SI APPENDIX

SI METHODS

Sequencing data analysis

Raw sequencing data were split by index using a dual-index barcode splitter (barcode_splitter.py) from L. Parsons (Princeton University). To analyze the bulk-segregant fitness data we first generated a list of potential mutations that segregated within the bulk-segregant pool using our previous whole-genome, whole-population time-course sequencing data (1) to identify mutations within the evolved population at the time of clone isolation (Dataset S3). For each mutation, we generated four search terms, corresponding to the ancestral and evolved alleles in both forward and reverse orientation. Each search term consisted of the known mutation as well as ten nucleotides immediately upstream and downstream of the mutation. In most instances, the search term was specific to a single locus within the genome. If the initial search term lacked absolute specificity in the reference genome, the search term was extended by five to ten nucleotides in each direction. For each FASTQ file, the number of reads containing the search terms was recorded, providing output in the form of ancestral and evolved allele counts at each locus. The search criteria required read accuracy/precision since only reads that possessed no mismatches/errors within the immediate vicinity of the mutation were counted. For comparison, the sequence reads were mapped to a corrected W303 genome (1) using BWA version 0.7.12 (2) with default parameters except "Disallow an indel within INT bp towards the ends" set to 0 and "Gap open penalty" set to 5. Mutations were then called using FreeBayes version v0.9.21-24-g840b412 (3) with default parameters for pooled samples. The two approaches were often harmonious (SI Appendix, Fig. S10), and any significant discrepancies resulted due to FreeBayes calling the same mutation under several different call variants, particularly at the end of reads, an issue stemming from forced alignment to the designated ancestral reference. To detect copy number variants (CNV's), including aneuploidies, we generated genome-wide coverage plots from the eleven populations examined in this study. This detection method has previously been effective in identifying CNV's in our laboratory-evolved haploids and diploids. No CNV's were detected amongst the eleven evolved haploid populations investigated here.

The background-averaged fitness effect of each mutation was calculated as the linear regression of the log ratio of allele frequency ($N_{evolved}/N_{ancestral}$) over the 100-generation experiment. Data corresponding to any time point at which either the ancestral or evolved allele was undetected were removed from analysis. Standard error of the regression and 95%

confidence intervals were determined using MATLAB™ (MathWorks®). Mutations were classified as neutral if the confidence interval encompassed zero. Mutations with confidence intervals entirely above or below zero were characterized as beneficial or deleterious, respectively. Few driver mutations exhibit genetic linkage; a notable exception is *cne1* and *gpb2* (separated by 2.6 kb), which is treated as a single mutation in downstream analyses. To confirm that the log ratio of allele frequencies follows a linear trend over time, we employed the Durbin-Watson test to detect the presence of autocorrelation in residuals from the linear model (SI Appendix, Fig. S11A; Dataset S3). Of the 118 total mutations, 18 exhibited a significant correlation in their residuals ($p < 0.01$), indicating that a linear model is an appropriate estimate of fitness for most mutations. Four driver mutations appear to exhibit a non-linear trend per the Durbin-Watson test (SI Appendix, Fig. S11C), most notably *kel1* which we know interacts synergistically with *hsl7.* Since sampling error is intrinsic to the sequencing-based bulk-segregant fitness assay, the Durbin-Watson test is susceptible to type I error as autocorrelated residuals could simply result from the random deviations in allele frequencies between time points (SI Appendix, Fig. S11D). We tested for type I error by comparing the variance in the hitchhiker mutations that tested positive and those that did not (SI Appendix, Fig. S11B). The similarity in variance between the two groups suggests that these alleles are truly hitchhiker mutations.

Identification of additional allelic variants segregating in bulk-segregant pools

The genotyped bulk-segregant individuals from the BYS2E01 cross were analyzed for the presence of previously unrecognized genetic variants – alleles that differed between the *MAT***a** ancestor and *MAT*α parent, low frequency mutations present in the evolved clone, or mutations that arose during the construction of the bulk-segregant pool. VCF files corresponding to all BYS2E01 individuals were merged and scanned for calls that fit the following criteria: 1) the mutation is only called in a fraction of individuals, 2) the mutation, when called, is present near 100% in most clones, 3) the mutation is called at less than 100% in the individuals acknowledged to be non-clonal mixed samples, and 4) the mutation does not have another allelic variant. A total of seven genetic variants were identified in the BYS2E01 pool. All pools were then screened for these seven variants using the search term approach (SI Appendix, Fig. S12).

IQG1 **allele replacement**

All individuals in the BYS2E01-745 segregant pool contain the evolved allele of *iqg1*. To determine if the *iqq1* allele imparts a fitness effect, we replaced the *iqg1* mutation with the wild-

