nature communications

Article

https://doi.org/10.1038/s41467-026-77469-1

Pervasive weak incompatibilities shape phenotypic variation and genome evolution in yeast hybrids

Received: 25 January 2026

Accepted: 24 August 2026

Published online: 09 September 2026

Check for updates

Artemiza A. Martínez & Gregory I. Lang

Species barriers limit gene flow and maintain co-adapted genomes. Interspecific hybridization can break down species barriers, generating phenotypic variation and altering evolutionary potential. Although the phenotypic and genomic consequences of hybridization have been extensively studied in many systems, far less is known about the longer-term evolutionary dynamics of highly divergent, mosaic genomes and the extent to which genetic incompatibilities shape their adaptation. By suppressing anti-recombination genes in meiosis, we generate 20 haploid recombinant hybrids from a cross between *Saccharomyces cerevisiae* and *Saccharomyces paradoxus*. Across conditions, these hybrids are, on average, less fit than either parent and show broad phenotypic variation. Patterns of protein-complex inheritance reveal no evidence of strong pairwise lethality but support pervasive weak negative genetic interactions. Using laboratory evolution, we show that each recombinant genome follows a distinct adaptive trajectory, and that a small subset of hybrid protein complexes and loci show hybrid-specific mutational targeting. Species-of-origin alleles also bias evolutionary outcomes by reshaping selection on interacting genes. Together, these results indicate that strong pairwise incompatibilities are rare, while weak, background-dependent interactions are widespread and shape both phenotypic variation and genome evolution in hybrid genomes.

Understanding how genetic background influences adaptation is a central goal in evolutionary biology. Natural hybrids are widespread across taxa and offer powerful systems in which to study the genetic basis of adaptation and speciation. Hybridization is a double-edged sword. On one hand, hybridization can disrupt interactions between coevolved genes, leading to incompatibilities that affect fitness or viability. This outcome of hybridization is well studied and is widely acknowledged as a key contributor to the maintenance and reinforcement of species barriers[1,2]. On the other hand, hybridization can create or potentiate novel phenotypes absent from either parent[3,4]. Although still debated[5], this idea raises the possibility that hybridization can facilitate rapid adaptation to novel environments[3]. More broadly, hybridization creates mosaic genomes in which different combinations of parental alleles may occupy distinct positions on the fitness landscape. Prior work has shown that even within species, different recombinant backgrounds can vary substantially in their fitness, adaptability, and evolutionary trajectories[6]. However, how highly mosaic hybrid genomes respond to selection over many generations, and the extent to which weak genetic incompatibilities contribute to these differences, remains less well understood.

Interspecific hybridization is frequent between yeast in the genus *Saccharomyces* due to weak pre-zygotic barriers[7,8]. Diploid F1 hybrids, formed from haploid strains of different species, are easily generated in the laboratory and are widespread in natural environments where they are valued for advantageous traits including altered metabolism, stress tolerance, and enhanced fermentation performance[9–12].

Department of Biological Sciences, Lehigh University, Bethlehem, PA, USA. e-mail: glang@lehigh.edu

Numerous studies have shown that evolution of diploid hybrid genomes, both in nature and in the laboratory, is driven largely by loss of heterozygosity and genomic rearrangement[13–16].

While diploid F1 hybrids of *Saccharomyces* yeasts can be propagated mitotically and are often more fit than either parent due to hybrid vigor (heterosis), strong post-zygotic species barriers prevent the recovery of viable progeny following meiosis. Despite many attempts to identify Dobzhansky–Muller incompatibilities in yeast using genetic screens, chromosome-replacement experiments, genome-wide analyses of rare viable spores, and quantitative trait mapping in partially fertile crosses, no strong pairwise nuclear-nuclear incompatibilities have been identified[17–22]. While karyotype differences exist between some *Saccharomyces* species and these have been shown to contribute to species barriers[23,24], the primary post-zygotic barrier is sequence divergence, which reduces meiotic viability by impairing proper chromosome pairing[25].

This post-zygotic barrier can be partially overcome by eliminating anti-recombination mechanisms that typically prevent mispairing in meiosis[25]. A recent study showed that placing two genes (*MSH2* and *SGS1*) under the control of a mitotic-specific promoter that is repressed in meiosis enables recovery of viable haploid recombinant hybrids (segregants)[26]. Unlike diploid F1 hybrids, these segregants fix allelic combinations to a single parental species at each locus, making it possible to associate phenotypic variation with mosaic patterns of parental inheritance and to test how hybrid genomic background shapes subsequent evolution[17,22,27,28].

Here, we generate, phenotype, and evolve a panel of 20 segregants between *S. cerevisiae* and its sister species *S. paradoxus*[29]. Across conditions, these segregants are, on average, less fit than either parent. By combining phenotyping with genetic mapping, we identify both Mendelian and quantitative traits in the segregants, including a defect in mother–daughter cell separation that maps to the mitotic regulator *AMN1*. To test for pairwise genetic incompatibilities, we classify the species-of-origin composition of all annotated protein complexes. Every two-subunit complex forms an interspecific (hybrid) assembly in at least one segregant, indicating that lethal pairwise incompatibilities are rare or absent; however, two-subunit hybrid complexes are underrepresented, indicating broad constraints on interspecific assembly. Finally, we evolve the segregants and their corresponding diploid or higher-ploidy derivatives for 10,000 generations and track their fitness gains and patterns of genomic change. The resulting populations follow distinct adaptive trajectories, showing that the rate and genetic basis of adaptation depend strongly on hybrid background. A small subset of genes and two-subunit complexes accumulate mutations more often than expected in hybrid genomes, identifying candidate targets of hybrid-specific selection; these include recurrent mutations in *HSP104* that arise in half of the hybrid lineages. We further identify the *S. cerevisiae* allele of the prion factor *SUP35* as a potentiating variant that increases the likelihood that *HSP104* becomes a repeated target of selection. Together, these results support a model in which strong pairwise incompatibilities are rare, but weak, background-dependent interactions are widespread and shape both phenotype and evolution in recombinant hybrid yeast genomes. Similar landscapes of weak incompatibility may therefore be a general feature of other recently formed hybrid lineages.

Results

Meiotic recombination produces balanced interspecific hybrid genomes

S. cerevisiae and *S. paradoxus* are sister species that diverged ~5 million years ago. The two species are ~85% identical at the amino acid level, yet the genomes are nearly completely syntenic (Supplementary Fig. 1)[16,30,31]. Of ~6600 annotated genes in *S. cerevisiae*, 5366 have unambiguous one-to-one nuclear orthologs in *S. paradoxus*, and 99% of orthologs are not significantly different in size (Wilcoxon signed-rank, p value < 0.001, Supplementary Data 1). These two species are separated by strong post-zygotic barriers due to improper chromosome pairing and segregation in meiosis, leading to spore viability of less than 1%.

We constructed prototrophic *S. cerevisiae* S288C and *S. paradoxus* CBS432 strains in which the mismatch repair gene *MSH2* and the helicase gene *SGS1* are under the control of the meiotically repressed *CLB2* promoter. Transient suppression of Msh2/Sgs1 during meiosis relaxes mismatch-repair–mediated anti-recombination between the highly diverged homologs, increasing crossovers and improving chromosome segregation. While this manipulation reduced spore viability in the intraspecific crosses, it increased spore viability in the interspecific hybrid cross to 23.5% and raised the recovery of fully viable tetrads from 0 to 3%, consistent with previous reports[26]. Using this approach, we isolated 20 segregants from five four-spore viable tetrads (Fig. 1a, b and Supplementary Fig. 2a).

We confirmed by whole-genome sequencing that the segregants carry mosaic nuclear genomes with nearly-equal contributions from both parents, together with tetrad-specific mitochondrial inheritance patterns (Fig. 1c and Supplementary Fig. 2b). Crossovers were markedly reduced. We find 17.6 crossovers per meiosis, ~4.5 times fewer than the ~70–90 crossovers that are typical in *S. cerevisiae*[32]. Approximately one quarter of chromosomes are uniparental throughout their length (Fig. 1d and Supplementary Fig. 2b), and smaller chromosomes are more likely to be uniparental (Supplementary Fig. 2c).

These recombinant genomes show both Mendelian and quantitative trait variation. Among these traits, clump formation is the most visually striking and reflects a parental difference in cellular morphology: *S. cerevisiae* primarily grows as dispersed single cells, whereas *S. paradoxus* forms small aggregates (Fig. 2a). Among 20 segregants, this phenotype segregated approximately 2:2 (Fig. 2a and Supplementary Fig. 3a). We identified a ~ 222 kb region on Chromosome II whose segregation pattern matched the clumping phenotype in all but one spore. This interval contains 112 annotated genes, including *AMN1*, which has been previously implicated in mother–daughter separation[33]. Using CRISPR–Cas9, we exchanged the parental alleles of *AMN1* in both species and in four segregants derived from a single tetrad. The *S. paradoxus* allele ($AMN1^{368D}$) inhibited mother–daughter separation, leading to cell clumps, whereas the *S. cerevisiae* allele ($AMN1^{368V}$) restored growth as single cells (Fig. 2a). Thus, clumpiness segregates as a Mendelian trait in this interspecific cross, whereas most other phenotypes in the panel show quantitative, continuous variation.

Recombinant hybrids are on average less fit than either parent

We next asked how the segregants varied for quantitative traits affecting growth and reproduction. We quantified traits linked to fitness in the segregant panel, including thermotolerance, growth rate, competitive fitness, and sporulation efficiency. Consistent with previous observations[34], *S. cerevisiae* exhibits greater thermotolerance than *S. paradoxus*. The diploid F1 hybrid exceeded the growth of both parents. By contrast, the 20 segregants showed a broad range of temperature responses, with most segregants growing well at 30 °C, but failing to form robust colonies at 37 °C (Fig. 2b). Quantification of growth at 37 °C relative to 30 °C shows that only a few segregants approach the thermotolerance of *S. cerevisiae*. Segregants 2 A, 4B and 1B retained approximately 80–85% of their own 30 °C growth at 37 °C, but remained less thermotolerant than *S. cerevisiae* (two-sided Welch's two-sample t-tests, Benjamini–Hochberg-adjusted, p value < 0.0001 for each; Fig. 2c). Similar variability was also observed at 16 °C. Direct growth-rate measurements corroborated these findings, showing intermediate phenotypes with a marked decline at 37 °C, consistent with the spot assay results (Supplementary Fig. 3b, c). Together, these results show that thermotolerance is a highly variable trait in the segregants.

