Originally published 9 March 2017; revised 20 March 2017

www.sciencemag.org/content/355/6329/1040/suppl/DC1

Supplementary Materials for

Design of a synthetic yeast genome

Sarah M. Richardson,* Leslie A. Mitchell,* Giovanni Stracquadanio,* Kun Yang,*
Jessica S. Dymond, James E. DiCarlo, Dongwon Lee, Cheng Lai Victor Huang,
Srinivasan Chandrasegaran, Yizhi Cai, Jef D. Boeke,† Joel S. Bader†

*These authors contributed equally to this work.
†Corresponding author. Email: jef.boeke@nyumc.org (J.D.B.); joel.bader@jhu.edu (J.S.B.)

Published 10 March 2017, *Science* **355**, 1040 (2017)
DOI: 10.1126/science.aaf4557

This PDF file includes

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/355/6329/1040/suppl/DC1)

Movie S1

Revision (20 March 2017): In this revision, the figures and references have been inserted. The conclusions of the Research Article remain unchanged. The originally posted supplementary materials are available here.

Materials and Methods

Sc2.0 consortium
Consortium members all signed formal agreements requiring adherence to certain principles, such as a commitment of each team to raise funds for their chromosome project(s), as well as a Statement of Principles (*37*). Step-by-step protocols and training kits were used to help teams relatively unfamiliar with yeast genetics implement the SwAP-In protocol, perform regular fitness testing of the resulting yeast, perform whole genome sequencing as validation and make all relevant data publicly available. As chromosomes are finished, the more voluminous tables and datasets are made available on the Sc2.0 web site (www.syntheticyeast.org) under the Sc2.0 data section, including design diagrams, PCRtag tables, feature tables, primer tables, and variant tables of physical strains compared to the final design.

Source of reference sequences
The reference sequence was downloaded from SGD (http://downloads.yeastgenome.org/curation/chromosomal_feature/saccharomyces_cerevisiae.gff), whose sequence was last updated by SGD on Feb. 3, 2011.

RE site selection for segmentation
Our chunk/megachunk design strategy incorporates selectable marker and flanking homology sequences in the rightmost chunk of each megachunk during synthesis. The design strategy mandates megachunks between ~30 and 60 kb long, with RE sites that leave unique, non-palindromic overhangs placed every 10 kb. Enzymes may be reused within a megachunk as long as they are unique to the chunk they define and the two flanking chunks. Enzymes that can cut to leave distinct overhangs (e.g. *Sfi*I) can be reused at both ends of the same chunk (as well as flanking chunks). All overhangs must be mutually incompatible within the megachunk to assure proper chunk assembly order and orientation. Each megachunk ends in a marker and a stretch of wild type targeting sequence, both of which will ultimately be overwritten in the next SwAP-In step; for this reason, no megachunk selectable marker is permitted to interrupt an essential gene, a gene required for normal growth speed, or the corresponding UTRs. Enzymes must be carefully chosen to avoid any that appear in the wild type targeting sequence or the UTRs of the appended markers; but those that appear in the coding region of the marker can be removed by recoding. A special challenge is created by the requirement that the selectable marker in the rightmost chunk of each megachunk must be flanked by two different RE sites ("*L*" and "*R*" in main text) used to release the left end of megachunk *m*+1 and the right end of megachunk *m* (Fig. 2, main text). We initially planned to locate the RE sites 500 bp to the left and right of the selectable marker, which itself must be substituted for a nonessential ORF. However, these constraints proved impossible to satisfy when sequence edits were limited to synonymous recoding. We solved this problem by using RE sites recognized by Type IIB enzymes. These unusual enzymes are ideal in this context because they excise their recognition site by cleaving on both sides of it. While the precise cutting by these enzymes is likely limited by the "long reach" between the recognition and cut sites, the function of these particular cuts is simply to release a sequence that can cleanly recombine with endogenous target sequences. Thus the precision of the cut site is not an important factor for these sites (unlike the sites used for joining chunks by ligation, where precision is critical). Throughout the RE selection process, emphasis is placed on prioritizing less expensive and more effective restriction enzymes as these are in general more reliable.

Incorporation of 30-60 kb megachunks into S. cerevisiae cells

Sc2.0 chromosomes are built by SwAP-In, which relies on three screening steps to identify "winners". As an example, consider transformation of a megachunk encoding *URA3* at its rightmost end, intended to overwrite a pre-existing *LEU2* marker in the genome, corresponding to the leftmost end of the megachunk. Following transformation of the megachunk, transformants are plated on medium lacking uracil to select for cells in which the rightmost end of the incoming DNA carrying the marker has integrated. Next, transformation plates are replica plated onto medium lacking leucine to identify colonies that cannot grow, indicating that the incoming DNA has "overwritten" the pre-existing *LEU2* marker. Finally, the subset of $Ura3^{+}/Leu2^{-}$ colonies is subjected to PCRTag analysis across the newly integrated megachunk to identify colonies that exclusively produce synthetic PCRTag amplicons. On average, from 100 primary transformants, 10% of them encode the correct "plate-based phenotype" ($Ura3^{+}/Leu2^{-}$) and 10% of those produce only synthetic PCRTag amplicons. Thus, 1% of screened colonies are winners. A typical transformation yields thousands of colonies, and thus multiple winners are identified for each megachunk integration step. These frequencies can vary substantially across megachunks.

A potential improvement on the SwAP-In design

One unanticipated consequence of the SwAP-In design strategy is that the knockout corresponding to the site of selectable marker insertion occasionally produces an unwanted phenotype such as slow growth or poor transformation efficiency, reducing the efficiency of the next round of megachunk incorporation. This is a temporary problem, as the gene is restored upon the integration of the next megachunk. Very clear examples of this can be seen in the SOM of Shen et al. (*10*). In future projects, we will insert the selectable markers downstream of an ORF, incorporating a "recyclable" 3′ UTR region for the target gene along with the selection marker so as not to disturb the endogenous gene's function. This change should improve this shortcoming of the current iteration of SwAP-In.

PCRTag design – further considerations

The synthetic sequence of each PCRTag primer pair is generated by running the wild-type sequence through GENEDESIGN's "most different" algorithm (*7*), which maximizes nucleotide differences between primers synonymously. To minimize translational effects, changes to codons with extremely low relative synonymous codon usage value (*38*) are prohibited. Cross-priming is minimized by forcing primers to begin and end in the wobble position so that the first and last bases of each primer in the pair are guaranteed to differ. Moreover, primer sequences must differ overall by at least 33%. All primers have a melting temperature within a narrow range to permit a single set of PCR conditions, and all primers are unique in the wild-type genome according to BLAST (*39*). Amplicons specified by two primer pairs, one matching the native sequence and one the synthetic, are chosen such that each set of primer pairs cannot amplify any sequence in the genome under one kilobase long except their design target. Amplicons are restricted to 200–500 bp to create electrophoresis bands in the same size range and thus are easily verified visually. Shorter amplicons also allow relatively "dirty" DNA to be employed as template. Given that the 5′ regions of CDSs are known to have special codon usage requirements (*40, 41*), PCRTags are prohibited from the first 100 bases of a gene to avoid unintentional disruption of regulatory sequence or mRNA structure. To maximize genome coverage, one amplicon is chosen per kilobase of gene sequence, and the amplicons are designed to minimize overlap.

Constructing a chimeric IXL-synIXR linear chromosome (yLM461)
Similar to virtually all wild-type eukaryotic nuclear chromosomes, Sc2.0 designer chromosomes are intended to be linear in structure with terminal telomere sequences. An exceptional case is s*ynIXR*, originally built as a circular molecule including ~10 kb of bacterial artificial chromosome (BAC) sequence in addition to the 90 kb synthetic yeast chromosome arm (*5*) (fig. S7A). The circular format of "*synIXR-BAC*" facilitated both construction and transfer of the chromosome arm from yeast into *E. coli* for characterization. To generate a strain expressing *synIXR* as a hybrid synthetic/native linear chromosome attached to native *IXL*, we digested the *synIXR-BAC* with two unique restriction enzymes flanking the *synIXR* sequence, *Not*I and *Asi*SI. The resulting 90,994 bp fragment encoded the entire synthetic chromosome arm, encompassing *LEU2* followed by *YIL002C* through to *YIR044C*. The digestion product was extracted once with phenol/chloroform and then ethanol precipitated. A "left end cap" (LEC) construct (pJS315), encoding ~1.1kb of native chromosome *IXL* sequence upstream of *YIL002W-A* was digested with *BsaI* and *Asi*SI. A universal telomere cap (UTC) construct (pJS160) encoding ~350bp of *S. cerevisiae* telomere repeats (($TG_{(1-3)}$)) was digested with *Not*I and *Bsa*I. Both the LEC and UTC fragments were subjected to gel purification prior to ligation in equimolar quantity with the *synIXR* digestion product. The ligation product was transformed into competent yeast in which *YIL001W* and *YIR039C* were deleted with *URA3* and *KanMX*, respectively (yJS698) (fig. S7B). Transformants were plated on SC–Leu and replica plated onto SC–Ura or YPD G418 medium. Leu^+, Ura^-, $G418^s$ transformants were selected for further testing by PCRTag analysis. One clone, yLM062, was found to amplify all synthetic PCRTags but no wild type PCRTags. To build an unmarked *IXL-synIXR*, we converted the *LEU2* marker to *URA3*, which was then overwritten with a spanning PCR product and selection carried out on 5-FOA to generate yLM461. Marker "swapper" plasmids, designed for quickly converting any strain from *URA3* to *LEU2* (pLM090) and vice versa (pLM091), were constructed.

