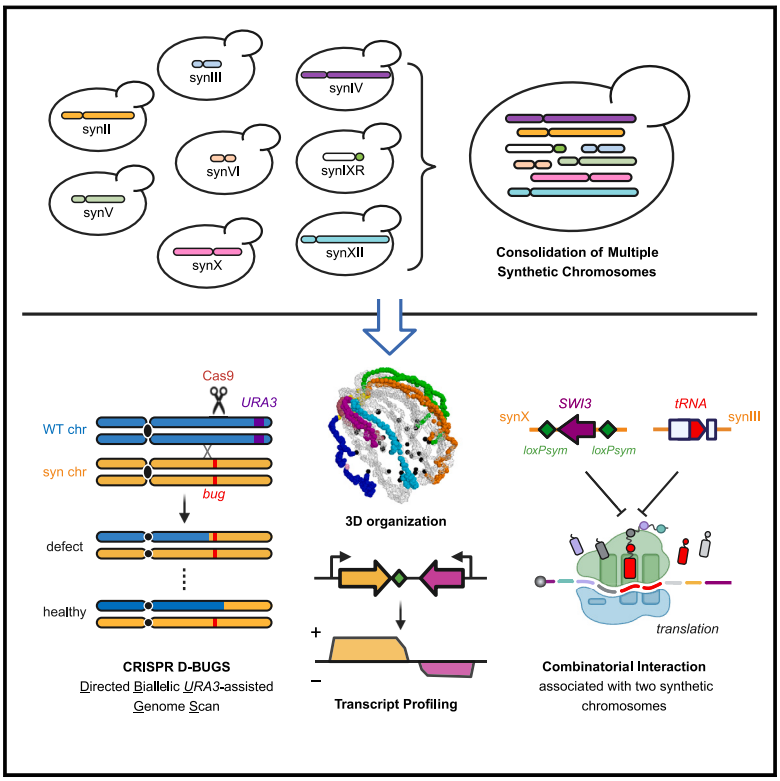
Cell

Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions

Graphical abstract



Highlights

- Consolidation of synthetic chromosomes using advanced endoreduplication intercrossing
- CRISPR D-BUGS to map "bugs" of designer modifications
- Combinatorial interaction links transcription, inositol metabolism, and tRNA abundance
- Next-generation chromosome consolidation strategy to finish Sc2.0

Authors

Yu Zhao, Camila Coelho, Amanda L. Hughes, ..., Lars M. Steinmetz, Yizhi Cai, Jef D. Boeke

Correspondence

jef.boeke@nyulangone.org

In brief

Multiple synthetic chromosomes in Sc2.0 were consolidated together into a single yeast strain, revealing combinational genetic interactions through CRISPR D-BUGS.

Zhao et al., 2023, Cell *186*, 5220–5236
November 22, 2023
https://doi.org/10.1016/j.cell.2023.09.025

CellPress

CellPress
OPEN ACCESS

Cell

Article

Debugging and consolidating multiple synthetic chromosomes reveals combinatorial genetic interactions

Yu Zhao,[1] Camila Coelho,[1,13] Amanda L. Hughes,[2,13] Luciana Lazar-Stefanita,[1,13] Sandy Yang,[1,13] Aaron N. Brooks,[2,15] Roy S.K. Walker,[3,18] Weimin Zhang,[1] Stephanie Lauer,[1] Cindy Hernandez,[1] Jitong Cai,[4] Leslie A. Mitchell,[1] Neta Agmon,[1,14] Yue Shen,[5,6] Joseph Sall,[7] Viola Fanfani,[8,16] Anavi Jalan,[9] Jordan Rivera,[9] Feng-Xia Liang,[7] Joel S. Bader,[4] Giovanni Stracquadanio,[8] Lars M. Steinmetz,[2,10,17] Yizhi Cai,[11] and Jef D. Boeke[1,12,19,*]

[1]Institute for Systems Genetics and Department of Biochemistry and Molecular Pharmacology, NYU Langone Health, New York, NY 10016, USA
[2]European Molecular Biology Laboratory (EMBL), Genome Biology Unit, 69117 Heidelberg, Germany
[3]School of Engineering, Institute for Bioengineering, the University of Edinburgh, Edinburgh EH9 3BF
[4]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA
[5]BGI, Shenzhen, Beishan, Industrial Zone, Shenzhen 518083, China
[6]Guangdong Provincial Key Laboratory of Genome Read and Write, BGI, Shenzhen, Shenzhen 518120, China
[7]Microscopy Laboratory, NYU Langone Health, New York, NY 10016, USA
[8]School of Biological Sciences, the University of Edinburgh, Edinburgh EH9 3BF
[9]Department of Biology, New York University, New York, NY, USA
[10]Department of Genetics and Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94304, USA
[11]Manchester Institute of Biotechnology, the University of Manchester, 131 Princess Street, Manchester M1 7DN, UK
[12]Department of Biomedical Engineering, NYU Tandon School of Engineering, Brooklyn, New York, NY 11201, USA
[13]These authors contributed equally
[14]Present addresses: Alagene, Inc. Tel Aviv, Israel
[15]Present addresses: ANB, Inscripta, Inc. Boulder CO 80301, USA
[16]Present addresses: VF, Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[17]Present addresses: LAM, Neochromosome, Inc., Long Island City, NY, USA
[18]Present addresses: RW, ARC Centre of Excellence in Synthetic Biology, School of Natural Sciences, Macquarie University, Sydney, Australia
[19]Lead contact
*Correspondence: jef.boeke@nyulangone.org
https://doi.org/10.1016/j.cell.2023.09.025

SUMMARY

The Sc2.0 project is building a eukaryotic synthetic genome from scratch. A major milestone has been achieved with all individual Sc2.0 chromosomes assembled. Here, we describe the consolidation of multiple synthetic chromosomes using advanced endoreduplication intercrossing with tRNA expression cassettes to generate a strain with 6.5 synthetic chromosomes. The 3D chromosome organization and transcript isoform profiles were evaluated using Hi-C and long-read direct RNA sequencing. We developed CRISPR Directed Biallelic *URA3*-assisted Genome Scan, or "CRISPR D-BUGS," to map phenotypic variants caused by specific designer modifications, known as "bugs." We first fine-mapped a bug in synthetic chromosome II (*synII*) and then discovered a combinatorial interaction associated with *synIII* and *synX*, revealing an unexpected genetic interaction that links transcriptional regulation, inositol metabolism, and $tRNA_{Ser}^{CGA}$ abundance. Finally, to expedite consolidation, we employed chromosome substitution to incorporate the largest chromosome (*synIV*), thereby consolidating >50% of the Sc2.0 genome in one strain.

INTRODUCTION

Rapid advances in DNA synthesis technology enable transitioning from genome reading and editing to genome writing. The field of synthetic genomics has achieved several milestones with regard to synthetic genomes, including the synthesis of the genomes of various viruses and bacteria such as *Mycoplasma genitalium* and *Escherichia coli* (*E. coli*).[1–5] Synthetic designer genomes provide a versatile platform for addressing previously unapproachable biological questions,[6,7] such as rapid vaccine response to pandemics,[8] minimal viability requirements for a bacterial cell,[9] potential for codon swaps,[5,10,11] incorporation of non-standard amino acids, and virus resistance.[4,12]

In the synthetic yeast project (Sc2.0), we are designing and synthesizing a eukaryotic genome *in silico* through a bottom-up approach.[13] This Sc2.0 genome is based on *Saccharomyces*

cerevisiae (*S. cerevisiae*), a unicellular eukaryotic model organism widely used in basic research and industrial fermentation. Thousands of genome-wide edits were introduced, including deletion of mobile elements and introns, relocation of tRNAs, and swapping of stop codons from TAG to TAA, freeing up TAG to potentially encode a non-standard amino acid.[13] We also developed a watermark system called the PCRtag by synonymously recoding snippets of open reading frames (ORFs) to specify pairs of primers selectively recognizing either wild-type or synthetic chromosomes, allowing facile distinction between synthetic and native genomic content by PCR.[14] We inserted a palindromic loxPsym site into the 3′ UTR of each nonessential gene, enabling a Cre recombinase-mediated genome-wide shuffling system called SCRaMbLE.[15–22] These modifications are designed to help increase genome stability, generate various phenotypes, and tackle biological questions. Importantly, synteny of protein-coding genes is conserved during assembly, whereas synteny of tRNA genes is utterly destroyed by design in the synthetic chromosomes described here. The SCRaMbLE method provides a means to rearrange the synteny of protein-coding genes on demand, as a "sequel" to initial chromosome design and synthesis.

Previously, six complete chromosomes (*synII*, *synIII*, *synV*, *synVI*, *synX*, and *synXII*) and one chromosome arm (*synIX* right arm) out of 16 chromosomes (abbreviated 6.5 hereafter) had been successfully synthesized.[14,23–28] Since then, assembly of all synthetic chromosomes has been completed.[20,29–37] Each synthetic chromosome was synthesized separately by teams comprising the Sc2.0 consortium. Consequently, each haploid yeast strain produced contains only one synthetic chromosome, leaving most of the genome native. A challenge is to consolidate every chromosome into one fully synthetic Sc2.0 strain. Here, we used an "endoreduplication intercross" strategy to consolidate all previously constructed 6.5 synthetic chromosomes (*synII, synIII, synV, synVI, synIXR, synX*, and *synXII*) into a single strain, referred to as syn6.5. However, the original endoreduplication intercross strategy presented a number of challenges that prevented a simple scale-up to facilitate combining many synthetic chromosomes into one high-fitness strain.

One Sc2.0 designer feature is the ultimate relocation of all tRNAs to a neochromosome, requiring removal of all endogenous tRNAs.[13] Native *S. cerevisiae* contains 275 genomic tRNAs. In the original 6.5 strains, a total of 97 of these tRNA genes had been deleted, collectively accounting for one-third of tRNA pools (Table S1). This feature created a practical challenge for consolidating chromosomes—before neochromosome assembly and delivery are complete, the available number of tRNA genes and thus derived tRNA molecules will drop as more and more synthetic chromosomes are consolidated. To avoid possible growth defects due to reduced tRNA abundance during consolidation, we integrated a corresponding "tRNA array" into each synthetic chromosome. Each tRNA array consists of the synthetic counterparts of all tRNAs originally deleted from their native locus on that chromosome, strung together in tandem, and inserted at a single position on each corresponding synthetic chromosome. Addition of tRNA arrays was crucial for reproducibly obtaining the many endoreduplication intercross strains needed to consolidate all these chromosomes, and addition of tRNA arrays to each synthetic chromosome defines a key enhancement—the "advanced endoreduplication intercross." A further limitation to the original endoreduplication intercross involved its scalability. We discovered that attempts to destabilize more than one chromosome at a time often met with failure, presumably because when more than one chromosome is destabilized simultaneously, the genic imbalance is great, leading to serious fitness defects. Thus, it was necessary to develop an elaborate multistep intercrossing procedure to successfully combine the 6.5 synthetic chromosomes in a single, relatively high-fitness strain.

Finally, occasional designer features can result in unexpected fitness defects, and we refer to these genomic modifications as "bugs.[26]" Precisely and systematically mapping these variants has been challenging and laborious. Inspired by diverse CRISPR applications in yeast genomic editing, regulation, and mapping,[38–43] we developed a highly reliable bug-mapping method known as CRISPR Directed Biallelic *URA3*-assisted Genome Scan (CRISPR D-BUGS). We successfully repaired bugs identified in single synthetic chromosomes, including a mitochondria-related defect caused by two loxPsym sites in the 3′ UTR of *SHM1*. CRISPR D-BUGS was also expanded to map a combinatorial defect associated with an essential tRNA gene *SUP61* in *synIII* and *SWI3* in *synX*. In this case, neither variant alone caused a fitness defect, but the combination of the two caused a severe defect.

To further probe effects of designer modifications on three-dimensional (3D) genome organization and transcriptional regulation of the compact yeast genome, we used chromosome conformation capture (Hi-C) and nanopore direct RNA sequencing[44] to characterize a strain with 6.5 synthetic chromosomes. Finally, to expedite consolidation, we used a next-generation method, chromosome substitution,[33] to transfer the largest single synthetic chromosome, *synIV*, into the yeast strain already carrying 6.5 synthetic chromosomes, thereby consolidating more than half of the Sc2.0 genome and producing the syn7.5 strain.

RESULTS

Synthetic chromosome consolidation using an advanced endoreduplication intercross strategy

The Sc2.0 consortium assembled each of the 16 synthetic chromosomes (*synI–synXVI*) in discrete haploid strains, and we previously established a consolidation strategy called endoreduplication intercrossing.[26,45] Briefly, two strains with different synthetic chromosomes and opposite mating type are mated, generating heterozygous diploid strain carrying two synthetic chromosomes along with their native counterparts. After destabilizing native chromosomes with centromere-adjacent pGAL promoters, sporulating, and screening the resulting spore clones, we obtain haploid strains with two or more synthetic chromosomes. However, the original endoreduplication intercross strategy needed enhancement to accommodate multiple large chromosomes in a single cell. We thus developed an advanced endoreduplication intercross strategy in which tRNA arrays (see below) were restored to each synthetic chromosome to avoid such deficits. Following several rounds of intercross consolidation in which one synthetic chromosome was consolidated per round, we obtained a single strain, YZY1178, with 6.5

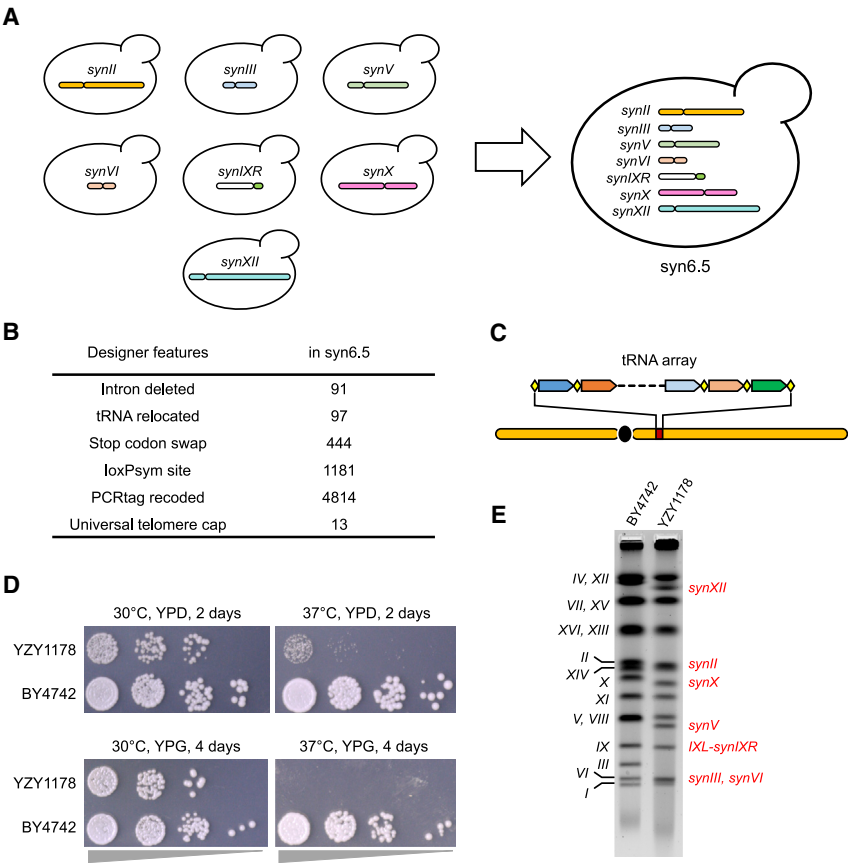


Figure 1. Consolidation of multiple synthetic chromosomes

(A) All previously assembled synthetic chromosomes were consolidated using endoreduplication intercrosses, generating one haploid strain, syn6.5.

(B) Sc2.0 designer features carried in the syn6.5 strain.

(C) A tRNA array was integrated into each synthetic chromosome. Each tRNA gene was flanked with rox recombination sites (yellow diamonds). The detailed anatomy of these arrays is shown in Figure S1.

(D) Fitness assays for draft syn6.5 strain, YZY1178, following consolidation. Spots are from yeast media with 10-fold dilutions.

(E) Pulsed-field gel electrophoresis (PFGE) to evaluate electrophoretic karyotype of a syn6.5 strain. See also Figure S1.

synthetic chromosomes (*synII*, *synIII*, *synV*, *synVI*, *synIXR*, *synX*, and *synXII*), representing all synthetic chromosomes assembled and published prior to the publication of the current set of Sc2.0 papers (Figure 1A). In this strain, ~31% of the genome is synthetic, and thus it carries thousands of designer features (Figure 1B).

As an interim solution for fitness defects caused by tRNA deletions prior to neochromosome introduction, we designed tRNA arrays (Figure 1C), thus maintaining tRNA gene balance as additional synthetic chromosomes are incorporated. Briefly, each tRNA array was released from its host plasmid and integrated using homologous recombination (Figure S1). All tRNA genes were also flanked by a pair of rox sites, which can be recognized by Dre (but not Cre) recombinase, enabling a chromosomal tRNA-specific rearrangement system.

The "draft" syn6.5 strain grows slightly slower on rich medium (YPD) but grows comparably on plates with non-fermentable glycerol (YPG) (Figure 1D). Unlike the parent strains, the syn6.5 strain also shows an obvious growth defect at 37°C, suggesting the existence of a "combinatorial bug" resulting from genetic interactions between designer variants introduced on more than one synthetic chromosome, analogous to the phenomenon of synthetic lethality/fitness. The karyotype was confirmed using pulsed-field gel analysis (PFGE), with synthetic chromosomes showing expected faster migration due to their shorter lengths (Figure 1E). Each chromosome has euploid genome coverage in whole-genome sequencing (WGS) (Figure S1). No mutations or genome arrangements appeared during consolidation when compared with the original parental strains with single synthetic chromosomes, suggesting that mutation rates are not elevated in the presence of multiple synthetic chromosomes.

Mapping fitness defects using CRISPR D-BUGS

With thousands of modifications, growth defects resulting from designer bugs have been observed in most synthetic chromosomes.[26] Identifying these bugs is important for restoring cell fitness, and understanding their mechanisms may illuminate new biological insights. We developed a systematic and efficient bug-mapping strategy exploiting loss of heterozygosity (LOH) in diploids called CRISPR Directed Biallelic *URA3*-assisted Genome Scan, or CRISPR D-BUGS. CRISPR D-BUGS exploits heterozygous diploid strains bearing a synthetic chromosome and a native chromosome in which one telomere bears a *URA3* marker gene. In such diploid strains, homologous recombination between two chromatids can be enhanced by a targeted chromosomal double-strand break (Figure 2A).[41] After cell division, daughter cells will carry a pair of chromosomes that are homozygous for synthetic DNA from the recombination site to the telomere region but retain heterozygosity in the remainder of the chromosome, and these LOH events can be readily selected for by plating on 5-FOA (5-fluoro-orotic acid). By using gRNAs that target different PCRtag sequences, a series of yeast strains with various homozygous synthetic regions can be generated (Figure 2B). We checked the fitness of such strains and subsequently mapped the "fitness boundary" at which derivative strains shift phenotypically from unhealthy to healthy under specific conditions, provided the bug is recessive.

Absent other mapping information, screening is begun within the resolution of a chromosome arm. Subsequently, the search continues with a series of gRNAs to map the bug more precisely.

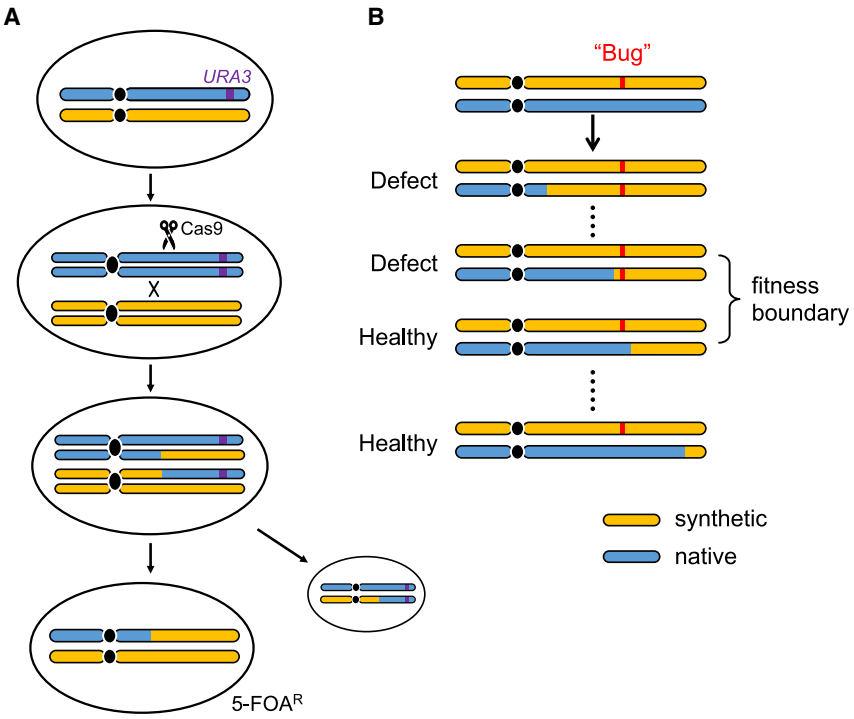


Figure 2. Fitness mapping using CRISPR D-BUGS

(A) General outline. A *URA3* marker is integrated into native allele (blue), which is cleaved by Cas9 targeted by an sgRNA selected specifically to cut at one wild-type PCRtag. Following mitotic recombination, strains homozygous for the synthetic region (orange) are selected on 5-FOA.