type *IQG1* allele in 82 of the 192 segregant individuals. We synthesized a gBlock® gene fragment (Integrated DNA Technologies) containing a guide RNA (AGAAAATATTATGAAGTTTT) targeting the *iqg1* mutation and the adjacent PAM sequence (CGG). The fragment was first cloned into the *URA3*-marked plasmid p426-SNR52p-gRNA-SUP4t (Addgene #43803). We then inserted the *HIS3* gene into the *BtgI* restriction sites in *URA3* creating a partial *ura3* deletion. The resulting *HIS3* gRNA plasmid, a *TRP1*-marked plasmid containing a constitutively expressed Cas9 (Addgene #43802), and a *URA3*-marked plasmid containing the wild-type *IQG1* allele (4) were co-transformed into 82 *iqg1* segregants. Transformants were selected on media utilizing the aforementioned auxotrophic markers, and all strains were plasmid-cured following verification of the allele swap. To screen for convertants, we utilized the *SspI* restriction site that was introduced through the evolved *iqg1* mutation (AGTATT → AATATT). Following PCR and *SspI* digest, transformants that yielded an intact 505 bp PCR product were presumed to be *IQG1* convertants. We performed Sanger sequencing on a subset of converted clones to confirm the allele swap and ensure that no other mutations were introduced into the 505 bp region around the Cas9 cut site. Overall efficiency of allele replacement ranged from 15% to 85%.
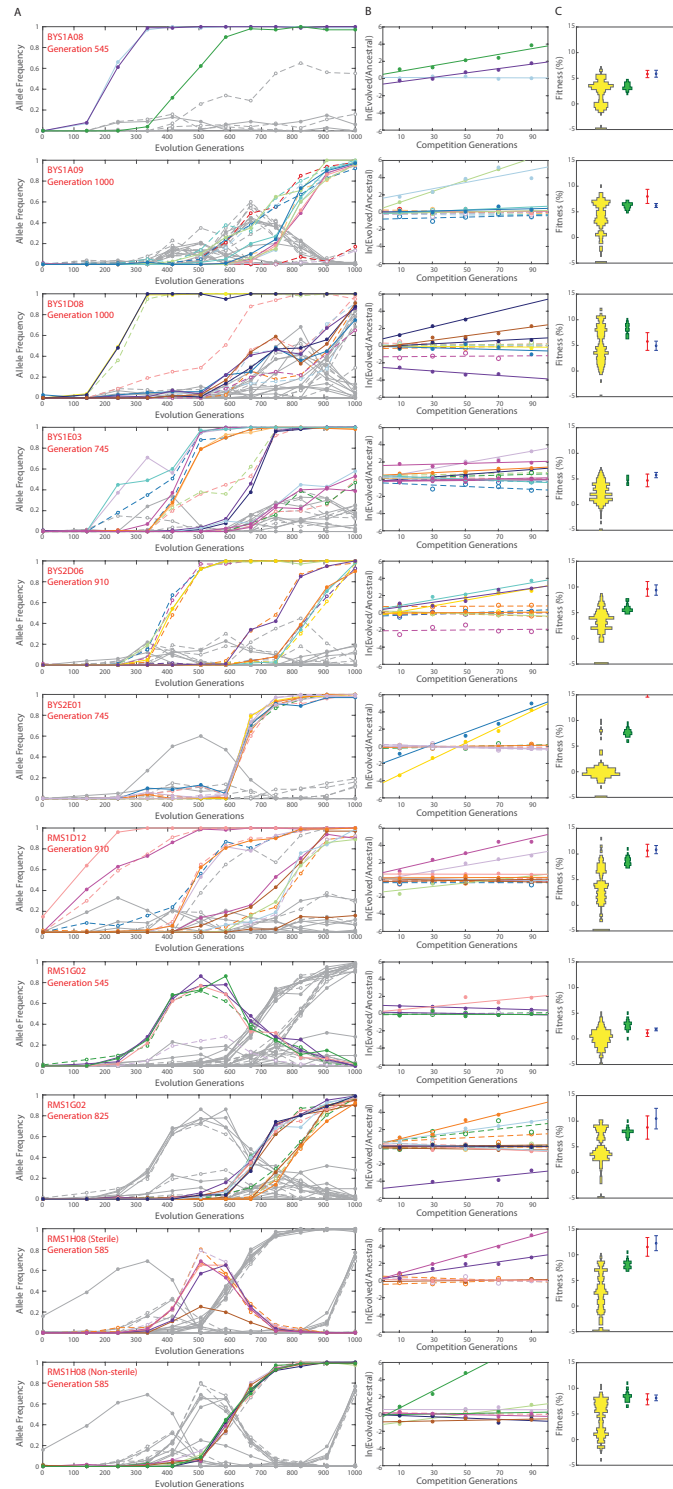
Fig. S1 Genetic dissection of mutations from all eleven evolved clones. *(A)* Genome evolution of each population was previously tracked through time-course, whole-genome sequencing (1). An evolved clone was isolated from each population at defined time points. Each trajectory represents a unique mutation, colored by chromosome, within the isolated clone, whereas gray trajectories indicate mutations detected in competing lineages within a population. *(B)* The background-averaged