The linear structure of *IXL-synIXR* was interrogated by pulsed field gel electrophoresis; *IXL-synIXR* migrated identically to native chromosome *IX* (fig. S7C). In comparison, the *synIXR-BAC* strain did not penetrate the pulsed field gel. For size reference, a linearized version of *synIXR-BAC* ~110 kb in size and expressed independently from native *IXL* (*42*) migrated faster than all native yeast chromosomes (fig. S7C). Growth of the *IXL-synIXR* strain was indistinguishable from wild type on all media tested (see (*9*), fig. S8). Moreover, colony size and morphology were identical to wild type (fig. S7D). The chimeric *IXL-synIXR* chromosome provides a useful template for construction of full-length *synIX*.

Endoreduplication backcrossing and intercrossing
Destabilization of a chromosome by inducibly expressing the galactose promoter in *cis* (*29*) is a useful means to expel native chromosomes from native/synthetic chromosome heterodiploid strains. This is important to circumvent the generation of patchwork wild type/synthetic chromosomes in meiotic progeny as a result of recombination (*5*). We define an endoreduplication backcross as the destabilization of a single native chromosome, producing a strain 2n-1 for the retained synthetic chromosome. This is also a useful strategy to debug synthetic chromosomes, in particular if mutations have accumulated elsewhere in the genome of a strain encoding a synthetic chromosome. We define an endoreduplication intercross as the simultaneous destabilization of two native chromosomes, producing a strain 2n-2 for two

retained synthetic chromosomes. This is useful to combine synthetic chromosomes into a single strain.

Here we demonstrate the utility of endoreduplication intecrossing to build poly-synthetic strains (Fig. 3A, fig. S6). The growth of four spore tetrads for *synVI synIXR* was consistent with endoduplication of both synthetic chromosomes prior to sporulation. Two spore tetrads for *synIII synVI* and *synIII synIXR* was also consistent with endoduplication; viable spores resulted from co-segregation of *synIII* with the essential *SUP61* tRNA integrated at the *HO* locus on chromosome 4, and inviable spores presumably died in the absence of this tRNA gene.

pGAL1-CEN3::URA3(Kl), *pGAL1-CEN6::URA3(Kl)*, and *pGAL-CEN9::URA3(Kl)* constructs were generously provided by Rodney Rothstein (*30*). In each case, the prepped construct was digested with *NotI* to drop out the insert (typically ~5kb) and used directly for yeast transformation followed by selection on synthetic medium lacking uracil (SC–Ura). Integration was verified using a *URA3(Kl)* specific primer paired with a locus specific primer that annealed outside of the region of integration. Synthetic strains used for integration of *pGAL-CENx* were as follows: (i) *synIII*: yLM422, where the essential *SUP61* tRNA is integrated into the HO locus on chromosome *IV*; (ii) *synVI*: yLM402 (*MAT**a***) and yLM399 (*MATα*), where in both strains the two nucleotides of *MOB2* had been corrected to the designed sequence although the glycerol negative growth suppression defect had not yet been corrected (*9*); (iii) *synIXR*: yLM461, where *synIXR* has been converted from a bacterial artificial chromosome to a linear, synthetic/native hybrid chromosome attached to native *IXL* (SOM, main text).

Following mating, selection and single colony purification, heterozygous diploid strains (encoding two synthetic chromosomes plus the *pGAL-CENx* alleles integrated into the corresponding native chromosomes) were grown in liquid YP+galactose (2%) medium for 12-18 hours at 30°C with rotation. 200 µL of culture was spread on SC medium supplemented with 5-fluoroorotic acid (5-FOA) and extra uracil to select cells that no longer expressed *URA3*. After incubation for 2 days at 30°C, 5-FOA resistant colonies were streaked on 5-FOA and ~20 unrelated single colonies patched on YPD. For crosses involving *synIII*, which encodes the *MATα* locus (*synVI pGAL1-CEN3* x *synIII pGAL1-CEN6* and *synIXR pGAL1-CEN3* x *synIII pGAL1-CEN9*), patches were pre-screened to identify those with a gain-of-function A-mater phenotype, resulting from loss of the *MATα* locus, consistent with loss native chromosome *III*. Genomic DNA was prepared from 12 isolates for each putative 2n-2 strain and the exclusive amplification of synthetic as compared to wild type PCRTags was confirmed using a subset of PCRTags (one every ~10-20 kb for each chromosome). 2n-2 strains verified by PCRTagging were subjected to sporulation at room temperature for 5 days followed by dissection. For *synIII synVI 2n-2* and *synIII synIXR 2n-2* strains the *MAT**a*** locus, an essential feature to drive the meiotic program, was expressed from plamsid pLM329. The double synthetic strains (*synIII synVI* yLM684; *synVI synIXR* yLM654 and yLM655) were then "corrected" at the *PRE4* locus by incorporating a *WT-SYN.PRE4* allele (*9*). The final double synthetic strains are *synIII synVI* (yLM890), *synIII synIXR* (yLM758, *MATα*), *synVI synIXR* (yLM892, *MAT**a***), *synVI synXIR* (yLM894 *MATα*). The presence of all synthetic PCRTags and absence of all native PCRTags was verified for each of the four strains.

To generate the triple synthetic strain, into the uncorrected *PRE4 synIII synVI* and *synVI synIXR* strains were integrated *pGAL-CEN9* (yLM669) and *pGAL-CEN3* (yLM672), respectively. Integration was verified by PCR as described above. Mating, conditional chromosome destabilization, sporulation in the presence of plasmid-based *MAT**a***, dissection, and final PCRTag analysis was carried out as described above. The *synIII synVI synIXR* strain (yLM759) was subsequently corrected at the *PRE4* locus using the *PRE4.SYN-WT* allele to generate yLM896.

*Cloning the MAT**a** locus*

*MAT***a** was PCR amplified with primers designed to anneal ~1kb up and downstream of the locus and additionally encode 30bp of homology to the terminal ends of a *Sma*I-digested pRS413 vector (*34*). Genomic DNA extracted from BY4741 was used as template in a PCR reaction with Phusion polymerase (NEB; M0430L). The PCR product was purified using the PureLink Quick PCR Purification Kit (Invitrogen; K310002) and subjected to isothermal assembly (*35*) with *Sma*I-digested pRS413 that had been gel purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research; D4008). Following transformation into competent *E. coli* cells (Top10), assemblies were verified from miniprepped plasmids digested with *PvuII*. The subcloned *MAT***a** locus (pLM329) was functionally validated by transforming the construct into the MATα BY4742 strain and confirming an inability to mate with BY4741 as well as successful sporulation of *MATα/MATα synIII/synIII* strains generated by chromosome destabilization.

BIOSTUDIO global and local edits

Global edits refer to edits that are made in batch across a chromosome or that involve non-local constraints, such as the uniqueness constraints for PCRTags. Local edits refer to changes that do not require attention to other regions. Note, however, that local edits almost always affect the coordinates of downstream features. Insertions and deletions therefore require careful attention to the annotation of every downstream feature, and even some upstream features. We use a simple model of feature position logic to determine how to update annotations and to describe feature modifications (fig. S5). Special attention must be paid to certain types of feature deletions; for example, if an intron is deleted, the adjacent CDS features must be merged and not remain annotated as two separate, contiguous features.

Typically, during the local editing phase, BIOSTUDIO was employed by the yeast biologist to delete specific features in a specific order (see Table 2 for a typical sequence of edit types performed for a chromosome). That is, the individual performs all edits of a particular type, for example deletion of introns, by systematically moving from left to right through the sequence, and then returns to the left end to initiate the next type of edit. See Movie S1 for a graphical view of the steps involved in designing *synV (12)*.

