








Synthetic hybrids of six yeast species

David Peris ^{1,2,3}✉, William G. Alexander ^{1,2,4}, Kaitlin J. Fisher ¹, Ryan V. Moriarty^{1,2}, Mira G. Basuino ⁴, Emily J. Ubbelohde ⁴, Russell L. Wrobel ^{1,2} & Chris Todd Hittinger ^{1,2}✉

Allopolyploidy generates diversity by increasing the number of copies and sources of chromosomes. Many of the best-known evolutionary radiations, crops, and industrial organisms are ancient or recent allopolyploids. Allopolyploidy promotes differentiation and facilitates adaptation to new environments, but the tools to test its limits are lacking. Here we develop an iterative method of Hybrid Production (iHyPr) to combine the genomes of multiple budding yeast species, generating *Saccharomyces* allopolyploids of at least six species. When making synthetic hybrids, chromosomal instability and cell size increase dramatically as additional copies of the genome are added. The six-species hybrids initially grow slowly, but they rapidly regain fitness and adapt, even as they retain traits from multiple species. These new synthetic yeast hybrids and the iHyPr method have potential applications for the study of polyploidy, genome stability, chromosome segregation, and bioenergy.

¹Laboratory of Genetics, J. F. Crow Institute for the Study of Evolution, Wisconsin Energy Institute, Center for Genomic Science Innovation, University of Wisconsin-Madison, Madison, WI, USA. ²DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, USA. ³Department of Food Biotechnology, Institute of Agrochemistry and Food Technology (IATA), CSIC, 46980 Paterna, Valencia, Spain. ⁴Department of Biology, Truman State University, Kirksville, MO 63501, USA. ✉email: david.perisnavarro@gmail.com; cthittinger@wisc.edu

Polyploidy generates diversity by increasing the number of copies of each chromosome¹. Allopolyploidy instantly adds chromosomal variation from multiple species through hybridization, while autopolyploidy leads to variation as gene copies from a single species diverge during evolution. Allopolyploidy facilitates differentiation and adaptation to new environments². Many plants, animals, fungi, and other eukaryotes are ancient or recent allopolyploids, including some of the best-known industrial organisms, crops, and evolutionary radiations^{3,4}.

Phylogenomic analyses support an ancient allopolyploid origin for the baker's yeast *Saccharomyces cerevisiae*⁵. *S. cerevisiae* has been one of the most important model organisms to study polyploidy in the context of evolution⁶, its effects on mutation rate⁷, and as a model of how cancer progresses as clonal populations adapt through driver mutations⁸. Despite the decreased fitness of newly generated polyploids⁹, experimental evolution assays in *S. cerevisiae* and comparisons of the genomes of industrial *Saccharomyces* interspecies hybrids have shown that they return to high fitness through the generation of aneuploidies, chromosomal rearrangements, and loss-of-heterozygosity^{10,11}.

Genome rearrangements are common in *Saccharomyces* allopolyploids used to make fermented beverages^{12,13}, but experimental tools to test the limits of polyploidy and genome rearrangements are lacking. Random chromosomal aberrations can be easily generated by using techniques from synthetic biology, such as SCRaMbLE¹⁴. However, SCRaMbLE is currently only available in single, partly synthetic *S. cerevisiae* strain, limiting the genomic diversity that can be explored.

Saccharomyces species have similar genome content, identical numbers of chromosomes ($n = 16$), and genomes that are mostly syntenic¹⁵. Since they have limited pre-zygotic barriers, interspecies hybrids can be generated easily when haploid strains of opposite mating types encounter each other. Much more rarely, diploid yeast cells can become competent to mate by inactivating or losing one *MAT* idiomorph or undergoing gene conversion at the *MAT* locus¹⁶. To attempt the generation of allododecaploid (base ploidy of $12n$) hybrids of six species, here we develop an iterative Hybrid Production (iHyPr) method. iHyPr combines traits from multiple species, such as temperature tolerance, and through adaptive laboratory evolution (ALE), facilitates rapid adaptation to new environments. This method enables basic research on polyploidy and chromosome biology. iHyPr is applicable to research on bioenergy and synthetic biology where it can harness genomic diversity to generate more efficient strains that produce new bioproducts¹⁷ or to combine industrially useful traits from multiple species^{18,19}.

Results

Hybrids of six yeast species can be generated with iHyPr. iHyPr allowed us to experimentally test the limits of chromosome biology and allopolyploidy by constructing a series of higher-order interspecies hybrids (Supplementary Fig. 1). First, we used two differentially marked HyPr plasmids, which each encode a drug-inducible *HO* gene (homothallic switching endonuclease) that promotes mating-type switching, to efficiently generate and select for two-species hybrids as done previously²⁰. Next, using two newly created, differentially marked HyPr plasmids, we crossed these two-species hybrids to construct three-species and four-species hybrids. The construction of higher-order synthetic hybrids has not been reported previously. Finally, we constructed six-species hybrids using three different crossing schemes (Fig. 1, Supplementary Fig. 2).

In all three schemes, diploid genomes were successfully introduced from each of the six parent species (Fig. 2b, Supplementary Fig. 3).

During hybrid construction, as more and more genomes were introduced, the frequency of successful matings decreased (Spearman rank-sum test $R = -0.89$, p value = 1.1×10^{-5} , Fig. 3a, Supplementary Data 3), and the fitness of synthetic hybrids declined (Spearman rank-sum test $R = -0.77$, p value = 7.6×10^{-4} , Fig. 3c). The fitness decrease may be due to the increased cell area, which was correlated with the increased genome size (Spearman rank-sum test $R = 0.97$, p value = 1.4×10^{-5} , Fig. 3b). One model proposes that increased ploidy prolongs spindle assembly and results in more frequent spindle checkpoint-associated mitotic arrest²¹. Genetic incompatibilities could also play a role in the decreased fitnesses of interspecies hybrids.

Genome size and stability limits. The largest synthetic hybrid expanded its genome size 3.3 times (from 24 Mb to ~80 Mbp) (Supplementary Fig. 5A, Supplementary Data 2), and its cell area was 2.3 times larger than a diploid cell (Fig. 3b). The increases to cell volume, estimated from cell area, closely matched the increases to genome size. For example, 3.3-fold increase in genome size of the largest synthetic hybrid corresponded to a 3.4-fold increase in cell volume (Supplementary Data 2). These data fit well with the long-held observation that cell size is correlated with genome size within taxa and cell types²², as well as recent work suggesting that the DNA:cytoplasm ratio imposes the primary limitation on cell size due to fitness defects²³.

Some species contributed many fewer chromosomes than others to the synthetic hybrids of six species (defined as the ancestor hybrids) (Fig. 2b). During construction, chromosome losses were widespread and outnumbered gains (two-sided t -test, $t = -3.4408$, d.f. = 6, p value < 1.37×10^{-2} , Supplementary Data 2). These aneuploidies rose dramatically as the number of species donating genomes increased (linear regression $r^2 = 0.79$, p value = 3.01×10^{-6}) (Fig. 2a). Complete chromosomal aneuploidies were much more common than aneuploidies caused by unbalanced translocations or deletions (97.31% versus 2.69% of total detected chromosomal aberrations). Aneuploidies involving chromosome III, where 88.9% of translocations or deletions in this chromosome were unbalanced (Fig. 2b, Supplementary Fig. 3, Supplementary Note 1), were especially common because it contains the *MAT* locus being cut by the *Ho* endonuclease during iHyPr. These results suggest that, in addition to the expected gene conversion events, iHyPr generated mating-competent cells via partial chromosome losses.

Although our goal was to generate true allododecaploid ($12n$) six-species hybrids, a ploidy level acquired by only a few organisms^{24,25}, we obtained at most eight copies of some chromosomes. The maximum ploidy reached was ~ $7n$, and none six-species hybrids were euploid (Fig. 2b). Due to massive chromosome loss, we inferred the six-species hybrid with the largest genome had an average of ~7 copies of each chromosome (i.e. $12n - 88$) when estimated using genome sequencing and bioinformatic tools (visual inspection of *sppIDer* plots) or an average of ~8 copies of each chromosome (i.e. $12n - 64$) when total DNA content was estimated using flow cytometry (Fig. 2b, Supplementary Figs. 3 and 5C, Supplementary Data 2).

Mitochondrial inheritance affects genotype and phenotype.

