

Published in final edited form as:

*Nature*. ; 477(7365): 471–476. doi:10.1038/nature10403.

## Synthetic chromosome arms function in yeast and generate phenotypic diversity by design

Jessica S. Dymond<sup>1</sup>, Sarah M. Richardson<sup>1,2</sup>, Candice E. Coombes<sup>1</sup>, Timothy Babatz<sup>1</sup>, Héloïse Müller<sup>4</sup>, Narayana Annaluru<sup>4</sup>, William J. Blake<sup>5,6</sup>, Joy Wu Schwerzmann<sup>4,7</sup>, Junbiao Dai<sup>1,8</sup>, Derek L. Lindstrom<sup>9,10</sup>, Annabel C. Boeke<sup>1,11</sup>, Daniel E. Gottschling<sup>9</sup>, Srinivasan Chandrasegaran<sup>4</sup>, Joel S. Bader<sup>1,3</sup>, and Jef D. Boeke<sup>1</sup>

<sup>1</sup>High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>2</sup>McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>3</sup>Whiting School of Engineering, Johns Hopkins University, Baltimore, MD

<sup>4</sup>Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

<sup>5</sup>Codon Devices, Cambridge, MA

<sup>9</sup>Fred Hutchinson Cancer Research Center, Seattle, WA

Recent advances in DNA synthesis technology have enabled the construction of novel genetic pathways and genomic elements, furthering our understanding of system-level phenomena<sup>1–7</sup>. The ability to synthesize large segments of DNA allows one to engineer pathways and genomes according to arbitrary sets of design principles. Here we describe a synthetic yeast genome project, Sc2.0, and the first partially synthetic eukaryotic chromosomes, *Saccharomyces cerevisiae* chromosome *synIXR*, and *semi-syn VII*. We defined design principles for a synthetic genome predicted to result in I) (near) wild-type phenotype and fitness II) a genome lacking destabilizing elements such as tRNA genes or transposons<sup>8,9</sup>, and III) genetic flexibility to facilitate future studies (Box 1). The synthetic genome features multiple systemic modifications complying with the design principles, including an inducible evolution system, SCRaMbLE (Synthetic Chromosome Rearrangement and Modification by LoxP-mediated Evolution). We show the utility of SCRaMbLE as a novel method of combinatorial mutagenesis, capable of generating complex genotypes and a broad variety of phenotypes. When complete, the fully synthetic yeast genome will allow massive restructuring of the yeast genome, and may open the door

Correspondence and requests for materials should be addressed to J.D.B. (jboeke@jhmi.edu).

<sup>6</sup>GreenLight Biosciences, Medford, MA

<sup>7</sup>Battelle Memorial Institute, Atlanta, GA

<sup>8</sup>Tsinghua University, Beijing, China

<sup>10</sup>Agilent Laboratories, Santa Clara, CA

<sup>11</sup>Bowdoin College, Brunswick, ME

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Author Contributions** J.S.D., S.M.R., S.C., J.S.B., and J.D.B. designed experiments. J.S.D., S.M.R., C.E.C., T.B., H.M., N.A., J.W.S., J.D., and A.C.B. performed experiments. W.J.B. built the *synIXR* chromosome. D.L.L. and D.E.G. generated the integrated CRE-EBD cassette. J.S.D., S.M.R., J.S.B., and J.D.B. analyzed data and wrote the manuscript.

**Author information** *SynIXR* and *semi-syn VII* sequences were deposited at GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>; *synIXR*: JN020955; *semi-syn VII*: JN020956). Microarray data were submitted to GEO (<http://www.ncbi.nlm.nih.gov/geo/>).

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

The authors declare no competing financial interests.

to a new type of combinatorial genetics based entirely on variations in gene content and copy number.

## Design and synthesis

We designed the right arm of chromosome *IX* (*IXR*) according to the three principles outlined above and in Box 1. *IXR* is the smallest chromosome arm in the genome and features multiple genomic elements of interest (Fig. 1a), making this chromosome arm suitable for a pilot study. The designed sequence, *synIXR*, is based on a native *IXR* sequence extending from open reading frame (ORF) *YIL002W* through the centromere and the remainder of chromosome *IXR*, an 89,299 bp sequence (native *IXR* position 350585-438993<sup>10</sup>). In accordance with design principle II, a tRNA gene, a Ty1 LTR, and telomeric sequences were removed. The final *synIXR* sequence, 91010 bp, is slightly longer than the native sequence due to the inclusion of 43 loxPsym sites, replacing 20.3% of the native chromosome. A 30 kb telomeric segment of the left arm of chromosome *VI* (*semi-syn VII*) was similarly designed (Fig. 1b, Supplement “*Syn VII* design and incorporation”), and replaced 15.7% of the native chromosome. The two synthetic segments comprise 17% of original sequence lengths that were 1) changed by base substitution, 2) deleted, or 3) inserted (Table S1); sequences were submitted to GenBank (Supplement sequence files “*SynIXR*” and “*Semi-syn VII*”; *synIXR*: JN020955; *semi-syn VII*: JN020956).

