Published in final edited form as:

Nature. 2019 February ; 566(7743): 275–278. doi:10.1038/s41586-019-0887-y.

Heterozygous mutations cause genetic instability in a yeast model of cancer evolution

Miguel C. Coelho[1,2,3], **Ricardo M. Pinto**[3,4], **Andrew W. Murray**[1,2]

[1]FAS Center for Systems Biology, NW 469, 52 Oxford St., Cambridge, MA, 02138, USA;

[2]Molecular and Cell Biology, Harvard University, NW 469, 52 Oxford St., Cambridge, MA, 02138, USA;

[3]Center for Genomic Medicine, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA;

[4]Department of Neurology, Harvard Medical School, Boston, MA 02115, USA

Abstract

Genetic instability, a heritable increase in the mutation rate, accelerates evolutionary adaptation[1] and is widespread in cancer[2,3]. In mammals, instability can arise from damaging both copies of genes involved in DNA metabolism and cell cycle regulation[4] or from inactivating one copy of a gene whose product is present in limiting amounts (haploinsufficiency[5]), but determining the relative importance of these two mechanisms is difficult. In *E. coli*[6], applying repeated, strong selection enriches for genetic instability. We used this approach to evolve genetic instability in diploid cells of the budding yeast, *Saccharomyces cerevisiae* and isolated clones with increased rates of point mutation, mitotic recombination, and chromosome loss. We identified candidate, heterozygous, instability-causing mutations and engineering these mutations, as heterozygotes, into the ancestral diploid strain caused genetic instability. Mutations that inactivate one copy of haploinsufficient genes are more common than those that dominantly alter the function of the mutated gene copy. The mutated genes are enriched for genes functioning in transport, protein quality control, and DNA metabolism, and reveal new targets for genetic instability[7–11], including

Correspondence and requests for materials should be addressed to M.C. (costacoelho@fas.harvard.edu/ mcc150@gmail.com) or A.W.M. (awmurray@fas.harvard.edu).

Author contributions M.C., R.M.P. and A.W.M. designed the experiments, M.C. and R.M.P. performed the experiments, M.C., R.M.P. and A.W.M. performed data analysis and wrote the manuscript.

The authors declare no competing financial interests.

Supplementary information is available in the online version of the paper (Supplementary Tables 1–4 and Extended Data Figures 1–8).

Materials availability.
The authors declare that the yeast strains (Supplementary Table 4) and HAP1 human cell lines created (Extended Data Fig. 8c) in this study are available, upon reasonable request.

Code availability.
The authors declare that the Python scripts used for the analysis of yeast genomic sequences and frequency of segregation are available online (https://github.com/koschwanez/mutantanalysis) or upon request.

Data availability.
The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. Source data for all figures are provided with the paper (online). Genomic sequencing data is available from the authors, upon reasonable request.

essential genes. Although only a minority (10 out of 57 genes with orthologs or close homologs) of the targets we identified have homologous human genes implicated in cancer[2], the remainder are candidates to contribute to human genetic instability. To test this hypothesis, we inactivated six examples in a near haploid human cell line: five of these mutations increased instability. We conclude that single genetic events cause genetic instability in diploid yeast cells, and propose that similar, heterozygous mutations in mammalian homologs initiate genetic instability in cancer.

Cancer requires multiple changes in cellular properties, and the accumulation of the mutations that produce these changes can be accelerated by genetic instability. The origin and mechanism of genetic instability is difficult to study in cancer making it hard to assess the relative importance of evolutionary trajectories that inactivate both copies of a tumor suppressor gene (Knudson's two-hit model[12]) and those that mutate only one of two copies of a gene, exposing haploinsufficiency[5] or dominantly altering a protein's function. Sequencing the genomes of human tumors reveals mutations that are likely to cause genetic instability[2,13–18], but cannot easily reveal whether these mutations are homozygous or heterozygous. By directly selecting for genetic instability in a more experimentally tractable organism, budding yeast, we assessed the relative importance of the one- and two-hit routes to genetic instability.

Studies on model organisms have primarily sought recessive mutations[7–11] that cause genetic instability. To avoid this restriction, we evolved genetic instability in diploid budding yeast by selecting for sequential inactivation of three sets of growth suppressor genes: genes that can prevent cell proliferation under certain conditions. Because these mutants are recessive, diploids should inactivate both copies of a growth suppressor gene to survive. Formally, the genes that mutate to cause instability and the growth suppressor genes are analogous to two classes of tumor suppressor genes: instability-causing mutations identify genes analogous to tumor suppressors that guard the stability of the genome, such as p53, whereas growth suppressor genes are analogous to tumor suppressor genes that regulate cell growth and proliferation, such as the retinoblastoma gene. By selecting for sequential inactivation of growth suppressor genes, we selected for genetic instability, by using random mutagenesis, we avoided the restrictions imposed by systematic collections of deletion mutants, and by using diploids containing homozygous, dominant growth suppressors, we mimicked the selection to inactivate tumor suppressors in cancer evolution.

We searched for genetic instability by selecting clones that survived and proliferated after sequentially exposing diploid yeast to three drugs that suppress growth of wild-type cells (Fig. 1a). We used a reporter cassette containing the *URA3* gene to monitor the level and type of genetic instability in diploid strains (Fig. 1b), distinguishing three classes of mutation: point mutation (local mutations that affect the *URA3* gene), mitotic recombination (the sum of mitotic recombination and chromosome truncation), and chromosome loss (Fig. 1c). We tested the reporter using environmental[19] (Extended Data Fig. 1a) or genetic[10] (Extended Data Fig. 1b) perturbations that cause different types of genetic instability. The triple selection selected more strongly for genetic instability than any of the single selection steps (Extended Data Fig. 2a).

Starting from mutagenized populations, we selected and characterized 89 genetically unstable diploid clones (Fig. 1b, Supplementary Table1). All had an increase in at least one class of genetic instability (Fig. 1d), 44 had an increase in two of the three classes of mutation, 27 had an increase in all three, and 46 had at least a 10-fold elevation in one class of mutation rate. Increases in chromosome loss and mitotic recombination were correlated within individual clones (Extended Data Fig. 2b) and all three mutation classes were correlated with the global mutation rate (Extended Data Fig. 2c). We conclude that selecting for growth suppressor inactivation selects for genetic instability, mimicking an important feature of cancer evolution.

We asked how many events were necessary to trigger genetic instability by examining genetic instability in the four haploid products of meiosis from our diploid, evolved clones; we made the assumption that most mutations that increase genetic instability in diploid strains would also cause instability in haploids (Extended Data Fig. 3a). In 59 out of 61 clones we tested, some of the haploid spores are genetically unstable. We first assessed the minimum number of mutations required for instability (Fig. 2a). In most clones (41/59), genetic instability segregated as expected if a single heterozygous mutation was enough to cause instability; in the remaining clones instability appeared to require the simultaneous presence of two or more heterozygous mutations (16/59), or a single homozygous mutation (2/59). We also asked which fraction of the spores had the maximum genetic instability (Extended Data Fig. 3b): 29/59 clones contained only one mutation that affects genetic instability, 17/59 showed the 1:3 pattern expected if they contain two mutations that affect genetic instability, and 11/59 show the 1:7 pattern expected if they contained three mutations that affect genetic instability.

To identify mutations that correlate with genetic instability, we performed whole genome sequencing of the 89 evolved clones. For genes which were mutated in ≥ 4 independent lineages (25 genes), or carried nonsense mutations (10 genes), we deleted one copy in the ancestor and found this sufficed to cause an increase in genetic instability (Fig. 2b and Extended Data Fig. 4a–d).

To identify less frequent mutations that correlated with genetic instability, we performed bulk segregant analysis[20,21] on the 48 diploid clones that exhibited the highest genetic instability: we pooled and whole genome sequenced genetically unstable spores, arguing that mutations that caused genetic instability should be at high frequency in these pools, whereas other, non-causative mutations should show a mean frequency of 0.5. 39 clones had one to three heterozygous alleles whose frequency was above the cutoff frequency of 0.67 in the selected, genetically unstable pools (Extended Data Fig. 5a). We identified a total of 74 candidate genes that had mutated to cause genetic instability (Extended Data Fig. 5b, Supplementary Table2). Based on the frequency with which different candidate genes were mutated, we estimate that there are roughly 150 genes that can mutate to produce dominant genetic instability in diploids. Seven genes were identified as both frequently mutated and by bulk segregant analysis.

We were surprised at the prevalence of heterozygous mutations that segregated with genetic instability. We analyzed seventeen candidate heterozygous, genetic-instability causing

mutations in more detail. We engineered each mutation, as a heterozygote (mut/+), into the ancestral diploid and measured its effect on mutation rate. For all seventeen genes, engineering one copy of the mutant allele into a diploid strain increased the frequency of at least one form of genetic instability from 2- to 82-fold (Fig. 2c and Extended Data Fig. 6a–c).

By comparing the phenotype of strains with one mutant and one wild type allele (mut/+) versus one deleted and one wild type copy of the gene (Δ/+), we can distinguish two ways in which the mutant allele causes genetic instability. The first is haploinsufficiency: if the amount of the gene product is limiting in a diploid cell, removing one copy will cause genetic instability and the phenotypes of the mut/+ and Δ/+ strains should be similar (Fig. 2d, top). The second is that the mutant copy of the gene actively interferes with chromosome metabolism predicting that the phenotype of the mut/+ strain should be stronger than that of the Δ/+ strain (Fig. 2d, middle). For 10 out of 17 genes, at least one form of instability exceeded that of the wild-type strain by at least two standard deviations in mut/+ than the Δ/+ analog, indicating that dominant mutations are altering protein function (Fig. 2c and Extended Data Fig. 6a–c).

We tested the effect of restoring the wild type version of a mutated allele in three types of genetically unstable strains: evolved clones, strains engineered to contain one copy of a mutation shown to cause genetic instability, and strains that contained heterozygous mutations of genes that are haploinsufficient for genetic stability[22]. For each type of strain, we tested three different genes and for eight out of nine genes restoring homozygosity for the wild-type allele reduced genetic instability, supporting the argument that a single heterozygous mutation is sufficient to cause instability (Extended Data Fig. 6d).

The 92 genes we identified are associated with nine high level biological processes (Fig. 3a and Extended Data Fig. 7a): transport genes are significantly enriched, even after Bonferroni correction, and protein quality control and DNA metabolism genes pass a 5% false discovery cutoff. 22% of the candidate genes we identified as targets for genetic instability are essential (Supplementary Table 2). There was little overlap between genes identified in deletion screens and our evolution screen (Extended Data Fig. 7b and Supplementary Table 3). Of the yeast genes that we identified, 54 have human orthologs, 3 have close homologs, and 10 of these have been implicated in cancer[2,13–18] (Fig. 3b, Extended Data Fig. 7c–f and Supplementary Table 3).

We asked if inactivating six genes whose homologs lacked known cancer associations caused genetic instability in a near-haploid human cell line (HAP1)[23]. Two of these genes were identified both in our experiments and previous systematic screens for genetic instability in yeast (*MUS81* and *TAO3*) and four were unique to our work. The products of these genes participate in recombination (*MUS81*), metabolism (*DYRK1B)*, mitochondria (*DDHD1*) the cytoskeleton (*FRYL*, *GRID2IP*), and autophagy (*ATG2B)* in human cells. We used hypoxanthine-guanosine phosphoribosyl transferase gene (*HPRT1*) inactivation to monitor genetic instability, by exposing cells to 6-thioguanine (6TG, Extended Data Fig. 8a). As a control, we inactivated *SMARCA4*, which is implicated in human cancer and whose mutation has been shown to cause genetic instability[24], and which is a homolog of *SNF2*, a

yeast gene involved in chromatin remodeling. *SMARCA4* inactivation resulted in 150-fold increase in mutation rate relative to wildtype (Fig. 3c). Individual inactivation of all but one of the other six genes caused an elevation in mutation rate from 4- to 440-fold in human cells (Fig. 3c and Extended Data Fig. 8b,c). Inactivating human homologs of genes identified by selecting for genetic instability in diploid yeast cells causes genetic instability and suggests that mutations in these genes are candidates to contribute to human cancer.

