

Cite this: *Chem. Sci.*, 2021, 12, 5381Received 19th December 2020  
Accepted 2nd March 2021

DOI: 10.1039/d0sc06924h

rsc.li/chemical-science

## Debugging: putting the synthetic yeast chromosome to work

Ze-Xiong Xie,<sup>†</sup> Jianting Zhou,<sup>†</sup> Juan Fu<sup>†</sup> and Ying-Jin Yuan<sup>ID\*</sup>

Synthetic genomics aims to *de novo* synthesize a functional genome redesigned from natural sequences with custom features. Designed genomes provide new toolkits for better understanding organisms, evolution and the construction of cellular factories. Currently maintaining the fitness of cells with synthetic genomes is particularly challenging as defective designs and unanticipated assembly errors frequently occur. Mapping and correcting bugs that arise during the synthetic process are imperative for the successful construction of a synthetic genome that can sustain a desired cellular function. Here, we review recently developed methods used to map and fix various bugs which arise during yeast genome synthesis with the hope of providing guidance for putting the synthetic yeast chromosome to work.

### Introduction

The discovery of the chemical synthesis of urea and secretin opened the door to new organic chemicals and protein synthesis.<sup>1</sup> The discovery of the DNA double helix structure along with more recent advances in DNA sequencing provided the impetus for the synthesis of genomes.<sup>2–4</sup> Synthetic biologists are no longer satisfied with just copying a natural genome, and are ambitious to create new versions of genomes.<sup>5–19</sup> Computer-aided simulation allows the redesign of genomes with specific functions and, in compliance with the most fundamental principle of genomic design, maintaining cell viability,<sup>7,11,12,20</sup> custom genetic features can be introduced to increase the genome flexibility. For example, recoding, the introduction of recombination sites and watermark sequences,<sup>9</sup> and the deletion of repetitive sequences and unstable elements may all be achieved.<sup>12,20</sup> The newly designed genome sequence is hierarchically divided into oligonucleotides,<sup>7,9,21</sup> and subsequently assembled into “short”,<sup>22,23</sup> “medium”,<sup>24–26</sup> and “long”<sup>13,20</sup> DNA segments both *in vivo* and *in vitro*. Finally, the chemically synthesized DNA is transplanted into a bacterial or yeast cell replacing the native genetic material.<sup>11,27</sup>

Synthetic genomes have been made for a wide range of cell types, from unicellular prokaryotes<sup>7,11,12,18</sup> and eukaryotes<sup>9,10,13–17</sup> to multicellular plants and animals.<sup>28</sup> However, generating a synthetic genome that is capable of maintaining cell fitness is still challenging.<sup>29</sup> Although the genomes of *Synechocystis* PCC6803 and *Mycoplasma genitalium* have been synthesized and assembled in yeast cells, the step to replace the corresponding native genomes failed.<sup>5,19</sup> Design flaws often occur due to

limited biological knowledge or incomplete design principles,<sup>7,13,15</sup> whilst spontaneous DNA variations will occur during cell division and genome assembly.<sup>7,14,16,17,30</sup> For example, these variations include single-nucleotide variation (SNV), insertion/deletion (InDel), structural variation (SV) and copy number variation (CNV). Thus, researchers have to detect the bugs and debug following every assembly and sometimes even have to redesign the genome in order to maintain cell fitness (Fig. 1A).<sup>7,10</sup> Genome synthesis is associated with debugging throughout the process,<sup>29</sup> for example, about 15 years and US\$40 million were spent on the synthesis and debugging of the *Mycoplasma mycoides* genome.<sup>29</sup> Even the presence of a single base variant could have delayed the process for months.<sup>7</sup> During our work on the synthetic yeast chromosome V,<sup>16</sup> our group organized 5 sub-teams and spent more than 6 months debugging a lethal defect (unpublished data). Genome synthesis is tedious work and the most difficult task is to locate and debug errors. Here, we review recent progress in mapping and correcting bugs, revealing insights into genome synthesis.

### Bug mapping

Phenotypes, genotypes and omics can reflect the presence of bugs in synthetic genome sequences.<sup>13–17</sup> Genome reassembly with semisynthetic DNA constructs,<sup>7</sup> or iterative meiotic recombination and tetrad dissection have been used to narrow down possible defective regions (Fig. 1B).<sup>13,16</sup> Using semisynthetic genome constructs, the failure to transplant a synthetic *M. mycoides* genome was attributed to the 811–900 construct.<sup>7</sup> A yeast strain carrying a synthetic yeast chromosome VI (synVI) exhibited a respiratory deficient. In this case a meiotic recombination strategy and an “endoreduplication backcross” strategy were used to track the defect to the synthetic PCRTag closer to the *PRE4* 3' end (Table 1).<sup>13</sup>

Frontiers Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300072, PR China. E-mail: yjyuan@tju.edu.cn  
<sup>†</sup> Equal contribution.



