

## RESEARCH ARTICLE SUMMARY

## SYNTHETIC BIOLOGY

# Deep functional analysis of *synII*, a 770-kilobase synthetic yeast chromosome

Yue Shen,\* Yun Wang,\* Tai Chen,\* Feng Gao, Jianhui Gong, Dariusz Abramczyk, Roy Walker, Hongcui Zhao, Shihong Chen, Wei Liu, Yisha Luo, Carolin A. Müller, Adrien Paul-Dubois-Taine, Bonnie Alver, Giovanni Stracquadanio, Leslie A. Mitchell, Zhouqing Luo, Yanqun Fan, Baojin Zhou, Bo Wen, Fengji Tan, Yujia Wang, Jin Zi, Zexiong Xie, Bingzhi Li, Kun Yang, Sarah M. Richardson, Hui Jiang, Christopher E. French, Conrad A. Nieduszynski, Romain Koszul, Adele L. Marston, Yingjin Yuan, Jian Wang, Joel S. Bader, Junbiao Dai, Jef D. Boeke, Xun Xu, Yizhi Cai,† Huanming Yang†

**INTRODUCTION:** Although much effort has been devoted to studying yeast in the past few decades, our understanding of this model organism is still limited. Rapidly developing DNA synthesis techniques have made a “build-to-understand” approach feasible to reengineer on the genome scale. Here, we report on the completion of a 770-kilobase synthetic yeast chromosome II (*synII*). *SynII* was characterized using extensive Trans-Omics tests. Despite considerable sequence alterations, *synII* is virtually indistinguishable from wild type. However, an up-regulation of translational machinery was observed and can be reversed by restoring the transfer RNA (tRNA) gene copy number.

**RATIONALE:** Following the “design-build-test-debug” working loop, *synII* was successfully

designed and constructed in vivo. Extensive Trans-Omics tests were conducted, including phenomics, transcriptomics, proteomics, metabolomics, chromosome segregation, and replication analyses. By both complementation assays and SCRaMBLE (synthetic chromosome rearrangement and modification by *loxP*-mediated evolution), we targeted and debugged the origin of a growth defect at 37°C in glycerol medium.

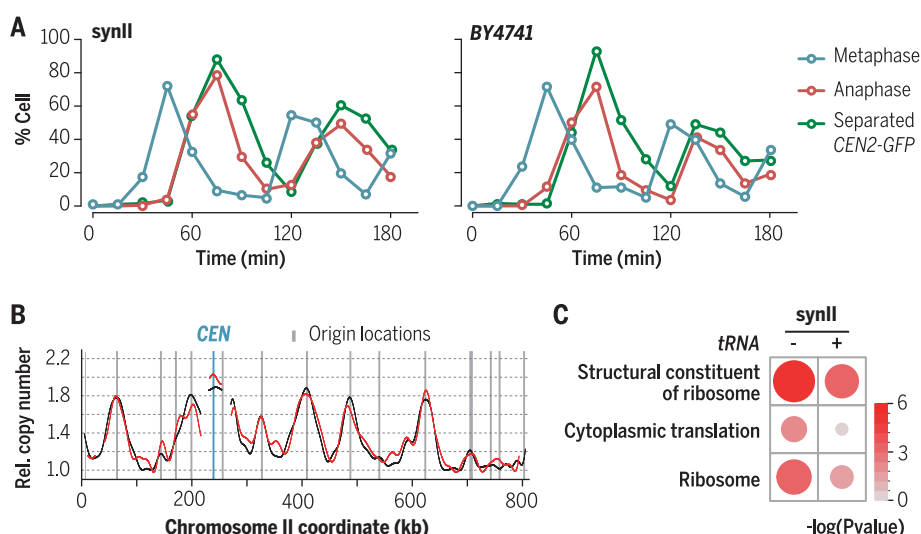
**RESULTS:** To efficiently construct megabase-long chromosomes, we developed an *I-SceI*-mediated strategy, which enables parallel integration of synthetic chromosome arms and reduced the overall integration time by 50% for *synII*. An *I-SceI* site is introduced for generating a double-strand break to promote targeted homologous recombination during mitotic growth. Despite

hundreds of modifications introduced, there are still regions sharing substantial sequence similarity that might lead to undesirable meiotic recombinations when intercrossing the two semi-synthetic chromosome arm strains. Induction of the *I-SceI*-mediated double-strand break is otherwise lethal and thus introduced a strong selective pressure for targeted homologous recombination. Since our strategy is designed to generate a markerless *synII* and leave the *URA3* marker on the wild-type chromosome, we observed a tenfold increase in *URA3*-deficient colonies upon *I-SceI* induction, meaning that our strategy can greatly bias the crossover events toward the designated regions.

By incorporating comprehensive phenotyping approaches at multiple levels, we demonstrated that *synII* was capable of powering the growth of yeast indistinguishably from wild-type cells (see the figure), showing highly consistent biological processes comparable to the native strain. Meanwhile, we also noticed modest but potentially significant up-regulation of the translational machinery. The main alteration underlying this change in expression is the deletion of 13 tRNA genes.

A growth defect was observed in one very specific condition—high temperature (37°C) in medium with glycerol as a carbon source—where colony size was reduced significantly. We targeted and debugged this defect by two distinct approaches. The first approach involved phenotype screening of all intermediate strains followed by a complementation assay with wild-type sequences in the synthetic strain. By doing so, we identified a modification resulting from PCRTag recoding in *TSC10*, which is involved in regulation of the yeast high-osmolarity glycerol (HOG) response pathway. After replacement with wild-type *TSC10*, the defect was greatly mitigated. The other approach, debugging by SCRaMBLE, showed rearrangements in regions containing HOG regulation genes. Both approaches indicated that the defect is related to HOG response dysregulation. Thus, the phenotypic defect can be pinpointed and debugged through multiple alternative routes in the complex cellular interactome network.

**CONCLUSION:** We have demonstrated that *synII* segregates, replicates, and functions in a highly similar fashion compared with its wild-type counterpart. Furthermore, we believe that the iterative “design-build-test-debug” cycle methodology, established here, will facilitate progression of the Sc2.0 project in the face of the increasing synthetic genome complexity. ■



**SynII characterization.** (A) Cell cycle comparison between *synII* and BY4741 revealed by the percentage of cells with separated CEN2-GFP dots, metaphase spindles, and anaphase spindles. (B) Replication profiling of *synII* (red) and BY4741 (black) expressed as relative copy number by deep sequencing. (C) RNA sequencing analysis revealed that the significant up-regulation of translational machinery in *synII* is induced by the deletion of tRNA genes in *synII*.

The list of author affiliations is available in the full article online.

\*These authors contributed equally to this work.

†Corresponding authors: Email: yizhi.cai@ed.ac.uk (Y.C.);

yanghm@genomics.cn (H.Y.)

Cite this article as Y. Shen et al., *Science* 355, eaaf4791 (2017). DOI: 10.1126/science.aaf4791

## RESEARCH ARTICLE

## SYNTHETIC BIOLOGY

# Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome

Yue Shen,<sup>1,2,10,11\*</sup> Yun Wang,<sup>1,11\*</sup> Tai Chen,<sup>1,11\*</sup> Feng Gao,<sup>1</sup> Jianhui Gong,<sup>1</sup> Dariusz Abramczyk,<sup>2</sup> Roy Walker,<sup>2</sup> Hongcui Zhao,<sup>1</sup> Shihong Chen,<sup>1</sup> Wei Liu,<sup>2</sup> Yisha Luo,<sup>2</sup> Carolin A. Müller,<sup>3</sup> Adrien Paul-Dubois-Taine,<sup>2</sup> Bonnie Alver,<sup>2</sup> Giovanni Stracquadanio,<sup>4,5,12</sup> Leslie A. Mitchell,<sup>6</sup> Zhouqing Luo,<sup>7</sup> Yanqun Fan,<sup>1</sup> Baojin Zhou,<sup>1</sup> Bo Wen,<sup>1</sup> Fengji Tan,<sup>1</sup> Yujia Wang,<sup>1</sup> Jin Zi,<sup>1</sup> Zexiong Xie,<sup>8</sup> Bingzhi Li,<sup>8</sup> Kun Yang,<sup>4</sup> Sarah M. Richardson,<sup>4,5</sup> Hui Jiang,<sup>1</sup> Christopher E. French,<sup>2</sup> Conrad A. Nieduszynski,<sup>3</sup> Romain Koszul,<sup>9</sup> Adele L. Marston,<sup>2</sup> Yingjin Yuan,<sup>8</sup> Jian Wang,<sup>1</sup> Joel S. Bader,<sup>5,6</sup> Junbiao Dai,<sup>7</sup> Jef D. Boeke,<sup>6</sup> Xun Xu,<sup>1,11</sup> Yizhi Cai,<sup>2†</sup> Huanming Yang<sup>1,10†</sup>

Here, we report the successful design, construction, and characterization of a 770-kilobase synthetic yeast chromosome II (synII). Our study incorporates characterization at multiple levels—including phenomics, transcriptomics, proteomics, chromosome segregation, and replication analysis—to provide a thorough and comprehensive analysis of a synthetic chromosome. Our Trans-Omics analyses reveal a modest but potentially relevant pervasive up-regulation of translational machinery observed in synII, mainly caused by the deletion of 13 transfer RNAs. By both complementation assays and SCRaMBLE (synthetic chromosome rearrangement and modification by *loxP*-mediated evolution), we targeted and debugged the origin of a growth defect at 37°C in glycerol medium, which is related to misregulation of the high-osmolarity glycerol response. Despite the subtle differences, the synII strain shows highly consistent biological processes comparable to the native strain.

**Y**easts are among the most important eukaryotic microorganisms, both scientifically and economically, with at least 1500 species identified to date. In the past few decades, a number of different yeast species have been sequenced, and comprehensive genomic

comparison studies have revealed answers to several complex evolutionary questions, such as whole-genome duplication and horizontal gene transfer during their history (1–5). However, despite the increased knowledge and understanding of the natural genome, our ability to engineer entire genomes is still very limited. Rapidly developing DNA synthesis techniques have made a “build-to-understand” approach feasible, which allows synthetic biologists to reengineer and construct genes, pathways, and even entire genomes (6–12).

Under the collaborative framework of the Sc2.0 project ([www.syntheticyeast.org](http://www.syntheticyeast.org)), we report the completion of yeast synthetic chromosome II, termed synII. The sequence of native *Saccharomyces cerevisiae* chromosome II was determined more than two decades ago. This native chromosome is 807,888 base pairs (bp) in length and includes 410 open reading frames, 13 tRNAs, and 30 introns (13). SynII was designed based on the native chromosome II following previously reported rules (12, 14, 15), resulting in a “designer” chromosome 770,035 bp in length, 43,149 bp shorter than the native sequence. SynII was initially synthesized as minichunks (~3 kb), assembled into chunks (~10 kb) and integrated into the yeast genome to replace the native chromosome. Extensive Trans-Omics tests were conducted, including phenomics, transcriptomics, proteomics, chromosome segre-

gation, and replication analyses, which indicated that synII yeast, despite significant sequence alterations, is virtually indistinguishable from wild type. However, our analyses reveal that up-regulation of many components of the translational machinery (for example, ribosomal proteins at both the RNA and protein levels) is a typical feature of synthetic chromosome replacement strains, and the main cause for this subtle defect was shown to be the reduction tRNA gene copy number; it was reversed by restoring the tRNA gene copy number.