fitness effect of each evolved mutation is measured through a bulk-segregant fitness assay where a segregant pool is propagated in the selective environment and allele frequencies are tracked by whole-genome, time-course sequencing. Fitness is calculated as the linear regression of the natural log ratio of evolved to ancestral allele frequency over time. The color scheme remains consistent between the evolutionary trajectories and bulk-segregant fitness assay. *(C)* Individual clones isolated from a bulk-segregant pool are assayed for fitness against an ancestral reference in a flow cytometry-based competition to determine how fitness segregates in the cross (yellow). The fitness distribution of the individual segregants is compared to the fitness of the evolved clone from which they arose (green). Shown is the additive expectation of all driver mutations (blue) and all detected mutations (red) as measured by bulk-segregant fitness assays.
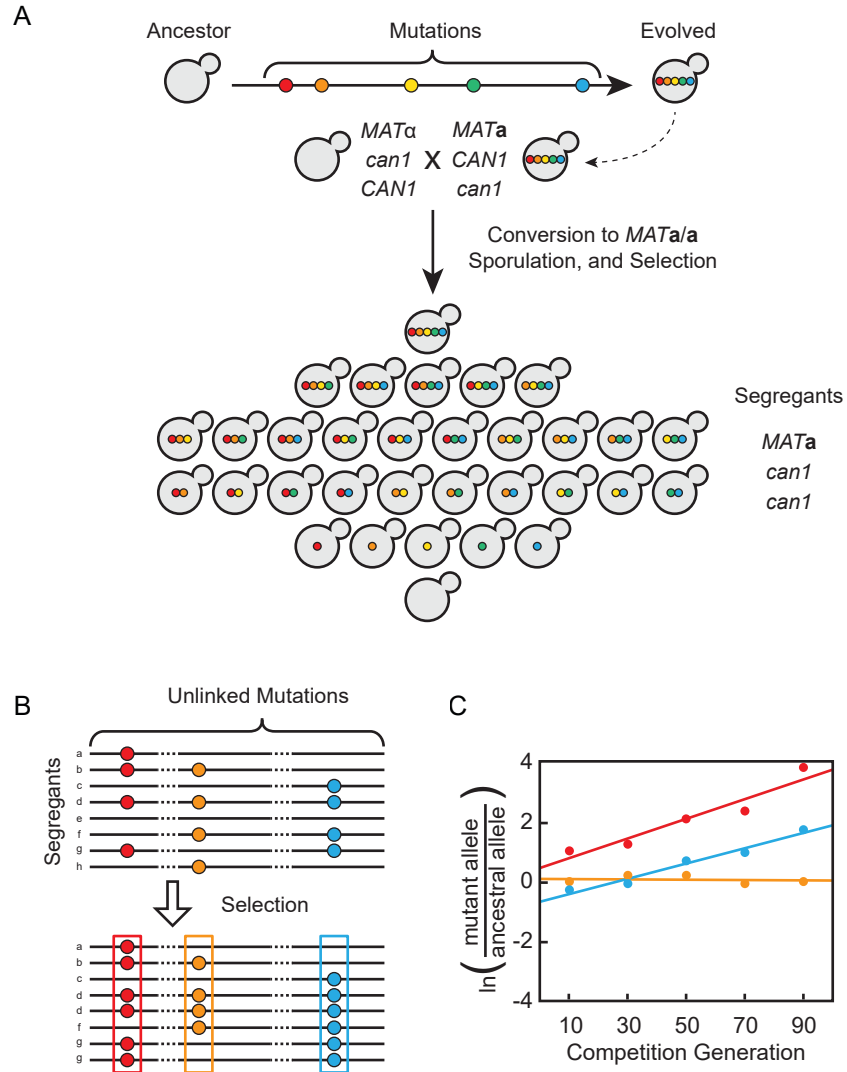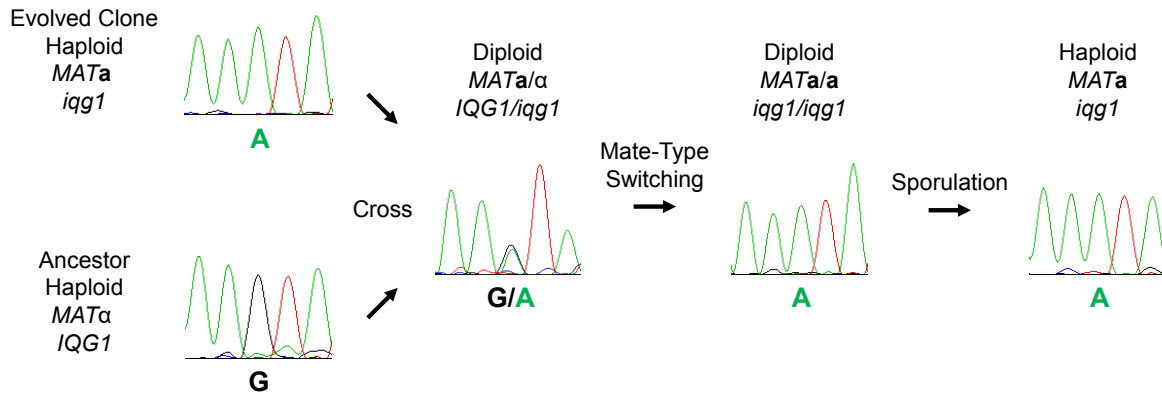
Fig. S2: Construction of bulk-segregant pools and implementation of bulk-segregant fitness assays. *(A)* To uncover how evolved mutations impact fitness, an evolved clone is selected for bulk-segregant analysis. First, the evolved *MAT***a** clone is crossed to a *MAT*α version of the ancestor engineered to possess a functional *CAN1* gene and non-functional *can1* gene at unlinked loci. The resulting *MAT***a**/α diploid is heterozygous at all loci mutated in the evolved clone. The diploid is then mating-type switched to *MAT***a**/**a** to ensure that only *MAT***a** progeny are produced. To sporulate, the diploid is complemented with a plasmid harboring the *MAT*α2 gene, and isolation on canavanine-containing media selects against any unsporulated diploids. Each segregant contains a random combination of evolved mutations, and the pool, collectively, contains