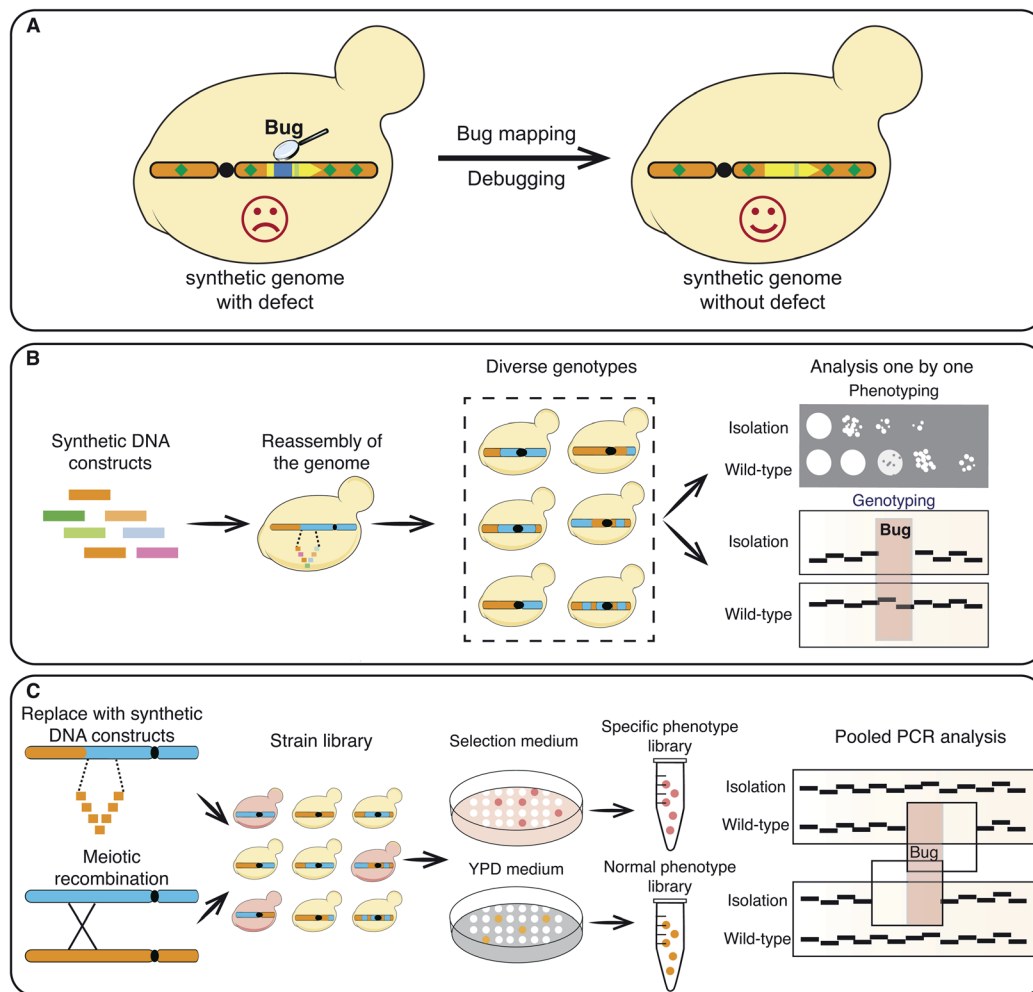


Fig. 1 Bug mapping of synthetic yeast chromosomes. (A) Mapping and debugging unanticipated design flaws or sequence variations on synthetic yeast chromosomes can recover cell growth fitness. (B) Bug mapping with a genome reassembly strategy. Synthetic DNA constructs are used to replace the corresponding native genomic sequences to generate colonies with diverse genotypes. Colonies with growth defects are isolated for genomic analysis to narrow down potential defective loci. (C) Bug mapping by PoPM strategy. A genome reassembly strategy or a meiotic recombination strategy are used to generate candidate colonies with diverse genotypes. Then, colonies are separated into two libraries based on their fitness, with one being robust and the other being defective in growth. The two libraries undergo pooled PCR analysis to track potential loci that cause growth defects.

Both the aforementioned strategies require multiple steps that have to be carried out repeatedly over numerous rounds, and thus are extremely time consuming and labor-intensive. A pooled PCRTag mapping (PoPM) strategy was recently developed to locate bugs on the synthetic yeast chromosome X (synX).<sup>15</sup> Utilizing this strategy one can rapidly locate bugs on a synthetic yeast chromosome (Fig. 1C).<sup>15</sup> Yeast colonies were divided into two libraries based on their fitness, with one being normal and the other being defective in growth. Subsequently, PCR analysis was performed on both libraries using pre-designed PCRTags based on synthetic and wild-type chromosomes. The specific regions only detected in the normal library but not in the defective library were considered potential defective loci. Using a PoPM strategy, the causes of growth defects were quickly pinpointed to the genomic loci including the deletion of *tR(CCU)*, the inserted loxPsym site at the 3'UTR of *YJR120W* and a specific synonymous recoding of the essential gene *FIP1* (Fig. 2A–C).<sup>15</sup>