Introns were deleted using the "delete" command after selecting the intron feature, resulting in precise intron deletion. In contrast, tRNA genes were removed using the "delete and insert loxPsym site" command. In most cases, tRNA genes were adjacent to one or more repeat elements such as Ty LTR sequences. In such cases, the rule applied was to delete everything between the farthest LTR or Ty sequence and the tRNA gene as well as the tRNA gene itself, provided that no "verified ORFs" lay between such features. This led to the deletion of

substantial additional non-tRNA, non-repeat sequence from the chromosomes. In one case, two tRNA genes and all associated repeats were deleted with a single command.

Universal telomere caps (UTC3; pLM258) consisting of a telomere seed sequence and a consensus core X sequence, kindly provided by Ed Louis of Oxford University, were inserted at each end of the chromosome in the appropriate orientation. Centromere sequences were not modified except to ensure that they were flanked by *loxPsym* sites lying between them and the nearest ORF on each side.

BIOSTUDIO/GBROWSE graphical interface conventions
BIOSTUDIO offers a "skin" for GBROWSE features to enhance editing in the form of a customized configuration file. Default preferences for Sc2.0 are to render essential genes in red, important but not essential genes in purple, and other genes in varying shades of blue according to their status in the annotation.

Version control
In addition to providing an interface for editing, we developed BIOSTUDIO to provide a version and revision control system. As described below, individual stages in the design are stored as checkpoints in GFF format, permitting rollback to a previous version. We also developed software to compare versions to create a list of differences, similar to difference finding for text documents. Accept/reject decisions for changes identified in this manner for DNA sequences were made manually, similar to the requirement for manual decisions for merging text documents. While individual chromosomes were edited in parallel, we did not in general have multiple editors working simultaneously and independently on a single chromosome, making merge functionality less relevant. The repository is centralized like CVS (*43*) and SVN (*44*), rather than distributed like GIT (*45*), and is specialized for chromosomes annotated using the GFF standard.

The annotation log includes the complete information required to effect each change: the chromosome, the chromosome version, and the affected sequence coordinates. Each change is then "compiled" from a human-readable form to a coded command that can be applied or removed. Version control was also extended to cover the various "living" versions of chromosomes.

Version control rules applied
Chromosome versions are stored as GFF files, which contain a header region, a table of feature annotations, and a FASTA sequence.

Chromosome versions are denoted yeast_chrCC_M.LL.
CC is the one or two digit chromosome number: 1, 2, 3, ..., 16.
The single digit M is incremented for each major editing stage.
M = 0 is the wild type sequence and should be linked to a public accession number and strain identifier.
The batch edit of adding PCRTags increments M to 1.
The batch edit of Stop codon swaps increments M to 2.
The batch edit of loxPsym site insertions increments M to 3.

Following each increment of M, the minor edit digits are reset to 00 and then increment with additional changes to the chromosome sequence.

Thus, using chromosome 5 as an example, the wild-type sequence is version 5_0.00; adding PCRTags increments to 5_1.00; swapping Stop codons increments to 5_2.00; adding *loxPsym* sites increments to 5_3.00. Please note that to date there have never been minor version increments for major versions 1 and 2.

Following 5_3.00, editing involves corrections and adjustments to the batch-edited sequence. These include deletions of repetitive sequences and Ty elements, removal of tRNA genes, removal of introns, and editing of telomeres and sub-telomeric regions. These edits are typically done by an expert moving from left to right along the chromosome design. Final edits are performed to permit segmentation. These include synonymous recoding to add or remove RE sites. The final sequence ordered for synthesis was 5_3.43.

The header region of the GFF should include a brief description of the differences between the current and previous version. These descriptions should be retained with each new version so that the header of the final version provides the full history of changes back to the wild-type chromosome.

In some cases, the annotation underlying the final ordered sequence has minor errors. For example, the location of a feature such as a *loxPsym* site might have coordinates slightly shifted from the actual location in the synthetic design. These changes require a change to the annotation table in a GFF file, but not to the FASTA sequence that is also part of the GFF. In these cases, the annotation table is corrected and a comment is added to the GFF header. The GFF version is incremented by 1 for each batch of corrections of annotation errors.

If synthesis products deviate from the designed sequence and the errors are detected prior to integration into yeast, either the errors must be fixed, or the designed sequence must be updated to match the physical sequence. A good example is the allowance of a 10% variation in the length of homopolymer runs outside of coding regions. If the decision is to update the designed sequence to the physical sequence, the version number is incremented and the changes must be reflected in the annotations.

The chromosome version corresponding to the first complete synthetic chromosome that replaces the corresponding wild-type chromosome in a living yeast cell receives the major version number 9 and the minor version reverts to 01. For chromosome 5, for example, the designed and ordered version 5_3.43 becomes 5_9.01 when successfully and completely integrated into yeast. Usually integration proceeds incrementally along a chromosome, and then additional recombination steps are required to remove the wild type in favor of the synthetic chromosome. These intermediate strains are tracked in workflow, but they should not receive minor version number increments. Ideally, the sequence of the CC_9.01 living chromosome should identically match the final CC_3.LL version. Any differences should be noted in the header of the CC_9.01 GFF, which should also reference the final CC_3.LL design.

Finally, once a given version of any synthetic chromosome is built, it may be introduced into multiple genetic backgrounds. When this happens, although the strain number changes, the version number does not change. The version number should change only when an intentional change is made to the sequence of a "living" synthetic chromosome.

BIOSTUDIO requirements

BIOSTUDIO requires PERL 5.18.0 or higher, GENEDESIGN 5.0 (*7*), and BIOPERL-LIVE (*36*). If a graphical user interface is desired, GBROWSE (*27*) is required; if a wiki is desired, FosWiki is required.

BIOSTUDIO can be installed on an Amazon Web Services (AWS) instance using an AWS image from the AWS MarketPlace (AMI ID: ami-dcb8f4b6):

Open the Amazon EC2 console at https://console.aws.amazon.com/ec2/
From the console dashboard, choose Launch Instance
Search and choose BIOSTUDIO AWS image from AWS MarketPlace Community AMIs.
Then configure and launch the instance.

BIOSTUDIO can also be installed on any Linux and Mac OS X following the instruction below.

Before installing BIOSTUDIO, dependencies must be installed. To install dependencies on Linux, run this command:

sudo apt-get install git ncbi-blast+ libcairo2 libcairo-perl libfreetype6-dev emboss

To use Homebrew to install dependencies on Mac OS X, run these commands:

brew tap homebrew/science
brew install gcc blast GD cairo emboss pkg-config

To make sure that the perl libraries for Cairo are install properly on Mac OS X, linking .pc may be needed:

sudo cp /opt/X11/lib/pkgconfig/*. pc /usr/local/lib/pkgconfig/

Then using the following commands to download and install BioPerl on Linux or Mac OS X:

```
git clone https://github.com/bioperl/bioperl-live.git
cd bioperl-live/
sudo perl Build.PL
sudo ./Build installdeps
sudo ./Build test
sudo ./Build install
cd ../
git clone https://github.com/bioperl/bioperl-run.git
cd bioperl-run/
sudo perl Build.PL
sudo ./Build installdeps
sudo ./Build install
```

To install GENEDESIGN:

```
git clone https://bitbucket.org/notadoctor/genedesign-dev
cd genedesign-dev/
sudo perl Build.PL
sudo ./Build installdeps
sudo ./Build test
sudo ./Build install
```

If a graphical user interface is needed, GBROWSE can be installed following the instructions from http://gmod.org/wiki/GBrowse_2.0_Install_HOWTO

If a wiki is needed, FosWiki can be downloaded and installed from http://foswiki.org

Once the dependencies have been installed, install BIOSTUDIO:

```
git clone https://bitbucket.org/notadoctor/biostudio-dev
cd biostudio-dev/
sudo perl Build.PL
```

At this point BIOSTUDIO will begin to ask questions to configure itself.

Where should configuration files be installed?
BIOSTUDIO needs write access to a directory where it can store customization and configuration files.

Where should scripts be installed?
This is a directory you will want to add to your $PATH.

Where should BIOSTUDIO write tmp files?
Many temporary files are generated by certain processes; by granting access to an automatically cleaned directory

Enable GBrowse interaction?
Address of GBrowse server?
Only asked if the GBrowse prerequisites have been met. If yes, BIOSTUDIO plug-ins will be added to the GBrowse installation site. The address of GBrowse installation is using is needed (usually localhost).

Enable SGE job farming?
This option should only be used when a grid engine is available. BIOSTUDIO will automatically submit certain embarrassingly parallel computations.

Enable BLAST+ support?
Only asked if the BLAST+ prerequisites have been met.

Enable Cairo graphics support?
Only asked if the Cairo prerequisites have been met.