## Design and synthesis of synII

Complete chromosome redesign of chromosome II was performed based on the Sc2.0 project standards (12) using the BioStudio design suite (14) to yield synII. Relative to the wild-type reference sequence, synII has 33 deletions (53,605 bp in total), 269 insertions (10,456 bp in total) and 14,949 single-nucleotide substitutions. Although synII represents a 5.3% reduction in sheer size compared to the native chromosome II, hundreds of designer features were encoded throughout this synthetic chromosome (table S1 and supplementary materials and methods). The extensive changes made to the chromosome naturally raise the question of whether there are any subtle phenotypic effects that may be missed using approaches described in previous studies.

Compared with previous synthetic strategies (12, 15), we devised an alternative modular assembly method for the construction of synII. This new approach enables parallel integration to speed assembly, while at the same time providing an effective route to debug potential phenotypic defects (see below). We divided synII into 25 megachunks (~30 kb each) using BioStudio, and each megachunk was then segmented by an in-house software “Segman” into minichunks (~3 kb) (fig. S1A), which are compatible with the Gibson assembly method (16) (fig. S1B) for construction into chunks (detailed information on megachunks, chunks, and minichunks can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)). Instead of performing step-by-step megachunk integration (SwAP-In) (15) from one end (left or right arm), we performed integration in parallel from both ends with two opposite mating type parental strains, YS000-L and YS000-R (fig. S1C), bearing synII-L and synII-R, respectively, which overlap by 30 kb, effectively reducing the overall integration time by almost 50% by parallelizing the integration effort. These two strains were crossed to produce a heterozygous diploid. Despite hundreds of modifications being introduced in synII, regions sharing substantial sequence similarity to the native chromosome remain that provide a template for unwanted meiotic recombination events. Following meiosis, the resulting haploid strains, many of which would be synthetic-wild-type hybrids, would require tremendous screening efforts to identify cells carrying full-length synII chromosomes. Therefore, we adopted an I-SceI-mediated strategy (17) to break synII-L and synII-R at a designed site, promoting mitotic recombination between them to generate the fully synthetic synII (fig. S1D). Induction of the lethal I-SceI-mediated double-strand DNA breaks at the junctions of

<sup>1</sup>BGI-Shenzhen, Shenzhen 518083, China. <sup>2</sup>School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK. <sup>3</sup>Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE, UK. <sup>4</sup>High-Throughput Biology Center, School of Medicine, Johns Hopkins University, Baltimore, MD 21205, USA. <sup>5</sup>Department of Biomedical Engineering, School of Engineering, Johns Hopkins University, Baltimore, MD 21218, USA. <sup>6</sup>Institute for Systems Genetics, New York University Langone Medical Center, ACLSW Room 503, 430 East 29th Street, New York, NY 10016, USA. <sup>7</sup>Key Laboratory of Industrial Biocatalysis (Ministry of Education), Center for Synthetic and Systems Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China. <sup>8</sup>Key Laboratory of Systems Bioengineering (Ministry of Education), SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300072, PR China. <sup>9</sup>Department of Genomes and Genetics, Institut Pasteur/CNRS UMR3525, 25-28, Rue du Docteur Roux, 75724 Paris Cedex 15, France. <sup>10</sup>James D. Watson Institute of Genome Sciences, Hangzhou 310058, China. <sup>11</sup>BGI-Qingdao, Qingdao 266555, China. <sup>12</sup>School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.

\*These authors contributed equally to this work.

†Corresponding authors: Email: yizhi.cai@ed.ac.uk (Y.C.); yanghm@genomics.cn (H.Y.)

synthetic and wild-type sequence introduced a strong selective pressure for the two semisynthetic segments to recombine, via a homology-directed repair mechanism. Because our recombination strategy is designed to generate a markerless synII and leave the *URA3* marker on wild-type chromosomes, we observed a 10-fold increase in the *URA3*-deficient colony numbers upon I-SceI induction, as judged by marker loss screening, suggesting that our strategy can greatly bias the crossover events toward the designated regions. The resulting synthetic chromosome II strain (version yeast\_chr02\_9.01) was identified by the PCRTag method (figs. S2 and S3).

### Sequencing revealed structural variations in synII

Compared with the designed sequence, 61 variations belonging to four classes were identified by deep sequencing: 50 single-nucleotide variations (SNVs), 5 missing loxPsym sites, 4 deletions, and 2 structural variations (SVs) (table S2). Twenty-eight of the 50 SNVs were found to correspond exactly to the genotype of the native chromosome sequence at these positions, suggesting that these represent residual homologous recombination “patchworks” as seen in synIII and other synthetic chromosomes (12, 15, 18–21). When such patchwork regions are very short (i.e., lying entirely between two sets of PCRTags), they can be missed by the PCRTag analysis. For the remaining 22 SNVs, 3 were found to preexist in synthetic minichunk DNA, indicating that these mutations were introduced during synthesis. The remaining SNVs map to overlapping regions between minichunks or megachunks, suggesting that these SNVs were likely introduced during minichunk assembly or megachunk integration. Because none of these SNVs are in coding regions or noticeably altered phenotypes, they were not corrected.

As previously seen in other larger synthetic chromosomes (synV and synX) (19, 20), we observed two complex SVs: the first of these was a ~15-kb tandem duplication in megachunk L with chunk L2–L4 duplicated and a loxPsym site located between the duplications (fig. S5, A and B); another SV, identified in megachunk T, was fully characterized by PGM sequencing (Life Technologies) and is a >30-kb complex DNA sequence including multiple copies of chunks T4 and T5 and a partial-chunk backbone plasmid pSBGAK (fig. S5, A and C). We hypothesize that the 34-bp loxPsym sequence can serve as a homologous region during homologous recombination-mediated integration, albeit at a very low frequency, which led to the formation of the first SV. The mechanism of formation of the second SV remains unknown.

### Correction of structural variants in synII

We designed a straightforward strategy to repair the two structural variations, again by applying the I-SceI system (17). The 18-bp I-SceI recognition sequence was designed to carry a selectable marker (*URA3*) and overlap with both ends of the tandem repeat junction as a donor fragment (Fig. 1A). Upon induction of I-SceI digestion, the synII chromosome was broken and thereafter was repaired

through homology-directed recombination between the repeat sequences, and as a consequence it effectively looped out the duplicated region.

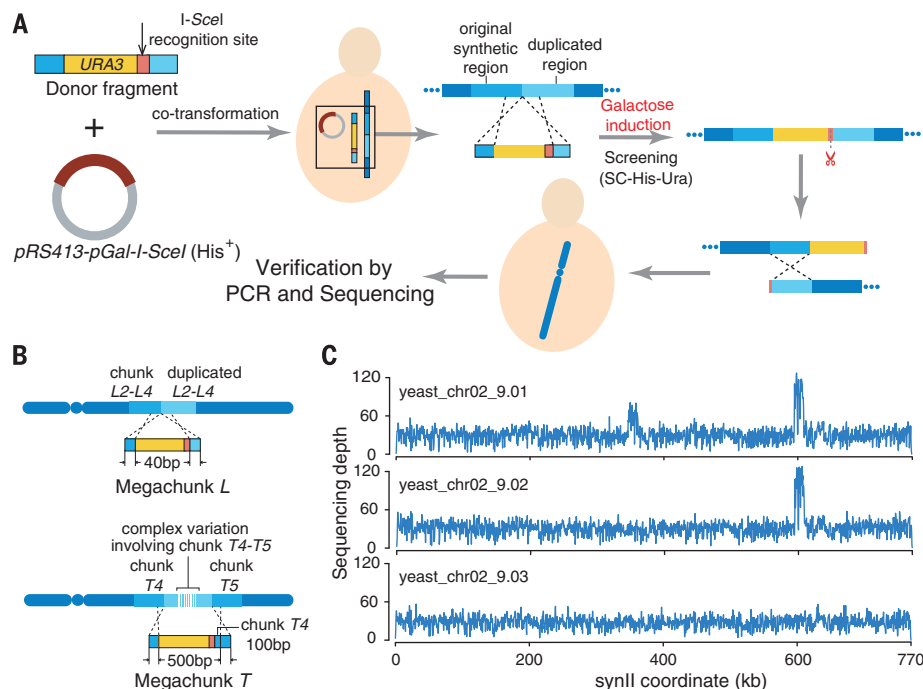
Using this strategy, we sequentially repaired the two SVs in synthetic chromosome II, yielding strains yeast\_chr02\_9.02 and yeast\_chr02\_9.03, respectively (Fig. 1B). Polymerase chain reaction (PCR) analysis (fig. S5D) and deep sequencing (Fig. 1C) validated the successful repair of the duplicated regions in synII. Duplications were also observed in other synthetic chromosomes (synV, synX, and synXII). Thus, the I-SceI-mediated strategy could serve as an efficient strategy, especially suitable for repair of large duplications in synthetic chromosome construction.

In addition to the SVs, pulsed field gel electrophoresis (PFGE) revealed an abnormal chromosome karyotype of native chrXIII and chrXVI (fig. S6). This was further confirmed by DNA sequencing and Hi-C analysis (22). Because similar instances also have been found during the construction of synIII (15), we surmise that this represents a spontaneous low-frequency event occurred at some

point during assembly. PFGE shows that the chrXIII and chrXVI karyotype of the synII strain with SVs (yeast\_chr02\_9.01) is correct. Therefore, we repaired this chromosomal crossover by first switching the mating type of the repaired synII (yeast\_chr02\_9.03) and then back crossing to synII strain (yeast\_chr02\_9.01) to avoid crossover between synthetic chromosome II and native chromosome II. After tetrad dissection, the correct karyotype was verified by PFGE analysis (fig. S6).