SNVs, InDels, SVs and CNVs on synthetic yeast chromosomes have been detected by various methods including pulsed field gel electrophoresis (PFGE), whole genome sequencing (WGS), and PCR and qPCR analysis (Fig. 2D and E).<sup>9,10,13,14,16,17</sup> CNVs on synthetic chromosomes are often associated with introduced loxPsym sites or cohesive termini resulting from restriction enzyme digestion during chromosome assembly.<sup>15,16</sup> For example, a CNV on synX involved three fragments of 12.8, 27.7 and 5.6 kb in length, which were ligated by the loxPsym sequence (Fig. 2E).<sup>15</sup> The quadruplication of a 2 kb sequence on synV was tandemly ligated by a short cohesive-end-like "GCGGCGC" sequence.<sup>16</sup>

## Debugging

Reverting synthetic sequences at the defective loci to the corresponding native ones is a typical debugging method.<sup>13,17</sup> For example, in the synthesis of synX, the PCRTag in *FIP1*,



Table 1 The pros and cons of various debugging methods

Methods	Pros	Cons
HR-based gene replacement	<ul style="list-style-type: none"> <li>• The method is simple and based on two HR procedures.</li> <li>• The method can be used to repair SNVs,<sup>13,15</sup> InDels<sup>14,15,17</sup> and CNVs.<sup>16</sup></li> </ul>	<ul style="list-style-type: none"> <li>• The selection process is laborious.</li> <li>• False positive rates are relatively high.</li> </ul>
I-SceI endonuclease mediated method	<ul style="list-style-type: none"> <li>• DSBs are introduced in specific genomic loci by I-SceI.</li> <li>• SNVs,<sup>48</sup> SVs,<sup>14</sup> and CNVs<sup>14</sup> can be seamlessly modified with high efficiency.</li> </ul>	<ul style="list-style-type: none"> <li>• An extra step to incorporate the I-SceI cassette is necessary.</li> </ul>
Meiotic crossover	<ul style="list-style-type: none"> <li>• This method can be used to repair extra copies of synthetic DNA that misintegrated on the synthetic chromosome.<sup>15</sup></li> <li>• The insertion of selective markers can promote the selection efficiency.</li> </ul>	<ul style="list-style-type: none"> <li>• The strain carrying the synthetic chromosome must be mated with a native strain of the opposite mating-type.</li> <li>• The chromosomal crossover randomly occurs throughout the whole chromosome, thus the selection process is laborious.</li> </ul>
Endoreduplication backcross	<ul style="list-style-type: none"> <li>• This method can avoid crossover between synthetic chromosomes and corresponding native chromosomes during meiosis.</li> <li>• CNVs of long DNA sequences that misintegrated on other chromosome(s),<sup>16</sup> 2-micron plasmid deletions<sup>15</sup> and mitochondria deficiencies<sup>13</sup> can be efficiently repaired by using this method.</li> </ul>	<ul style="list-style-type: none"> <li>• The strain carrying the synthetic chromosome must be mated with a native strain of the opposite mating-type, in which the corresponding native chromosome needs to be engineered with a specific cassette of <i>GAL1p-CEN::KL.URA3</i>.</li> </ul>
CRISPR/Cas9-mediated genomic editing	<ul style="list-style-type: none"> <li>• DSBs are introduced in specific genomic loci by CRISPR/Cas9.</li> <li>• SNVs and InDels can be repaired in one step by subsequent homologous recombination of the donor DNA.<sup>16</sup></li> </ul>	<ul style="list-style-type: none"> <li>• The method relies on the presence and precise positioning of PAM site in the target sequence and therefore is not applicable to the entire genome.</li> <li>• Other problems of this method include off-target cleavage and variable efficiencies.</li> </ul>
Dual-labelling correction	<ul style="list-style-type: none"> <li>• The PAM site and corresponding protospacer are inserted to make the target loci applicable to CRISPR/Cas9.</li> <li>• DSBs are introduced in specific genomic loci by CRISPR/Cas9 to facilitate the repair of SNVs and InDels.<sup>16</sup></li> <li>• The false positive rate is decreased by double selection of both an antibiotic and an auxotroph.</li> <li>• Multiple targets can be repaired simultaneously.</li> </ul>	<ul style="list-style-type: none"> <li>• It is necessary to incorporate a doubly selective cassette containing both an antibiotic gene and an auxotroph gene adjacently to a specific target for repair.</li> </ul>