We systematically introduced two systematic sets of changes *in silico* using the genome editing suite BioStudio (described elsewhere): introduction of TAG/TAA stop codon swaps and PCRTags sequences (see Supplement “Sequence design/editing algorithms”). In recognition of design principle III, the elimination of the TAG stop codon by recoding to TAA frees a codon for future genetic code expansion (e.g. by adding a 21st unnatural amino acid<sup>11, 12</sup>) and could serve as a future mechanism of reproductive isolation and control. PCRTags are short pairs of recoded sequences, unique to either the wild-type or synthetic genome, and serve as convenient, low cost and closely spaced genetic markers for verifying introduction of the synthetic sequence and removal of the native sequence. The inclusion of PCRTags allows design of PCR primers specific to either the native or synthetic sequence to rapidly evaluate the presence of synthetic and absence of native sequences, critical for evaluating incorporation of synthetic DNA (see below and Supplement, “*Syn VII* design and incorporation”). PCRTags, designed *in silico*, were tested in triplicate to verify specificity (Fig. S1, Tables S2 and S3).

LoxPsym sequences are nondirectional loxP sites capable of recombining in either orientation<sup>13</sup>, and theoretically produce inversions or deletions with equal probability. Under design principle III, these sites form the substrate for the inducible SCRaMbLE system and are intended to generate combinatorial diversity. We inserted loxPsym sites 3 bp after each nonessential gene's stop codon and at major landmarks, such as sites of LTR and tRNA deletions, flanking *CEN9*, and adjacent to telomeres (Fig. 1; Supplement “LoxPsym site insertion”). LoxPsym sites inserted at equivalent positions genome-wide will allow formation of many structurally distinct genomes.

## Synthetic chromosome introduction and construction of “arm swap” strains

The *synIXR* chromosome, cloned in a circular BAC vector, includes all sequences needed for propagation in yeast and bacteria (Fig. 1a). We introduced *synIXR* into a diploid strain by transformation; typically, about 10-15% of the *synIXR* transformants obtained were positive for all PCRTags pairs tested (Fig. 2d). We choose one such transformant, strain A (Fig. 2a) and truncated one native *IXR* homolog (*IX R*) by transforming with a suitably

designed linear DNA fragment<sup>14</sup> introducing a selectable marker (*URA3*) and a telomere seed sequence, generating strain C (Fig. 2b). Chromosome truncation was confirmed by pulsed field gel electrophoresis analysis (Fig. 2c), and strain C was sporulated to generate haploids carrying *synIXR* and *IX<sup>R</sup>*. We observed more spore lethality than in control crosses, presumably due to segregation of *synIXR* away from *IX<sup>R</sup>*; cells bearing only *synIXR* or only *IX<sup>R</sup>* would lack many essential genes and not survive. PCRTags analysis of 14 *synIXR* candidate “arm swap” strains revealed 10 haploids with all synthetic PCRTags and no native PCRTags present (Fig. 2d, Fig. S2). The remaining four strains carried BACs with “patchworks” of synthetic and native sequences suggestive of meiotic gene conversion events (Fig. S2). Sanger sequencing and structural analyses (Table S4; Fig. S3; Supplement, “DNA sequence analysis”) of recovered *synIXR* BACs revealed no mutations had occurred in the synthetic chromosome. Thus the synthetic sequence is replicated faithfully.

Whereas *synIXR* was incorporated in a circular form, we used an alternate strategy to integrate the *semi-synVIL* chromosome fragment into native chromosome *VI* (Fig. S4): a linear synthetic fragment marked with *LEU2* was transformed into a *YFL054C::kanMX* strain. Approximately 13% (75/586) had the *Leu<sup>+</sup>* G418<sup>S</sup> phenotype expected for the desired integrant. PCRTags analysis showed that 10 of 12 such strains contained only synthetic PCRTags, as expected for full replacement (Fig. S5).

