SYNTHETIC BIOLOGY

Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly

Kaihang Wang*, Daniel de la Torre, Wesley E. Robertson, Jason W. Chin†

The design and creation of synthetic genomes provide a powerful approach to understanding and engineering biology. However, it is often limited by the paucity of methods for precise genome manipulation. Here, we demonstrate the programmed fission of the *Escherichia coli* genome into diverse pairs of synthetic chromosomes and the programmed fusion of synthetic chromosomes to generate genomes with user-defined inversions and translocations. We further combine genome fission, chromosome transplant, and chromosome fusion to assemble genomic regions from different strains into a single genome. Thus, we program the scarless assembly of new genomes with nucleotide precision, a key step in the convergent synthesis of genomes from diverse progenitors. This work provides a set of precise, rapid, large-scale (megabase) genome-engineering operations for creating diverse synthetic genomes.

Efforts to minimize (*1*, *2*), refactor (*3*), recode (*4*, *5*), and reorganize (*2*, *6*) chromosomes and genomes are providing new insights and opportunities. However, in *Escherichia coli*, the workhorse of synthetic biology, the methods necessary to realize a complete set of operations for synthetic genome design are missing. These operations include (i) the iterative replacement of genomic DNA with synthetic DNA, (ii) deletion of genomic DNA, (iii) translocation of large genomic sections, and (iv) inversion of large genomic sections as well as (v) methods for combining large genome sections from distinct strains for the convergent assembly of synthetic genomes. Each operation should be scarless and programmed with nucleotide precision so that genome designs can be precisely and rapidly realized.

Efficient, precise, and robust methods for iterative replacement (>100 kb per step) and deletion of genome sections have been reported (*7*); however, there has been less progress on creating methods for generating precisely programmed inversions or translocations in *E. coli*, with most current methods for inversions relying on sequence-specific recombinases. Moreover, methods for combining large (e.g., 0.5-Mb) sections from distinct genomes rely on classical conjugation (*8*) and its derivatives (*5*, *9*). Although these methods can be useful (*5*, *9*), they are fundamentally limited because (i) they require large regions of homology [commonly at least 3 kb, and sometimes up to 400 kb, between the donor and recipient genomes (*5*)], (ii) undesired chimeras between the two genomes may result, and (iii) the site of crossover between the two genomes is not precisely specified. Indeed, in favorable cases, crossovers are only selected with kilobase resolution.

Chromosome fission and fusion have occurred in natural evolution (*10*, *11*), and these processes may have accelerated evolution (*10*, *12*, *13*). The synthetic splitting and fusion of chromosomes have been explored to a limited extent, primarily in naturally recombinogenic organisms (*13*–*18*). One report excised up to 720 kb from a single region of the *E. coli* genome (*19*) by using natural homologous recombination in *E. coli*. Because the recombination frequency in *E. coli* is generally low (*20*), this approach is presumably very inefficient. A protelomerase of bacteriophage N15 and a *Vibrio* origin of replication were used to divide the circular *E. coli* chromosome into two linear subchromosomes. However, only one characterized arrangement was viable (*21*). Thus, the limited methods for splitting the *E. coli* genome are not general or efficient.

Here, we demonstrate that an *E. coli* genome, without any prior modification, can be efficiently split, by single-step programmed fission, into pairs of synthetic chromosomes. The resulting synthetic chromosomes enable precise, programmed fusions, genomic inversions, and translocations; moreover, they provide a route to assemble new genomes through the precise, convergent assembly of large genomic fragments from distinct strains.