(B) Series of strains with different synthetic regions generated by gRNAs targeting different wild-type PCRtag loci allows mapping of fitness boundary. See also Figure S2.

Resolution can be increased with two or more rounds of mapping until a group of colonies generated from the same single gRNA shows two distinct levels of fitness. This variability results from mitotic recombination occurring within a ~10-kb window from the cleavage site, and such a region may include multiple designer features, such as PCRtags or loxPsym insertions.[41] Using WGS, the bug is mapped at high resolution by defining the locations of the synthetic genome modifications within the recombination interval. Using the same principles, CRISPR D-BUGS can be utilized to map dominant bugs, in which a single-allele modification results in a defect in a heterozygous diploid strain bearing both synthetic and natural chromosomes (Figure S2).

To test this approach, we first tried to map a perplexing bug on a previously synthesized chromosome. The original *synII* strain (chr02_9_03) showed a recessive growth defect on YPG medium at 37°C (Figure 3A), which appeared after megachunk X was assembled.[24] To map the *synII* bug using CRISPR D-BUGS, we constructed a *synII*/+ heterozygote and selected gRNAs targeting synthetic PCRtags within megachunk X. The colonies generated from gRNA targeting at *YBR256C* PCRtags (called gRNA.YBR256C) and gRNA.YBR261C all showed the defect, whereas colonies generated from gRNA.YBR270C and gRNA.YBR275C were all healthy on YPG medium at 37°C (Figures 3B and S3). This mapped the bug between *YBR261C* and *YBR270C*. In a second round of bug mapping, single colonies generated using gRNA.YBR265W showed a mixture of two fitness levels. Using WGS, the recombination interval of each colony was precisely mapped relative to synthetic sequence variants, linking each variant to strain fitness (Figure 3C). This strategy helped identify two adjacent loxPsym sites between *YBR263W* (*SHM1*) and *YBR264C* (*YPT10*) as potentially responsible for the fitness defect. Deleting both loxPsym sites successfully restored strain fitness, as in strain YZY166 (chr02_9_04) (Figure 3A).

The loxPsym sites were integrated 3 bp downstream of the stop codon of *SHM1* and *YPT10* (Figure S3B). *SHM1* encodes the mitochondrial serine hydroxymethyltransferase, and its deletion results in impaired respiratory function, consistent with the observed *synII* fitness defect on YPG.[46] By contrast, *ypt10* deletion showed no fitness defects under various conditions including different temperatures or carbon sources.[47] These genes are convergent and closely spaced, such that integration of loxPsym sites produces *SHM1* transcripts containing two loxPsym sequences in their 3′ UTRs. These two loxPsym sequences are predicted to form a stem-loop structure in the *SHM1* 3′ UTR, which we hypothesize may affect mRNA stability (Figure S3). Consistent with this hypothesis, deletion of both or either loxPsym site(s) significantly recovered transcript abundance and successfully rescued the growth defect, strongly implying that formation of a stem loop in the RNA leads to loss of RNA abundance and the fitness defect (Figures 3D and S3E). We also found that *Shm1p* level was reduced in the presence of two loxPsym sites and recovered upon their removal (Figure S3F).

Since the two loxPsym sites are located in the 3′ UTR, we also wondered about their effects on transcript properties. To answer this, we used nanopore direct RNA sequencing to evaluate full-length native transcripts of *SHM1* and *YPT10* directly and measured transcript end site (TES) distributions.[48] In the original *synII* (chr02_9_03) strain, the majority of *SHM1* TESs were extended by ~66 nt, matching length of two transcribed loxPsym sites (68 nt), indicating that *SHM1* transcript termini were not significantly affected and were extended by the expected length (Figure 3E). There were also around 10% transcripts extended by ~160 nt, forming a second peak specific to *synII*, suggesting that transcription termination could be slightly affected by the loxPsym sequences. The removal of both loxPsym sites (chr02_9_04) successfully recovered the TES distribution, overlapping with the wild-type peak. For the *SHM1* TES in YZY363 and YZY374, in which the individual loxPsym sites were deleted,

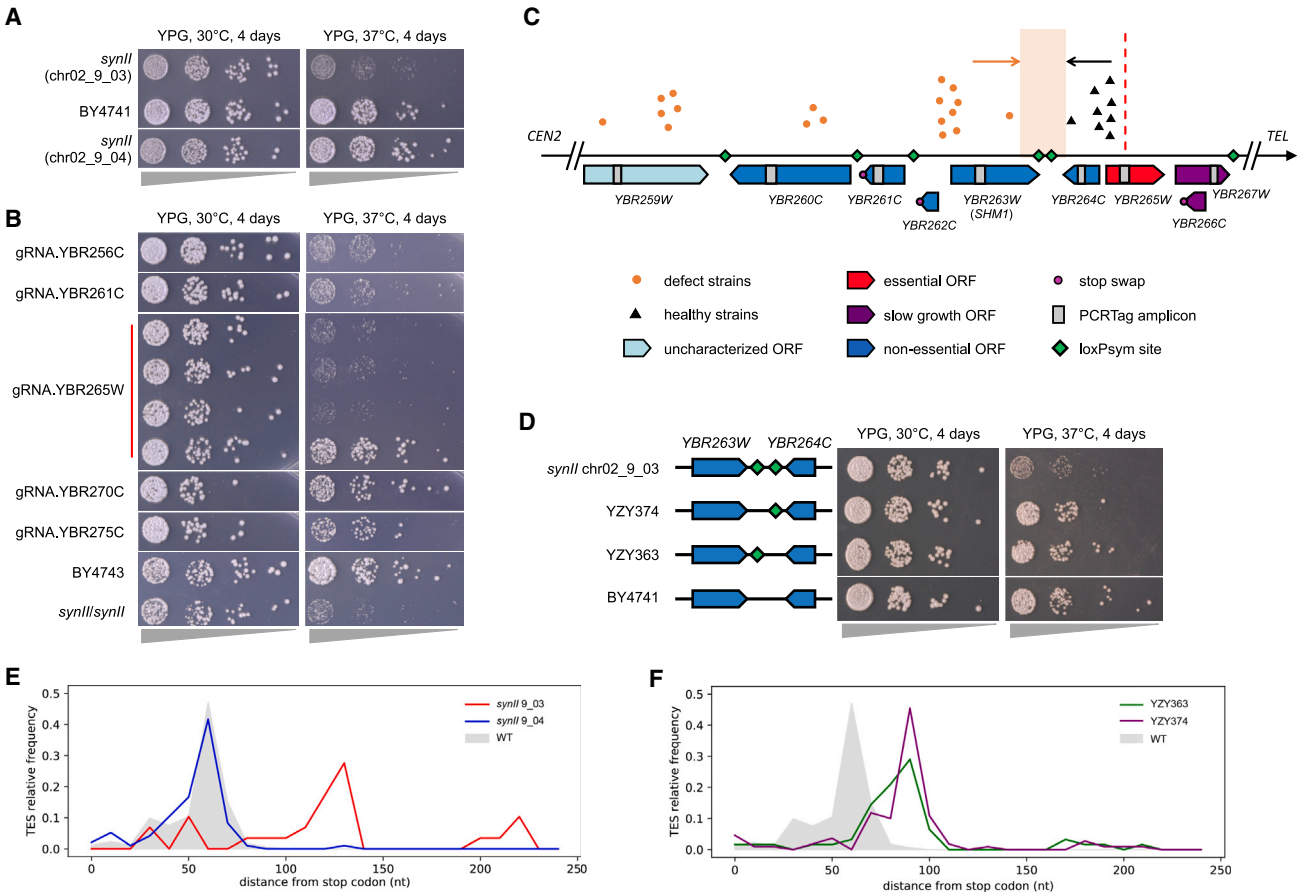


Figure 3. CRISPR D-BUGS mapping in *synII*

(A) Fitness assay on YPG plates for a strain with the original *synII* (chr02_9_03). In strain YZY166 (chr02_9_04), the bug was fixed by deleting two loxPsym sites downstream of *SHM1*.

(B) The CRISPR D-BUGS colonies generated using gRNAs labeled on the left side. For each gRNA, at least four colonies were tested and showed the same fitness except gRNA.YBR265W (additional colonies in Figure S3).

(C) Recombination sites in gRNA.YBR265W colonies indicating fitness level are aligned with *synII* designer features. Red dashed line indicates locus in *synII* corresponding to cleavage site in native counterpart. Original fitness assay is shown in Figure S3.

(D) Fitness assay for the strains deleted with either loxPsym site.

(E) Transcript end site (TES) distributions of *SHM1* transcripts from original *synII* (red) and updated *synII* (blue), compared with wild type (gray).

(F) The same measurements in strains with either loxPsym site deleted (YZY363 and YZY374 as in D).

See also Figure S3.

a single peak was formed and extended by ~32 nt, matching the expected length of a single loxPsym site (Figure 3F). We also measured the *YPT10* TESs in these strains and observed similar patterns (Figure S3). In summary, the incorporation of two loxPsym sites, presumably capable of forming a stem loop in the 3′ UTRs, primarily impacted the expression levels but not the transcription end sites for *SHM1* and *YPT10*.

Mapping a *synI* bug to an unexpected deletion

CRISPR D-BUGS was also applied to map the growth defect of a *synI* strain. A special feature of *synI* is that it is fused with *synIII*.[29] The draft strain (yJL671, chr01_9_02) showed a recessive growth defect even on rich medium (YPD), which is not caused by chromosome fusion (Figure S4). Using CRISPR D-BUGS, we successfully mapped the bug to a window of 5 ORFs between *YAL055W* and *YAL049C* in a single round of mapping (Figure S4C). Using WGS, we found an unexpected ~10-kb deletion in yJL671, arising from unequal recombination between the two loxPsym sites flanking *YAL062W* and *YAL054C*, which became homozygous in the low-fitness diploids, but remained heterozygous in healthy strains (Figures 4A and S4E). A review of *synI*'s assembly history revealed that this deletion resulted from an unexpected off-target recombination between two loxPsym sites during CRISPR-mediated repair of a missense mutation in strain yJL663, which contained an earlier draft *synI* version (Figure 4B). By repairing this deletion using SpCas9-NG, a final healthy strain with the complete *synI* sequence was obtained as yCTC002 (Figures 4C and S4).[49] In summary, CRISPR D-BUGS was used to quickly map distinct bug types in *synI* and *synII*.

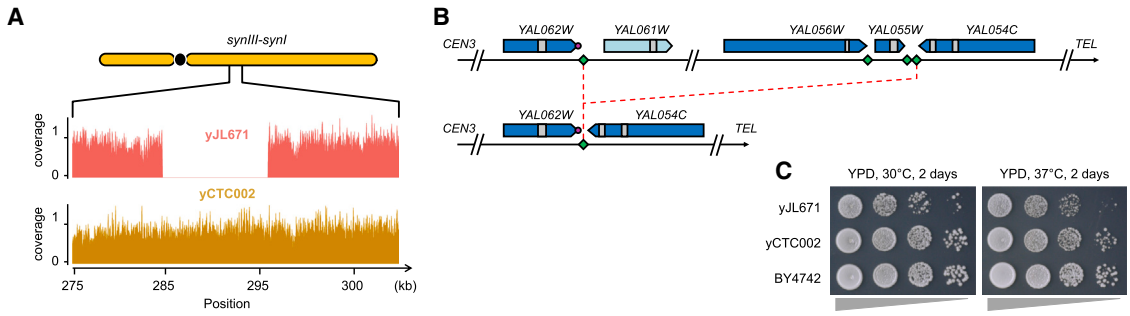


Figure 4. CRISPR-D-BUGS mapping in *synIII-synI* fusion chromosomes

(A) Deletion detected in *synIII-synI* strain (yJL671), which was repaired in final version (yCTC002). Fitness assays are shown in Figure S4.

(B) The diagram of *synI* in design (upper) and actual strain of yJL671 (bottom).

(C) Fitness assay for final *synI* strain (yCTC002) with bug repaired.

See also Figure S4.

A combinatorial bug associated with *synIII* and *synX*

Although strains with single synthetic chromosomes are healthy, combinatorial defects can still occur due to combinations of sequence changes that by themselves have no phenotypes, owing to genetic interactions between variants on two (or theoretically more than two) synthetic chromosomes. Although strains containing *synIII* (chr03_9_02) and *synX* (chr10_9_01) alone are healthy, a combinatorial 37°C defect in a strain containing both *synIII* and *synX* and no other synthetic chromosomes (Figure 5A). We first used CRISPR D-BUGS to map the bug in *synX* (Figure S5). In the first round of "rough" mapping, we successfully mapped a fitness boundary between *YJL097W* and *YJL210W* at the left arm. Fine mapping revealed that single colonies generated from a gRNA targeting *YJL176C* showed mixed fitness levels. WGS mapped the bug to the loxPsym site integrated downstream of *YJL175W*, representing the boundary separating healthy and temperature-sensitive strains (Figure 5B).

YJL175W is a "dubious ORF," and overlaps the 5′ end of *YJL176C* (*SWI3*), an important named gene. The loxPsym inserted into the 3′ UTR of *YJL175W* is transcribed as a part of the *SWI3* 5′ UTR (Figure 5C). Consequently, there are two loxPsym sequences in the transcript of *SWI3*, which might therefore form a looped secondary RNA structure. Interestingly, the *synX SWI3* transcript level was increased 3-fold, and this transcriptional phenotype was restored to wild-type levels by deleting the 5′ UTR loxPsym but not by deleting the 3′ UTR loxPsym (Figure 5D). Paradoxically, the Swi3p level was reduced in the presence of 5′ UTR loxPsym and restored upon its removal (Figure 5E). The most parsimonious explanation for these results is that the inverted repeat within the loxPsym in the 5′ UTR forms a stem loop that stabilizes the RNA isoform and blocks translation. Insertion of the loxPsym in the 3′ UTR, as in all other nonessential Sc2.0 genes, had minimal to no effect on protein levels. The temperature-sensitive strain containing both *synIII* and *synX* (but not the parental strains) is consistent with temperature sensitivity of *swi3* null alleles.[50] As an essential component of SWI/SNF chromatin remodeling complex, Swi3p is required for transcription of many genes, including *INO1*, and *swi3* null mutants are viable but auxotrophic for inositol.[51–53] Remarkably, the *synX* strain displayed partial inositol auxotrophy, largely restored by removing the 5′ UTR loxPsym site (but not the 3′ UTR loxPsym site; Figure S5), consistent with the proposed mechanism. A similar pattern was observed in a *synVII* bug (Figure 5C), wherein a similarly "misplaced" loxPsym site in the *NSR1* 5′ UTR led to increased mRNA but dramatically reduced protein level.[31] As in the case above, deletion of the 5′ loxPsym site restored normal transcript and protein levels and rescued the *synVII* growth defect.

Following the same principles, we initially mapped the bug in *synIII* (chr03_9_02) with CRISPR D-BUGS to the right arm, then roughly between *YCR057C*/*YCR067C* (Figure S6A), and finally fine-mapped it to two loxPsym sites between *YCR061W* and *YCR065W* (Figure 5F). By restoring them to wild type, we found that the loxPsym site upstream of *YCR063W* caused the defect (Figure S6B). This loxPsym site marked the deletion of *SUP61*, an essential single copy $tRNA_{Ser}^{CGA}$ gene, which decodes the rare UCG serine codon (Figure 5G). Importantly, unlike the examples above, this loxPsym site was not embedded inside any transcribed region, suggesting that it might rather be deletion of the tRNA that was responsible for the defect. Before introducing the complete tRNA neochromosome, all strains containing *synIII* have a synthetic version of *SUP61* (*synSUP61*) integrated in the *HO* locus on *chrIV* to temporarily provide its essential function. By itself, *synSUP61* suffices for cell survival and health.[23] Like all of the synthetic tRNAs in Sc2.0, *synSUP61* is flanked by 500 bp 5′ and 40 bp 3′ of *Ashbya gossypii* tRNA flanking sequences and has a precise intron deletion. Introducing a single copy of *SUP61* in the strain complemented the defect (Figure S6B), suggesting that *synSUP61* is too lowly expressed or otherwise incapable of providing full functionality. To test this hypothesis, we examined expression of *synSUP61* by northern blotting and observed that it produced about half the normal amount of tRNA and a surprisingly large amount of 5′ pre-tRNA, suggesting inefficient processing of this tRNA, relative to *SUP61* (Figure 5H). This appears to be associated with replacing the tRNA 5′ flanking region with the *Ashbya* sequence in *synSUP61* and not from intron deletion or the 3′ flanking region swap (Figure S6C).

As both parental *synIII* and *synX* strains are healthy, we conclude that the combinatorial bug results from an unexpected

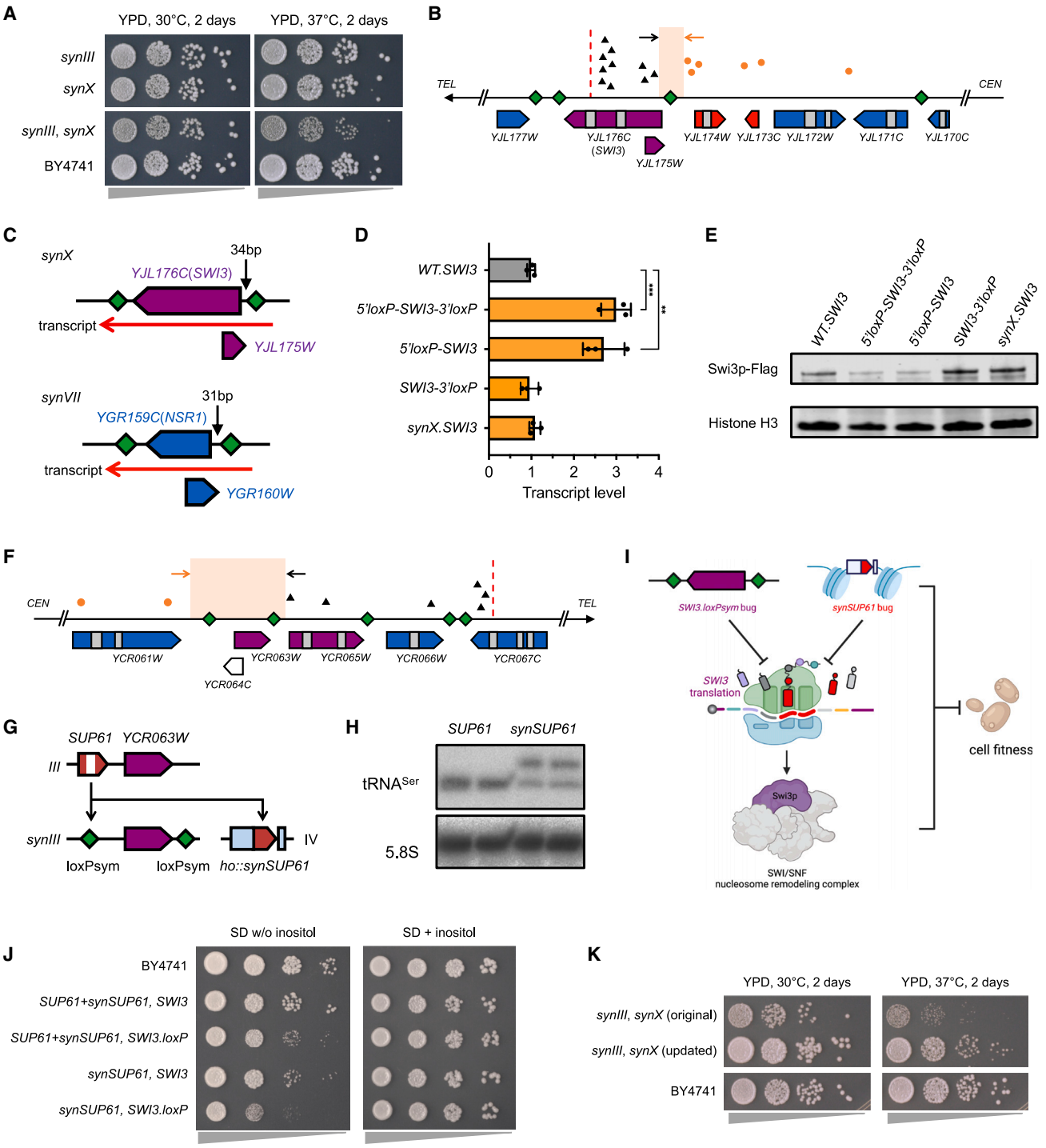


Figure 5. Combinatorial defect between *synIII* and *synX*

(A) Fitness assay showing a combinatorial defect in *synIII*, *synX* context.