## Phenotypic and genetic analysis of synthetic haploids

Design principle I prioritizes a wild-type phenotype and high fitness level despite the incorporated modifications. *SynIXR* has a designed sequence alteration approximately every 500 bp, 2.64% of total sequence is altered, and it carries 43 loxPsym sites. To check for negative effects of these modifications on fitness, we 1) examined colony size and morphology under various conditions, and 2) performed transcript profiling. We examined colony size and morphology of *synIXR* swap strains under six distinct growth conditions. It was impossible to distinguish swap strains from the wild-type (BY4741) under these conditions, suggesting that any fitness defect attributable to *synIXR* is modest; fitness tests on *semi-synVIL* gave similar results (Fig. S6).

Synonymous substitutions, loxPsym site introduction, or other changes might change gene expression. We performed transcript profiling experiments on swap strains *synIXR-1D*, *synIXR-6B*, and *synIXR-22D* (Supplement, “Transcriptional Profiling”); these studies revealed interesting but predictable trends (Fig. 3). As expected, genes present in two copies (*YIL001W* and *YIL002C*, present on both *synIXR* and *IX<sup>R</sup>*) were approximately doubled in transcript abundance. Most genes showed no significant expression change, although a few showed modest decreases; however, the subtelomeric genes *YIR039C* and *YIR042C* showed increased expression. We speculate that in the circular synthetic chromosome these are released from telomeric silencing, resulting in their overexpression. Overall *synIXR* genes show relatively normal expression, suggesting that loxPsym sites and PCRTags minimally effect expression. Similarly, no significant changes were observed by RNA blotting (Fig. S7a). To detect possible compensatory transcriptome changes, we profiled transcripts genome-wide. Except for trivial differences attributable to slightly different selectable marker configurations in the strains, there were no consistent statistically significant differences seen outside of *IXR* itself (Fig. S7b). Thus, modifications present in *synIXR* and *semi-synVIL* do not produce major fitness effects, nor compensatory transcriptomic alterations.

## Conditional Genome Instability

The design principles dictate that SCRaMbLE be available for use on demand, yet lie dormant until intentional Cre recombinase induction, at which point generation of genetic

diversity is desirable. To complete the SCRaMbLE toolkit, we incorporated an engineered Cre recombinase fused to the murine estrogen binding domain (EBD). This recently described Cre-EBD variant<sup>15</sup> is estradiol inducible, has low basal activity, and is controlled by daughter cell-specific promoter *SCW11* (Fig. S8). *pSCW11*-Cre-EBD should produce a pulse of recombinase activity once and only once in each cell's lifetime, and depend on estradiol exposure. The uninduced integrated construct is well tolerated even in swap strains, which, with 43 loxPsym sites, is expected to be Cre-hypersensitive. Upon estradiol addition, rearrangements are induced at the loxPsym sites and viability dropped by 100-fold in *synIXR* strains (Fig. 4a, Fig. S9). This loss of viability likely results from loss of *synIXR* essential genes. In contrast, viability in *semi-synVIL*, which lacks essential genes, is not affected by Cre induction (Fig. 1b, Fig. S9d).

*Semi-synVIL* contains just five loxPsym sites, including one immediately adjacent to the telomeric TG<sub>1-3</sub> repeats (Fig. 1b). This simple configuration allows comprehensive PCR-based mapping of rearrangements of four of the loxPsym sites in SCRaMbLEd strains. A SCRaMbLEd *semi-synVIL* population was analyzed for most of the possible rearranged configurations by PCR, revealing a large variety of deletions and inversions (Fig. 4a); most predicted rearrangements were readily detected.

The symmetry of loxPsym sites allows alignment in two orientations, theoretically giving rise to deletions and inversions with equal frequency. *SynIXR* contains 43 loxPsym sites, allowing over 3600 potential pairwise interactions between *synIXR* loxPsym sites. We reasoned that SCRaMbLEd *synIXR* clones should display high phenotypic diversity. Indeed, SCRaMbLEd swap strains show more growth rate heterogeneity than wild-type controls (Fig. 4c, Fig. S10). These SCRaMbLE clones show many different phenotypes (Supplement, "SCRaMbLE Analysis", Fig. S11). In summary, SCRaMbLE is sufficient to generate significant genetic heterogeneity and complex phenotypes.

