011;6:
929 26599. doi:10.1371/journal.pone.0026599

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

932 64. Wiser JM, Ribeck N, Lenski RE. Long-Term Dynamics of Adaptation in Asexual Populations.
933 Science (80-). 2013;342: 1364–1367. doi:10.1126/science.1243357

934 65. Chou H, Chiu H, Delaney NF, Segrè D, Marx CJ. Diminishing Returns Epistasis Among Beneficial
935 Mutations Decelerates Adaptation. Science (80-). 2011;332: 1190–1193.
936 doi:10.1126/science.1203799

937 66. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial
938 mutations in an evolving bacterial population. Science (80-). 2011;332: 1193–1196.
939 doi:10.1126/science.1203801

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

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

946 69. Marad DA, Buskirk SW, Lang GI. Altered access to beneficial mutations slows adaptation and
947 biases fixed mutations in diploids. Nat Ecol Evol. 2018;2: 882–889. doi:10.1038/s41559-018-
948 0503-9

949 70. Zeyl C, Vanderford T, Carter M. An Evolutionary Advantage of Haploidy in Large Yeast
950 Populations. Science (80-). 2003;299: 555–558. doi:10.1126/SCIENCE.1078417

951 71. Von Borstel RC, Cain KT, Steinberg CM. Inheritance of spontaneous mutability in yeast. Genetics.
952 1971;69: 17–27. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1212686/

953 72. Kunz BA, Peters MG, Kohalmi SE, Armstrong JD, Glattke M, Badiani K. Disruption of the RAD52

954 gene alters the spectrum of spontaneous SUP4-o mutations in Saccharomyces cerevisiae.
955 Genetics. 1989;122: 535–542. Available: /pmc/articles/PMC1203727/?report=abstract

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

959 74. Rolef Ben-Shahar T, Heeger S, Lehane C, East P, Flynn H, Skehel M, et al. Eco1-dependent cohesin
960 acetylation during establishment of sister chromatid cohesion. Science (80-). 2008/07/26.
961 2008;321: 563–566. doi:10.1126/science.1157774

962 75. Panizza S, Tanaka T, Hochwagen A, Eisenhaber F, Nasmyth K. Pds5 cooperates with cohesion in
963 maintaining sister chromatid cohesion. Curr Biol. 2000;10: 1557–1564. doi:10.1016/S0960-
964 9822(00)00854-X

965 76. Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature
966 separation of sister chromatids. Cell. 1997/10/23. 1997;91: 35–45. doi:doi: 10.1016/s0092-
967 8674(01)80007-6.

968 77. Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Nasmyth K. Cohesin's binding to
969 chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Mol Cell.
970 2000/07/06. 2000;5: 243–254. doi:10.1016/S1097-2765(00)80420-7

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

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

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

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

982 82. Loeillet S, Herzog M, Puddu F, Legoix P, Baulande S, Jackson SP, et al. Trajectory and uniqueness
983 of mutational signatures in yeast mutators. Proc Natl Acad Sci U S A. 2020;117: 24947–24956.
984 doi:10.1073/pnas.2011332117

985 83. Fisher KJ, Buskirk SW, Vignogna RC, Marad DA, Lang GI. Adaptive genome duplication affects
986 patterns of molecular evolution in Saccharomyces cerevisiae. Otto SP, editor. PLOS Genet.
987 2018;14: e1007396. doi:10.1371/journal.pgen.1007396

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

991 85. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, et al. Characteristic genome
992 rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A.
993 2002;99: 16144–16149. doi:10.1073/pnas.242624799

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

997 87. Lauer S, Avecilla G, Spealman P, Sethia G, Brandt N, Levy SF, et al. Single-cell copy number variant
998 detection reveals the dynamics and diversity of adaptation. de Visser JAGM, editor. PLOS Biol.
999 2018;16: e3000069. doi:10.1371/journal.pbio.3000069

1000 88. Tsaponina O, Barsoum E, Åström SU, Chabes A. Ixr1 is required for the expression of the
1001 ribonucleotide reductase Rnr1 and maintenance of dNTP pools. PLoS Genet. 2011;7.
1002 doi:10.1371/journal.pgen.1002061