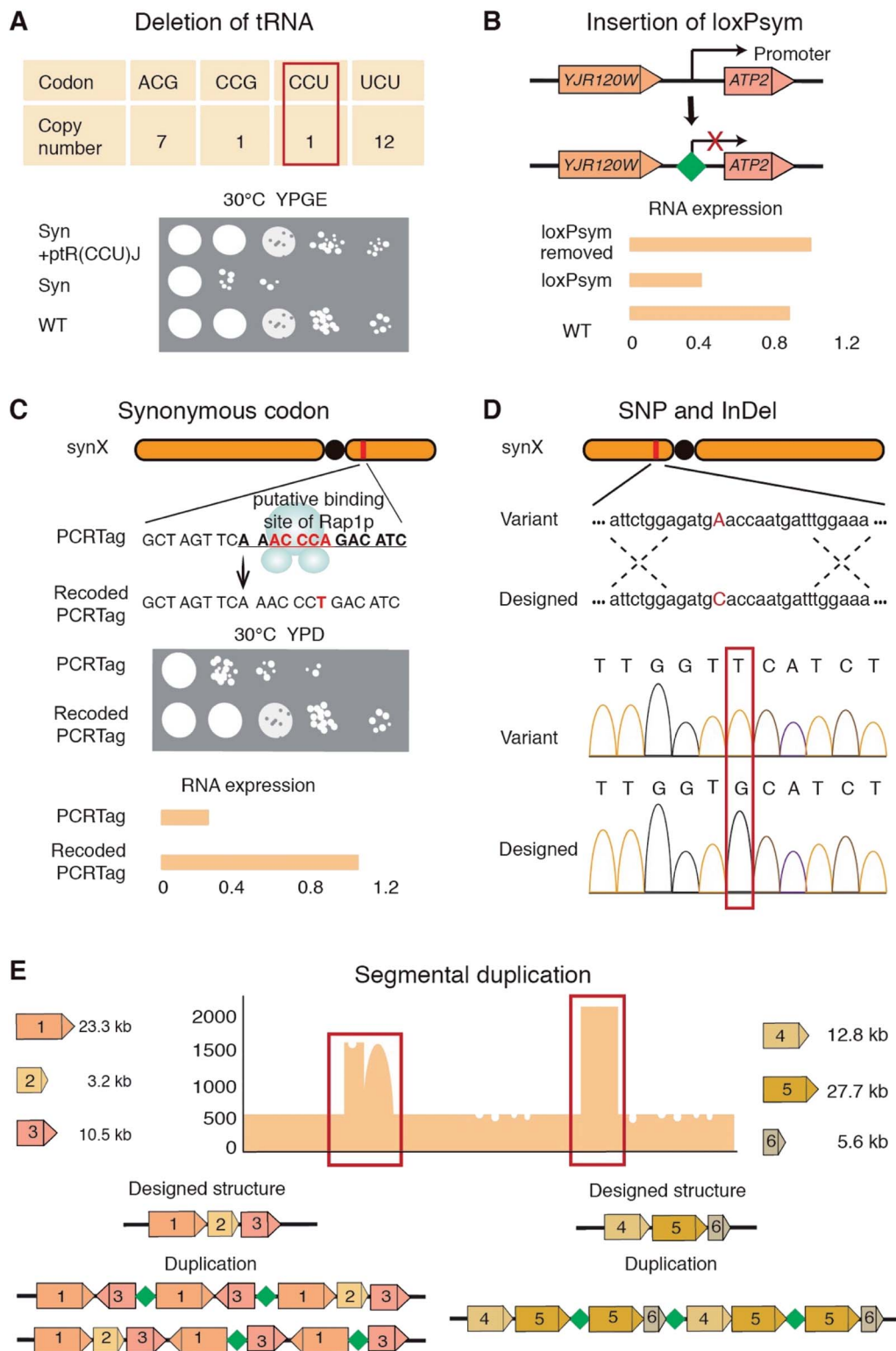
a synonymously recoded sequence, was designed and introduced to resemble the binding site of Rap1p. Consistent with Rap1p being a transcriptional repressor, the authors observed that *FIP1* expression in the *synX* strain was down-regulated. Strikingly, merely reverting the base pairs in the putative Rap1p binding site was sufficient to completely repair the growth deficiency (Fig. 2C).<sup>15</sup> Another debugging example from the *synX* project was at the *ATP2* locus. In a qPCR screen, *ATP2* expression in the *synX* strain was found to be significantly lower than the wild type level. An loxPsyn site inserted downstream of the neighboring ORF *YJR120W* was hypothesized to alter the *ATP2* promoter function. Deletion of this loxPsyn site restored the *ATP2* expression and recovered the fitness of the strain.<sup>15</sup>

Chromosome abnormalities, SVs and CNVs, are not uncommon in synthetic yeast chromosomes and have been fixed to recover cell fitness. The “endoreduplication backcross” method was an efficient way to repair the CNVs of long sequences (Fig. 3A and Table 1).<sup>20,31</sup> Using this method, an extra copy of two