To further characterize the utility of SCRaMbLE, we performed a mutagenesis study. *SynIXR* encodes both *MET28* and *LYS1*, genes required for amino acid biosynthesis<sup>16, 17</sup>. Null mutants result in auxotrophy, and can be easily detected by replica-plating. We introduced episomal Cre-EBD (*pSCW11-Cre-EBD-URA3MX* cloned in a CEN plasmid) into strain C previously made *LYS2*<sup>+</sup> (strain "D", yJS587) and performed SCRaMbLE. We screened 20,242 colonies and 3% (604/20,242) were candidate *lys1* and/or *met28* auxotrophs. Of 360 candidates tested more rigorously, 295 (81.9%) were confirmed: we found 212 Lys<sup>-</sup> auxotrophs (1.37%), 66 Met<sup>-</sup> auxotrophs (0.43%), and interestingly, 17 Lys<sup>-</sup>Met<sup>-</sup> double auxotrophs (0.11%). PCRTag profiles of 24 Met<sup>-</sup> auxotrophs, 35 Lys<sup>-</sup> auxotrophs, and 7 double auxotrophs (Fig. 4d) showed that all Met<sup>-</sup> auxotrophs had deletions in the loxPsym-flanked segment containing *MET28* and *YAP5*, whereas all Lys<sup>-</sup> auxotrophs had deletions in the loxPsym-flanked segment containing *LYS1*. The deletion profiles of many SCRaMbLEd auxotrophs were highly variable, often with more than one segment missing.

To formally confirm the observed SCRaMbLE phenotypes resulted solely from deletions in *synIXR*, we recovered the *synIXR* chromosome from two Met<sup>-</sup> auxotrophs into *E. coli*, then introduced them to a clean genetic background. In both cases the auxotrophic phenotype was associated with the presence of the SCRaMbLEd chromosomes (Fig. S12; Supplement, "*MET28* and *LYS1* SCRaMbLE Mutagenesis"). Thus the SCRaMbLE system is a highly effective method of mutagenesis, giving rise to mutants with different genetic backgrounds, and generating a wide variety of double mutants.

## Perspectives

We have shown there does not appear to be any significant theoretical impediment to extending the design strategy outlined here to the entire yeast genome, apart from the challenge of 12 Mb of DNA synthesis. Whether or not fitness defects will accumulate as design and synthesis are scaled up further remains to be seen; however, the overall high fitness of the swap strains described here validates the design strategy. Furthermore, the iterative bottom-up approach used will allow identification of potential “problem regions” in synthetic sequences as synthesis moves forward. If a given swap experiment results in only transformants with reduced fitness (or no transformants are obtainable), the underlying defect can be mapped by introducing sub-segments, facilitated by strategic placement of unique restriction sites throughout synthetic chromosome arms. Also, since a subset of transformants consist of patchworks of native and synthetic sequence (Fig. S2, S5), analysis of such strains can in principle be used to rapidly map phenotypic defects. The stability and sequence fidelity of large circular chromosomes observed here and elsewhere<sup>5-7</sup> bode well for use of yeast as a host platform for synthetic biology.

SCRaMbLE may become a useful general strategy for analyzing genome structure, content, and function. One important feature of SCRaMbLE is its potential to be customized; expression of different Cre-EBD variants from various promoters at distinct inducer (estradiol) levels should produce distinct SCRaMbLE dynamics. Use of weaker promoters than *pSCW11*, promoters expressed at different cell cycle phases, performing SCRaMbLE in diploids, and lowering the inducer concentration should all contribute to decreased lethality of SCRaMbLE strains, an important consideration as additional segments of the genome are replaced with synthetic counterparts and the proportion of essential genes that can be lost by SCRaMbLEing increases. As shown here, SCRaMbLE mutagenesis is efficient and generates mutants with a wide variety of different genetic backgrounds. It is possible that different combinations of gene deletions will give rise to a variety of subtly different phenotypes that can be rapidly mapped by PCRTAG analysis; more extensive analysis by deep sequencing will reveal changes in both genome structure and content. As the synthetic yeast genome grows, opportunities for genome rearrangement will increase exponentially. In principle, changes in chromosome number, ploidy, content, and structure are all possible, increasing the utility of the SCRaMbLE system. For example, there may be many different routes to a minimal genome, and exploring all of them by a hit or miss predictive approach is impractical and unlikely to yield comprehensive results. Using SCRaMbLE, many independent routes of genome minimization can be explored at one time, under manifold environmental conditions, for example by growing yeast cells long-term either in serially transferred batch cultures, or in a chemostat or turbidostat under conditions where Cre is minimally active. Such an approach may also lead to derivatives that are more fit than the parent, e.g. by gene duplication events facilitated by the Cre-EBD/loxPsym system employed here.