1003 89. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H. GINS, a novel multiprotein
1004 complex required for chromosomal DNA replication in budding yeast. Genes Dev. 2003;17: 1153–
1005 1165. doi:10.1101/gad.1065903

1006 90. Orr HA, Otto SP. Does diploidy increase the rate of adaptation? Genetics. 1994;136: 1475–1480.
1007 Available: /pmc/articles/PMC1205926/?report=abstract

1008 91. Tlsty TD, Albertini AM, Miller JH. Gene amplification in the lac region of E. coli. Cell. 1984;37:
1009 217–224. doi:10.1016/0092-8674(84)90317-9

1010 92. Slack A, Thornton PC, Magner DB, Rosenberg SM, Hastings PJ. On the Mechanism of Gene
1011 Amplification Induced under Stress in Escherichia coli. Matic I, editor. PLoS Genet. 2006;2: e48.
1012 doi:10.1371/journal.pgen.0020048

1013 93. Reddy G, Desai MM. Global epistasis emerges from a generic model of a complex trait. bioRxiv.
1014 2020; 2020.06.14.150946. doi:10.1101/2020.06.14.150946

1015 94. Wünsche A, Dinh DM, Satterwhite RS, Arenas CD, Stoebel DM, Cooper TF. Diminishing-returns
1016 epistasis decreases adaptability along an evolutionary trajectory. Nat Ecol Evol. 2017;1: 0061.
1017 doi:10.1038/s41559-016-0061

1018 95. Schoustra S, Hwang S, Krug J, de Visser JAGM. Diminishing-returns epistasis among random
1019 beneficial mutations in a multicellular fungus. Proc R Soc B Biol Sci. 2016;283: 20161376.
1020 doi:10.1098/rspb.2016.1376

1021 96. Lyons DM, Zou Z, Xu H, Zhang J. Idiosyncratic epistasis creates universals in mutational effects
1022 and evolutionary trajectories. Nat Ecol Evol. 2020;4: 1685–1693. doi:10.1038/s41559-020-01286-
1023 y

1024 97. Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, et al.
1025 Mechanisms of Haploinsufficiency Revealed by Genome-Wide Profiling in Yeast. Genetics.
1026 2005;169: 1915–1925. doi:10.1534/GENETICS.104.036871

1027 98. Gerstein AC, Otto SP. Ploidy and the Causes of Genomic Evolution. J Hered. 2009;100: 571–581.
1028 doi:10.1093/jhered/esp057

1029 99. Otto SP, Gerstein AC. The evolution of haploidy and diploidy. Current Biology. Elsevier; 2008. pp.
1030 R1121–R1124. doi:10.1016/j.cub.2008.09.039

1031 100. Schacherer J, de Montigny J, Welcker A, Souciet JL, Potier S. Duplication processes in
1032 Saccharomyces cerevisiae haploid strains. Nucleic Acids Res. 2005;33: 6319–6326.

1033 doi:10.1093/nar/gki941

1034 101. Coïc E, Feldman T, Landman AS, Haber JE. Mechanisms of Rad52-independent spontaneous and
1035 UV-induced mitotic recombination in Saccharomyces cerevisiae. Genetics. 2008;179: 199–211.
1036 doi:10.1534/genetics.108.087189

1037 102. Prado F, González-Barrera S, Aguilera A. RAD52-dependent and -independent homologous
1038 recombination initiated by Flp recombinase at a single FRT site flanked by direct repeats. Mol
1039 Gen Genet. 2000;263: 73–80. doi:10.1007/PL00008677

1040 103. Payen C, Koszul R, Dujon B, Fischer G. Segmental Duplications Arise from Pol32-Dependent
1041 Repair of Broken Forks through Two Alternative Replication-Based Mechanisms. Haber JE, editor.
1042 PLoS Genet. 2008;4: e1000175. doi:10.1371/journal.pgen.1000175