all possible genotypes. *(B)* Each bulk-segregant pool is propagated in the selective environment for 100 generations, during which time allele frequencies change based on their fitness effect. Samples are collected every 20 generations and analyzed by whole-genome sequencing. *(C)* The background-averaged fitness effect of a mutation is measured as the change in the natural log ratio of allele frequency over time. Beneficial mutations (red and blue) increase in frequency while neutral mutations (orange) remain steady over time.
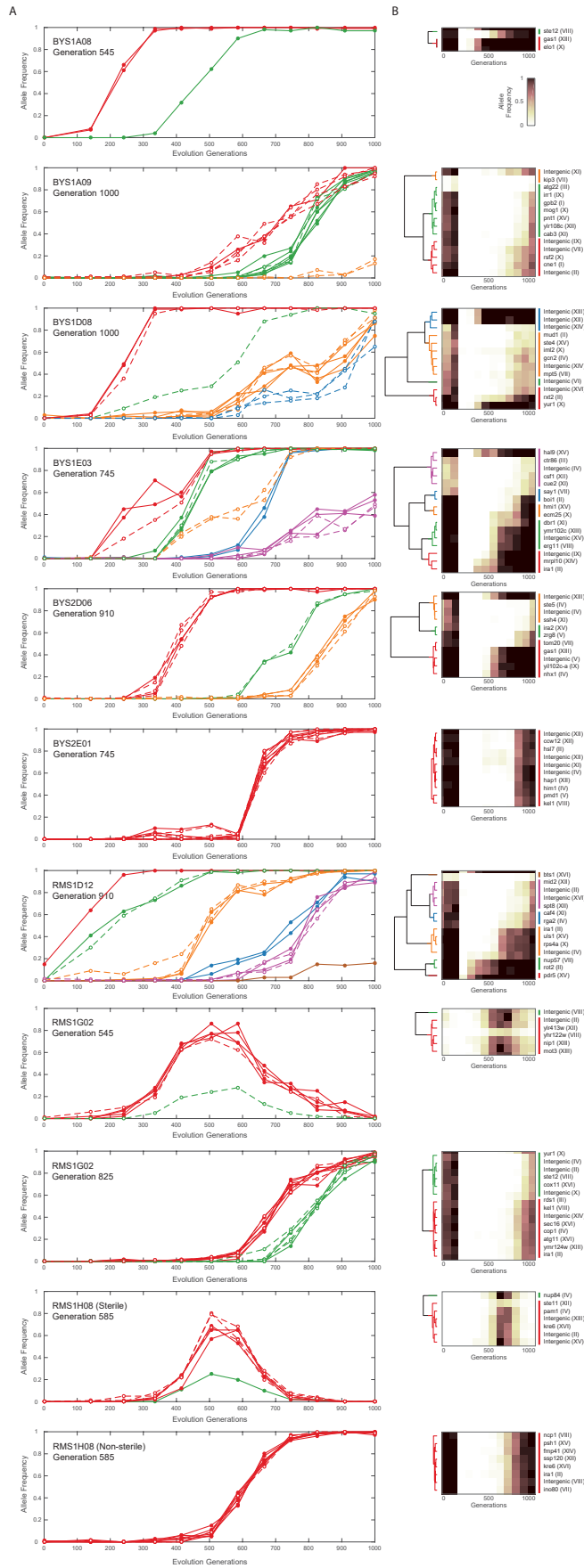
Fig. S3: Mutational cohort clustering and dynamics of selected evolved lineages. Mutations were assigned into cohorts through grouping of evolved alleles based on a hierarchical clustering approach that leverages the known evolutionary dynamics. The heat maps reflect allele frequencies, and dendrograms display the distance and relationship between mutations. Mutations are colored by cohort designation, in accordance with both the dendrogram and evolutionary trajectories.
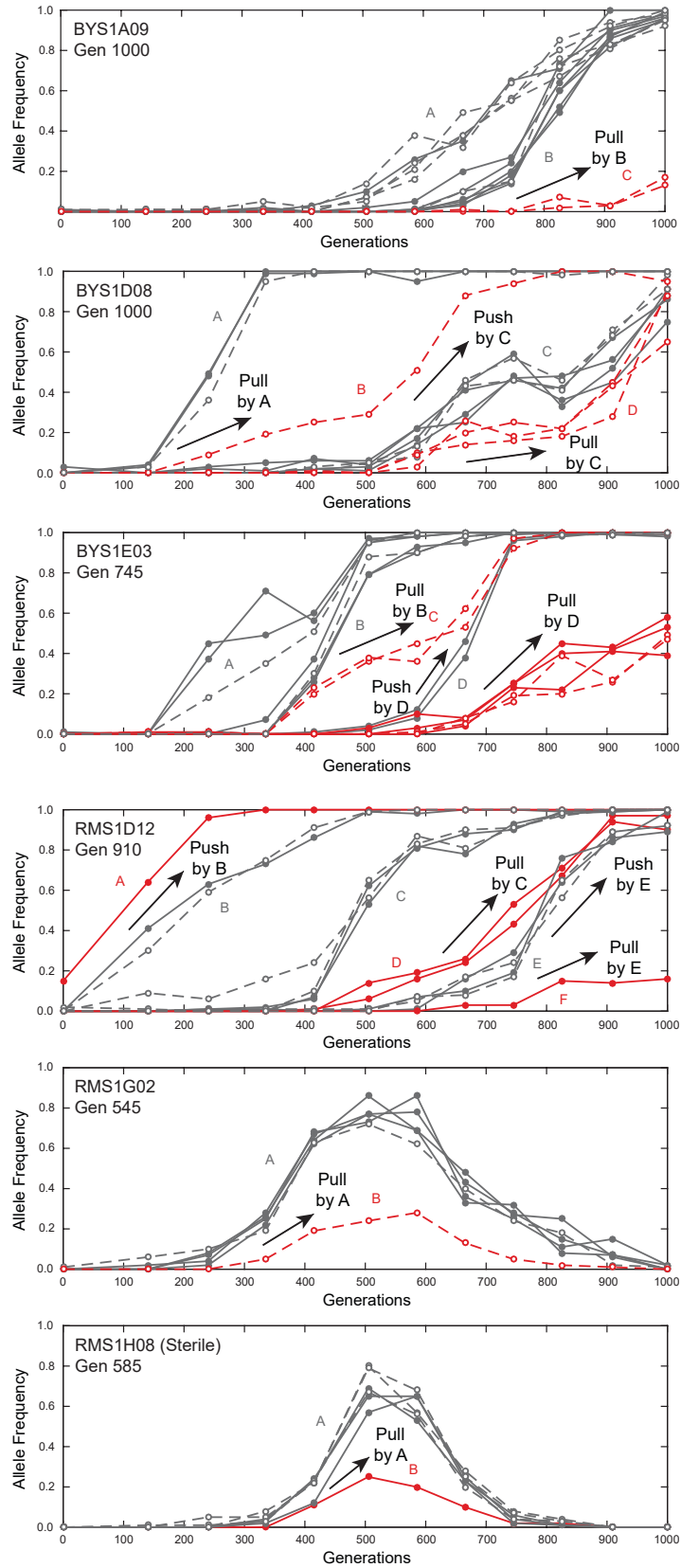
Fig. S4: Driverless cohorts rise due to one or more rounds of genetic associations with adaptive cohorts. A significant number of mutational cohorts are devoid of detectable driver mutations. These driverless cohorts (red) often contain few mutations and exhibit evolutionary trajectories that mimic the neighboring adaptive cohorts (gray). We propose that mutations within a driverless cohort arise on the background of an adaptive cohort while at low frequency and, subsequently, are "pulled" up in frequency. Once pulled to an appreciable frequency, the fate of a driverless cohort is then dependent upon the next beneficial mutation and the genetic background in which it occurs. If the beneficial mutation arises in the background of the driverless cohort, the driverless cohort will be "pushed" up in frequency.