During interspecific hybridization, hybrids can inherit one of the two parent mitotypes or a recombinant version depending on the budding location²⁶. In general, one of the parent mitotypes was quickly fixed during the generation of our hybrids, except for three cases: the allotetraploid *Saccharomyces kudriavzevii* × *Saccharomyces mikatae* yHRWh4, the allotetraploid *S. cerevisiae* × *Saccharomyces uvarum* yHRWh10, and the six-species hybrid yHRWh36, which were all heteroplasmic

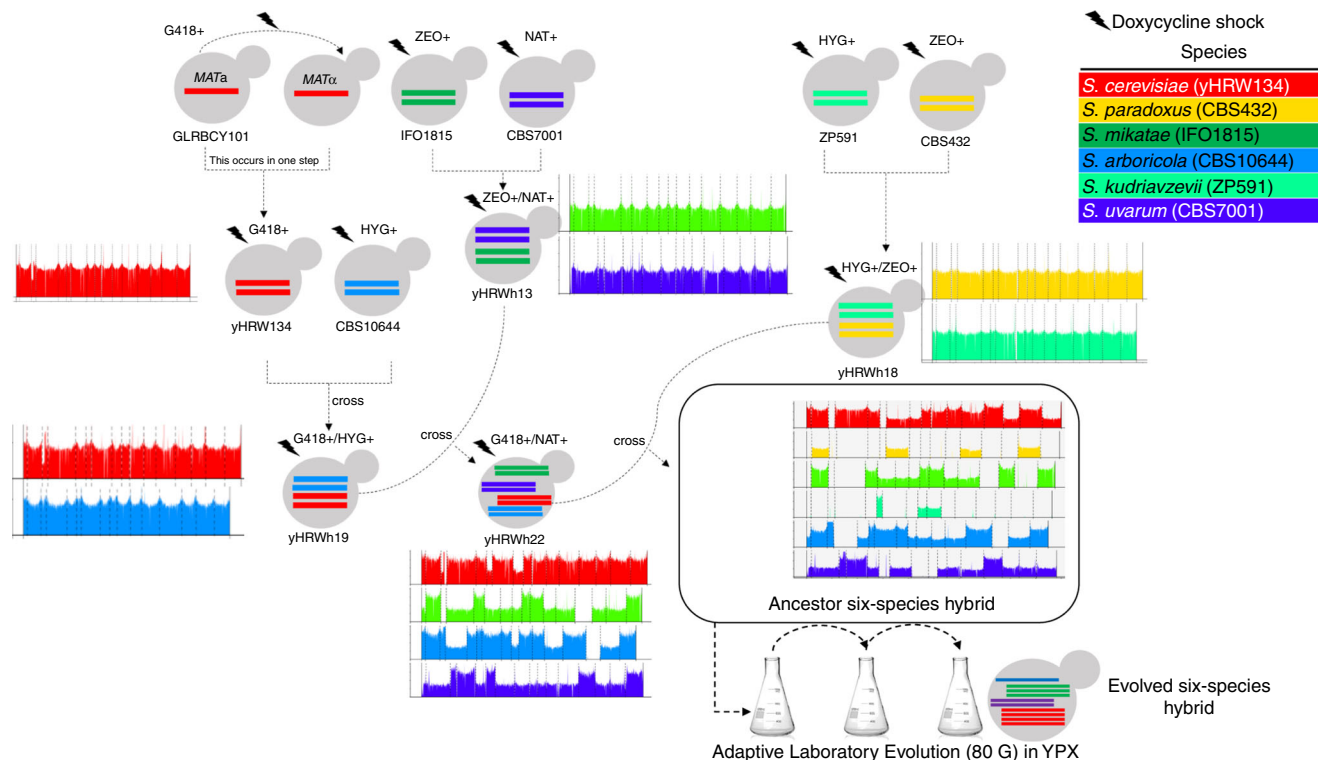


Fig. 1 The generation of ancestor and evolved six-species hybrids. Synthetic hybrid generation scheme using the iHyPr method. The example shown is the six-species hybrid yHRWh39. Chromosomes were colored according to their species designation, with height representing copy number, using the *sppIDer* pipeline⁴⁸. For an extended explanation of iHyPr, including the other two crossing schemes, see Supplementary Figs. 1 and 2. Arrows mark hybridization steps. For additional intermediate and six-species hybrid nuclear and mitochondrial genomes with higher resolution, see Supplementary Figs. 3 and 4. Ancestor six-species hybrids underwent ALE for 80 generations.

(Supplementary Fig. 4). In rich medium at 20 °C, for strains with similar numbers of hybridized species, hybrids with a *S. cerevisiae* mitochondrial genome (mtDNA) grew 7–15% faster than the hybrids with the mtDNA of another species (Fig. 3c, d). mtDNA inheritance was also significantly correlated with nuclear genome retention (ANOVA multifactor F value = 19.9, d.f. = 1, p value = 7.77×10^{-4}), with the mtDNA donor tending to contribute more nuclear chromosomes (Fig. 4, Supplementary Fig. 4). These results are consistent with recent observations in hybrids used in the fermented beverage industry¹².

Trait combination, adaptive evolution, and genome stability. Higher-order synthetic hybrids allow investigators to rapidly combine traits from many different parents, such as differences in sugar consumption and temperature preferences. To determine if the inherent chromosomal instability of these six-species hybrids could be harnessed as a diversity generator, we tested how these new six-species hybrids altered their kinetic parameters during ALE. ALE was performed for an estimated 80 generations in a medium containing glucose (as a control) or xylose, a sugar poorly metabolized by most *Saccharomyces* species²⁷. To provide baseline xylose metabolic capability upon which to improve, we chose a *S. cerevisiae* parent strain that had been engineered by inserting xylose utilization genes into Chromosome IV^{28,29}. Ancestor six-species hybrids grew slowly, and despite differing from each other in chromosomal composition (Figs. 2b and 4a), single-colony isolates of all 12 ALE replicates (three replicates for the two ancestor hybrids retaining the chromosome IV in two ALE conditions) outperformed their ancestors in culturing conditions identical to the ALE (one-sided Wilcoxon rank-sum test, p value = 3.51×10^{-4}). Many evolved strains even outperformed the *S. cerevisiae* reference strain (Fig. 5a), but we detected no

significant differences between hybrids evolved in YPD and YPX (one-sided Wilcoxon rank-sum test, p value > 9.39×10^{-2}). However, in microtiter plate culturing conditions where more replicates could be achieved, evolved hybrid populations grew as much as 71% faster on xylose than the reference *S. cerevisiae* strain, and populations evolved on xylose outperformed those evolved on glucose (one-sided Wilcoxon rank-sum test, p value = 1.29×10^{-2}) (Supplementary Fig. 6, Supplementary Data 6). Importantly, all our evolved hybrids grew well at low-temperature conditions (4 °C) where the *S. cerevisiae* parent could not grow (Fig. 5b), demonstrating that the cold tolerance of the other parents^{30,31} had been retained through hybridization and ALE.

Since maximum growth rate on xylose improved considerably regardless of whether hybrids were evolved on xylose or glucose, we hypothesized several factors that could be responsible, such as xylose cassette amplification or genome stabilization. Neither chromosome IV nor the xylose utilization genes themselves were selectively amplified in either condition (Fig. 2b, Supplementary Fig. 7, Supplementary Data 7). Evolved hybrids with more reduced genome sizes tended to have slightly higher fitness among ALE replicates, but the correlation was not significant (Spearman rank-sum test, p value 0.054) (Supplementary Fig. 8). These results suggest that the regions of the genome that are lost or amplified matter more for adaptation than total size. Although genome instability increased after each step during the construction of the six-species hybrids (Fig. 2a), genome instability decreased after 80 generations of ALE (Fig. 6a). Nonetheless, genome sequencing of a random selection of colonies from one of the evolved six-species hybrids demonstrated that genomic diversity was still being generated at a prodigious rate (Fig. 6b, c). Thus, genome stabilization was ongoing, and future studies are needed to