We found many mutations that cause genetic instability as heterozygotes, demonstrating that a single mutational event can accelerate evolutionary adaptation in diploid cells and that strong, directional selection can select for such mutations. The 92 genes we identified as targets for instability include 21 genes involved in DNA and chromosome metabolism, whose altered function is expected to cause genetic instability. The remainder have roles in other essential aspects of cellular behavior, suggesting that perturbations in these pathways interfere with DNA and chromosome metabolism enough to cause genetic instability. In both yeast and bacteria, strong directional selection leads to the evolution of genetic instability[6,25]. In yeast, selecting for genetic instability yields one-hit, heterozygous, instability-causing mutations much more often than, two-hit, homozygous mutations. We argue that the selection for multiple changes in cellular properties that occurs in cancer also selects for genetic instability and that, as they do in yeast, heterozygous mutations can lead to genetic instability. Both in yeast and in cancer, further selection could enhance the initial instability of heterozygous mutations by eliminating the remaining wild-type copy of the genes, or selecting additional heterozygous, dominant mutations in other genes (Fig. 3d).

Of the 57 genes with human homologs, 10 have been implicated in human cancer. We hypothesize that the remaining 47 that have close human homologs are candidates to mutate to cause genetic instability in cancer. For five of the six genes we tested, eliminating these genes triggered genetic instability suggesting that mutations in many of the human homologs of the genes that we identified are candidates to play a role in the initiation, progression and therapy resistance of human cancer.

Methods

Strains and growth media (yeast).

Sacharomyces cerevisiae W303 was used as a base strain to build the genetic instability reporter cassette in our ancestor strains. The strains we used are corrected for the *rad5–535* mutation found in most W303 strains. Knock-out and reconstruction of evolved mutations were performed on our ancestor diploid strain (yMC003) or generated by transforming the α −mating type haploid (yMC001) and crossing with the untransformed a-mating type haploid (yMC002, see Extended Data table 4). Strains were grown in YPD (10 g/l yeast extract, 20 g/l peptone, and 2% [w/v] dextrose) or CSM (10g/l YNB, 2% [w/v] dextrose) –URA liquid media, or on CSM-5FOA, CSM-ADE, CSM-HIS 2% agar plates.

HAP1 cell culture and CRISPR/Cas9-mediated gene inactivation (human cells).

Near-haploid, HAP1 human cells were obtained directly from Horizon Discovery and cultured according to standard cell culture protocols in Iscove's Modified Dulbecco's

Medium (IMDM), supplemented with 10% FBS and 1% PenStrep. Isogenic cell lines deficient for the human homologs of yeast instability genes (*ATG2B*, *SMARCA4*, *DYRK1B*, *GRID2IP*, *DDHD1*) were generated using CRISPR/Cas9-mediated inactivation. In summary, single-guide RNAs (sgRNAs) were designed against the target genes (Extended Data Fig. 8c), while attempting to maximize on-target activity and minimizing off-target activity[26]. Individual sgRNAs were cloned into a spCas9- and GFP-expressing plasmid (pX458)[27] and transfected into wildtype HAP1 cells using Lipofectamine 3000. Individual clones were generated from single GFP-positive cells isolated by FACS (MoFlo XDP Sorter, Beckman Coulter). Sanger sequencing was used to identify isogenic clones for each target gene, carrying small frameshift insertions/deletions in early exons, which were predicted to generate no active protein products (Extended Data Fig. 8c). The Δ*MSH2* and Δ*FRYL* HAP1 lines were obtained directly from Horizon Discovery, while the Δ*MUS81* HAP1 lines[28] were a gift from the Matos lab (IBC-ETH, Zurich).

Construction of a genetic instability reporter.

We wanted to compare the genetic instability in our ancestor and evolved strains. We engineered a *URA3* reporter cassette that could report the global mutation rate and distinguish between point mutation, recombination or truncation and chromosome loss events. A copy of the *HIS3* gene linked directly to *URA3* on the left arm of chromosome I and the only wild type copy of the *ADE1* gene lies on the opposite side of the centromere of the chromosome. The phenotypes for growth on 5FOA, -HIS, and -ADE media distinguish the different types of mutational events. To create the reporter, we first transformed a haploid *MAT***a** strain auxotrophic for histidine (*his3–11,15*) by deleting the *URA3* coding region seamlessly *(ura3Δ0)*, and replacing *ADE1* with the Hph:MX cassette (Hygromycin resistance) – strain yMC001. In a haploid *MATα* strain auxotrophic for histidine (*his3–11,15) we* deleted *URA3* seamlessly *(ura3Δ0)*, leaving the ADE1 locus on the right arm of chromosome I intact (yMC002). Then we inserted a cassette containing the budding yeast *URA3* and heterologous *HIS3* genes into chromosome I of yMC002 in the following arrangement, reading out from the centromere: the endogenous ARS106 replication origin locus, *URA3* and *KlHIS3* (homolog from *Saccharomyces Kluyveri*). We generated the diploid ancestor (yMC003) by crossing yMC001 with yMC002. The diploid is homozygous for all loci that contribute to growth suppression (*TRP1-TRP5, LYS2, LYS5* and *CAN1*, and heterozygous at chromosome I, where *URA3, KlHIS3* and *ADE1* are present on one copy and the other lacks all three genes (Fig. 1c). Histidine auxotrophs following *URA3* inactivation represent recombination or truncation events, while histidine and adenine double-auxotrophs identify chromosome loss events.

Perturbation experiments to test genetic instability reporter.

We perturbed our ancestor with environmental stimuli that increased the frequency of point mutations, recombination and chromosome loss. Cells were treated with 1% Ethyl MethaneSulfonate (EMS) in PBS for 30min, (to induce point mutations), neutralized with 5% Sodium thiosulfate [w/v] and allowed to recover for 2h in CSM-URA, or irradiated with 5.1 mJ/cm^2 of Ultraviolet (UV, λ=254nm) on a CSM-URA plate (to induce point mutations and mitotic recombination), or treated with 30μg/ml Benomyl for 2h (to induce chromosome loss). We also tested if mutations known to increase genetic instability[29,30] would generate a

detectable signal in our reporter. We transformed our ancestor strain to introduce a *POL3-L523D/POL3* heterozygous mutation (increased point mutation due to inability to proof-read[30]), or deleted both copies of either *RAD27* (whose deletion increases loss of heterozygosity[29]) or *MAD2* (whose deletion increases chromosome loss[31]). Cells were grown in CSM-URA, plated in 5FOA ($0.5–1\times10^7$cells) and surviving colonies (*URA3* mutants) were replica plated into CSM-HIS or CSM-ADE.

Mutagenesis.

To introduce heterozygous mutations and increase the genetic variability in our starting populations, we used UV and EMS mutagenesis (at levels which produced ~200 mutations per diploid ancestor clone). The diploid strain yMC003 was grown to saturation (10^8cells/ml), plated on YPD plates (10^6cells) exposed to 5.1mJ/cm^2 of UV radiation (~90% killing haploids, ~60% killing diploids) in YPD plates, grown in the dark (to prevent photo-lyase mediated pirimidine dimer repair) for two days, and the surviving colonies were washed off with liquid YPD and stored in aliquots at −80C in YPD plus 20% glycerol. The same wash and storage procedure was performed following chemical mutagenesis produced by incubating ancestor cells in 1% EMS in 0.1M Sodium Phosphate buffer, pH 7.0 for 30min. EMS was inactivated with 5% Na-Thiosulfate. Aliquots of 10^7cells from UV or EMS mutagenesis were used as starting populations to sample genetic diversity of our evolution experiment.

Selecting for sequential inactivation of growth suppressor genes.

To select for inactivation of growth suppressor genes, we used drugs that are toxic to cells expressing the active version of the gene, thus selecting for cells that have inactivated the gene. The *CAN1* gene encodes an arginine permease that allows canavanine, a toxic arginine analog, to enter yeast cells and block their growth[32], 5-fluoro-anthranilate (5-FAA) is converted into the toxic analog 5-fluoro-tryptophan by the action of the *TRP1-TRP5* genes[33], and α-aminoadipate is lethal when used as a sole nitrogen source unless cells have inactivated *LYS2* or *LYS5*34. Because these mutants are recessive, diploids should inactivate both copies of a growth suppressor gene to survive. Selecting for repeated inactivation of growth suppressor genes should select for genetic instability[35]. To start our experiment with clones that have functional growth suppressor genes we first plated the mutagenized ancestor in medium lacking tryptophan, and lysine. We pooled the surviving colonies into liquid YPD. The first selection step was performed by growing cells in YPD, then plating on 5-FAA. Colonies that grew on 5-FAA were replica-plated to YPD to increase the number of cells. The second selection step was performed by replica-plating these colonies from YPD onto α-aminoadipate. Survivors on this medium were replica-plated to YPD and then these amplified colonies were replica-plated to canavanine (selecting for inactivation of *CAN1*). In the starting diploid, heterozygous deletions for the growth suppressor genes are sensitive to the growth suppressors. At the end of the evolution experiment, approximately 70% of surviving clones (n=105) were resistant to canavanine and unable to grow in CSM–TRP and CSM–LYS media.

Genetic instability quantification (yeast).

Clones that survived selection (evolved clones) were tested for their genetic instability. We measure the mutation frequency as the *URA3* inactivation frequency determined by plating cells on 5FOA-containing medium. To avoid the effects of mutations that occur early in the exponential expansion of a culture (jackpots), cultures were grown in CSM-URA so that cells that inactivated *URA3* would be unable to continue proliferating, Cultures were grown until they reached exponential growth phase ($1–5\times10^7$ cells/ml), sonicated, counted using a Coulter counter and plated onto 5FOA plates ($5\times10^{6-7}$cells). To verify the specificity of the 5FOA selection, we took 20 colonies of the ancestor (yMC003) that survived on 5FOA plates, tested their ability to grow on CSM-URA and analyzed their *URA3* genes. All 20 failed to grow on CSM-URA, 19 had lost the *URA3* gene and one had a nonsynonymous point mutation (475C>A) in *URA3.* To determine how cells became 5FOA-resistant, we used our genetic instability reporter (Fig. 1c) and replica-plated surviving colonies from 5FOA into CSM-HIS and CSM-ADE, thus distinguishing point mutations (His^+ Ade^+), mitotic recombination (His^- Ade^+), and chromosome loss (His^- Ade^-).

Genetic instability quantification (human cells).