We designed and synthesized a system to precisely split the unmodified genome into two user-defined, circular chromosomes (Fig. 1A) and tested our approach by splitting the *E. coli* MDS42 (*1*) genome (data file S1) into a 3.43- and a 0.56-Mb chromosome. To achieve this, we first introduced Cas9 with appropriate spacers (table S1), the lambda-red recombination machinery, and a fission bacterial artificial chromosome (BAC) (data file S2) into cells. We implemented six Cas9-directed cuts in the DNA of these cells; two of these cuts target the genome, and four of these cuts target the fission BAC (data files S3 and S4). The two cuts in the genome create fragment 1 and fragment 2, and the four cuts in the fission BAC release linker sequence 1 and linker sequence 2. Chromosome 1 (3.43 Mb) containing the genomic origin of replication (*oriC*) was formed through lambda-red–mediated recombination between genomic fragment 1 and linker sequence 1, by virtue of their 50–base pair (bp) regions of homology (table S2). Similarly, chromosome 2 (0.56 Mb) was formed through lambda red–mediated recombination between genomic fragment 2 and linker sequence 2 (Fig. 1A and fig. S1); this linker sequence contained its own replication and segregation machinery.

In the prefission strain, the fission BAC is nonessential and contains a *SacB-Cm*^{R} double selection cassette (this confers resistance to chloramphenicol and sensitivity to sucrose, but cells can grow on sucrose by losing the fission BAC), the *luxABCDE* operon (conferring luminescence), and *rpsL* (conferring sensitivity to streptomycin). After successful fission, the *rpsL* gene is lost, cells are resistant to streptomycin, the *luxABCDE* operon is removed from a strong promoter to chromosome 1 (leading to weaker luminescence), and the *SacB-Cm*^{R} double selection cassette becomes part of chromosome 2 and cannot be lost. Thus, correct postfission cells are selectively sensitized to sucrose.

After execution of the fission protocol, we enriched for cells that had undergone genome fission to generate two chromosomes, through growth on streptomycin and chloramphenicol (table S3). This selects for both loss of *rpsL* and maintenance of *Cm*^{R} in the *SacB-Cm*^{R} double selection cassette and therefore kills cells containing the fission BAC but allows growth of cells that have undergone programmed genome fission.

We characterized individual postfission clones by several independent methods. First, we examined the luminescence of cells and their growth on selective media (Fig. 1B). Successful clones had decreased luminescence with respect to prefission controls, gained sucrose sensitivity, and gained the ability to grow when challenged simultaneously with both chloramphenicol and streptomycin. Second, we performed polymerase chain reactions (PCRs) across the new junctions resulting from fission. Successful postfission clones exhibited bands of the expected size that were not present in prefission clones (Fig. 1C). This confirmed that both fission junctions were as expected. Third, we confirmed the expected restriction enzyme digestion pattern for the postfission genome by pulsed-field gel electrophoresis (fig. S2). Finally, we determined the replicon organization of the genome by de novo assembly; we achieved this by combining the results of short-read (300-bp paired end) and long-read (N50 of ~8.3 kb) sequencing in Unicycler (*22*) to generate one contig per replicon. The postfission assembly formed two circular contigs, which corresponded to the chromosomes expected from fission (fig. S3 and table S4). The

Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, England, UK.
*Present address: Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA.
†Corresponding author. Email: chin@mrc-lmb.cam.ac.uk

copy number of each chromosome was as expected (table S5).

We demonstrated the scope and generality of fission by programming the splitting of the genome into five additional distinct and diverse pairs of chromosomes (Fig. 2 and figs. S2 to S4). These included a pair in which chromosome 1 is 2.44 Mb and chromosome 2 is 1.55 Mb. Because chromosome 2 has BAC-derived replication and segregation machinery, our data are consistent with BACs being able to maintain megabases of DNA. The only constraints we imposed on the choice of fission sites were that they contained a protospacer adjacent motif (PAM) for Cas9 and lay greater than 30 bp outside any gene. Although a single 2-Mb fission test failed (fig. S5 and table S3), all other experiments we tried led to successful fission (figs. S1 to S4). Fission had only modest effects on the growth of cells (fig. S6). We observed that the genome fissions were present after approximately 105 generations of continuous growth (fig. S7).