Fig. S5: Sequence confirmation of the gene conversion event at the *IQG1* locus during construction of the BYS2E01 bulk-segregant pool. To generate the BYS2E01 bulk-segregant pool, the evolved clone (*iqg1*) was crossed to the ancestor (*IQG1*). The resulting *MAT**a**/α* diploid exhibited the expected *IQG1/iqg1* genotype. However, after mate-type switching, the resultant *MAT**a**/**a*** diploid possessed only the *iqg1* allele, indicative of a gene conversion event. Following sporulation of the *MAT**a**/**a*** diploid, all haploid segregants possessed the *iqg1* allele.
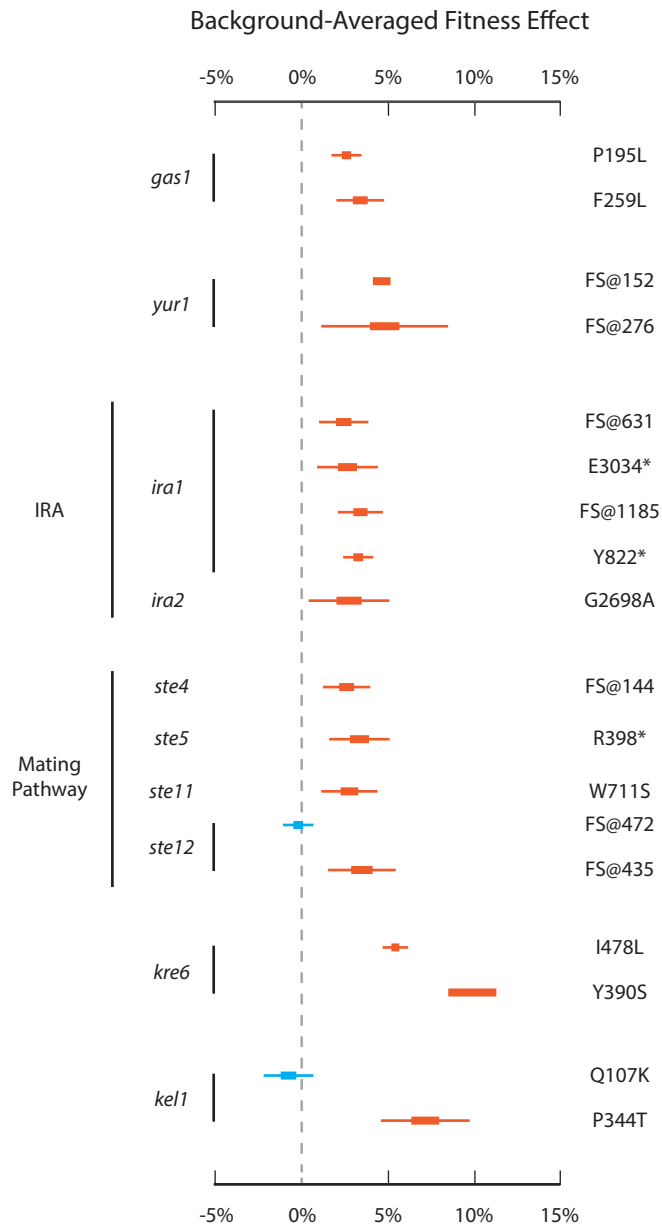
Fig. S6: Evolved mutations in common targets of selection confer similar fitness effects. We compared the background-averaged fitness effects of mutations in common targets of selection, defined as genes or genetic pathways that were mutated more than once within the eleven evolved lineages in this study. Thick bars refer to the standard error and thin bars refer to the 95% confidence interval of the background-averaged fitness effect of each mutation as determined through propagation and sequencing of each bulk-segregant pool (see Fig. 2). Both driver mutations (orange) and neutral mutations (blue) are represented. Each mutation is labeled by its predicted coding change (FS: frameshift, *: stop codon).
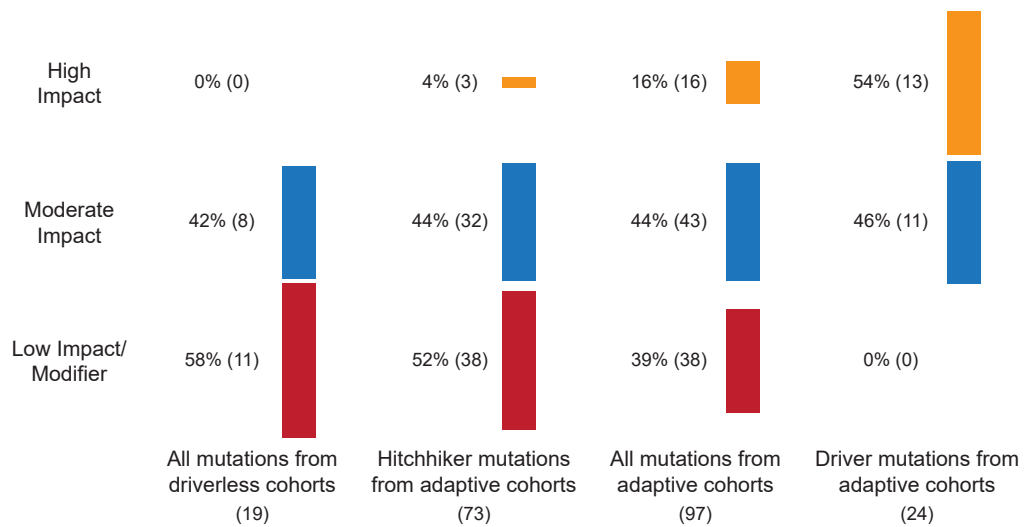
Fig. S7: Mutations in driverless cohorts resemble hitchhiker mutations. Mutations were divided into categories based upon their protein coding effect (high impact: frameshift and nonsense, moderate impact: missense, low impact/modifier: synonymous/intergenic/intronic; SNPeff, (5)). The mutational spectrum of driverless cohorts closely resembles the spectrum of hitchhiker mutations from adaptive cohorts (p = 0.90) while it is significantly different from the spectrum of driver mutations (p = 0.09) and adaptive cohorts (p < 10-6), Fisher-Freeman-Halton exact test).
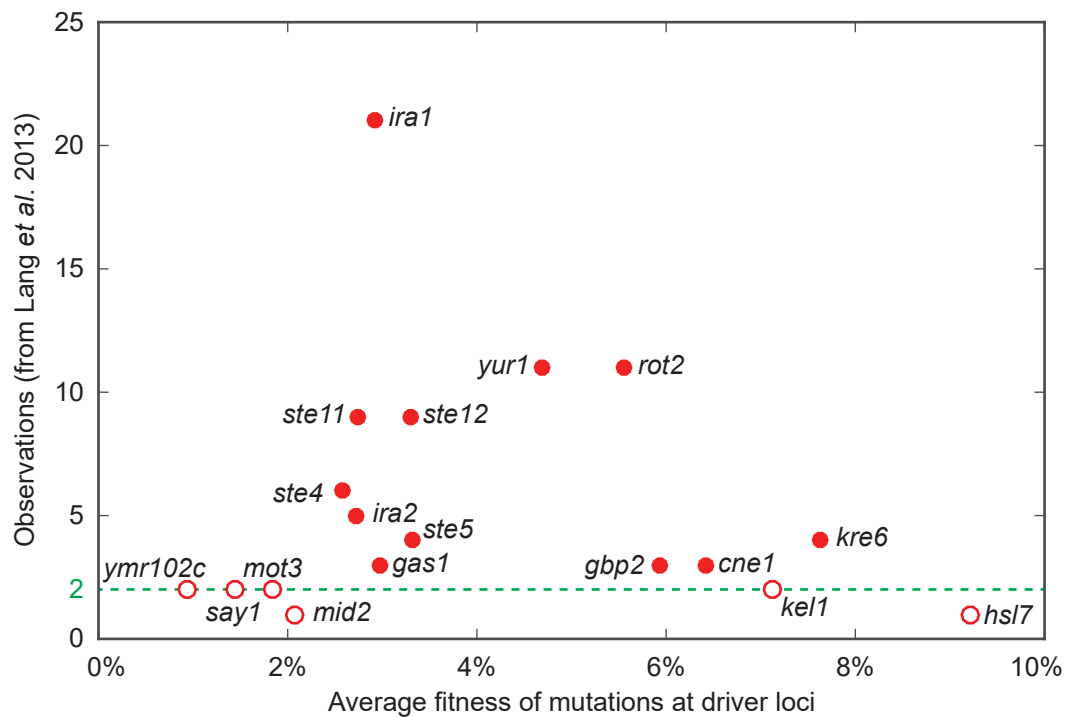