https://doi.org/10.1038/s41467-026-77469-1

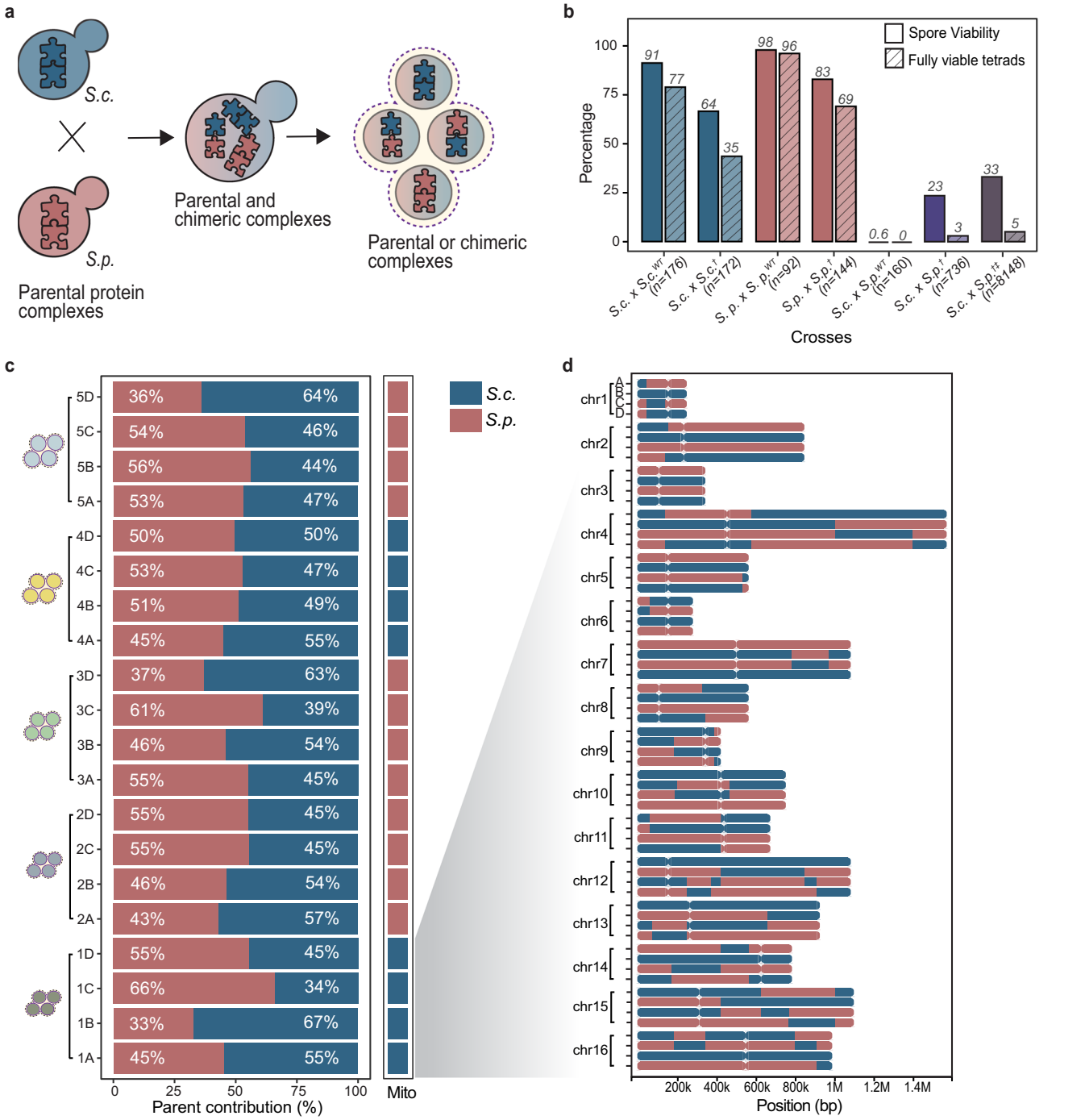
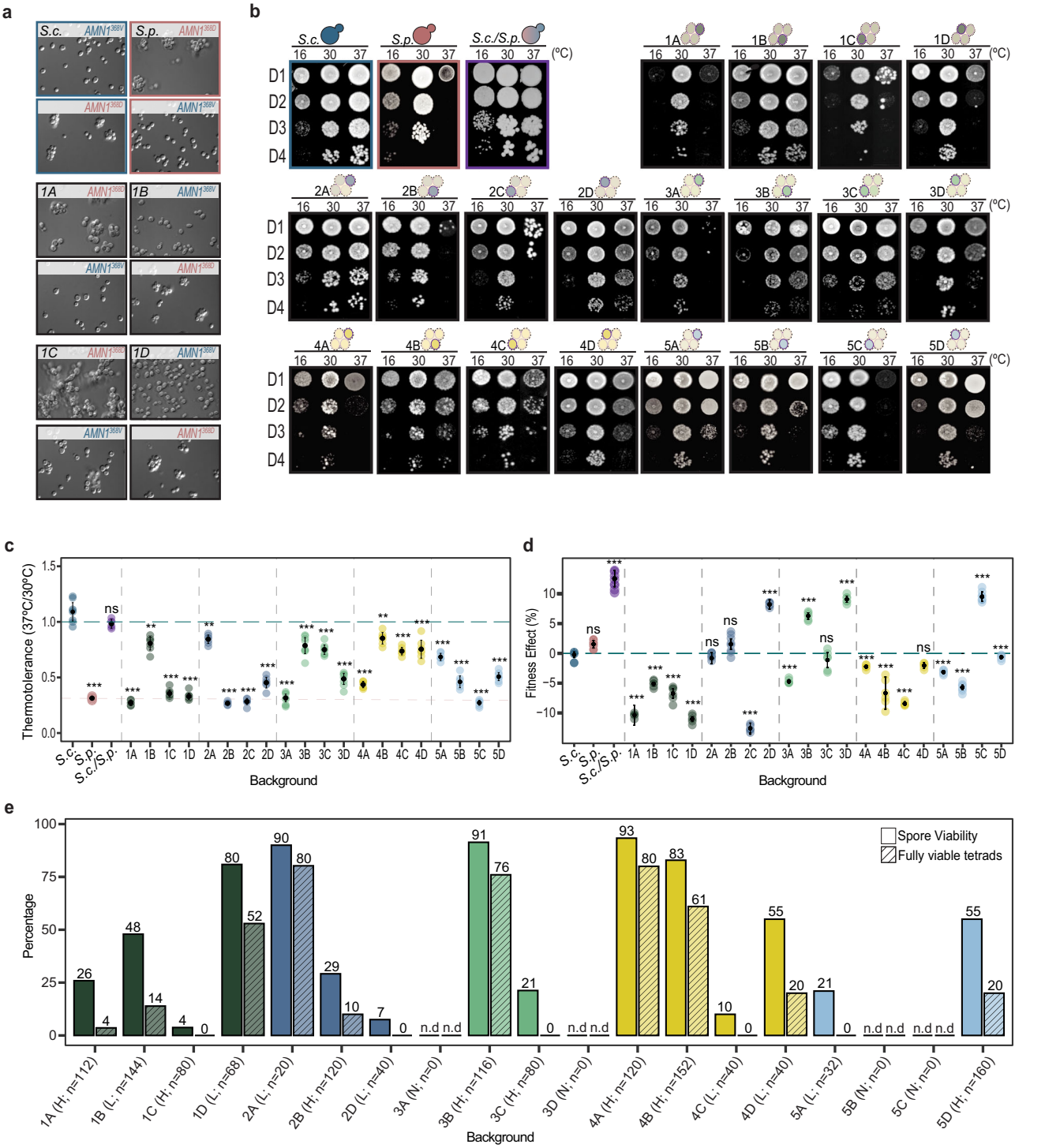


Fig. 1 | Construction and genomic architecture of recombinant haploid hybrids. a Haploid recombinant hybrids (segregants) were generated by placing the anti-recombination factors *SGS1* and *MSH2* under control of the meiosis-repressed *CLB2* promoter, transiently suppressing mismatch-repair–mediated anti-recombination during meiosis, increasing viable spore recovery and yielding genomes with mosaic nuclear ancestry. Recombination reshuffles parental alleles so that multi-subunit protein complexes can assemble in either parental or chimeric configurations; some chimeric assemblies are likely suboptimal and thus constitute potential targets of selection for compensatory change. **b** Spore viability, calculated as the percentage of all dissected spores that germinated, and fully viable tetrads (hatched bars), calculated as the percentage of dissected tetrads that yielded four viable spores, for intra- and interspecific crosses of *S. cerevisiae* (*S. c.*) and *S. paradoxus* (*S. p.*), with and without *CLB2*-driven *SGS1/MSH2* suppression. Parentheses indicate total numbers of spores dissected. †, temporal suppression of *SGS1/MSH2*; ‡, data from Bozdag et al., 2021. **c** Nuclear ancestry across 20 segregants (five tetrads), shown as the percentage contribution from *S. c.* (blue) and *S. p.* (pink) in each genome. Adjacent bars indicate the parental mapping assignment of the representative mitochondrial marker *ATP6*, which was consistent among spores from the same tetrad. **d** Chromosome-wide recombination maps for segregants A–D of Tetrad 1, showing balanced 2:2 segregation of parental alleles along each chromosome (colors as in panel c). Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-026-77469-1



To quantify overall performance, we next measured competitive fitness at 30 °C in rich glucose medium. At this temperature, the two parental species show comparable growth, whereas the diploid F1 hybrid displays clear heterosis. In contrast, most segregants are less fit than one or both parents, and only a few exceed both (Fig. 2d). Competitive fitness correlates moderately with growth rate ($r = 0.56$), but with several strong outliers (Supplementary Fig. 3d), consistent with contributions from additional traits beyond simple growth rate.

Given the pronounced variation in vegetative growth, we next asked whether sexual fitness is similarly affected. We measured sporulation efficiency for each segregant after constructing segregant-derived homozygous diploids. We observed wide variation in both sporulation efficiency and spore viability (Fig. 2e), indicating that meiotic performance also varies substantially across hybrid backgrounds. Both growth and sporulation are sensitive to mitochondrial function, and nuclear–mitochondrial incompatibilities have been described previously in Saccharomyces yeasts[19,20]. However, no clear respiratory defect was observed among segregants carrying the *S. cerevisiae MRS1* allele, including those showing *S. paradoxus*-like mapping support at individual mitochondrial loci (Supplementary Fig. 3e).

Fig. 2 | Segregants exhibit broad phenotypic diversity, reduced fitness and variable meiotic competence. a Clumping phenotype segregates 2:2 within each tetrad. Segregants carrying the *S. paradoxus* (*S. p.*) $AMN1^{368D}$ allele display mother–daughter separation defects, whereas those with the *S. cerevisiae* (*S. c.*) $AMN1^{368V}$ allele separate normally. CRISPR–Cas9 allele swaps in parental strains and segregants confirm the causal effect of the *AMN1* allele. **b** Spot-dilution assays at 30 °C, 37 °C and 16 °C. Temperature-dependent growth profiles reveal both thermal sensitivity and emergent thermotolerance among segregants. The diploid F1 hybrid shows heterosis under all three conditions (D1–D4 indicate four successive 1:32 serial dilutions). **c** Thermotolerance was estimated from spot assays using $n = 7$ independently measured cultures per background, except for the F1 hybrid ($n = 4$), 1 A and 2B ($n = 6$ each), and 1 C, 1D, 3A–3D, 5 C and 5D ($n = 8$ each). Growth at 37 °C was normalized to growth at 30 °C using the equal-average method. Asterisks mark thermotolerance values that differ significantly from the *S. cerevisiae* parent (two-sided Welch's two-sample t-tests on R = S37 °C/S30 °C, with Benjamini–Hochberg correction across all comparisons; $*p_adj < 0.05$, $**p_adj < 0.01$, $***p_adj < 0.001$; ns, not significant). Error bars show mean ± s.e.m. **d** Competitive fitness of the F1 hybrid and 20 segregants relative to the *S. cerevisiae* parent. The *S. paradoxus* parent did not differ significantly from *S. cerevisiae*. Differences in competitive fitness were assessed using two-sided Z-tests comparing grouped selection-coefficient estimates and their regression standard errors, with Bonferroni correction: $*p_adj < 0.05$, $**p_adj < 0.01$,$***p_adj < 0.001$; ns, not significant. Error bars represent the standard error of the grouped regression estimate. Sample sizes were $n = 8$ independent biological competition assays per background, except for 1 A and 4B ($n = 6$ each) and 3 C and 4 A ($n = 7$ each). Exact sample sizes are also provided in the Source Data file. **e** Meiotic competence of segregant-derived diploids, quantified as spore viability and recovery of fully viable tetrads (hatched bars). Values were calculated across all dissected tetrads (n = total spores dissected). Sporulation categories are indicated as H (high), L (low) or N (no spores). For strains scored as N, spore viability and fully viable tetrad values could not be determined (n.d.). No diploid derivative was obtained for segregant 2 C. Source data are provided as a Source Data file.