We demonstrated that the programmed fusion of synthetic chromosomes, generated by fission, enables the generation of precisely rearranged genomes (Fig. 3). We applied fission to a cell in which ~0.54 Mb, section C (Fig. 2A and figs. S1 and S3), of the genome is watermarked by 2521 synonymous codon changes (*5*) (data file S5); this brought the total number of successful fissions to 7 (Fig. 2A and fig. S1). The resulting cell contained chromosome 1 (3.45 Mb) and a watermarked chromosome 2 (0.54 Mb). After fission, we replaced the *SacB-Cm*^{R} double selection cassette in chromosome 2 with an *oriT-pheS*-Kan*^{R} cassette (table S6). This cell provided a common intermediate for diverse fusions.

We first used fusion to regenerate the original genome. We prepared chromosome 1 for fusion

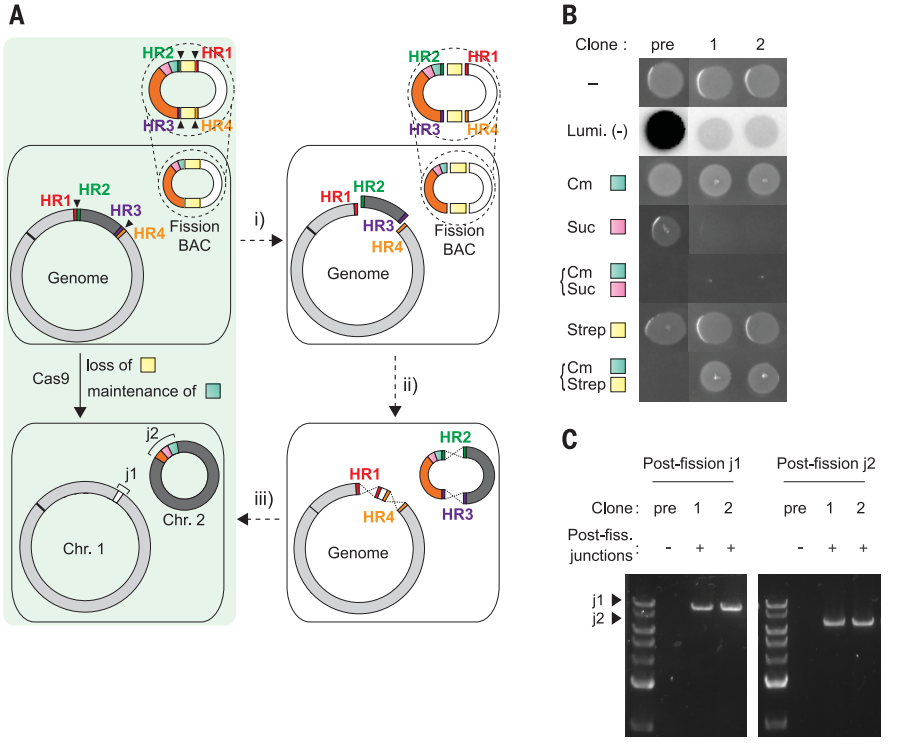

Fig. 1. Programmed genome fission splits the *E. coli* genome into two chromosomes. (**A**) *E. coli* harbors a fission BAC containing a double selection cassette (*sacB-Cm*^{R} shown as pink and green, respectively), *rpsL* (yellow), a *luxABCDE* operon (white), and the BAC replication machinery (orange). During fission, (i) Cas9 induces six cuts (black triangles), splitting the genome into fragment 1 (light gray, containing *oriC* indicated by black line) and fragment 2 (dark gray) and the fission BAC into four pieces (linker sequence 1, linker sequence 2, and two copies of *rpsL*). (ii) Homology regions (HRs) between fragments and their cognate linkers. (iii) Lambda red recombination joins fragments and linkers to yield chromosomes 1 and 2 (Chr. 1 and Chr. 2). Junctions 1 and 2 (j1 and j2) are new junctions. (**B**) Growth and luminescence (Lumi.) of prefission (pre) and postfission (1 and 2) clones are consistent with the generation of two chromosomes (Chr. 1, ~3.43 Mb and Chr. 2, ~0.56 Mb). Cells were stamped in plain LB agar (-), 20 μg/ml chloramphenicol (Cm), 7.5% sucrose (Suc), 100 μg/ml streptomycin (Strep), or the indicated combination. (**C**) PCR of postfission (Post-Fiss.) clones across j1 and j2.