Fig. S8: Fitness effect alone fails to predict the number of occurrences of each driver mutation. The number of observed mutations in a given gene across forty replicate populations (1) is not correlated to its background-averaged fitness effect (ρ = 0.11, p = 0.66, Spearman, two-tailed). For *gas1*, *kre6*, and *yur1*, the fitness values are the average of two measurements from independent populations. Similarly, for *ira1* the fitness value is the average of four measurements from independent populations. For *kel1* the fitness value is only from clone BYS2E01-745 and does not include the neutral *kel1* mutation in population clone RMS1G02-825. Open circles are driver mutations that were missed by our previous recurrence-based methods that identified putative driver mutations based on three or more observations across forty replicate populations (1). These include modest-effect mutations (~1-2%) as well as the large-effect interacting mutations, *kel1* and *hsl7*.
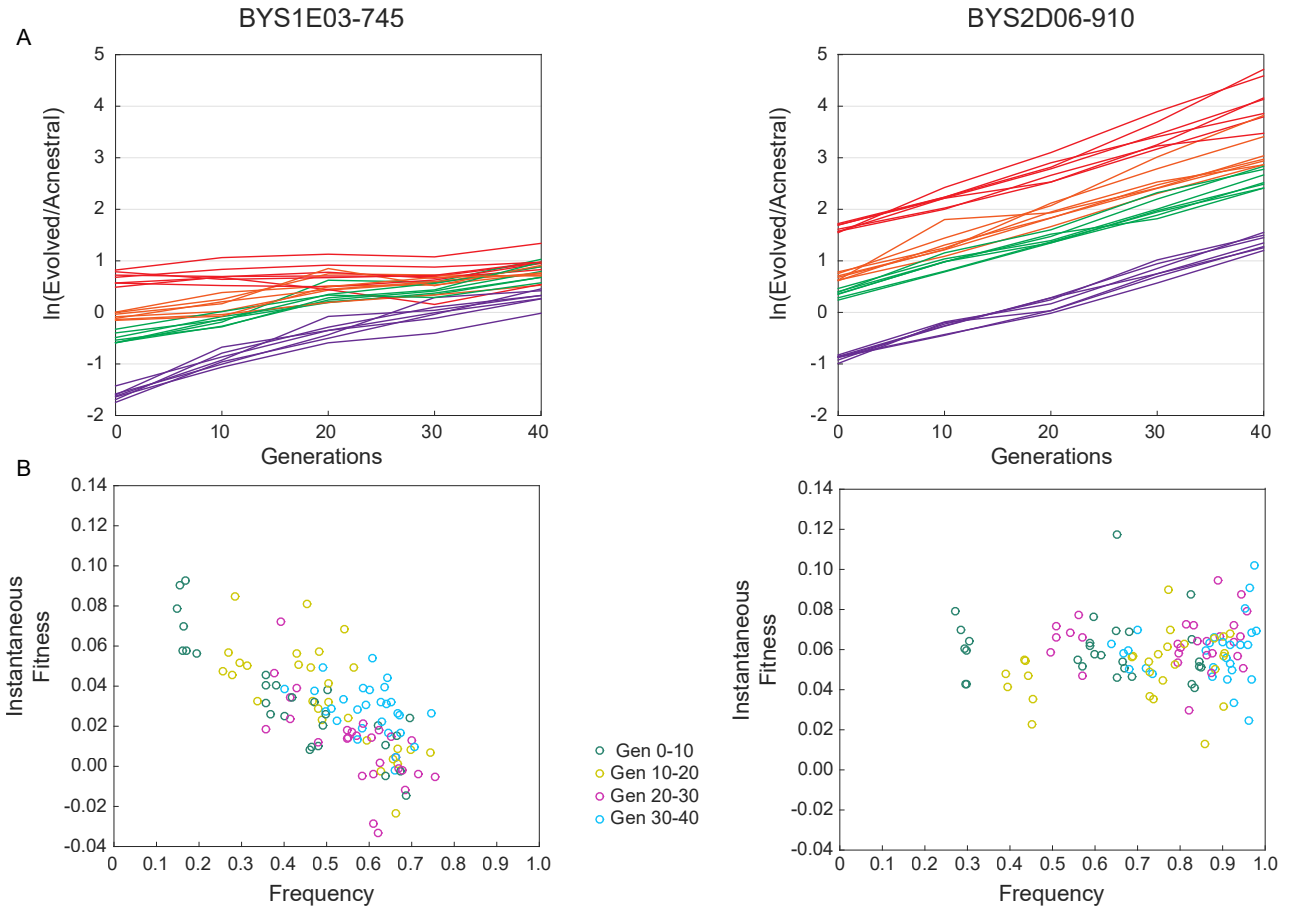
Fig. S9: Evolved clone from population BYS1E03 exhibits negative frequency-dependent fitness. Each evolved clone was assayed for fitness against an ancestral reference strain in seven replicate competitions. *(A)* In each competition, the evolved clone was seeded at four different starting frequencies (purple: 0.25, green: 0.50, orange: 0.75, and red: 0.90). The BYS2D06-910 competitions converge to a common ratio regardless of initial frequency, and subsequently its fitness (i.e. slope) decreases over time. In contrast, the fitness of BYS1E03-745 and all other evolved clones (Dataset S1) remains constant over time. *(B)* The instantaneous fitness values of clones BYS2D06-910 and BYS1E03 are calculated as the change in the log ratio of allele frequency across two consecutive timepoints (n and n+10). The fitness of clone BYS2D06 is inversely proportional to competition frequency (ρ = -0.72; Pearson correlation), whereas all other clones are independent of frequency, including BYS1E03 shown here (ρ = 0.10, Pearson correlation).