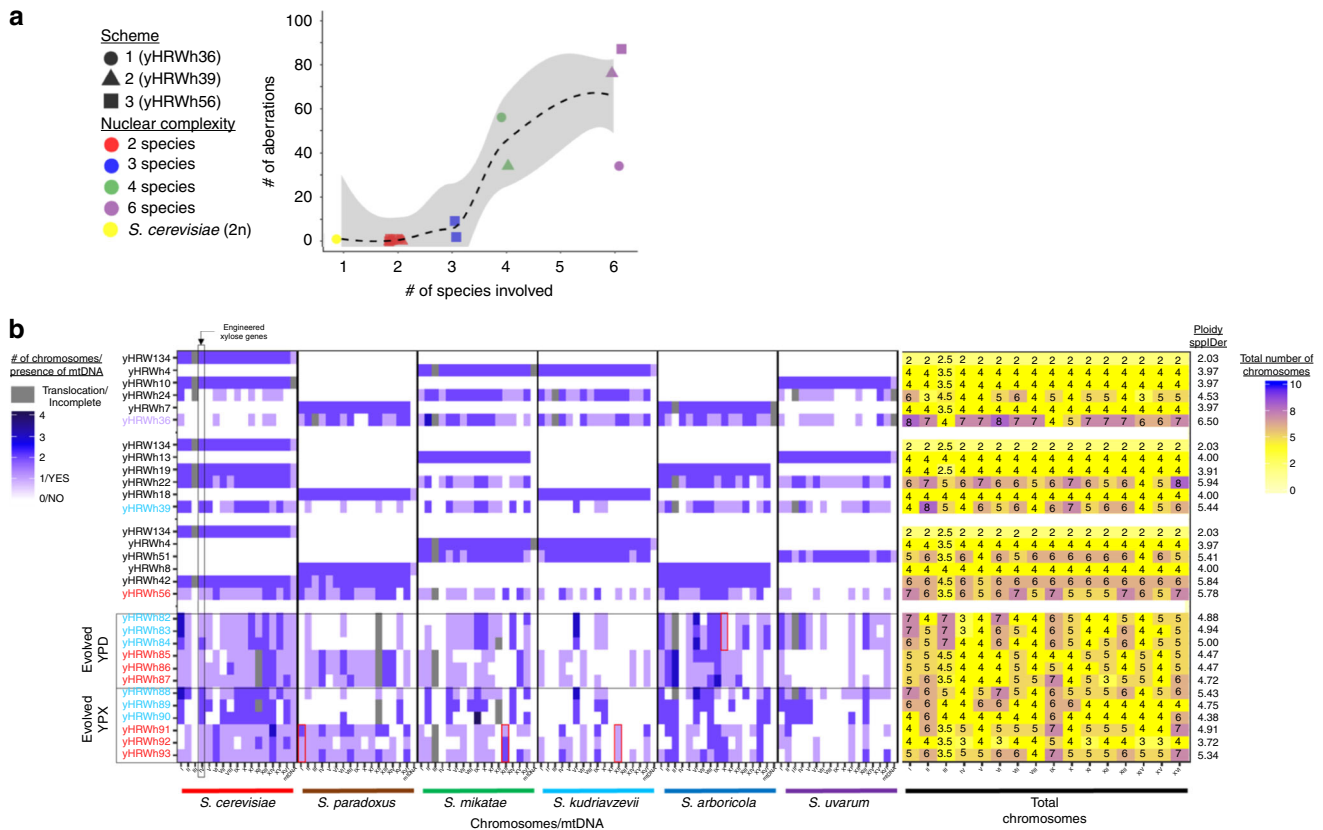


Fig. 2 Genome contributions to synthetic hybrids. The numbers and sources of chromosomes for each synthetic hybrid were inferred from *sppIDer* plots (Supplementary Fig. 3), which were corrected based on flow cytometry ploidy estimations. **a** The number of chromosomal aberrations was inferred for each synthetic hybrid as new translocation, gain, and loss events not seen in the preceding hybrid (Supplementary Fig. 3). Chromosomal aberrations involving parts of chromosomes were conservatively counted only in cases of clear fusion of entire arms, whereas smaller loss-of-heterozygosity events were not counted. The synthetic hybrids generated from each independent scheme are represented with different shapes. Color points are colored according to the number of species genomes contributing to the strain. A LOESS regression line and the 95% confidence interval of the fit are represented with a discontinuous black color and gray shadow, respectively. **b** Chromosome content is colored according to the species donor. Mitochondrial inheritance was inferred using *mitosppIDer* (Supplementary Fig. 4). The numbers of chromosomes for each species are colored according to the left heatmap legend. Incomplete and recombinant mtDNA are colored in gray. Total number of chromosomes is shown in the right part, which is colored according to the right legend. Ploidy estimates based on de novo genome assemblies, which correlates with flow cytometry (Spearman rank-sum test $R = 0.88$, p value = 7.5×10^{-8} , Supplementary Fig. 5c), are indicated at the right side. Synthetic hybrids are reported in the order constructed (Supplementary Fig. 2). yHRW134 and yHRWh4 are shown multiple times because of their use in multiple crossing schemes. Evolved hybrids are grouped based on the conditions in which they were evolved, and they are colored according to their ancestor hybrid. Red squares highlight chromosomes that were retained or lost in all hybrids evolved in the same condition when compared to their siblings evolved in the other condition. *S. cerevisiae* chromosome IV, where the xylose utilization genes were inserted, is indicated by the black square. Considerable karyotypic diversity continued to be generated during 80 generations of ALE (Fig. 6), but each evolved strain is easily recognized as more similar to its ancestor six-species hybrid. Source data are provided in the Source Data file and at <http://bit.ly/2v1rq1T>.

determine whether and when the stabilization of the hybrid genomes would be sufficient to apply to industry.

Discussion

iHyPr is a powerful addition to the yeast synthetic biology toolkit that can be used to generate and select for genome diversity, while combining industrially relevant traits from multiple parents. Here, we used iHyPr to combine xylose utilization from a biofuel strain of *S. cerevisiae* with cold tolerance, a trait critical for the production of many fermented beverages^{32–34}.

Previous efforts to generate higher ploidy *Saccharomyces* cells were arduous. A documented autohexaploid *S. cerevisiae* strain was produced by using a complex combination of auxotrophic intermediates⁹, and allotetraploids of *S. cerevisiae* × *S. kudriavzevii* have been generated by using protoplast fusion and rare-mating³⁵. Recently, a CRISPR/Cas9 system was developed to switch mating types and generate tetraploid yeast cells in a

manner similar to HyPr³⁶. We show here that iHyPr can be used to produce higher-order hybrids iteratively without additional transformations. Our six-species hybrids were unstable and quickly lost chromosomes, a result nicely modeled with polyploid yeasts. As the number of chromosomes increases, the time required for spindle assembly increases, generating chromosomes without attached spindles, which results in chromosome loss during anaphase²¹.

The levels of allopolyploidy reached in our study will facilitate the understanding of the cellular fitness consequences in eukaryotes. The increase in ploidy was associated with a short-term fitness defect. However, ALE rapidly improved fitness, while allowing multiple parent traits, such as cold tolerance and xylose utilization, to be retained. Since both ALE selection regimes improved performance on xylose, most improvements were likely not in direct response to the environment. Instead, they may have been driven by endogenous factors, including possibly genome

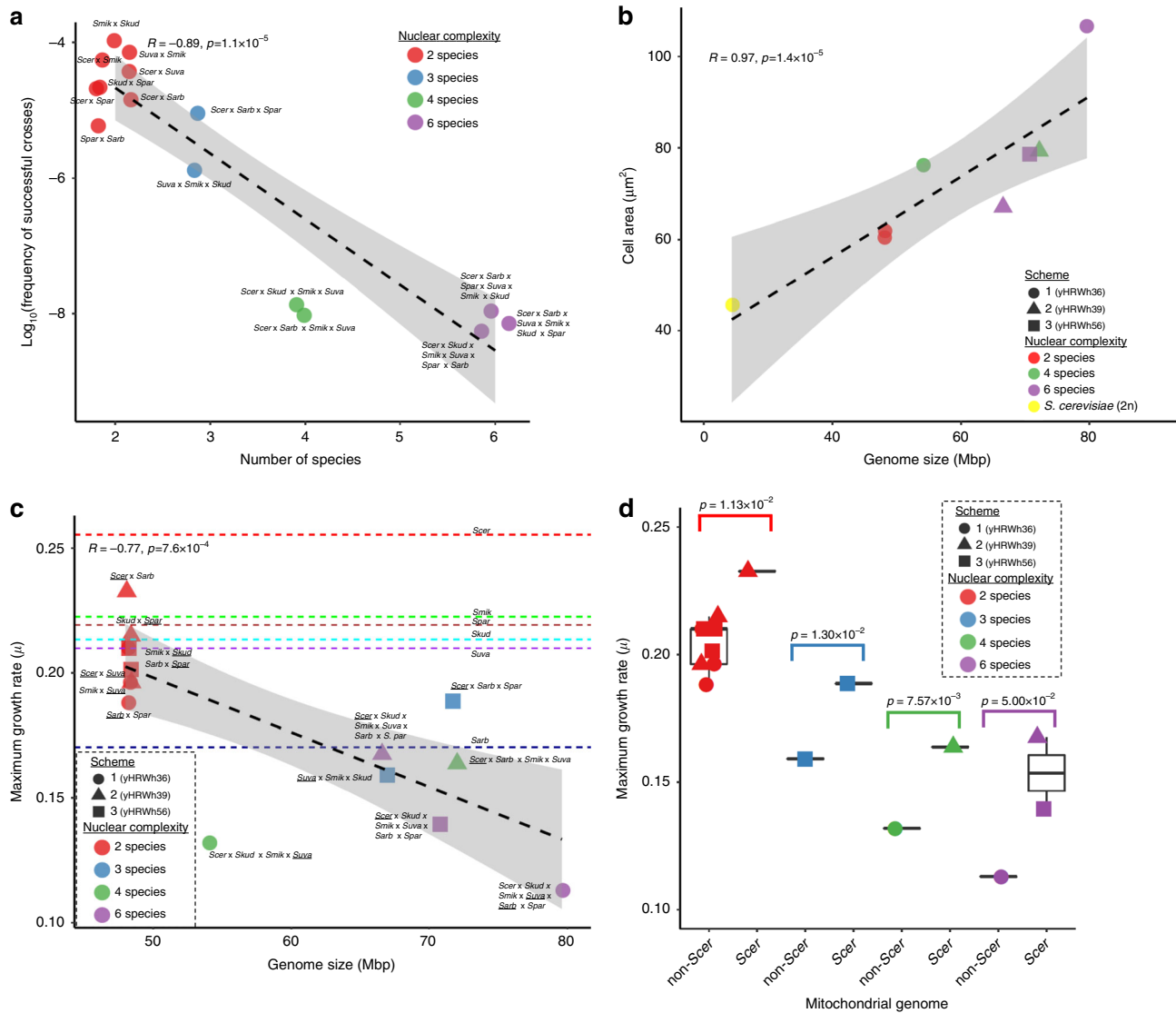


Fig. 3 Characteristics of six-species hybrids. **a** The number of species contributing genomes to synthetic hybrids is inversely correlated with the frequency of successful matings ($n = 2$ cross attempts). **b** Genome size is correlated with average cell area (average $n = 63$ counted cells). **c** Genome size (Supplementary Data 2) versus the average maximum growth rate (μ ($n = 6$ independent biological replicates), defined as $(\ln(OD2) - \ln(OD1)) / (T2 - T1)$) in rich medium at 20 °C (Supplementary Data 4). Dashed lines are the μ for the parent species indicated close to the line. For *S. uvarum*, the average of two strains with different HyPr plasmids is shown. **d** The maximum specific growth rate (μ , defined as $(\ln(OD2) - \ln(OD1)) / (T2 - T1)$) in rich medium at 20 °C is higher in interspecies hybrids inheriting *S. cerevisiae* mtDNA. Median values are represented by a horizontal line inside the box, and the upper and lower whiskers represent the highest and lowest values of the 1.5 × IQR (interquartile range), respectively. Colors correspond to the number of species contributing genomes to each strain. Synthetic hybrids generated from independent schemes are represented by different shapes in panels **b–d**. The Spearman rank-sum test R and p values are displayed. A linear regression and its 95% confidence interval are represented with a black dashed line and gray shadow, respectively. The mtDNA donor is underlined in the names in panel **c**. Species composition abbreviations are *Scer*, *S. cerevisiae*; *Spar*, *S. paradoxus*; *Smik*, *S. mikatae*; *Sarb*, *S. arboricola*; *Skud*, *S. kudriavzevii*; and *Suva*, *S. uvarum*. Source data are provided in the Source Data file and at <http://bit.ly/2v1rq1T>.

stabilization or the removal of interspecies genetic incompatibilities. The capacity of this approach to generate extensive karyotypic and phenotypic diversity will be of great interest for many industrial applications.