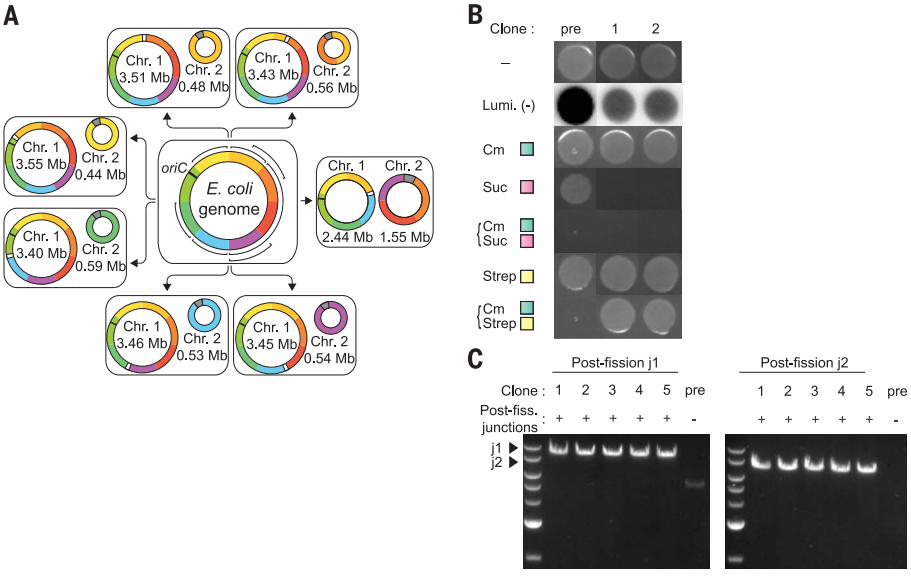

Fig. 2. Fission can be performed throughout the *E. coli* genome. (**A**) Successful fissions performed. Each color on the *E. coli* genome corresponds to ~0.5 Mb. We named the sections A to H. A is dark orange, and the other sections are labeled alphabetically in a clockwise sequence. Linker sequence 1, white; *oriC*, black bar; linker sequence 2, gray. Boundaries and homologies of each fission experiment are provided in table S2. Seven fissions are shown, including the 3.43, 0.56 Mb fission (Fig. 1). The 3.45, 0.54 Mb fission (purple Chr. 2) was performed by using an *E .coli* genome in which a ~0.54-Mb section had been recoded (Fig. 3). (**B**) Growth and luminescence for the generation of the 2.44, 1.55 Mb fission; annotation as in Fig. 1B. Data for other fissions are shown in fig. S4. (**C**) PCR of clones across new junctions for 2.44, 1.55 Mb fission. Postfission clones (1 to 5) exhibit products of the expected size, whereas the prefission control does not. Junction PCRs for other fissions are in fig. S4.