*synV* sequences (291 217–299 980 bp and 305 355–325 235 bp) on other chromosome(s) was removed, and the conditional growth defects of the strain were consequently repaired.<sup>16</sup> Meiotic crossover is a frequently employed strategy to repair CNVs on the same chromosome (Fig. 3B). To remove the CNV on *synX*, haploids of opposite mating types carrying *synX* and semi-*synX*, with and without the CNV, respectively, were crossed. Following tetrad dissection, a haploid strain carrying a *synX* without the CNV was isolated using meiotic recombination.<sup>15</sup> To facilitate the repair of SVs and CNVs, I-SceI endonuclease was employed to generate a double-strand break (DSB) at the target region and to promote homologous recombination (HR) (Fig. 3C).<sup>16,17</sup>

It is hard to predict the effects of each of the numerous SNVs and InDels that occur during genome synthesis. These may interfere with the evaluation of the design principles. Therefore, perfect matching of the synthesized sequence to the designed sequence is crucial and remains a major challenge in the field.<sup>16</sup> Recently a combination of various methods including HR-based





**Fig. 2** Bugs on synthetic yeast chromosomes. (A) The deletion of *tR(CCU)J* leads to a growth defect in the synX strain under selective conditions. (B) The insertion of a loxPsym site at the 3'UTR of *YJR120W* disrupts the expression of *ATP2*. (C) The PCRTag in *FIP1* is recoded to resemble the binding site of Rap1p, which down-regulates *FIP1* expression. Spontaneously SNVs/InDels (D) and CNVs and SVs (E) are detected on synthetic yeast chromosomes. YPGE, yeast extract peptone glycerol ethanol. YPD, yeast extract peptone dextrose. Green diamond, loxPsym site.

gene replacement, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated genomic editing and dual-labelling correction have

been employed to fix SNVs and InDels. This resulted in the successful synthesis of a synV that “perfectly” matches the designed sequence (Table 1).<sup>16</sup>