Weak selection acts against hybrid protein complexes

Having established that many recombinant genotypes show reduced performance across multiple environments, we next asked whether any signal of incompatibility could be detected at the level of protein–protein interactions[12,27,35,36]. One plausible source of such incompatibilities is the misassembly of protein complexes, in which subunits from different species are forced to interact. If such interactions were strongly deleterious, we would expect viable recombinant genomes to be depleted of hybrid assemblies and enriched for uniparental complexes, especially among small complexes where each subunit makes a larger contribution to overall function.

We curated a set of 652 protein complexes whose subunits could be confidently identified as one-to-one orthologs (Supplementary Data 2) and, for each of the 20 segregants, classified every complex as "hybrid" (containing ≥1 subunit from each species) or "uniparental" (all subunits from a single species). As expected, smaller complexes (2–4 subunits) were hybrid in ~62% of complex–segregant combinations, whereas larger assemblies (≥5 subunits) were hybrid in ~96% (Supplementary Fig. 4a). Across all 13,040 complex–segregant combinations, the probability that a complex assembles as a hybrid increases by ~2.3-fold with each additional subunit (logistic regression, p value < 10^{-16}), consistent with a combinatorial expectation that larger assemblies are increasingly likely to contain subunits from both species. This strong size effect indicates that larger complexes are informative mainly at a descriptive level, whereas tests for excess or depletion of hybrid assembly are most interpretable in smaller complexes.

To test whether hybrid protein complexes are under- or over-represented relative to a neutral expectation, we compared the observed frequencies of hybrid and uniparental two-subunit protein complexes against a null model that takes into account that spores from a single tetrad are not independent. We therefore focused on two-subunit complexes ($n = 230$), where hybrid and uniparental configurations have equal baseline probability and null expectations are more straightforward to define. We excluded protein complexes in which both subunits are encoded by genes on the same chromosome, because linkage constrains their segregation.

Every two-subunit hybrid complex was recovered in at least two independent segregants, showing that there are no pairwise lethal incompatibilities. We observe, however, a slight skew toward uniparental complexes in the overall distribution, consistent with weak selection against hybrid assembly (Fig. 3a). We performed 10,000 randomizations of chromosomes within each of the five tetrads and recalculated the distribution of hybrid complexes. Fewer than 1% of the reshuffling trials produced a greater excess of complexes observed as hybrids in only 4 or 6 of the segregant genomes. In addition, complexes that occur by chance in the majority of segregant genomes are more likely to be nonessential (Fig. 3b).

Hybrid genomes follow unique evolutionary trajectories

Having established that recombinant hybrid genomes differ widely in fitness and other phenotypes, we next asked how these differences shape long-term adaptation. We therefore evolved haploid, diploid, and higher-ploidy hybrid backgrounds for 10,000 generations and tracked fitness over time relative to a fluorescently labeled *S. cerevisiae* reference strain (see Methods). The initial fitness of the haploid segregants and their corresponding diploid or higher-ploidy derivatives was only moderately correlated ($r = 0.63$, two-sided $p = 0.0040$; Fig. 4a), with some backgrounds either gaining or losing fitness after the shift in ploidy. Segregant-founded lines evolved heterogeneously: some backgrounds improved steadily, whereas others showed much greater divergence among replicates, with trajectories that separated over time or showed transient declines before later recovery. These patterns demonstrate strong genome background effects on adaptive trajectories. Across the panel, low-fitness founders had the most rapid initial rate of adaptation (Fig. 4b,c).

Although evolution occurred at 30 °C, we used thermotolerance at 37 °C as a stress-responsive trait that might reveal correlated changes in proteostasis or stress-buffering capacity. Because negative genetic interactions are often exacerbated at elevated temperatures, gains in growth at 37 °C could provide one indication that some backgrounds acquire mutations that improve performance under latent stress.

Among the parental lines, *S. cerevisiae* already grew robustly at 37 °C and showed little change over evolution (gen 5,000 p value = *0.70*; gen 9000 p value = *0.19*), whereas *S. paradoxus* remained poorly thermotolerant and even showed a transient reduction at generation 5,000 (p value = *0.002*), with one replicate improving (Supplementary Fig. 5). Segregant-founded populations, in contrast, displayed heterogeneous trajectories. Comparing evolved populations to their own ancestors, 10/20 hybrids showed a significant decrease in thermotolerance at generation 5000, and across all lineages the change in thermotolerance between generation 0 and 5000 was on average −0.13 ± 0.23 (mean ± s.d., p value < 10^{-12}), indicating a substantial global reduction early in evolution. By generation 9000, 7/20 hybrids remained significantly less thermotolerant than their ancestor, but globally the mean change had attenuated to −0.05 ± 0.59 and was no longer significantly different from zero (p value = *0.26*). At the lineage level, strong improvers were relatively rare (~20% of populations) and were strongly enriched in a small subset of backgrounds (notably 5 C and 5D). Thus, despite frequent modest losses, adaptation at 30 °C did produce clear increases in thermotolerance in a subset of hybrids, showing that high-temperature performance can improve even without direct selection.

https://doi.org/10.1038/s41467-026-77469-1

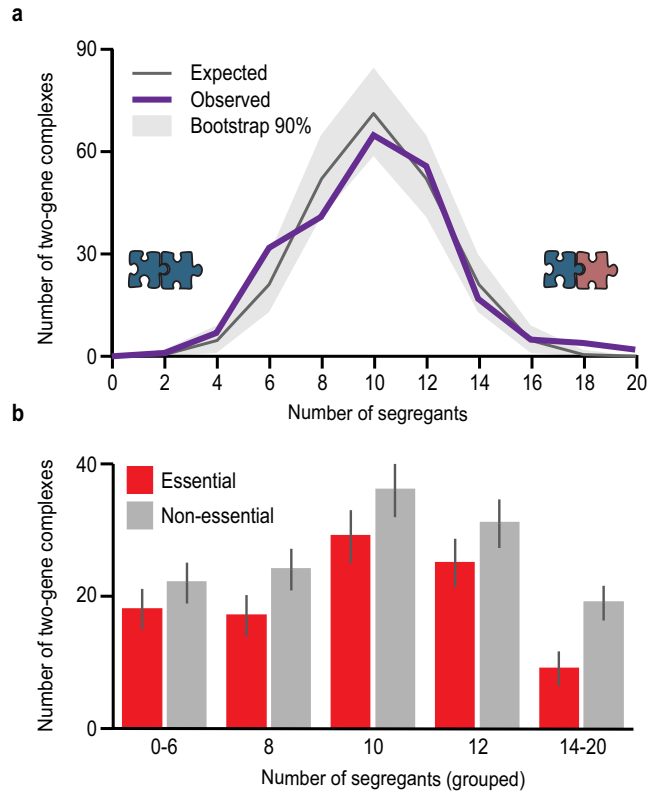


Fig. 3 | Two-subunit protein complexes show a slight but significant skew towards uniparental origin. a For each of 230 two-subunit complexes, we calculated the distribution of the number of segregant genomes in which the complex assembles as hybrid versus uniparental. We excluded complexes that are genetically linked on the same chromosome. The purple distribution shows the observed hybrid-complex frequencies across segregants, whereas the gray distribution shows the expectation under the null model. The observed distribution deviated significantly from the expected distribution ($p < 0.001$, chi-square goodness-of-fit test). To assign significance, we performed 10,000 randomization iterations by computationally shuffling the chromosomes within each of the five tetrads and recalculating the distribution of hybrid complexes. The gray shaded band represents the central 90% of the distributions obtained from these randomizations. The randomized data matched the expected distribution: 91.2% of randomization experiments had chi-square p value > 0.01, and only 8 of the 10,000 randomization experiments showed a stronger deviation from the expected distribution than the observed data. **b** Essential two-subunit complexes are observed less frequently as hybrid complexes relative to non-essential complexes. We define essential complexes as those in which either protein is classified as essential by the Saccharomyces Genome Database (yeastgenome.org). Complexes were grouped into five hybrid-frequency categories of the 20 segregants. Observed counts are shown for $n = 98$ essential and $n = 132$ non-essential complexes ($n = 230$ two-subunit complexes in total). Error bars are the standard deviation expected due to binomial sampling. Source data are provided as a Source Data file.

Genome background shapes the evolutionary outcomes of recombinant hybrids

To identify the genetic targets of selection, we sequenced evolved populations from diploid or higher-ploidy segregant backgrounds at generation 2400 and from segregant-founded haploid backgrounds at generation 5000. In total, we analyzed 328 independently evolved populations, including 38 parental populations, 15 F1 diploid hybrid populations, and 275 segregant-derived hybrid populations (Supplementary Fig. 7a and Supplementary Data 4). Across the 142 diploid and higher-ploidy founded populations, we identified 14,246 variants, most of which (87.3%) were called as heterozygous (Supplementary Fig. 6a). Across the 171 haploid-founded populations, we identified 21,933 variants (Supplementary Fig. 6b).

Autodiploidization is a common outcome of laboratory evolution experiments[37,38], and our haploid-founded lines were no exception: 78.5% of mutations in these populations are heterozygous (Supplementary Fig. 6a, c). We confirmed by flow cytometry that >90% of our haploid-founded lines underwent autodiploidization. Interestingly, nine out of the 14 populations founded from segregants 4 C (4/8) and 5D (5/6) remained haploid throughout 5000 generations (Fisher's exact p value = *0.003* and p value = *0.0002*, respectively; Supplementary Fig. 6d), indicating that the propensity to autodiploidize also varies across hybrid backgrounds.

Across hybrid genomes, we detected 29,650 mutations (38.6% intergenic, 32.2% missense, 15.0% synonymous, 11.4% frameshift, 2.6% nonsense), comparable to parental distributions (Supplementary Fig. 7b). Because we replaced the native promoter of *MSH2*, we anticipated a slight elevation in mutation rate. Through 5000 generations, we observe an average of 110 ± 66 mutations per lineage (Supplementary Fig. 7c, d), above the ~30 mutations we expected based on prior evolution experiments[37,39,40]. This excess is driven largely by frameshift mutations, which are ~4-fold more frequent than in earlier studies. Interestingly, populations founded from two hybrid backgrounds (3 C and 5B) have an excess of mutations relative to the other segregants, suggesting possible background-specific mutator phenotypes (Supplementary Fig. 7c).

We next asked which loci repeatedly acquired more nonsynonymous mutations than expected by chance. Among 13,283 nonsynonymous mutations (Fig. 5a), of which 9,139 are nuclear missense mutations (Supplementary Fig. 8a), shared targets—*IRA2*, *UPC2*, and *PTR2*—were enriched for nonsynonymous mutations in both hybrids and parents (Supplementary Fig. 8b, c), whereas *HSP104* and *BSC1* are strongly enriched only in hybrids despite a small number of mutations in the parental backgrounds (Supplementary Fig. 8d, e).