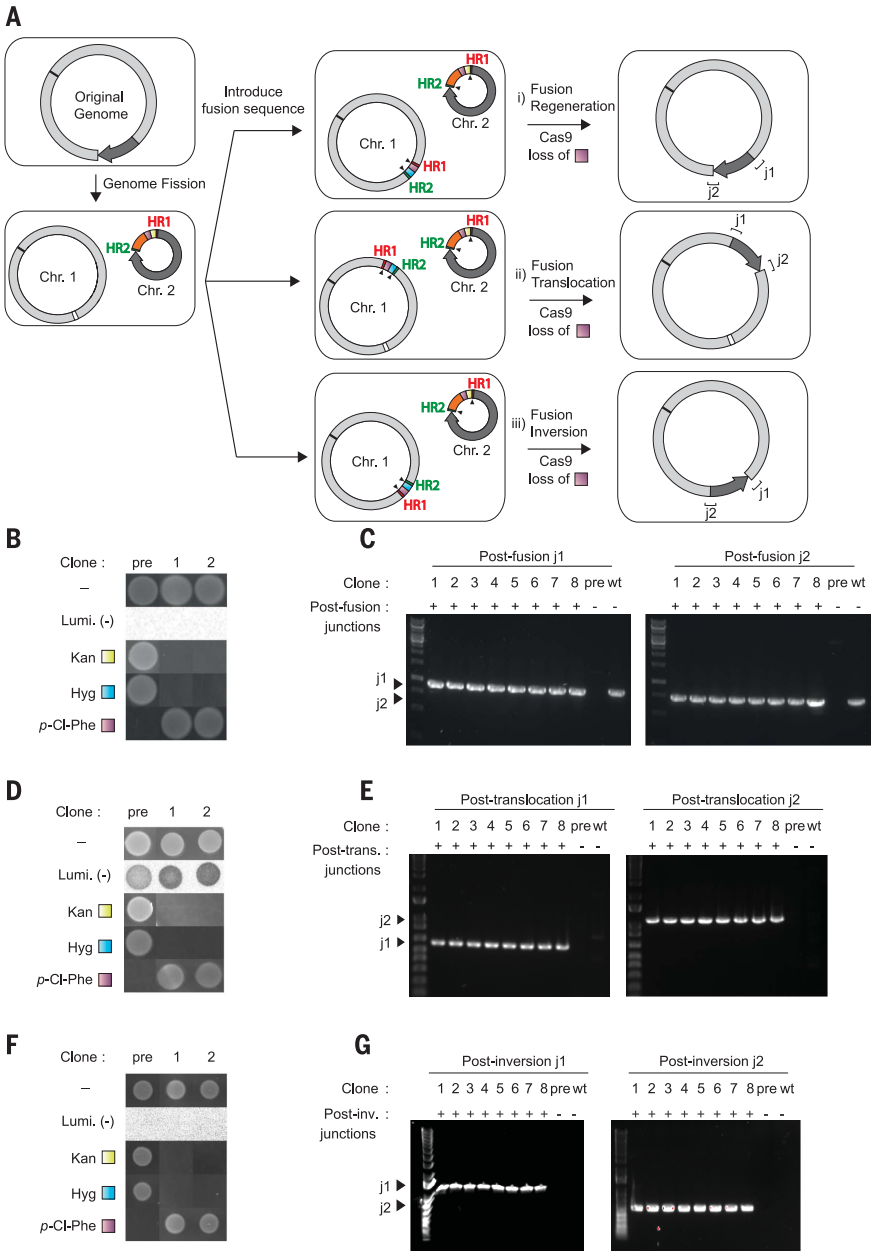

Fig. 3. Programmed chromosomal fusion enables translocations and inversions of large genomic segments from common fission intermediates. (**A**) *E. coli* with two chromosomes (Chr. 1 ~3.45 Mb and Chr. 2 ~0.54 Mb) was generated by fission. The sequence of Chr. 2 is watermarked as described in the text. The color-coding is as in Fig. 1A; a *pheS*-Kan*^{R} double selection cassette (purple and yellow, respectively) is shown. A fusion sequence, consisting of a *pheS*-Hyg*^{R} (purple and blue, respectively) double selection cassette flanked by HR1 and HR2, is introduced in the indicated positions and orientation in Chr. 1 by lambda-red recombination. Cas9 spacer-directed cleavage (black arrows), lambda-red recombination, and selection for fusion products through the loss of *pheS** on 4-chloro-phenylalanine yield the indicated products. (i) Regenerating the original genomic arrangement, (ii) translocation of the 0.54-Mb segment 700 kb away from its original position, and (iii) inversion of the 0.54-Mb segment. (**B**) Growth and luminescence of pre- and postfusion regeneration (1 and 2) clones. Hyg, hygromycin; Kan, kanamycin; *p*-Cl-Phe, 4-chloro-phenylalanine. (**C**) PCR of clones across new junctions for fusion regeneration. Postfusion clones (1 to 8) exhibit products of the expected size, whereas the pre-fusion control does not. wt, wild type. (**D** and **E**) As in (B and C) but for fusion translocation (trans.). (**F** and **G**) As in (B and C) but for fusion inversion (inv.).