We selected a set of 6 genes based on how frequently they were mutated in independently evolved yeast lineages, the magnitude of the increase in genetic instability in reconstructed yeast clones, and to sample different functional classes (ATG2, mutated in 10 lines, 5-fold elevation in overall mutation rate, metabolism; DDL1, mutated in 2 lines, 3-fold elevation in overall mutation rate, mitochondria; YAK1 mutated in 6 lines, 2-fold elevation in overall mutation rate, cell-cycle; TAO3 mutated in 11 lines, 2.5-fold elevation in overall mutation rate, transcription; BNI1 mutated in 8 lines, 9-fold elevation in overall mutation rate, cytoskeleton; MUS81 mutated in 2 lines, 7-fold elevation in overall mutation rate, DNA). To quantify mutation rates in human cells we used the hypoxanthine-guanine phosphoribosyl transferase (HPRT1) loss of function (LOF) frequency assay[36]. In the presence of 6-Thioguanine (6-TG), HPRT1-expressing cells die while HPRT1 LOF mutants survive. By comparing the frequency of survival between the CRISPR knock-out cell lines and the wildtype we can estimate the relative increase in mutation rate. In order to eliminate pre-existing HPRT1 mutations, cells were treated for 2–3 days with hypoxanthine-aminopterin-thymidine (HAT) media. The optimal concentration for 6-TG in our assay was established using a dose-effect curve on wildtype (6-TG sensitive) and Δ*MSH2* (6-TG resistant) cell lines, with concentrations ranging from 0.1 to 100μM of 6-TG. The effective concentration of 3.75μM 6-TG killed all wildtype cells without inducing toxicity in the control, Δ*MSH2* line. Cells were counted using a Countess II FL cell counter, and 10^7 cells were seeded in a single-well plate followed by treatment with 3.75μM 6-TG for 14 days, with media changes performed every 1–3 days. Surviving colonies were stained with Hoechst 33342 (NucBlue Live Ready Probes, ThermoFisher) and the entire surface of the plate ($95cm^2$) was imaged at 10X magnification on an inverted fluorescence microscope (Leica DMi8). The total number of colonies per CRISPR KO cell line (n=3–6 replicates) was counted using the cell counter plugin in Fiji, using a fixed binary threshold and a minimal colony area cutoff (6 cells on day 10 or 12 cells on day 14) for image segmentation.

Genomic DNA preparation.

To prepare genomic DNA, cultures were pelleted and resuspended in 50 µl of 1% Zymolyase in 0.1 M, pH 8.0 NaEDTA. The cells were incubated for 30 min at 37°C to digest the cell wall, and then the cells were lysed by adding 50 µl 0.2 M NaEDTA, 0.4 M pH 8.0 Tris, 2% SDS and incubated at 65°C for 30 min. 63 µl of 5 M potassium acetate was added, and the mixture was incubated for 30 min on ice. The insoluble residue was then pelleted, and 750 µl of ice-cold ethanol was added to 300 µl of the supernatant to precipitate the DNA. The DNA was pelleted, and the pellet was resuspended in 0.2 mg/ml RNAase A. After 1 hr of incubation at 37°C, 2 µl of 20 mg/ml Proteinase K was added, and the solution was incubated for an additional 2 hr at 37°C. The DNA was again precipitated by adding 130 µl isopropanol. The DNA was pelleted, briefly washed with 70% ethanol, repelleted, and resuspended in 100 µl 10 mM Tris, pH 8.0.

Whole genome sequencing.

We used the standard Nextera library preparation kit (Illumina, California), optimized to generate 300–600 bp dsDNA fragments after tagmentation (adapted for small volume library preparation[37]). DNA was analyzed using QuBit BR dsDNA quantification, TapeStation (Agilent HS D1000) electrophoretic analysis and qPCR (Agilent BioAnalyzer 2100). Double paired-end 125–150bp reads were obtained from the HiSeq2000 BauerCore facility sequencer. Evolved clones were sequenced at a depth 10–20x, while bulk segregant pools were sequenced at a depth 30–120x.

Bulk segregant analysis of genetic instability.

To sporulate the diploid strains, cultures were grown to saturation in YPD, and then diluted 1:50 into YP 2% acetate. The cells were grown in acetate for 12–24 hr and then pelleted and resuspended in 2% acetate. After 2–4 days of incubation on a roller drum at 25°C, sporulation was verified by observing the cells in a microscope. To digest ascii, 1 ml of the sporulated culture was pelleted and resuspended in 50 µl 10% Zymolyase (www.zymoresearch.com) for 1 hr at 30°C. 400 µl of water and 50 µl of 10% Triton X-100 were added, and the digested spores were sonicated for 5–10 s to separate the tetrads. Tetrad separation was verified by observing the cells in a microscope. The separated tetrads were then spun down slowly (6000 rpm) and resuspended in water and plated in YPD. Replica plating to CSM-URA allowed us to confirm correct segregation of the URA3 marker (2:2). Individual URA+ haploid segregants were tested (~400/evolved clone) for URA3 mutation frequency in 5FOA, and the top 10% of genetically unstable segregants (~40/evolved clone), which should contain mutations that cause instability, were pooled for gDNA extraction.

Sequence analysis.

DNA sequences were aligned to the S288C reference genome r64 (downloaded from the Saccharomyces Genome Database, www.yeastgenome.org) using the Burrows-Wheeler Aligner[38] (bio-bwa.sourceforge.net). The resulting SAM (Sequence Alignment/Map) file was converted to a BAM (binary SAM) file, sorted, indexed, and made into a pileup format file using the samtools software[38] (samtools.sourceforge.net). Variants including single-nucleotide polymorphisms and insertions/deletions were called from the pileup file using the

Varscan software[39] (varscan.sourceforge.net). To perform the segregation analysis, we wrote a custom sequencing pipeline in Python (www.python.org), using the Biopython (biopython.org) and pysam (code.google.com/p/pysam/) modules, that finds sequence variants between the ancestor and clone, classifies each variant as a nonsynonymous coding region, synonymous coding region, or promoter mutation, and ranks each mutation by its segregation frequency[40].

Strain reconstruction (yeast).

Alleles were replaced by transformation of the ancestor (yMC003) with a linearized DNA fragment carrying the mutated version of the allele with a selectable marker downstream of the gene and followed by intergenic DNA to direct recombination (reconstructed (mut/+); *KAN:MX or NAT:MX*). Transformants were selected in appropriate media (2xG418 or Clonat). Alleles were disrupted by replacing the full ORF with a Kan:MX or a Nat:MX cassete and selected in a similar way (Δ/+). Colonies that grew were streaked out on YPD, then transferred a second time to the selective media. The target sequence was PCR amplified and Sanger sequenced to verify correct insertion. DNA polymerase, polynucleotide kinase, and DNA ligase were purchased from New England Biolabs (www.neb.com). In eight cases, the reconstructed (mut/+) strain had at least one type of mutation rate (either global, PM, RT or CL) higher than the evolved clone that harbored the causative mutation (Fig. 2c, Extended Data Fig. 6a–c). Possible explanations include differences in cell physiology due to the accumulation of unselected mutations in the evolved clones, and selection for mutations that partially suppress genetic instability after the inactivation of the last growth suppressor gene.

Venn diagram construction for genetic instability targets comparison.

We first grouped the human driver genes from 6 independent cancer driver studies[41–46]. Then we identified by reciprocal blast analysis 273 yeast genes that are orthologous or closely homologous to of 1310 human cancer drivers. To be included in this set the yeast and human genes had to satisfy one of three criteria: 1) they were reciprocal best blast hits, 2) the E value for the hit of the yeast gene to the cancer gene was <1E-10 and the difference between the best hit in the human genome and the hit to the cancer gene is <1E10-fold, or 3) the E value for the best hit in the human genome is 0 and the E value for the hit to the cancer gene is <1E-100. We compared three groups of yeast genes: genes whose mutations led to genetic instability in our work (92 genes), genes that were identified as conferring genetic instability in three systematic screens of the yeast gene deletion collection[29,47,48], and random genes in the yeast genome. To test the overlap between the latter two groups and the cancer-homologous dataset, we randomly generated 6 subgroups of 92 genes for the 648 genes that confer instability when deleted, and 6 subgroups for the 5150 annotated yeast genes (SGD). In this way each group of yeast genes had the same number of genes. The average and standard deviation of the overlap between the 6 subgroups and the genetic instability genes and cancer homolog genes is presented inside the Venn diagram in Extended Data Fig. 7b.

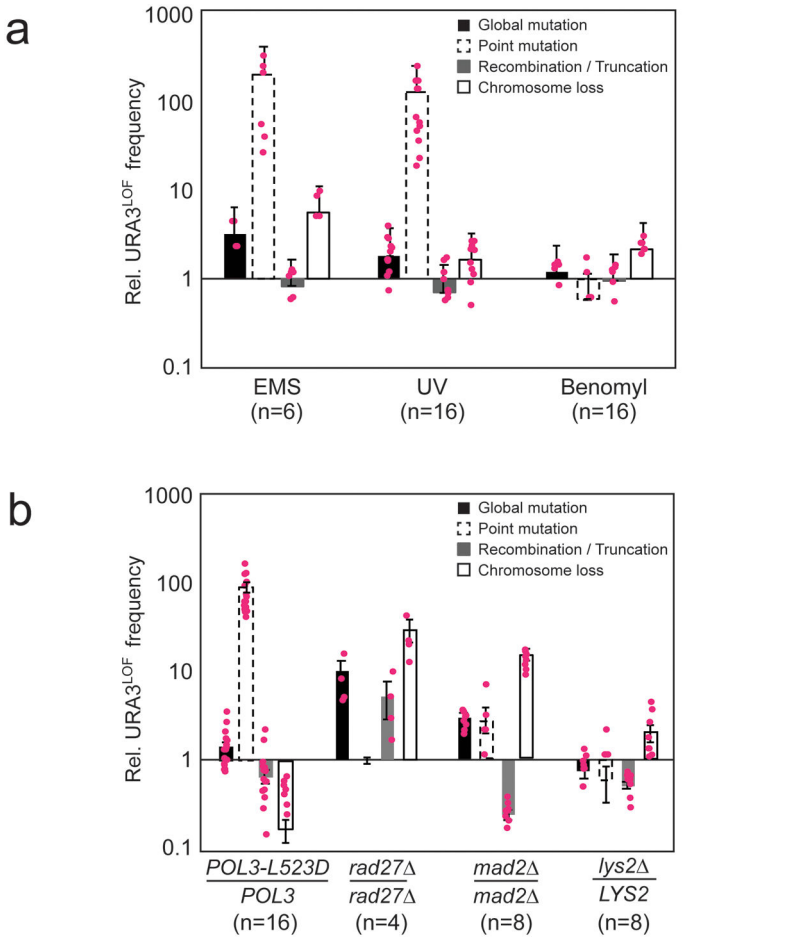
Gene ontology enrichment for functional classes.

We curated the functions of the genetic instability genes using the SGD database, GO entries, and available literature, and divided the genes into 9 major functional classes: cell cycle, cytoskeleton, DNA, metabolism, mitochondria, protein quality control, transcription, translation and transport. We performed two statistical tests to determine if there was a significant enrichment in any of the functional classes, when comparing with the functional distribution of all yeast genes: whether the probability of the number of observed hits was < 0.05 using the Bonferroni correction, and whether hits passed a 5% false discovery cutoff.

Figures and data analysis.

Unless otherwise noted, data analysis was performed on Excel (Microsoft), Matlab (Mathworks) or in the R programming language (www.r-project.org) and plots were generated using the R library ggplot2 (*Wickham, Elegant Graphics for Data Analysis, Springer, 2009*). Adobe Illustrator was used to format and assemble the plotted data into figures.