(B) Recombination sites in gRNA.YJL176C colonies indicating fitness level were aligned to *synX* designer features (labels as in Figure 3C). Fitness assays are shown in Figure S5.

(C) Diagram of *YJL176C* (*SWI3*) loxPsym pattern, compared with *YGR159C* (*NSR1*) from *synVII*.[31]

(D) *SWI3* transcript levels in wild-type background (gray bar) and *synX* strains (orange bars) with both loxPsym sites (*5′loxP-SWI3-3′loxP*), 3′ loxPsym deleted (*5′loxp-SWI3*), 5′ loxPsym deleted (*SWI3-3′loxP*), and no loxPsym site (*synX.SWI3*). Error bars represent the mean ± SD of three replicates. $p < 0.001$ (***), or 0.01 (**).

(E) Immunoblotting of Swi3p-Flag in strains with loxPsym deleted from 5′ and/or 3′ UTR.

(legend continued on next page)

interaction between *synSUP61* and *SWI3*, both of which appear to be under-expressed relative to native counterparts. The single copy essential *SUP61* gene produces the only tRNA decoding the rare UCG serine codon. Interestingly, the *SWI3* transcript has above-average usage of UCG for serine, and importantly, it includes two tandem UCG codons (Figure S6D). Tandem rare codons can cause translational pausing or even arrest due to "starvation" for charged cognate tRNAs.[54–56] Based on this, we hypothesized that reduced abundance of $tRNA_{Ser}^{CGA}$ reduces translation of *SWI3* below the already lower than normal level caused by 5′ UTR ectopic loxPsym site. This is predicted to result in an even lower level of functional SWI/SNF complex and the resulting pronounced growth defect (Figure 5I). To test this hypothesis, we repaired either or both bugs and measured inositol auxotrophy (Figure 5J). Interestingly, either deletion of the *SWI3* loxPsym site or addition of *SUP61* individually were able to partially rescue auxotrophy, suggesting that ultimately, the observed phenotypes result from low *Swi3* protein levels. When the two "buggy" components were both restored to their native forms, the fitness of the strain was successfully rescued. To further probe this combinatorial interaction mechanism, we mutated either of the *SWI3* tandem serine codons from rare UCG to common UCU, with *SWI3* loxPsym site deleted (Figure S6E). Consistently, strains with either mutation showed significantly improved growth on plates without inositol. Notably, the inositol auxotroph was still not completely rescued. Similar results were also observed in *synX* strain, in which the removal of the 5′ *SWI3* loxP site significantly but not completely rescued inositol auxotrophy, suggesting that additional synthetic modifications may also affect inositol biosynthesis (Figure S6F).

By repairing the *SWI3* and *synSUP61* bugs, the fitness of the *synIII*, *synX* strain was largely rescued at both 30°C and high temperature (Figure 5K). For multiple synthetic chromosomes (*synII, synIII, synV, synVI, synIXR, synX, synXII*), we repaired all known bugs, including the *SHM1* bug in *synII* and the combinatorial bug between *SWI3* in *synX* and *synSUP61* in *synIII*. As expected, the growth defect was dramatically improved, albeit with minor residual growth defects at 37°C (Figure S6G).

Characterization of multiple synthetic chromosomes in syn6.5 strains

We are curious about how multiple synthetic chromosomes would affect genome organization and whether the large numbers of designer features would affect transcription. Therefore, we used genomic chromosome conformation capture (Hi-C) to investigate the nuclear organization of all 6.5 synthetic chromosomes (STAR Methods).[57,58] Previous work has shown that the Sc2.0 design improved mappability as a result of deletion of repetitive regions, especially Ty elements.[58] In our strain with 6.5 synthetic chromosomes, we calculated the spatial contact frequency and generated a heatmap of genomic interactions (Figure 6A), generating a 3D map showing predicted conformations of the synthetic and native yeast chromosomes (Figures 6B and S7A). Similar to the wild type, all centromeres of synthetic and native chromosomes interacted *in trans* due to their tethering near the spindle pole body (SPB, the microtubule organizing center of yeast), as well as their telomeres, which are clustered at the nuclear envelope.[59] To detect differences in internal folding of the chromatin, we calculated decay of contact frequency or probability (p) as a function of the genomic distance (STAR Methods). The two plots are nearly overlapping, with the exception of long distances greater than 100 kb because of ever-increasing noise, resulting from the overall short chromosomal arms of yeast as previously described.[58] We conclude that genetic modifications spanning approximately half of the entire yeast genome do not cause differences in the overall structure of the chromosomes (Figure 6C). These results indicated that Sc2.0 designer modifications have minor effects on chromosomal organization, even when multiple synthetic chromosomes are combined.

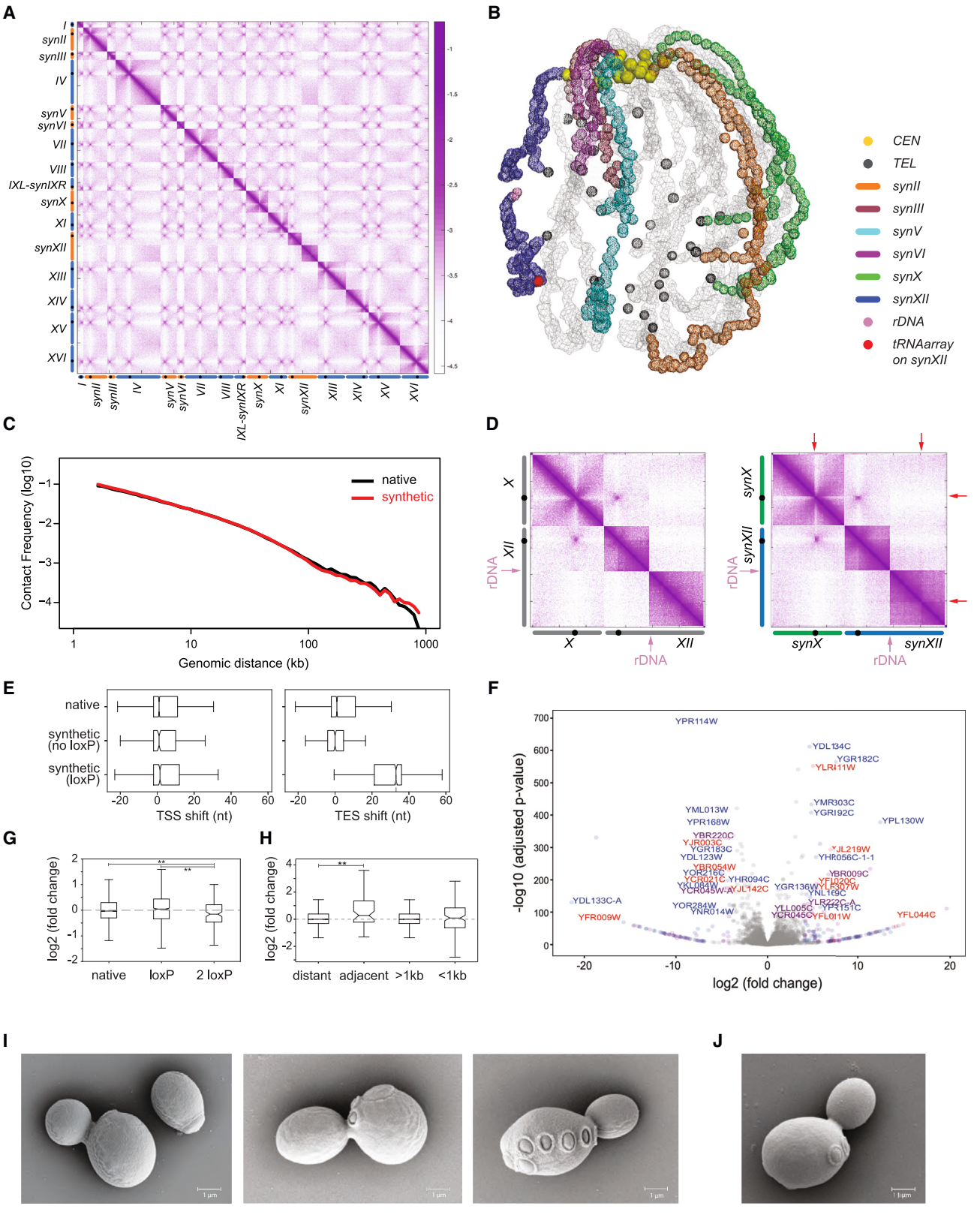
For several synthetic chromosomes (*synII*, *synIII*, *synX*, and *synXII*), we noted a sharp boundary formed at one position in their contact frequency maps (Figures 6D and S7B). These boundaries exactly match locations of the tRNA arrays, such as the *synX* tRNA array integrated on the left arm close to *CEN10* (~12 kb) and the *synXII* tRNA array distal from *CEN12* (~609 kb) and downstream of the rDNA locus (~319 kb). Due to the repeated nature of each array, exclusively integrated in synthetic chromosomes, these unmappable structures are invisible to Hi-C, resulting in the formation of artifactual sharp boundaries formed by juxtaposing their upstream and downstream flanking sequences.

Transcript profiling using RNAseq

To determine whether transcript boundaries were affected by the incorporation of synthetic design features, we mapped transcript isoforms from the syn6.5 strain using nanopore long-read direct RNA sequencing. As expected, transcript start sites were not affected by the inclusion of 3′ loxPsym sites. Neither transcripts arising from genes on native chromosomes nor those without flanking loxPsym sites on synthetic chromosomes showed end site alterations; however, addition of loxPsym sites at the 3′ end of genes increased the length of their transcripts by 34 nt on average (Figure 6E). This is consistent with incorporation of the loxPsym site into the transcript without altering its cleavage/polyadenylation site.

To assess the effects of synthetic genome design on gene expression levels, we performed stranded mRNA sequencing. Some genes on the native and synthetic chromosomes, both with and without flanking loxPsym sites, showed significantly

(F) For *synIII* bug mapping, recombination sites in gRNA.YCR067C single colonies were aligned to *synIII* designer features. Fitness assays are shown in Figure S6.
(G) Relocation of *synSUP61* to *HO*, and gray blocks indicate flanking *Ashbya gossypii* sequences. White band indicates *SUP61* intron removed in the synthetic version.
(H) Northern blot to check quality and level of $tRNA^{Ser}$ expressed from native and synthetic *SUP61*.
(I) Proposed combinatorial interactions between *synSUP61* bug and *SWI3.loxP* bug.
(J) Inositol auxotrophy analysis, with wild-type *SUP61* integrated and/or *SWI3* bug repaired.
(K) Fitness assay of final *synIII, synX* strain with both bugs fixed.
See also Figures S5 and S6.



(legend on next page)

altered expression levels (Figure 6F), indicating that the synthetic genome design caused some, but not widespread, indirect effects on gene expression levels. LoxPsym-flanked genes were not significantly affected compared with genes on the native chromosomes; however, genes with transcripts that incorporate two loxPsym sites within their 3′ UTRs tended to experience a slight decrease in transcript abundance (Figure 6G), potentially indicating decreased stability of these transcripts on average. This was consistent with the *synII* growth defect caused by two tandem loxPsym sites in the *SHM1* 3′ UTR.

The relocation of tRNAs led to a major alteration in the 3D organization of the synthetic chromosomes (Figure 6D). We therefore compared expression of genes adjacent to tRNAs on the native and synthetic chromosomes and saw that the removal of tRNA genes in the synthetic genome appeared to be associated with increased expression of their former neighbors (Figure 6H), consistent with previous studies of tRNA gene mediated silencing.[33,60–62] This observation did not hold true for slightly more distant genes (<1 kb), suggesting that tRNA expression only affects the most proximal genes. The removal of introns from genes in the synthetic genome also did not appear to greatly affect their expression levels (Figure S7C). Overall, the design features of the syn6.5 genome appear to have only modest effects on transcript isoform boundaries and expression levels.

Morphology of yeast cells with syn6.5

To evaluate the cell morphology of yeast strains with multiple synthetic chromosomes, we visualized dividing syn6.5 cells with scanning electron microscopy (Figures 6I and 6J). They showed active multiplication and normal cell morphology, length, and shape. The budding of daughter cells left ring-shaped bud scars on mother cells. We observed cells at various stage of budding, including an aged mother cell with seven bud scars, still actively budding. Thus, rewriting multiple chromosomes does not appear to markedly affect cell morphology or cellular replicative lifespan.

Mutation frequency

Null mutants in the *CAN1* gene result in resistance to canavanine. To compare forward mutation frequencies of syn6.5 and wild-type BY4742 strains, we roughly measured the Can^R frequency and found that the syn6.5 strain showed a Can^R frequency of 5.5×10^{-6}, perhaps modestly higher (ns, $p = 0.06$) than wild-type BY4742 (4.5×10^{-6}) (Table S2). Our main conclusion from this is that the semisynthetic 6.5 strain does not have a significantly elevated mutation rate. Further experiments will be needed to more accurately measure the mutation rate and investigate the underlying mechanisms.

Transferring *synIV* into synthetic chromosome strains using chromosome substitution

Currently, all yeast chromosomes have been synthesized separately in their individual host strains. In order to consolidate them more efficiently in the current syn6.5 strain, we deployed a next-generation consolidation strategy developed in our laboratory: chromosome substitution.[33] We hoped to develop a method to directly transfer individual chromosomes to a recipient haploid strain that already carries multiple synthetic chromosomes. In yeast, *kar1-1* or *kar1-Δ15* mutations prevent nuclear fusion during mating when they are present in either parent.[63] Thus, most progeny of these crosses remain haploid but have a mixed cytoplasm. In these cells, chromosomes are occasionally transferred from the donor to the recipient strain. This process, called "exceptional cytoduction[64]" or chromoduction,[65] results in an n+1 cell that can be selected for using proper auxotrophic and drug-resistance markers.

Based on this phenomenon, a chromosome substitution method was developed entailing two steps: (1) introduction of the chromosome of interest into a recipient strain by chromoduction, resulting in an n+1 strain, and (2) destabilization of the native chromosome by inducing transcription through its centromere in the n+1 strain.[33] To demonstrate how chromosome substitution could be deployed in synthetic chromosome consolidation, we picked the largest synthetic chromosome, *synIV*, as a "worst-case scenario" proof of principle (Figure 7). The efficiency of chromoduction is inversely correlated with the chromosome size, and using this approach, each subsequent smaller synthetic chromosome substitution is predicted to be even more efficient.[64]

First, in the recipient strain (syn6.5 as described above), we sequentially introduced (1) a *can1Δ0* deletion, (2) a *cyh2*

Figure 6. Characterization of the strain with multiple synthetic chromosomes

(A) The heatmap shows the contact probability (log_{10}) between pairs of chromosomal sites. The corresponding chromosomes were labeled as horizontal and vertical axes.

(B) The 3D chromosome trajectories of multiple synthetic chromosomes. Gray, all other native chromosomes. The 3D structures are shown as a movie in Data S1.

(C) Contact frequencies as a function of the genomic distance.

(D) Heatmaps for native (left) and synthetic (right) chromosomes *X* and *XII*. The tRNA array integration loci are highlighted with red arrows.

(E) Change in distribution of transcript start sites (TSSs) and transcript end sites (TESs) in transcript isoforms arising from genes on the native chromosomes (native, n = 3,590) and synthetic chromosomes. Synthetic, loxP (n = 1,271) and synthetic, no loxP, (n = 690) indicate the genes on synthetic chromosomes with, or without 3′ loxPsym sites within 300 bp of their stop codons, respectively.

(F) Volcano plot of gene expression in the multiple synthetic chromosome strain, compared with wild type. Lists of significantly up- and downregulated genes are presented in Table S3. Blue: genes on native chromosomes. Purple: genes on synthetic chromosomes. Red: genes with loxPsym site incorporated.

(G) The expression levels of genes with two loxPsym sites in their 3′ UTRs (n = 131) compared with genes on native chromosomes and genes incorporated with only one loxPsym site (n = 1,165). Error bars represent the mean ± SD. $p < 0.01$ (**).

(H) Comparison of expression level changes in the synthetic strain, in which tRNAs are relocated, for the genes closest to native tRNAs (adjacent, n = 62) and all other genes (distant, n = 8,380), as well as genes either >1kb (n = 8,317) or <1kb (n = 130) away from native tRNAs. The adjacent genes are listed in Table S4. **$p < 0.01$ determined by Mann-Whitney U tests.

(I) SEM pictures of single yeast cells with multiple synthetic chromosomes, compared with wild-type cells as in (J).

See also Figure S7.

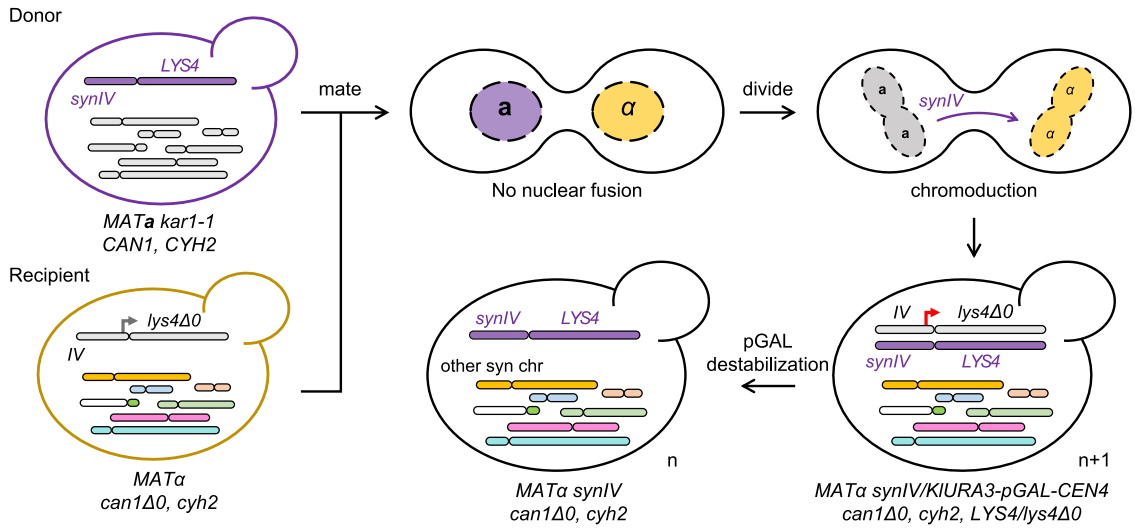


Figure 7. Consolidation of *synIV* into syn6.5 strains using chromosome substitution

Donor: A strain carrying a synthetic chromosome(s) to be introduced with a selectable marker (*synIV* and *LYS4* in this example). Recipient: the strain of the opposite mating type already containing one or more multiple synthetic chromosomes but retaining the native counterpart of the target (native *chrIV* and *lys4Δ0* here), which is tagged with *KlURA3* and pGAL-CEN (black hooked arrow). Transfer of *synIV* by chromoduction into recipient haploid progeny can be selected using the appropriate auxotrophic and drug-resistance markers. Finally, native *chrIV* is destabilized by induction of the pGAL promoter (red hooked arrow), which can be selected as 5-FOA^R due to the loss of *KlURA3*, completing the process of chromosome substitution.

mutation, and (3) a *lys4Δ0* deletion in *chrIV* using CRISPR/Cas9, making the strain resistant to canavanine and cycloheximide and auxotrophic for lysine, respectively (strain YZY402). The *can1Δ0* and *cyh2* markers are used to select against non-mating donor strains and occasionally formed diploid zygotes. The *lys4Δ0* can be complemented by *LYS4*$^+$ in *synIV* from the donor strain. Next, the *kar1-1* mutation was introduced using CRISPR/Cas12a. After mating and selection for chromoductants, haploid progeny with both native and synthetic chromosome *IV* were successfully obtained. Finally, native *chrIV* was destabilized upon induction of the pGAL promoter and selection for 5-FOA^R.