Once configuration is over, install any missing libraries (if necessary):

```
sudo ./Build installdeps
```

And finally, test and install.

```
sudo ./Build test
sudo ./Build install
```

Genome design with BIOSTUDIO

This section describes how to recreate the process that was followed for the batch editing phase of Sc2.0 synthetic chromosome design.

The annotated *Saccharomyces cerevisiae* S288C genome may be obtained as a single GFF file from the Saccharomyces Genome Database (*26*):

```
curl -O http://downloads.yeastgenome.org/curation/ chromosomal_feature/saccharomyces_cerevisiae.gff
```

Each chromosome should be separated into its own GFF file and entered into BIOSTUDIO version control:

```
BS_AddFromGFF.pl –SPECIES Saccharomyces_cerevisiae --INPUT saccharomyces_cerevisiae.gff
```

The new chromosomes in the genome repository can then be added to GBrowse:

```
BS_FirstGBrowse.pl –c Saccharomyces_cerevisiae_chrI_0_00
```

Add PCRTags to each chromosome and increment the genome version number:

```
BS_PCRTagger.pl --CHROMOSOME Saccharomyces_cerevisiae_chrI_0_00 --EDITOR notadoctor --MEMO "pcr tagging" --ITERATE genome --MINTAGMELT 58 --MAXTAGMELT 60 --MINPERDIFF 33 --MINTAGLEN 19 --MAXTAGLEN 28 --MINAMPLEN 200 --MAXAMPLEN 500 --MAXAMPOLAP 25 --MINORFLEN 501 --FIVEPRIMESTART 101 --MINRSCUVAL 0.06
```

Swap the stop codons to remove the TAG stop codon and increment the genome version number:

```
BS_CodonJuggler.pl --CHROMOSOME Saccharomyces_cerevisiae_chrI_1_00 --EDITOR notadoctor --MEMO "stop swapping" --ITERATE genome --FROM TAG --TO TAA —DUBWHACK
```

Insert *loxPsym* sites three bases pairs downstream of the stop codon of every verified gene on each chromosome and increment the version number:

```
BS_ChromosomeSplicer.pl --CHROMOSOME Saccharomyces_cerevisiae_chrI_2_00 --EDITOR notadoctor --MEMO "first loxPsym seed" --ITERATE genome --ACTION featflank --INSERT loxPsym --FEATURE gene --DISTANCE 3 --DIRECTION 3
```

Then each chromosome may be individually modified by experts, incrementing its chromosome version number.

Supplemental Figures

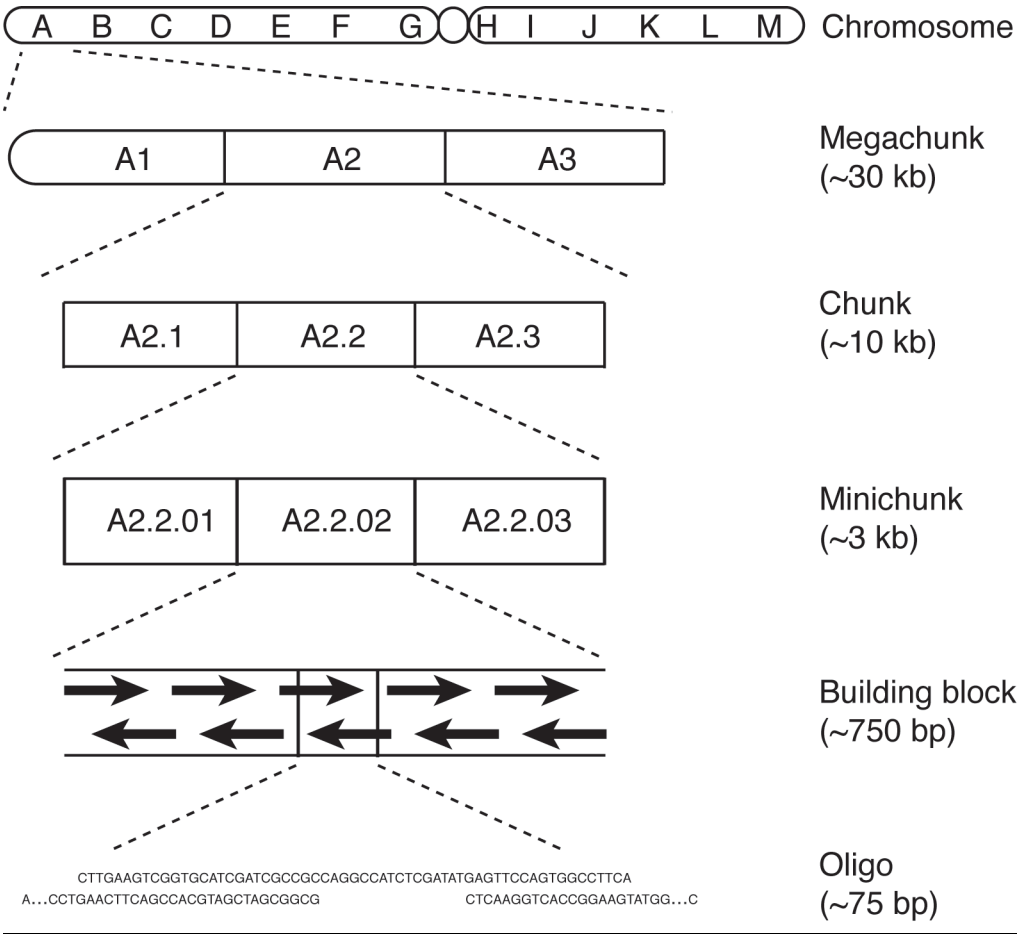

fig. S1. Assembly hierarchy of Sc2.0. Names of DNA molecules used in assembly schemes and the size ranges of each are indicated. Megachunks can range as large as 60 kb.

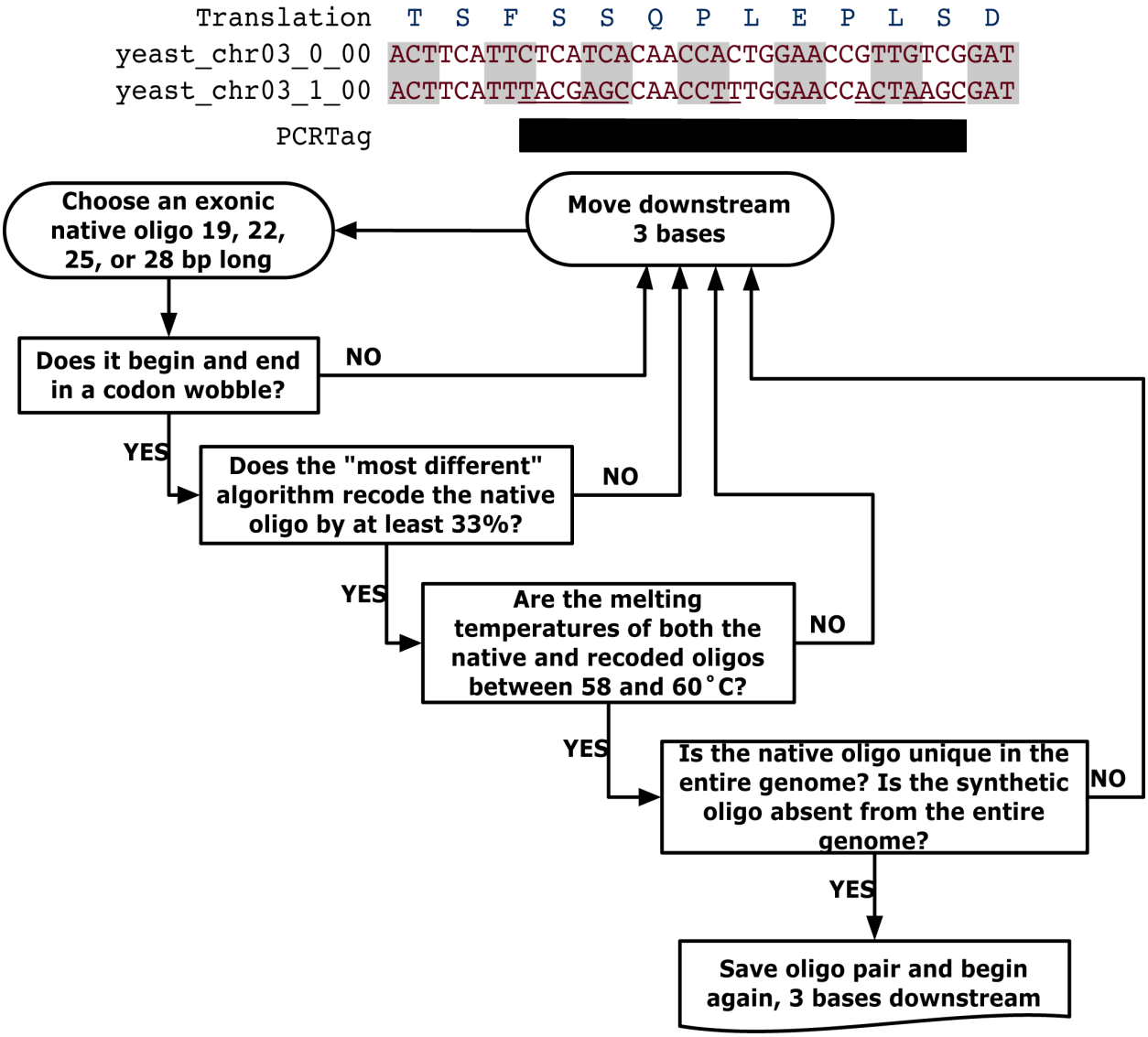

fig. S2. PCRTags are watermarking sequences embedded in open reading frames by recoding. A PCRTag is designed to both begin and end in the wobble base of a codon and to be recoded as much as possible without changing the translation of the gene or violating a minimum relative synonymous codon usage value standard. Changes to the sequence between the native and the recoded version are underlined. Adapted from Dymond et al. (*5*).