Because each hybrid genome is a distinct mosaic of *S. cerevisiae* and *S. paradoxus* alleles, we next asked whether the same targets were favored uniformly across backgrounds. In the absence of strong background effects, the same loci should recur similarly across hybrid founders. However, if genomic mosaicism leads to widespread and background-specific genetic interactions, evolutionary outcomes will be idiosyncratic, with selection acting on different loci in each hybrid background, even when the hybrids experience identical environments. To distinguish between these scenarios, we measured the variance-to-mean ratio of mutation counts across hybrid backgrounds. Among the 25 most frequently mutated genes, six had relative variance scores greater than 2, indicating a strongly non-uniform distribution of mutations across hybrid backgrounds (Fig. 5b).

Most genes identified as targets of selection carry mutations in both parental alleles, although some showed pronounced allelic bias (*IRA1*, 33/46 mutations in the *S. paradoxus* allele; *BSC1*, 19/21 mutations in the *S. paradoxus* allele). To infer whether selection favors complete loss-of-function or alteration/attenuation-of-function, we applied a Bayesian model to the observed mutational spectrum at each target gene (Fig. 5b). *IRA2* and *PTR2* accumulate missense, nonsense, and frameshift mutations, consistent with selection for loss-of-function. In contrast, *HSP104*, *UPC2*, and *BSC1* accumulate mostly missense mutations, suggesting that selection is acting to alter or attenuate the function of these genes.

As a complementary analysis, we tested whether these targets showed an excess of nonsynonymous mutations after accounting for synonymous and nonsynonymous mutational opportunities. *HSP104*, *UPC2*, and *PTR2* showed elevated dN/dS values and significant nonsynonymous enrichment, consistent with positive selection acting on protein-altering mutations. In contrast, *BSC1* showed a strong skew toward the *S. paradoxus* allele and mostly missense mutations but did not show significant nonsynonymous enrichment in the pooled dN/dS analysis, suggesting a more allele-specific or background-dependent pattern (Supplementary Data 5).

https://doi.org/10.1038/s41467-026-77469-1

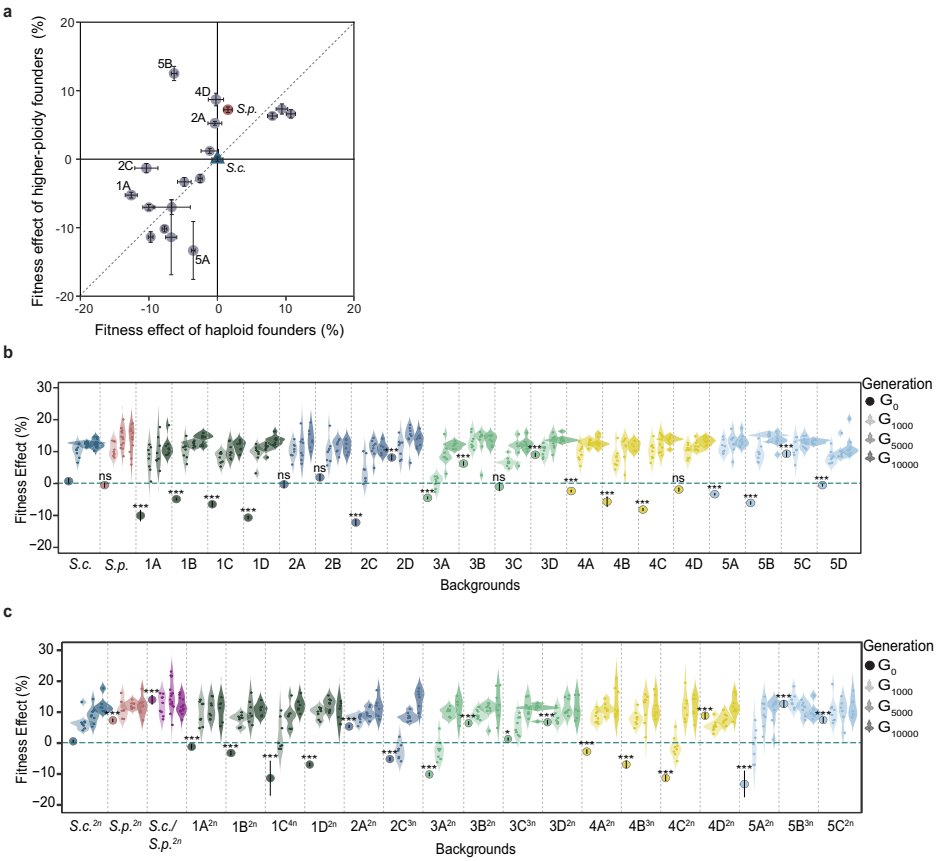


Fig. 4 | Long-term fitness trajectories of segregants and their derivatives. **a** Relationship between fitness effects, measured as selection coefficients, of segregants in the haploid state (x-axis) and their diploid or higher-ploidy derivatives (y-axis) at 30 °C. Each point represents one paired background ($n = 19$); error bars in both dimensions indicate the standard errors of the grouped selection-coefficient estimates. Fitness of the diploid and higher-ploidy founders was estimated from $n = 8$ independent biological competition assays per background, and haploid fitness estimates were taken from Fig. 2d. The solid line shows the linear fit, and the shaded band indicates the 95% confidence interval of the fit (Pearson's $r = 0.63$, 95% CI 0.24-0.84, two-sided p value = 0.0040). Segregants 1 A, 2 A, 2 C, 4D, 5 A, and 5B show the largest ploidy-associated fitness shifts, with differences greater than 5 percentage points in selection coefficient. **b** Relative competitive fitness of the haploid parental strains and segregants at generations 0, 1,000, 5,000 and 10,000, expressed relative to the *S. cerevisiae* ancestor. Generation-0 circles show the grouped selection-coefficient estimates reported in Fig. 2d; error bars indicate the corresponding regression standard errors. Violin plots show the distributions of independently evolved replicate populations at each subsequent time point, with each observation representing one independently evolved population. For generation 0, differences in competitive fitness relative to the *S. cerevisiae* parent were assessed using two-sided Z-tests comparing grouped selection-coefficient estimates and their regression standard errors, with Bonferroni correction across 22 comparisons: $^{*}p_adj < 0.05$, $^{**}p_adj < 0.01$, $^{***}p_adj < 0.001$; ns, not significant. Error bars for generation-0 estimates represent the standard errors of the grouped regressions. **c** Relative competitive fitness of the derived diploid and higher-ploidy founders and their evolved populations, displayed as in b. Generation-0 circles show the grouped selection-coefficient estimates reported in Fig. 4a. Each observation within a violin represents one independently evolved population. Ploidy states are indicated as 2n, 3n, and 4n. Although diploid derivatives were initially sought, some backgrounds yielded triploid or tetraploid derivatives; these available higher-ploidy derivatives were retained as founders for the evolution experiment. Generation-0 differences relative to the *S. cerevisiae* parent were assessed using the same two-sided Z-test approach described in b, with Bonferroni correction across 19 comparisons. Source data are provided as a Source Data file.

Finally, we asked whether hybrid protein complexes are preferential targets of compensatory evolution. Among two-subunit complexes, deviations from the null expectation were rare: only three complexes showed some of the largest positive mutation excesses and were retained as exploratory candidates for hybrid-enriched mutational targeting: *RSP5–BUL1*, *YKU70–YKU80*, and *DNF3–CRF1* (Fig. 5c). Thus, although hybrid-specific targets of selection do emerge, mutational targeting of hybrid protein-complex assembly appears to be uncommon. These results support a model in which interspecies protein interactions are broadly tolerated, with only subtle, complex-specific constraints.

The *S. cerevisiae SUP35* haplotype is associated with repeated selection on *HSP104*

HSP104 was among the most frequently mutated genes in the hybrid populations and remained a strong candidate after multiple statistical tests for selection (32 nonsynonymous mutations, length-weighted Poisson test p value < 10^{-13}; relative variance score = 3.25; dN/dS = 7.974, p value = 0.006; G-score = 8.429, p value = 0.004) across backgrounds (Fig. 5b). The missense mutations were not randomly distributed across the protein: 23 of 28 unique mutated residues mapped to the ATPase-containing NBD1/NBD2 domains. This enrichment was significant relative to expectation based on domain length (one-sided binomial test, p value = 0.0009; permutation test with 10,000 randomizations), supporting the interpretation that recurrent *HSP104* evolution preferentially targets domains central to ATP-dependent disaggregase activity (Supplementary Fig. 9).

Within the hybrid panel, *HSP104* mutations did not show a significant allele-specific bias (two-sided binomial test, p value = 0.345). In contrast, parental evolved populations showed a strong background-specific pattern: *HSP104* mutations were recovered in 9 of 22 independently evolved *S. cerevisiae* populations but in none of the 16 *S.*