This transfer produced a yeast strain with more than half of its genome synthetic. This strain (*synII*, *synIII*, *synIV*, *synV*, *synVI*, *synIXR*, *synX*, and *synXII*) was confirmed as a haploid synthetic strain by WGS and flow cytometry (Figures S7D and S7E). It was characterized by slower growth, with a G1 arrest, suggesting a cell cycle defect, and will require further bug mapping to generate a high-fitness derivative. Combined with CRISPR D-BUGS, chromosome substitution is a NextGen readily scalable strategy to consolidate and debug incoming synthetic chromosomes into the genetic background of other synthetic ones.

DISCUSSION

The goal of Sc2.0 is to build a eukaryotic organism with a completely redesigned and human-synthesized genome. Thousands of genome-wide features facilitate a variety of applications such as genome minimization, non-standard amino acid incorporation, and SCRaMbLE. As an inducible evolution system, SCRaMbLE has been used to exploit genomic structural variation, biosynthesis pathway engineering, and host strain improvement.[15–19,66–68] Recently, synthetic chromosomes and SCRaMbLE have been used together with deep transcript isoform profiling to study genome architecture and its contribution to transcriptional regulation.[22,48] Genome-wide SCRaMbLE and rearrangements also provide insights into 3D spatial organization and chromatin accessibility.[20,22] With more synthetic chromosomes incorporated, SCRaMbLE can also be deployed to study additional biological questions such as yeast phenotypes and minimized chromosomes and genomes,[69] making Sc2.0 a platform to understand eukaryotic genomes and develop industrial applications.

Native *S. cerevisiae* contains ~295 introns belonging to 280 protein-coding genes,[70] and 91 of these introns are absent from syn6.5 strains. Previous studies have demonstrated that intron accumulation might aid yeast starvation response by sequestering available splicing factors and affecting splicing of RNAs encoded by other genes.[70,71] Most recently, it was shown that introns may also lead to improved phenotypic plasticity of ribosomes with a fitness advantage.[72] As splicing machinery may be more available in the syn6.5 strain, it would be interesting to evaluate the impact reduced intron content may have on splice isoforms and ribosome regulation in the future.

The introduction of these designer edits also results in unexpected bugs in the form of a wide variety of growth defects. Interestingly, several bugs found in this and companion studies map to loxPsym sites introduced downstream of what are now classified as dubious ORFs, which end up damaging promoters and 5′ UTRs of authentic genes. Thus, this is a type of bug that results from the rules of ORF annotations, which were adopted early in the Sc2.0 project at a time when dubious ORFs were not yet well defined and annotated.

This debugging method will also be helpful during further consolidation of additional synthetic chromosomes. Using

CRISPR D-BUGS, we first mapped a *synII* growth defect to two repetitive loxPsym sites in the *SHM1* 3′ UTR, which led to reduced expression by an unknown mechanism. *SHM1* encodes mitochondrial serine hydroxymethyltransferase (SHMT), an enzyme that can interconvert serine and glycine and produce 5,10 methylene tetrahydrofolates (CH2-THF). Shm1p only comprises about 5% of total SHMT activity, whereas its cytosolic isoform Shm2p comprises the majority.[73] Shm1p and Shm2p are conserved from bacteria to humans, and both enzymes are essential components in the one-carbon metabolism cycle, which provides a crucial substrate for mitochondria initiator tRNA formylation and other reactions required for biosynthesis of nucleotides, amino acids, and lipids.[74,75] The *shm1Δ* strain was shown to have a near-wild-type level of formylation and mitochondrial protein expression, suggesting that cytosolic SHMT activity produces a sufficient supply of one carbon units in the absence of Shm1p.[46]

A further clue about this came from observing that an extra copy of *TSC10* suppressed the *synII* growth phenotype. We found that overexpressing *TSC10* either from a plasmid or via ectopic integration at the *HO* locus rescued the growth defect of *synII* (Figures S7F and S7G). *TSC10* is an essential gene, encoding a 3-ketosphinganine reductase catalyzing the second step in phytosphingosine synthesis, using serine as the key precursor molecule.[76] This pathway is the basis for all sphingolipids made in yeast, including ceramide, a major component of the mitochondrial membrane. Interestingly, this pathway is highly involved in heat stress response and activated immediately upon onset of heat stress.[77] Several sphingolipids were also implicated as secondary messengers involved in signaling pathways that regulate the heat stress response.[78] Based on this, we speculate that in the original *synII* strain (chr02_9_03), sphingolipid/ceramide biosynthesis was affected when grown on YPG medium due to a deficit of cytosolic serine. This is supported by the fact that serine is mainly synthesized from glycine via SHMT activity of Shm1p and Shm2p in a "gluconeogenic" pathway on non-fermentable glycerol but not on glucose.[79,80] In short, we hypothesize that *SHM1* under-expression leads to insufficient ceramide biosynthesis, and at high temperature, this reduces mitochondrial function.

During consolidation, we noted the existence of a combinatorial bug revealed when combining *synIII* and *synX*. tRNA abundance and codon usage are closely linked, with "rich" codons corresponding to abundant tRNAs overrepresented in highly expressed genes.[81,82] Optimal codon usage is also predicted to ensure the proper speed of translation elongation for efficiency, accuracy, and correct protein folding.[56,83] Under stressful conditions, yeast can alter tRNA abundance to facilitate selective stress-related translation, and interestingly, rare codons corresponding to low-abundance tRNAs are enriched in stress-responsive genes potentially allowing for efficient and sensitive regulation.[84] From a synthetic chromosome consolidation perspective, we discovered a completely unexpected connection between abundance of a single copy tRNA, inositol auxotroph, and, potentially, chromatin dynamics, consistent with regulation via codon usage and tRNA pool adjustments. In the current stage of Sc2.0, the tRNAs have been relocated into chromosome-specific tRNA arrays but will eventually be consolidated into a single neochromosome. Each tRNA gene in the arrays is flanked by rox recombination sites that can be recognized by Dre recombinase. This site-specific recombinase system can be deployed to perform a tRNA gene-specific form of SCRaMbLE, but it can also be deployed in the future to simultaneously remove all of the chromosome-specific tRNA arrays once a future version of the tRNA neochromosome, namely one entirely lacking rox sites, is introduced into the Sc2.0 progenitor strain.

Limitations of the study

In this study, we have constructed 21 strains with all pairwise combinations of 7 previously completed synthetic chromosomes (Figure S7H). Happily, the vast majority of these show no growth defect at 30°C in YPD medium. However, we noted modest growth defects at 37°C suggesting that additional but mild combinatorial bugs may exist. It is certainly possible that there may be other as yet undiscovered phenotypic variations, such as altered genome evolutionary plasticity. The current syn7.5 strain containing these synthetic chromosomes still shows a growth defect, although the fitness has improved dramatically since the integration of *synIV* tRNA array (data not shown). It is also possible to recover growth fitness through adaptive laboratory evolution (ALE).[20,22] Sequencing of the evolved genomes has revealed numerous variants, and further investigation is necessary to determine which of these variants contribute to the observed improvement in fitness. However, we believe that performing ALE before completing chromosome consolidation is likely to lead to very complex networks of mutation/suppressor mutation relationships that may be incompatible with additional synthetic chromosomes, so we prefer to perform ALE only after all of the synthetic chromosomes are combined. Debugging combinatorial defects is more challenging than single bugs, but we are confident that CRISPR D-BUGS will greatly facilitate deciphering these.

CONSORTIA

This work is part of the international Synthetic Yeast Genome (Sc2.0) consortium. The chromosome design and building consortium includes research groups worldwide: Boeke Lab at Johns Hopkins University and New York University (led chromosomes I, III, IV, VI, VIII, and IX); Chandrasegaran lab at Johns Hopkins (led chromosomes III and IX); Cai Lab at University of Edinburgh and University of Manchester (led chromosomes II and VII and tRNA neochromosome); Yue Shen's team at BGI-Research SHENZHEN (led chromosomes II, VII, and XIII); Y.J. Yuan's team at Tianjin University (led chromosomes V and X); Dai Lab at Tsinghua University and Shenzhen Institute of Advanced Technology, CAS (led chromosome XII); Ellis Lab at Imperial College London (led chromosome XI); Sakkie Pretorius's team at Macquarie University (led chromosomes XIV and XVI); Matthew Wook Chang's team at National University of Singapore (led chromosome XV); Bader and Boeke Labs at Johns Hopkins University (led design and workflow); and Build-A-Genome undergraduate teams at Johns Hopkins University and Loyola University Maryland (contributed to chromosomes I, III, IV, VIII, and IX). The Sc2.0 consortium includes numerous

other participants who are acknowledged on the project website, www.syntheticyeast.org.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Synthetic chromosome versions
- METHOD DETAILS
  - Yeast media, growth and transformation
  - Consolidation using advanced endoreduplication intercross
  - tRNA array design and integration
  - CRISPR D-BUGS
  - Genomic editing using CRISPR/SpCas9-NG
  - Genomic editing using CRISPR/Cas9 and Cas12a
  - Pulsed-field gel electrophoresis (PFGE)
  - Whole genome sequencing and alignment
  - GFP tagging and immunoblotting
  - Real-time PCR (RT-PCR)
  - Genomic chromosome conformation capture (Hi-C) library preparation
  - Hi-C sequencing alignment and analysis
  - RNA extraction for transcript profiling
  - Direct RNA sequencing
  - Directional mRNA sequencing
  - Base calling, quality-filtering, and long-read alignment
  - Gene expression quantification
  - Chromosome substitution to consolidate *synIV*
  - DNA content assay
  - Scanning electron microscopy
  - *CAN1* resistance frequency
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantitative RT-PCR analysis for transcript levels
  - Quantitative genome coverage
  - Quantitative Hi-C heat map and contact frequency
  - Quantitative analysis of RNA transcriptomes

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.09.025.

ACKNOWLEDGMENTS

We thank Meru J. Sadhu, Pan Cheng, Guangbin Shi, and Michael Pacold for helpful suggestions. We thank Megan Hogan, Raven Luther, Hannah Ashe, and Gwen Ellis from the Matthew Maurano lab and Runyu Hong from the David Fenyö lab for assistance with WGS analyses. We give special thanks to the Sc2.0 consortium for many discussions and collaborations. This work is part of the Sc2.0 project (http://syntheticyeast.org/), supported by NSF grants MCB-1026068, MCB-1443299, MCB-1616111, and MCB-1921641 to J.D.B. The NYU Microscopy Laboratory is partially supported by the Laura and Isaac Perlmutter Cancer Center support grant NIH/NCI P30CA016087, and the Zeiss Gemini300SEM was purchased with NIH S10 OD019974. Transcriptome profiling was funded by a Volkswagen Stiftung grant (94769) to L.M.S. Figure 5I and the graphical abstract were created with BioRender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.Z., L.A.M., and J.D.B.; experiments and data analysis, Y.Z., C.C., A.L.H., L.L.-S., S.Y., A.N.B., R.S.K.W., W.Z., C.H., Y.S., J.S., and V.F.; genomic data organization, J.C. and J.S.B.; manuscript writing, Y.Z., A.L.H., S.Y., S.L., and J.D.B.; supervision, L.A.M., N.A., F.-X.L., G.S., L.M.S., Y.C., and J.D.B.; C.C., S.Y., C.H., A.J., and J.R. are students from the NYU "Build-A-Genome" class.

DECLARATION OF INTERESTS

J.D.B. is a founder and director of CDI Labs, Inc., a founder of and consultant to Neochromosome, Inc., and a founder, SAB member of, and consultant to ReOpen Diagnostics, LLC and serves or served on the Scientific Advisory Board of the following: Sangamo, Inc., Logomix Inc., Modern Meadow, Inc., Rome Therapeutics, Inc., Sample6, Inc., Tessera Therapeutics, Inc., and the Wyss Institute.

Received: May 5, 2022
Revised: January 3, 2023
Accepted: September 25, 2023
Published: November 8, 2023

REFERENCES

1. Cello, J., Paul, A.V., and Wimmer, E. (2002). Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science *297*, 1016–1018. https://doi.org/10.1126/science.1072266.
2. Smith, H.O., Hutchison, C.A., 3rd, Pfannkoch, C., and Venter, J.C. (2003). Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. Proc. Natl. Acad. Sci. USA *100*, 15440–15445. https://doi.org/10.1073/pnas.2237126100.
3. Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., et al. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. Science *329*, 52–56. https://doi.org/10.1126/science.1190719.
4. Lajoie, M.J., Rovner, A.J., Goodman, D.B., Aerni, H.R., Haimovich, A.D., Kuznetsov, G., Mercer, J.A., Wang, H.H., Carr, P.A., Mosberg, J.A., et al. (2013). Genomically recoded organisms expand biological functions. Science *342*, 357–360. https://doi.org/10.1126/science.1241459.
5. Fredens, J., Wang, K., de la Torre, D., Funke, L.F.H., Robertson, W.E., Christova, Y., Chia, T., Schmied, W.H., Dunkelmann, D.L., Beránek, V., et al. (2019). Total synthesis of Escherichia coli with a recoded genome. Nature *569*, 514–518. https://doi.org/10.1038/s41586-019-1192-5.
6. Zhang, W., Mitchell, L.A., Bader, J.S., and Boeke, J.D. (2020). Synthetic genomes. Annu. Rev. Biochem. *89*, 77–101. https://doi.org/10.1146/annurev-biochem-013118-110704.
7. Venter, J.C., Glass, J.I., Hutchison, C.A., 3rd, and Vashee, S. (2022). Synthetic chromosomes, genomes, viruses, and cells. Cell *185*, 2708–2724. https://doi.org/10.1016/j.cell.2022.06.046.
8. Dormitzer, P.R., Suphaphiphat, P., Gibson, D.G., Wentworth, D.E., Stockwell, T.B., Algire, M.A., Alperovich, N., Barro, M., Brown, D.M., Craig, S., et al. (2013). Synthetic generation of influenza vaccine viruses for rapid response to pandemics. Sci. Transl. Med. *5*, 185ra68. https://doi.org/10.1126/scitranslmed.3006368.
9. Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L.,

et al. (2016). Design and synthesis of a minimal bacterial genome. Science *351*, aad6253. https://doi.org/10.1126/science.aad6253.

10. Ostrov, N., Landon, M., Guell, M., Kuznetsov, G., Teramoto, J., Cervantes, N., Zhou, M., Singh, K., Napolitano, M.G., Moosburner, M., et al. (2016). Design, synthesis, and testing toward a 57-codon genome. Science *353*, 819–822. https://doi.org/10.1126/science.aaf3639.
11. Isaacs, F.J., Carr, P.A., Wang, H.H., Lajoie, M.J., Sterling, B., Kraal, L., Tolonen, A.C., Gianoulis, T.A., Goodman, D.B., Reppas, N.B., et al. (2011). Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. Science *333*, 348–353. https://doi.org/10.1126/science.1205822.
12. Robertson, W.E., Funke, L.F.H., de la Torre, D., Fredens, J., Elliott, T.S., Spinck, M., Christova, Y., Cervettini, D., Böge, F.L., Liu, K.C., et al. (2021). Sense codon reassignment enables viral resistance and encoded polymer synthesis. Science *372*, 1057–1062. https://doi.org/10.1126/science.abg3029.
13. Richardson, S.M., Mitchell, L.A., Stracquadanio, G., Yang, K., Dymond, J.S., DiCarlo, J.E., Lee, D., Huang, C.L., Chandrasegaran, S., Cai, Y., et al. (2017). Design of a synthetic yeast genome. Science *355*, 1040–1044. https://doi.org/10.1126/science.aaf4557.
14. Dymond, J.S., Richardson, S.M., Coombes, C.E., Babatz, T., Muller, H., Annaluru, N., Blake, W.J., Schwerzmann, J.W., Dai, J., Lindstrom, D.L., et al. (2011). Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. Nature *477*, 471–476. https://doi.org/10.1038/nature10403.
15. Blount, B.A., Gowers, G.F., Ho, J.C.H., Ledesma-Amaro, R., Jovicevic, D., McKiernan, R.M., Xie, Z.X., Li, B.Z., Yuan, Y.J., and Ellis, T. (2018). Rapid host strain improvement by in vivo rearrangement of a synthetic yeast chromosome. Nat. Commun. *9*, 1932. https://doi.org/10.1038/s41467-018-03143-w.
16. Gowers, G.F., Chee, S.M., Bell, D., Suckling, L., Kern, M., Tew, D., McClymont, D.W., and Ellis, T. (2020). Improved betulinic acid biosynthesis using synthetic yeast chromosome recombination and semi-automated rapid LC-MS screening. Nat. Commun. *11*, 868. https://doi.org/10.1038/s41467-020-14708-z.
17. Liu, W., Luo, Z., Wang, Y., Pham, N.T., Tuck, L., Pérez-Pi, I., Liu, L., Shen, Y., French, C., Auer, M., et al. (2018). Rapid pathway prototyping and engineering using in vitro and in vivo synthetic genome SCRaMbLE-in methods. Nat. Commun. *9*, 1936. https://doi.org/10.1038/s41467-018-04254-0.
18. Luo, Z., Wang, L., Wang, Y., Zhang, W., Guo, Y., Shen, Y., Jiang, L., Wu, Q., Zhang, C., Cai, Y., et al. (2018). Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. Nat. Commun. *9*, 1930. https://doi.org/10.1038/s41467-017-00806-y.
19. Shen, M.J., Wu, Y., Yang, K., Li, Y., Xu, H., Zhang, H., Li, B.Z., Li, X., Xiao, W.H., Zhou, X., et al. (2018). Heterozygous diploid and interspecies SCRaMbLEing. Nat. Commun. *9*, 1934. https://doi.org/10.1038/s41467-018-04157-0.
20. Williams, T.C., Kroukamp, H., Xu, X., Wightman, E.L.I., Llorente, B., Borneman, A.R., Carpenter, A.C., Wyk, N.V., Espinosa, M.I., Daniel, E.L., et al. (in revision). Parallel laboratory evolution and rational debugging reveal genomic plasticity to Saccharomyces cerevisiae synthetic chromosome XIV defects and rearrangements. Cell Genomics.
21. Zhang, H., Fu, X., Gong, X., Wang, Y., Zhang, H., Zhao, Y., and Shen, Y. (2022). Systematic dissection of key factors governing recombination outcomes by GCE-SCRaMbLE. Nat. Commun. *13*, 5836. https://doi.org/10.1038/s41467-022-33606-0.
22. Zhou, S., Wu, Y., Zhao, Y., Zhang, Z., Jiang, L., Liu, L., Zhang, Y., Tang, J., and Yuan, Y.J. (2023). Dynamics of synthetic yeast chromosome evolution shaped by hierarchical chromatin organization. Natl. Sci. Rev. *10*, nwad073. https://doi.org/10.1093/nsr/nwad073.
23. Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadanio, G., Richardson, S.M., Dymond, J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. Science *344*, 55–58. https://doi.org/10.1126/science.1249252.
24. Shen, Y., Wang, Y., Chen, T., Gao, F., Gong, J., Abramczyk, D., Walker, R., Zhao, H., Chen, S., Liu, W., et al. (2017). Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. Science *355*, eaaf4791. https://doi.org/10.1126/science.aaf4791.
25. Xie, Z.X., Li, B.Z., Mitchell, L.A., Wu, Y., Qi, X., Jin, Z., Jia, B., Wang, X., Zeng, B.X., Liu, H.M., et al. (2017). "Perfect" designer chromosome V and behavior of a ring derivative. Science *355*, eaaf4704. https://doi.org/10.1126/science.aaf4704.
26. Mitchell, L.A., Wang, A., Stracquadanio, G., Kuang, Z., Wang, X., Yang, K., Richardson, S., Martin, J.A., Zhao, Y., Walker, R., et al. (2017). Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond. Science *355*, eaaf4831. https://doi.org/10.1126/science.aaf4831.
27. Wu, Y., Li, B.Z., Zhao, M., Mitchell, L.A., Xie, Z.X., Lin, Q.H., Wang, X., Xiao, W.H., Wang, Y., Zhou, X., et al. (2017). Bug mapping and fitness testing of chemically synthesized chromosome X. Science *355*, eaaf4706. https://doi.org/10.1126/science.aaf4706.
28. Zhang, W., Zhao, G., Luo, Z., Lin, Y., Wang, L., Guo, Y., Wang, A., Jiang, S., Jiang, Q., Gong, J., et al. (2017). Engineering the ribosomal DNA in a megabase synthetic chromosome. Science *355*, eaaf3981. https://doi.org/10.1126/science.aaf3981.
29. Luo, J., Mercy, G., Vale-Silva, L.A., Sun, X., Agmon, N., Zhang, W., Yang, K., Stracquadanio, G., Thierry, A., Ahn, J.Y., et al. (2023). Synthetic chromosome fusion: effects on genome structure and function. in revision. Cell Genomics.
30. Zhang, W., Lazar-Stefanita, L., Yamashita, H., Shen, M.J., Mitchell, L.A., Kurasawa, H., Haase, M.A.B., Sun, X., Jiang, Q., Lauer, S.L., et al. (2023). Manipulating the 3D organization of the largest synthetic yeast chromosome. Molecular Cell. https://doi.org/10.1016/j.molcel.2023.10.015.
31. Shen, Y., Gao, F., Wang, Y., Wang, Y., Zheng, J., Gong, J., Zhang, J., Luo, Z., Schindler, D., Deng, Y., et al. (2023). Dissecting aneuploidy phenotypes by constructing Sc2.0 chromosome VII and SCRaMbLEing synthetic disomic yeast. Cell Genomics. https://doi.org/10.1016/j.xgen.2023.100364.
32. Lauer, S., Luo, J., Lazar-Stefanita, L., Zhao, Y., Zhang, W., Fanfani, V., Lobzaev, E., Haase, M.A.B., Easo, N., McCulloch, L.H., et al. (2023). Context-dependent neocentromere activity in synthetic yeast chromosome VIII. Cell Genomics.. https://doi.org/10.1016/j.xgen.2023.100437.
33. McCulloch, L.H., Sambasivam, V., Hughes, A.L., Annaluru, N., Ramalingam, S., Fanfani, V., Lobzaev, E., Mitchell, L., Cai, J., Class, T.B.A.G., et al. (2023). Consequences of a telomerase-related fitness defect and chromosome substitution technology in yeast synIX strains. Cell Genomics. https://doi.org/10.1016/j.xgen.2023.100419.
34. Blount, B.A., Lu, X., Driessen, M.R.M., Jovicevic, D., Sanchez, M.I., Ciurkot, K., Zhao, Y., Lauer, S., McKiernan, R.M., Gowers, G.-O.F., et al. (2023). Synthetic yeast chromosome XI design provides a testbed for studying extrachromosomal circular DNA. Cell Genomics. https://doi.org/10.1016/j.xgen.2023.100418.
35. Zhou, C., Wang, Y., Huang, Y., An, Y., Wang, Y., Zhang, J., Mitchell, L.A., Bader, J.S., Cai, Y., Boeke, J.D., et al. (2023). Screening and characterization of aging regulators using synthesized yeast chromosome XIII. Cell Genomics.
36. Foo, J.L., Kitano, S., Susanto, A.V., Jin, Z., Lin, Y., Luo, Z., Huang, L., Liang, Z., Mitchell, L.A., Yang, K., et al. (2023). Establishing chromosomal design-build-test-learn through combinatorial reconfiguration of a synthetic chromosome. Cell Genomics. https://doi.org/10.1016/j.xgen.2023.100435.
37. Schindler, D., Walker, R.S.K., Jiang, S., Brooks, A.N., Wang, Y., Müller, C.A., Cockram, C., Luo, Y., García, A., Schraivogel, D., et al. (2023). Design, construction, and functional characterization of a tRNA neochromosome in yeast. Cell. https://doi.org/10.1016/j.cell.2023.10.015.

38. DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343. https://doi.org/10.1093/nar/gkt135.

39. Jacobs, J.Z., Ciccaglione, K.M., Tournier, V., and Zaratiegui, M. (2014). Implementation of the CRISPR-Cas9 system in fission yeast. Nat. Commun. *5*, 5344. https://doi.org/10.1038/ncomms6344.

40. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759–771. https://doi.org/10.1016/j.cell.2015.09.038.

41. Sadhu, M.J., Bloom, J.S., Day, L., and Kruglyak, L. (2016). CRISPR-directed mitotic recombination enables genetic mapping without crosses. Science *352*, 1113–1116. https://doi.org/10.1126/science.aaf5124.

42. Zhao, Y., and Boeke, J.D. (2018). Construction of designer selectable marker deletions with a CRISPR-Cas9 toolbox in Schizosaccharomyces pombe and new design of com mon entry vectors. G3 (Bethesda) *8*, 789–796. https://doi.org/10.1534/g3.117.300363.

43. Zhao, Y., and Boeke, J.D. (2020). CRISPR-Cas12a system in fission yeast for multiplex genomic editing and CRISPR interference. Nucleic Acids Res. *48*, 5788–5798. https://doi.org/10.1093/nar/gkaa329.

44. Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., et al. (2018). Highly parallel direct RNA sequencing on an array of nanopores. Nat. Methods *15*, 201–206. https://doi.org/10.1038/nmeth.4577.

45. Reid, R.J., Sunjevaric, I., Voth, W.P., Ciccone, S., Du, W., Olsen, A.E., Stillman, D.J., and Rothstein, R. (2008). Chromosome-scale genetic mapping using a set of 16 conditionally stable Saccharomyces cerevisiae chromosomes. Genetics *180*, 1799–1808. https://doi.org/10.1534/genetics.108.087999.

46. May, A.I., Prescott, M., and Ohsumi, Y. (2020). Autophagy facilitates adaptation of budding yeast to respiratory growth by recycling serine for one-carbon metabolism. Nat. Commun. *11*, 5052. https://doi.org/10.1038/s41467-020-18805-x.

47. Louvet, O., Roumanie, O., Barthe, C., Peypouquet, M.F., Schaeffer, J., Doignon, F., and Crouzet, M. (1999). Characterization of the ORF YBR264c in Saccharomyces cerevisiae, which encodes a new yeast Ypt that is degraded by a proteasome-dependent mechanism. Mol. Gen. Genet. *261*, 589–600. https://doi.org/10.1007/s004380050001.

48. Brooks, A.N., Hughes, A.L., Clauder-Münster, S., Mitchell, L.A., Boeke, J.D., and Steinmetz, L.M. (2022). Transcriptional neighborhoods regulate transcript isoform lengths and expression levels. Science *375*, 1000–1005. https://doi.org/10.1126/science.abg0162.

49. Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science *361*, 1259–1262. https://doi.org/10.1126/science.aas9129.

50. Auesukaree, C., Damnernsawad, A., Kruatrachue, M., Pokethitiyook, P., Boonchird, C., Kaneko, Y., and Harashima, S. (2009). Genome-wide identification of genes involved in tolerance to various environmental stresses in Saccharomyces cerevisiae. J. Appl. Genet. *50*, 301–310. https://doi.org/10.1007/BF03195688.

51. Peterson, C.L., and Herskowitz, I. (1992). Characterization of the yeast Swi1, Swi2, and Swi3 genes, which encode a global activator of transcription. Cell *68*, 573–583. https://doi.org/10.1016/0092-8674(92)90192-F.

52. Yoshinaga, S.K., Peterson, C.L., Herskowitz, I., and Yamamoto, K.R. (1992). Roles of SWI1, SWI2, and SWI3 proteins for transcriptional enhancement by steroid receptors. Science *258*, 1598–1604. https://doi.org/10.1126/science.1360703.

53. Villa-García, M.J., Choi, M.S., Hinz, F.I., Gaspar, M.L., Jesch, S.A., and Henry, S.A. (2011). Genome-wide screen for inositol auxotrophy in Saccharomyces cerevisiae implicates lipid metabolism in stress response signaling. Mol. Genet. Genomics *285*, 125–149. https://doi.org/10.1007/s00438-010-0592-x.

54. Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr. Opin. Biotechnol. *6*, 494–500. https://doi.org/10.1016/0958-1669(95)80082-4.

55. Wang, Y., Li, C., Khan, M.R., Wang, Y., Ruan, Y., Zhao, B., Zhang, B., Ma, X., Zhang, K., Zhao, X., et al. (2016). An engineered rare codon device for optimization of metabolic pathways. Sci. Rep. *6*, 20608. https://doi.org/10.1038/srep20608.

56. Guimaraes, J.C., Mittal, N., Gnann, A., Jedlinski, D., Riba, A., Buczak, K., Schmidt, A., and Zavolan, M. (2020). A rare codon-based translational program of cell proliferation. Genome Biol. *21*, 44. https://doi.org/10.1186/s13059-020-1943-5.

57. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293. https://doi.org/10.1126/science.1181369.

58. Mercy, G., Mozziconacci, J., Scolari, V.F., Yang, K., Zhao, G., Thierry, A., Luo, Y., Mitchell, L.A., Shen, M., Shen, Y., et al. (2017). 3D organization of synthetic and scrambled chromosomes. Science *355*, eaaf4597. https://doi.org/10.1126/science.aaf4597.

59. Taddei, A., and Gasser, S.M. (2012). Structure and function in the budding yeast nucleus. Genetics *192*, 107–129. https://doi.org/10.1534/genetics.112.140608.

60. Hull, M.W., Erickson, J., Johnston, M., and Engelke, D.R. (1994). tRNA genes as transcriptional repressor elements. Mol. Cell. Biol. *14*, 1266–1277.

61. Good, P.D., Kendall, A., Ignatz-Hoover, J., Miller, E.L., Pai, D.A., Rivera, S.R., Carrick, B., and Engelke, D.R. (2013). Silencing near tRNA genes is nucleosome-mediated and distinct from boundary element function. Gene *526*, 7–15. https://doi.org/10.1016/j.gene.2013.05.016.

62. Hamdani, O., Dhillon, N., Hsieh, T.S., Fujita, T., Ocampo, J., Kirkland, J.G., Lawrimore, J., Kobayashi, T.J., Friedman, B., Fulton, D., et al. (2019). tRNA genes affect chromosome structure and function via local effects. Mol. Cell. Biol. *39*, e00432-00418. https://doi.org/10.1128/MCB.00432-18.

63. Conde, J., and Fink, G.R. (1976). A mutant of Saccharomyces cerevisiae defective for nuclear fusion. Proc. Natl. Acad. Sci. USA *73*, 3651–3655. https://doi.org/10.1073/pnas.73.10.3651.

64. Dutcher, S.K. (1981). Internuclear transfer of genetic information in kar1-1/KAR1 heterokaryons in Saccharomyces cerevisiae. Mol. Cell. Biol. *1*, 245–253. https://doi.org/10.1128/mcb.1.3.245-253.1981.

65. Ji, H., Moore, D.P., Blomberg, M.A., Braiterman, L.T., Voytas, D.F., Natsoulis, G., and Boeke, J.D. (1993). Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. Cell *73*, 1007–1018. https://doi.org/10.1016/0092-8674(93)90278-x.

66. Shen, Y., Stracquadanio, G., Wang, Y., Yang, K., Mitchell, L.A., Xue, Y., Cai, Y., Chen, T., Dymond, J.S., Kang, K., et al. (2016). SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. Genome Res. *26*, 36–49. https://doi.org/10.1101/gr.193433.115.

67. Luo, Z., Wang, L., Wang, Y., Zhang, W., Guo, Y., Shen, Y., Jiang, L., Wu, Q., Zhang, C., and Cai, Y. (2018). Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. Nat. Commun. *9*, 1–10.

68. Zhao, M., Zhao, Y., Yao, M., Iqbal, H., Hu, Q., Liu, H., Qiao, B., Li, C., Skovbjerg, C.A.S., Nielsen, J.C., et al. (2020). Pathway engineering in yeast for synthesizing the complex polyketide bikaverin. Nat. Commun. *11*, 6197. https://doi.org/10.1038/s41467-020-19984-3.

69. Dai, J., Boeke, J.D., Luo, Z., Jiang, S., and Cai, Y. (2020). Sc3.0: revamping and minimizing the yeast genome. Genome Biol. *21*, 205. https://doi.org/10.1186/s13059-020-02130-z.

70. Parenteau, J., Maignon, L., Berthoumieux, M., Catala, M., Gagnon, V., and Abou Elela, S. (2019). Introns are mediators of cell response to starvation. Nature *565*, 612–617. https://doi.org/10.1038/s41586-018-0859-7.

71. Morgan, J.T., Fink, G.R., and Bartel, D.P. (2019). Excised linear introns regulate growth in yeast. Nature *565*, 606–611. https://doi.org/10.1038/s41586-018-0828-1.

72. Lukačišin, M., Espinosa-Cantú, A., and Bollenbach, T. (2022). Intron-mediated induction of phenotypic heterogeneity. Nature *605*, 113–118. https://doi.org/10.1038/s41586-022-04633-0.

73. Kastanos, E.K., Woldman, Y.Y., and Appling, D.R. (1997). Role of mitochondrial and cytoplasmic serine hydroxymethyltransferase isozymes in de novo purine synthesis in Saccharomyces cerevisiae. Biochemistry *36*, 14956–14964. https://doi.org/10.1021/bi971610n.

74. Lee, J.C., Tsoi, A., Kornfeld, G.D., and Dawes, I.W. (2013). Cellular responses to L-serine in Saccharomyces cerevisiae: roles of general amino acid control, compartmentalization, and aspartate synthesis. FEMS Yeast Res. *13*, 618–634. https://doi.org/10.1111/1567-1364.12063.

75. Piper, M.D., Hong, S.P., Ball, G.E., and Dawes, I.W. (2000). Regulation of the balance of one-carbon metabolism in Saccharomyces cerevisiae. J. Biol. Chem. *275*, 30987–30995.

76. Beeler, T., Bacikova, D., Gable, K., Hopkins, L., Johnson, C., Slife, H., and Dunn, T. (1998). The Saccharomyces cerevisiae TSC10/YBR265w gene encoding 3-ketosphinganine reductase is identified in a screen for temperature-sensitive suppressors of the Ca2+-sensitive csg2Δ mutant. J. Biol. Chem. *273*, 30688–30694.

77. Chen, P.W., Fonseca, L.L., Hannun, Y.A., and Voit, E.O. (2015). Dynamics of the heat stress response of ceramides with different fatty-acyl chain lengths in baker's yeast. PLoS Comp. Biol. *11*, e1004373. https://doi.org/10.1371/journal.pcbi.1004373.

78. Jenkins, G.M., Richards, A., Wahl, T., Mao, C., Obeid, L., and Hannun, Y. (1997). Involvement of yeast sphingolipids in the heat stress response of Saccharomyces cerevisiae. J. Biol. Chem. *272*, 32566–32572. https://doi.org/10.1074/jbc.272.51.32566.

79. Albers, E., Laizé, V., Blomberg, A., Hohmann, S., and Gustafsson, L. (2003). Ser3p (Yer081wp) and Ser33p (Yil074cp) are phosphoglycerate dehydrogenases in Saccharomyces cerevisiae. J. Biol. Chem. *278*, 10264–10272. https://doi.org/10.1074/jbc.M211692200.

80. Melcher, K., and Entian, K.D. (1992). Genetic analysis of serine biosynthesis and glucose repression in yeast. Curr. Genet. *21*, 295–300. https://doi.org/10.1007/BF00351686.

81. Frumkin, I., Lajoie, M.J., Gregg, C.J., Hornung, G., Church, G.M., and Pilpel, Y. (2018). Codon usage of highly expressed genes affects proteome-wide translation efficiency. Proc. Natl. Acad. Sci. USA *115*, E4940–E4949.

82. Wei, Y., Silke, J.R., and Xia, X. (2019). An improved estimation of tRNA expression to better elucidate the coevolution between tRNA abundance and codon usage in bacteria. Sci. Rep. *9*, 3184. https://doi.org/10.1038/s41598-019-39369-x.

83. Liu, Y. (2020). A code within the genetic code: codon usage regulates co-translational protein folding. Cell Commun. Signal. *18*, 145. https://doi.org/10.1186/s12964-020-00642-6.

84. Torrent, M., Chalancon, G., de Groot, N.S., Wuster, A., and Madan Babu, M. (2018). Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. Sci. Signal. *11*, eaat6409. https://doi.org/10.1126/scisignal.aat6409.

85. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498. https://doi.org/10.1038/ng.806.

86. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

87. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

88. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419. https://doi.org/10.1038/nmeth.4197.

89. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

90. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

91. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

92. Xie, Z.X., Mitchell, L.A., Liu, H.M., Li, B.Z., Liu, D., Agmon, N., Wu, Y., Li, X., Zhou, X., Li, B., et al. (2018). Rapid and efficient CRISPR/Cas9-based mating-type switching of Saccharomyces cerevisiae. G3 (Bethesda) *8*, 173–183. https://doi.org/10.1534/g3.117.300347.

93. Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115–132. https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.

94. Hill, A., and Bloom, K. (1987). Genetic manipulation of centromere function. Mol. Cell. Biol. *7*, 2397–2405. https://doi.org/10.1128/mcb.7.7.2397-2405.1987.

95. Swiat, M.A., Dashko, S., den Ridder, M., Wijsman, M., van der Oost, J., Daran, J.M., and Daran-Lapujade, P. (2017). FnCpf1: a novel and efficient genome editing tool for Saccharomyces cerevisiae. Nucleic Acids Res. *45*, 12585–12598. https://doi.org/10.1093/nar/gkx1007.

96. Luo, J., Sun, X., Cormack, B.P., and Boeke, J.D. (2018). Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. Nature *560*, 392–396. https://doi.org/10.1038/s41586-018-0374-x.

97. Weill, U., Yofe, I., Sass, E., Stynen, B., Davidi, D., Natarajan, J., Ben-Menachem, R., Avihou, Z., Goldman, O., Harpaz, N., et al. (2018). Genome-wide SWAp-Tag yeast libraries for proteome exploration. Nat. Methods *15*, 617–622. https://doi.org/10.1038/s41592-018-0044-9.

98. Ikushima, S., Zhao, Y., and Boeke, J.D. (2015). Development of a tightly controlled off switch for Saccharomyces cerevisiae regulated by camphor, a low-cost natural product. G3 (Bethesda) *5*, 1983–1990. https://doi.org/10.1534/g3.114.012765.

99. Teste, M.A., Duquenne, M., François, J.M., and Parrou, J.L. (2009). Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae. BMC Mol. Biol. *10*, 99. https://doi.org/10.1186/1471-2199-10-99.

100. Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi-C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268–276. https://doi.org/10.1016/j.ymeth.2012.05.001.

101. Lazar-Stefanita, L., Scolari, V.F., Mercy, G., Muller, H., Guérin, T.M., Thierry, A., Mozziconacci, J., and Koszul, R. (2017). Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during

the cell cycle. EMBO J. *36*, 2684–2697. https://doi.org/10.15252/embj.201797342.

102. Lazar-Stefanita, L., Luo, J., Montagne, R., Thierry, A., Sun, X., Mercy, G., Mozziconacci, J., Koszul, R., and Boeke, J.D. (2022). Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts. Cell Genom. *2*, None. https://doi.org/10.1016/j.xgen.2022.100163.

103. Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat. Methods *9*, 999–1003. https://doi.org/10.1038/nmeth.2148.

104. Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R., and Mozziconacci, J. (2012). Normalization of a chromosomal contact map. BMC Genomics *13*, 436. https://doi.org/10.1186/1471-2164-13-436.

105. Lesne, A., Riposo, J., Roger, P., Cournac, A., and Mozziconacci, J. (2014). 3D genome reconstruction from chromosomal contacts. Nat. Methods *11*, 1141–1143. https://doi.org/10.1038/nmeth.3104.