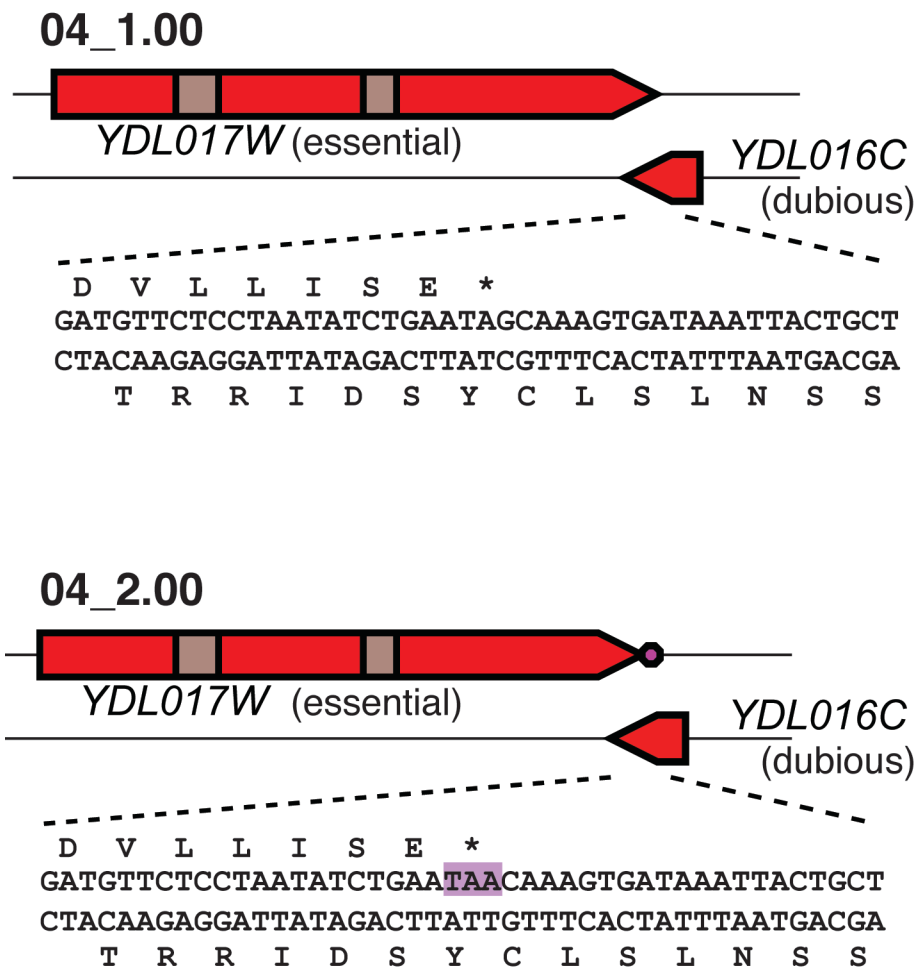

fig. S3. Codon changes in overlapping ORFs. An example from *synIV* is shown of overlapping verified (*YDL017W*) and dubious (*YDL016C*) ORFs. In such cases the change is always made in favor of the verified ORF; in this case, TAG re-coding of the *YDL017W* stop codon fortuitously does not alter the coding sequence of the dubious *YDL016C*.

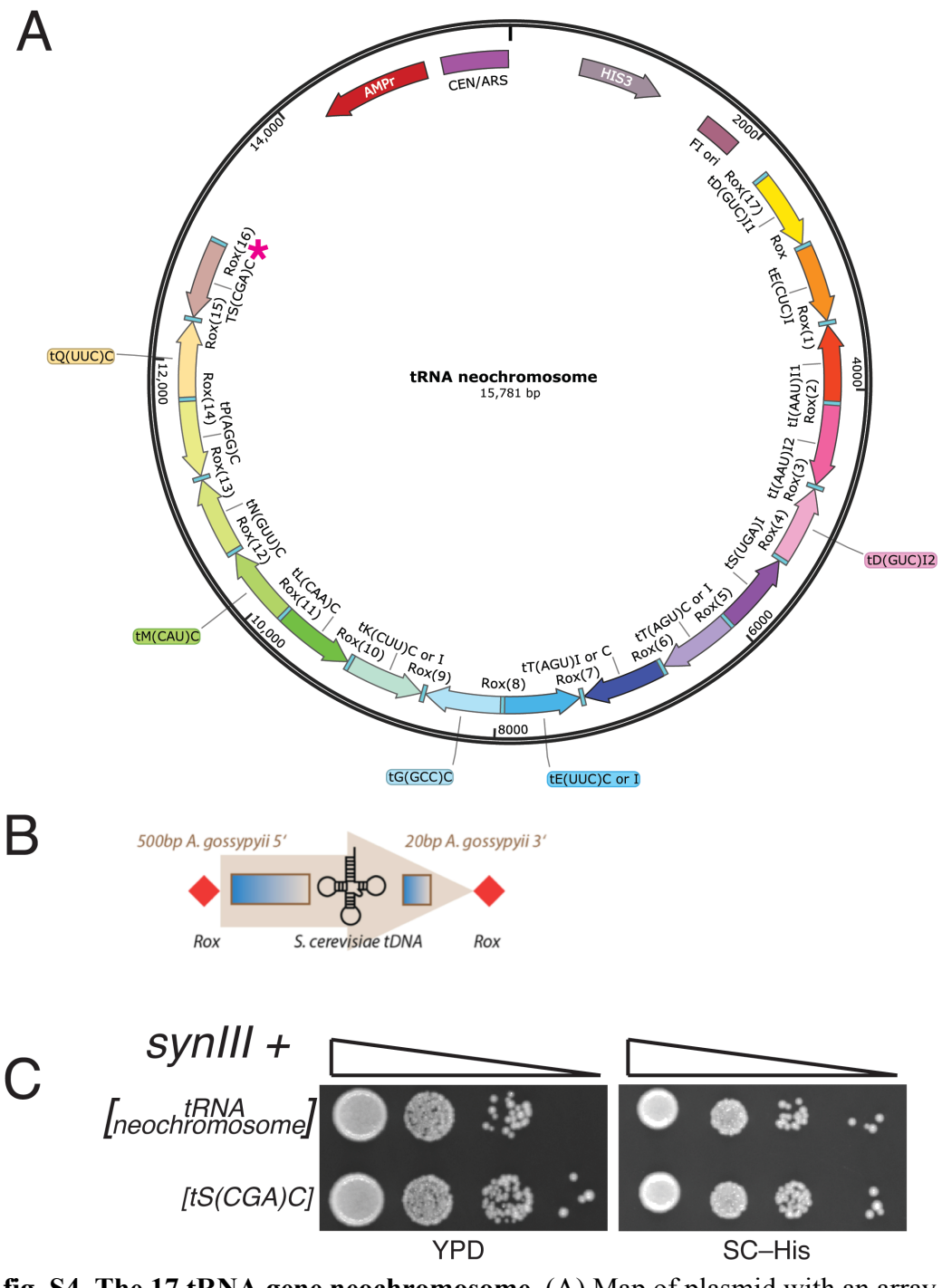

fig. S4. The 17 tRNA gene neochromosome. (A) Map of plasmid with an array of 17 tRNA genes. (B) Design of individual synthetic tRNA genes. Rox sites are recognized by the Dre recombinase and enable orthogonal SCRaMbLE to Cre/lox. (C) Growth of *synIII* strain with above tRNA gene neochromosome or a refactored version of the only essential tRNA gene on *synIII*, *tS(CCGA)C*, both are on a *HIS3 CEN* backbone. Identical colony size suggests tRNA expression is not affected by being on the array of tRNA genes and that the presence of the tRNA array is not deleterious.