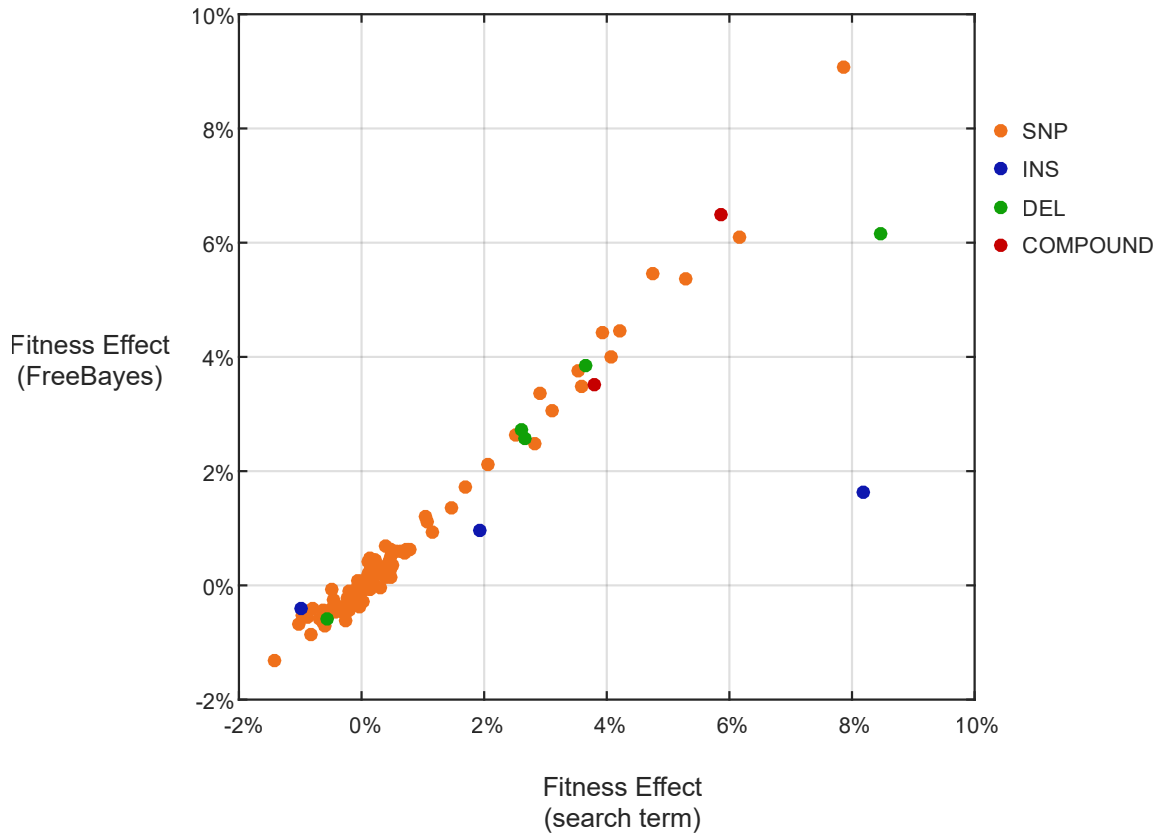
Fig. S10: Comparison of methods for determining allele counts within the bulk-segregant fitness assay. Allele frequencies for the bulk-segregant fitness assay were determined through the analysis of Illumina sequencing data using two independent approaches: 1) BWA/FreeBayes and 2) our in-lab grep-based 'search term' approach. In both cases, the fitness effect of each evolved allele was calculated as the change in log ratio of allele frequency over time. Data from the first round of sequencing, which represents ~50% of total read coverage, were used to compare approaches. Manually investigation of the discrepancies validated the 'search term' approach, which was then used exclusively for all downstream analysis.
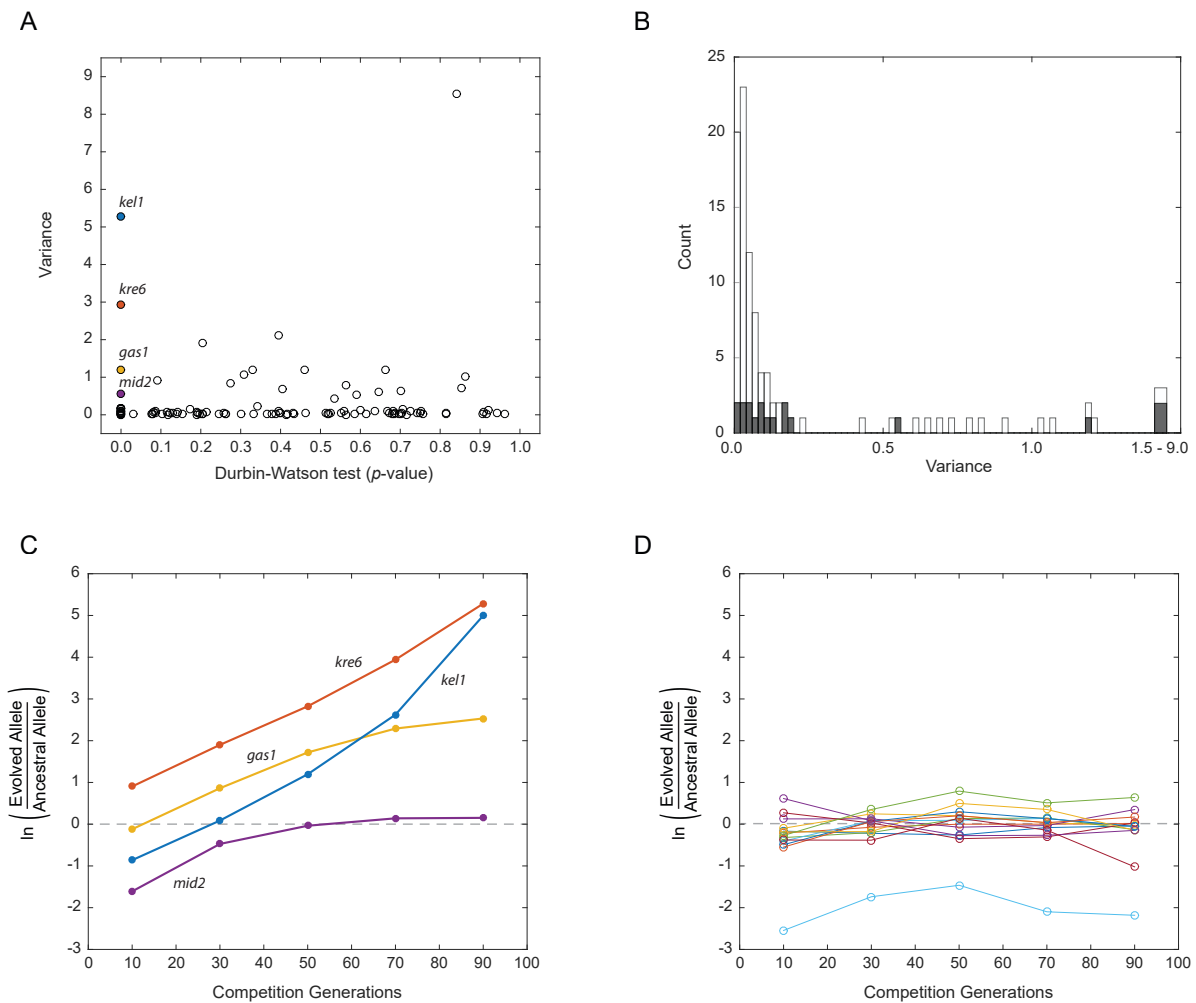
Fig. S11: Fitness estimates are consistent throughout the duration of the bulk-segregant fitness assay. The Durbin-Watson statistic was utilized to test the assumption of linearity in our regression model when estimating fitness from the bulk-segregant fitness assays. *(A)* For each allele, the Durbin-Watson p-value was plotted against the variance in the log ratio of allele frequencies. The null hypothesis of no autocorrelation in residuals was rejected at the $\alpha = 0.01$ level for 18 alleles. Amongst these statistically significant alleles are four drivers (*kel1*: blue, *kre6*: orange, *gas1*: yellow, *mid2*: purple). *(B)* Alleles were divided into two groups based upon their significance by the Durbin-Watson test (white: $p > 0.01$, black: $p < 0.01$). Visualized as a histogram, the distribution of variance is similar for the two groups, suggesting a high rate of type I error in the Durbin-Watson test. *(C)* Shown are the dynamics of the bulk-segregant fitness assay for the four driver mutations from panel *A*. Two driver mutations appear to increase in fitness (*kel1* and *kre6*) while two appear to decrease in fitness (*gas1* and *mid2*) over the course of the bulk-segregant fitness assay. The most extreme change