106. Haase, S.B., and Lew, D.J. (1997). Flow cytometric analysis of DNA content in budding yeast. Methods Enzymol. *283*, 322–332. https://doi.org/10.1016/s0076-6879(97)83026-1.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse-anti-GFP	Roche	Cat# 11814460001, RRID:AB_390913
Rabbit-anti-Histone H3	Abcam	Cat# ab1791, RRID:AB_302613
IRDye Goat anti-Mouse IgG	LI-COR Biosciences	Cat# 926-32210, RRID:AB_621842
IRDye Goat anti-Rabbit IgG	LI-COR Biosciences	Cat# 926-68071, RRID:AB_10956166
Mouse-anti-FLAG	Sigma-Aldrich	Cat# F1804, RRID:AB_262044
Qubit dsDNA HS and BR Assay kit	Thermo Fisher Scientific	Cat# Q32851
Bacterial and Virus Strains		
Turbo Competent *E. coli*	NEB	Cat# C2984
Chemicals, Peptides, and Recombinant Proteins		
Glycerol	Thermo Fisher Scientific	Cat# G33-4
myo-Inositol	Sigma-Aldrich	Cat# I5125
D-(+)-Galactose	Sigma-Aldrich	Cat# G0625
5-FOA	US Biological	Cat# F5050
Pulsed Field Certified Agarose	Bio-Rad	Cat# 1620137
Yeast Nitrogen Base w/o AA, Carbohydrate & w/o AS, Inositol (YNB)	US Biological	Cat# Y2030-01
L-Canavanine Sulfate	Sigma-Aldrich	Cat# C9758
Cycloheximide	Sigma-Aldrich	Cat# C7698
Propidium Iodide	Thermo Fisher Scientific	Cat# P3566
RNase A	Thermo Fisher Scientific	Cat# EN0531
Critical Commercial Assays		
NEBNext Ultra II FS DNA Library Prep Kit for Illumina	NEB	Cat# E7805L
Fungi/Yeast Genomic DNA Isolation Kit	NORGEN	Cat# 27300
RNeasy Mini Kit	Qiagen	Cat# 74106
SuperScript IV Reverse Transcriptase	Thermo Fisher Scientific	Cat# 18090050
Lightcycler 480 SYBR Green I Master Mix	Lightcycler	Cat# 04887352001
MasterPure yeast RNA purification kit	Lucigen	Cat# MPY03100
Agilent RNA 6000 Nano kit	Agilent	Cat# 5067-1511
Qubit RNA High Sensitivity kit	Thermo Fisher Scientific	Cat# Q32852
Direct RNA Sequencing kit	Oxford Nanopore Technology	Cat# SQK-RNA002
MinION flow cell	Oxford Nanopore Technology	FLO-MIN106D, R9 version
NEBNext Poly(A) mRNA magnetic isolation module	NEB	Cat# E7490
ERCC RNA Spike-In control mix	Thermo Fisher Scientific	Cat# 4456740
NEBNext Ultra II directional RNA library prep kit	NEB	Cat# E7765
Agilent High Sensitivity DNA kit	Agilent	Cat# 5067-4626
Q5 High-Fidelity 2X Master Mix	NEB	Cat# M0492L
Deposited Data		
NCBI BioProject	Saccharomyces cerevisiae strain: multiple Genome sequencing	PRJNA351844
Sc2.0 project	http://syntheticyeast.org/	N/A
Experimental Models: Organisms/Strains		
For yeast strain information, see Table S4	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
For oligonucleotide sequence information, see Tables S5 and S6	This study	N/A
Recombinant DNA		
pX330-SpCas9-NG	Addgene	Addgene# 117919
pYZ555-pGal-Cas9	This study	Addgene# 186161
Software and Algorithms		
PicardTools (v1.140)	picard.sourceforge.net	RRID: SCR_006525
Genome Analysis Toolkit (GATK, v3.6)	DePristo et al.[85]	RRID: SCR_001876
Trimmomatic (v0.38)	Bolger et al.[86]	RRID: SCR_011848
BEDTools (v2.29.2)	Quinlan et al.[87]	RRID: SCR_006646
Salmon (v1.6.0)	Patro et al.[88]	RRID: SCR_017036
DESeq2	Love et al.[89]	RRID: SCR_015687
BD Biosciences C6 Accuri Flow Cytometer	BD Biosciences	RRID: SCR_019591
Bowtie2 (v2.2.9)	Langmead et al.[90]	RRID: SCR_005476
Illumina NextSeq 500	Illumina	RRID: SCR_014983
Thermo Fisher Qubit fluorimeter	Thermo Fisher Scientific	RRID: SCR_018095
Agilent 2100 Bioanalyzer Instrument	Agilent	RRID: SCR_019389
Microscopy Laboratory Core Facility	New York University Langone Health	RRID: SCR_017934
Samtools (v1.9)	Li et al.[91]	RRID: SCR_002105

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jef D. Boeke (jef.boeke@nyulangone.org).

Materials availability

The plasmid pYZ555 used in CRISPR D-BUGS was submitted to Addgene (ID# 186161). All synthetic chromosome strains are available upon request.

Data and code availability

All genomic raw data, including synthetic chromosome sequences in syn6.5, RNAseq, Hi-C, and Nanopore direct RNA sequencing, are available at BioProject PRJNA914659 under the overarching Sc2.0 umbrella BioProject (PRJNA351844). More information about each synthetic chromosome and summaries can be accessed on the Sc2.0 website (www.syntheticyeast.org). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Synthetic chromosome versions

We started the consolidation using strains containing individual synthetic chromosome and when necessary, we switched the mating type using CRISPR.[92] Detailed information on intermediate strain names, genotypes and version numbers are listed in Table S5. Briefly, they are *synII* yeast_chr02_9_03,[24] *synIII* yeast_chr03_9_02,[23] synIV yeast_chr04_9_03,[30] *synV* yeast_chr05_9_22,[25] *synVI* yeast_chr06_9_03,[26] *synIXR* genebank JN020955,[14] *synX* yeast_chr10_9_01[27] and *synXII* yeast_chr12_9_02.[28]

METHOD DETAILS

Yeast media, growth and transformation

Yeast strains were cultured in YPD-rich medium or defined SC media with appropriate components dropped out. All yeast transformations were performed using standard lithium acetate protocols.[93] To check *synII* growth defects, YPG plates contained 3% glycerol. For inositol auxotrophy tests, the inositol free YNB medium was prepared using yeast nitrogen base w/o inositol (US Biological

Y2030-01), and supplemented with 5 g/L ammonium sulfate, 2% dextrose and the necessary amino acid supplements. Control plates were supplemented with 76 mg/L myo-inositol (Sigma I5125).

Consolidation using advanced endoreduplication intercross

Two separate synthetic chromosomes were consolidated using advanced endoreduplication intercross from their host strain with opposite mating types.[26] First, the two host haploid strains were engineered: the *KlURA3*-pGAL-CENx module was integrated into the native counterparts of the target synthetic chromosomes to be lost.[94] Our early experiments suggested that destabilizing two native chromosomes per diploid by this method worked more reliably than trying to destabilize larger numbers of native chromosomes at the same time. After mating and inoculation in YP+Galactose (2%) medium overnight, the relevant heterozygous diploid strain was screened on SC+5-FOA plates for successful destabilization of both target chromosomes, generating a 2n–2 strain. After growth for 24 hours allowing for endoreduplication in YPD, the strain was cultured in sporulation medium at room temperature. Finally, from tetrad dissection, spore clones with more than both target synthetic chromosomes were obtained. This process was continued to consolidate more synthetic chromosomes.

tRNA array design and integration

For each synthetic chromosome, a tRNA array containing all the synthetic version of tRNA genes from its host chromosome was constructed (Figure S1C). Each syn-tRNA contains 500 bp 5' and 40 bp 3' flanking sequences from *Eremothecium* (*Ashbya*) *gossypii* or *Eremothecium cymbalariae*. The tRNA arrays were integrated as shown in Figure S1. The arrays were released from plasmids using restriction enzyme digestion (Table S1). To integrate tRNA arrays, we constructed junction DNAs with 500 bp homology arms to the target genomic locus, 500 bp homology arms to the linearized tRNA array and the *KlURA3* selection marker at one end. Integrations were selected on SC–Ura plates and confirmed by colony PCR. Afterwards, the *KlURA3* marker was deleted using CRISPR/Cas9 and a gRNA.KlURA3 (ACCAGTAACCCCGTGGGCGT), provided with a flanking donor DNA.

CRISPR D-BUGS

In this study, we developed CRISPR D-BUGS to quickly and reliably map the bug on a synthetic chromosome. Step one in this process is to determine whether the fitness defect is recessive (most cases) or dominant. Assuming that the defect to be mapped is recessive, we first created a diploid strain of yeast heterozygous for the target chromosome arm with a *URA3* marker integrated in an intergenic region close to the telomere (~25 kb) of the native chromosome. This heterozygous diploid strain was obtained by mating a haploid strain carrying target synthetic chromosomes with another haploid strain with corresponding native chromosomes. Then several gRNAs targeting WT PCRtags at different regions along the chromosome were selected (Table S6). For the initial round of CRISPR D-BUGS, it is good to have gRNAs targeting near the telomeres, near the middle of the left or right arms, and on either side of the centromere (at least ~10 kb away). The gRNA was assembled into a CRISPR/pGAL-Cas9 plasmid backbone (pYZ555 with *LEU2* marker, Addgene# 186161) using Golden Gate cloning.[43]

Afterwards, the heterozygous diploid strain was transformed with the CRISPR plasmid and selected on SC–Leu dextrose plates (Cas9 OFF). A single colony was inoculated in SC–Leu galactose medium and incubated at 30℃ overnight (Cas9 ON). The medium was diluted and plated on 5-FOA plates to select the single colonies with successful mitotic recombination, further confirmed using PCRtag assays. Finally, we assessed fitness by single colony formation spot tests to identify the fitness boundary. Once the fitness boundary is rough-mapped, further intermediate gRNAs can be chosen for fine mapping until a gRNA that produces a mix of fit and unfit clones is identified. WGS of the fit and unfit clones can then be deployed to fine-map the fitness boundary.

Genomic editing using CRISPR/SpCas9-NG

The CRISPR/SpCas9-NG system was used to repair an accidental single base pair mutation in *synI YAL061W*. The SpCas9-NG ORF was subcloned from pX330-SpCas9-NG obtained from Addgene (#117919), and assembled with *TEF1* promoter and *CYC1* terminator.[30] Briefly, the gRNA (GGTCCATGTGCTACACACAC) targeting at *YAL061W* with CG as the PAM was used to repair the mutation in yJL663 with a draft version of *synI*. The donor DNA was the PCR product from wild-type genomic DNA containing 140 bp homology arm on each side of the target mutation. We got 3 out of 11 positive colonies where the mutation was repaired.

Genomic editing using CRISPR/Cas9 and Cas12a

We also used CRISPR/Cas9 and Cas12a (also called Cpf1) to repair the mapped bugs or introduce variants. We followed the protocols as described previously.[38,95] All targets, gRNA and PAM sequences for this study are listed in Table S6.

Pulsed-field gel electrophoresis (PFGE)

To evaluate the multiple synthetic chromosomes in syn6.5 strains, chromosome plugs were prepared the separated by clamped homogeneous electric field (CHEF) gel electrophoresis using the CHEF-DR III Pulsed-Field Electrophoresis System (Bio-Rad), as previously described.[96] The following program was used, temperature: 14 °C, voltage: 6 V/cm, switch time: 60 s to 120 s, run time: 20 h, included angle: 120°, using 0.5 × Tris-Borate-EDTA buffer and a 1% gel with low melting point agarose (Bio-Rad #1620137). Gels were stained with 5 μg/ml ethidium bromide in water after electrophoresis for 30 min, de-stained in water for 30 min, and then imaged.

Whole genome sequencing and alignment

The yeast genomic DNA samples for sequencing were prepared using a Norgen Biotek fungi/yeast genomic DNA isolation kit (Cat# 27300). The library was prepared using NEBNext Ultra II FS DNA library prep kit (NEB E7805L) with 500 ng genomic DNA as input. The whole genome sequencing was performed using an Illumina NextSeq 500 system using pair-end 36 bp protocol. All raw reads were trimmed to remove adaptor sequence using Trimmomatic,[86] and subsequently mapped to synthetic chromosome sequences using Bowtie2 software,[90] with Samtools.[91] The variants were analyzed with Genome Analysis Toolkit (GATK).[85] The coverage for each locus was calculated using BEDTools and normalized to average genome-wide coverage.[87]

GFP tagging and immunoblotting

To quantify protein expression level of *SHM1* by immunoblotting, we first tagged it with GFP. Since the loxPsym sites are inserted in 3' UTR close to the stop codon, we integrated the tag at N-terminal instead of C-terminal. We used the same sequence and design of SWAT library.[97] We first isolated the strain of *SHM1* tagged with GFP at N terminal and its native promoter from the SWAT library. Then, we PCR amplified the region containing GFP and 500 bp homology arms on each side from its genomic DNA as the donor DNA, which was transformed with CRISPR/Cas9, using one gRNA (GACTAGCGATTGTGCACCAC). Successful integration was confirmed with colony PCR and Sanger sequencing. Notably, the mitochondria targeting signal of Shm1p was not affected.

We tagged *SHM1* with GFP in original *synII* (9.03) and fixed *synII* (9.04) strain, generating YZY516 and YZY517, respectively. The original wild-type strain from SWAT library, YZY208, was used as a control. These strains were cultured in YPD medium overnight, and then diluted to produce a log phase culture. The cell lysate was prepared and run on SDS-PAGE as previously described.[98] Briefly, 200 μl of yeast cells (OD_{600} ~0.6) were collected and washed with water. These cells were then resuspended in 200 μl of lysis buffer containing 20 mM HEPES (pH 7.4), 0.1% Tween-20, 2 mM MgCl2, 300 mM NaCl, and 1× protease inhibitor cocktail (Roche 11873580001). Approximately 100 μl of 0.5 mm glass beads were added to the suspension, and the tubes were subjected to vortexing using the MP FastPrep-24 5G Homogenizer with 6 cycles of shaking for 30 seconds and cooldown for 30 seconds. Subsequently, the mixture was centrifuged at 12,000×g for 10 minutes at 4°C. 30 μl of the supernatant were transferred to a new tube and mixed with 10 μl of NuPAGE LDS Sample Buffer (Invitrogen NP0007). Samples were then heated at 70°C for 10 minutes. 10 μl of each sample was loaded onto NuPAGE 4-12% Bis-Tris Gels and subjected to electrophoresis at 80 V for 15 minutes, followed by 120 V for 1 hour. The Semi-Dry and Rapid Blotting System (Bio-Rad Laboratories, Inc, Hercules, CA) was used to transfer proteins to a PVDF membrane (Bio-Rad Laboratories, Inc, Hercules, CA). Anti-GFP antibody from mouse (Cat# 11814460001; Roche) and TUBB2A from rabbit (Cat# AV40177; Sigma-Aldrich) were mixed and used as the primary antibody. IRDye Goat anti-Mouse IgG (Cat# 926-32210; LI-COR Biosciences) and Goat anti-Rabbit IgG (Cat# 926-68071; LI-COR Biosciences) were used as secondary antibodies, respectively. The fluorescence signal was detected on an Odyssey CLx Imager from LI-COR (Lincoln, Nebraska).

Proteins were transferred to a PVDF membrane for blocking, antibody binding and imaging. Anti-GFP antibody from mouse (Roche 11814460001) at 1:1000 dilution and anti-H3 antibody from rabbit (Abcam ab1791) at 1:2500 dilution were used as the primary antibodies. IRDye Goat anti-mouse IgG (LI-COR Biosciences 926-32210) and Goat anti-rabbit IgG (LI-COR Biosciences 926-68071) were used as the secondary antibodies, respectively. The fluorescence signal was detected on an Odyssey CLx Imager from LI-COR.

To check protein expression level of *SWI3* with loxPsym in 5' UTR and/or 3' UTR at *synX*, we first tagged it with 3×Flag tags (DYKDHDGDYKDHDIDYKDDDDK) and a GS linker $(GGGGS)_3$ at C-terminus. The immunoblotting was performed with the same method as above. Anti-Flag antibody from mouse (Sigma F1804) at 1:2000 dilution was used as the primary antibodies. The same internal control and secondary antibodies were used as above.

Real-time PCR (RT-PCR)

We used RT-PCR to check the expression level of *SHM1* as previously described.[26] Briefly, from 3 single colonies as triplicates, the RNA was prepared using RNeasy Mini kit (Qiagen 74106). First strand cDNA was prepared using SuperScript IV Reverse Transcriptase (Invitrogen 18090050) and oligo $d(T)_{20}$ primer. The expression level was tested using Lightcycler 480 SYBR Green I Master Mix (Lightcycler 04887352001) in a 10 μl reaction system. The qPCR was performed and analyzed using the LightCycler 480 System. For *SHM1*, the forward primer (GCTCTGGAACTGTACGGATTA) and reverse primer (ACGTTCATGATAGCGGAGTAAA) were designed by IDT PrimerQuest Tool. The *TAF10* was used as the internal control.[99]

Genomic chromosome conformation capture (Hi-C) library preparation

We prepared the libraries and performed the analysis as described previously.[100–102] Three independent syn6.5 yeast colonies were inoculated into 5 mL YPD medium overnight, then subcultured into 100 mL YPD for 3 h growth at 30°C. Cells were crosslinked by 3% [v/v] formaldehyde for 20 min at room temperature and subsequently quenched with 350 mM glycine for 15 min at 4°C on a rocking platform. The crosslinked cells were harvested by centrifugation at 4°C for 5 min at 1500 ×g, washed with cold medium to remove any remaining traces of formaldehyde and resuspended in 10 mL spheroplast buffer (1 M sorbitol, 5 mM DTT, 250 U zymolyase 100T (US Biological, Z1004)) for 40 min at 30°C. Spheroplasts were harvested by centrifugation at 4°C for 10 min at 2500 ×g, washed with 10 mL of cold 1 M sorbitol and resuspended in 2 mL of 0.5% SDS at 65°C for 20 min. 125 Units of MboI (NEB, R0147M) were used for chromatin fragmentation in a final reaction volume of 3 mL and an incubation time of up to 16 hours at 37°C. The digested chromatin was centrifuged at 18000×g for 20 min and the pellet was resuspended in 200 μL of cold water. 5' MboI sticky ends were blunted, using a dNTPs mix that contains biotin-14-dCTP (Invitrogen, 19518018), by Klenow enzyme (NEB, M0210L) at 37°C for 80 min.

Biotinylated DNA fragments were then ligated using 60 Units T4 ligase (Thermo Scientific, 10621441) in 1.2 mL ligation reaction volume at room temperature for 2 hours on a gently rotating wheel. Ligation product was reversed cross-linked by 0.5 mg/mL proteinase K (Thermo Scientific, EO0492) in 0.5% SDS and 25 mM EDTA at 65°C for 4 hours, then, the genomic DNA was recovered using ethanol precipitation and purified using DNA recovery kit (Zymo Research, D4046). The sequencing library was prepared following the protocol of Illumina sequencing library prep kit (NEB, E7805), where the DNA was fragmented using NEBNext Ultra II FS Enzyme mix and incubation for 18 min at 37°C. DNA library was sequenced using an Illumina NextSeq 500 and MidOutput-150 kit with the 75,75 paired pair-end protocol.

Hi-C sequencing alignment and analysis

To generate contact maps: ~40-50 M 75 bp long paired-end reads were processed using the HICLib algorithm[103] adapted for *S. cerevisiae* genome, S288C available on SGD (https://www.yeastgenome.org/strain/s288c). Read-pairs were independently mapped using Bowtie 2[90] (mode: –very-sensitive –rdg 500,3 –rfg 500,3) on the corresponding reference sequence (S288C and/or with the corresponding versions of the synthetic chromosomes) indexed for MboI restriction site. The unwanted restriction fragments (RFs) were filtered out (e.g., loops, non-digested fragments),[104] whereas, the valid RFs were binned into units of fixed size of 5 kb. Additional filtering was done to remove bins with a high variance in contact frequency (<1.5 S.D. or 1.5–2 S.D.), to minimize biases in the contact map coming from the uneven distribution of restriction sites, GC content and mappability. The filtered contact maps were normalized using the sequential component normalization procedure (SCN).[104] Approximately 10 million valid contacts were used to generate a genomic contact map for each technical triplicate.

Contact probability between genomic loci was calculated as a function of their distance along the genome, using the parameter p(*s*).[101] Hi-C contact probability (p) decreases as the genomic distance (*s*) between restriction fragments increases. p(*s*) plots were computed on intrachromosomal read pairs from which self-circularizing and uncut events were discarded.[104] The retained reads were log-binned in function of their distance along chromosome arms. In other words, p(*s*) shows the distribution of the sum of contacts weighted by both bin-size $1.1^{(1+bin)}$ and chromosome length (*s*). Comparison of the degree of p(*s*) decay is indicative of a change in polymer state.

For visualization, we generate the 3D representations using the "Shortest-path Reconstruction in 3D" (ShRec3d)[105] algorithm, with the exact specifications described before.[102] Briefly, the algorithm first computes a distance matrix by inverting element-wise the corresponding 5 kb-binned contact map that was previously filtered and normalized as described earlier. The resulting matrix is supplemented by calculating the shortest path distances. This procedure removed infinite values and yielded the derived values that satisfy the triangular inequality. To extract the 3D coordinates (x, y, z), the distance matrix was subjected to the Sammon mapping algorithm. Subsequently, the "mat2pdb" function available in Matlab 2018 was employed to generate the pdb files. Visualization of these files was accomplished using PyMOL (Molecular Graphics System, Version 2.0, Schrödinger, LLC). It's important to note that the 3D structures derived are averaged representations of the contact maps. As such, they do not showcase exact structures identifiable in individual cells. Interpretations of these maps should be made considering the contact frequencies across vast cell populations.