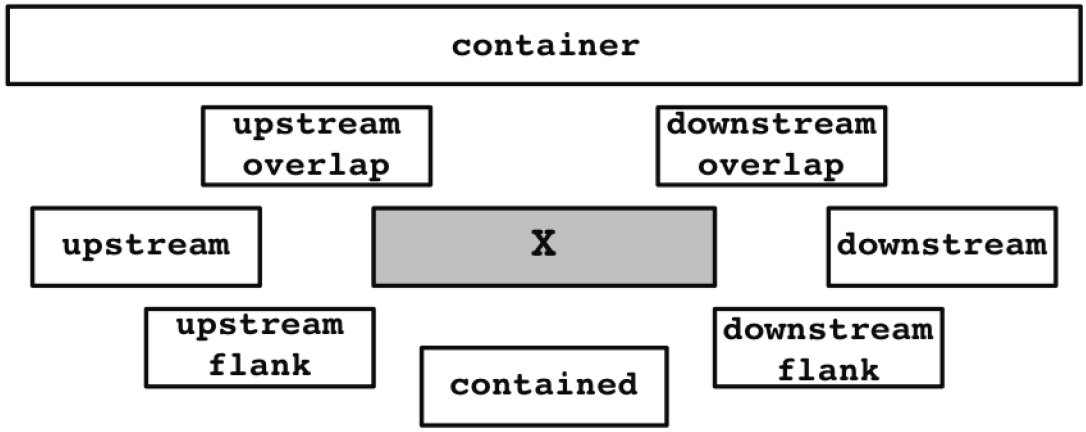

fig. S5. Feature position logic. There are eight possible relationships between a feature **X** and all other features annotated on a chromosome. Deleting **X** will cause start position changes in the downstream features, stop position changes in the container, the upstream overlap and all of the downstream features, and the complete deletion of all contained features. If **X** is an intron, the flanking features are CDSs and the annotations of these must be merged.

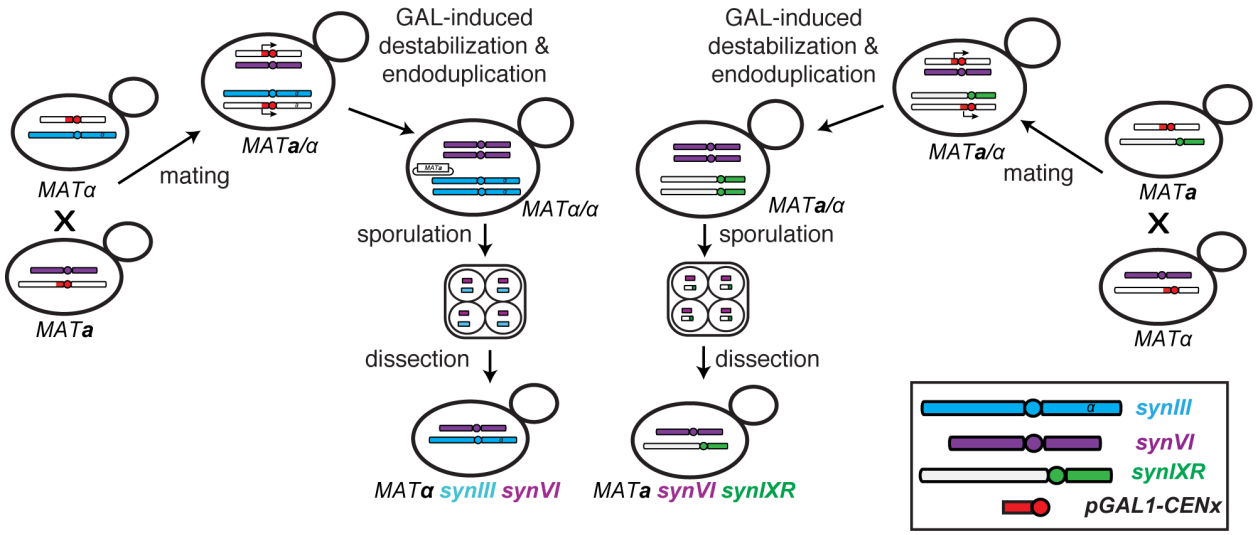

fig. S6. Constructing double-*syn* strains. The *pGAL-CENx* construct was integrated into the appropriate synthetic chromosome strain (*synIII*, blue; *synVI*, purple, *synIXR*, green), which were then mated to generate heterozygous diploid cells (*synIII/III pGAL-CEN6 synVI/VI pGAL-CEN3* and *synVI/VI pGAL-CEN9 synIXR/IX pGAL-CEN6*). Following growth in galactose to induce destabilization of the specified native chromosomes and selection on FOA medium, the 2n-2 state was confirmed by PCRTag analysis. Double-*syn* chromosome strains (*synIII synVI* and *synVI synIXR*) were generated by sporulation and dissection. An episomal copy of *MAT**a*** was introduced to permit sporulation in the *synIII synVI* strain. Similarly, a *synIII synIXR* strain was constructed (not pictured here).

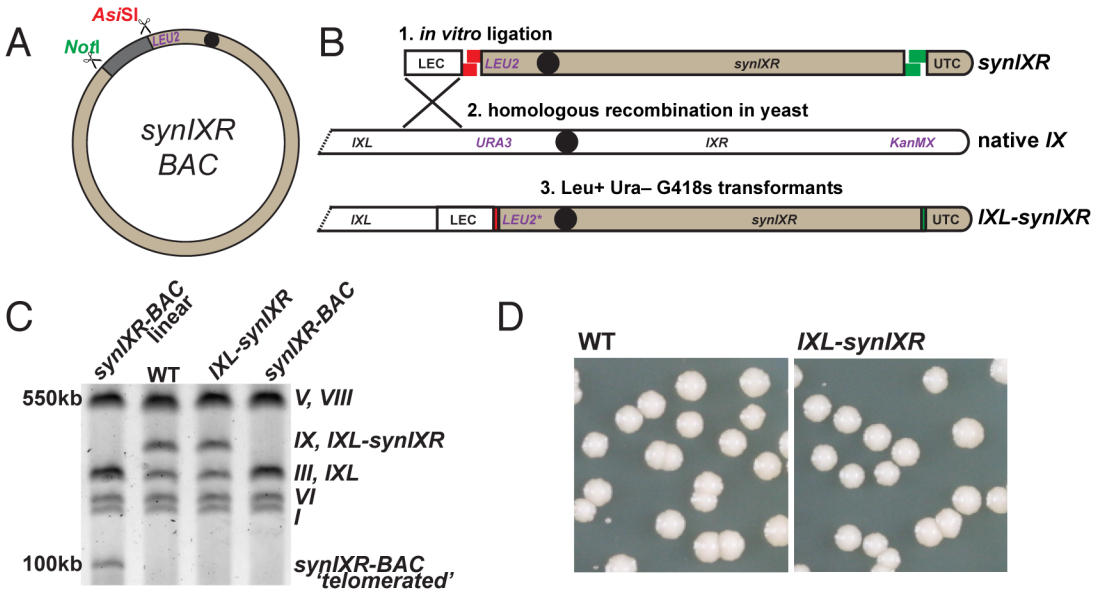

fig. S7. Constructing a *IXL-synIXR* chimeric linear chromosome (A) Structure of *synIXR-BAC*. The s*ynIXR* chromosome arm (brown) encodes a *LEU2* selectable marker, a centromere (black circle), plus unique *Not*I and *Asi*SI restriction enzyme sites flanking the *BAC* sequence (gray). **(B)** Chimeric chromosome construction. *SynIXR* was excised from the BAC by digestion with *Asi*SI (red) and *Not*I (green) and subsequently ligated *in vitro* to a left end cap (LEC) fragment and a universal telomere cap (UTC) fragment. Upon transformation into competent yeast previously engineered to express *URA3* and *KanMX* markers from the loci indicated, the LEC fragment enabled homologous recombination (X) with native chromosome. From transformants able to grow on medium lacking leucine (Leu^+) but not on medium lacking uracil (Ura^-) or supplemented with G418 ($G418^s$), a strain encoding all synthetic but no wild-type PCRTags was identified. The *LEU2* marker in this strain (*LEU2**) was subsequently deleted as described in the SOM text. **(C)** The structure of the *IXL-synIXR* chimeric linear chromosome was evaluated by pulsed-field gel electrophoresis. While the circular *synIXR-BAC* does not penetrate the gel, the *IXL-synIXR* molecule migrates identically to native IX under these conditions. In contrast, a linear derivative of *synIXR-BAC* engineered with the telomerator (*42*) migrates faster than all yeast chromosomes (*synIXR-BAC* 'telomerated'). The native *IXL* chromosome arm co-migrates with native *III*. **(D)** Colony size and morphology of wild type (WT) cells is indistinguishable from cells expressing the chimeric *IXL-synIXR* chromosome.