https://doi.org/10.1038/s41467-026-77469-1

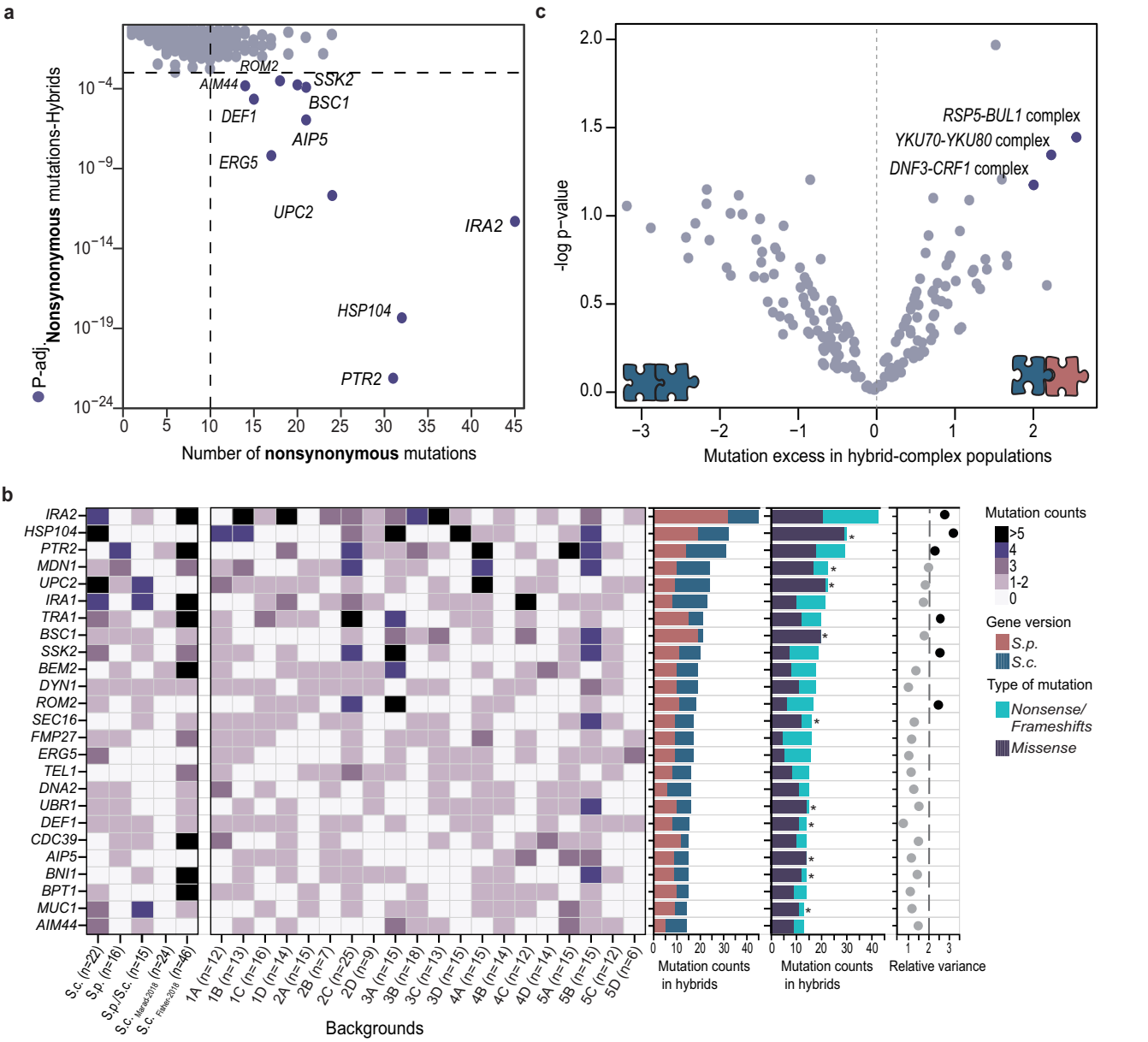


Fig. 5 | Gene- and complex-level targets of selection in evolved hybrid genomes. **a** Gene-level enrichment of nonsynonymous coding mutations across ~5800 CDSs in hybrids, evaluated under a Poisson expectation. The x-axis shows the number of nonsynonymous mutations per gene; the y-axis shows the probability of observing that number by chance, weighted by the average CDS length across species. Genes exceeding the multiple-testing threshold (Benjamini–Hochberg-adjusted p value $< 10^{-4}$) and with at least 10 nonsynonymous mutations are highlighted in purple. **b** Recurrently mutated genes. Left, heat map of the top 25 genes ranked by total nonsynonymous counts across hybrid populations, parental controls and published datasets (Fisher et al., 2018; Marad et al., 2018) (darker shading indicates higher mutation counts). Center, bar plots partitioning counts by allele of origin—*S. paradoxus* (pink, S. p.) versus *S. cerevisiae* (blue, *S. c.*). Right, mutation class and selection signal: bars show missense (purple) versus nonsense/frameshift (light blue) mutations. Posterior probabilities from a Bayesian model were used to assess whether the observed mutational spectrum was more consistent with selection for non-loss-of-function versus loss-of-function changes. An asterisk indicates strong posterior support for the non-loss-of-function model (posterior probability > 0.995). Prior probabilities were based on seven independent yeast evolution experiments (Martínez & Lang, 2023). Adjacent dot plot: relative variance scores quantify deviation from neutral expectations; genes with scores ≥ 2 (dashed line) show strongly non-uniform and repeatable targeting across hybrid genetic backgrounds, consistent with background-dependent targeting. **c** Enrichment of missense mutations in 250 two-subunit protein complexes classified as hybrid or uniparental across evolved hybrid populations. The x-axis shows mutation excess and the y-axis shows $-\log_{10}(P)$, where P is the chi-square-test p value. Highlighted complexes represent exploratory candidates for hybrid-enriched mutational targeting (dark purple). Source data are provided as a Source Data file.

paradoxus populations, despite 99.3% amino acid identity between the two Hsp104 orthologs (Fisher's exact test, one-sided p value = *0.0031*). This pattern argues against divergence at *HSP104* itself as the primary driver of selection and instead points to the surrounding genetic background. Consistent with this interpretation, the propensity to acquire *HSP104* mutations often segregated 2:2 within tetrads, such that replicate populations founded from two spores repeatedly acquired *HSP104* mutations, whereas those founded from the other two did not. Together, these observations suggest that a segregating *S. cerevisiae*-derived locus contributes disproportionately to the selective context in which *HSP104* becomes a repeated target of selection.

A genome-wide scan of parental ancestry revealed a single over-represented region: nearly all hybrids with *HSP104* mutations retained a conserved *S. cerevisiae* haplotype (~318 kb) on Chromosome IV

containing the interactors *SUP35* (96.5% identity) and *HSP42* (85.7% identity). The only exception (5 A) lacked this block but carried a mutation in *HSP42* (Fig. 6a). Additional mutations in *SUP35* itself (4B, 4D) further implicated this region as a candidate genetic context influencing the repeated evolution of *HSP104*.

SUP35 encodes a prion-forming translation termination factor. We assayed the prion state using an *ade1*-W244* translational read-through reporter, in which non-prion colonies are red and prion colonies are pink/white. The *S. cerevisiae* parent exhibits a weak prion phenotype, whereas *S. paradoxus* is non-prion, and the recombinant hybrids span the full range from strong prion (1B, 3 A) to weak prion (1 A, 1D, 3 C, 3D) to non-prion (1 C, 3B) (Fig. 6b). Inhibiting Hsp104 activity with GdnHCl shifted most lines toward the non-prion state, but two hybrids (3D, 1D) remained white and grew robustly on −Ade, indicating genetic restoration of adenine biosynthesis rather than simple prion curing.

To test how evolved *HSP104* mutations affect prion state, we reconstructed two evolved mutations in the *S. cerevisiae* parent and in two segregants (1 A and 3 A) that carry both the *S. cerevisiae* SUP35 allele and the *S. cerevisiae HSP104* allele; direct editing of the *S. paradoxus HSP104* allele was unsuccessful. The *hsp104*-D408Y substitution, which originally arose in a population founded from segregant 1B where the resident allele is *S. paradoxus*, caused only modest shifts in prion strength when introduced into *S. cerevisiae* and the segregants. By contrast, *hsp104*-S409Y, which arose in a 3A-founded population carrying the *S. cerevisiae* allele, consistently cured the prion phenotype when reconstructed in the *S. cerevisiae* parent and in both segregants (Fig. 6c). Overall, these results show that the *S. cerevisiae SUP35* haplotype redirects selection onto *HSP104* and that the consequences of *HSP104* mutations for prion state are strongly conditioned by hybrid genetic background.

Discussion

Meiosis fails in interspecific yeast diploids due to defects in chromosome pairing and segregation[25]. Suppressing anti-recombination genes partially overcomes this strong post-zygotic barrier[7]. Using this strategy, we isolated 20 viable recombinant hybrids between *S. cerevisiae* and *S. paradoxus*, whose nuclear genomes are a balanced mosaic of the two evolutionarily divergent yet syntenic parental genomes. These segregants provide a powerful system to dissect how hybridization reshapes the fitness landscape and influences subsequent adaptive evolution. The segregants display a range of phenotypic variation, from clear 2:2 segregation of a cell separation phenotype that maps to the *AMN1* locus, to quantitative, continuous variation in growth rate, thermotolerance, and sporulation efficiency. While some segregants have higher fitness, on average, they are less fit than either parent across tested conditions, consistent with fitness-reducing genetic interactions arising when coadapted genomes are recombined.

At the level of protein complexes, all two-subunit hybrid assemblies are observed at least once, establishing that these combinations are compatible with viability. This is in line with proteomics studies in diploid interspecific hybrids, which show that chimeric protein complexes form with similar stoichiometries to uniparental complexes[12,35], and is consistent with previous work that also found no evidence for classic Bateson–Dobzhansky–Muller incompatibilities between *Saccharomyces* yeasts[21,41,42]. Yet we also find that two-protein complexes show a slight but significant overall bias toward uniparental composition among viable hybrids. We do not interpret this pattern as evidence for obligate lethal incompatibilities. Rather, it is consistent with weak and potentially background-dependent incompatibilities that reduce the probability of spore survival, germination, or recovery. Under this model, the viable segregants analyzed here represent a filtered subset of meiotic products, in which hybrid configurations are tolerated but are underrepresented overall because their fitness effects are incompletely penetrant.

Unlike nuclear-nuclear incompatibilities, several nuclear-mitochondrial incompatibilities are known to exist[17,22,43]. Surprisingly, however, we do not observe the previously identified incompatibility between an *S. cerevisiae* nuclear-encoded gene, *MRS1*, and the *S. paradoxus* mitochondrial genome[17], likely reflecting strain-specific differences between the *S. paradoxus* backgrounds.

Each hybrid segregant is a unique mosaic of *S. cerevisiae* and *S. paradoxus* alleles, and therefore, each occupies a different position on the fitness landscape. Hybrid genotypes, therefore, displayed not only substantial variation in their ancestral fitness but also variation in their rates of adaptation. Some backgrounds evolved in highly parallel ways across replicate populations, whereas others showed much greater divergence, indicating that adaptive trajectories depend strongly on hybrid genomic background. The recombinant panel functions not only as a collection of hybrid genotypes, but also as a mapping framework for background-dependent evolution.

Most genetic targets of selection are not determined simply by the species of origin of the gene itself. The one exception is *BSC1*, which was observed as a significant target of selection only in hybrids that carried the *S. paradoxus* version of this gene (Fig. 5b). *BSC1* is a highly diverged, repeat-rich locus. The *S. paradoxus* version of the gene is significantly larger than the *S. cerevisiae* version due to a fusion event (Supplementary Fig. 1b and Supplementary Fig. 9b, c). This gene is also highly variable within *S. cerevisiae*[44]. The mutational signatures are consistent with intra-genic rearrangements between repeat sequences, rather than de novo SNPs (Supplementary Fig. 10b,c), suggesting that *BSC1* is a structurally labile locus whose involvement in adaptation depends on hybrid background.

Of all recurrent targets of selection, *HSP104* showed the strongest background dependence. The 32 nonsynonymous mutations in *HSP104* arose in just 10 of the 20 hybrid backgrounds, as well as in the *S. cerevisiae* but not the *S. paradoxus* parent. Each of the five tetrads from which the hybrid lineages were derived show 2:2 segregation of the propensity to acquire *HSP104* mutations, suggesting that a single locus from *S. cerevisiae* is imposing selective pressure on *HSP104*. Our ancestry mapping and prion assays point to *SUP35* as the most likely candidate. The *S. cerevisiae SUP35* allele is associated with greater prion-forming potential, and Hsp104 is a key regulator of Sup35/[PSI +] propagation. Together, these observations support a model in which the *S. cerevisiae SUP35* haplotype potentiates repeated selection on *HSP104* in a subset of hybrid backgrounds. Although the mechanistic basis of this effect remains unresolved, this result shows that the hybrid panel can be used to map background-specific genetic interactions that shape where selection acts. More broadly, it illustrates how recombinant hybrid genomes can be used not only to identify repeated targets of selection, but also to map segregating loci that influence when and where those targets are favored.

Thermotolerance provides an additional window into how hybrid backgrounds respond to selection. We initially treated growth at 37 °C as a proxy for temperature-sensitive genetic interactions, reasoning that elevated temperature should exacerbate defects in proteostasis and other fragile cellular processes. However, long-term evolution at 30 °C produced highly heterogeneous thermotolerance outcomes. Most hybrids showed modest losses or no net change in thermotolerance, whereas a minority of backgrounds produced strong improver lineages whose final high-temperature growth rose well above their ancestral distribution. Thus, adaptation at 30 °C did not generally resolve temperature-sensitive defects across the panel. Instead, correlated improvements in thermotolerance were restricted to a subset of genetic backgrounds, indicating that only some hybrid genomes can access adaptive routes that improve both baseline fitness and stress tolerance.