Extended Data

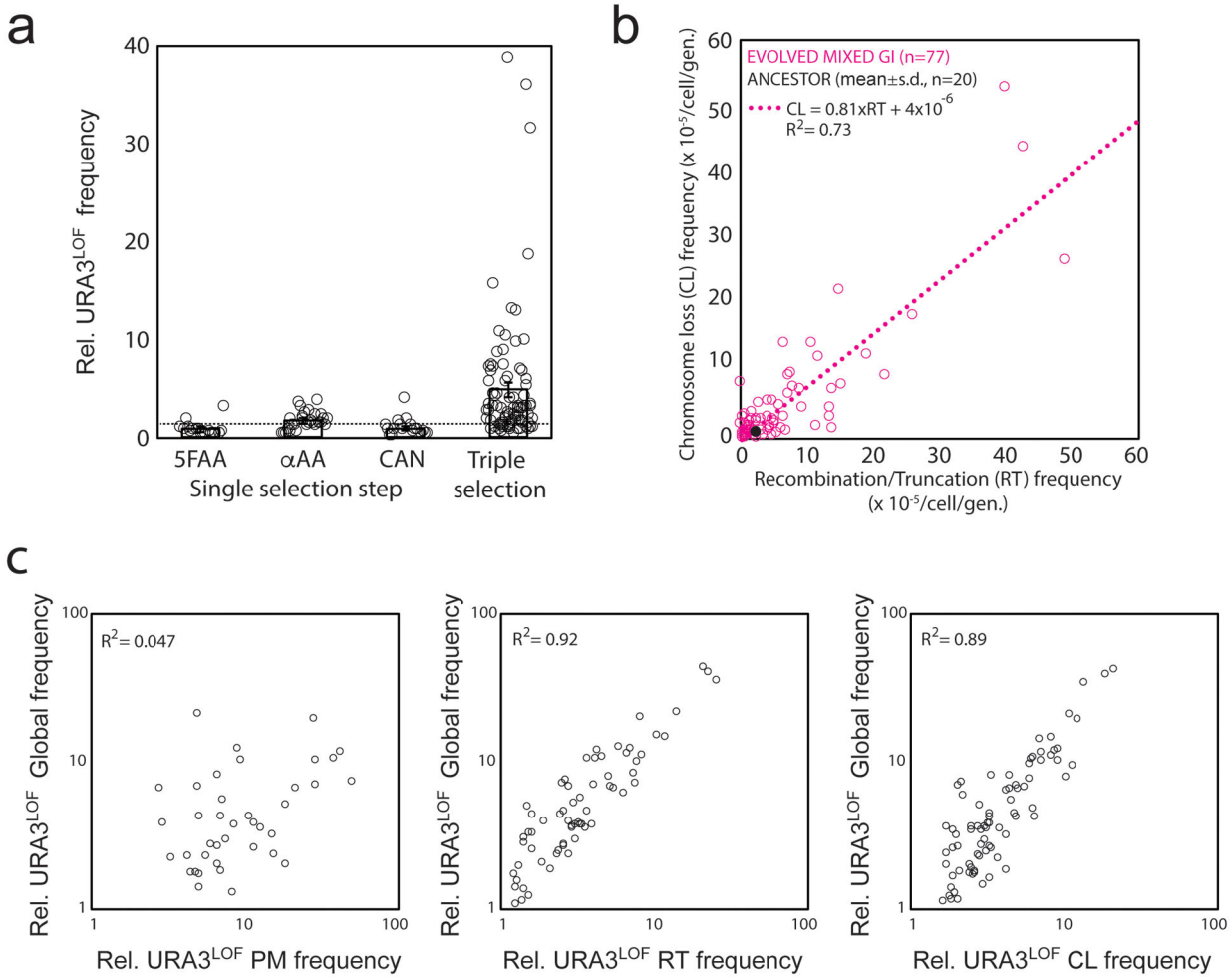


POL3-L523D, proofreading-defective allele of polymerase delta
rad27Δ, deletion of protein required for base excision repair and recombination
mad2Δ, deletion of component of the spindle checkpoint
lys2Δ, deletion of one of the selection markers using KAN-R cassette to test if *LYS2* inactivation or the presence of KAN-R are mutagenic.
*Note: Datapoints = 0 were not plotted, but were used to determine mean and SEM.

Extended Data Figure 1. The genetic instability reporter responds to perturbations that induce point mutations, recombination/truncation events and chromosome loss.

a) Cells were treated with three agents: ethyl methanesulfonate (EMS) a chemical mutagen that primarily induces point mutations, ultra-violet radiation (UV), which induces a mixture of events, and benomyl, a microtubule depolymerizing drug, which primarily induces chromosome loss. The graph shows the frequency with which *URA3* was inactivated or lost from our ancestral diploid strain upon each of the treatments. We show the global rate of loss or inactivation of *URA3* as well as the rates for point mutation, mitotic recombination/ chromosome truncation, and chromosome loss. **b)** Genetic perturbations were applied to a diploid strain carrying the reporter construct: a proof-reading mutation in polymerase delta (*POL3-L523D)* primarily induces point mutations, removing *RAD27*, which encodes a 5' flap endonuclease, primarily elevates mitotic recombination, and removing *MAD2*, which

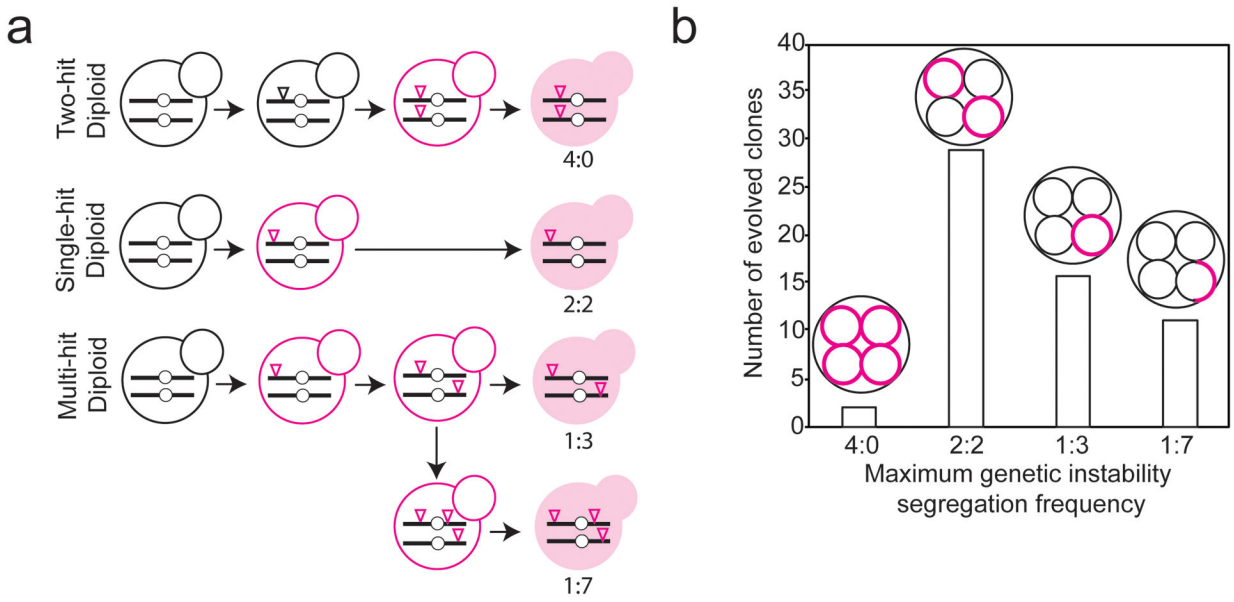
encodes a component of the spindle checkpoint, induces chromosome loss. As a control, inactivation of a single copy of *LYS2*, using the Kanamycin resistance cassette (KanR) did not result in increased instability. The graph shows the frequency with which *URA3* was inactivated or lost from each of the genetically engineered strains. We show the global rate of loss or inactivation of *URA3* as well as the rates for point mutation, mitotic recombination/chromosome truncation, and chromosome loss ("n" represents biologically independent experiments per condition, bars represent mean and error bars represent SEM). Boxplots display center line (median), upper and lower limits (quartiles) and whiskers are 1.5x interquartile range.



Extended Data Figure 2. Triple selection evolved higher levels of genetic instability than single selection and many evolved strains show a correlated increase in recombination and chromosome loss frequencies.

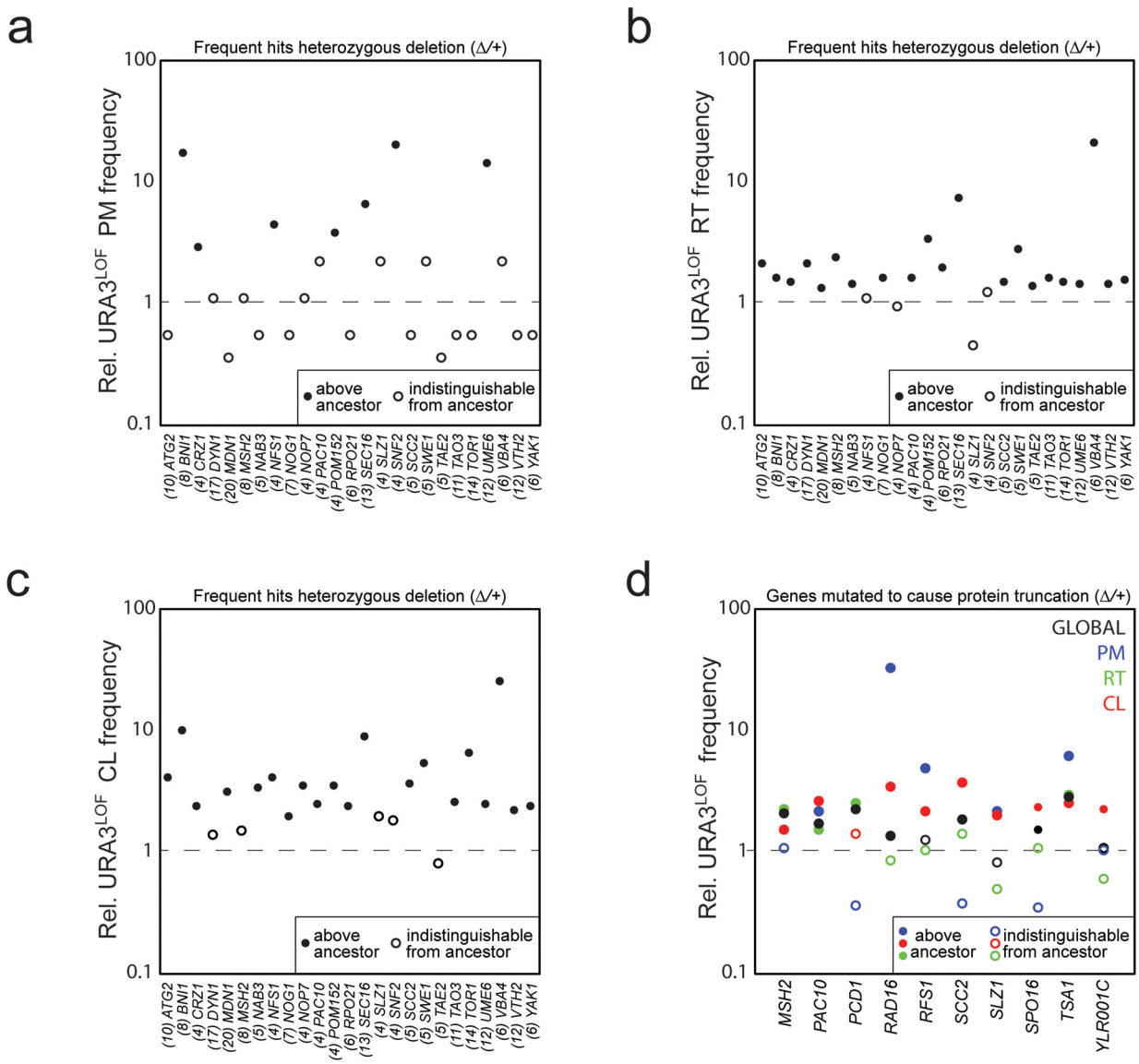
a) *URA3* mutation frequency in evolved clones after a single selection step for resistance to 5-fluoroanthranilic acid (5FAA, n=18 biologically independent samples), α-aminoadipate (αAA, n=29 biologically independent samples), or canavanine (CAN, n=18 biologically independent experiments), and after triple, sequential selection (Triple selection, n=93 biologically independent experiments) as assayed by testing clones for the frequency with which they gave rise to resistance to 5-FOA normalized by the ancestor (mean±s.e.m.). A

single selection step does not promote as much instability as the triple selection (p-values for one sided Student's t-test comparing each single to triple selection pair: p_{5FAA}=0.004, $p_{\alpha\text{-}AA}$=0.009, p_{CAN}=0.002). **b)** Correlation between recombination/truncation (RT) and chromosome loss (CL) mutation frequency for n=77 mutants which did not show a specific elevation in one class of mutational event (linear quadratic regression). **c)** Correlation between rates of the three classes of mutational events and global mutation rates (n_{PM}=51, n_{RT}=54, n_{CL}=80 biologically independent samples, linear quadratic regression). R value corresponds to the Pearson correlation coefficient.