by replacing its linker sequence 1 with a fusion sequence for chromosome 2 (oligonucleotide sequences are provided in table S6). This contained a *pheS*-Hyg*^{R} double selection cassette flanked by Cas9 cut sites and homology to fragment 2 in chromosome 2 (Fig. 3A). Fusion was initiated by Cas9-mediated cleavage at either side of the *pheS*-Hyg*^{R} cassette in chromosome 1 and at the ends of the watermarked sequence in chromosome 2, and the resulting homologous ends were joined through lambda red–mediated recombination. We selected the fusion product on 4-chloro-phenylalanine.

We characterized postfusion clones by several independent methods. Successful clones were no longer sensitive to 4-chloro-phenylalanine or resistant to kanamycin or hygromycin (Fig. 3B). PCR across the new junctions generated by fusion led to bands of the correct size that were not present in the prefusion clones (Fig. 3C). We further demonstrated successful fusion by de novo assembly of short-read (300-bp paired end) and long-read (N50 of ~20 kb) sequencing. The prefusion genome formed two circular contigs, whereas all postfusion assemblies formed a single circular contig, which corresponds to the expected fusion product (fig. S3).

We demonstrated that inserting the fusion sequence at different positions in chromosome 1 (500 or 700 kb away from linker sequence 1) (Fig. 3A and fig. S8), followed by initiation of fusion with chromosome 2, enables the selection of genomes bearing defined translocations (Fig. 3, D and E, and figs. S3 and S8). We also demonstrated that inserting the fusion sequence into chromosome 1 in an inverted orientation (Fig. 3A), followed by initiation of fusion with chromosome 2, enables the selection of genomes bearing defined inversions (Fig. 3, F and G, and fig. S3). An attempt at fusion 1.8 Mb away from the linker sequence did not lead to a stable translocation (fig. S8 and table S3).

Next, we combined genome fission, conjugative transplant, and chromosome fusion to precisely combine defined sections of distinct genomes (Fig. 4A). This is a key step in the precise assembly of synthetic genomes from strains containing synthetic sections.

We began with two strains, each containing a different watermarked genomic section [section C or section A (Fig. 2 and fig. S1)], with the target of combining the watermarked sections in a single, chimeric genome. We defined one strain as the donor (data file S5) and the other strain as the recipient (data file S6). We performed fission on the genome of the donor to capture its watermarked sequence in chromosome 2. We then replaced the *SacB-Cm*^{R} double selection cassette in chromosome 2 with an *oriT-pheS*-Kan*^{R} cassette (table S6) and transformed a nontransferable F′ plasmid (*5*) into the donor strain. These steps prepare the donor strain for transplant of chromosome 2 to the recipient.

In parallel, we performed fission on the genome of the recipient to split its genome, at the same position as the donor, into two synthetic