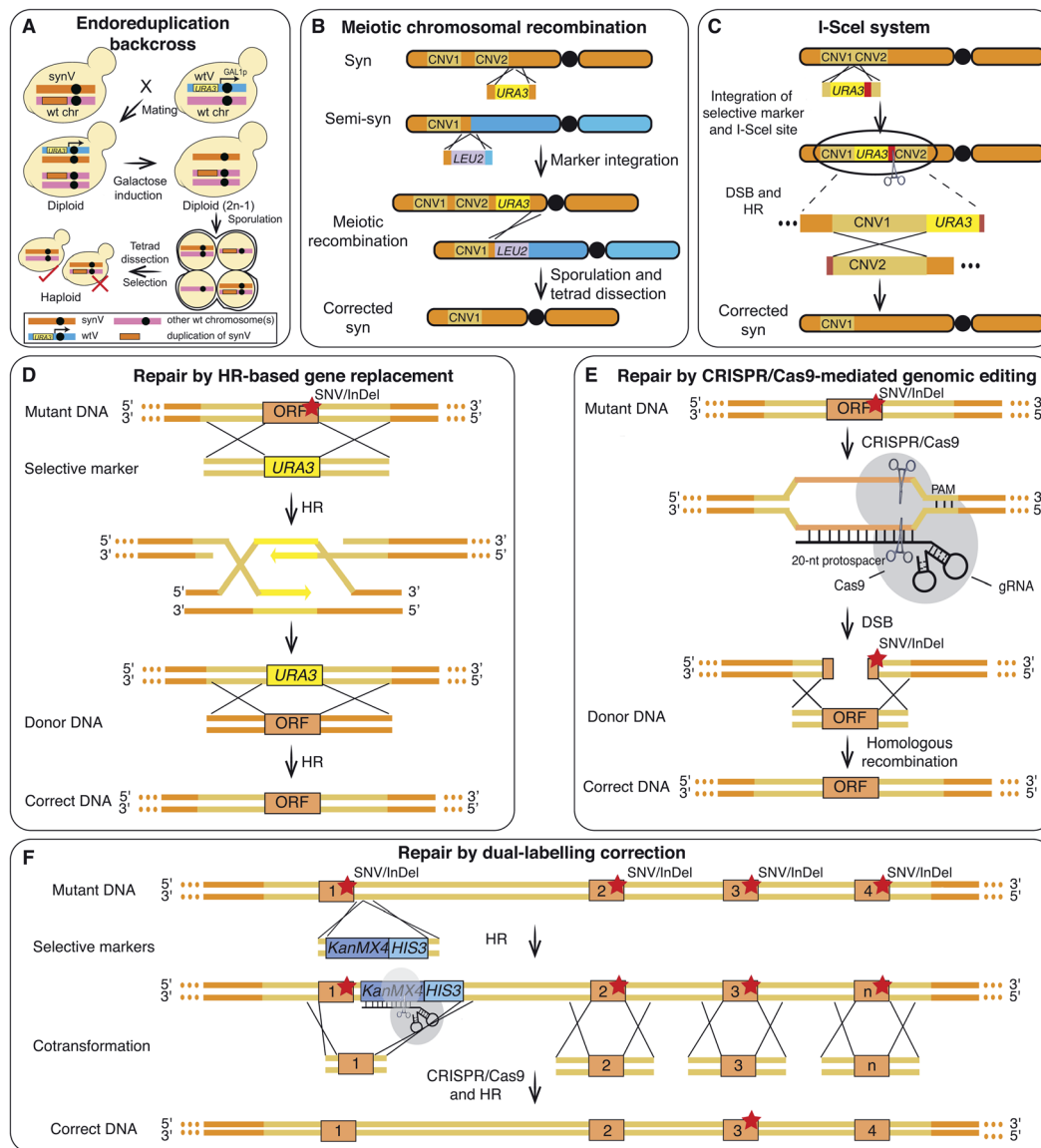


Fig. 3 Debugging of CNVs, SNVs and InDels. (A) An “endoreduplication backcross” strategy was used to repair large-scale duplications of synV that misintegrated on other chromosome(s). (B) A Meiotic crossover strategy was used to restore the duplication that misintegrated on the synthetic chromosome. (C) An I-SceI mediated repair strategy was used to correct SVs. The red segments denote I-SceI recognition sites. SNVs/InDels on synthetic yeast chromosomes are repaired using HR-based gene replacement (D), CRISPR/Cas9-mediated genomic editing (E) and dual-labelling correction (F) strategies. The orange rectangles denote the target regions that need to be corrected and the red stars represent the SNVs/InDels.

For the initial synV, SNVs and InDels involving 3331 bp were revealed using WGS.<sup>16</sup> The HR-based gene replacement, requiring two steps of homologous recombination and selection (Fig. 3D), is laborious. A CRISPR/Cas9-mediated one-step method was then developed. It has been reported that the use of the CRISPR/Cas9 genomic editing tool to generate DSBs resulted in a 5–130 fold increase in the HR efficiency.<sup>32</sup> Therefore, CRISPR/Cas9 was adapted in an effort to eliminate the SNVs and InDels during the construction of synV (Fig. 3E).<sup>16,33</sup> However, this method relies on the presence and precise positioning of the NGG sequence, the protospacer adjacent motif (PAM) site in the target sequence and therefore is not applicable

to the entire genome. Other problems with this method include off-target cleavage and variable efficiencies.

The problems associated with the aforementioned methods have hindered the complete removal of errors in synV. Therefore, a dual-labelling correction strategy (Fig. 3F) was developed to repair the remaining errors. A doubly selective cassette containing both an antibiotic gene (*e.g.* *KanMX4*) and an auxotroph gene (*e.g.* *HIS3*) was inserted adjacent to a specific target with the double selection minimizing false positive rates. The insertion of the cassette also introduced a protospacer and a PAM site. Subsequently, a pool of donor DNA fragments covering multiple target sites was co-transformed for multiplex correction. The predesigned artificial protospacer was cut using





Cas9 and thus increasing the HR efficiency.<sup>16</sup> The isolated candidate clone with the most correction events confirmed by sequencing was subjected to the next round of correction. After 24 rounds, all errors, including two new mutations, were corrected resulting in a synV with a sequence as designed and exhibiting wild-type fitness.<sup>16,30</sup>

## Discussion and outlook

Advances in DNA assembly have enabled the synthesis of both prokaryotic and eukaryotic genomes. The “bottom-up” genome synthesis opens the door to many imaginative changes at the genomic scale. These include creating a functional genome of minimal size,<sup>11</sup> data storage,<sup>34</sup> recoding at the genomic scale,<sup>8,18,38–42</sup> and directed genome evolution.<sup>10,35–38</sup> However, design flaws and spontaneous variants often cause malfunctions of the synthetic chromosome, which then require bug finding and removal.<sup>39</sup> We have summarized strategies that have been applied to the bug mapping and debugging process, which typically involve genotypically and phenotypically massive screening and are labor-intensive and time-consuming. Future strategies incorporating automation and high-throughput screenings are highly sought after. The endeavors by the Global Biofoundry Alliance (GBA) have made great advances in automation, which has been used widely for different purposes including genome assembly,<sup>40</sup> directed enzyme evolution, and liquid-based high-throughput analysis.<sup>41–46</sup> Furthermore, we expect that robotic systems and workflows will be automatically executed in the future *via* artificial intelligence and machine learning to facilitate the bug mapping and debugging process.<sup>47</sup>

## Author contributions

Z. X. X. and Y. J. Y. conceived and designed the manuscript. Z. X. X., J. Z., J. F., and Y. J. Y. drafted the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was funded by the National Key Research and Development Program of China (2020YFC1316500), the National Natural Science Foundation of China (21907074) and the Natural Science Foundation of Tianjin (20JCQNJC02090).