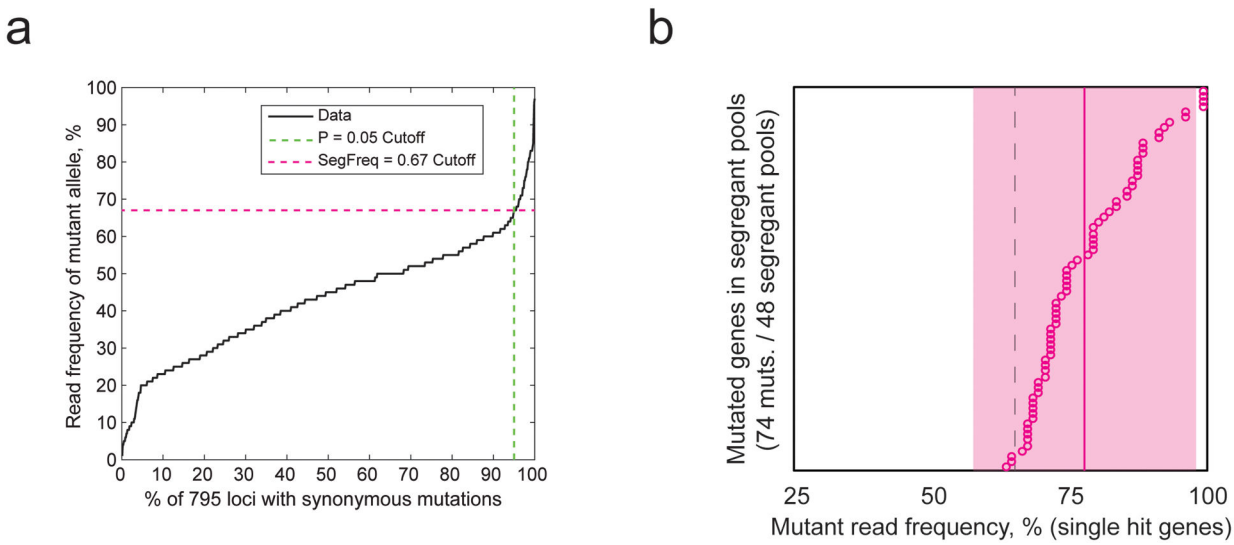
Extended Data Figure 3. Meiotic segregation analysis of maximum genetic instability.

a) Possible scenarios for the evolution of genetic instability. In diploids, a recessive mutation can only cause genetic instability if both alleles of the gene are inactivated (Two-hit Diploid), whereas a dominant mutation need only occur in one copy of the gene (Single-hit Diploid) and some mutators require heterozygous mutations in two or more genes (Multi-hit Diploid). The number and nature of causative mutations will be reflected in the meiotic segregation pattern, assuming the mutations also cause instability in haploids. **b)** Meiotic segregation for the maximum level of genetic instability (measured as the mutation frequency to 5FOA resistance) in haploid spores derived from different evolved diploid clones. Genetically unstable spores are shown in magenta.



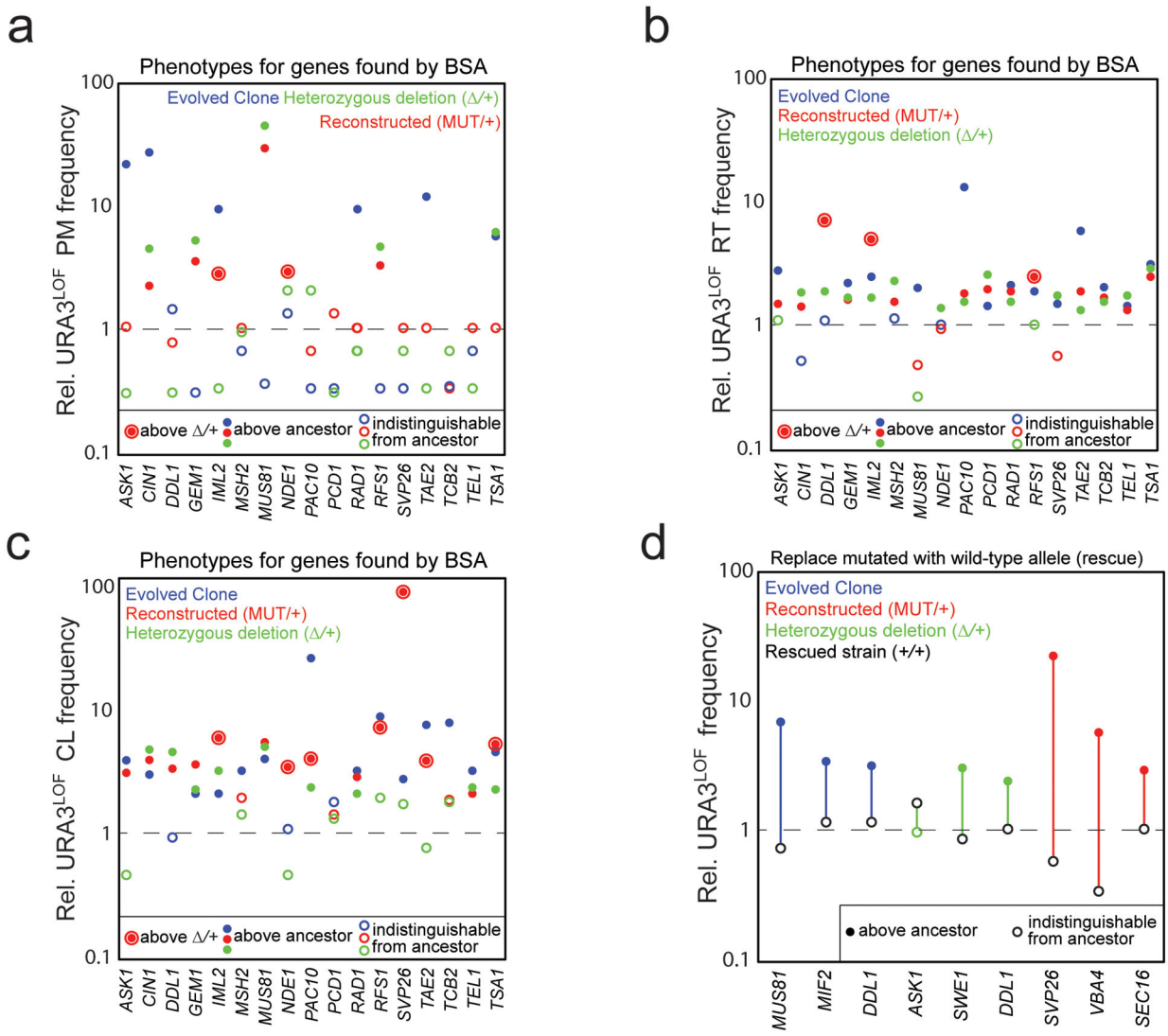
Extended Data Figure 4. Heterozygous deletion of frequently mutated genes affects individual genetic instability types.

Heterozygous deletions (n=25) were engineered in the ancestral strains (*Δ/WT*) carrying our reporter construct and used to measure instability. **a)** The point mutation **b)** recombination/ truncation and **c)** chromosome loss rates in heterozygous deletion mutants. **d)** Mutation rates for global (GLOBAL) and individual (PM, point mutation; RT, recombination or truncation; CL, chromosome loss) instability types in strains carrying heterozygous deletions (*Δ/+*) of genes whose candidate genetic-instability causing mutations were nonsense mutations leading to the production of truncated proteins, which are likely to be non-functional. Filled circles in **a-d** correspond to strains whose instability exceeds that of the ancestor by two standard deviations (n = 3 biologically independent experiments).



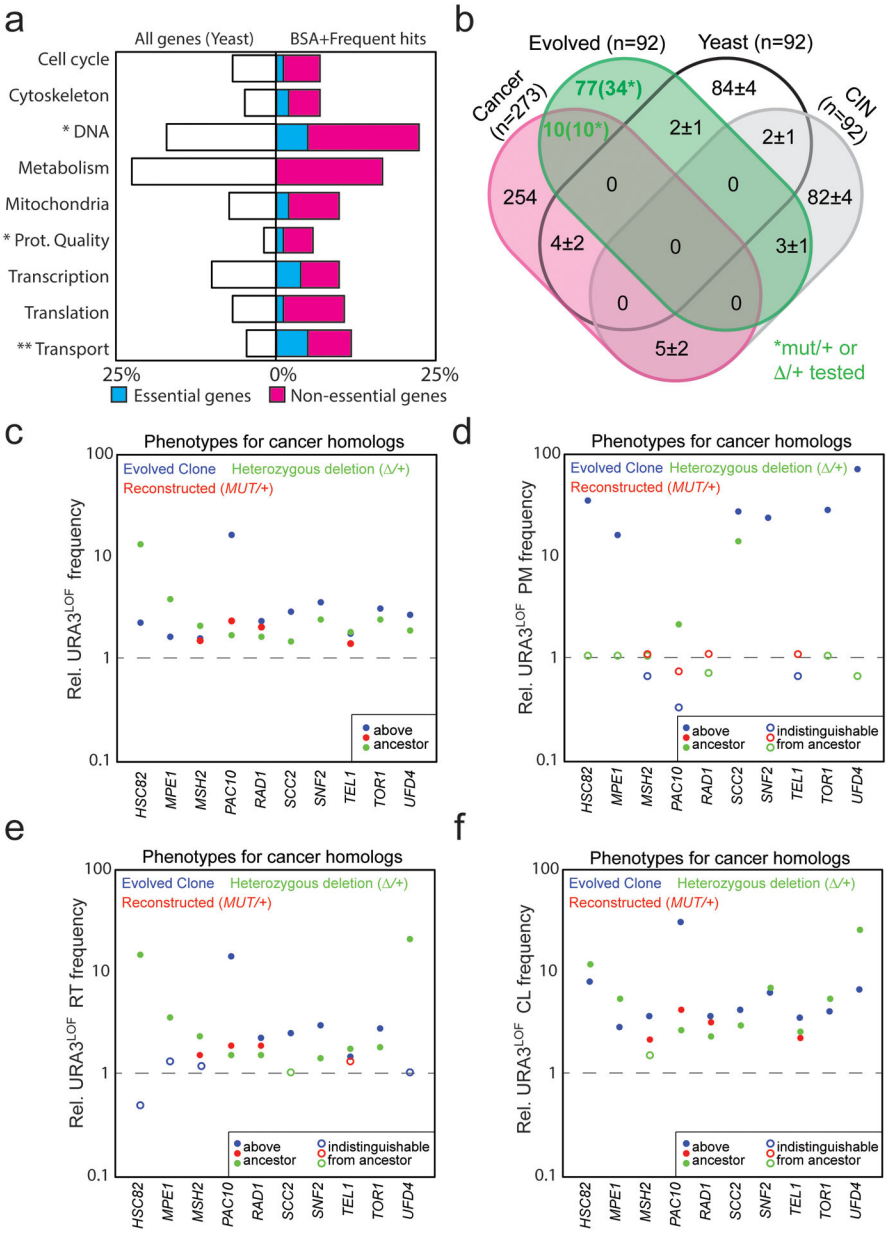
Extended Data Figure 5. Mutant allele frequency in pools of haploid segregants.

a) The graph shows the frequency distribution of synonymous mutations at 795 loci, summed across multiple sequenced pools of segregants (n=48 biologically independent samples). 95% of synonymous mutations have read frequencies of less than 0.67 allowing us to define this as the cutoff for putative causative mutations. Only those loci with 20 or more reads were included in the analysis. **b)** Mutant read frequency for lineage-specific, putative causative mutations for genetic instability (magenta). The dotted line represents the cutoff frequency (0.67) and the magenta line represents the average frequency of the 74 mutant alleles identified (magenta box= ±2s.d.). Three genes (*RAD1, TCB2,* and *CIN1*) that have values slightly below the cutoff were also tested, and in all three cases, reconstruction demonstrated that mutations in these genes increases genetic instability.