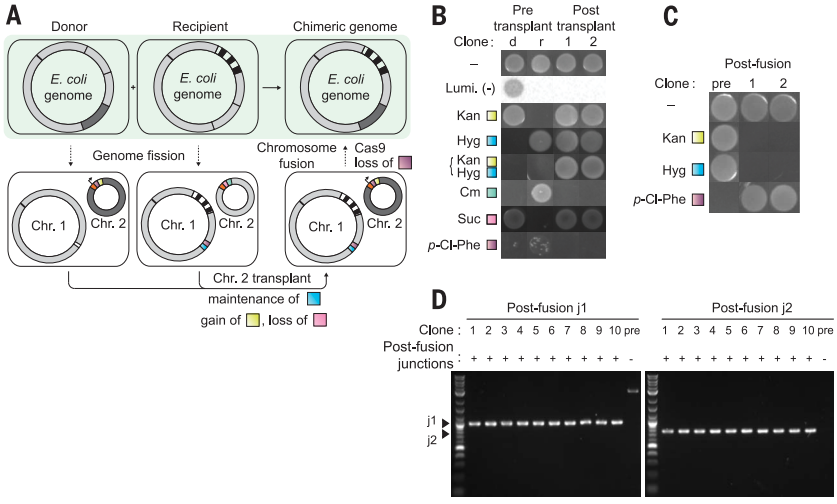

Fig. 4. Precise genome assembly from genomic segments of distinct strains. (**A**) Precisely combining the watermarked region 1 (dark gray) from a donor strain and a watermarked region 2 (black striped) from a recipient strain into a single strain. Fission is performed in parallel in the donor and recipient strains. The resulting donor strain contains a watermarked Chr. 2 containing an *oriT* (black arrow) and a *pheS*-Kan*^{R} double selection cassette (purple and yellow); the remainder of linker sequence 2 is orange. The resulting recipient strain contains an analogous nonwatermarked Chr. 2, with a *sacB-Cm*^{R} cassette (pink and green). The linker sequence 1 (white) is replaced with a fusion sequence containing a *pheS*-Hyg*^{R} cassette (purple and blue) in preparation for fusion. The donor cell is provided with a nontransferable F′ plasmid. Mixing of donor and recipient cells facilitates conjugative transplant of Chr. 2 from the donor to the recipient; selection for *Kan*^{R} and against *sacB*-mediated sucrose sensitivity enables the isolation of cells that have gained a watermarked Chr. 2 and lost the nonwatermarked Chr. 2. Subsequent genome fusion generates a strain in which the watermarked regions 1 and 2 have been precisely combined in a single chromosome. (**B**) Following the process of chromosomal transplant by growth on selective media and luminescence. d, the pretransplant donor; r, pretransplant recipient. (**C**) Following the process of chromosomal fusion through growth on selective media. (**D**) PCR across the new junctions generated by chromosomal fusion yields products of the expected size in the postfusion clones (1 to 10) but not in the prefusion control.

chromosomes. This created a recipient containing a nonwatermarked chromosome 2 (the fission BAC used in the recipient, and therefore chromosome 2, contains a *sacB-Cm*^{R} cassette and does not contain *oriT*) and chromosome 1 that contains the second watermarked region. The linker sequence 1 in chromosome 1 of the recipient was then replaced with a fusion sequence containing a *pheS*-Hyg*^{R} cassette flanked by regions of homology to the fragment of the original genome captured in chromosome 2.

To generate cells that contain both watermarked regions, we mixed donor and recipient cells. We selected for transfer of chromosome 2 from the donor to the recipient and recipient cells in which chromosome 2 from the donor had replaced the endogenous chromosome 2; we termed this overall process chromosome transplant. The resulting recipient cells contained chromosome 2 from the donor and chromosome 1 from the recipient. We generated a single, chimeric genome that contains both the watermarked sequences by fusion of the donor chromosome 2 and the recipient chromosome 1 (Fig. 4, B to D; fig. S3; and data file S7). All attempts at genome assembly were successful.

We demonstrated the efficient programmed, single-step fission of the unmodified *E. coli* genome into diverse megabase-scale chromosomes. These chromosomes provide a common intermediate for the facile creation of diverse genomes. The chromosomes in a single cell can be fused into a single genome to effect precise genomic translocations or precise and scarless inversions. This facilitates the realization of reorganized genome designs and the exploration of modular, synthetic syntenies that may be more amenable to engineering (*2*). Moreover, the transplant of chromosomes between cells, followed by fusion, enables the precise convergent assembly of new genomes. Our work provides the necessary set of precise, rapid, large-scale genome-engineering operations for creating diverse synthetic genomes.