## References

- 1 M. Rugnetta, *Synthetic biology*, *Encyclopedia Britannica*, 2016, accessed 11 March 2021, <https://www.britannica.com/science/synthetic-biology>.
- 2 J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737–738.
- 3 E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford,

- J. Howland, L. Kann, J. Lehoczy, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordtsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki and C. International Human Genome Sequencing, *Nature*, 2001, **409**, 860–921.



- 4 J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. Zhu, *Science*, 2001, **291**, 1304–1351.
- 5 D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison III and H. O. Smith, *Science*, 2008, **319**, 1215–1220.
- 6 M. Itaya, K. Fujita, A. Kuroki and K. Tsuge, *Nat. Methods*, 2008, **5**, 41–43.
- 7 D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison III, H. O. Smith and J. C. Venter, *Science*, 2010, **329**, 52–56.
- 8 D. G. Gibson, H. O. Smith, C. A. Hutchison III, J. C. Venter and C. Merryman, *Nat. Methods*, 2010, **7**, 901–903.
- 9 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 and J. D. Boeke, *Nature*, 2011, **477**, 471–476.
- 10 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. 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 and S. Chandrasegaran, *Science*, 2014, **344**, 55–58.
- 11 C. A. Hutchison III, R. Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z. Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson and J. C. Venter, *Science*, 2016, **351**, aad6253.
- 12 N. Ostrov, M. Landon, M. Guell, G. Kuznetsov, J. Teramoto, N. Cervantes, M. Zhou, K. Singh, M. G. Napolitano, M. Moosburner, E. Shrock, B. W. Pruitt, N. Conway, D. B. Goodman, C. L. Gardner, G. Tyree, A. Gonzales, B. L. Wanner, J. E. Norville, M. J. Lajoie and G. M. Church, *Science*, 2016, **353**, 819–822.
- 13 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. Fenyo, J. Dai, J. S. Bader and J. D. Boeke, *Science*, 2017, **355**, eaaf4831.