Movie S1. Movie depicting the editing process for *synV* (*12*). The movie depicts a series of 39 steps involved in the design of Sc2.0 synthetic chromosome *synV* starting from the reference chromosome. The overall diagram shows the main features such as protein coding genes (boxed arrows) annotated using a color scheme in which nonessential genes are blue, essential genes are red, and other important genes (e.g. "slow growth" genes) are purple. The first three editing steps are "global" to the chromosome in which the following edit types are made: (1) PCRTags (watermarks) are edited into the appropriate open reading frame regions; (2) TAG stop codons are converted to TAA codons; (3) *loxPsym* sites are inserted in the 3′ UTRs of each nonessential gene. Following these three global steps, a series of human-directed but software-aided steps are performed, each step sweeping from left to right across the chromosome. These steps include the deletion of repeats and tRNA genes, introns, and subtelomeric sequences, possible replacement of deleted features with *loxPsym* sites, and other custom modifications.

References and Notes

1. J. M. Cherry, C. Adler, C. Ball, S. A. Chervitz, S. S. Dwight, E. T. Hester, Y. Jia, G. Juvik, T. Roe, M. Schroeder, S. Weng, D. Botstein, SGD: *Saccharomyces* Genome Database. *Nucleic Acids Res.* **26**, 73–79 (1998). doi:10.1093/nar/26.1.73 Medline

2. A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver, Life with 6000 genes. *Science* **274**, 546–567 (1996). doi:10.1126/science.274.5287.546 Medline

3. F. J. Isaacs, P. A. Carr, H. H. Wang, M. J. Lajoie, B. Sterling, L. Kraal, A. C. Tolonen, T. A. Gianoulis, D. B. Goodman, N. B. Reppas, C. J. Emig, D. Bang, S. J. Hwang, M. C. Jewett, J. M. Jacobson, G. M. Church, Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333**, 348–353 (2011). doi:10.1126/science.1205822 Medline

4. J. Dymond, J. Boeke, The *Saccharomyces cerevisiae* SCRaMbLE system and genome minimization. *Bioeng. Bugs* **3**, 168–171 (2012). Medline

5. J. S. Dymond, S. M. Richardson, C. E. Coombes, T. Babatz, H. Muller, N. Annaluru, W. J. Blake, J. W. Schwerzmann, J. Dai, D. L. Lindstrom, A. C. Boeke, D. E. Gottschling, S. Chandrasegaran, J. S. Bader, J. D. Boeke, Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**, 471–476 (2011). doi:10.1038/nature10403 Medline

6. L. A. Mitchell, N. A. Phillips, A. Lafont, J. A. Martin, R. Cutting, J. D. Boeke, qPCRTag analysis—A high throughput, real time PCR assay for Sc2.0 genotyping. *J. Vis. Exp.* **99**, e52941 (2015). Medline

7. S. M. Richardson, P. W. Nunley, R. M. Yarrington, J. D. Boeke, J. S. Bader, GeneDesign 3.0 is an updated synthetic biology toolkit. *Nucleic Acids Res.* **38**, 2603–2606 (2010). doi:10.1093/nar/gkq143 Medline

8. N. Annaluru, H. Muller, L. A. Mitchell, S. Ramalingam, G. Stracquadanio, S. M. Richardson, J. S. Dymond, Z. Kuang, L. Z. Scheifele, E. M. Cooper, Y. Cai, K. Zeller, N. Agmon, J. S. Han, M. Hadjithomas, J. Tullman, K. Caravelli, K. Cirelli, Z. Guo, V. London, A. Yeluru, S. Murugan, K. Kandavelou, N. Agier, G. Fischer, K. Yang, J. A. Martin, M. Bilgel, P. Bohutski, K. M. Boulier, B. J. Capaldo, J. Chang, K. Charoen, W. J. Choi, P. Deng, J. E. DiCarlo, J. Doong, J. Dunn, J. I. Feinberg, C. Fernandez, C. E. Floria, D. Gladowski, P. Hadidi, I. Ishizuka, J. Jabbari, C. Y. L. Lau, P. A. Lee, S. Li, D. Lin, M. E. Linder, J. Ling, J. Liu, J. Liu, M. London, H. Ma, J. Mao, J. E. McDade, A. McMillan, A. M. Moore, W. C. Oh, Y. Ouyang, R. Patel, M. Paul, L. C. Paulsen, J. Qiu, A. Rhee, M. G. Rubashkin, I. Y. Soh, N. E. Sotuyo, V. Srinivas, A. Suarez, A. Wong, R. Wong, W. R. Xie, Y. Xu, A. T. Yu, R. Koszul, J. S. Bader, J. D. Boeke, S. Chandrasegaran, Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55–58 (2014). doi:10.1126/science.1249252 Medline

9. L. A. Mitchell, A. Wang, G. Stracquadanio, Z. Kuang, X. Wang, K. Yang, S. Richardson, J. A. Martin, Y. Zhao, R. Walker, Y. Luo, H. Dai, K. Dong, Z. Tang, Y. Yang, Y. Cai, A. Heguy, B. Ueberheide, D. Fenyö, J. Dai, J. S. Bader, J. D. Boeke, Synthesis, debugging,

and effects of synthetic chromosome consolidation: *synVI* and beyond. *Science* **355**, aaf4831 (2017).

10. Y. Shen, Y. Wang, T. Chen, F. Gao, J. Gong, D. Abramczyk, R. Walker, H. Zhao, S. Chen, W. Liu, Y. Luo, C. A. Müller, A. Paul-Dubois-Taine, B. Alver, G. Stracquadanio, L. A. Mitchell, Z. Luo, Y. Fan, B. Zhou, B. Wen, F. Tan, Y. Wang, J. Zi, Z. Xie, B. Li, K. Yang, S. M. Richardson, H. Jiang, C. E. French, C. A. Nieduszynski, R. Koszul, A. L. Marston, Y. Yuan, J. Wang, J. S. Bader, J. Dai, J. D. Boeke, X. Xu, Y. Cai, H. Yang, Deep functional analysis of *synII*, a 770-kilobase synthetic yeast chromosome. *Science* **355**, aaf4791 (2017).

11. X. Li, M.-Z. Ding, D. Liu, L. Zhang, B.-L. Liu, X.-L. Wu, F.-F. Li, X.-T. Dong, B. Jia, W.-Z. Zhang, G.-Z. Jiang, Y. Liu, X. Bai, T.-Q. Song, Y. Chen, S.-J. Zhou, R.-Y. Zhu, F. Gao, Z. Kuang, X. Wang, M. Shen, K. Yang, G. Stracquadanio, S. M. Richardson, Y. Lin, L. Wang, R. Walker, Y. Luo, P.-S. Ma, H. Yang, Y. Cai, J. Dai, J. S. Bader, J. D. Boeke, Y.-J. Yuan, Bug mapping and fitness testing of chemically synthesized chromosome *X*. *Science* **355**, aaf4706 (2017).

12. Z.-X. Xie, B.-Z. Li, L. A. Mitchell, Y. Wu, X. Qi, Z. Jin, B. Jia, X. Wang, B.-X. Zeng, H.-M. Liu, X.-L. Wu, Q. Feng, W.-Z. Zhang, W. Liu, M.-Z. Ding, X. Li, G.-R. Zhao, J.-J. Qiao, J.-S. Cheng, M. Zhao, Z. Kuang, X. Wang, J. A. Martin, G. Stracquadanio, K. Yang, X. Bai, J. Zhao, M.-L. Hu, Q.-H. Lin, W.-Q. Zhang, M.-H. Shen, S. Chen, W. Su, E.-X. Wang, R. Guo, F. Zhai, X.-J. Guo, H.-X. Du, J.-Q. Zhu, T.-Q. Song, J.-J. Dai, F.-F. Li, G.-Z. Jiang, S.-L. Han, S.-Y. Liu, Z.-C. Yu, X.-N. Yang, K. Chen, C. Hu, D.-S. Li, N. Jia, Y. Liu, L.-T. Wang, S. Wang, X.-T. Wei, M.-Q. Fu, L.-M. Qu, S.-Y. Xin, T. Liu, K.-R. Tian, X.-N. Li, J.-H. Zhang, L.-X. Song, J.-G. Liu, J.-F. Lv, H. Xu, R. Tao, Y. Wang, T.-T. Zhang, Y.-X. Deng, Y.-R. Wang, T. Li, G.-X. Ye, X.-R. Xu, Z.-B. Xia, W. Zhang, S.-L. Yang, Y.-L. Liu, W.-Q. Ding, Z.-N. Liu, J.-Q. Zhu, N.-Z. Liu, R. Walker, Y. Luo, Y. Wang, Y. Shen, H. Yang, Y. Cai, P.-S. Ma, C.-T. Zhang, J. S. Bader, J. D. Boeke, Y.-J. Yuan, "Perfect" designer chromosome *V* and behavior of a ring derivative. *Science* **355**, aaf4704 (2017).