REFERENCES AND NOTES

1. G. Pósfai *et al.*, *Science* **312**, 1044–1046 (2006).
2. C. A. Hutchison 3rd *et al.*, *Science* **351**, aad6253 (2016).
3. L. Y. Chan, S. Kosuri, D. Endy, *Mol. Syst. Biol.* **1**, 0018 (2005).
4. M. J. Lajoie *et al.*, *Science* **342**, 357–360 (2013).
5. J. Fredens *et al.*, *Nature* **569**, 514–518 (2019).
6. J. S. Dymond *et al.*, *Nature* **477**, 471–476 (2011).
7. K. Wang *et al.*, *Nature* **539**, 59–64 (2016).
8. J. Lederberg, E. L. Tatum, *Nature* **158**, 558 (1946).
9. F. J. Isaacs *et al.*, *Science* **333**, 348–353 (2011).
10. D. W. Burt, *Cytogenet. Genome Res.* **96**, 97–112 (2002).
11. G. Giannuzzi *et al.*, *Genome Res.* **23**, 1763–1773 (2013).
12. V. S. Cooper, S. H. Vohr, S. C. Wrocklage, P. J. Hatcher, *PLOS Comput. Biol.* **6**, e1000732 (2010).
13. J. A. Escudero, D. Mazel, *Int. Microbiol.* **20**, 138–148 (2017).
14. D. Ausiannikava *et al.*, *Mol. Biol. Evol.* **35**, 1855–1868 (2018).
15. M. Itaya, T. Tanaka, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5378–5382 (1997).
16. J. Luo, X. Sun, B. P. Cormack, J. D. Boeke, *Nature* **560**, 392–396 (2018).
17. Y. Shao *et al.*, *Nature* **560**, 331–335 (2018).
18. Y. Ueda *et al.*, *J. Biosci. Bioeng.* **113**, 675–682 (2012).
19. Y. Yamaichi, H. Niki, *EMBO J.* **23**, 221–233 (2004).
20. C. Raeside *et al.*, *mBio* **5**, e01377–e14 (2014).
21. X. Liang, C. H. Baek, F. Katzen, *ACS Synth. Biol.* **2**, 734–740 (2013).
22. R. R. Wick, L. M. Judd, C. L. Gorrie, K. E. Holt, *PLOS Comput. Biol.* **13**, e1005595 (2017).

ACKNOWLEDGMENTS

We thank J. Houseley and J. Ajioka for providing equipment for pulsed-field gel electrophoresis and J.E. Sale for helpful comments on the manuscript. **Funding:** This work was supported by the Medical Research Council (MRC), UK (MC_U105181009 and MC_UP_A024_1008), and an ERC Advanced Grant SGCR, all to J.W.C. **Author contributions:** K.W. designed and implemented the genome manipulation processes reported. K.W., D.d.l.T., and W.E.R. demonstrated scope. D.d.l.T. implemented the de novo assembly approach. J.W.C. defined the direction of research, supervised the project, and wrote the paper with the other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The sequences for de novo genome assemblies and DNA sequencing data have been deposited in NCBI's GenBank and SRA databases, and their accession numbers are listed in table S4. All other data needed to evaluate the conclusions of the study are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/365/6456/922/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S6
Data Files S1 to S7
References (*23*–*27*)

17 May 2019; accepted 2 August 2019
10.1126/science.aay0737

Science

Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly

Kaihang Wang, Daniel de la Torre, Wesley E. Robertson and Jason W. Chin

Science **365** (6456), 922-926.
DOI: 10.1126/science.aay0737

Programmable genome engineering

The model bacterium *Escherichia coli* has a single circular chromosome. Wang *et al.* created a method to fragment the *E. coli* genome into independent chromosomes that can be modified, rearranged, and recombined. The efficient fission of the unmodified *E. coli* genome into two defined, stable pairs of synthetic chromosomes provides common intermediates for large-scale genome manipulations such as inversion and translocation. Fusion of synthetic chromosomes from distinct cells generated a single genome in a target cell. Precise, rapid, large-scale genome engineering operations are useful tools for creating diverse synthetic genomes.

Science, this issue p. 922