## Methods Summary

### DNA preparation

BAC DNA was prepared using the Qiagen Plasmid Midi kit or alkaline lysis<sup>18</sup>. The following protocol modifications were made: cells were diluted 1:100 from an overnight culture into 50 ml grown in LB plus 50 µg/ml carbenicillin, and grown at 30°C for 14-16 hours. Qiagen purified DNA was treated with 60 µg/ml proteinase K at 37°C overnight, then phenol/chloroform extracted. DNAs prepared without a column were phenol/chloroform extracted, and then RNase treated immediately prior to use.



Yeast genomic DNA for use in PCRTag analysis was prepared by standard methods<sup>19</sup>. DNA preparation for recovery of the *synIXR* BAC into bacteria was as previously reported<sup>20</sup>.

### PCR conditions

PCRTags were amplified using Taq polymerase (New England Biolabs). Template concentrations were 1 ng/μl for genomic DNA and 10 pg/μl for purified BAC DNA. The following program was used: 94°C 3 min; 30 cycles of 94°C 30 sec, 65°C 30 sec, 72°C 30 sec; 72°C 3 min.

### RNA analysis

Total RNA was isolated by hot acid phenol extraction. Microarray hybridization and data analysis were performed at the Johns Hopkins Microarray Core Facility ([www.microarray.jhmi.edu](http://www.microarray.jhmi.edu)). Dubious ORFs and pseudogenes were omitted from *synIXR* transcript analysis.

### Pulsed field gels

DNAs were prepared as described elsewhere<sup>21</sup>. Identity of the chromosomes was inferred from the known molecular karyotype of WT (BY4743) and from lambda ladders run on the same gel.

## Methods

### Yeast strains, transformation and tetrad analysis

Strains ABY7 and ABY8 were derived from strain BY4743<sup>22</sup>; ABY7 (MATa) and ABY7 (MAT<sup>-</sup>) otherwise share the genotype: *his3 1 leu2 0 ura3 0 lys2 0 met15 0 yil001: :URA3 yir039: :kanMX*.

BY4743 spheroplasts were transformed with *synIXR*. The strain *YFL054C: :kanMX*<sup>23</sup> was transformed with *synVIL* restriction fragments by standard lithium acetate transformation.

Strains *synIXR*-1D and others were backcrossed to strains ABY7 and ABY8; resultant diploids were sporulated and genotyped to identify *synIXR* segregants.

### Phenotypic screening

Single colonies were picked into 96-well plates and grown for 48 h in YPD at 30°C. (SCRaMbLE strains were grown 72 h in YPD at 30°C, diluted 1:10 and grown 4 h prior to plating.) Ten-fold dilutions were spotted on various agar medium types/selective conditions in OmniTrays (NUNC), as described<sup>24</sup>. Most cells were grown 72 hours (except YPGE plates, grown for 108 h), scored for growth, and photographed.

### Yeast growth and media

Unless otherwise indicated, all experiments were performed at 30°C. YPGE was supplemented with 2% ethanol and 2% glycerol. Concentrations of drugs were as follows: hydroxyurea, 0.2M; methylmethane sulfonate, 0.05%; 6-azauracil, 100 μg/ml; benomyl, 15 μg/ml; hydrogen peroxide, 1 mM; cycloheximide, 10 μg/ml. Cycloheximide and hydrogen peroxide resistance were assayed by growing cells in treated medium for two hours, then plating on YPD. Other phenotypes were assayed by growing cells to mid-log phase in rich media then spotting ten-fold dilutions on selective media.

## Colony size measurements

Cells were plated at various dilutions so that similar numbers of colonies were observed on control and experimental (estradiol-treated) plates. Colony size was measured using ImageJ software<sup>25</sup>, and normalized against the total number of colonies on each plate. Sample sizes for data presented in Fig. 4c are as follows: WT, n = 488 colonies; WT+C+E, n = 486; 2.2.1D, n = 395; 2.2.1D +C, n = 251; 2.2.1D+E, n = 416; 2.2.1D +C+E, n = 394.

## SynIXR BAC Sequence Analysis

The original *synIXR* BAC was sequenced by the manufacturer, Codon Devices<sup>26</sup>. *SynIXR* BACs were recovered into bacteria and sequenced by Agencourt (Beckman Coulter Genomics), using sequencing primers listed in Table S5. Repetitive sequences, including the highly internally repetitive *MUC1* open reading frame, were PCR-amplified prior to sequencing where necessary.

## Pulsed Field Gels

Samples were run on a 1.0% agarose gel in 0.5× TBE pH 8.0 for 20 hours at 14°C on a CHEF apparatus. The voltage was 3.5 V/cm, at an angle of 120° and 60-120 second switch time ramped over 20 hours.