Extended Data Figure 6. Effect of engineering causative mutations into the ancestral diploid on point mutation, recombination/truncation and chromosome loss frequencies.

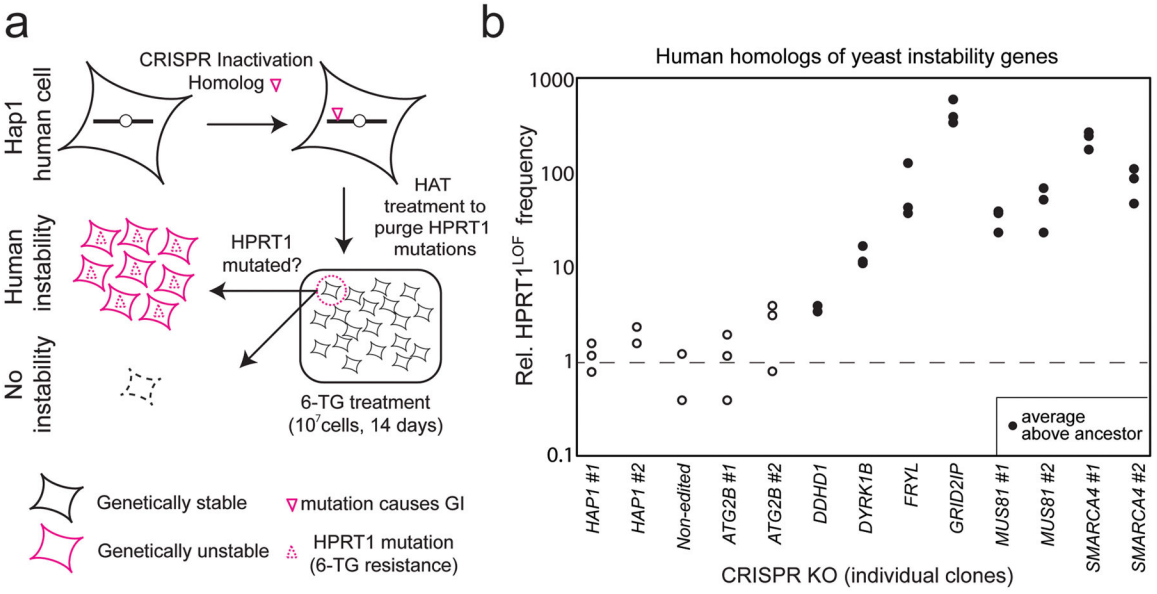
Putative causative mutations from genes identified by bulk segregant analysis were engineered into the ancestral diploid strains as heterozygotes by replacing one of the wild type alleles (mut/+) and our reporter construct was used to measure instability compared to that of the evolved clone and a heterozygous deletion for the gene that harbored the putative causative mutation (*Δ/+*). We also tested a putative causative mutation in PAC10, a gene that was frequently mutated. **a)** point mutation **b)** recombination/truncation and **c)** chromosome loss rates. **d)** Global mutation rates for strains where the causative heterozygous allele (black circle) was replaced with the wildtype allele (green circle, rescue). Filled circles in A-D correspond to strains whose instability exceeds that of the ancestor by two standard deviations (n = 3 biologically independent experiments).



Extended Data Figure 7. Enrichment in protein quality control and transport genes as heterozygous drivers of instability.

a) Frequency of functional classes within genes where we identified mutations that caused genetic instability. Probabilities were calculated from the binomial distribution,** $p < 0.05$ with Bonferroni correction, 10 hypotheses tested, * significant at 2.5% false discovery rate using the Benjamini-Hochberg procedure (left frequency in yeast genome; right, genes whose mutation we identified as leading to genetic instability, n=92 independent genes). Blue represents the fraction of essential genes, magenta represents the fraction of non-essential genes. **b)** Overlap between genes in which we selected mutations that produced genetic instability (Evolved, green), yeast genes homologous to human cancer genes (Cancer, magenta), random subsets of yeast genes (Yeast, black), and randomly selected subsets of genes previously identified in yeast screens for genetic instability (CIN, grey).

Quantification of **c) global, d)** point mutation, **e)** recombination/truncation and **f)** chromosome loss rates in heterozygous point mutants or heterozygous deletion mutants for yeast genetic instability genes with cancer driver homologs. Data for the mut/+ reconstructed strain showed for 4 strains (*MSH2, PAC10, RAD1* and *TEL1*). The data for five of these genes, *MSH2, PAC10, SCC2, SNF2,* and *TOR1* is also shown in Extended Data Fig. 4. Filled circles in **c-f** correspond to strains whose instability exceeds that of the ancestor by two standard deviations (n = 3 biologically independent experiments).



c

Gene	gRNA*	fs mutation	Sequence (Sanger)
ATG2B #1	GCTCGTACACCTTCTAGCCG	22bp del. exon 8	ACAGCTCGTACACCTTCTAGCCGTGGTAA
ATG2B #2	GCTCGTACACCTTCTAGCCG	10bp del. exon 8	ACCTTCTAGCCGTGGTAAGCAACTCAAAC
DDHD1	GCTTCTGCCAGAGTACGACG	1bp ins. exon 1	TCTGCCAGAGTACGGACGGGGCACGAGC
DYRK1B	GCTGATGAACCAGCATGACA	1bp ins. exon 5	GAACCAGCATGGACACGGAGATGAA
FRYL	GTCAAACAAGGTGCGAAGTA	1bp ins. exon 6	CTCCCTTCCTTACTTTCGCACCTTGTTTG
GRID2IP	CGGGCCTGTCTGGATCGAGT	80bp ins. exon4	CGGCCACGGGT-(76bp)-TCGCCTGTCTGG
MSH2	GCGCCGTATAGAAGTCGCCC	2bp del. exon1	TTTCGACCGGGGCGACTTCTATACGGCG
MUS81 #1,2	TACTGGCCAGCTCGGCACTC	7bp del. exon 4	TACTGGCCAGCTCGGCACTCAGGAGCC
SMARCA4 #1,2	GCAGCAGACAGACGAGTACG	1bp ins. exon10	GCAGACAGACGAGTTACGTGGCTAACCT

fs - frameshift; del. - deletion; ins. - insertion; red - deleted sequence, green - inserted sequence
*cloned into p458x-GFP

Extended Data Figure 8. Mutation rates of HPRT1 in HAP1 human cells carrying inactivated homologs of genes mutated during yeast evolution to trigger genetic instability.

a) Targets selected from yeast evolution experiments were tested in human HAP1 cells, by using CRISPR-mediated gene inactivation of the corresponding homologs. HAP1 knockout cell lines were incubated with 6-thioguanine (6-TG) (after treatment with hypoxanthine-aminopterin-thymidine (HAT) medium to eliminate jackpots) and colony formation was quantified and compared to the wildtype. **b)** Inactivation frequency of HPRT1, normalized

by the wildtype for different homolog deletions. HAP1 #1 and HAP1 #2 are the parental wildtype, Non-edited corresponds to a cell line that went through the transformation protocol for ATG2B editing, but did not acquire mutations in ATG2B, ATG2B#1 and ATG2B #2, did not show increased instability. SMARCA4 was known to increase HPRT1 inactivation frequency and was used as a positive control for our assay. Values are normalized by average HPRT1 inactivation rate of HAP1 wildtype, filled circles correspond to strains whose instability exceeds that of the ancestor by two standard deviations (n = 3 biologically independent experiments). **c)** Guide RNA target sequences and sequences of the mutated regions of genes after CRISPR-mediated gene inactivation for each of the cell lines tested.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Filipa Beça, Dario Cabrera, Beverly Neugeboren, Chiara Pogliano for developing reagents and conducting preliminary experiments, and the BauerCore facility at Harvard University for technical help. Gregg Wildenberg, John Koschwanez, Nichole Wespe and Marco Fumasoni provided technical help and discussed experiments and results. We thank Joao Matos for the kind gift of the MUS81 deletion CRISPR cell lines used in Fig. 3c and Extended Data Fig. 8. Vladimir Denic, Boris Shraiman, Michael Desai, Marco Fumasoni and Laura Bagamery provided critical comments on the manuscript. M.C. received a long-term fellowship from Human Frontiers Science Program (HFSP, LT 000694/2014-L). The work was supported by NIH/NIGMS award (R01/GM043987) granted to A.W.M.

References

1. Swings T et al. Adaptive tuning of mutation rates allows fast response to lethal stress in Escherichia coli. Elife 6, doi:10.7554/eLife.22939 (2017).
2. Lawrence MS et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495–501, doi:10.1038/nature12912 (2014). [PubMed: 24390350]
3. Drost J et al. Sequential cancer mutations in cultured human intestinal stem cells. Nature 521, 43–47, doi:10.1038/nature14415 (2015). [PubMed: 25924068]
4. Hanahan D & Weinberg RA Hallmarks of cancer: the next generation. Cell 144, 646–674, doi:10.1016/j.cell.2011.02.013 (2011). [PubMed: 21376230]
5. Inoue K & Fry EA Haploinsufficient tumor suppressor genes. Adv Med Biol 118, 83–122 (2017). [PubMed: 28680740]
6. Mao EF, Lane L, Lee J & Miller JH Proliferation of mutators in a cell population. J Bacteriol 179, 417–422 (1997). [PubMed: 8990293]
7. Huang ME, Rio AG, Nicolas A & Kolodner RD A genomewide screen in Saccharomyces cerevisiae for genes that suppress the accumulation of mutations. Proc Natl Acad Sci U S A 100, 11529–11534, doi:10.1073/pnas.2035018100 (2003). [PubMed: 12972632]
8. Yuen KW et al. Systematic genome instability screens in yeast and their potential relevance to cancer. Proc Natl Acad Sci U S A 104, 3925–3930, doi:10.1073/pnas.0610642104 (2007). [PubMed: 17360454]
9. Stirling PC et al. The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. PLoS Genet 7, e1002057, doi:10.1371/journal.pgen.1002057 (2011). [PubMed: 21552543]
10. Andersen MP, Nelson ZW, Hetrick ED & Gottschling DE A genetic screen for increased loss of heterozygosity in Saccharomyces cerevisiae. Genetics 179, 1179–1195, doi:10.1534/genetics.108.089250 (2008). [PubMed: 18562670]