- 14 Y. Shen, Y. Wang, T. Chen, F. Gao, J. Gong, D. Abramczyk, R. Walker, H. Zhao, S. Chen, W. Liu, Y. Luo, C. A. Muller, 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 and H. Yang, *Science*, 2017, **355**, eaaf4791.
- 15 Y. Wu, B. Z. Li, M. Zhao, L. A. Mitchell, Z. X. Xie, Q. H. Lin, X. Wang, W. H. Xiao, Y. Wang, X. Zhou, H. Liu, 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 and Y. J. Yuan, *Science*, 2017, **355**, eaaf4706.
- 16 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 and Y. J. Yuan, *Science*, 2017, **355**, eaaf4704.
- 17 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 and J. Dai, *Science*, 2017, **355**, eaaf3981.
- 18 J. Fredens, K. Wang, D. de la Torre, L. F. H. Funke, W. E. Robertson, Y. Christova, T. Chia, W. H. Schmied, D. L. Dunkelmann, V. Beranek, C. Uttamapinant, A. G. Llamazares, T. S. Elliott and J. W. Chin, *Nature*, 2019, **569**, 514–518.
- 19 M. Itaya, K. Tsuge, M. Koizumi and K. Fujita, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15971–15976.
- 20 S. M. Richardson, L. A. Mitchell, G. Stracquadanio, K. Yang, J. S. Dymond, J. E. DiCarlo, D. Lee, C. L. Huang, S. Chandrasegaran, Y. Cai, J. D. Boeke and J. S. Bader, *Science*, 2017, **355**, 1040–1044.
- 21 K. Kannan and D. G. Gibson, *Science*, 2017, **355**, 1024–1025.
- 22 R. W. Barnett and H. Erfle, *Nucleic Acids Res.*, 1990, **18**, 3094.
- 23 R. B. Ciccarelli, P. Gunyuzlu, J. Huang, C. Scott and F. T. Oakes, *Nucleic Acids Res.*, 1991, **19**, 6007–6013.
- 24 J. S. Dymond, L. Z. Scheifele, S. Richardson, P. Lee, S. Chandrasegaran, J. S. Bader and J. D. Boeke, *Genetics*, 2009, **181**, 13–21.
- 25 D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison III and H. O. Smith, *Nat. Methods*, 2009, **6**, 343–345.
- 26 Q. Lin, B. Jia, L. A. Mitchell, J. Luo, K. Yang, K. I. Zeller, W. Zhang, Z. Xu, G. Stracquadanio, J. S. Bader, J. D. Boeke and Y. J. Yuan, *ACS Synth. Biol.*, 2015, **4**, 213–220.
- 27 D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison III, H. O. Smith and J. C. Venter, *Science*, 2010, **329**, 52–56.
- 28 G. A. Logsdon, C. W. Gambogi, M. A. Liskovych, E. J. Barrey, V. Larionov, K. H. Miga, P. Heun and B. E. Black, *Cell*, 2019, **178**, 624–639.
- 29 M. Baker, *Nature*, 2011, **473**, 403–408.
- 30 Z. X. Xie, D. Liu, B. Z. Li, M. Zhao, B. X. Zeng, Y. Wu, Y. Shen, T. Lin, P. Yang, J. Dai, Y. Cai, H. Yang and Y. J. Yuan, *Chem. Soc. Rev.*, 2017, **46**, 7191–7207.
- 31 Z. X. Xie, L. A. Mitchell, H. M. Liu, B. Z. Li, D. Liu, N. Agmon, Y. Wu, X. Li, X. Zhou, B. Li, W. H. Xiao, M. Z. Ding, Y. Wang, Y. J. Yuan and J. D. Boeke, *G3: Genes, Genomes, Genet.*, 2018, **8**, 173–183.
- 32 J. E. DiCarlo, J. E. Norville, P. Mali, X. Rios, J. Aach and G. M. Church, *Nucleic Acids Res.*, 2013, **41**, 4336–4343.
- 33 L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang, *Science*, 2013, **339**, 819–823.
- 34 W. Chen, M. Han, J. Zhou, Q. Ge, P. Wang, X. Zhang, S. Zhu, L. Song and Y. Yuan, *Natl. Sci. Rev.*, 2021, DOI: 10.1093/nsr/nwab028.
- 35 W. Liu, Z. Luo, Y. Wang, N. T. Pham, L. Tuck, I. Perez-Pi, L. Liu, Y. Shen, C. French, M. Auer, J. Marles-Wright, J. Dai and Y. Cai, *Nat. Commun.*, 2018, **9**, 1936.
- 36 B. Jia, Y. Wu, B. Z. Li, L. A. Mitchell, H. Liu, S. Pan, J. Wang, H. R. Zhang, N. Jia, B. Li, M. Shen, Z. X. Xie, D. Liu, Y. X. Cao, X. Li, X. Zhou, H. Qi, J. D. Boeke and Y. J. Yuan, *Nat. Commun.*, 2018, **9**, 1933.
- 37 J. Wang, B. Jia, Z. X. Xie, Y. X. Li and Y. J. Yuan, *Front. Chem. Sci. Eng.*, 2018, **12**, 806–814.
- 38 J. Wang, Z. X. Xie, Y. Ma, X. R. Chen, Y. Q. Huang, B. He, B. Jia, B. Z. Li and Y. J. Yuan, *Nat. Commun.*, 2018, **9**, 3783.
- 39 S. Chen, Z. X. Xie and Y. J. Yuan, *FEMS Yeast Res.*, 2020, **20**, foaa012.
- 40 N. Hillson, M. Caddick, Y. Cai, J. A. Carrasco, M. W. Chang, N. C. Curach, D. J. Bell, R. Le Feuvre, D. C. Friedman, X. Fu, N. D. Gold, M. J. Herrgard, M. B. Holowko, J. R. Johnson, R. A. Johnson, J. D. Keasling, R. I. Kitney, A. Kondo, C. Liu, V. J. J. Martin, F. Menolascina, C. Ogino, N. J. Patron, M. Pavan, C. L. Poh, I. S. Pretorius, S. J. Rosser, N. S. Scrutton, M. Storch, H. Tekotte, E. Travnik,





- C. E. Vickers, W. S. Yew, Y. Yuan, H. Zhao and P. S. Freemont, *Nat. Commun.*, 2019, **10**, 2040.
- 41 R. Chao, J. Liang, I. Tasan, T. Si, L. Y. Ju and H. M. Zhao, *ACS Synth. Biol.*, 2017, **6**, 678–685.
- 42 T. Si, R. Chao, Y. Min, Y. Wu, W. Ren and H. Zhao, *Nat. Commun.*, 2017, **8**, 15187.
- 43 F. Ma, M. T. Chung, Y. Yao, R. Nidetz, L. M. Lee, A. P. Liu, Y. Feng, K. Kurabayashi and G. Y. Yang, *Nat. Commun.*, 2018, **9**, 1030.
- 44 G. O. F. Gowers, S. M. Chee, D. Bell, L. Suckling, M. Kern, D. Tew, D. W. McClymont and T. Ellis, *Nat. Commun.*, 2020, **11**, 868.
- 45 A. R. Abate, T. Hung, R. A. Sperling, P. Mary, A. Rotem, J. J. Agresti, M. A. Weiner and D. A. Weitz, *Lab Chip*, 2013, **13**, 4864–4869.
- 46 H. S. Moon, K. Je, J. W. Min, D. Park, K. Y. Han, S. H. Shin, W. Y. Park, C. E. Yoo and S. H. Kim, *Lab Chip*, 2018, **18**, 775–784.
- 47 S. H. M. Mehr, M. Craven, A. I. Leonov, G. Keenan and L. Cronin, *Science*, 2020, **370**, 101–108.
- 48 V. N. Noskov, T. H. Segall-Shapiro and R. Y. Chuang, *Nucleic Acids Res.*, 2010, **38**, 2570–2576.