If weak incompatibilities in hybrid complexes contribute broadly to reduced hybrid performance, then chimeric protein assemblies might be expected to emerge as preferential mutational targets during

https://doi.org/10.1038/s41467-026-77469-1

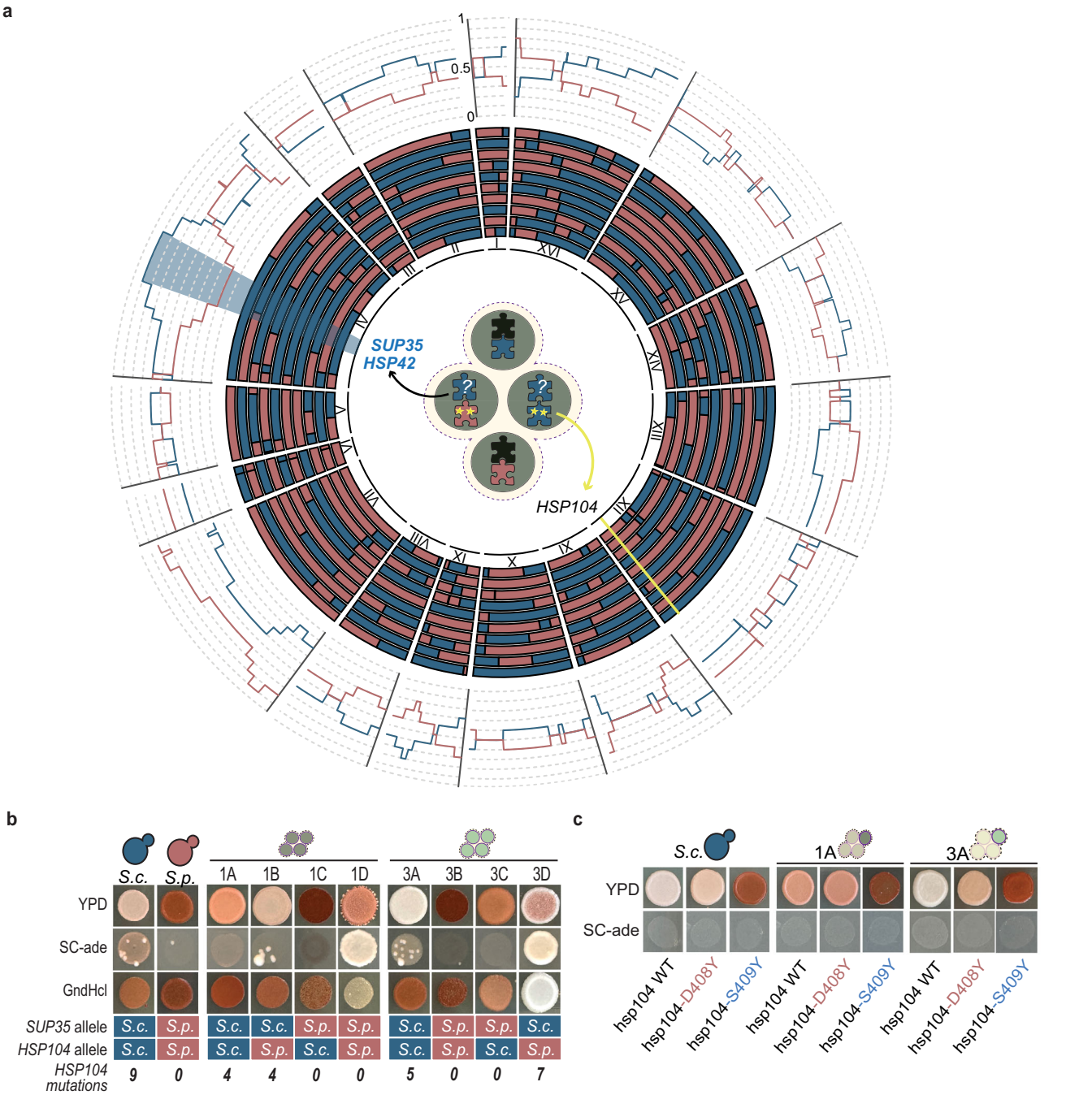


Fig. 6 | An *S. cerevisiae* chromosome IV block containing *SUP35* is associated with recurrent *HSP104* mutations in hybrid segregants. a Circos maps of segregants that gave rise to *HSP104* mutations in evolved populations (~50% of segregants). Hybrids whose populations acquired *HSP104* mutations consistently retain an *S. cerevisiae* (blue, *S. c.*) segment on chromosome IV; the only exception (5 A) carries the corresponding *S. paradoxus* (pink, S. p.) block. This interval spans *HSP42* and *SUP35*, genes functionally connected to Hsp104-mediated proteostasis and prion propagation. **b** Prion phenotyping of parents and segregants on YPD, SC−Ade and 5 mM guanidine-HCl, using the *ade1*-W244* translational readthrough reporter (non-prion colonies red; prion colonies pink/white). The *S. cerevisiae* parent exhibits a weak [PSI +] state, whereas the *S. paradoxus* parent is [psi −]. Segregants from tetrads 1 and 3 span parental-like and intermediate prion states, indicating differences in prion strength across hybrid backgrounds. The table below summarizes allele origin and mutational load: each square indicates the allele of origin for *SUP35* and *HSP104*, and the numbers indicate how many *HSP104* mutations arose in evolved populations founded from each segregant. Prion phenotyping was performed with four biological replicates per genotype. **c** Functional reconstructions of evolved Hsp104 variants. Two recurrent mutations—Hsp104-D408Y (c.1222 G > T; p.Asp408Tyr), which arose in a population founded from segregant 1B carrying the S.p. *HSP104* allele (pink), and Hsp104-S409Y (c.1226 C > A; p.Ser409Tyr), which arose in a 3A-founded population carrying the *S. c. HSP104* allele (blue)—were engineered into the *S. c. HSP104* allele in the *S. cerevisiae* parent and in segregants 1 A and 3 A. Prion phenotypes of reconstruction strains were assayed as in b; two independent transformants were tested for each genotype. Source data are provided as a Source Data file.

evolution. Overall, however, we find that such cases are rare. Among two-subunit complexes, only three—*RSP5–BUL1* (ubiquitin ligase), *YKU70–YKU80* (telomere end-binding), and *DNF3–CRF1* (P4-ATPase)—show excess mutational targeting when assembled in hybrid form, indicating potential selective pressures acting on these chimeric protein complexes. Notably, *YKU70–YKU80* has previously been shown to evolve rapidly and to generate incompatibilities even between *S. paradoxus* populations[45,46], reinforcing the view that divergence in these complexes can promote hybrid dysfunction. Three of the five mutations that arose during the evolution experiment occur in regions of *YKU70* or *YKU80* that are known to play a role in dimerization (Supplementary Fig. 10)[47].

Overall, our findings demonstrate how hybridization can both constrain and diversify adaptation; weak, distributed incompatibilities depress baseline fitness and subtly bias which genes and complexes are repeatedly targeted, while differences in genome composition create distinct adaptive landscapes across segregants. As a result, recombinant hybrid genomes vary not only in phenotype but also in the tempo, repeatability, and genetic basis of adaptation.

Methods

Strain construction

All strains derive from prototrophic and heterothallic versions of *S. cerevisiae* S288C/DBY12000 and *S. paradoxus* CBS432. The *S. paradoxus* CBS432 strain was obtained from the Phaff Yeast Culture Collection as yGIL2026. The *HO* locus was deleted by replacement with a *KanMX* cassette, which was subsequently excised to generate the marker-free heterothallic derivatives yGIL2092 and yGIL2093. In *S. cerevisiae*, the native promoters of *SGS1* and *MSH2* were replaced using the p*CLB2*::3xHA::KanMX cassettes from strain yGIL1791 from ref.[26], generating the parental strain yGIL2099 (*MAT*a, *S. cerevisiae*). Because direct transformation of the corresponding constructs into *S. paradoxus* was unsuccessful, the modified alleles were introduced into the yGIL2092 background by four successive backcrosses, generating yGIL2279 (*MAT*a, *S. paradoxus*). The diploid F1 hybrid yGIL2283 was generated by crossing these parents. Viable spores from yGIL2283 were isolated by tetrad dissection and archived. Strains were sporulated in SPO++ (1.5% potassium acetate with supplements) for ~6 days at room temperature, ascus walls were digested with Zymolyase, and tetrads dissected. We performed 184 dissections of yGIL2283 to isolate five four-spore tetrads (20 segregants). Haploid segregants were transformed with a galactose-inducible *HO* construct to generate diploid derivatives through mating-type switching. In hybrid backgrounds, however, this process was less efficient than in *S. cerevisiae*, often requiring prolonged induction and showing variable pheromone responses, consistent with reduced or poorly regulated mating efficiency. In the first round of induction, most segregants yielded the intended diploid derivative; however, several backgrounds instead produced higher-ploidy derivatives. Specifically, 1 C yielded a tetraploid derivative, whereas 2 C, 3 C, 4B and 5B yielded triploid derivatives. In the same round, no diploid derivative was recovered for 2B, 2D or 5D. To avoid delaying the start of the long-term evolution experiment, we used the derivatives available at that time as founders, including the higher-ploidy strains. A second round of *HO* induction was subsequently performed to recover the missing diploid derivatives for sporulation and spore viability assays. This was successful for all backgrounds except 2 C. These later-derived diploids were used only for characterization and were not included as founders of the evolution experiment. Ploidy assignments for each construction were based on a combination of mating-type PCR, complementation, sporulation, and DNA-content analysis by flow cytometry. A complete list of yeast strains used in this study is provided in Supplementary Data 6.

Growth/thermotolerance phenotyping (spot assays)

Overnight cultures were arrayed in 96-well plates and used to prepare four successive 1:32 serial dilutions. Dilutions were prepared on a Biomek Independent Span-8, and 5 μL from each dilution was pinned to YPD agar using the Biomek multichannel head. Plates were incubated 2 days at 30 °C or 37 °C, and 3 days at 16 °C, scanned under fixed settings, and spot intensities were quantified in Fiji/ImageJ with local background subtraction. For each sample, intensities across dilutions were averaged to produce S_{30} and S_{37}; thermotolerance was defined as $R = S_{37}/S_{30}$. Differences between each background and the *S. cerevisiae* parent in Fig. 2c were assessed using two-sided Welch's two-sample t-tests on replicate *R* values, with Benjamini–Hochberg correction across all displayed comparisons. The same assay was performed for evolved populations at generations 0, 5,000 and 9,000. Within each hybrid background, thermotolerance changes were assessed using paired, two-sided t-tests comparing lineage-matched *R* values at generations 5,000 or 9,000 with generation 0. Strong-improver lineages were defined as those with $R_{\mathrm{equal},9000} \geq \mu_0 + 3\sigma_0$, where μ_0 and σ_0 are the mean and standard deviation of ancestral R across lineages for that segregant founder.

Growth curve analysis

Yeast growth in YPD was monitored at 17 °C, 30 °C, and 37 °C using a Tecan Infinite 200 Pro plate reader in Corning 96-well flat-bottom transparent polystyrene plates (353072). OD_{600} was recorded at fixed intervals with orbital shaking (10 s, 1 mm amplitude) before each read and multiple reads per well (2 × 2 grid). Media-only wells were used for background subtraction. Background-subtracted curves were fit to a Gompertz growth model using scipy.optimize.curve_fit to estimate key parameters, including maximum specific growth rate (μ_max), carrying capacity (A), and lag time (λ).