occurs in *kel1*, which we have confirmed interacts epistatically with another driver mutation, *hsl7*. *(D)* Shown are the dynamics of the bulk-segregant fitness assay for the 14 hitchhiker mutations from panel *A*. These hitchhikers exhibit little variance and with minor deviations from zero, in contrast to the driver mutations in panel *C*.
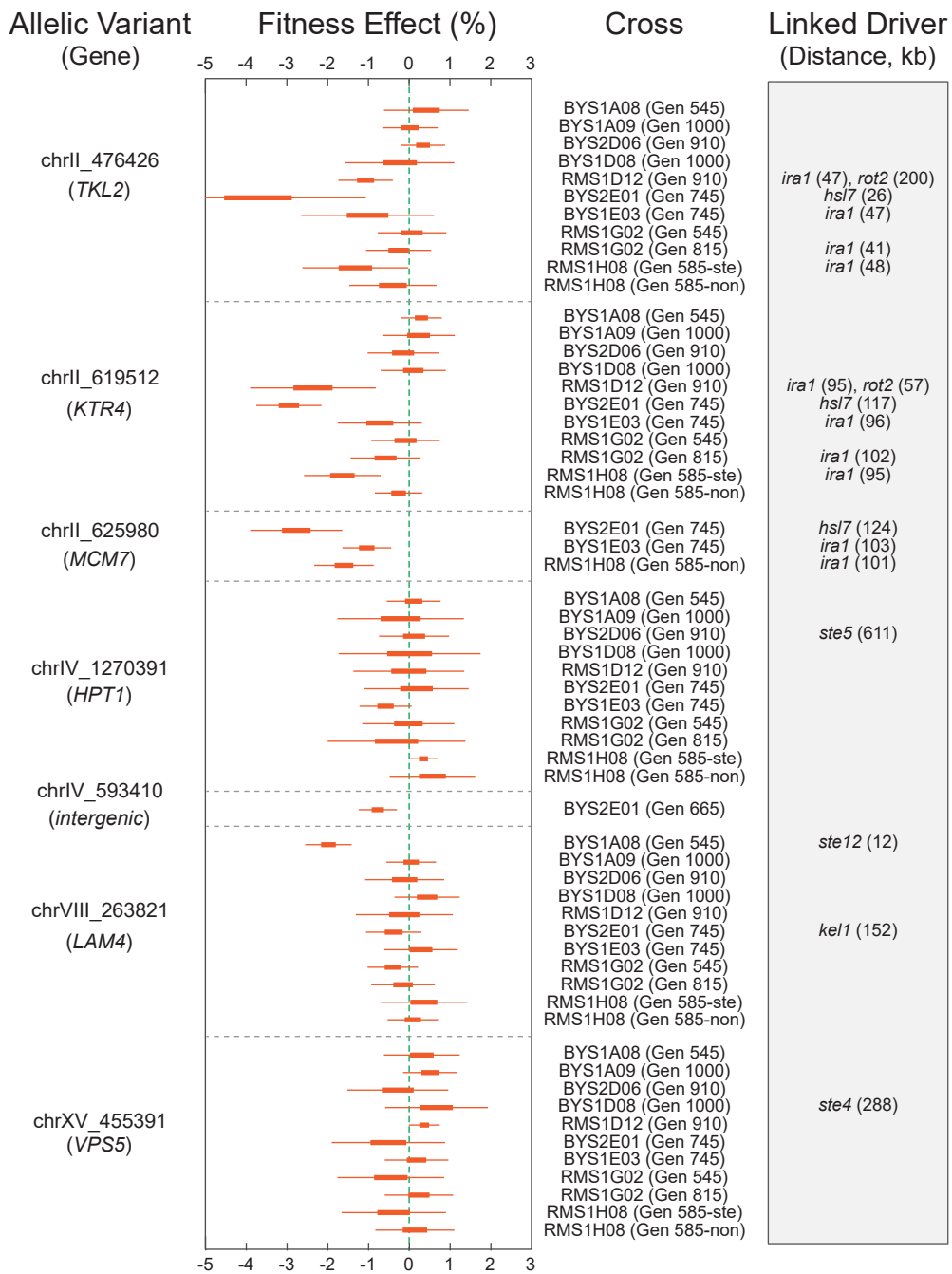
Fig. S12: Newly discovered allelic variants have minimal impact on fitness. Previously unknown allelic variants were identified through sequencing of the bulk-segregant individuals from the BYS2E01 cross. The frequencies of the seven discovered allelic variants were then tracked over time within all eleven bulk-segregant pools. Most variants are present in all pools, indicating a disparity between the *MAT**a*** ancestor and *MAT*α parent. An allele detected in only a fraction of the eleven pools (i.e. chrII_625980) was likely present at an intermediate frequency in the *MAT**a*** ancestral or *MAT*α parental culture. An allele specific to progeny from a single cross (i.e. chIV_593410) was either a low-frequency evolved mutation or a mutation that occurred during the construction of the bulk-segregant pool. Most genetic variants appear neutral by the bulk-segregant fitness assay. Variants that depart from neutral are explained via linkage to a known driver mutation.

SI REFERENCES

1. Lang GI, *et al.* (2013) Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500(7464):571-574.

2. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 25(14):1754-1760.

3. Garrison E & Marth G (2012) Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.

4. Ho CH, *et al.* (2009) A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. *Nature biotechnology* 27(4):369-377.

5. Cingolani P, *et al.* (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* 6(2):80-92.

6. Lang GI, Murray AW, & Botstein D (2009) The cost of gene expression underlies a fitness trade-off in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 106(14):5755-5760.