### Comprehensive characterization of synII (yeast\_chr02\_9.03) phenotype

Growth curves, phenotype, and morphology tests under various cultivation and stress conditions were performed to check the fitness of the repaired synII strain compared with the wild-type counterparts (BY4741 and BY4742) (Fig. 2, A and B). Results show that synII strains are largely on par with wild-type strains (figs. S7 to S9). We did observe a significant fitness defect after megachunk E integration. Further investigation revealed



**Fig. 1. Structure variation repair through chromosome breakage.** (A) I-SceI mediated repair strategy for synII structure variations. The donor fragment was designed to carry a *URA3* cassette (yellow) and I-SceI recognition site (red), with both ends overlapping the structural variation sequences observed in synII (yeast\_chr02\_9.01). The donor fragment was integrated into synII between the two tandem repeats through homologous recombination, then an episomal plasmid pRS413-pGal-I-SceI was transformed into the cell. A double-strand break at the I-SceI site was induced in galactose medium, and the homologous recombination of two partial chromosomes of synII eliminated the duplication. (B) Structure variations in megachunks L and T and their corresponding donor sequence design. In megachunk L, a copy of chunk L2–L4 was observed following the original L2–L4 sequence that generated a tandem duplication. The donor fragment was inserted directly between two duplications. In megachunk T, a complicated variation involving sequence of chunk T4–T5 was inserted into synII between chunk T4 and T5, and the donor fragment was inserted to remove the complex variation. (C) Deep sequencing read depth analysis revealed the successful sequential removal of duplications. The starting synthetic chromosome synII (yeast\_chr02\_9.01) has both duplication regions, and after the first round of repair at megachunk L, the resulting chromosome synII (yeast\_chr02\_9.02) has only one duplication. The finished chromosome synII (yeast\_chr02\_9.03) was obtained after the second repair at the megachunk T region.



that the defect was introduced by the insertion of the selectable marker *URA3* (for integration selection) adjacent to the *NCL1* gene, which encodes a tRNA:m5C-methyltransferase. Loss of function of the *NCL1* gene is previously reported to be involved in slow growth and temperature sensitivity (23, 24), which explains the observed phenotypic defect. This auxotrophic marker interference phenotype was reversed in the next step of SwAp-IN, in which the *URA3* marker was overwritten by an intact *NCL1* gene. Similar instances of such reversible fitness defects were also found after integration of megachunks B, U, and W (fig. S8).

In addition, a defect at high temperature (37°C) in medium with glycerol as the carbon source

was observed: Colony size is much smaller than wild type. We confirmed that the slow-growth phenotype is recessive by crossing *synII* with wild type (BY4742) (fig. S10A). To ascertain the origin of this defect, we examined the phenotype of each intermediate strain under this particular condition and found that the defect appeared after the integration of megachunk X (fig. S10B). Using complementation assays, we identified a modification resulting from PCRTag recoding in *YBR265W* (fig. S10C). As an essential gene, *YBR265W* encodes 3-ketosphinganine reductase (Tsc10p) that catalyzes the second step in the pathway for sphingolipid synthesis in *S. cerevisiae* (25). Sphingolipids are reportedly involved in the regulation of the yeast high-osmolarity glycerol (HOG) response

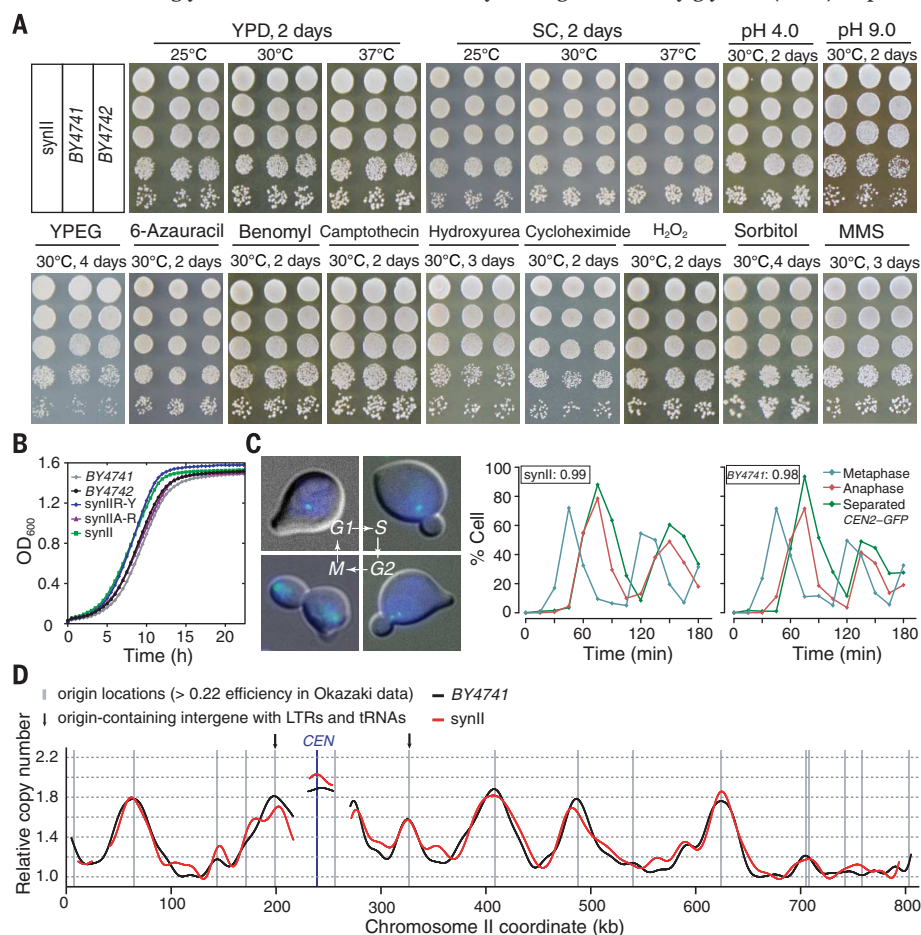
pathway (26). We reasoned that the defect might be caused by modifications in *YBR265W*. On replacement with the wild-type gene, the defect was greatly mitigated (fig. S10D). This highlights the effectiveness of the genome debugging mechanism provided by the modular chromosome construction strategy in the Sc2.0 project. Given that some defects might involve more than one bug, the aforementioned method could be inadequate. In this case, SCRaMbLE (synthetic chromosome rearrangement and modification by *loxP*-mediated evolution) may be a powerful alternative approach for targeting the defect. We have demonstrated the feasibility of this approach by quickly evolving the *synII* strain to identify and overcome a design problem. By turning on the SCRaMbLE system in *synII* cells and directly plating the SCRaMbLED population on the defect-causing condition, we identified a few large colonies that resemble wild type (fig. S11A). From two large colonies that we isolated for sequencing, no rearrangement was observed within the *YBR265W* region. However, all genomic rearrangements on *synII* involve genes coding for regulatory proteins that have either genetic or physical interactions with HOG regulatory proteins (fig. S11B). Interestingly, we observed that a phenotypic defect can be recovered through a number of alternative routes in the complex cellular interaction network.

As more and more individual synthetic chromosomes are completed and will be merged into a single haploid cell, we successfully consolidated *synII* and *synIII* using methods established previously (18). The resulting *synII/III* cell grows just like native cells (fig. S12), indicating that *synII* can be successfully incorporated into another synthetic background without introducing growth defects.

### Segregation and replication of *synII*

Once *synII* was constructed and phenotypically profiled, a logical next question is to ask how this synthetic chromosome replicates and segregates. Therefore, to further investigate whether the modifications to *synII* might introduce subtle chromosome damage not detected in large-scale phenotypic assays, we examined and compared DNA replication and segregation processes during the cell cycle in *synII* and WT (BY4741) strains by tagging *synII* with an array of *tet* operators and using the previously described TetR-GFP method (27). Here, *synII* and WT cells were released from a *G1* block synchronously into the cell cycle (fig. S13). The progression of cells from metaphase to anaphase was identified by spindle morphology. We scored separation of sister chromatids at *CEN2* by counting the percentage of cells in which two GFP dots were visible in both *synII* and BY4741 strains and found them to be comparable (Fig. 2C). Calculation of the overall ratio of metaphase to anaphase spindles and flow cytometry analysis of DNA content provided further evidence for normal cell cycle progression in the *synII* strain (Fig. 2C and fig. S14).

Early activating replication origins are frequently adjacent to tRNAs and transposable elements



**Fig. 2. Phenotypic profiling of *synII*.** (A) Phenotype tests of *synII* on different media. Ten-fold serial dilutions of overnight cultures of *synII* and wild-type (BY4741 and BY4742) strains were used for plating. From left to right: YPD at 25°C, 30°C, and 37°C; SC at 25°C, 30°C, and 37°C; low pH YPD (pH 4.0) and high pH YPD (pH 9.0); YPEG; SC+6-azauracil; YPD+benomyl; YPD+camptothecin; YPD+hydroxyurea; YPD+cycloheximide (10 µg/ml, 2 hours pretreatment); YPD+H<sub>2</sub>O<sub>2</sub> (1 mM, 2 hours pretreatment); YPD+sorbitol; YPD+MMS, (YPD, yeast extract peptone dextrose; YPEG, yeast extract peptone glycerol ethanol; MMS, methyl methane sulfone; SC, synthetic complete). (B) Growth curves of *synII*A-R, *synII*R-Y, and *synII* strains compared with those of BY4741 and BY4742 strains in YPD at 30°C. (C) Cell cycle comparison between *synII* and BY4741. Images show cell morphology at different stages during the cell cycle after release from *G1* block. DNA staining is shown in blue; CEN2-GFP in *G1*, *S*, *G2*, and *M* phase are shown in green. Graphs show the percentage of *synII* cells with separated CEN2-GFP dots, metaphase spindles, and anaphase spindles during the cell cycle. For each time point, at least 200 cells were counted. The inset numbers of 0.99 indicate the overall ratio of metaphase to anaphase cells throughout the time course for *synII* and BY4741 strains. (D) *synII* (red) and BY4741 (black) replication time expressed as relative copy number by deep sequencing.

containing long terminal repeats (LTRs) (28). In synII, two early origins are no longer linked with tRNAs and LTRs, because they were removed by design. Therefore, synII offers an opportunity to investigate whether this correlation represents a functional relationship. We determined the replication dynamics for synII and demonstrated that the synthetic chromosome showed identical replication dynamics to the wild-type chromosome II (Fig. 2D). This finding is consistent with the linkage between early activating origins, tRNAs and LTRs not being a functional requirement for early origin activation.

Our results show that synII exerts no gross negative effect on key cell cycle transitions, including chromosome replication (S phase) and segregation of sister chromatids (anaphase). In addition, Hi-C analysis of synII revealed no substantial changes between synthetic and native chromatin, suggesting that the designed sequence has little or no effect on the average global folding of the chromosome (22).