RNA extraction for transcript profiling

Total RNA was extracted from 50mL flash-frozen cell pellets grown to mid-log (OD_{600} ~0.65-0.85) using MasterPure yeast RNA purification kit (Lucigen Cat# MPY03100) including a DNaseI treatment step. RNA (diluted 1:10) quality and concentration were measured by Agilent 2100 Bioanalyzer with the Agilent RNA 6000 Nano Kit (Cat# 5067-1511) and Qubit RNA High Sensitivity Kit (Thermo Fisher Cat# Q32852), respectively.

Direct RNA sequencing

Poly(A) mRNA was enriched from 93.75 μg total RNA on 250 μL Dynabeads oligo$(dT)_{25}$ beads. The Direct RNA Sequencing kit (SQK-RNA002, Oxford Nanopore Technology) was used to generate libraries from 500 ng poly(A) RNA. An optional reverse transcription was performed at 50°C for 50 min using SuperScript IV reverse transcriptase (Invitrogen Cat# 18090050) in between the ligation of the RTA and RMX adaptors. Following reverse transcription the RNA:cDNA was cleaned up with 1.8 volumes of Agencourt RNAclean XP beads and washed with 70% ethanol. Following RMX ligation only 1 volume of beads were used in the clean-up, and wash buffer (in the kit SQK-RNA002) was used in the wash steps. Direct RNA libraries (typically 150-200 ng) were loaded onto primed (EXP-FLP001) MinION flow cells (FLO-MIN106D, R9 version) in RRB buffer and run on the GridION with MinKNOW 3.1.8 for up to 72 hours.

Directional mRNA sequencing

NEBNext Poly(A) mRNA magnetic isolation module (NEB Cat# E7490) was used to enrich poly(A) mRNA from 500 ng total RNA with 5 μL 1:500 diluted ERCC RNA Spike-In control mix (Thermo Fisher Cat# 4456740) in 50 μL. The NEBNext Ultra II directional RNA library prep kit for Illumina with sample purification beads (NEB Cat# E7765) was used to prepare stranded mRNA sequencing libraries from the poly(A) RNA. Libraries were amplified for 11 cycles with i7 index primers (NEB Cat# E7500S). Libraries were individually cleaned-up with 0.9 volumes of sample purification beads and concentration and size distributions were measured by Qubit dsDNA high sensitivity kit and by Agilent 2100 Bioanalyzer with the Agilent High Sensitivity DNA kit (Cat# 5067-4626). Equimolar

amounts were combined of all samples, cleaned-up on 0.9 volumes of sample purification beads, and submitted for 150 bp paired end sequencing on the NextSeq 500 (Illumina) at the EMBL Genomics Core.

Base calling, quality-filtering, and long-read alignment

Nanopore long reads were base-called, trimmed of adapter sequences, and filtered for quality, retaining only those with the best alignment scores for multi-mapping reads, as previously described.[48]

Gene expression quantification

NEBNext Ultra II directional mRNA was quantified at the transcript-level by Salmon v1.6.0,[88] aligning to known transcripts as well as those identified in the long-read sequencing. When observing mature mRNA transcripts, reads were aligned against a database of transcripts with and without introns. Salmon was run with sequence and position bias modelling enabled. Differential gene expression analysis was performed in DESeq2.[89]

Chromosome substitution to consolidate *synIV*

We developed a method to directly consolidate individual chromosome with other multiple synthetic chromosomes. In the recipient strain with 6.5 synthetic chromosomes, *can1Δ0*, *lys4Δ0* with ORF deleted, *cyh2* mutation (Q38K) were introduced using stepwise CRISPR/Cas9 editing with donor DNA provided. Native *chrIV* was targeted with the *KlURA3-pGAL-CEN4* module. The *kar1-1* mutation (P150S) was introduced using CRISPR/Cas12a, generating the final recipient strain YZY402.

The donor strain (yWZ675, carrying *synIV*, yeast_chr04_9_03) and recipient strain (YZY402, carrying 6.5 synthetic chromosomes) were prepared as fresh patches (~2 cm in diameter) on separate YPD plates incubated overnight at 30°C, and then mated together by replica plating. After incubation at 30℃ for 12 h, the mating plate was replica-plated to a selection plate of SC–Lys+Can (60 ng/μl, Sigma Cat# C9758) +cycloheximide (10 ng/μl, Sigma Cat# C7698). After incubation at 30°C for a week, haploid progeny with both *chrIV* and *synIV* (n+1) were successfully obtained and re-streaked to fresh selection plates, which were then checked by PCRtag assays. The efficiency for *synIV* transfer was around 10% (2 out of 23 screened). Finally, the strain was incubated in YP+galactose medium to destabilize the native chromosome and selected on 5-FOA plates, generating the final haploid strain.

DNA content assay

We used a previously described DNA content assay.[106] Briefly, about 5×10^6 cells were fixed in ethanol 70% for 1 h at room temperature, then pelleted, washed, and incubated in 10 mM Tris pH 7.5 with RNase A (0.1 mg/ml, Thermo Fisher EN0531) for 2 h at 37°C. Cells were pelleted, resuspended in 10 mM Tris pH7.5 with propidium iodide (5 μg/ml, Thermo Fisher P3566), and incubated for 1 h in dark at 4°C. Finally, the cells were pelleted and resuspended in 0.5 ml 50 mM Tris pH 7.5, and analyzed using a BD AccuriTM C6 flow cytometer.

Scanning electron microscopy

Cultured yeast cells were plated on 12mm poly-l-lysine coated glass coverslip in 24 well dish and fixed with 2.5% glutaraldehyde in PBS for one hour. After washing with PBS, the yeast cells were post fixed in 1% Osmium tetroxide for one hour, dehydrated in a series of ethanol solutions (30%, 50%, 70%, 85%, 95%, 100%), and dried with a Tousimis autosamdsri 931 (Rockville, MD) critical point dryer. The cover slips were put on SEM stubs, sputter coated with gold/palladium by DESK V TSC HP Denton Vacuum (Moorestown, NJ), and imaged by a Zeiss Gemini300 FESEM (Carl Zeiss Microscopy, Oberkochen, Germany) using secondary electron detector at 3 kV at a working distance of 10 mm.

CAN1 resistance frequency

We added 100 μl of overnight culture containing approximately 2×10^7 cells of syn6.5 or BY4742 (as determined by a hemocytometer) to SC–Arg w/ canavanine plates (60 ng/μl) and counted the number of colonies appearing after two days of incubation at 30℃, with three biological replicates. An unpaired t-test was used to determine statistical significance.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantitative RT-PCR analysis for transcript levels

Quantitative RT-PCR was used to assess the transcript levels of *SWI3* (Figure 5D) and *SHM1* (Figure S3E). Detailed experimental procedures are provided in the STAR Methods section. Three single colonies were tested as triplicates and the expression levels were analyzed using the LightCycler 480 System. Ct values were normalized to *TAF10* as the internal control.[99] Error bars represent mean ± SD of three replicates. $p<0.001$ (***), or 0.01 (**) were calculated by an unpaired t-test.

Quantitative genome coverage

We calculated the WGS coverage as shown in Figures 4A, S1D, and S4E. Following the alignment (described in STAR Methods), we calculated the sequencing coverage for each base pair using BEDTools, normalized to the mean value of genome-wide coverage and then plotted using ggplot.

Quantitative Hi-C heat map and contact frequency

We used Hi-C to generate a heatmap showing the contact probability (log_{10}) between pairs of chromosomal sites (Figure 6). The detailed calculation procedures are described in the STAR Methods section under Hi-C sequencing alignment and analysis.

Quantitative analysis of RNA transcriptomes

Quantitative analysis details for transcript start sites (TSSs) and transcript end sites (TESs) (Figure 6E), expression levels (Figures 6F, 6G, and 6H) of genes on synthetic chromosomes compared to native chromosomes, are described in corresponding figure legends.

CellPress
OPEN ACCESS

Supplemental figures

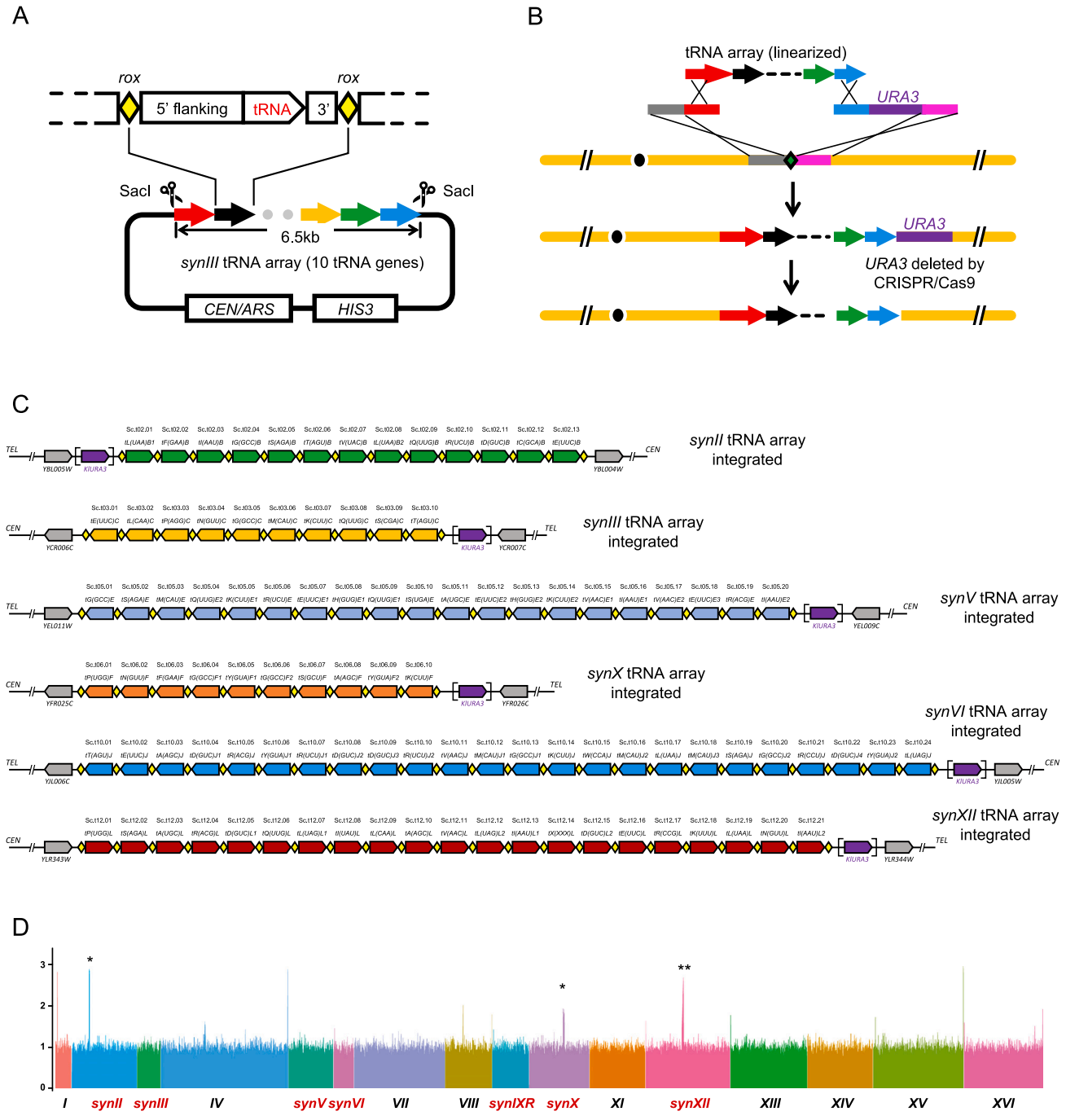


Figure S1. Integration of tRNA arrays into each synthetic chromosome, related to Figure 1

(A) Each tRNA array contains all the tRNA genes from its native chromosome. The array is stored in a plasmid and can be released by restriction enzyme digestion. The *synIII* tRNA array is shown as an example here.

(B) The tRNA array was integrated into the synthetic chromosome using a two-step homologous recombination method. First, it was integrated and selected with *URA3* marker. Second, the *URA3* marker was deleted using CRISPR/Cas9 with donor provided. The restriction enzymes to linear tRNA array, and integration loci are listed in Table S1.

(legend continued on next page)

(C) Anatomy of tRNA arrays integrated in synthetic chromosomes. Each tRNA gene is assigned a unique ID, showing its host chromosome and gene number. For example, Sc.t02.01 indicates the first synthetic tRNA gene in *S. cerevisiae* from chromosome *II*. Each tRNA array was integrated into its host synthetic chromosomes. *KlURA3* (purple in square bracket) was used as an integration marker, which was then deleted by CRISPR/Cas9. The rox sites (yellow diamond) were also integrated, enabling tRNA array rearrangements in future applications.

(D) Sequencing coverage in the WGS for multiple synthetic chromosomes. The presence of complete multiple synthetic chromosomes (*synII*, *synIII*, *synV*, *synVI*, *synIXR*, *synX*, and *synXII*) was confirmed by WGS. * Duplications of tRNAarray happened during their initial integration, which was later repaired by CRISPR/Cas9. ** Higher coverage of rRNA repeats was detected in *synXII*.

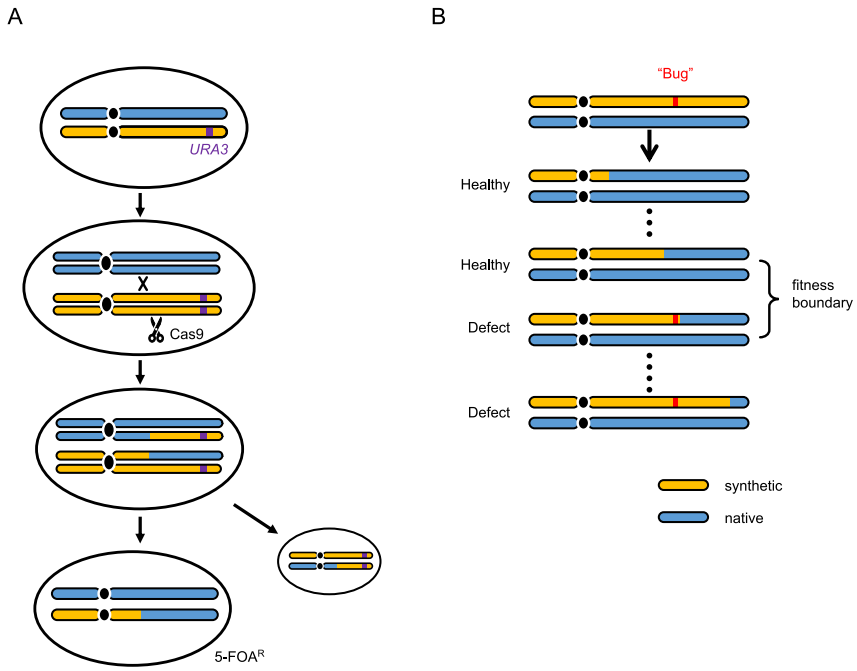


Figure S2. CRISPR D-BUGS can also be used to map dominant bugs, related to Figure 2

(A) Compared with Figure 2, *URA3* marker (purple) is pre-integrated into the synthetic allele (orange), instead of the wild-type allele. gRNAs are selected to target Cas9 cleavage at synthetic PCRtags.

(B) A series of strains is generated, in which the defect phenotype indicates the presence of the dominant bug in the synthetic sequence.

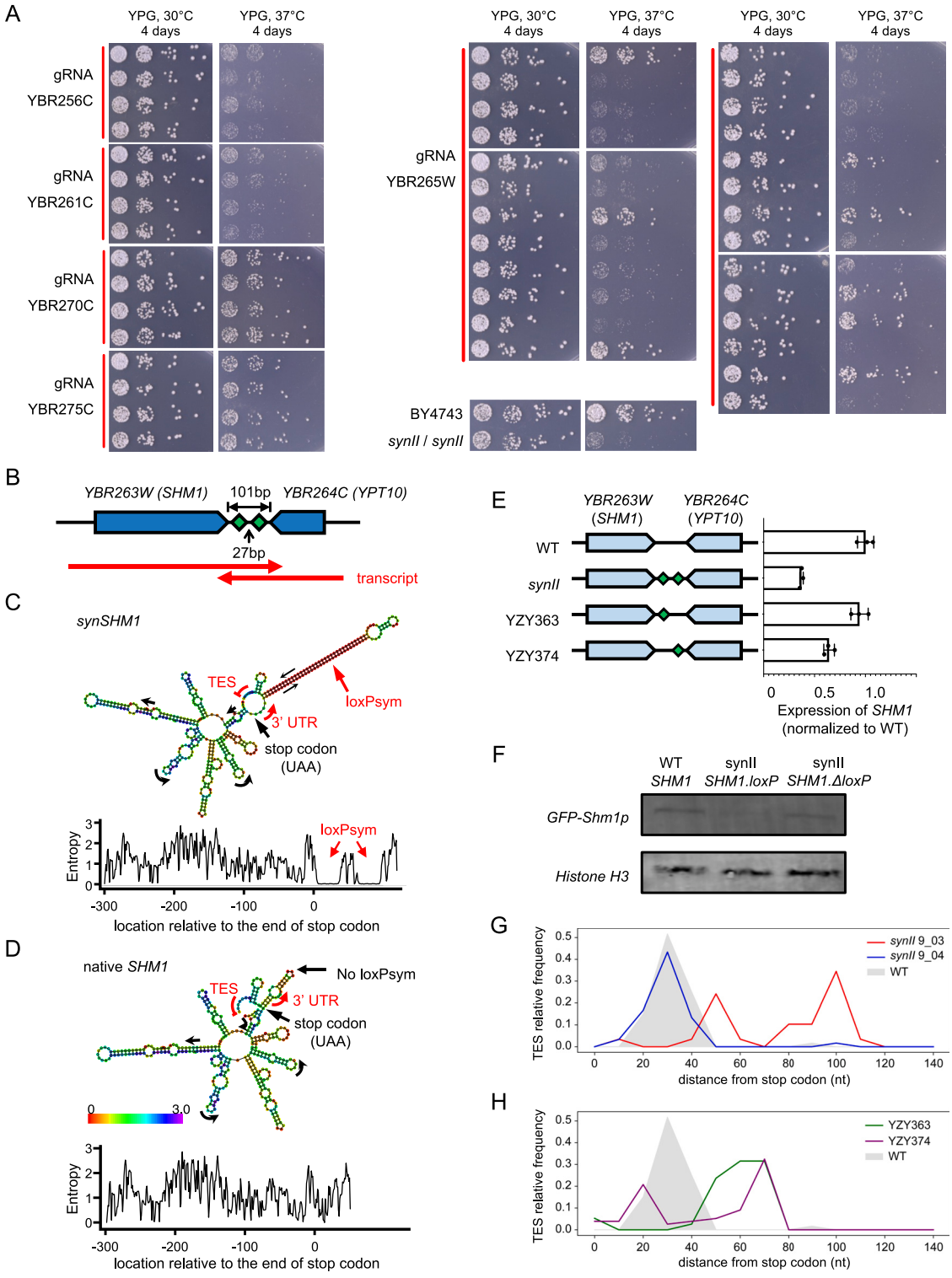


Figure S3. Bug mapping for *synII* with CRISPR D-BUGS, related to Figure 3

(A) Fitness of single colonies generated from *synII* bug mapping. The gRNA targets are labeled on the left side.

(B) Two loxPsym sites between *SHM1* and *YPT10*, with overlapping transcripts.

(C) The predicted RNA secondary structure (left) and its entropy (right) for *synSHM1* with two loxPsym sites in its 3′ UTR. The calculation was using ViennaRNA package based on minimum free energy (MFE) model. The sequence from 300 nt upstream of stop codon to the transcript end was used. TES, transcript end site.

(legend continued on next page)

(D) Same calculation performed using native *SHM1* without loxPsym sites.‖
(E) Real-time PCR to check the transcript level of *SHM1*. YZY363, *synII* strain with only left side loxPsym site. YZY374, *synII* strain with only right side loxPsym site.
(F) Western blot to check level of N-terminally GFP-tagged Shm1p. Histone H3 was used as the internal control.
(G) The TES distributions of the *YPT10* transcript, in original *synII* with two loxPsym sites (9_03) and updated version with both loxPsym sites deleted (9_04).
(H) The same measurements for the *YPT10* transcript with either of loxPsym sites, in the *synII* of YZY363 and YZY374, are also shown in Figure 3D.