*NotI* (Promega) digests were performed on whole chromosomes embedded in agarose plugs. Agarose plugs were removed from the 0.5 M EDTA storage buffer, washed with 0.05 M EDTA for one hour at room temperature, and then washed with 0.1× followed by 1× restriction enzyme buffer under the same conditions.

## Northern and Southern Blots Probe Preparation

Probes were prepared using the Prime-It II kit (Stratagene), and hybridized using Ultrahyb hybridization solution (Ambion) according to the manufacturer's instructions.

## SCRaMBLE

Cre activity was induced by exposure to 1 μM -estradiol (Sigma-Aldrich) in rich media for either 48 hours (integrated Cre) or 4 hours (episomal Cre) except where indicated otherwise. PCRTag analysis of Met<sup>-</sup> and Lys<sup>-</sup> auxotrophs was performed with a non-redundant array using one primer pair per loxPsym-flanked segment.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank George Church for suggesting the global substitution of TAG codons with TAA codons, Carla Connelly for sharing technical expertise and Victor Huang for generating a sequence visualizer. We are grateful to Brendan Cormack, Geraldine Seydoux and Jeremy Nathans for offering helpful advice, Patrick Cai and Jean Peccoud for suggesting methods to validate the sequence data, and Ed Louis for providing expert advice on telomeres. Supported by National Science Foundation grant MCB0718846 to J.D.B., J.S.B. and S.C., by a grant from Microsoft to J.S.B. and J.D.B., by Department of Energy Fellowship DE-FG02097ER25308 to S.M.R., National Institutes of Health grant AG023779 to D.E.G., and by a fellowship from Fondation pour la Recherche Médicale to H.M.

## References

1. Han JS, Boeke JD. A highly active synthetic mammalian retrotransposon. *Nature*. 2004; 429:314–318. [PubMed: 15152256]

2. Richardson SM, Wheelan SJ, Yarrington RM, Boeke JD. GeneDesign: rapid, automated design of multikilobase synthetic genes. *Genome Res.* 2006; 16:550–556. [PubMed: 16481661]
3. Chan LY, Kosuri S, Endy D. Refactoring bacteriophage T7. *Mol Syst Biol.* 2005; 1:2005.0018. [PubMed: 16729053]
4. Stricker J, et al. A fast, robust and tunable synthetic gene oscillator. *Nature.* 2008; 456:516–519. [PubMed: 18971928]
5. Gibson DG, 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.* 2008; 105:20404–20409. [PubMed: 19073939]
6. Gibson DG, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science.* 2010; 329:52–56. [PubMed: 20488990]
7. Lartigue C, et al. Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science.* 2009; 325:1693–1696. [PubMed: 19696314]
8. Admire A, et al. 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.* 2006; 20:159–173. [PubMed: 16384935]
9. Ji H, et al. Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. *Cell.* 1993; 73:1007–1018. [PubMed: 8388781]
10. Churcher C, et al. The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX. *Nature.* 1997; 387:84–87. [PubMed: 9169870]
11. Park H, et al. Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science.* 2011 In press.
12. Isaacs FJ, et al. Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. *Science.* 2011 In press.
13. Hoess RH, Wierzbicki A, Abremski K. The role of the loxP spacer region in P1 site-specific recombination. *Nucleic Acids Res.* 1986; 14:2287–2300. [PubMed: 3457367]
14. Vollrath D, Davis RW, Connelly C, Hieter P. Physical mapping of large DNA by chromosome fragmentation. *Proc Natl Acad Sci U S A.* 1988; 85:6027–6031. [PubMed: 3045811]
15. Lindstrom DL, Gottschling DE. The mother enrichment program: a genetic system for facile replicative life span analysis in *Saccharomyces cerevisiae*. *Genetics.* 2009; 183:413–22. ISI–13SI. [PubMed: 19652178]
16. Kuras L, Cherest H, Surdin-Kerjan Y, Thomas D. A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism. *EMBO J.* 1996; 15:2519–2529. [PubMed: 8665859]
17. Ogawa H, Fujioka M. Purification and characterization of saccharopine dehydrogenase from baker's yeast. *J Biol Chem.* 1978; 253:3666–3670. [PubMed: 418069]
18. Sambrook J, Russell DW. Isolation of BAC DNA from Small-scale Cultures. *Cold Spring Harb Protoc.* 2006; 2006.pdb.prot4006.
19. Hoffman CS. Preparation of yeast DNA. *Curr Protoc Mol Biol.* 2001; Chapter 13 Unit13.11.
20. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. Ty elements transpose through an RNA intermediate. *Cell.* 1985; 40:491–500. [PubMed: 2982495]
21. Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell.* 1984; 37:67–75. [PubMed: 6373014]
22. Brachmann CB, et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast.* 1998; 14:115–132. [PubMed: 9483801]
23. Winzeler EA, et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science.* 1999; 285:901–906. [PubMed: 10436161]
24. Hampsey M. A review of phenotypes in *Saccharomyces cerevisiae*. *Yeast.* 1997; 13:1099–1133. [PubMed: 9301019]
25. Abramoff MD, Magelhaes PJ, Ram SJ. Image Processing with ImageJ. *Biophotonics International.* 2004; 11:36–42.