### SynII Trans-Omics analysis reveals genome plasticity

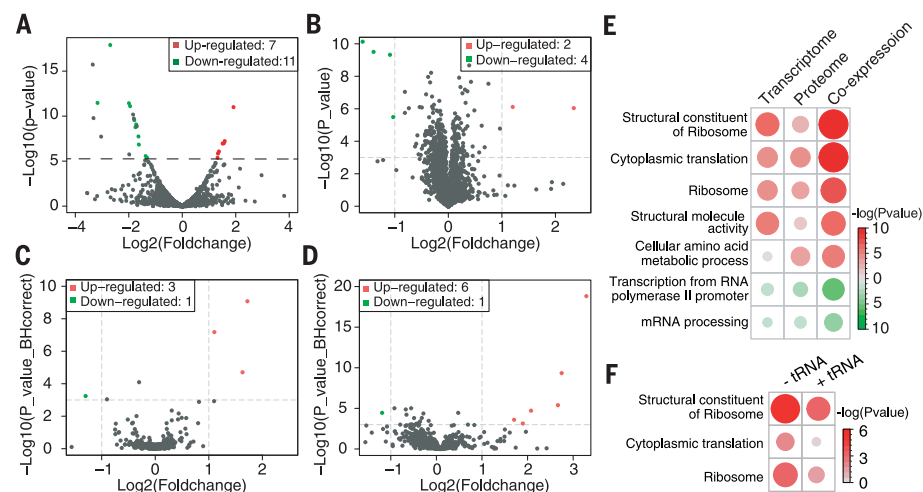
We used a systems biology approach (Trans-Omics analysis) to probe the effect of synII on the genomics, transcriptomics, proteomics, and metabolomics of *S. cerevisiae*. The insertion of 267 loxP sites in synII could affect genomic stability; therefore, we evaluated genome integrity and the loss frequency of chromosome segments in the absence of Cre expression. PCRTag analysis showed that no deletions were observed in 27 independent isolates derived from more than 130 mitotic generations from nine independent lineages (fig. S15, A and B). The overall loss rate was estimated at lower than  $5.9 \times 10^{-6}$ . In addition, the deep sequencing analysis of nine derived single strains showed that no mutation or genome rearrangement was observed, indicating that genome stability was maintained faithfully even after 100 generations of nonselective growth (fig. S15C). Transcriptome profiling identified only 18 out of 6561 genes as having differential mRNA expression in comparison to the wild type, with 7 up-regulated and 11 down-regulated ( $FDR < 0.01$  and  $P < 7.62 \times 10^{-6}$ ) (Fig. 3A and table S3). In the proteomics analysis, mass spectrometry was performed on synII and BY4741 strains. Mass spectrometry provided abundance data for 3965 out of 6682 protein-coding genes, and only six proteins showed substantial differential abundance in synII compared with wild-type (Fig. 3B and table S4). Finally, metabolic liquid chromatography-mass spectrometry (LC-MS) analysis was performed for metabolomics profiling. In total, 4941 and 6417 mass spectra peaks with  $CV < 30\%$  were identified in positive and negative mode, respectively, mapping to 1032 unique metabolites. Fewer than 0.78% of the metabolites were differentially represented (Fig. 3, C and D). Potential differentially regulated metabolites included sphingolipids, glycerophospholipids, steroids, and steroid derivatives (tables S5 and S6). Interestingly these are all molecules associated with membranes.

Moreover, we observed cellular perturbations that potentially arise from the Sc2.0 overall design principles. Compared with the WT strain, a subtle but potentially biologically significant up-regulation of genes with the Gene Ontology (GO) terms “ribosome” and “cytoplasmic translation” was observed in the synII strain on both the transcriptome and proteome level (Fig. 3E and fig. S16, A to E). Similar up-regulation was also observed in the synV and synX transcriptome (fig. S17). As previously reported, deletion of multicopy tRNA genes can lead to increased expression of translation machinery (29); thus, we reasoned that the up-regulation of translational function might be caused by the 13 tRNA genes deleted in synII, which are all from multicopy tRNA gene families. RNA sequencing (RNA-seq) analysis of synII carrying a synthetic tRNA array encoding the 13 deleted tRNA genes (yeast\_chr02\_3\_9.03, strain YCy1193) verified our hypothesis: The introduction of tRNA array to synII greatly diminished up-regulation of translational functions (Fig. 3F).

In summary, despite these very specific and subtle differences, the Trans-Omics analysis provides explicit evidence that biological processes within the synII strain are highly consistent with the wild-type strain. Therefore, the yeast genome displays a great degree of plasticity and can readily cope with the large degree of editing encoded into synII.

### Conclusions

In previous reports, the entire *Mycoplasma genitalium* genome was synthesized from oligonucleotides, assembled in budding yeast into a complete genome (9, 30). The genome transplantation method heavily depends on accurate design of a viable genome, and the low success rate of the transplantation step (and a lack of knowledge about how generically it can be done) makes it challenging to restore fitness when defects arise. In contrast, the Sc2.0 consortium uses a modular construction approach to progressively swap the wild-type chromosome with its designer counterpart, as illustrated for the first time (15), and many finished chromosomes (18–21). The modular approach provides an important mechanism to systematically dissect and repair phenotypic defects, exemplified by the elimination of a growth defect on a specific carbon source that we found to be caused by a 25-bp PCRTAG sequence in a 770-kb chromosome. Similar debugging success has been reported in other synthetic chromosomes (18, 19, 21). Another advantage of modular assembly is that it enables multiple teams to collaborate and construct synthetic chromosomes in parallel. The modular approach also permits parallel construction of individual chromosomes. In this study, we have constructed synthetic chromosome II from both ends. Other innovative approaches with the shared goal of parallelized construction are possible; synXII was constructed in six parallel



**Fig. 3. SynII strain Trans-Omics profile (BY4741 as reference) demonstrates that the synthetic chromosome design has minimal effect on cell physiology, despite a modest up-regulation of translational machinery triggered by tRNAs removal.** (A to D) Identified dysregulated genetic features at (A) transcriptome level, (B) proteome level, and (C and D) metabolome level (metabolic and lipid profiling in LC-MS positive mode, respectively) of synII cells, compared with BY4741 cells. The detailed methods of these comparisons are described in the materials and methods section. The total number of differentially expressed ( $P < 0.001$ ) features in transcriptome, proteome, and metabolome are also presented. Up-regulated and down-regulated features are labeled in red and green, respectively. (E) Enriched pathways and the coexpression profile revealed by transcriptome and proteome in yeast GO terms. Up-regulated features are labeled in red, and down-regulated features are labeled in green. (F) RNA-seq analysis of synII with/without tRNA array. By adding back the tRNA array of synII, the up-regulation of translational functions is greatly mitigated. For (E) and (F), the significance level is indicated by heat-map color intensities and symbol sizes. Up-regulated and down-regulated features are labeled in red and green, respectively.



production strains (a method termed meiotic recombination-mediated assembly), which can further promote the construction efficiency (27). Future efforts can be planned through the attempt to join the complementary advantage of our approach to this method to develop more efficient construction methods.

Our comprehensive phenotypic assays of synII strains revealed the extensive plasticity of the yeast genome. An important Sc2.0 design goal was to be conservative by avoiding changes likely to affect fitness-related phenotypes. Results with synII and other single synthetic chromosome strains suggest that the design has been successful in avoiding major fitness defects. However, interesting subtle patterns of differential gene expression were revealed by looking for enriched GO terms in the transcriptome and proteome analyses. It may be premature to suggest that the Sc2.0 design is too conservative, because strains hosting multiple synthetic chromosomes may have more profound phenotypic differences from the wild type due to the interactions among designer features encoded by different chromosomes. A Trans-Omics approach will be a powerful approach to capture subtle changes in the interaction network at various levels. Our design objective for the Sc2.0 genome is not to create a strain that encodes a particular phenotype; rather, our goal is to create a robust, high fitness, engineerable chassis for unbiased exploration of the viable genotype-to-phenotype space using SCRaMbLE induction under various conditions. This goal is in contrast to more standard synthetic biology goals to engineer a single, specific phenotype or metabolic capability. Previous pilot studies provide strong evidence that SCRaMbLE works as designed as an effective approach to generate an unbiased exploration of the enormous design space for synthetic chromosomes (11, 31).

In conclusion, we have demonstrated that synII segregates, replicates, and functions (at the Trans-Omics level) in a very similar way to its wild-type counterpart, which has naturally evolved over millions of years. We conclude that synII fits the design objective and is ready to be integrated into the final Sc2.0 genome with the joint effort of the entire Sc2.0 community in the near future.

## Materials and methods

### SynII design

We designed synII based on the version of S288C-derived strains available as of [30/11/2011] on SGD ([http://downloads.yeastgenome.org/sequence/S288C\\_reference/genome\\_releases/](http://downloads.yeastgenome.org/sequence/S288C_reference/genome_releases/), Version ID: S288C\_reference\_genome\_R64-1-1\_20110203). Genome editing suite BioStudio (14) was used to conduct the in silico design of synII, with the final version being yeast\_chr02\_3.25. Compared to native chromosome II, the resulting 770,035 bp synII sequence is altered on average once in every 467 bp, and has a total of ~10.26% sequence alteration. 22 of 30 introns from protein-coding genes were removed from synII, with the remaining 8 retained because they are either known to lead to a fitness defect when deleted or they are ribosomal proteins genes and deletion of such introns may lead

to fitness defects (32). More design details can be found in table S1 and on the synthetic yeast project website ([www.syntheticyeast.org](http://www.syntheticyeast.org)).

### SynII segmentation

Once synII design was completed, segmentation of synII into megachunks (~30 kb) and chunks (~10 kb) was performed using BioStudio (14). By applying an in-house developed program called SegMan, each chunk was segmented into minichunks (~3 kb). Minichunks were designed to be assembled into chunks by Gibson assembly (16) (see “Minichunk assembly” section), and therefore 40 bp overlaps between each adjacent minichunks were added by SegMan. Terminal restriction endonuclease sites were also added on both ends of each minichunk to allow excision from the plasmid with the requirement that only 5' sticky or blunt ends could be generated by digestion. The 40 bp overlap regions were chosen based on following criteria: minimal free energy > -3 kcal/mol, and melting temperature ( $T_m$ ) of 68 ± 4°C. Higher free energy of overlap sequence leads to reduced probability of self-folding of single-stranded DNA and results in higher efficiency for overlap-based in vitro assembly methods (33), such as Gibson assembly (16) and USER assembly (34). Here the RNAfold program of the ViennaRNA Package 2.0 (35) was used to calculate the minimal free energy of the DNA sequence. A simplified formula (36) was applied to estimate the melting temperature:  $T_m = 945 \cdot \Delta H / [\Delta S + R \cdot \log(0.0001)] - 273.15$ , in which  $\Delta H$  is enthalpy (kJ/mol),  $\Delta S$  is entropy,  $R$  is the molar gas constant (1.9872 cal/mol-K). In addition, to reduce synthesis cost, *LEU2* and *URA3* markers were designed separately as independent cassettes and used repeatedly in minichunk assembly.

### SynII synthesis and assembly Minichunk assembly

Synthesis of ~3 kb minichunks was outsourced to Invitrogen, Genscript, and BGI Tech. pSBGAA or pSBGAK (sequence information can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)) was chosen as the accepting vector of chunks. Minichunks were excised from the plasmids using the terminal restriction sites. *Bam*HI was used to linearize the chunk-accepting vectors pSBGAA (Amp<sup>R</sup>) and pSBGAK (Kan<sup>R</sup>). The chunks were assembled using Gibson assembly strategy (16) with a modified recipe: 1 µL of Taq ligase (New England Biolabs) was added in the final volume of 20 µL for each reaction. The molar ratio of accepting vector to minichunk DNA was 1:5. After thorough mixing, a one-hour incubation at 50°C was performed. Then 10 µL of the reaction mixture was transformed to 50 µL *Escherichia coli* DH5α competent cells (TAKARA). For assembly verification, single colonies were selected for overnight culture at 37°C. After miniprep, restriction digestion of the terminal restriction sites was performed to verify the assembly result.