Long-term experimental evolution

We founded 384 lineages (4 × 96-well plates) from saturated YPD cultures with genotypes randomized by position. The intended design was eight replicate populations per background, including the 20 haploid segregants, the two parental strains, the heterozygous diploid F1 hybrid, and the diploid or higher-ploidy derivatives available at the start of the experiment. In a small number of cases, additional populations were unintentionally seeded and retained. Each day at the same time, cultures were serially diluted 1:32 twice into fresh YPD supplemented with ampicillin and tetracycline, following the same dilution and passaging regimen described in Martínez et al. (2022), yielding ~10 generations per day and propagating populations for up to 10,000 generations.

Competitive fitness assays

Fitness was measured by head-to-head competition against a fluorescent reference: yGIL519 (*MAT***a**), yGIL699 (*MAT*α), and yGIL702 (*MAT***a**/α). Experimental and reference cultures were grown separately to saturation in YPD medium and mixed at an approximately 1:1 initial ratio. Competitions were propagated at 30 °C in 96-well plates under the same serial-transfer regime used for experimental evolution. Every 24 h, cultures were diluted 1:1,024 into fresh YPD medium, corresponding to approximately 10 generations of growth per transfer. Samples were collected at the initial time point and every 10 generations for a total of 50 generations. For flow cytometry (BD FACSCanto II), 4 μL of culture was diluted into 60 μL PBS. A minimum of 30,000 events were recorded per sample. The selection coefficient was estimated as the slope of the linear regression of the natural logarithm of this ratio against elapsed generations and is reported per 10 generations[48]. Generation-0 fitness estimates were obtained by fitting grouped regressions across independent biological competition assays for each background; exact sample sizes are provided in the Source Data file. Differences between each background and the *S.*

cerevisiae parent were assessed using two-sided Z-tests comparing grouped selection-coefficient estimates and their regression standard errors, with Bonferroni correction across the comparisons displayed in each panel. For the comparison between haploid backgrounds and their corresponding diploid or higher-ploidy derivatives, only backgrounds with valid fitness estimates in both ploidy states were included. Pearson correlation was calculated across 19 paired backgrounds; 2B, 2D and 5D were excluded because fitness data were not available for the corresponding derived founders used in Fig. 4a. Plots were generated with ggplot2 3.5.2.

Synteny analysis and gene size variation

Synteny between *S. cerevisiae* S288C and *S. paradoxus* CBS432 was inferred with SynChro[49] using protein sequences; input FASTA/GFF files were pre-formatted with a custom Perl script (courtesy of L. Morales and I. Sedeño, LIIGH). From GFF annotations, we extracted CDS lengths for the ~5,800 well-annotated *S. cerevisiae* genes and from the *S. paradoxus* reference. We identified 5,366 one-to-one orthologs, 434 *S. cerevisiae*-only, and 112 *S. paradoxus*-only genes. For orthologs, absolute CDS length differences were computed and standardized (Z-scores across genes); outliers were defined as $|Z| > 2$ (≥ 2 SD from the mean).

DNA extraction, library preparation, and sequencing

From each evolved population, we isolated a single colony; genomic DNA from this clone was used for sequencing. Genomic DNA was extracted by phenol–chloroform, ethanol-precipitated, and treated with RNase A. Multiplexed libraries were prepared in 96-well format using a modified Nextera protocol. Briefly, 10 ng of genomic DNA per sample was fragmented at 55 °C for 10 min, and libraries were amplified using dual N5/N7 index primers and KAPA polymerase. PCR products were purified using AMPure XP beads. Libraries were quantified, pooled at equal concentration and size-selected for fragments between 500 and 800 bp by 2% agarose-gel electrophoresis and purification using the QIAquick Gel Extraction Kit (Qiagen). Fragment-size distributions were assessed using a BioAnalyzer 2100 with a High Sensitivity DNA chip (Agilent)[48]. Sequencing was performed by the Lewis-Sigler Institute for Integrative Genomics (Princeton University) core facility. Sequencing was done in three batches: (i) parental recombinant-hybrid progeny and their parental strains; (ii) diploid and higher-ploidy evolved populations at generation 2,400; and (iii) haploid evolved populations at generation 5,000.

Sequencing analysis pipeline

Whole-genome sequencing data were processed using a pipeline adapted from ref.[50]. with minor adaptations for hybrid genomes. Illumina 150 bp paired-end reads were trimmed with Trimmomatic v0.36 and aligned with BWA-MEM v0.7.15 to a concatenated reference comprising the *S. cerevisiae* S288C and *S. paradoxus* CBS432 nuclear genomes plus both mitochondrial genomes (downloaded May 2022 from the Yeast PacBio resource: [https://yjx1217.github.io/Yeast_PacBio_2016/data]. Alignments were sorted, indexed and processed with SAMtools v1.4, and PCR duplicates were removed before calculating per-site sequencing depth. Variants were filtered to remove low-quality calls, variants shared among multiple ancestral or evolved samples, and calls located in low-complexity regions or near ambiguous recombination breakpoints. Shared background variants were identified using the vcf-isec function in VCFtools v0.1.15, and retained variants were functionally annotated with SnpEff v5.0. Variants with an allele frequency ≥0.9 and an associated p value < 0.001 were classified as homozygous. For ancestry inference, gene-level coverage was calculated for each genomic region listed in the positions file using samtools coverage -H, iterating over all regions for each BAM file. Gene-level coverage values were normalized independently for the *S. cerevisiae* and *S. paradoxus* parental references. For each gene and recombinant segregant, a log-ratio statistic comparing the relative support for the two parental alleles was calculated and converted into a binary genotype call, with positive values assigned as 1 (*S. cerevisiae*) and negative values assigned as 0 (*S. paradoxus*) (Source Data). Regions spanning recombination breakpoints that remained ambiguous were assigned to the species with higher mapped read support, assuming 2:2 segregation across contiguous gene blocks. Complete analysis scripts and software parameters are available in the associated GitHub repository.

Mitochondrial parental support was evaluated by mapping reads to the *S. cerevisiae* and *S. paradoxus* mitochondrial genomes and comparing sequencing coverage at three annotated loci, *ATP6*, *COX1* and *COX2*[51]. Locus assignments were based on the parental reference receiving the greater relative mapping support. Because some loci showed intermediate or variable support, potentially reflecting differences in sequence divergence, gene structure or cross-mapping between the parental references, *ATP6* was used as the representative mitochondrial marker in the main figure, as it showed the clearest and most consistent assignment among spores from the same tetrad. Coverage-based results for all examined loci are provided in the Source Data.

Hybrid complex composition and mutation enrichment analysis

We constructed a complex–gene map by integrating annotations from SGD, Complex Portal and published compendia[27,36,52] mapping systematic gene names and their associated protein complexes and allowing individual genes to participate in multiple complexes (Supplementary Data 2). Of 5366 one-to-one orthologs, 2121 genes mapped to 652 obligate complexes and 3,245 genes were classified as non-complex.

For each segregant, we read allele calls for all subunits and labeled the complex hybrid if it contained at least one *S. cerevisiae* and at least one *S. paradoxus* allele; otherwise, non-hybrid. This yielded a complex-by-strain hybrid-status matrix (Supplementary Data 2). To contextualize the observed hybridization patterns, we generated null expectations under random inheritance, assuming independent per-gene segregation, and compared the observed and expected distributions of hybrid-background frequency, defined as the number of segregants in which a complex is classified as hybrid. Analyses focused on two-subunit complexes ($n = 250$). To minimize the effects of physical linkage, we repeated the analysis using only unlinked complexes, defined as complexes whose subunit-encoding genes are located on different chromosomes ($n = 230$). We define essential complexes as those in which either protein is classified as essential by the *Saccharomyces* Genome Database (yeastgenome.org).

For mutation enrichment, we compared the observed number of missense mutations in complexes classified as hybrid versus uniparental with expectations based on the number of evolved populations carrying each complex state. Effects were summarized as mutation excess, calculated as the observed minus expected number of missense mutations in hybrid-complex populations, together with the corresponding −log10-transformed p values.

Identification of common targets

Candidate targets of selection were identified using a gene-length-weighted Poisson model adapted from ref.[50]. Briefly, each annotated *S. cerevisiae* coding sequence (approximately 5,800 genes) was tested for enrichment of nonsynonymous mutations, including missense, nonsense and frameshift variants, and in a separate analysis for enrichment of missense mutations only. For each gene, the expected number of mutations was calculated from the total number of mutations in the corresponding dataset, scaled by the gene's contribution to the total coding-sequence length analyzed. To approximate mutational opportunity in the hybrid genome, gene length was defined as the mean CDS length of the *S. cerevisiae*–*S. paradoxus* ortholog pair. *S.*

paradoxus-specific ORFs lacking an annotated *S. cerevisiae* ortholog were excluded. The probability of observing the recorded number of mutations in each gene was calculated using the upper tail of the Poisson distribution. Genes with adjusted p value $< 10^{-4}$ and ≥10 nonsynonymous or missense mutations were classified as common targets. To distinguish hybrid-specific from general targets, we applied the same analysis to our evolved parental populations and to published *S. cerevisiae* evolution datasets[37,39,40] and compared gene-wise enrichment across datasets (Supplementary Data 4).

Gene-level dN/dS analysis

To test whether candidate genes showed an excess of protein-altering substitutions, we performed a gene-level dN/dS-style analysis on recurrently mutated genes from our experimental evolution dataset. Mutations were classified based on their annotated coding consequence. Because mutations occurred in both the *S. cerevisiae* and *S. paradoxus* subgenomes, orthologs were analyzed separately by species whenever possible. For each gene, we extracted the coding sequence from the hybrid reference genome and enumerated all possible single-nucleotide substitutions, classifying each as synonymous or nonsynonymous. This allowed us to calculate the number of synonymous sites (Ssites) and nonsynonymous sites (Nsites), representing the mutational opportunities for each class. Observed synonymous and nonsynonymous mutations were counted as Sobs and Nobs. Raw dN/dS was calculated by normalizing observed mutation counts by the corresponding number of available sites. Because some genes had no synonymous mutations, we also calculated a pseudo-corrected dN/dS using a 0.5 continuity correction (Supplementary Data 5). To test for nonsynonymous enrichment, we used a one-sided exact binomial test, with the expected probability of a nonsynonymous mutation defined by the fraction of nonsynonymous sites. We also calculated pooled estimates across both orthologs by summing observed counts and mutational opportunities across the *S. cerevisiae* and *S. paradoxus* copies. As a complementary analysis, we calculated a likelihood-ratio G-score comparing observed synonymous and nonsynonymous counts to expected counts based on Ssites and Nsites.

Sequence alignment of candidate targets of selection

Protein sequences of candidate targets were aligned with MAFFT. Alignments (FASTA) were parsed with a custom Python script (Biopython 1.83) to compute per-gene statistics: percent identity, numbers of matches/mismatches/indels, total alignment length, and effective gene size. For position-wise divergence, alignments were reshaped to a residue-level table classifying each position as a match, mismatch, or gap. Divergence tracks were visualized in R version 4.4.2 (ggplot2 3.5.2) using line/tile glyphs along the protein coordinate.

LOF/non-LOF selection (Bayesian analysis)

We assessed whether selection in hybrids preferentially favored loss-of-function (LOF) versus non-LOF changes using a Bayesian model, which uses priors derived from a meta-analysis of laboratory evolution targets[50]. From our hybrid dataset, we took the 100 most recurrently mutated genes across segregants and classified coding mutations as LOF-consistent (nonsense, frameshift, start-loss, stop-gain) or non-LOF-consistent (missense); synonymous and intergenic variants were excluded.