26. Blake WJ, et al. Pairwise selection assembly for sequence-independent construction of long-length DNA. *Nucleic Acids Res.* 2010; 38:2594–2602. [PubMed: 20194119]

**Box 1****Modifications in synthetic sequence**

Elements removed: Retrotransposons (presumed nonessential)

Subtelomeric repeats (presumed nonessential)

Introns (none on *IXR*; presumed nonessential)

Elements relocated to extrachromosomal array: tRNA genes (genome-destabilizing<sup>7,8</sup>)

Element replaced: TAG Stop codons replaced by TAA (future code manipulation)

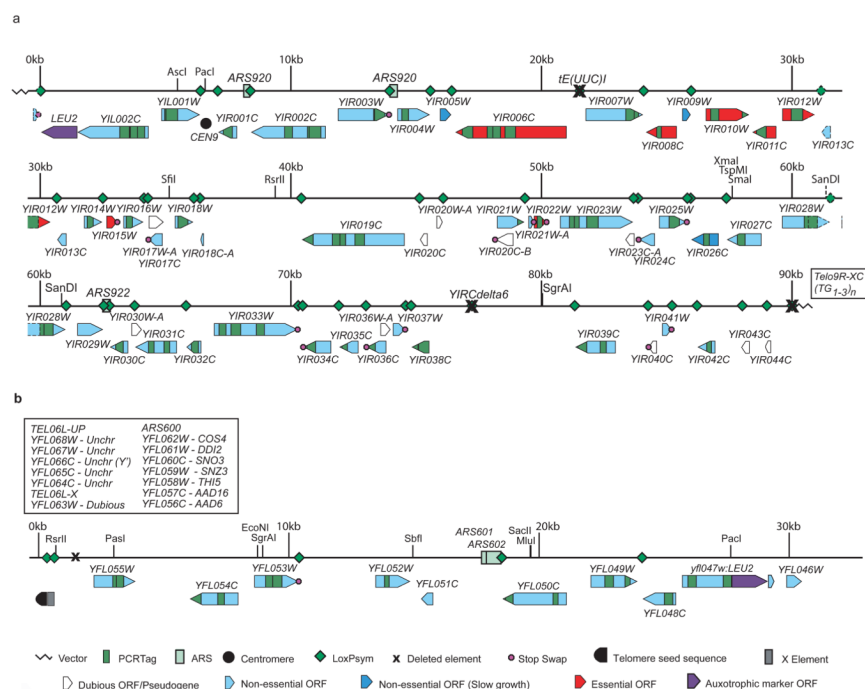
Individual synonymous codons (site engineering)

Strings of synonymous codons (PCRTags/markers)

Element introduced: LoxPSym sites (conditional genome instability; SCRaMbLE)

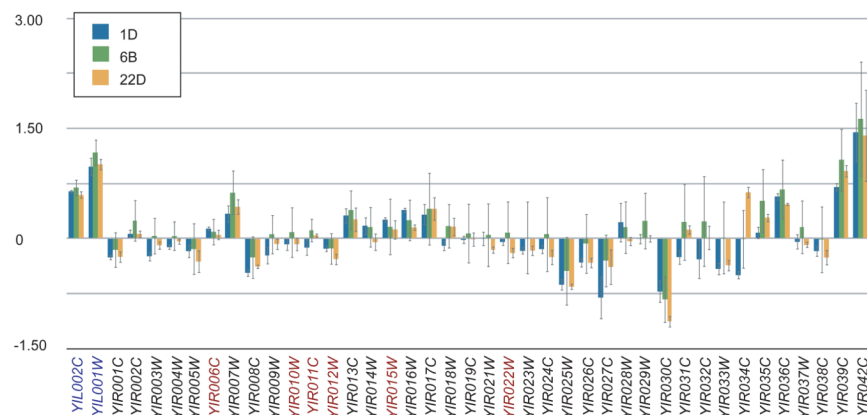
Not changed: Gene order

Noncoding regions (except as noted)



**Figure 1.** Map of a) *synIXR* and b) *semi-syn VIL*. Boxed text indicates elements deleted in the synthetic chromosomes. “Vector” in part a is circular. Vertical green bars inside ORFs indicate PCRTag amplicons; only sequences at the outside edges of these are recorded.