### Chunk and junction preparation

Once verified by restriction digestion, the component chunks of each megachunk were digested

using the terminal restriction sites and gel purified. For each megachunk, 1 µL of each chunk DNA, 0.5 µL of T4 DNA ligase (New England Biolabs) and ddH<sub>2</sub>O were added to a final volume of 10 µL and mixed well for an overnight ligation at 16°C. Then the reaction product was diluted 1:10 and used as template to amplify ~1 kb junction fragments with the corresponding primers (primer sequence information can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)) using Phusion DNA polymerase (New England Biolabs). After gel verification, junction fragments were purified using a PCR cleanup kit (Axygen).

### Replacement of WT yeast chromosome II with synthetic chunks

BY4741 (*MATa his3Δ1 leu2Δ0 LYS2 met15Δ0 ura3Δ0*) containing the KanMX marker (strain ID: YS000-L) and BY4742 (*MATα*) containing the *URA3* marker (strain ID: YS000-R) were used as the initial strains for synthetic megachunk replacement from the left arm and right arm, respectively. The *URA3* and *LEU2* selectable markers were used iteratively for replacing the native sequences of chromosome II with synthetic chunks. Together with the ~1 kb junction fragments (consisting of ~500 bp overlaps with two adjacent chunks), the chunks constituting each megachunk were co-transformed using the LiOAc transformation protocol (15) with 300 ng of each chunk DNA and 200 ng of each junction DNA added. The transformation products were resuspended in 100 µL 5 mM CaCl<sub>2</sub> and plated on an appropriate selectable media (SC-Ura or SC-Leu) with serial dilutions where appropriate. After 18 successive rounds (left arm initiated) and 8 successive rounds (right arm initiated) of replacements, two semisynthetic chromosomes were successfully constructed, with one containing synthetic megachunks A to R (synIIA-R, strain ID: YS018) and the other containing synthetic megachunks R to Y (synIIR-Y, strain ID: YS026). The selectable marker *URA3* in synIIR-Y was removed through yeast transformation with the corresponding markerless synthetic fragment and screened with 5-FOA (37).

### Integration of parallel constructed semisynthetic chromosomes

An I-SceI (38)-mediated method was developed to combine the two semisynthetic chromosomes, synIIA-R and synIIR-Y. For synIIA-R, a fragment (I-SceI-*URA3*) containing the I-SceI recognition sequence and *URA3* was designed to have a 40 bp overlap upstream and downstream of the *LEU2* marker on synIIA-R. A previously described method (39) was used to integrate the I-SceI-*URA3* cassette into synIIA-R to replace *LEU2* marker (strain ID: YS027). Following the same strategy, the I-SceI-*URA3* cassette was inserted into synIIR-Y with the *URA3* marker residing in wild-type sequence at the upstream end of megachunk R. In addition, a *LEU2* marker was inserted into synIIR-Y at the upstream of the centromeric region of synIIR-Y within gene *YBL005W* (strain ID: YS028). Mating of the synIIA-R strain (*MATa*, strain ID: YS027) and synIIR-Y (*MATα*, strain ID: YS028) was performed by overnight co-culturing

in 3 mL YPD medium at 30°C. The I-SceI expression vector pRS413-pGAL-I-SceI was transformed into the diploid cells followed by plating on SC-His plate. Single colonies were selected for overnight culture in 3 mL SC-His/glucose medium. The overnight culture was then added to 20 mL SC-His/raffinose (2% raffinose and 0.1% glucose) to reach an initial OD<sub>600</sub> of 0.1 and incubated until OD<sub>600</sub> reached 0.4 (~4-hour). Cells were harvested by centrifugation at 6,200 rcf for 5 min and resuspended in 20 mL SC-His/galactose (2% galactose). After a 2-hour I-SceI induction on galactose media, 20 µL cells were inoculated into 3 mL YPD medium, followed by overnight culture at 30°C. Overnight culture (1 mL) was harvested by centrifugation at 13,800 rcf for 1 min. The pellets were resuspended in 200 µL ddH<sub>2</sub>O, patched on a SPOR plate and incubated at room temperature for 1 day. Then 5-7 days incubation at 30°C was performed until a significant number of tetrads was observed.

Cells were resuspended in 25 µL Zymolyase-20T (25 mg/mL in 1 M sorbitol), then incubated at 37°C, with shaking at 800 rpm for 60 min. Then 500 µL ddH<sub>2</sub>O was added and mixed thoroughly, followed by plating on 5-FOA plates and incubating at 30°C for 3 days. Then replica plating on SC-Leu and YPD plates was performed, followed by 24 hours incubation at 30°C. 5FOA<sup>R</sup> Leu<sup>-</sup> colonies were selected from YPD plates, inoculated into 3 mL YPD medium, and cultured overnight. To verify the integration of semisynthetic chromosomes, one pair of synthetic PCRTag and wild-type PCR tag were chosen from each megachunk (in total 25 pairs) to perform PCRTag analysis. Then sequencing was performed for further verification.

### Yeast genomic DNA preparation for PCRTag analysis

Yeast cells were collected from 0.5 mL of overnight culture by centrifugation at 13,800 rcf for 1 min. Pellets were resuspended in 100 µL breaking buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.1 M NaCl, 2% Triton x-100, 1% SDS). Glass beads (0.1 g, 0.5 mm of diameter) and 200 µL PCI (Phenol: Chloroform: Isoamyl alcohol= 25:24:1) were added and the mixture was vortexed at maximum speed for 3 min. Then 100 µL ddH<sub>2</sub>O was added and the mixture was centrifuged at 13,800 rcf for 5 min. Supernatant (150 µL) was transferred to a microcentrifuge tube.

### PCRTags analysis reaction setting

PCRTag amplification was performed using rTaq Polymerase (TAKARA). Forward and reverse PCRTag primer pairs (400 nM each, detailed information of PCRTag primers can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)), 1 µL BY4741/synII genomic DNA and 2.25 µL ddH<sub>2</sub>O were added to a final reaction volume of 12.5 µL. The PCR thermal-cycler program setting was as follows: 94°C/5 min, 30 cycles of (94°C/30 s, 55°C/30 s, 72°C/30 s), and a final extension of 72°C/5 min. Detection of PCRTags was carried out by gel electrophoresis. PCR product (3 µL) was loaded onto a 2% agarose gel, and electrophoresis was performed at 180 V for 20 min.

### Yeast genomic DNA preparation for DNA sequencing

Yeast cells were grown in 5 mL YPD medium using a 14 mL round bottom tube for 2 days till saturation. Pellets were collected by centrifugation at 13,800 rcf for 1 min. Breaking buffer (400 µL) was added to resuspend the pellet. Glass beads (0.2 g, 0.5 mm of diameter) and 400 µL PCI (Phenol:Chloroform:Isoamyl alcohol=25:24:1) were added, the resuspension was vortexed at maximum speed for 3 min, and then centrifuged at 13,800 rcf for 10 min. 400 µL of the aqueous layer was transferred to a new 1.5 mL tube. The genomic DNA was precipitated by adding 400 µL of isopropyl alcohol and kept at room temperature for 5 min. Then genomic DNA was pelleted by centrifugation at 13,800 rcf for 10 min. The pellet was washed with 500 µL 70% ethanol, followed by 5 min drying at 37°C. The genomic DNA was resuspended in 50 µL TE buffer (10 mM Tris-HCl pH8.0, 1 mM EDTA) with RNase (25 µg/mL) and incubated at 37°C for 30 min.

### Nucleotide sequence analysis of synII with Hiseq2500 sequencing platform Library preparation and whole genome sequencing

Paired-end whole genome sequencing was performed for the synII (yeast\_chr02\_9.01, strain ID: YS029) on the HiSeq2500 platform. A 500-bp library was prepared according to standard Illumina DNA preparation protocols.

### Sequencing quality control and mapping to genome sequence

Before mapping of reads, quality control of sequencing reads was performed. Reads with adapters or shorter than 90 bp were removed. Reads with more than one base having a Phred-score below 10 or with more than one unknown base were removed, leaving 673 Mbp cleaned paired-end reads, with 55.6-fold sequencing depth of the genome. Cleaned reads of each sample were mapped to yeast reference sequences (the original sequence of chromosome II being replaced by synthetic chromosome II sequence) using BWA 0.5.6 (40) with standard settings. For each alignment result, local realignment was performed with GATK 2.7-2 RealignerTargetCreator and IndelRealigner tools (41) to clean up mapping artifacts caused during reads mapping on the edges of indels. The resulting files in BAM format were then prepared for initial SNV/indel calling.

### Identification of SNVs and indels

Both Samtools (42) and GATK 2.7 (41) pipelines were used to identify the SNVs and indels using default parameters. The variants were filtered by the criteria of QUAL < 50 or PV4 < 0.1, 0.1, 0.1, 0.1 or MQ < 10 or DP < 10 or DP4 < 0, 0, 3, 3 for Samtools results and QUAL < 50 or FS > 3 or BaseQRankSum > 3 or MQRankSum > 3 or ReadPosRankSum > 3 or MQ < 10 or DP < 10 for GATK results. The variants identified by either tool were merged with CombineVariants implemented in GATK. The merged variants in the synthetic chromosome were checked manually using Tablet (43) to

exclude false-positive results caused by sequencing or mapping errors. Annotation was performed for observed variants, with the following types: synonymous type, nonsynonymous type, frame-shifts, and variant outside coding region (table S2).

### Identification of loxPsym sites

In synIII, loxPsym sites can be absent from expected locations (15). To check whether all loxPsym sites are present in synII, sequencing reads containing loxPsym sites were extracted from the whole read library for identifying the presence of all expected loxPsym sites. For each loxPsym site, the read mapped span the loxPsym site of upstream and downstream flanking sequence >10 nucleotides and the sequenced loxPsym site with bases of >=24 matching or mismatching loxPsym sequence was recognized as a loxPsym read and bases of >=15 deleted was counted as a "loxPsym\_lost" read (table S2). The loxPsym site with  $P < 0.001$ , which is estimated by Poisson mode, and supporting read number  $\geq 5$  was identified as a loss of loxPsym site.

### Structural variant detection with the PGM sequencing platform Library preparation and whole genome sequencing

A 400-bp DNA library of a synII (yeast\_chr02\_9.01, strain ID: YS029) was prepared for single-end whole genome sequencing according to the Life Tech standard preparation protocol using the Ion Xpress Barcode Adapter 1-96 Kit (Cat. no. 4474517) and sequenced on the Ion PGM platform. Quality control of sequencing reads was performed. Reads shorter than 30 bp or duplicated were removed. Reads with more than 1% of bases having a Phred-based quality score <10 or with unknown bases were trimmed first to meet the filtering criteria and removed if the first trimming step failed, leaving 488 Mbp cleaned reads, with 32.6-fold sequencing depth of the genome and 248 bp average length. Cleaned reads of each sample were mapped to synthetic yeast genome sequences using bowtie2-2.0.0 (44) with standard settings.