11. Strome ED, Wu X, Kimmel M & Plon SE Heterozygous screen in Saccharomyces cerevisiae identifies dosage-sensitive genes that affect chromosome stability. Genetics 178, 1193–1207, doi:10.1534/genetics.107.084103 (2008). [PubMed: 18245329]
12. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68, 820–823 (1971). [PubMed: 5279523]
13. Cho A et al. MUFFINN: cancer gene discovery via network analysis of somatic mutation data. Genome Biol 17, 129, doi:10.1186/s13059-016-0989-x (2016). [PubMed: 27333808]
14. Davoli T et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell 155, 948–962, doi:10.1016/j.cell.2013.10.011 (2013). [PubMed: 24183448]
15. Futreal PA et al. A census of human cancer genes. Nature reviews. Cancer 4, 177–183, doi:10.1038/nrc1299 (2004). [PubMed: 14993899]
16. Tamborero D et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Sci Rep 3, 2650, doi:10.1038/srep02650 (2013). [PubMed: 24084849]
17. Zehir A et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 23, 703–713, doi:10.1038/nm.4333 (2017). [PubMed: 28481359]
18. Gao J et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 6, pl1, doi:10.1126/scisignal.2004088 (2013). [PubMed: 23550210]
19. Yin Y & Petes TD Genome-wide high-resolution mapping of UV-induced mitotic recombination events in Saccharomyces cerevisiae. PLoS Genet 9, e1003894, doi:10.1371/journal.pgen.1003894 (2013). [PubMed: 24204306]
20. Brauer MJ, Christianson CM, Pai DA & Dunham MJ Mapping novel traits by array-assisted bulk segregant analysis in Saccharomyces cerevisiae. Genetics 173, 1813–1816, doi:10.1534/genetics.106.057927 (2006). [PubMed: 16624899]
21. Segre AV, Murray AW & Leu JY High-resolution mutation mapping reveals parallel experimental evolution in yeast. PLoS Biol 4, e256, doi:10.1371/journal.pbio.0040256 (2006). [PubMed: 16856782]
22. Choy JS et al. Genome-wide haploinsufficiency screen reveals a novel role for gamma-TuSC in spindle organization and genome stability. Molecular Biology of the Cell 24, 2753–2763, doi:10.1091/mbc.E12-12-0902 (2013). [PubMed: 23825022]
23. Carette JE et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature 477, 340–343, doi:10.1038/nature10348 (2011). [PubMed: 21866103]
24. Huang HT, Chen SM, Pan LB, Yao J & Ma HT Loss of function of SWI/SNF chromatin remodeling genes leads to genome instability of human lung cancer. Oncol Rep 33, 283–291, doi:10.3892/or.2014.3584 (2015). [PubMed: 25370573]
25. Sniegowski PD, Gerrish PJ & Lenski RE Evolution of high mutation rates in experimental populations of E. coli. Nature 387, 703–705, doi:10.1038/42701 (1997). [PubMed: 9192894]
26. Doench JG et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 34, 184–191, doi:10.1038/nbt.3437 (2016). [PubMed: 26780180]
27. Ran FA et al. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308, doi:10.1038/nprot.2013.143 (2013). [PubMed: 24157548]
28. Duda H et al. A Mechanism for Controlled Breakage of Under-replicated Chromosomes during Mitosis. Dev Cell 39, 740–755, doi:10.1016/j.devcel.2016.11.017 (2016). [PubMed: 27997828]
29. Andersen MP, Nelson ZW, Hetrick ED & Gottschling DE A genetic screen for increased loss of heterozygosity in Saccharomyces cerevisiae. Genetics 179, 1179–1195, doi:10.1534/genetics.108.089250 (2008). [PubMed: 18562670]
30. Jin YH et al. The multiple biological roles of the 3'-->5' exonuclease of Saccharomyces cerevisiae DNA polymerase delta require switching between the polymerase and exonuclease domains. Mol Cell Biol 25, 461–471, doi:10.1128/MCB.25.1.461-471.2005 (2005). [PubMed: 15601866]
31. Chen RH, Shevchenko A, Mann M & Murray AW Spindle checkpoint protein Xmad1 recruits Xmad2 to unattached kinetochores. J Cell Biol 143, 283–295 (1998). [PubMed: 9786942]
32. Rosenthal GA The biological effects and mode of action of L-canavanine, a structural analogue of L-arginine. Q Rev Biol 52, 155–178 (1977). [PubMed: 331385]

33. Toyn JH, Gunyuzlu PL, White WH, Thompson LA & Hollis GF A counterselection for the tryptophan pathway in yeast: 5-fluoroanthranilic acid resistance. Yeast 16, 553–560, doi:10.1002/(SICI)1097-0061(200004)16:6<553::AID-YEA554>3.0.CO;2-7 (2000). [PubMed: 10790693]
34. Chattoo BB et al. Selection of lys2 Mutants of the Yeast SACCHAROMYCES CEREVISIAE by the Utilization of alpha-AMINOADIPATE. Genetics 93, 51–65 (1979). [PubMed: 17248969]
35. Mao EF, Lane L, Lee J & Miller JH Proliferation of mutators in a cell population. J Bacteriol 179, 417–422 (1997). [PubMed: 8990293]
36. Johnson GE Mammalian cell HPRT gene mutation assay: test methods. Methods Mol Biol 817, 55–67, doi:10.1007/978-1-61779-421-6_4 (2012). [PubMed: 22147568]
37. Baym M et al. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One 10, e0128036, doi:10.1371/journal.pone.0128036 (2015). [PubMed: 26000737]
38. Li H & Durbin R Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760, doi:10.1093/bioinformatics/btp324 (2009). [PubMed: 19451168]
39. Koboldt DC et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22, 568–576, doi:10.1101/gr.129684.111 (2012). [PubMed: 22300766]
40. Koschwanez JH, Foster KR & Murray AW Improved use of a public good selects for the evolution of undifferentiated multicellularity. Elife 2, e00367, doi:10.7554/eLife.00367 (2013). [PubMed: 23577233]
41. Davoli T et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell 155, 948–962, doi:10.1016/j.cell.2013.10.011 (2013). [PubMed: 24183448]
42. Futreal PA et al. A census of human cancer genes. *Nature reviews*. Cancer 4, 177–183, doi:10.1038/nrc1299 (2004). [PubMed: 14993899]
43. Lawrence MS et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495–501, doi:10.1038/nature12912 (2014). [PubMed: 24390350]
44. Tamborero D et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Sci Rep 3, 2650, doi:10.1038/srep02650 (2013). [PubMed: 24084849]
45. Zehir A et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 23, 703–713, doi:10.1038/nm.4333 (2017). [PubMed: 28481359]
46. Cho A et al. MUFFINN: cancer gene discovery via network analysis of somatic mutation data. Genome Biol 17, 129, doi:10.1186/s13059-016-0989-x (2016). [PubMed: 27333808]
47. Stirling PC et al. The complete spectrum of yeast chromosome instability genes identifies candidate CIN cancer genes and functional roles for ASTRA complex components. PLoS Genet 7, e1002057, doi:10.1371/journal.pgen.1002057 (2011). [PubMed: 21552543]
48. Yuen KW et al. Systematic genome instability screens in yeast and their potential relevance to cancer. Proc Natl Acad Sci U S A 104, 3925–3930, doi:10.1073/pnas.0610642104 (2007). [PubMed: 17360454]

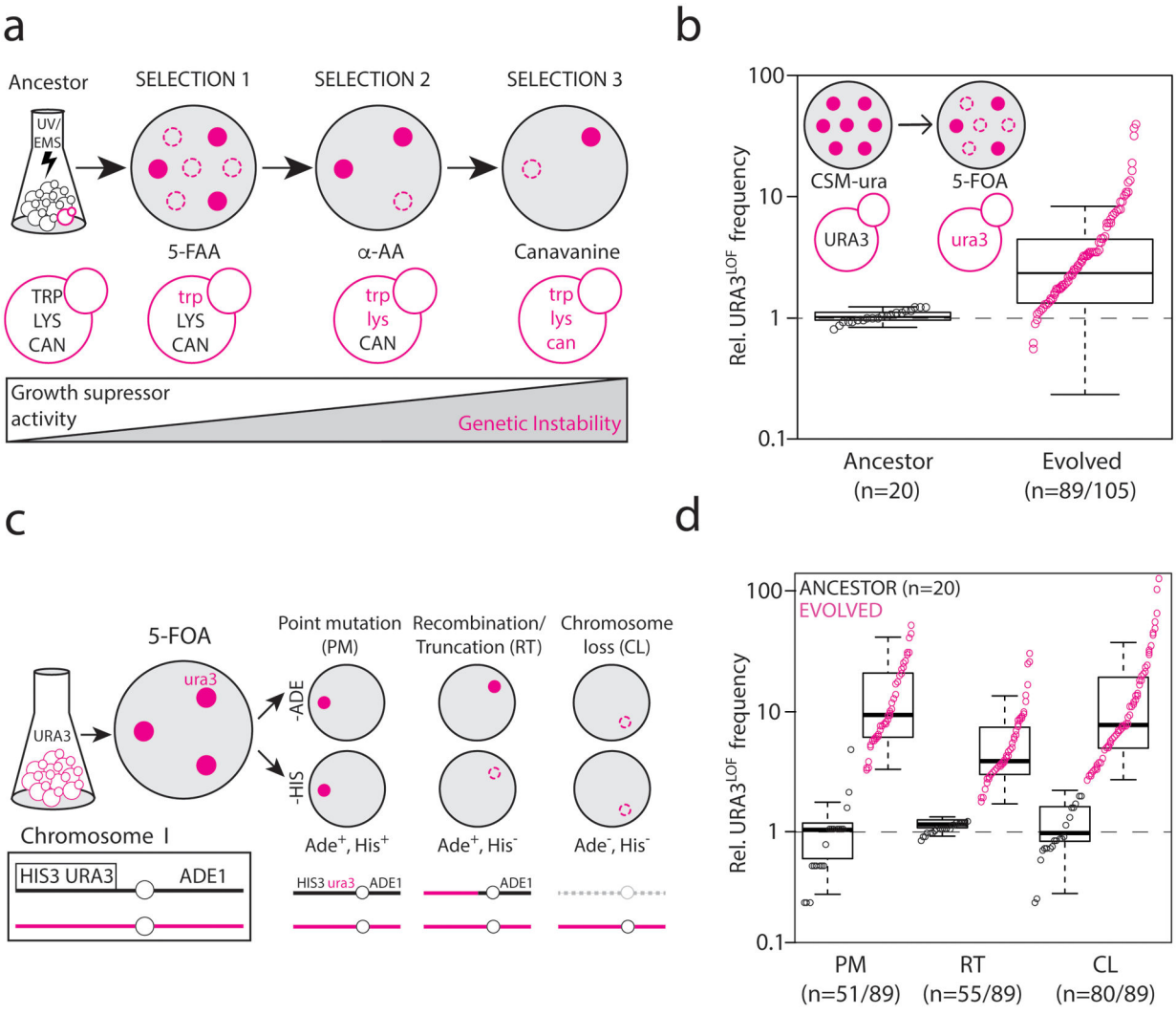


Figure 1. Sequential inactivation of growth suppressors selects for the evolution of genetic instability.

a) Selecting genetically unstable clones: clones derived from mutagenized diploid cells survived selection by sequentially disrupting 3 sets of growth suppressor genes: one of five genes required for tryptophan biosynthesis (*TRP*, SELECTION 1, growth on 5-Fluoroanthranilate), one of two genes required for lysine biosynthesis (*LYS*, SELECTION 2, growth on alpha-aminoadipate) and the *CAN1* gene (SELECTION 3, growth on canavanine). **b)** Mutation frequency of a single copy of the URA3 was elevated in evolved diploid clones when compared to ancestor (one-sided Student's t-test comparing ancestor and evolved datasets collectively, $p=3.9\times10^{-8}$), growth on 5-fluoroorotic acid. Points shown for 89 clones that showed an elevation for any one following mutation rates, global, point mutation (PM), mitotic recombination/chromosome truncation (RT), or chromosome loss (CL). **c)** Genetic instability reporter in diploid strains consisting of three single-copy genes on chromosome I: *URA3* coupled to the *HIS3* gene 70 kb from the left telomere, and the *ADE1* gene on the opposite side of the centromere (*URA3*, *HIS3* and *ADE1* are inactivated elsewhere in the genome). This architecture allows distinguishing three different *URA3*

inactivation modes in 5FOA-resistant diploids: point mutation (His^+ and Ade^+), mitotic recombination or loss of the left arm of chromosome I (His^- and Ade^+), and chromosome loss (His^-, and Ade^-). **d)** Frequencies of 5FOA-resistant point mutation (PM), recombination/truncation (RT) and complete chromosome loss (CL) events were significantly elevated in evolved clones (one-sided Student's t-test; $p_{PM}=3.2\times10^{-9}$, $p_{RT}=1.0\times10^{-6}$, $p_{CL}=4.1\times10^{-8}$). Values above the ancestor are shown. Mean elevations for the different mutation rates (mean±s.d.): Global=5.2±6.7, and PM=8.0±10.1, RT=3.4±4.4, CL=13.5±20.2. *URA3* inactivation rate (/cell/generation) in ancestor: Global=2.4×10^{-5}, and PM=2.0×10^{-7}, RT=2.0×10^{-5}, CL=4.1×10^{-6}. Boxplots display center line (median), upper and lower limits (quartiles) and whiskers show 1.5x interquartile range).