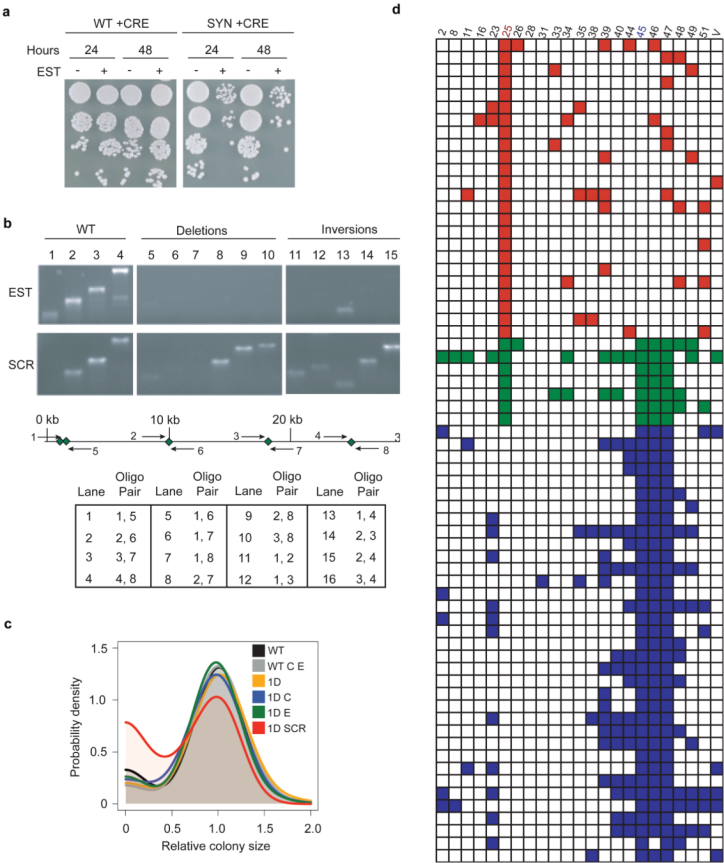
Strain construction and verification. a) Generation of *synIXR* haploids. (I) *synIXR* BAC (L) was transformed into BY4743 to generate “A” (II). (III) One copy of native *IXR* in “A” was replaced with a *URA3*-telomere seed cassette (U), generating *IX R* in “B”. (IV) “B” was sporulated to produce haploids. Circle, centromere; small square, *LEU2*. b) *IX R* structure. c) Top: Electrophoretic karyotype. \*, native *IXR*; \*\*, *IX R*. Bottom: Blot of *NotI*-digest. Linearized *synIXR* migrates as discrete ~100 kb band. Probe (*YIL002C*) detects all chromosome *IX* isoforms. WT, BY4743. d) PCRTag analysis. SYN, *synIXR* BAC; V, vector amplicon.



**Figure 3.**

Transcript profiling of wild-type and *synIXR* strains. Transcript profiling of *synIXR*-1D, -6B, and -22D. Log<sub>2</sub> ratio of RNA abundance relative to wild type (BY4741 or BY4742) is shown. ORFs labeled in blue exist in two copies. Essential genes are labeled in red. Error bars, standard deviation.





**Figure 4.** SCRaMbLE rearranges genomes. a) Cre induction decreased fitness in *synIXR* strain (SYN) but not wild type (WT; BY4741). EST, estradiol; Hours, exposure time. b) PCR analysis of *semi-synVIL* SCRaMbLE. Map, primer positions. Amplicon 13 is spurious (wrong size). C, Cre; E, Estradiol. c) Shifted colony size distribution in survivors. d) PCRTag analysis of Met<sup>-</sup> (red), Lys<sup>-</sup> (blue), and Met<sup>-</sup> Lys<sup>-</sup> (green) auxotrophs using PCRTags. PCRTag pair (top; in Table S2). Each row represents one clone. *MET28*, pair 25; *LYS1*, pair 45. Shaded boxes indicate presumed deletions. a-c, Integrated Cre-EBD; d, Episomal Cre-EBD.