### Structural variant reconstruction

To identify whether structural variations exist in synII, reads that did not map to the reference genome were split to pairwise ends (split reads, at least 30 bp) by scanning over all intermediate positions at least 30 bp from the ends of the read. Then these pairwise ends were then aligned to the reference using Bowtie2 (44) by single-end mapping with parameter -k 100. The pairwise ends that matched parental sequence were analyzed for breakpoints to provide direct evidence for structural variants. One read was assigned as the most probable mapping type based on five priorities: no recombination events (top priority), intra-chromosome recombination, inter-chromosome recombination within wild-type chromosomes and external chromosome recombination between synthetic and wild-type chromosomes, and single end mapping. For each identified breakpoint, a 5 bp error range was allowed. These reads at a breakpoint site with at least 3 supporting reads

were locally assembled using PolyPhred (45), and then the assembly results were annotated by Blastn against the reference containing the selection marker and acceptor vector and wild-type chromosome II sequence to identify the potential extrinsic sequence. Combining the split-read mapping, sequencing depth and local assembly contigs, the structural variations were reconstructed (fig. S5).

### SynII structural variant correction

An I-SceI mediated transformation strategy was designed to correct the two structural variations observed in megachunks L and T of synII (yeast\_chr02\_9.01, strain ID: YS029). For the duplication in megachunk L, a donor fragment containing the *URA3* sequence and an I-SceI site (TAGGGATAACAGGGTAAT) was designed to carry 40 bp overlapping regions at the end of the original chunk L4 and the beginning of duplicated chunk L2 (Fig. 1B). The donor fragment was produced by PCR using primers including the I-SceI site and overlap sequence. I-SceI expression vector pRS413-pGAL-I-SceI and the donor fragment were co-transformed into the synII strain (strain ID: YS029). After I-SceI induction on galactose media, cells were plated directly on 5-FOA plates directly and incubated at 30°C for 3 days. Then 5FOA<sup>R</sup> colonies were selected and cultured overnight in YPD medium at 30°C. Confirmation of repair was performed by PCR using a pair of primers designed to amplify across the L4–L2 breakpoint (primer sequence information can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)).

For the complex variation in megachunk T, a donor sequence was designed to contain a *URA3* cassette and the I-SceI site, with the left end overlapping with the right end of chunk T4 by 500 bp, and the right end overlapping with the left end of chunk T5 by 500 bp (Fig. 1B). Both 500 bp overlapping sequences and the I-SceI-*URA3* cassette were produced by PCR with 40 bp overlaps with each other and assembled together through the Gibson assembly method (16) to form the donor fragment. An additional 100 bp of chunk T4 was added between the I-SceI-*URA3* cassette and 500 bp overlapping sequence of T5 to facilitate the homologous recombination. The donor fragment and I-SceI expression vector pRS413-pGAL-I-SceI were then co-transformed into the synII (yeast\_chr02\_9.02, strain ID: YS030) in which the structural variation in megachunk L had been repaired. I-SceI induction was performed on galactose-containing media and colonies were screened on 5-FOA-containing plates. PCR verification was then performed with 8 pairs of verification primers designed across each breakpoint observed in this structural variation (fig. S5D). The synII (yeast\_chr02\_9.03, strain ID: YS031) was sequence-verified again using the PGM sequencing platform, with library preparation and sequencing method as described in “Structural variant detection with the PGM sequencing platform.”

### Mating type switch of yeast

pJDI38 (pGAL-HO) was transferred into the synII strain (yeast\_chr02\_9.03, strain ID: YS031) using

the LiOAc transformation protocol (15), followed by overnight culture in 5 ml SC-Ura/galactose (2% galactose) liquid at 30°C. The mating type was then determined by crossing with *MATa* and *MATa* tester strains and confirmed by microscopy.

### Mating and sporulation

A synII (yeast\_chr02\_9.01, strain ID: YS029) and a second synII (yeast\_chr02\_9.03) strain that had its mating type switched to alpha (strain ID: YS032) were recovered on YPD plates. SynII (yeast\_chr02\_9.01) colonies and synII (yeast\_chr02\_9.03) colonies were picked and patched in the same section (mating section) on a new YPD plate with an inoculating loop, followed by overnight culture at 30°C. Then diploid colonies were patched on SPOR plates [10 g potassium acetate (Sigma), 1.25 g yeast extract (Oxoid), 1 g glucose (Sangon Biotech), 20 g agar (Sangon Biotech), ddH<sub>2</sub>O up to 1 L] and left at room temperature for one day, followed by 8 days incubation at 30°C. Then cells were taken from SPOR plates and resuspended in 25 µL zymolyase solution (MP Biomedicals, 1 mg/ml Zymolyase in 1M sorbitol), followed by 25 min incubation at 37°C. Cells were checked under the microscope to determine zymolyase efficiency. After tetrads were appropriately digested, 500 µL ddH<sub>2</sub>O was added and the suspension mixed vigorously by vortexing to break up the tetrads. Then a series of 10-fold dilutions were made before plating on YPD plates. The plates were incubated at 30°C for 2–3 days. Single colonies were picked and streaked on fresh plates, followed by overnight incubation. Mating type of random spores was confirmed by crossing with *MATa* and *MATa* tester strains and confirmed by microscopy. Spore-derived colonies were selected for PCR verification. The nine pairs of primers used for identifying breakpoints in synII (yeast\_chr02\_9.01) were used here to verify the correct synII sequence (primer sequence information can be found at [www.syntheticyeast.org](http://www.syntheticyeast.org)). Pulsed-field gel electrophoresis (PFGE) was performed for karyotype analysis. Finally, candidates were further verified using the PGM sequencing platform, with library preparation and sequencing method as described in “Structural variant detection with the PGM sequencing platform” above.

### Pulsed-field gel electrophoresis

Samples were prepared for pulsed-field gel electrophoresis (46). Identification of chromosomes was inferred from the karyotype of WT (BY4741, BY4742) on the same gel. Samples were analyzed on a 1.0% agarose gel in 1×TAE (pH 8.0) for 22 hours at 14°C on a CHEF apparatus. The voltage was set to 5 V/cm, at an angle of 120°. Switch time was set to 60 – 90 s ramped over 22 hours.

### Genome stability analysis

The synII (yeast\_chr02\_9.03, strain ID: YS031) was streaked on YPD plate for 2 days and incubated at 30°C. Nine single colonies were selected for successive subculture in YPD medium for ~130 generations, followed by plating on YPD plates overnight. Three single colonies of each initial isolate were selected. Genomic DNA preparation

was performed for all 27 isolated single colonies. 48 pairs of PCRTags were chosen to perform PCRTag assay for all 27 isolated single colonies. Sequencing on the PGM platform was conducted for 9 of the 27 isolated single colonies (one from each initial isolate), with library preparation and sequencing method as described in “Structural variant detection with PGM sequencing platform” above. Sequence data were analyzed according to the method described in “Nucleotide sequence analysis of synII with Hiseq2500 sequencing platform” above.

### Cre SCRaMBLe of synII

BY4741 and synII (yeast\_chr02\_9.03, strain ID: YS033) with pSCW11-CreEBD-His3 plasmid (47) were cultured in SC-His liquid media overnight at 30°C, with duplicates for each strain. Cells were inoculated into SC-His liquid medium with 1 µM EST and cultured for 24 hours at 30°C. Then cells were washed, serial dilution and spotting were performed on YPG plates, followed by 4 days incubation at 37°C.

### Yeast total RNA isolation for RNA sequencing

The synII (yeast\_chr02\_9.03, strain ID: YS031) and BY4741 strain both with 3 biological replicates were cultured overnight in 3 mL YPD medium at 30°C. The cultures were added to 10 mL fresh YPD medium and incubated until the OD<sub>600</sub> reached ~0.8. The cells were harvested by centrifugation at 900 rcf for 5 min. Total RNA was isolated using the RiboPure-yeast kit (Ambion) according to the manufacturer's instructions.

### RNA-seq analysis of synII

The 330 bp cDNA libraries were prepared according to the manufacturer's instructions (Illumina Inc.) and were paired-end sequenced using the Illumina Hiseq2500. Raw read data were filtered using the following criteria: no N bases, no adaptor sequences, minimum read length 100, bases of low quality (<10) was no more than 1% in a read. An average of 7.0 M and 8.6 M clean reads were obtained from BY4741 and synII (yeast\_chr02\_9.03, strain ID: YS031) samples respectively, with 95.7% and 95.2% mapped to the corresponding genome reference. Cleaned reads were mapped to genomes by TopHat v2.0.10 (48), with the parameter -r (-mate-inner-dist) of 130. After reads counting, RPKM was calculated for each gene. Differential gene expression was analyzed by DEseq v1.20.0 (49), with the no replicates scenario (parameters were: -method blind, -sharingMode fit-only, -fitype, local). For each gene, we obtained a raw p-value and adjusted p-value using the Benjamini-Hochberg procedure. Genes were assessed for statistical significance by rejecting the null hypothesis if the adjusted P value (FDR) was <0.01 and if the raw P value fell below the threshold of the 5% Family Wise Error Rate (FWER) after Bonferroni correction (threshold = 7.62 × 10<sup>-6</sup>). False positive results were inferred by the following rules and removed: dubious genes, transposable genes, genes with low coverage (<60%) for native and synthetic strains.



## Proteome and metabolome analysis of synII

BY4741 and synII (yeast\_chr02\_9.03, strain ID: YS031) strains with 3 biological replicates were cultured overnight in 3 mL YPD medium at 30°C and re-inoculated in 10 mL fresh YPD medium for further incubation until the OD<sub>600</sub> reached ~0.8. The cells were collected for both proteome and metabolome analysis. Proteins from the yeast cell were extracted with Urea, reduced, alkylated, digested with trypsin and iTRAQ labeled BY4741 (113, 115, 117, 121) and synII (114, 116, 118 and 119) (AB SCIEX, Framingham, MA, USA). After labeling, the peptides were fractionation with SCX method and analysis by an Orbitrap Q Exactive mass spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled with an online HPLC. Mascot and IQuant (50) software were used for protein identification and quantification.

For the global metabolomics, metabolites were extracted with buffer (50% methanol and 50% water), separated with a BEH C18 column. For the lipidomics, metabolites were extracted with buffer (75% dichloromethane and 25% methanol), separated with a CSH C18 column. The metabolites were detected by a XEVO-G2XS QTOF mass spectrometer (Waters, Manchester, UK) and the raw spectrums were processed by Progenesis Q1 2.0 software (Nonlinear Dynamics, Newcastle, UK) for peak picking, alignment, normalization and identification. Further statistical analysis was performed on the resulting normalized peak intensities using in-house developed software metaX. Pathway analysis was conducted using the MetaboAnalyst pathway tool (51).