Mitochondrial inheritance also greatly influenced the genotypes and phenotypes of our synthetic hybrids. Even though a homoplasmic mtDNA state was quickly reached in most cases, a heteroplasmic state was detected in three exceptions that were all part of the same crossing scheme, and we offer a set of related possible explanations. The presence of selfish elements, such as homing endonucleases, could explain why multiple mitotypes were retained in yHRWh4. In this case, a portion of *S. mikatae* COX1, a gene with a high number of introns invaded by homing

endonuclease genes³⁷, seems to have been introduced into the *S. kudriavzevii* mtDNA (Supplementary Fig. 4A, C). An even more intriguing result occurred while generating the yHRWh10 hybrid, which remained in a heteroplasmic state and retained most of the mtDNAs of both parents (*S. cerevisiae* and *S. uvarum*) (Supplementary Fig. 4A). We recently demonstrated that, during the formation of *S. cerevisiae* × *S. uvarum* hybrids, the frequency of strains without a functional mtDNA was higher when the hybrid inherited a *S. uvarum* mtDNA, but introgression of the F-*Sc*III homing endonuclease gene restored normal mitochondrial retention³⁸. Therefore, the absence of F-*Sc*III in yHRWh10 may have influenced the loss of mtDNA in its descendants, such as the six-species hybrid yHRWh36, which retained only small regions

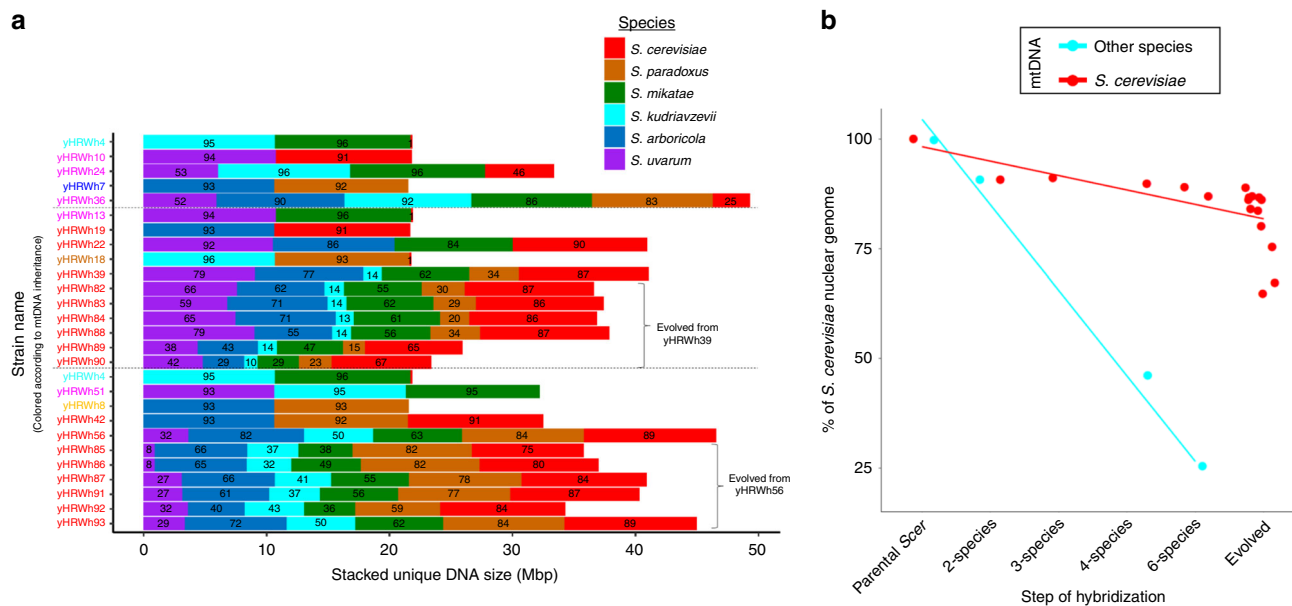


Fig. 4 Genome reduction during hybrid construction and adaptive laboratory evolution. **a** The genome contribution of each *Saccharomyces* species is stacked, and the percentage of retention is indicated inside the bar plot for each synthetic hybrid. Presence is reported, not copy number. For example, the red stacked bar for the hybrid yHRWh24 indicates that 46% of the total unique DNA for the *S. cerevisiae* parent (12.07 Mbp) was detected in the hybrid, which is ~5.55 Mbp of unique DNA. Synthetic hybrids are displayed in the order constructed (Supplementary Fig. 2). yHRWh4 is shown multiple times because of its use in two crossing schemes. We did not expect 100% genome contribution for each *Saccharomyces* species, even for recently created hybrids, because some genomic regions (e.g. repeats) are not unambiguously detectable with Illumina sequencing data. Genome size bars are colored according to each species' contribution. The strain names are colored based on the mtDNA inheritance inferred from *mitospIDer* (Supplementary Fig. 4), with two or more mtDNAs or regions shown as a gradient. **b** The nuclear compositions of the *S. cerevisiae* parent, synthetic hybrids, and evolved hybrids are plotted according to mtDNA inheritance ($n=1$ sequenced strain). Hybrids with *S. cerevisiae* mtDNA or with other mtDNA are colored in red and light blue, respectively. Source data are in the Source Data file and at <http://bit.ly/2v1rq1T>.

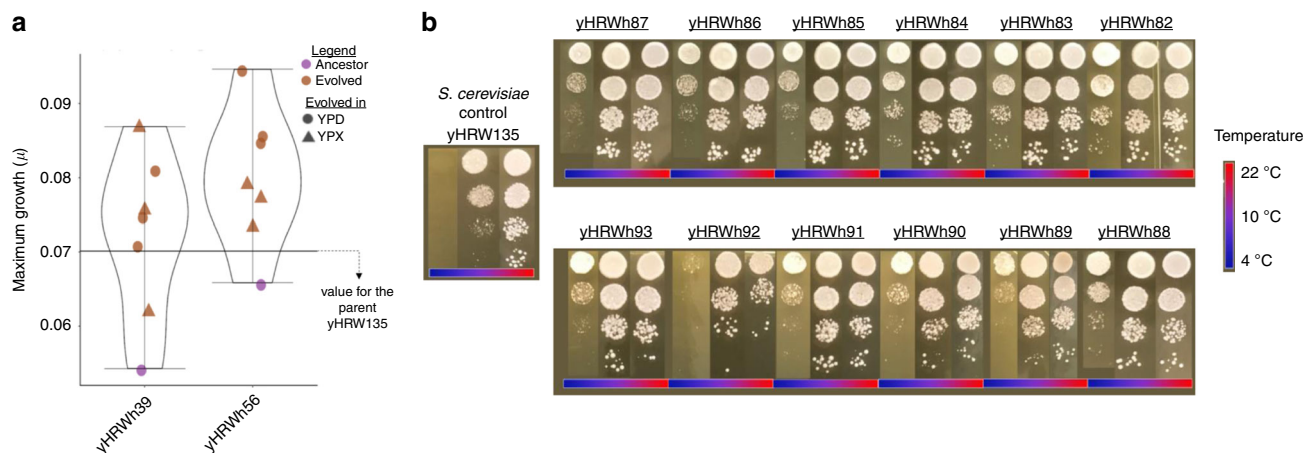


Fig. 5 Trait combination and improvement by adaptive laboratory evolution. **a** Box plots for the individual evolved colonies isolated from YPX or YPD plates after ALE and their synthetic hybrid ancestors. Kinetic parameters were tested in 3 ml YPX on a rotating culture wheel, identically to how they were evolved for 80 generations. The average values ($n=6$) of maximum specific growth rates (μ , defined as $(\ln(OD_2) - \ln(OD_1)) / (T_2 - T_1)$) for the *S. cerevisiae* reference strain (black line, yHRW135 was derived from yHRW134 by plasmid loss), ancestor six-species hybrids (purple dots), and evolved six-species hybrids (brown dots) are shown (Supplementary Data 5). Different shapes indicate the media in which the synthetic six-species hybrids were evolved. Additional kinetic parameters from microtiter plate experiments performed on evolved populations are shown in Supplementary Fig. 6 and Supplementary Data 6. Median values are represented by a horizontal line inside the box, and the upper and lower whiskers represent the highest and lowest values of the $1.5 \times IQR$ (interquartile range), respectively. **b** Spot tests for three temperatures (22, 10, and 4 °C) are displayed for the evolved strains and the *S. cerevisiae* reference strain yHRW135. Evolved six-species hybrids retained the ability to grow at 4 °C, a trait not possessed by *S. cerevisiae*, despite the fact that it was not selected during ALE. Source data are provided at <http://bit.ly/2v1rq1T>.