1043 104. Naiman K, Pagès V, Fuchs RP. A defect in homologous recombination leads to increased
1044 translesion synthesis in E. coli. Nucleic Acids Res. 2016;44: 7691–7699. doi:10.1093/nar/gkw488

1045 105. LaBar T, Phoebe Hsieh Y-Y, Fumasoni M, Murray AW. Evolutionary Repair Experiments as a
1046 Window to the Molecular Diversity of Life. Curr Biol. 2020;30: R565–R574.
1047 doi:10.1016/J.CUB.2020.03.046

1048 106. Russo M, Crisafulli G, Sogari A, Reilly NM, Arena S, Lamba S, et al. Adaptive mutability of
1049 colorectal cancers in response to targeted therapies. Science (80-). 2019;366: 1473–1480.
1050 doi:10.1126/science.aav4474

1051 107. Thiagalingam S, Laken S, Willson JKV, Markowitz SD, Kinzler KW, Vogelstein B, et al. Mechanisms
1052 underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci U S A.
1053 2001;98: 2698–2702. doi:10.1073/pnas.051625398

1054 108. D'Angiolo M, Chiara M De, Yue J-X, Irizar A, Stenberg S, Persson K, et al. A yeast living ancestor
1055 reveals the origin of genomic introgressions. Nature. 2020; 1–6. doi:10.1038/s41586-020-2889-1

1056 109. Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, et al. Genome evolution across
1057 1,011 Saccharomyces cerevisiae isolates. Nature. 2018;556: 339–344. doi:10.1038/s41586-018-
1058 0030-5

1059 110. Steenwyk JL, Opulente DA, Kominek J, Shen X-X, Zhou X, Labella AL, et al. Extensive loss of cell-
1060 cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. Kamoun S, editor.
1061 PLOS Biol. 2019;17: e3000255. doi:10.1371/journal.pbio.3000255

1062 111. Ghosal G, Chen J. DNA damage tolerance: a double-edged sword guarding the genome.
1063 Translational Cancer Research. 2013. pp. 107–129. doi:10.3978/j.issn.2218-676X.2013.04.01

1064 112. Hohl M, Mojumdar A, Hailemariam S, Kuryavyi V, Ghisays F, Sorenson K, et al. Modeling cancer
1065 genomic data in yeast reveals selection against ATM function during tumorigenesis. Symington
1066 LS, editor. PLOS Genet. 2020;16: e1008422. doi:10.1371/journal.pgen.1008422

1067 113. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation
1068 and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22:
1069 568–76. doi:10.1101/gr.129684.111

1070 114. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1:
1071 Protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res.

1072 2013;41: D808. doi:10.1093/nar/gks1094

1073 115. Song W, Dominska M, Greenwell PW, Petes TD. Genome-wide high-resolution mapping of
1074 chromosome fragile sites in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2014;111:
1075 E2210-8. doi:10.1073/pnas.1406847111

1076 116. Capra JA, Paeschke K, Singh M, Zakian VA. G-Quadruplex DNA Sequences Are Evolutionarily
1077 Conserved and Associated with Distinct Genomic Features in Saccharomyces cerevisiae. Stormo
1078 GD, editor. PLoS Comput Biol. 2010;6: e1000861. doi:10.1371/journal.pcbi.1000861

1079 117. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, et al. The Transcriptional
1080 Landscape of the Yeast Genome Defined by RNA Sequencing. Science (80-). 2008;320: 1344–
1081 1350.

1082 118. Azvolinsky A, Dunaway S, Torres JZ, Bessler JB, Zakian VA. The S. cerevisiae Rrm3p DNA helicase
1083 moves with the replication fork and affects replication of all yeast chromosomes. Genes Dev.
1084 2006;20: 3104–16. doi:10.1101/gad.1478906

1085 119. Szilard RK, Jacques P-E, Laramée L, Cheng B, Galicia S, Bataille AR, et al. Systematic identification
1086 of fragile sites via genome-wide location analysis of gamma-H2AX. Nat Struct Mol Biol. 2010;17:
1087 299–305. doi:10.1038/nsmb.1754

1088