## Yeast KEGG pathway and GO enrichment

To identify differentially active biological processes between synII (yeast\_chr02\_9.03, strain ID: YS031) and BY4741 strains, gene enrichment and co-expression enrichment analyses were performed using yeast KEGG pathways and yeast GO annotations integrating transcriptomics, proteomics and metabolomics data respectively. Genes and metabolites with differential expression of log<sub>2</sub> (fold-change) > 0 and down-regulation of log<sub>2</sub> (fold-change) < 0 were considered as up-regulated and down-regulated for enrichment analysis. The significance of each KEGG pathway and GO term in genes, metabolites and co-expression was individually identified using the hyper-geometric test and Chi-squared test with false discovery rate (FDR) correction (52) and the threshold *P* value < 0.001. The KEGG category and interaction with genes of metabolites were classified based on the Yeast Metabolome Database (53) ([www.ymdb.ca/system/downloads/current/ymdb.json.zip](http://www.ymdb.ca/system/downloads/current/ymdb.json.zip)).

## Growth curve assay

The growth curve analysis of BY4741, BY4742, and synII (yeast\_chr02\_9.03, strain ID: YS031) was carried out using a Bioscreen C system (Oy Growth Curves Ab Ltd). Overnight cultures with 3 replicates were sub-cultured into 300 µL of the different media to final OD<sub>600</sub> of ~0.1 in the 96 well microplates. Medium was also added as blank.

## Serial dilution assay on various types of media

BY4741, BY4742, all intermediate strains (strain ID: YS001-YS026), and synII (yeast\_chr02\_9.03, strain ID: YS031) were inoculated in YPD medium at 30°C overnight. Then 1:10 serial dilutions were performed using these samples onto different medium plates, which including: YPD, YPD with MMS (testing for defective DNA repair), YPD with benomyl (microtubule inhibitor), YPD with camptothecin (topoisomerase inhibitor), YPD with hydroxyurea (testing for defective DNA replication), YPD with various final concentrations of sorbitol (osmotic stress 0.5 M, 1 M, 1.5 M, 2 M, to cause osmotic stress), synthetic complete (SC) medium, SC with 6-azauracil (testing for defective transcription elongation), YPEG (respiratory defects) containing 2% glycerol and 2% ethanol (testing for respiratory defects), and YPD adjusted to pH 4.0 and pH 9.0 with HCl and NaOH (testing for vacuole formation defects). Two specific drugs, hydrogen peroxide (testing for oxidative stress) and cycloheximide (testing for defective protein synthesis), were added to overnight cultures to treat the cells for two hours. Yeast cells were collected by centrifugation and resuspended in water before serial dilution and plating. For different medium types, plates were incubated at 25°C /30°C /37°C for 2/3/4 days (fig. S8).

## Cell morphology

Cells were grown to log phase in YPD at 30°C. DIC Images were collected using a Nikon microscope Ti-E (100X) with an Andor Zyla 5.5 camera (fig. S9).

## CEN2-GFP strain construction and culturing

To create the *CEN2-GFP* strain, a 950 bp region to the right of *CEN2* was cloned into Amp327 (27) and integrated into the synII strain (yeast\_chr02\_9.03, strain ID: YS031) and BY4741 carrying *tetR-GFP*, generating a GFP label 15 kb to the right of *CEN2* (Strain ID: YCy922, YCy1018). Cell cycle synchronization of *CEN2-GFP* strains was performed by adding alpha factor in YEP medium. Alpha factor was removed by filtration and washing with YEP medium. Three hour mitotic time courses were performed at room temperature unless otherwise stated.

## Microscopy

Methods of fixing cells for GFP-labeled chromosome visualization and indirect immunofluorescence described previously were used in this study (27). Tubulin was visualized using an anti-rat antibody at a dilution of 1:50 in PBS/BSA and an anti-rat FITC antibody at a dilution of 1:16.67 in PBS/BSA. Microscopy was performed on a Zeiss Axioplan 2 microscope and images were captured using a Hamamatsu camera operated through Axiovision software. To score GFP dots and spindles in mitosis, cells (Strain ID: YCy922, YCy1018) were co-stained with DAPI to visualize nuclear morphology (fig. S13). For each sample obtained at each 15 min time point, 200 cells were counted in the field to score GFP dots and spindles.

## FACS analysis

Methods of flow cytometry described previously were used in this study (27). Cells were fixed in 70% ethanol at 4°C overnight. Cells (Strain ID: YCy922, YCy1018) were then treated with RNase (20 mg/ml, Sigma), and digested with proteinase K (20 mg/ml, Amresco). Samples were briefly sonicated before analysis. FACS analysis was performed according to the manufacturer's instructions (BD FACS Calibur). Data were analyzed using FlowJo software (fig. S14).

## SynII replication timing profiles analysis

Replication profiles were obtained using methods described previously (54). Asynchronous cells were fixed with 70% ethanol, treated with RNaseA and Proteinase K and stained with 10× Sytox. The MoFlo XDP was used to enrich for replicating and nonreplicating cells. Sorted cells were treated with Zymolase, RNaseA and Proteinase K. DNA was purified using phenol-chloroform extraction and ethanol precipitation. Single end Illumina HiSeq2500 sequencing yielded a minimum of 29.6 million uniquely mapped reads per sample. The ratio between uniquely mapped reads from the nonreplicating and the replicating sample were calculated for every 1000 bp window. The ratio was normalized to a baseline of one to control for differences in number of reads between the samples.

## REFERENCES AND NOTES

- O. Ozier-Kalogeropoulos, A. Malpertuy, J. Boyer, F. Tekia, B. Dujon, Random exploration of the *Kluyveromyces fragilis* genome and comparison with that of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **26**, 5511–5524 (1998). doi: 10.1093/nar/26.23.5511; pmid: 9826779
- R. B. Langkjaer, M. L. Nielsen, P. R. Daugaard, W. Liu, J. Piskur, Yeast chromosomes have been significantly reshaped during their evolutionary history. *J. Mol. Biol.* **304**, 271–288 (2000). doi: 10.1006/jmbi.2000.4209; pmid: 11090273
- V. Wood et al., The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**, 871–880 (2002). doi: 10.1038/nature724; pmid: 11859360
- S. Brachat et al., Reinvestigation of the *Saccharomyces cerevisiae* genome annotation by comparison to the genome of a related fungus: *Aschysphaeria gossypii*. *Genome Biol.* **4**, R45 (2003). doi: 10.1186/gb-2003-4-7-r45; pmid: 12844361
- M. Kellis, B. W. Birren, E. S. Lander, Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617–624 (2004). doi: 10.1038/nature02424; pmid: 15004568
- H. G. Menzella et al., Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005). doi: 10.1038/nbt1128; pmid: 16116420
- D. K. Ro et al., Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006). doi: 10.1038/nature04640; pmid: 16612385
- G. Pósfai et al., Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044–1046 (2006). doi: 10.1126/science.1126439; pmid: 16645050
- D. G. Gibson et al., Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008). doi: 10.1126/science.1151721; pmid: 18218864
- C. Lartigue et al., Genome transplantation in bacteria: Changing one species to another. *Science* **317**, 632–638 (2007). doi: 10.1126/science.1144622; pmid: 17600181
- J. Dymond, J. Boeke, The *Saccharomyces cerevisiae* SCRaBLE system and genome minimization. *Bioeng. Bugs* **3**, 168–171 (2012). pmid: 22572789
- J. S. Dymond et al., Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**, 471–476 (2011). doi: 10.1038/nature10403; pmid: 21918511