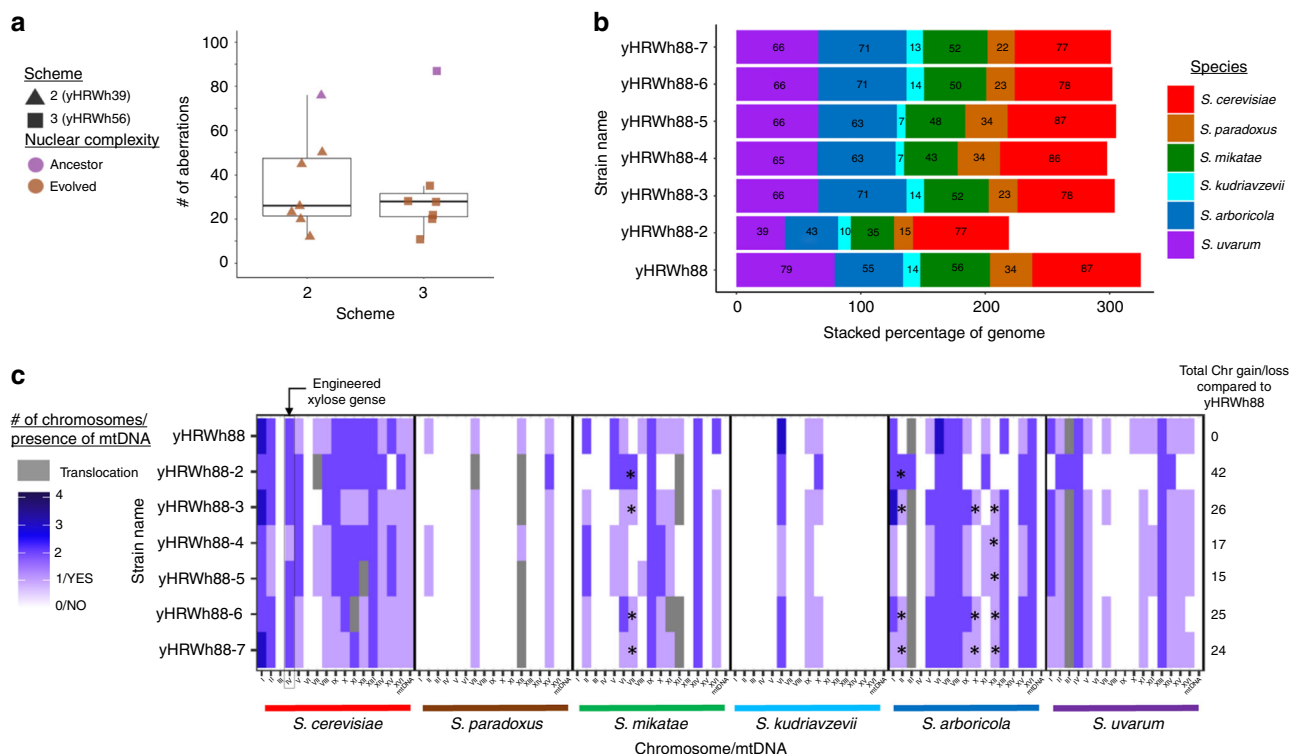


Fig. 6 Synthetic hybrids as a tool to study genome instability. **a** Box plots of the number of chromosomal aberrations inferred for ancestor and evolved synthetic hybrids (Fig. 2b, Supplementary Fig. 3) ($n = 1$ sequenced strain). Synthetic hybrids generated from each independent scheme are represented with different shapes. Purple and brown color points represent whether six-species hybrids were ancestor or evolved, respectively. Median values are represented by a horizontal line inside the box, and the upper and lower whiskers represent the highest and lowest values of the $1.5 \times$ IQR (interquartile range), respectively. **b** For each colony isolated from the population sample of the evolved synthetic hybrid yHRWh88, the genome contribution of each *Saccharomyces* species is stacked, and the percentage of retention is indicated inside the bar plot. The percentage of each species' contribution are colored according to the legend. **c** The number of chromosomes were inferred from *sppIDer* plots and corrected based on flow cytometry. The chromosome content was colored according to the species donor. The numbers of chromosomes for each species are colored according to the heatmap legend. Recombinant chromosomes are colored in gray. Asterisks indicate chromosomes that were retained in a particular colony but were not observed in the evolved yHRWh88 population sample, highlighting the instability of these hybrids. *S. cerevisiae* chromosome IV, where the xylose utilization genes were inserted, is indicated by the black square. Source data are provided in the Source Data file and at <http://bit.ly/2v1rQ1T>.

of *Saccharomyces arboricola* and *S. uvarum* mtDNAs (Supplementary Fig. 4C). In another recent study, mtDNA inheritance was dominated by one parent due to nuclear-mitochondrial interactions, rather than occurring stochastically³⁹. The loss of mtDNAs in particular hybrid combinations, as well as the unusually high or low coverage of specific regions in others, merits further study (Supplementary Note 2).

In summary, we generated and extensively characterized two-, three-, four-, and six-species synthetic hybrids, using our tool for yeast synthetic biology, iHyPr. We also improved the fitness of evolved strains, where polyploidy drives genome instability and evolution. Our higher-order allopolyploids acquired genome aberrations involving multiple species as they rapidly adapted to new environmental conditions. This technology pushes the budding yeast cell toward its limits in pursuit of basic research questions in chromosome biology and evolutionary genetics, as well as potential industrial applications.

Methods

Yeast strains and maintenance. The reference strain chosen for improvement was GLBRCY101, a haploid derivative of the *Saccharomyces cerevisiae* GLBRCY73 strain, which had been engineered with xylose utilization genes from *Scheffersomyces (Pichia) stipitis* and aerobically evolved for the consumption of xylose^{27–29}. Representative strains were selected from five additional *Saccharomyces* species based on published nuclear and mtDNAs (Supplementary Data 1). These six parent strains were used to generate the six-species hybrids. Yeast strains were stored in cryotubes with YPD (1% yeast extract, 2% peptone, and 2% glucose)

and 15% glycerol at -80°C . Routine cultures were maintained in YPD plus 2% agar plates at 24°C .

Two new Hybrid Production (HyPr) plasmids. We previously published two HyPr plasmids with *natMX* (pHCT2) and *hphMX* (pHMK34) resistance cassettes²⁰. We amplified the *ble* (Zeocin resistance) and *nptII* (G418 resistance) coding regions for marker swaps to generate pHRW32 and pHRW40 plasmids, respectively (Supplementary Data 8). The new HyPr plasmids enabled complex, iterative crossing schemes without adding extra steps to remove one of the two HyPr plasmids between the hybridization steps (Supplementary Fig. 1).

Saccharomyces transformation with HyPr plasmids. Before transforming GLBRCY101 with a HyPr plasmid, we removed its nuclear *kanMX* cassette by swapping the *kanMX* marker to *tkMX*⁴⁰. Next, we transformed this strain using a short DNA fragment designed to allow the *tkMX* gene to be removed via homologous recombination and selecting for successful marker loss on synthetic complete (SC) + FuDR medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 0.2% complete drop out mix, 2% glucose, and 50 $\mu\text{g}/\text{ml}$ 5-fluorodeoxyuridine). *S. cerevisiae* yHWA85 and representative strains of *Saccharomyces paradoxus*, *S. mikatae*, *S. kudriavzevii*, *S. arboricola*, and *S. uvarum* were transformed with one of the four HyPr plasmid versions (Supplementary Data 1). The diploid parent strains contain levels of heterozygosity lower than 0.037% (Supplementary Data 1).

Transformation of yeast strains was done using the lithium acetate/PEG-4000/carrier DNA method⁴¹, adjusting for the temperature tolerances of species²⁰. *S. cerevisiae* yHWA85 was first diploidized using the HyPr plasmid pHRW40, creating yHRW134 for subsequent crosses. The generation of this diploid strain occurred in one step, which was confirmed by polymerase chain reaction (PCR) amplification of the *MAT* loci (see below). The experimental reference strain yHRW135 was derived from yHRW134 by screening for spontaneous plasmid loss.

The iHyPr method. Following the HyPr method to facilitate mating-type switch²⁰, we pre-cultured strains with differentially marked HyPr plasmids in the presence of doxycycline to express the endonuclease encoded by *HO*, which is under a Tet-ON promoter (Supplementary Fig. 1); each plasmid also contains the full machinery for inducible expression of the promoter. To generate the six-species hybrids yHRWh36 and yHRWh39, we first hybridized three separate pairs of species, generating two-species hybrids (Supplementary Fig. 2A, B). In each case, once the three two-species hybrids were generated, two of those two-species hybrids were themselves hybridized to create a four-species hybrid, which finally was hybridized with the last two-species hybrid to generate the six-species hybrid. To generate the six-species hybrid yHRWh56, two two-species hybrids were separately crossed with diploid *Saccharomyces* strains from other species to create two separate three-species hybrids, which were then mated to generate the six-species hybrid (Supplementary Fig. 2C). Before each cross, parent strains were transformed with differentially marked HyPr plasmids (Supplementary Data 1 and 8, Supplementary Figs. 1 and 2) and treated with doxycycline in YPD at room temperature, except for *S. cerevisiae* which was incubated at 30 °C. The doxycycline triggers the expression of the *HO* endonuclease²⁰, which cuts one or more *MATa/MATα* idiomorphs and generates mating-compatible strains that behave as either *MATa* or *MATα*. A sample of each culture was combined in a 1-ml Eppendorf tube and patched on a YPD plate. After 2–3 days, a sample was taken with a toothpick and streaked on a YPD plate supplemented with the corresponding drugs to select for successful matings. In contrast to the original HyPr method, we pre-cultured the new hybrid in YPD with one of the two drugs used during the selective medium step, and that hybrid was then crossed with another strain containing one or two of the other HyPr plasmids not used previously (Supplementary Figs. 1 and 2). During these subsequent steps, we expected (and phenotypically verified) the loss of the HyPr plasmid containing the drug-resistance cassette not under selection. This approach and the additional HyPr plasmids made for this study facilitated the iterative crosses required to make six-species hybrids by avoiding the steps of plasmid removal and minimizing the number of generations between crosses (Supplementary Fig. 1).