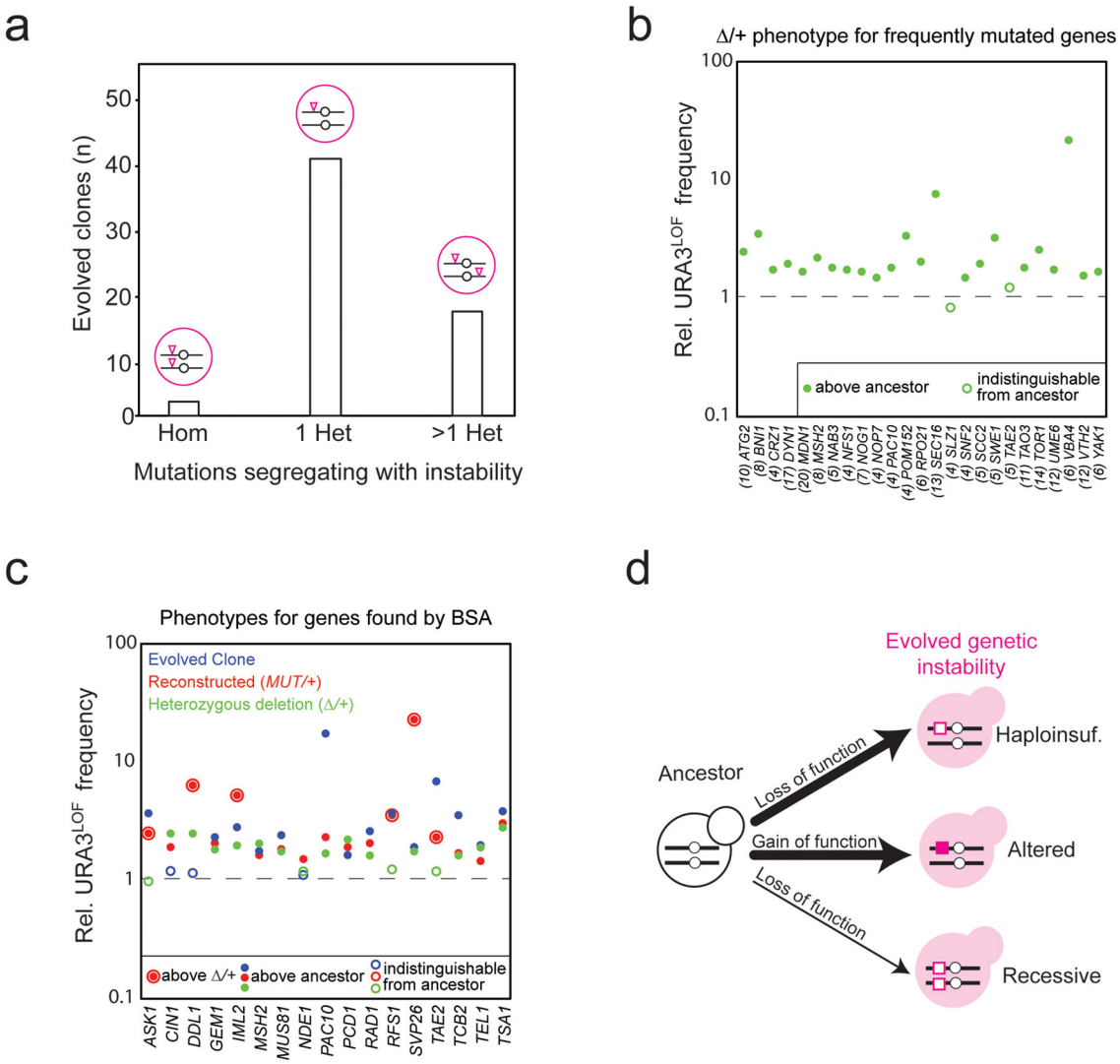
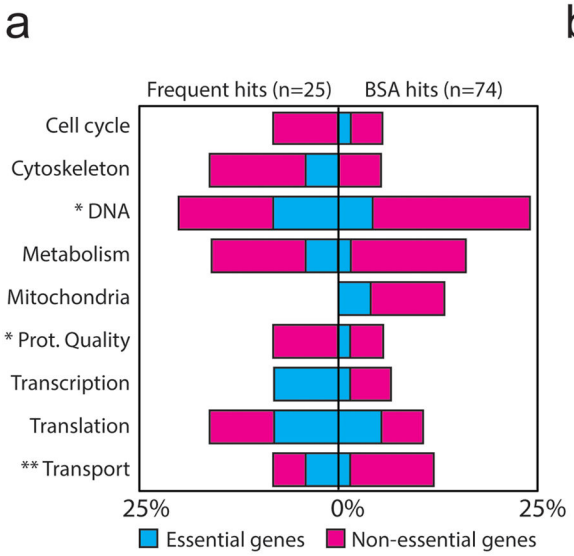


Figure 2. A single heterozygous mutation causes genetic instability in diploid cells.

a) Diagnosing the genetic basis of instability by following its meiotic segregation. Evolved diploid clones were classified as showing the segregation pattern corresponding to a homozygous (Hom) mutation, a single heterozygous mutation (1 Het) or multiple heterozygous mutations (>1 Het). **b)** Global mutation frequency for heterozygous deletions of each of 25 frequently mutated genes in our evolved diploid clones (mutated in ≥ 4 independent clones). Numbers show the number of lineages mutated for a given gene. Although heterozygous deletions of *SLZ1* and *TAE2* have a global mutation frequency similar to the wildtype, they have elevated point mutation and recombination/truncation rates, respectively (see Extended Data Fig. 4 A–C). **c)** Analysis of mutated genes identified from bulk segregant analysis. Global mutation frequencies for evolved clones (blue), reconstructed clones where the mutation from the evolved clone was introduced into the ancestor as a heterozygote (*MUT/+,* red), or the heterozygous deletion of the corresponding gene from the ancestor (*Δ/+*, green). Dotted lines in **b** and **c** corresponds to ancestor, filled circles in **b** and **c** correspond to strains whose instability exceeds that of the ancestor by two

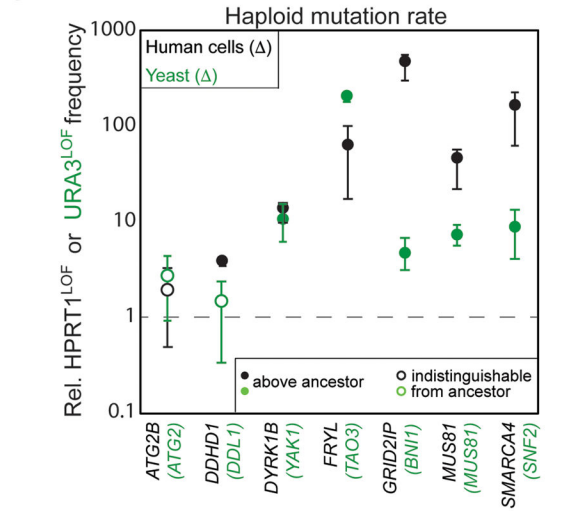
standard deviations in three biologically independent experiments; outer circles in **c** correspond to cases where the heterozygous mutation had higher instability than the heterozygous deletion. We also tested a putative causative mutation in PAC10, a gene that was frequently mutated. **d)** Scheme representing frequent events (thick arrows) that cause genetic instability: a single heterozygous mutation can cause loss of function (Haploinsuf.) or a dominantly altered function (Altered). Homozygous mutations are rare (thin arrow) and cause instability through loss of function mutations in genes that are not haploinsufficient (Recessive).



b

Evolved	Cancer	Evidence	Function
HSC82*	HSP90AB1	CD,M	Protein chaperone
MPE1*	RBBP6	CD,F	DNA replication
MSH2*	MSH2	CD,F,M	DNA mismatch repair
PAC10*	VBP1	CD,F,M	Protein chaperone
RAD1*	ERCC4	CD,F,M	DNA endonuclease
SCC2*	NIPBL	CD	DNA cohesion
SNF2*	SMARCA4	CD,F,M	Chromatin remodeling
TEL1*	ATM	CD,F,M	DNA repair
TOR1*	MTOR	CD	Cell growth / translation
UFD4*	TRIP12	CD,M	DNA repair

*mut/+ or Δ/+ are genetically unstable
CD - Cancer driver, F - Hereditary, M - Mouse model



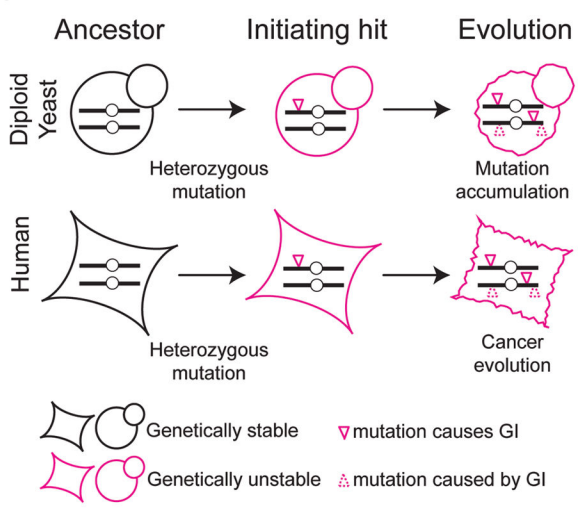


Figure 3. Genes selected to cause instability in yeast target different functional classes and homolog inactivation triggers instability in human cells.

a) Frequency of functional classes within genes where we identified mutations that caused genetic instability (left, frequent hits; right, hits enriched in bulk segregant analysis compared to the total number of genes in a given functional category), 7 genes were present in both groups (ATG2, MSH2, NFS1, NOP7, POM152, SEC16 and TAE2)); probabilities were calculated from the binomial distribution,** $p < 0.05$ with Bonferroni correction (10 hypotheses tested), * significant at 2.5% false discovery rate using the Benjamini-Hochberg procedure. Blue and magenta represent the fraction of essential and non-essential genes, respectively. **b)** Genes where we identified mutations that cause genetic instability and their human homologs that have been implicated in cancer in one of three ways: hereditary cancer predisposition, animal models, or the frequency with which they are mutated in cancer patients (cancer driver). **c)** HPRT1 mutation frequency in CRISPR-mediated knock-outs of homologs of yeast instability genes in the HAP1 cell lines (black dots), compared to *URA3* mutation frequency in yeast cells deleted for the corresponding homolog (green dots).

Closed circles, mutation rate significantly greater than wild type (center value corresponds to mean, error bars correspond to s.d.; filled circles correspond to strains whose instability exceeds that of the ancestor by two standard deviations; n = 3 biologically independent experiments). **d)** Scheme comparing genetic instability evolution in yeast and cancer evolution. Single-heterozygous mutations can initiate genetic instability and accelerate the accumulation of further mutations that alter cellular properties or increase genetic instability.