13. H. Feldmann *et al.*, Complete DNA sequence of yeast chromosome II. *EMBO J.* **13**, 5795–5809 (1994). PMID: 7813418
14. S. M. Richardson *et al.*, Design of a synthetic yeast genome. *Science* **355**, 1040–1044 (2017).
15. N. Annaluru *et al.*, Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55–58 (2014). doi: [10.1126/science.1249252](https://doi.org/10.1126/science.1249252); PMID: 24674868
16. D. G. Gibson *et al.*, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009). doi: [10.1038/nmeth.1318](https://doi.org/10.1038/nmeth.1318); PMID: 19363495
17. J. Parenteau *et al.*, Deletion of many yeast introns reveals a minority of genes that require splicing for function. *Mol. Biol. Cell* **19**, 1932–1941 (2008). doi: [10.1091/mbc.E07-12-1254](https://doi.org/10.1091/mbc.E07-12-1254); PMID: 18287520
18. L. A. Mitchell *et al.*, Synthesis, debugging, and effects of synthetic chromosome consolidation: synII and beyond. *Science* **355**, eaaf4831 (2017).
19. Y. Wu *et al.*, Bug mapping and fitness testing of chemically synthesized chromosome X. *Science* **355**, eaaf4706 (2017).
20. Z.-X. Xie *et al.*, “Perfect” designer chromosome V and behavior of a ring derivative. *Science* **355**, eaaf4704 (2017).
21. W. Zhang *et al.*, Engineering the ribosomal DNA in a megabase synthetic chromosome. *Science* **355**, eaaf3981 (2017).
22. G. Mercy *et al.*, 3D organization of synthetic and scrambled chromosomes. *Science* **355**, eaaf4597 (2017).
23. A. Alexandrov *et al.*, Rapid tRNA decay can result from lack of nonessential modifications. *Mol. Cell* **21**, 87–96 (2006). doi: [10.1016/j.molcel.2005.10.036](https://doi.org/10.1016/j.molcel.2005.10.036); PMID: 16387656
24. J. M. Dewe, J. M. Whipple, I. Chernyakov, L. N. Jaramillo, E. M. Phizicky, The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. *RNA* **18**, 1886–1896 (2012). doi: [10.1261/ma.033654.112](https://doi.org/10.1261/ma.033654.112); PMID: 22895820
25. S. K. Kim, Y. H. Noh, J. R. Koo, H. S. Yun, Effect of expression of genes in the sphingolipid synthesis pathway on the biosynthesis of ceramide in *Saccharomyces cerevisiae*. *J. Microbiol. Biotechnol.* **20**, 356–362 (2010). PMID: 20208441
26. M. Tanigawa, A. Kihara, M. Terashima, T. Takahara, T. Maeda, Sphingolipids regulate the yeast high-osmolality glycerol response pathway. *Mol. Cell. Biol.* **32**, 2861–2870 (2012). doi: [10.1128/MCB.06111-11](https://doi.org/10.1128/MCB.06111-11); PMID: 22586268
27. J. Fernius, A. L. Marston, Establishment of cohesion at the pericentromere by the Ctf19 kinetochore subcomplex and the replication fork-associated factor, Csm3. *PLOS Genet.* **5**, e1000629 (2009). doi: [10.1371/journal.pgen.1000629](https://doi.org/10.1371/journal.pgen.1000629); PMID: 19730685
28. S. C. Di Rienzi, D. Collingwood, M. K. Raghuraman, B. J. Brewer, Fragile genomic sites are associated with origins of replication. *Genome Biol. Evol.* **1**, 350–363 (2009). doi: [10.1093/gbe/evp034](https://doi.org/10.1093/gbe/evp034); PMID: 20333204
29. Z. Bloom-Ackermann *et al.*, A comprehensive tRNA deletion library unravels the genetic architecture of the tRNA pool. *PLOS Genet.* **10**, e1004084 (2014). doi: [10.1371/journal.pgen.1004084](https://doi.org/10.1371/journal.pgen.1004084); PMID: 24453985
30. D. G. Gibson *et al.*, One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20404–20409 (2008). doi: [10.1073/pnas.0811011106](https://doi.org/10.1073/pnas.0811011106); PMID: 19073939
31. Y. Shen *et al.*, SCRaMble generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* **26**, 36–49 (2016). doi: [10.1101/gr.193433.115](https://doi.org/10.1101/gr.193433.115); PMID: 26566658
32. J. Parenteau *et al.*, Preservation of gene duplication increases the regulatory spectrum of ribosomal protein genes and enhances growth under stress. *Cell Reports* **13**, 2516–2526 (2015). doi: [10.1016/j.celrep.2015.11.033](https://doi.org/10.1016/j.celrep.2015.11.033); PMID: 26686636
33. M. Müller, K. Ch. Daoulas, Calculating the free energy of self-assembled structures by thermodynamic integration. *J. Chem. Phys.* **128**, 024903 (2008). doi: [10.1063/1.2818565](https://doi.org/10.1063/1.2818565); PMID: 18205471
34. J. Bitinaite *et al.*, USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* **35**, 1992–2002 (2007). doi: [10.1093/nar/gkm041](https://doi.org/10.1093/nar/gkm041); PMID: 1734463
35. R. Lorenz *et al.*, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011). doi: [10.1186/1748-7188-6-26](https://doi.org/10.1186/1748-7188-6-26); PMID: 22115189
36. W. Rychlik, W. J. Spencer, R. E. Rhoads, Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Res.* **18**, 6409–6412 (1990). doi: [10.1093/nar/18.21.6409](https://doi.org/10.1093/nar/18.21.6409); PMID: 2243783
37. J. D. Boeke, F. LaCrute, G. R. Fink, A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**, 345–346 (1984). doi: [10.1007/BF00330984](https://doi.org/10.1007/BF00330984); PMID: 6394957
38. L. Colleaux, L. D'Auriol, F. Galibert, B. Dujon, Recognition and cleavage site of the intron-encoded omega transposase. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6022–6026 (1988). doi: [10.1073/pnas.85.16.6022](https://doi.org/10.1073/pnas.85.16.6022); PMID: 2842757
39. U. Güldener, S. Heck, T. Fielder, J. Beinhauer, J. H. Hegemann, A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res.* **24**, 2519–2524 (1996). doi: [10.1093/nar/24.13.2519](https://doi.org/10.1093/nar/24.13.2519); PMID: 8692690
40. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: [10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324); PMID: 19451168
41. A. McKenna *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: [10.1101/gr.107524.110](https://doi.org/10.1101/gr.107524.110); PMID: 20644199
42. H. Li *et al.*, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi: [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352); PMID: 19505943
43. I. Milne *et al.*, Using Tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **14**, 193–202 (2013). doi: [10.1093/bib/bbs012](https://doi.org/10.1093/bib/bbs012); PMID: 22445902
44. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009). doi: [10.1186/gb-2009-10-3-r25](https://doi.org/10.1186/gb-2009-10-3-r25); PMID: 19261174
45. D. A. Nickerson, V. O. Tobe, S. L. Taylor, PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res.* **25**, 2745–2751 (1997). doi: [10.1093/nar/25.14.2745](https://doi.org/10.1093/nar/25.14.2745); PMID: 9207020
46. J. Herschleb, G. Ananiev, D. C. Schwartz, Pulsed-field gel electrophoresis. *Nat. Protoc.* **2**, 677–684 (2007). doi: [10.1038/nprot.2007.94](https://doi.org/10.1038/nprot.2007.94); PMID: 17406630
47. D. L. Lindstrom, D. E. Gottschling, The mother enrichment program: A genetic system for facile replicative life span analysis in *Saccharomyces cerevisiae*. *Genetics* **183**, 413–422 (2009). doi: [10.1534/genetics.109.106229](https://doi.org/10.1534/genetics.109.106229); PMID: 19652178
48. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009). doi: [10.1093/bioinformatics/btp120](https://doi.org/10.1093/bioinformatics/btp120); PMID: 19289445
49. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010). doi: [10.1186/gb-2010-11-10-r106](https://doi.org/10.1186/gb-2010-11-10-r106); PMID: 20979621
50. B. Wen *et al.*, iQuant: An automated pipeline for quantitative proteomics based upon isobaric tags. *Proteomics* **14**, 2280–2285 (2014). doi: [10.1002/pmic.201300361](https://doi.org/10.1002/pmic.201300361); PMID: 25069810
51. J. Xia, D. S. Wishart, Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **6**, 743–760 (2011). doi: [10.1038/nprot.2011.319](https://doi.org/10.1038/nprot.2011.319); PMID: 21637195
52. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Meth.* **57**, 289–300 (1995).
53. T. Jewison *et al.*, YMDB: The Yeast Metabolome Database. *Nucleic Acids Res.* **40** (D1), D815–D820 (2012). doi: [10.1093/nar/gkr916](https://doi.org/10.1093/nar/gkr916); PMID: 22064855
54. C. A. Müller *et al.*, The dynamics of genome replication using deep sequencing. *Nucleic Acids Res.* **42**, e3 (2014). doi: [10.1093/nar/gkt878](https://doi.org/10.1093/nar/gkt878); PMID: 24089142

## ACKNOWLEDGMENTS

The project was funded by China National High Technology Research and Development Program (“863” Program: 2012AA02A708) to H.Y. and National Natural Science Foundation of China (21621004 and 21390203) to Y.Y.; a Chancellor’s Fellowship from the University of Edinburgh, a start-up fund from Scottish Universities Life Sciences Alliance and Biotechnology and Biological Sciences Research Council grants (BB/M005690/1, BB/M025640/1, and BB/M00029X/1) to Y.C.; the Wellcome Trust to A.M. (090903 and 092076); U.S. NSF grants MCB-1026068 and MCB-1158201 to J.D.B.; U.S. NSF grant MCB-1445545 to J.S.B.; and U.S. Department of Energy grant DE-FG02097ER25308 to S.M.R. Work at Tsinghua University was supported by National Science Foundation of China grant 31471254 and Chinese Ministry of Science and Technology grant 2012CB725201 to J.D. Funding from ERASynBio and Agence Nationale pour la Recherche (IESY ANR-14-SYNB-0001-03) to R.K. We thank W. Zhang for many comments on the manuscript. J.D.B. and J.S.B. are founders and directors of Neochromosome Inc. J.D.B. serves as a scientific advisor to Recombinetics Inc. and Sample6, Inc. These arrangements are reviewed and managed by the committees on conflict of interest at New York University Langone Medical Center (J.D.B.) and Johns Hopkins University (J.S.B.). Y.C., H.Y., and J.D.B. conceived the concept and designed the experiments. L.M., G.S., S.M.R., J.S.B., and J.D.B. designed the synII chromosome. Y.S., Y.W., T.C., J.G., Y.L., and Y.C. designed experiments. T.C., F.G., H.Z., Y.S., S.C., D.A., Z.L., R.W., W.L., and F.T. performed experiments. Y.W. and J.G. analyzed genomic and RNA-seq data. Y.F., B.Z., and J.G. analyzed mass spectrometry and LC-MS data. Y.S. and D.A. analyzed synII cell cycle data. C.M. and C.N. analyzed synII replication timing profiles. G.M., H.M., and R.K. analyzed synII 3C/3D structure. The manuscript was written by Y.S., Y.W., T.C., F.G., J.D., and Y.C., with input from C.M., C.N., R.K., C.F., A.M., J.S.B., J.D.B., and H.Y. and with comments provided by all other authors. SynII genomic sequencing data and RNA-seq data have been deposited in the Sequence Read Archive (SRA) sequence database ([www.ncbi.nlm.nih.gov/sra/](http://www.ncbi.nlm.nih.gov/sra/)), accession code SRP062892. Proteomics data have been submitted to the ProteomeXchange ([www.proteomexchange.org/](http://www.proteomexchange.org/)), accession code PXD003435. The metabolomics and lipidomics data have been deposited to the MetaboLights metabolomics repository (accession number MTBLS296; accessed via link [www.ebi.ac.uk/metabolights/MTBLS296](http://www.ebi.ac.uk/metabolights/MTBLS296)). The designed and sequenced synII information is uploaded with accession codes CP013607 and CP013608 on GenBank. Additional information [synII design diagram, PCRTag sequences, variants in physical strain (yeast\_chr02\_9\_01), PCR primers, and summary of megachunks, chunks, and minichunks] related to synII can be accessed on the Sc2.0 website ([www.syntheticyeast.org](http://www.syntheticyeast.org)).

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/355/6329/eaaf4791/suppl/DC1](http://www.sciencemag.org/content/355/6329/eaaf4791/suppl/DC1)  
Figs. S1 to S17  
Tables S1 to S7  
References

16 February 2016; accepted 1 February 2017  
10.1126/science.aaf4791



## Deep functional analysis of *synII*, a 770-kilobase synthetic yeast chromosome

Yue Shen, Yun Wang, Tai Chen, Feng Gao, Jianhui Gong, Dariusz Abramczyk, Roy Walker, Hongcui Zhao, Shihong Chen, Wei Liu, Yisha Luo, Carolin A. Müller, Adrien Paul-Dubois-Taine, Bonnie Alver, Giovanni Stracquadanio, Leslie A. Mitchell, Zhouqing Luo, Yanqun Fan, Baojin Zhou, Bo Wen, Fengji Tan, Yujia Wang, Jin Zi, Zexiong Xie, Bingzhi Li, Kun Yang, Sarah M. Richardson, Hui Jiang, Christopher E. French, Conrad A. Nieduszynski, Romain Koszul, Adele L. Marston, Yingjin Yuan, Jian Wang, Joel S. Bader, Junbiao Dai, Jef D. Boeke, Xun Xu, Yizhi Cai and Huanming Yang (March 9, 2017)

*Science* **355** (6329), . [doi: 10.1126/science.aaf4791]

Editor's Summary

---

This copy is for your personal, non-commercial use only.

---

- |                      |  |
|----------------------|--|
| <b>Article Tools</b> | Visit the online version of this article to access the personalization and article tools:<br><a href="http://science.sciencemag.org/content/355/6329/eaaf4791">http://science.sciencemag.org/content/355/6329/eaaf4791</a> |
| <b>Permissions</b>   | Obtain information about reproducing this article:<br><a href="http://www.sciencemag.org/about/permissions.dtl">http://www.sciencemag.org/about/permissions.dtl</a>  |

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.