The frequency of successful two-, three-, four-, and six-species hybrid generation were quantified in duplicates ($n = 2$) (Supplementary Data 3). The patch of co-culture was diluted in sterile H₂O, and a sample was spread onto both YPD plates and YPD supplemented with the appropriate drugs. The frequency of successful matings was calculated as the ratio between the number of colonies observed in YPD supplemented with the corresponding drugs and the number of colonies observed in YPD.

Mating-type and PCR-RFLP confirmation of strains. Diploidization of the *S. cerevisiae* strain was confirmed by PCR at the mating-type locus. Hybrid statuses were confirmed by restriction fragment length polymorphism (RFLP) analysis. We used the Standard Taq Polymerase (New England Biolabs, Ipswich, MA) and the primers listed in Supplementary Data 9. Genomic DNA was extracted using the phenol:chloroform method on a strain grown from pre-culture to saturation in YPD. Aliquots of 700 μ l of saturated culture were located in 1.5 ml microcentrifuge tubes that contained acid-washed beads. Each tube was centrifuged at maximum speed (21,130 \times G) for 5 min, and the supernatant was removed. Two hundred microliters of buffer EB (10 mM Tris-Cl, pH 8.0), 200 μ l of DNA lysis buffer (10 mM Tris pH 8.0, 1 mM EDTA, 100 mM NaCl, 1% SDS, 2% Triton X-100), and 200 μ l of phenol:chloroform were added to each tube. Vigorous vortexing was performed for 3–4 min, followed by 5 min of centrifugation at maximum speed. The top aqueous layer was transferred to 1 ml 100% EtOH. After an inversion mixture, DNA was precipitated at –80 °C for at least 10–15 min. A second centrifugation at maximum speed was performed, and the supernatant was discarded. We washed the pellet with 700 μ l of 70% EtOH, and we centrifuged again to remove any residue or trace of the supernatant. The pellet was dried and resuspended in 100 μ l of EB at 50–60 °C for 30 min. To remove RNA, we incubated the solution with 0.5 μ l of 10 mg/ml RNase A for 30 min at 37 °C. DNA was quantified with a Qubit 2.0 Fluorometer (ThermoFisher Scientific). For PCR-RFLP, resulting PCR products were digested with a restriction enzyme or a combination of multiple restriction enzyme assays able to discriminate among *Saccharomyces* species (New England Biolabs, Ipswich, MA). An extended PCR-RFLP pattern, developed in previous publications^{20,42} and this study, is detailed in Supplementary Data 10. Undigested PCR products were visualized on a 1.5% agarose gel, while digested PCR products were visualized on a 3% agarose gel.

Ploidy estimation by flow cytometry. Both asynchronous and hydroxyurea-arrested (G1/S arrested) mid-log cultures were prepared for each strain. Hydroxyurea-arrested strains were prepared to assist in the identification of G1 peaks in samples with broad and undefined cell cycle peaks. Briefly, cultures were grown to saturation and then diluted back 1:200. Back-diluted cultures were grown to a 0.4–0.6 optical density at 600 nm (OD₆₀₀). For each strain, 1 ml mid-log culture was transferred into 200 μ l 1 M hydroxyurea and incubated on a room temperature culture wheel for approximately half of the time that the respective strain took to grow from back-dilution to an OD₆₀₀ of 0.4–0.6. This ranged between 3 and 12 h. At the same time, 1 ml of asynchronous mid-log culture was harvested for fixation. All samples were fixed in 70% ethanol overnight, treated with RNase and Proteinase K, and finally stained with SYTOX Green dye

(Molecular Probes)⁴³. Stained cell suspensions were sonicated before flow cytometry. Fluorescence was measured with a BL1 laser (488 nm) on an Attune NxT (Invitrogen) flow cytometer at the lowest available flow rate. To accommodate for extremes in ploidy and cell size, voltage was adjusted to 250 for FITC dump channel (BL1) and FSC (Forward Scatter). All samples were run at the same voltage.

Flow cytometry data files were processed in FlowJo v10.4.2 (ref. 44). Samples were first gated on SSC (Side Scatter) and FSC to remove debris (Supplementary Fig. 9). Doublets were then removed by gating on BL1-A and FSC-A. A histogram of BL1-A values were then generated for remaining cells. Hydroxyurea peaks were identified and gated manually. Asynchronous G1 and G2 peaks were identified by applying a Watson (Pragmatic) Cell Cycle model and identifying G1 and G2 means. When cell cycle models did not fit the asynchronous sample data automatically, hydroxyurea samples were used to identify G1 peaks and these were manually gated to constrain G1 in asynchronous samples. Ploidy estimation were performed by comparing with fluorescence values of a haploid laboratory reference *S. cerevisiae* strain, S288C (Supplementary Data 1).

Cell size estimation. A subset of strains was used for microscopy analysis of cell size. Each strain was spotted from frozen stock onto YPD agar plates and grown at room temperature for 4 days. Water-cell suspensions were prepared for each strain, which were bright-field-imaged on an EVOS FL Auto 2.0 (Invitrogen) imaging system at $\times 400$. Cell area was analyzed in FIJI v2.0.0-rc-34/1.5a⁴⁵ using the Analyze Particles tool. Cell volume was estimated by assuming spherical cells and calculating radii (r) from cell area measurements.

Fitness quantification of the newly generated hybrids. To measure the impact of genome size increases on fitness, we performed a growth test in a rich medium. All parent species and the two-, three-, four-, and six-species hybrids were pre-cultured in 3 ml of YPD at room temperature. After pre-culture, 10 μ l of saturated culture was inoculated into a 96-well plate (Nunc, Roskilde, Denmark) containing 240 μ l of YPD. Spaces between the wells in the plates were filled with sterile H₂O to maintain the humidity of the plates and limit culture evaporation.

To monitor the growth of strains and populations in the different media conditions, the inoculated 96-well plate was placed in a BMG FLUOstar Omega (Ortenberg, Germany) at 20 °C. Absorbance at 595 nm was monitored every 15 min for 4 days. Background absorbance was subtracted from the average of nine negative controls containing the uninoculated medium being tested. Kinetic parameters for each condition were calculated in GCAT v6.3 (ref. 46). Median and standard deviations from six independent biological replicates, except yHRWh36 and yHRWh56 from which we obtained three replicates, were calculated in R (ref. 47) (Supplementary Data 4).

Genome sequencing and chromosome composition analyses. Genomic DNA (gDNA) samples from the diploidized *S. cerevisiae* strain and the two-, three-, four-, and six-species hybrids were submitted to the DOE Joint Genome Institute for paired-end Illumina sequencing. Evolved six-species hybrids and six individual colonies from yHRWh88 (see ALE section) were also submitted for sequencing. Libraries were constructed according to the manufacturer's instructions. Sequencing of the flow cell was performed on an Illumina MiSeq using MiSeq Reagent kits, following a 2 \times 150 nucleotide, indexed run recipe. Data were collected with Illumina RTA 1.18.54 and converted to fastq with Bcl2Fastq 2.20.0. Curated raw reads were submitted to the SRA database as Bioproject PRJNA476226 (Supplementary Data 11).

Genomic characterization was performed with sppIDer v1 (ref. 48). Our combined nuclear reference genome was built with the genome assemblies of *S. cerevisiae* GLBRCY22-3 (ref. 49), which is a close relative of the biofuel reference strain used here; *S. paradoxus* CBS432; *S. arboricola* CBS10644⁵⁰; *S. mikatae* IFO1815; *S. kudriavzevii* ZP591; *S. uvarum* CBS7001¹⁸; and *Saccharomyces eubayanus* FM1318 (ref. 51). Our combined mitochondrial reference genome was built with the mitochondrial assemblies of the aforementioned strains^{50–52}, except for CBS7001, whose mtDNA is still not completely assembled³⁸. Instead, we used the mtDNA of a close relative, *S. uvarum* CBS395⁵². Raw Illumina paired-end reads and the combined reference genomes were the input data of sppIDer, which is a wrapper that runs published tools to map the short reads to the combined reference genomes and creates several colorful and visually intuitive outputs⁴⁸. Here, we show depth of coverage plots from those species contributing genomes.

For each strain, the number of chromosomes and the ploidy were estimated from the sppIDer plots. This approximation gave a significant positive correlation with the ploidy estimated by flow cytometry (Spearman rank test $r = 0.91$, p value = 3.2×10^{-6}) (Supplementary Fig. 5C). The number of chromosomal aberrations was based on the number of gains, losses, or unbalanced translocations detected in the sppIDer plots (Supplementary Data 2). One chromosomal gain, loss, or unbalanced translocation was counted as one aberration. Aberrations observed in one hybrid and maintained in the offspring of subsequent crosses were not counted again; only new aberrations for each cross were reported in the aberration plot (Figs. 2a and 6a). Chromosomal aberrations involving parts of chromosomes were conservatively counted only in cases where there were clear fusions of entire chromosome arms.