13. W. Zhang, G. Zhao, Z. Luo, Y. Lin, L. Wang, Y. Guo, A. Wang, S. Jiang, Q. Jiang, J. Gong, Y. Wang, S. Hou, J. Huang, T. Li, Y. Qin, J. Dong, Q. Qin, J. Zhang, X. Zou, X. He, L. Zhao, Y. Xiao, M. Xu, E. Cheng, N. Huang, T. Zhou, Y. Shen, R. Walker, Y. Luo, Z. Kuang, L. A. Mitchell, K. Yang, S. M. Richardson, Y. Wu, B.-Z. Li, Y.-J. Yuan, H. Yang, J. Lin, G.-Q. Chen, Q. Wu, J. S. Bader, Y. Cai, J. D. Boeke, J. Dai, Engineering the ribosomal DNA in a megabase synthetic chromosome. *Science* **355**, aaf3981 (2017).

14. S. M. Richardson, B. S. Olson, J. S. Dymond, R. Burns, S. Chandrasegaran, J. D. Boeke, A. Shehu, J. S. Bader, Automated design of assemblable, modular, synthetic chromosomes. *Lect. Notes Comput. Sci.* **6068**, 280–289 (2010). doi:10.1007/978-3-642-14403-5_30

15. R. W. Barnett, H. Erfle, Rapid generation of DNA fragments by PCR amplification of crude, synthetic oligonucleotides. *Nucleic Acids Res.* **18**, 3094 (1990). doi:10.1093/nar/18.10.3094 Medline

16. R. B. Ciccarelli, P. Gunyuzlu, J. Huang, C. Scott, F. T. Oakes, Construction of synthetic genes using PCR after automated DNA synthesis of their entire top and bottom strands. *Nucleic Acids Res.* **19**, 6007–6013 (1991). doi:10.1093/nar/19.21.6007 Medline

17. H. Muller, N. Annaluru, J. W. Schwerzmann, S. M. Richardson, J. S. Dymond, E. M. Cooper, J. S. Bader, J. D. Boeke, S. Chandrasegaran, Assembling large DNA segments in yeast. *Methods Mol. Biol.* **852**, 133–150 (2012). doi:10.1007/978-1-61779-564-0_11 Medline

18. S. M. Richardson, S. J. Wheelan, R. M. Yarrington, J. D. Boeke, GeneDesign: Rapid, automated design of multikilobase synthetic genes. *Genome Res.* **16**, 550–556 (2006). doi:10.1101/gr.4431306 Medline

19. R. H. Hoess, A. Wierzbicki, K. Abremski, The role of the *loxP* spacer region in P1 site-specific recombination. *Nucleic Acids Res.* **14**, 2287–2300 (1986). doi:10.1093/nar/14.5.2287 Medline

20. Y. Shen, G. Stracquadanio, Y. Wang, K. Yang, L. A. Mitchell, Y. Xue, Y. Cai, T. Chen, J. S. Dymond, K. Kang, J. Gong, X. Zeng, Y. Zhang, Y. Li, Q. Feng, X. Xu, J. Wang, J. Wang, H. Yang, J. D. Boeke, J. S. Bader, SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* **26**, 36–49 (2016). doi:10.1101/gr.193433.115 Medline

21. M. J. Lajoie, S. Kosuri, J. A. Mosberg, C. J. Gregg, D. Zhang, G. M. Church, Probing the limits of genetic recoding in essential genes. *Science* **342**, 361–363 (2013). doi:10.1126/science.1241460 Medline

22. H. M. Salim, K. L. Ring, A. R. Cavalcanti, Patterns of codon usage in two ciliates that reassign the genetic code: *Tetrahymena thermophila* and *Paramecium tetraurelia*. *Protist* **159**, 283–298 (2008). doi:10.1016/j.protis.2007.11.003 Medline

23. A. Admire, L. Shanks, N. Danzl, M. Wang, U. Weier, W. Stevens, E. Hunt, T. Weinert, Cycles of chromosome instability are associated with a fragile site and are increased by defects in DNA replication and checkpoint controls in yeast. *Genes Dev.* **20**, 159–173 (2006). doi:10.1101/gad.1392506 Medline

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

25. C. Marck, 'DNA Strider': A 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res.* **16**, 1829–1836 (1988). doi:10.1093/nar/16.5.1829 Medline

26. D. G. Fisk, C. A. Ball, K. Dolinski, S. R. Engel, E. L. Hong, L. Issel-Tarver, K. Schwartz, A. Sethuraman, D. Botstein, J. M. Cherry; *Saccharomyces* Genome Database Project, *Saccharomyces cerevisiae* S288C genome annotation: A working hypothesis. *Yeast* **23**, 857–865 (2006). doi:10.1002/yea.1400 Medline

27. B. Berliner, in *Proceedings of the USENIX Winter 1990 Technical Conference*, Washington, DC, 22 to 26 January 1990 (USENIX Association, Berkeley, CA, 1990), pp. 341–352.

28. C. M. Pilato, B. Collins-Sussman, B. W. Fitzpatrick, *Version Control with Subversion* (O'Reilly Media, Sebastopol, CA, ed. 2, 2008).

29. S. Chacon, B. Straub, ProGit (SpringerLink Online service, ed. 2, 2014).

30. L. D. Stein, C. Mungall, S. Shu, M. Caudy, M. Mangone, A. Day, E. Nickerson, J. E. Stajich, T. W. Harris, A. Arva, S. Lewis, The generic genome browser: A building block for a model organism system database. *Genome Res.* **12**, 1599–1610 (2002). doi:10.1101/gr.403602 Medline

31. M. E. Skinner, A. V. Uzilov, L. D. Stein, C. J. Mungall, I. H. Holmes, JBrowse: A next-generation genome browser. *Genome Res.* **19**, 1630–1638 (2009). doi:10.1101/gr.094607.109 Medline

32. A. Hill, K. Bloom, Genetic manipulation of centromere function. *Mol. Cell. Biol.* **7**, 2397–2405 (1987). doi:10.1128/MCB.7.7.2397 Medline

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

34. K. J. Verstrepen, A. Jansen, F. Lewitter, G. R. Fink, Intragenic tandem repeats generate functional variability. *Nat. Genet.* **37**, 986–990 (2005). doi:10.1038/ng1618 Medline

35. J. Parenteau, M. Durand, G. Morin, J. Gagnon, J.-F. Lucier, R. J. Wellinger, B. Chabot, S. A. Elela, Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. *Cell* **147**, 320–331 (2011). doi:10.1016/j.cell.2011.08.044 Medline

36. J. Parenteau, M. Durand, S. Véronneau, A.-A. Lacombe, G. Morin, V. Guérin, B. Cecez, J. Gervais-Bird, C.-S. Koh, D. Brunelle, R. J. Wellinger, B. Chabot, S. Abou Elela, 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 Medline

37. R. S. Sikorski, P. Hieter, A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989). Medline

38. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009). doi:10.1038/nmeth.1318 Medline

39. J. E. Stajich, An introduction to BioPerl. *Methods Mol. Biol.* **406**, 535–548 (2007). Medline

40. A. Sliva, H. Yang, J. D. Boeke, D. J. Mathews, Freedom and responsibility in synthetic genomics: The Synthetic Yeast Project. *Genetics* **200**, 1021–1028 (2015). doi:10.1534/genetics.115.176370 Medline

41. P. M. Sharp, T. M. Tuohy, K. R. Mosurski, Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res.* **14**, 5125–5143 (1986). doi:10.1093/nar/14.13.5125 Medline

42. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, 421 (2009). doi:10.1186/1471-2105-10-421 Medline

43. S. Pechmann, J. Frydman, Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* **20**, 237–243 (2013). doi:10.1038/nsmb.2466 Medline

44. T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, Y. Pilpel, An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344–354 (2010). doi:10.1016/j.cell.2010.03.031 Medline

45. L. A. Mitchell, J. D. Boeke, Circular permutation of a synthetic eukaryotic chromosome with the telomerator. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17003–17010 (2014). doi:10.1073/pnas.1414399111 Medline