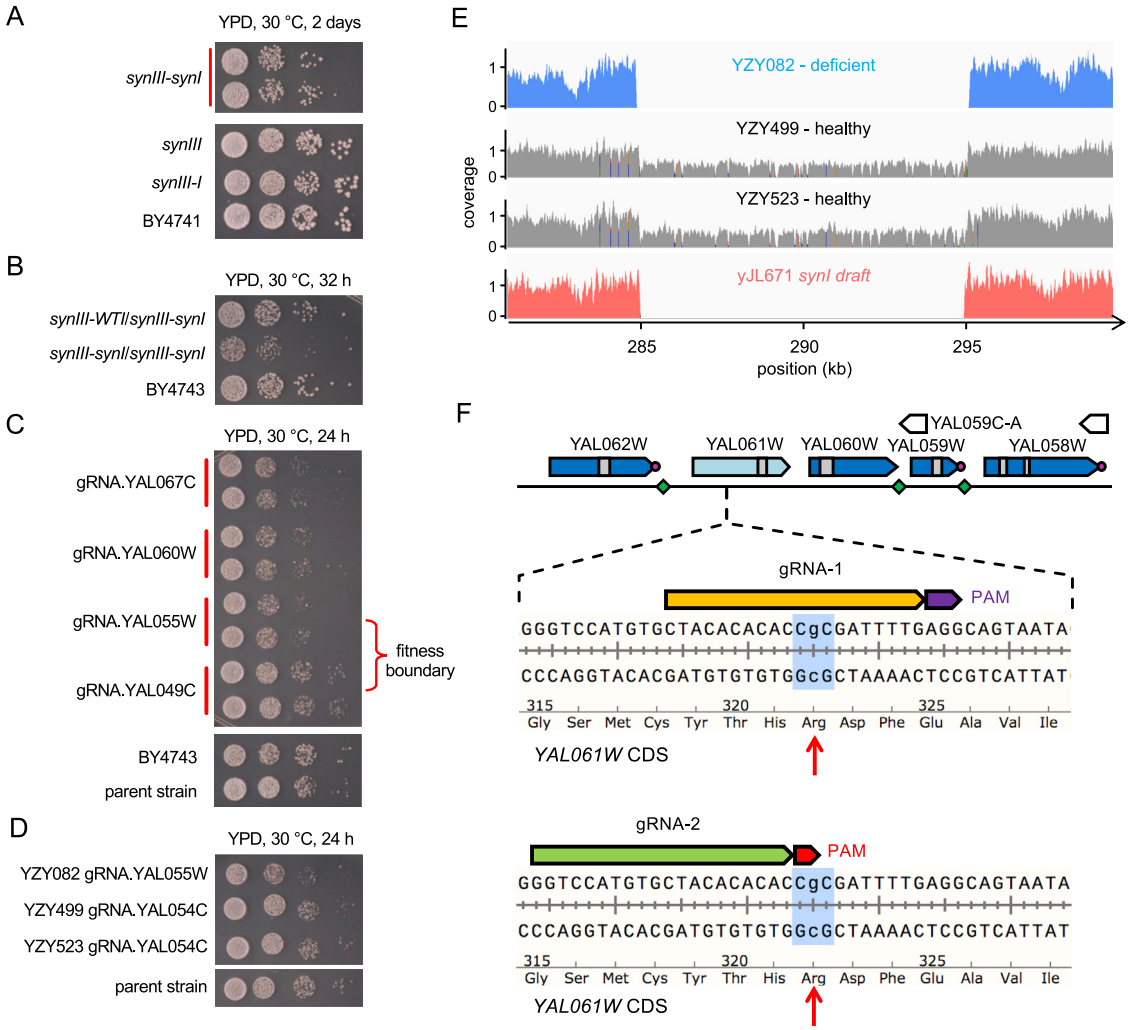


Figure S4. Bug mapping for *synI* with CRISPR D-BUGS, related to Figure 4

(A) The fitness assay showing growth of serially diluted draft *synIII-synI* strain (yJL671). *synIII*, the strain containing separate *synIII* and wild-type chr *I*. *synIII-I*, the strain containing *synIII* fused with wild-type chr *I*.

(B) The fitness assay performed on diploids to show that the defect is recessive.

(C) The fitness assay to find the fitness boundary. Two single colonies were tested for each gRNA selected.

(D) The spot assay to check the fitness of strains generating using gRNA.YAL055W (YZY082) and gRNA.YAL054C (YZY499 and YZY523). Parent strain, the diploid strain with heterozygous chr *I*.

(E) Coverage of *synI* genome from WGS of strains from (D). YZY082 has no coverage, whereas YZY499 and YZY523 only have half coverage and misaligned genome feature, indicating one allele is wild type. y axis, the coverage normalized to the average genomic coverage level. x axis, the position starting from the left end of *synIII-synI* fusion chromosome.

(F) The missense mutation in *YAL061W* CDS (red arrow) and the gRNA-1 used in the first trial to repair this mutation in *synI* draft strain.

(G) After the defect was mapped, we selected gRNA-2, compatible with SpCas9-NG to repair the mutation, generating the final and healthy *synI* strain yCTC002.

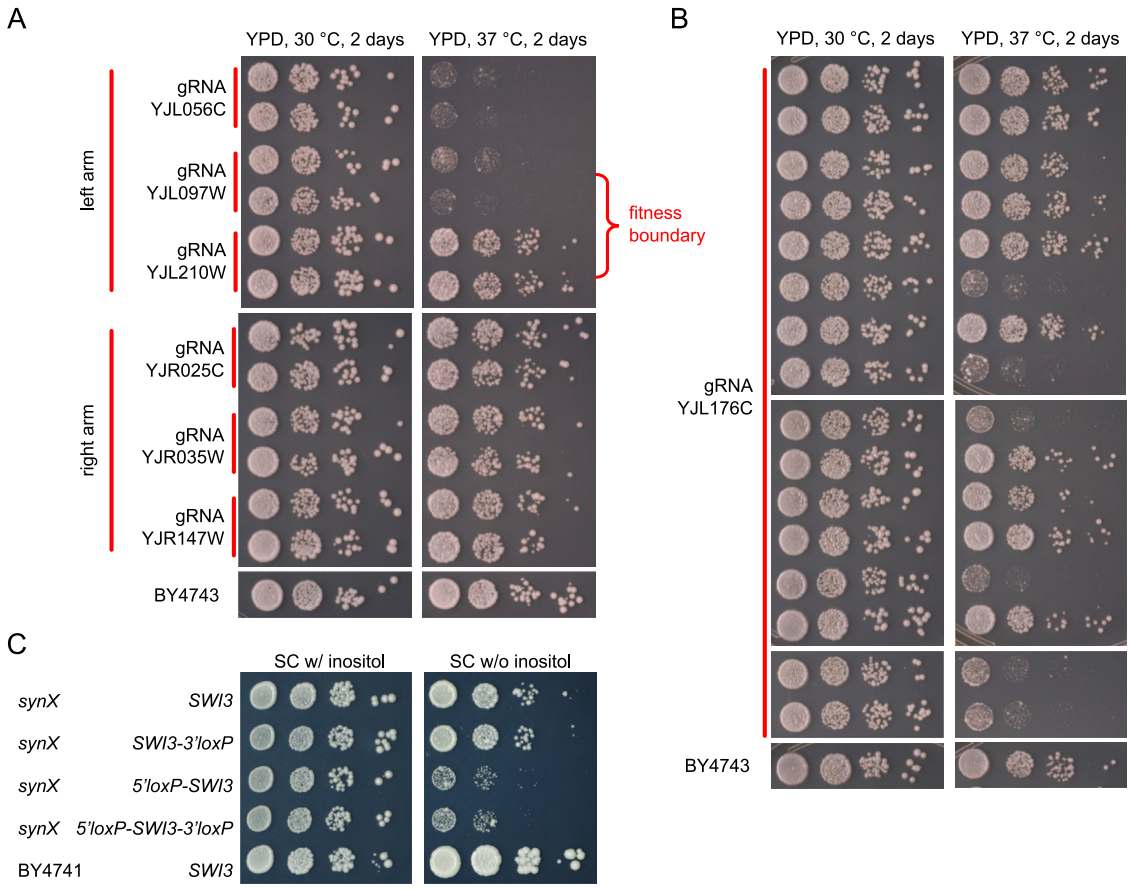
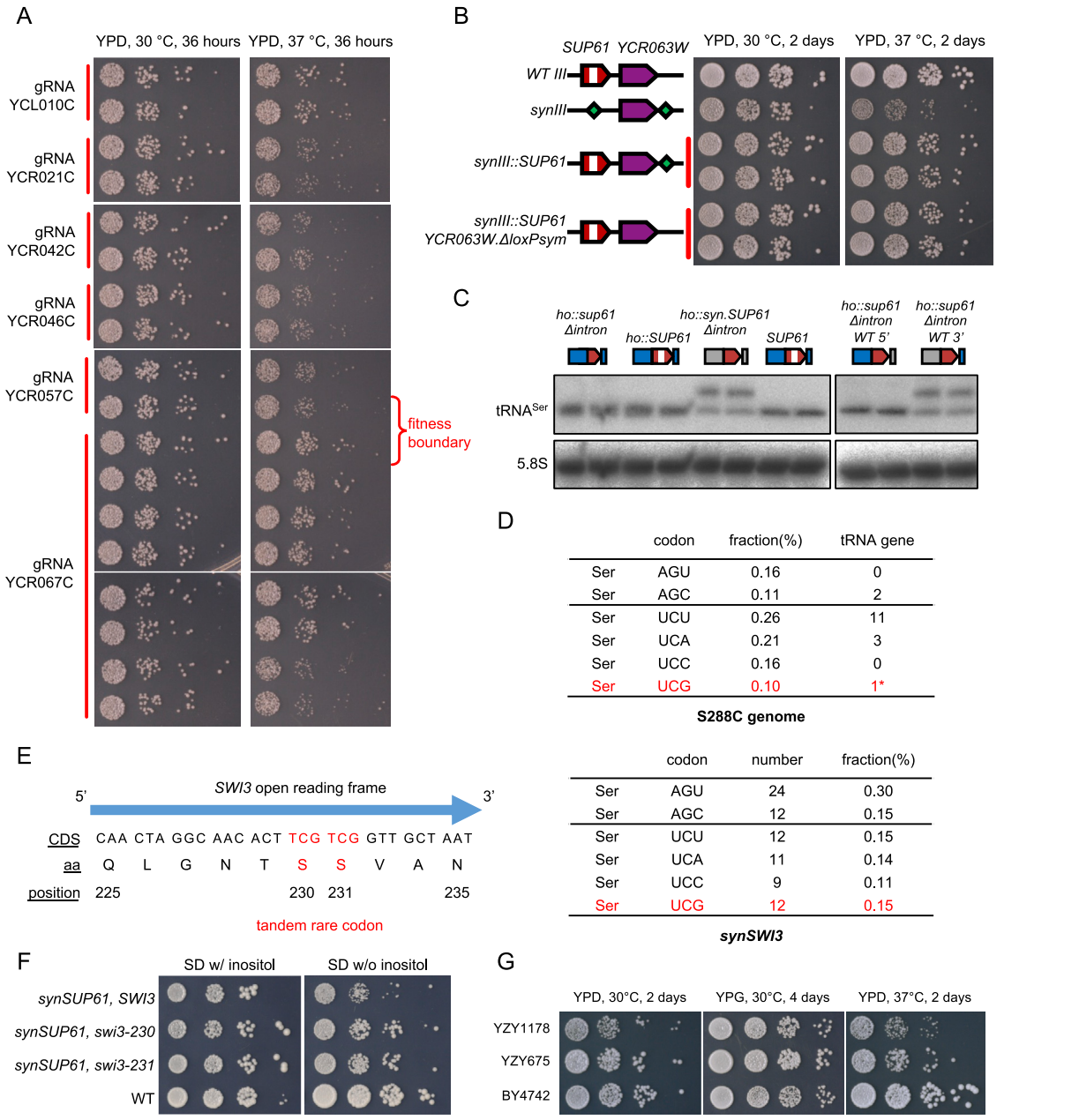


Figure S5. Bug mapping for *synX* with CRISPR D-BUGS, related to Figure 5

(A) We first selected 6 spatial distributed gRNAs targeting the left arm (left) and right arm (right) for CRISPR D-BUGS. The fitness assay shows that the bug is located in the left arm, and we found the fitness boundary between *YJL097W* and *YJL210W*.

(B) Fine mapping using one single gRNA targeting at wild-type *YJL176C*. Individual colonies were generated using the same gRNA.YJL176C but showed a diverse level of fitness. These colonies were whole-genome sequenced and aligned to *synX* as the reference to map their recombination sites as in Figure 5B.

(C) Yeast strains with original *synSWI3* in *synX* (5'loxP-SWI3-3'loxP), compared with the strain with 3'loxP deleted (*5'loxP-SWI3*), 5'loxP deleted (*SWI3-3'loxP*) or both deleted (*SWI3*). The plates were incubated for 3 days at 30°C.



	codon	fraction(%)	tRNA gene
Ser	AGU	0.16	0
Ser	AGC	0.11	2
Ser	UCU	0.26	11
Ser	UCA	0.21	3
Ser	UCC	0.16	0
Ser	UCG	0.10	1*

S288C genome

	codon	number	fraction(%)
Ser	AGU	24	0.30
Ser	AGC	12	0.15
Ser	UCU	12	0.15
Ser	UCA	11	0.14
Ser	UCC	9	0.11
Ser	UCG	12	0.15

synSWI3

Figure S6. Bug mapping for *synIII* with CRISPR D-BUGS and combinatorial interactions with *synX*, related to Figure 5

(A) Initial mapping located the bug to the right arm of *synIII*. By selecting more gRNAs, we found the fitness boundary between *YCR057C* and *YCR067C*. The single colonies generated from gRNA.YCR067C already showed a diverse level of fitness.

(B) CRISPR D-BUGS already mapped the bug to two loxPsym site (green diamond) in *synIII*. The left one was the landmark for *SUP61* deletion and the right one was integrated downstream of *YCR063W*. The integration of *SUP61* (*synIII::SUP61*) was sufficient to rescue the defect.

(C) The $tRNA_{Ser}^{CGA}$ species was expressed from different version of *syn.SUP61*. Two single colonies were tested in each group as replicates. Red, *SUP61* coding sequence. Gray, flanking sequences from *Eremothecium* (*Ashbya*) *gossypii*. Blue, the native sequences from wild-type *S. cerevisiae*. White, intron in *SUP61*.

(D) The codon usage for serine in wild-type yeast genome and *SWI3* at *synX*. In yeast, UCG is a rare codon (10% of serine) that is decoded by the only tRNA species expressed from *SUP61*, whereas UCU is a rich codon (26% of serine), and there are 11 copies of the tRNA gene that recognizes this codon.

(E) Either of tandem rare codons was mutated from UCG to UCU, named *swi3-230* and *swi3-231*, respectively.

(F) Spot assay to check the effect of codon swap on inositol auxotrophy, in the background strain of *synIII*, *synX* with *SWI3* loxPsym site deleted.

(G) The draft strain containing *synII*, *synIII*, *synV*, *synVI*, *synIXR*, *synX*, and *synXII* is YZY1178, which is also shown in Figure 1D. YZY675 was generated by repairing all know bugs, including the *SHM1* bug in *synII* and the combinatorial bug between *SWI3* in *synX* and *synSUP61* in *synIII*, generating YZY675.

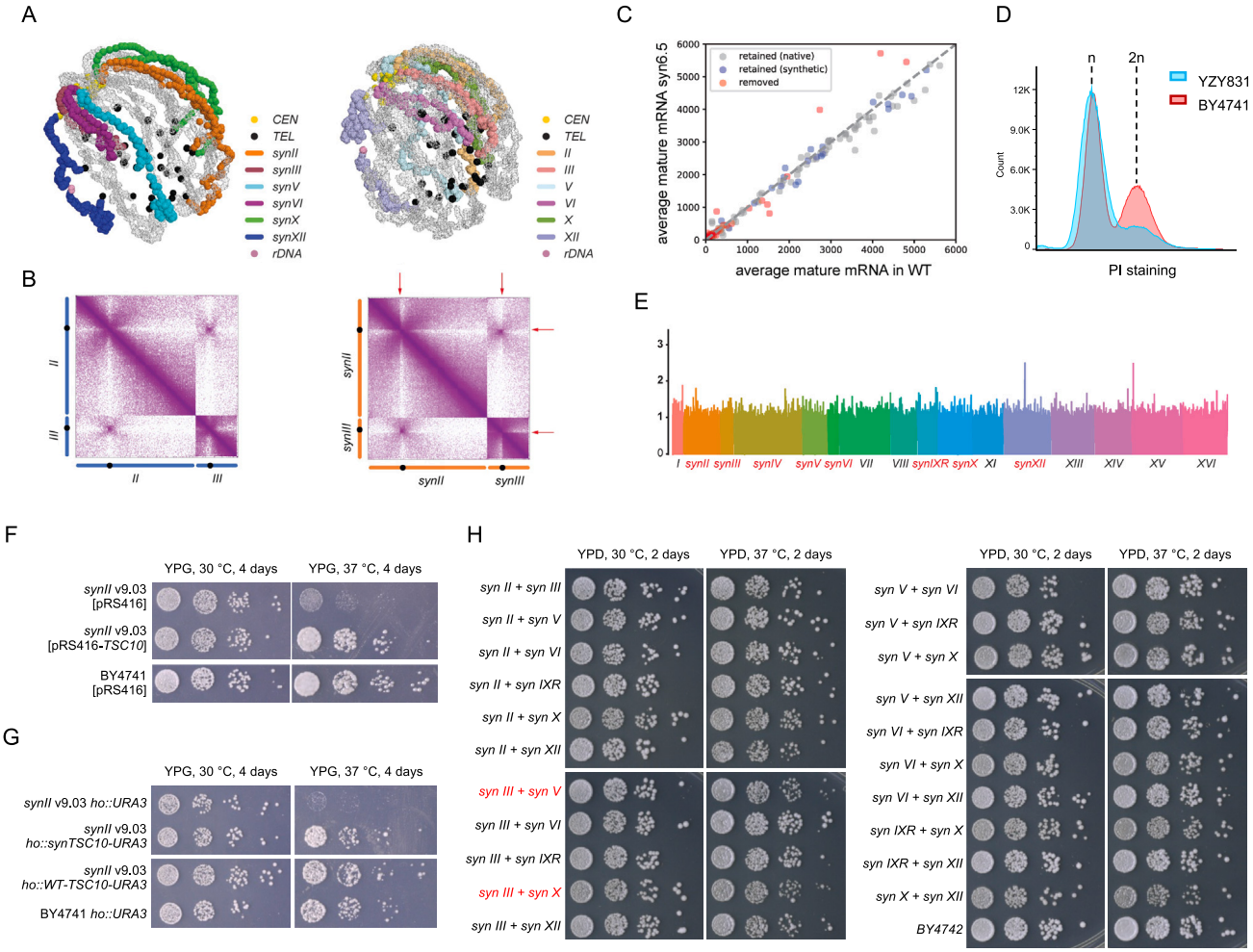


Figure S7. Characterization of multiple synthetic chromosomes, related to Figure 6

(A) The 3D chromosome trajectories of multiple synthetic chromosomes (left), compared with wild-type chromosomes (right). Gray, all other native chromosomes. A movie was also generated using the same labels in Data S1.

(B) Contact frequency heatmap for *synII* and *synIII*. Heatmaps for native (left) and synthetic (right) chromosomes *II* and *III*. The boundaries at the tRNA array integration loci are highlighted with red arrows.

(C) Expression levels of intron-containing genes in syn6.5 strain vs. wild-type control. Mean salmon quantification from three replicates of Illumina stranded mRNA sequencing of the syn6.5 and wild-type strains are compared for intron-containing genes. Genes that retained their introns in the syn6.5 strain on the native or synthetic chromosomes are colored in gray and blue, respectively, whereas genes with intron deletion are shown in red. x and y axis: transcripts per million.

(D) Check the DNA content and sequence of the strain YZY949 containing *synII*, *synIII*, *synIV*, *synV*, *synVI*, *synIXR*, *synX*, and *synXII*.

(E) Coverage from the WGS of YZY949, aligned to the reference sequence of multiple synthetic chromosomes.

(F) Spot assay for the original *synII* strain (v9.03) transformed with empty vector (pRS416) or second copy of wild-type *TSC10* under the control of its native promoter (pRS416-TSC10). BY4741 was transformed with the empty vector as the control.

(G) Spot assay for the *synII* v9.03 with either only *URA3*, a second copy of *TSC10* with synthetic PCRtags, or a copy of *TSC10* with wild-type PCRtags integrated at the *HO* locus. BY4741 was integrated with one copy of *URA3* at the *HO* locus as a control.

(H) Fitness assay for the strains with two synthetic chromosomes. Using endoreduplication intercross, we constructed strains with all possible combinations of two synthetic chromosomes and checked their growth on YPD. The *synIII*, *synV* and *synIII*, *synX* strain (in red) still showed a slight growth defect at high temperature.