Genome size and ploidy quantification. Two different approaches were performed to quantify the genome size of the sequenced strains. In the first approach, genome assemblies were performed using the collection of assemblers included in *iWGS* v1.1 (ref. 53). The assembly with the best assembly stats reported by *iWGS* was selected, and the genome size was reported (Supplementary Data 2). In the second approach, *sppIDer* coverage outputs (StrainName_winAvgDepth.d.txt) were parsed to quantify the percentage of each *Saccharomyces* nuclear genome retained in the hybrid, which was calculated as follows:

$$P_{\text{spp}} = \left(\frac{\text{Ct} \times \text{Ws}}{\text{Gs}} \right) \times 100,$$

where P_{spp} is the percentage for one of the parent species; Ct is the number of windows with a coverage mean value above 2; Ws is the window size; and Gs is the reference genome size for that parent species. These two calculations yielded a good approximation of the increased genome size, but both generated estimates that assumed the highly homozygous genome donated by each parent was haploid; *iWGS* and *sppIDer* plots were significantly correlated (Spearman rank test $r = 0.95$, p value = 2.2×10^{-16} , Supplementary Fig. 5D).

To get a better approximation of the genome size of each allopolyploid, we first determined the total number of copies of each chromosome contributed by each species, as quantified by *sppIDer*. Genome size was then calculated by multiplying the number of copies of each chromosome by its length and adding all these values together. Genome size and flow cytometry fluorescence were correlated (Spearman rank test $r = 0.93$, p value = 1.1×10^{-7} , Supplementary Fig. 5B).

Quantification of the number of copies of the xylose utilization cassette.

Illumina reads were extracted using the xylose utilization cassette sequence (8.7 kbp) as bait for *HybPiper* v1.2 (ref. 54). The generated bam files were viewed and sorted with *samtools* v1.4 (ref. 55), and the coverage for each nucleotide was quantified with *genomeCoverageBed*, which is included in *bedtools* v2.2.27 (ref. 56). The mean coverage of the coding sequence of the three engineered xylose utilization genes (*XYL1*, *XYL2*, and *XYL3*) (3.9 kbp) was calculated from the *genomeCoverageBed* output. For the chromosome IV, mean coverage values for windows of 3.9 kbp were calculated from the *genomeCoverageBed* output generated by *sppIDer*. The cassette value and chromosome distributions for each strain were compared by a one-sided Wilcoxon rank-sum test for a significant deviation from the expected ratio 1:1 (one copy of the cassette to one copy of chromosome IV) (Supplementary Data 7).

ALE and colony selection. Two of the three six-species hybrids (during construction, the third lost *S. cerevisiae* chromosome IV, where *Sch. stipitis* xylose utilization genes had been inserted) were evolved in triplicate at room temperature in tubes with two independent media conditions: 3.0 ml YPD or 3.0 ml YPX (1% yeast extract, 2% peptone, and 2% xylose). Three to five days of fermentation were performed to allow cells to consume the sugars, and an aliquot of each replicate was transferred at of 0.1 OD_{600} to a fresh medium until it reached approximately 80 generations. A colony from each independent ALE experiment, regardless of whether they were evolved in glucose or xylose, was selected on YPX plates (1% yeast extract, 2% peptone, 2% xylose, and 2% agar) and cryopreserved.

Microtiter plate growth curves. We compared the growth kinetics of the *S. cerevisiae* reference strain yHRW135, the ancestors of the two six-species hybrids retaining the chromosome IV (yHRWh39, yHRWh56), and populations of the evolved hybrids. Growth was tested in YPD and YPX at room temperature. Strains or populations were pre-cultured in 3.0 ml YPD or YPX, depending of the medium tested. After pre-culture, 10 μ l of saturated culture was inoculated into a 96-well plate (Nunc, Roskilde, Denmark) containing 240 μ l of identical medium as the pre-culture. Spaces between the wells in the plates were filled with sterile H_2O to maintain the humidity of the plates. The reference strain was cross-inoculated in all conditions; for example, yHRW135 pre-cultured in YPX was tested in both YPD and in YPX.

To monitor the growth of strains and populations in the different media, we inoculated 96-well plates and placed them in a BMG FLUOstar Omega at 20 °C. Absorbance at 595 nm was monitored every 15 min for 5 days. Background absorbance was subtracted from the average of three negative controls containing the uninoculated medium being tested. Kinetic parameters for each condition were calculated in *GCAT* v6.3 (ref. 46). Median and standard deviations from three independent biological replicates were calculated in R⁴⁷ (Supplementary Data 6). For each medium, parameters were normalized against the data generated by the reference strain yHRW135 when it was pre-cultured and grown in the medium tested.

Cold tolerance spot test. Temperature growth profiles are well known to vary among *Saccharomyces* species^{30,31}. In particular, *S. uvarum* and *S. kudriavzevii* are able to grow at low temperatures where *S. cerevisiae* cannot grow. To test if some phenotypic traits might be retained independently of the media regime, we performed spot tests in rich medium at different temperatures (22, 10, and 4 °C). The

S. cerevisiae reference strain (yHRW135) and the evolved six-species hybrids were compared. All strains were pre-cultured in liquid YPD medium at room temperature to saturation. Cultures were subjected to a series of 10-fold dilutions in YPD. Five microliters of each dilution was spotted onto three YPD agar plates identically. Plates were incubated in sealed plastic bags to keep them from drying out at the temperatures mentioned above. Each plate was photographed when most strains exhibited significant growth (4 days for 22 °C, 11 days for 10 °C, and 38 days for 4 °C).

Culture wheel growth curves. Strains isolated from single colonies from evolved hybrids, ancestor hybrids, and the reference strain (yHRW135) were pre-cultured in YPX and inoculated at an initial OD_{600} of 0.1 into 3 ml glass tubes containing YPX. Growth was monitored by measuring OD_{600} . Kinetic parameters were calculated as above. Median and standard deviations from six independent biological replicates were calculated as above. These experimental conditions most closely matched the conditions in which the strains were evolved, and they are reported in Fig. 5a and Supplementary Data 5.

Statistical analyses. Data analyses and plots were performed in R v3.6.1 (ref. 47) using the RStudio v1.1.383 platform. Raw data manipulation was done with *dplyr* v0.8.3 and *reshape2* v1.4.3. Linear models of regressions were added to the plots in Figs. 3 and 4a and Supplementary Figs. 5 and 8 using the *geom_smooth* option in the R package *ggplot2* v3.2.1. A LOESS regression line was added to the plot in Fig. 2a using the *geom_smooth* option in the R package *ggplot2*. For aberration data (Fig. 2a), r^2 and significance of regression were calculated with *summary(lm(y ~ x))*, where x was the number of species, and y was the number of observed aberrations. Correlations for ploidy and assembly comparisons were calculated in R using the *ggpubr* v0.2.3 package to apply a Spearman rank-sum test (Fig. 3, Supplementary Figs. 5 and 8), and plots were generated using *ggplot2*.

The impact of mitochondrial inheritance (Fig. 4b) in the retention of the nuclear genome of those hybrids involving *S. cerevisiae* was tested using a multifactor ANOVA in R, using *summary(aov(P ~ M * C))*, where P is the percentage retained of the *S. cerevisiae* nuclear genome; M is the mtDNA, which was encoded as a binary character (either as the *S. cerevisiae* mtDNA or that of another species); and C is the type of strain (i.e. classified as the *S. cerevisiae* parent; two-, three-, four-, or ancestor six-species hybrid; and evolved six-species hybrid).

t-tests for significant differences between frequency of chromosome gains and losses and Wilcoxon rank-sum tests for significant differences in the kinetic parameters shown in Figs. 3d and 5a and Supplementary Fig. 6, respectively, were performed in R.

Flow cytometry data were analyzed and plotted in R.

Correlations were tested in R using a Spearman rank-sum test and plotted using *ggplot2*.

The quantified numbers of chromosomes and mitochondrial inheritance inferred from *sppIDer* were represented by heatmaps using the R package *pheatmap* v1.0.12.

Statistics and reproducibility. iHyPr was tested three times using different schemes. The rest of experiments were tested once with biologically independent replicates according to the legend descriptions in each figure.

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

Data availability

Raw genome sequencing data have been deposited in NCBI's SRA database, Bioproject PRJNA467226. HyPr plasmids have been deposited in Addgene as deposit 77444. The source data underlying Figs. 2–6, Supplementary Figs. 3–8, and Supplementary Data 1–7 are provided in the Source Data file and at <http://bit.ly/2v1rq1T>. All other relevant data are available from the authors upon reasonable request.

Received: 27 May 2019; Accepted: 10 March 2020;

Published online: 29 April 2020

References

1. Van de Peer, Y., Mizrahi, E. & Marchal, K. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **18**, 411–424 (2017).
2. te Beest, M. et al. The more the better? The role of polyploidy in facilitating plant invasions. *Ann. Bot.* **109**, 19–45 (2012).
3. Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
4. Mable, B. K., Alexandrou, M. A. & Taylor, M. I. Genome duplication in amphibians and fish: an extended synthesis. *J. Zool.* **284**, 151–182 (2011).

5. Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997).
6. Albertin, W. & Marullo, P. Polyploidy in fungi: evolution after whole-genome duplication. *Proc. R. Soc. Lond. B Biol. Sci.* **279**, 2497–2509 (2012).
7. Selmecki, A. M. et al. Polyploidy can drive rapid adaptation in yeast. *Nature* **519**, 349–352 (2015).
8. Sprouffs, K. et al. Cancer in light of experimental evolution. *Curr. Biol.* **22**, R762–R771 (2012).
9. Mortimer, R. K. Radiobiological and genetic studies on a polyploid series (haploid to hexaploid) of *Saccharomyces cerevisiae*. *Radiat. Res.* **9**, 312–326 (1958).
10. Dunham, M. J. et al. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **99**, 16144–16149 (2002).
11. Heil, C. S. et al. Loss of heterozygosity drives adaptation in hybrid yeast. *Mol. Biol. Evol.* **34**, 1596–1612 (2017).
12. Langdon, Q. K. et al. Fermentation innovation through complex hybridization of wild and domesticated yeasts. *Nat. Ecol. Evol.* **3**, 1576–1586 (2019).
13. Gallone, B. et al. Interspecific hybridization facilitates niche adaptation in beer yeast. *Nat. Ecol. Evol.* **3**, 1562–1575 (2019).
14. Dymond, J. & Boeke, J. The *Saccharomyces cerevisiae* SCRaMBLE system and genome minimization. *Bioeng. Bugs* **3**, 170–173 (2012).
15. Scannell, D. R. et al. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3* **1**, 11–25 (2011).
16. Gunge, N. & Nakatomi, Y. Genetic mechanisms of rare matings of the yeast *Saccharomyces cerevisiae* heterozygous for mating type. *Genetics* **70**, 41–58 (1972).
17. Steensels, J. et al. Improving industrial yeast strains: exploiting natural and artificial diversity. *FEMS Microbiol. Rev.* **38**, 947–995 (2014).
18. Peris, D. et al. Reticulate evolution in the *Saccharomyces* genus. *International Specialized Symposium on Yeasts*, Bariloche, Argentina (2018).
19. Baker, E. P. et al. Mitochondrial DNA and temperature tolerance in lager yeasts. *Sci. Adv.* **5**, eaav1869 (2019).
20. Alexander, W. G. et al. Efficient engineering of marker-free synthetic allotetraploids of *Saccharomyces*. *Fungal Genet. Biol.* **89**, 10–17 (2016).
21. Jelenic, I. et al. Spindle dynamics model explains chromosome loss rates in yeast polyploid cells. *Front. Genet.* **9**, 296 (2018).
22. Gregory, T. R. Coincidence, coevolution, or causation? DNA content, cellsize, and the C-value enigma. *Biol. Rev.* **76**, 65–101 (2001).
23. Neurohr, G. E. et al. Excessive cell growth causes cytoplasm dilution and contributes to senescence. *Cell* **176**, 1083–1097 (2019).
24. Tymowska, J. “Polyploidy and Cytogenetic Variation in Frogs of the Genus *Xenopus*,” in *Amphibian Cytogenetics and Evolution* (eds Green, D. M. & Sessions, S. K.) 259–297 (Academic Press, 1991).
25. Ainouche, M. L. et al. Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *N. Phytol.* **161**, 165–172 (2003).
26. Berger, K. H. & Yaffe, M. P. Mitochondrial DNA inheritance in *Saccharomyces cerevisiae*. *Trends Microbiol.* **8**, 508–513 (2000).
27. Peris, D. et al. Hybridization and directed evolution of diverse *Saccharomyces* species for cellulosic biofuel production. *Biotechnol. Biofuels* **10**, 78 (2017).
28. Wohlbach, D. J. et al. Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. *Proc. Natl. Acad. Sci. USA* **108**, 13212–13217 (2011).
29. Sato, T. K. et al. Harnessing genetic diversity in *Saccharomyces cerevisiae* for improved fermentation of xylose in hydrolysates of alkaline hydrogen peroxide pretreated biomass. *Appl. Environ. Microbiol.* **8**, 540–554 (2014).
30. Salvadó, Z. et al. Temperature adaptation markedly determines evolution within the genus *Saccharomyces*. *Appl. Environ. Microbiol.* **77**, 2292–2302 (2011).
31. Gonçalves, P. et al. Evidence for divergent evolution of growth temperature preference in sympatric *Saccharomyces* species. *PLoS ONE* **6**, e20739 (2011).
32. Torija, M. J. et al. Effects of fermentation temperature and *Saccharomyces* species on the cell fatty acid composition and presence of volatile compounds in wine. *Int. J. Food Microbiol.* **85**, 127–136 (2003).
33. Beltran, G. et al. Integration of transcriptomic and metabolic analyses for understanding the global responses of low-temperature winemaking fermentations. *FEMS Yeast Res.* **6**, 1167–1183 (2006).
34. Beltran, G. et al. Effect of fermentation temperature and culture media on the yeast lipid composition and wine volatile compounds. *Int. J. Food Microbiol.* **121**, 169–177 (2008).
35. Pérez-Través, L. et al. Evaluation of different genetic procedures for the generation of artificial hybrids in *Saccharomyces* genus for winemaking. *Int. J. Food Microbiol.* **156**, 102–111 (2012).
36. Xie, Z. X. et al. Rapid and efficient CRISPR/Cas9-based mating-type switching of *Saccharomyces cerevisiae*. *G3* **8**, 173 (2018).
37. Wu, B. & Hao, W. Horizontal transfer and gene conversion as an important driving force in shaping the landscape of mitochondrial introns. *G3* **4**, 605–612 (2014).
38. Li, X. C. et al. Mitochondria-encoded genes contribute to the evolution of heat and cold tolerance among *Saccharomyces* species. *Sci. Adv.* **5**, eaav1848 (2019).
39. Verspohl, A., Pignedoli, S. & Giudici, P. The inheritance of mitochondrial DNA in interspecific *Saccharomyces* hybrids and their properties in winemaking. *Yeast* **35**, 173–187 (2018).
40. Alexander, W. G., Doering, D. T. & Hittinger, C. T. High-efficiency genome editing and allele replacement in prototrophic and wild strains of *Saccharomyces*. *Genetics* **198**, 859–866 (2014).
41. Gietz, R. D. & Woods, R. A. Genetic transformation of yeast. *Biotechniques* **30**, 816–831 (2001).
42. Peris, D. et al. The molecular characterization of new types of *S. cerevisiae* x *S. kudriavzevii* hybrid yeasts unveils a high genetic diversity. *Yeast* **29**, 81–91 (2012).
43. Haase, S. B. & Reed, S. I. Improved flow cytometric analysis of the budding yeast cell cycle. *Cell Cycle* **1**, 117–121 (2002).
44. Leland Stanford Jr. University. *FlowJo* (1996).
45. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
46. Bukhman, Y. et al. Modeling microbial growth curves with GCAT. *Bioenergy Res.* **8**, 1–9 (2015).
47. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2010).
48. Langdon, Q. K. et al. sppIDer: a species identification tool to investigate hybrid genomes with high-throughput sequencing. *Mol. Biol. Evol.* **35**, 2835–2849 (2018).
49. McIlwain, S. J. et al. Genome sequence and analysis of a stress-tolerant, wild-derived strain of *Saccharomyces cerevisiae* used in biofuels research. *G3* **6**, 1757–1766 (2016).
50. Yue, J. X. et al. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* **49**, 913–924 (2017).
51. Baker, E. et al. The genome sequence of *Saccharomyces eubayanus* and the domestication of lager-brewing yeasts. *Mol. Biol. Evol.* **32**, 2818–2831 (2015).
52. Sulo, P. et al. The evolutionary history of *Saccharomyces* species inferred from completed mitochondrial genomes and revision in the ‘yeast mitochondrial genetic code’. *DNA Res.* **24**, 571–583 (2017).
53. Zhou, X. et al. *In silico* Whole Genome Sequencer & Analyzer (iWGS): a computational pipeline to guide the design and analysis of de novo genome sequencing studies. *G3* **6**, 3655–3670 (2016).
54. Johnson, M. G. et al. HybPiper: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.* **4**, 1600016 (2016).
55. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
56. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

Acknowledgements

We thank Trey K. Sato for providing GLBRCY101, Srivatsan Raman for flow cytometry access, Amanda B. Hulfactor for assistance with Fig. 5b, the Joint Genome Institute (JGI) for providing Illumina Sequencing services, and Miguel Morard for feedback on preliminary figures. This material is based upon work supported in part by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494; the National Science Foundation under Grant Number DEB-1253634; the USDA National Institute of Food and Agriculture Hatch Project Number 1020204; and the Robert Draper Technology Innovation Fund from the Wisconsin Alumni Research Foundation (WARF). D.P. is a Marie Skłodowska-Curie fellow of the European Union’s Horizon 2020 research and innovation programme, grant agreement No. 747775. K.J.F. is a Morgridge Metabolism Interdisciplinary Fellow of the Morgridge Institute for Research. C.T.H. is a Pew Scholar in the Biomedical Sciences, a Vilas Early Career Investigator, and a H. I. Romnes Faculty Fellow, supported by the Pew Charitable Trusts, Vilas Trust Estate, and Office of the Vice Chancellor for Research and Graduate Education with funding from WARF, respectively. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231.

Author contributions

Conceived the experiments: D.P., W.G.A., K.J.F., R.L.W., and C.T.H. D.P. and C.T.H. mentored R.V.M., and W.G.A. mentored M.G.B. and E.J.U. Engineered strains: W.G.A., R.V.M., and R.L.W. Data generation: W.G.A., K.J.F., M.G.B., E.J.U., and R.L.W. Data analysis: D.P. and K.J.F. Supervised the study: D.P. and C.T.H. Wrote the paper with editorial input from all co-authors: D.P. and C.T.H.

Competing interests

The Wisconsin Alumni Research Foundation has filed a patent application entitled “Synthetic yeast cells and methods of making and using same” (describing the HyPr and iHyPr methods with W.G.A., D.P., and C.T.H. as inventors, US application #15/805950, filed 11/7/2017, pending). K.J.F., R.V.M., M.G.B., E.J.U., and R.L.W. declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-15559-4>.

Correspondence and requests for materials should be addressed to D.P. or C.T.H.

Peer review information *Nature Communications* thanks Tom Ellis and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission 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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020