Variance scores analysis

To identify candidate genes under functional constraint or directional selection, we analyzed the dispersion of nonsynonymous (missense-inclusive) mutation counts across all evolved hybrid populations. For each gene, we computed the mean, variance, and variance-to-mean ratio (VMR = variance/mean). Genes with VMR ≥ 2 were classified descriptively as showing strongly non-uniform targeting across backgrounds. This threshold was used to identify candidate background-dependent targets and was not treated as a formal significance test.

Genome editing with CRISPR–Cas9

Targeted edits were introduced into parental and recombinant hybrid strains using pGIL218, a derivative of pML104-NatMX that constitutively expresses Cas9 and contains a guide-RNA expression cassette with BclI and SwaI cloning sites together with a NatMX selectable marker. Species-specific single-guide RNAs were designed to target *AMN1*, *ADE1* and *HSP104* and were cloned into pGIL218. The resulting plasmids were co-transformed with 500-bp synthetic donor DNA fragments carrying the desired substitutions. Donor templates also contained synonymous substitutions within the PAM to prevent Cas9 re-cutting. Transformants were selected on nourseothricin-containing medium, and the edited loci were verified by Sanger sequencing. For *HSP104* reconstruction, edited clones were recovered only in the *S. cerevisiae* background; two guides designed for the *S. paradoxus* allele did not yield verified edited clones. Oligonucleotide sequences, guide sequences and plasmid information are provided in Supplementary Data 6. Newly generated strains and plasmids are available from the corresponding author upon reasonable request.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The short-read sequencing data reported in this study have been deposited to the NCBI SRA BioProject database. The parental genomes and the hybrid at generation 0 are in accession number PRJNA1330885. Sequencing of generation 2,400 of the diploid and higher-ploidy evolved populations is in accession number PRJNA1331785. Sequencing of generation 5,000 of the haploid evolved populations is in accession number PRJNA1331855. The data underlying the figures in this study are provided in the Source Data file associated with the article. Sequencing data have been deposited in the appropriate public repositories, with accession numbers provided in the manuscript. Source Data is provided as a Source Data file. Source data are provided with this paper.

Code availability

Custom scripts used for data analysis and figure generation are available at https://github.com/ArtemizaMartinez/yeast-hybrid-genome-evolution and are archived in Zenodo at https://doi.org/10.5281/zenodo.21842209.

References

1. Maheshwari, S. & Barbash, D. A. The genetics of hybrid incompatibilities. *Annu. Rev. Genet.* **45**, 331–355 (2011).
2. Presgraves, D. C. The molecular evolutionary basis of species formation. *Nat. Rev. Genet.* **11**, 175–180 (2010).
3. Abbott, R. et al. Hybridization and speciation. *J. Evolut. Biol.* **26**, 229–246 (2013).
4. Runemark, A., Vallejo-Marin, M. & Meier, J. I. Eukaryote hybrid genomes. *PLoS Genet.* **15**, e1008404 (2019).
5. Servedio, M. R., Hermisson, J. & van Doorn, G. S. Hybridization may rarely promote speciation. *J. Evolut. Biol.* **26**, 282–285 (2013).
6. Jerison, E. R. et al. Genetic variation in adaptability and pleiotropy in budding yeast. *Elife* **6**, e27167 (2017).
7. Bozdag, G. O. & Ono, J. Evolution and molecular bases of reproductive isolation. *Curr. Opin. Genet. Dev.* **76**, 101952 (2022).
8. Maclean, C. J. & Greig, D. Prezygotic reproductive isolation between Saccharomyces cerevisiae and Saccharomyces paradoxus. *BMC Evol. Biol.* **8**, 1 (2008).

9. Langdon, Q. K. et al. Fermentation innovation through complex hybridization of wild and domesticated yeasts. *Nat. Ecol. Evol.* **3**, 1576–1586 (2019).
10. Molinet, J. & Stelkens, R. The evolution of thermal performance curves in response to rising temperatures across the model genus yeast. *Proc. Natl. Acad. Sci. USA* **122**, e2423262122 (2025).
11. Peris, D. et al. Hybridization and adaptive evolution of diverse Saccharomyces species for cellulosic biofuel production. *Biotechnol. Biofuels* **10**, 78 (2017).
12. Piatkowska, E. M. et al. Chimeric protein complexes in hybrid species generate novel phenotypes. *PLoS Genet.* **9**, e1003836 (2013).
13. D'Angiolo, M. et al. A yeast living ancestor reveals the origin of genomic introgressions. *Nature* **587**, 420–425 (2020).
14. Heil, C. S. S. et al. Loss of heterozygosity drives adaptation in hybrid yeast. *Mol. Biol. Evol.* **34**, 1596 (2017).
15. Leducq, J.-B. et al. Speciation driven by hybridization and chromosomal plasticity in a wild yeast. *Nat. Microbiol.* **1**, 1–10 (2016).
16. Morales, L. & Dujon, B. Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. *Microbiol. Mol. Biol. Rev.* **76**, 721–739 (2012).
17. Chou, J.-Y. et al. Multiple molecular mechanisms cause reproductive isolation between three yeast species. *PLoS Biol.* **8**, e1000432 (2010).
18. Greig, D. A Screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. *PLOS Genet.* **3**, e21 (2007).
19. Greig, D. et al. Hybrid speciation in experimental populations of yeast. *Science* **298**, 1773–1775 (2002).
20. Hou, J. et al. Comprehensive survey of condition-specific reproductive isolation reveals genetic incompatibility in yeast. *Nat. Commun.* **6**, 7214 (2015).
21. Kao, K. C., Schwartz, K. & Sherlock, G. A genome-wide analysis reveals no nuclear Dobzhansky-Muller pairs of determinants of speciation between S. cerevisiae and S. paradoxus, but suggests more complex incompatibilities. *PLoS Genet.* **6**, e1001038 (2010).
22. Lee, H.-Y. et al. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. *Cell* **135**, 1065–1073 (2008).
23. Delneri, D. et al. Engineering evolution to study speciation in yeasts. *Nature* **422**, 68–72 (2003).
24. Fischer, G. et al. Highly variable rates of genome rearrangements between Hemiascomycetous yeast lineages. *PLOS Genet.* **2**, e32 (2006).
25. Hunter, N. et al. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* **15**, 1726–1733 (1996).
26. Bozdag, G. O. et al. Breaking a species barrier by enabling hybrid recombination. *Curr. Biol.* **31**, R180–R181 (2021).
27. Swamy, K. B., Schuyler, S. C. & Leu, J.-Y. Protein complexes form a basis for complex hybrid incompatibility. *Front. Genet.* **12**, 609766 (2021).
28. Tang, S. & Presgraves, D. C. Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. *Science* **323**, 779–782 (2009).
29. Tirosh, I. et al. A Yeast hybrid provides insight into the evolution of gene expression regulation. *Science* **324**, 659–662 (2009).
30. Kellis, M. et al. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
31. Liti, G. et al. Population genomics of domestic and wild yeasts. *Nature* **458**, 337–341 (2009).
32. Brion, C. et al. Variation of the meiotic recombination landscape and properties over a broad evolutionary distance in yeasts. *PLoS Genet.* **13**, e1006917 (2017).
33. Fang, O. et al. Amn1 governs post-mitotic cell separation in Saccharomyces cerevisiae. *PLoS Genet.* **14**, e1007691 (2018).
34. Gonçalves, P. et al. Evidence for divergent evolution of growth temperature preference in sympatric Saccharomyces species. *PloS one* **6**, e20739 (2011).
35. Dandage, R. et al. Frequent assembly of chimeric complexes in the protein interaction network of an interspecies yeast hybrid. *Mol. Biol. Evol.* **38**, 1384–1401 (2021).
36. Swamy, K. B. et al. Proteotoxicity caused by perturbed protein complexes underlies hybrid incompatibility in yeast. *Nat. Commun.* **13**, 4394 (2022).
37. Fisher, K. J. et al. Adaptive genome duplication affects patterns of molecular evolution in Saccharomyces cerevisiae. *PLoS Genet.* **14**, e1007396 (2018).
38. Harari, Y. et al. Spontaneous changes in ploidy are common in yeast. *Curr. Biol.* **28**, 825–835. e4 (2018).
39. Johnson, M. S. et al. Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations. *Elife* **10**, e63910 (2021).
40. Marad, D. A., Buskirk, S. W. & Lang, G. I. Altered access to beneficial mutations slows adaptation and biases fixed mutations in diploids. *Nat. Ecol. Evol.* **2**, 882–889 (2018).
41. Li, C., Wang, Z. & Zhang, J. Toward genome-wide identification of Bateson-Dobzhansky-Muller incompatibilities in yeast: a simulation study. *Genome Biol. Evol.* **5**, 1261–1272 (2013).
42. Xu, M. & He, X. Genetic incompatibility dampens hybrid fertility more than hybrid viability: yeast as a case study. *PLoS One* **6**, e18341 (2011).
43. Jhuang, H. Y., Lee, H. Y. & Leu, J. Y. Mitochondrial-nuclear co-evolution leads to hybrid incompatibility through pentatricopeptide repeat proteins. *EMBO Rep.* **18**, 87–101 (2017).
44. Kowalec, P. et al. Newly identified protein Imi1 affects mitochondrial integrity and glutathione homeostasis in Saccharomyces cerevisiae. *FEMS Yeast Res.* **15**, fov048 (2015).
45. Liti, G. et al. Segregating YKU80 and TLC1 alleles underlying natural variation in telomere properties in wild yeast. *PLoS Genet.* **5**, e1000659 (2009).
46. Sawyer, S. L. & Malik, H. S. Positive selection of yeast nonhomologous end-joining genes and a retrotransposon conflict hypothesis. *Proc. Natl. Acad. Sci. USA* **103**, 17614–17619 (2006).
47. Chen, H. et al. Structural insights into yeast telomerase recruitment to telomeres. *Cell* **172**, 331–343.e13 (2018).
48. Buskirk, S. W., Peace, R. E. & Lang, G. I. Hitchhiking and epistasis give rise to cohort dynamics in adapting populations. *Proc. Natl. Acad. Sci. USA* **114**, 8330–8335 (2017).
49. Drillon, G., Carbone, A. & Fischer, G. SynChro: a fast and easy tool to reconstruct and visualize synteny blocks along eukaryotic chromosomes. *PLoS ONE* **9**, e92621 (2014).
50. Martínez, A. A. & Lang, G. I. Identifying targets of selection in laboratory evolution experiments. *J. Mol. Evol.* **91**, 345–355 (2023).
51. Yue, J. X. et al. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* **49**, 913–924 (2017).
52. Laurent, J. M. et al. Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. *PLoS Biol.* **18**, e3000627 (2020).

Acknowledgements

We thank Robert Skibbens and Maitreya Dunham for helpful discussions and Gianni Liti, Amber Rice, Michael Desai, and members of the Lang Lab for lab comments on the manuscript.

Author contributions

Conceptualization: A.A.M. and G.I.L.; methodology: A.A.M. and G.I.L.; investigation: A.A.M.; data curation: A.A.M.; formal analysis: A.A.M.; visualization: A.A.M.; funding acquisition: G.I.L.; supervision: G.I.L.;

project administration: G.I.L.; writing—original draft: A.A.M. and G.I.L.; writing—review and editing: A.A.M. and G.I.L.

Funding

This study was supported by a grant from the National Institutes of Health (R35GM149540). Portions of this research were conducted on Lehigh University's Research Computing infrastructure partially supported by the National Science Foundation (Award 2019035).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-77469-1.

Correspondence and requests for materials should be addressed to Gregory I. Lang.

Peer review information *Nature Communications* thanks Duncan Greig and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026