Assembly of Radically Recoded *E. coli* Genome Segments

Julie E. Norville[1,2], Cameron L. Gardner[3,4], Eduardo Aponte[5], Conor K. Camplisson[1,2], Alexandra Gonzales[1], David K. Barclay[1], Katerina A. Turner[1], Victoria Longe[6], Maria Mincheva[7], Jun Teramoto[1,8,9], Kento Tominaga[10], Ryota Sugimoto[11], James E. DiCarlo[12], Marc Guell[2], Eriona Hysolli[1], John Aach[1], Christopher J. Gregg[13], Barry L. Wanner[14], George M. Church[1,2]*

[1]Department of Genetics, Harvard Medical School, Boston MA 02115, USA

[2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA 02138, USA

[3]National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Bethesda MD 20892, USA

[4]Nuffield Department of Medicine, Medical Sciences Division, University of Oxford, Headington, Oxford OX3 7BN, UK

[5]bioMerieux, Cambridge, MA 02142, USA

[6]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02143, USA

[7]Sofia University, Sofia 1000, Bulgaria

[8]On leave from Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

[9]Present address: Integrated Research Center of Kobe University, Chuo-ku, Kobe 650-0047, Japan

[10]Department of Biological Sciences, Graduate Schools of Science and Engineering, Tokyo Metropolitan University, Hachioji-shi, Tokyo 192-0397, Japan

[11]Laboratory for Chemistry and Life Science, Institute of Innovative Research, Tokyo Institute of Technology, R1-104, 4259 Nagatsutacho, Midori-ku, Yokohama 226-8503, Japan

[12]College of Physicians and Surgeons, Columbia University, New York City, NY 10032, USA

[13]GreenLight Biosciences, Inc., Medford, MA 02155

[14]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA

*Correspondence to: Email: gchurch@genetics.med.harvard.edu

Abstract

The large potential of radically recoded organisms (RROs) in medicine and industry depends on improved technologies for efficient assembly and testing of recoded genomes for biosafety and functionality. Here we describe a next generation platform for conjugative assembly genome engineering, termed CAGE 2.0, that enables the scarless integration of large synthetically recoded *E. coli* segments at isogenic and adjacent genomic loci. A stable *tdk* dual selective marker is employed to facilitate cyclical assembly and removal of attachment sites used for targeted segment delivery by site-specific recombination. Bypassing the need for vector transformation harnesses the multi Mb capacity of CAGE, while minimizing artifacts associated with RecA-mediated homologous recombination. Our method expands the genome engineering toolkit for radical modification across many organisms and recombinase-mediated cassette exchange (RMCE).

Introduction

Radically recoded organisms (RROs) are desirable for several reasons. For example, recoding enforces genetic and metabolic isolation (*1, 2*), allows facile use of non-standard amino acids (*1, 3*), and has the potential to enable multivirus resistance (*3, 4*). RROs can be used to produce proteins containing non-standard amino acids and are an enabling tool to build biocontained microbes for environmental remediation, industrial applications, and medicine (*5*). In addition, *E. coli* can be engineered to stably maintain non-native sequences from other organisms (*6-9*), support multiplex modification (*10-12*), and deliver DNA to other organisms including diverse bacteria, fungi, plants, and mammalian cells (*13-17*). The ability to transfer scarless large segments of modified DNA to genomes with or without recombination is a desirable feature as genome-scale modifications of organisms across biological kingdoms becomes routine (*9, 12, 18-31*).

In a separate study, we tested 55 of the 87 50-kb segments required to construct a recoded *E. coli* strain in which seven codons were reassigned to synonymous codons, denoted *rE.coli-57* (*5*). The 3.97 MB genome design of *rE.coli-57* has 62,214 codon replacements (*5*). A transformation-based assembly approach allowed us to evaluate each recoded segment individually for strain viability following deletion of the corresponding chromosomal region. To do this, we created a testing pipeline where synonymous codon replacements are validated for strain fitness of each segment (*5*). To construct the fully recoded *rE.coli-57*, we will need to replace all 87 segments throughout the genome at isogenic loci. Previously, we used Conjugative Assembly Genome Engineering (CAGE) to merge sets of genome modifications (*32*) to construct a

Genomically Recoded Organism (GRO) where all instances of the TAG stop codon had been replaced with TAA stop codons (*3*). Methods that allow replacement of natural sequences with synthetic DNA segments offer the potential of nearly limitless genome modifications, provided the recoded sequences are viable within living cells (*5*).

When constructing organisms with highly modified genome sequences, there are tradeoffs that will determine which genome assembly technique is optimal. One parameter is how much testing of recoded segments is required to generate optimal phenotypes. The requirement to validate strain fitness favors approaches where both sequences are initially maintained to allow for the systematic deletion of the corresponding natural region to identify deleterious sequences. Direct replacement of segments is supported, however, when assembling strains from prevalidated sequences. A second factor is the optimal size of the replacement segments, which can differ based on the number of mutations or design errors present in synthesized DNA, the need to test DNA segments as they are assembled, and the method used for delivery of the segments by electroporation, phage packaging, conjugation, or transplantation. A final consideration is the requirement for homologous recombination, which can facilitate particular strain assembly methodologies but can also generate undesirable artifacts. *E. coli* offers the ability of using recombinogenic or non-recombinogenic cells, and both the testing protocol (*5*) and CAGE 2.0 can benefit from approaches that minimize recombination.

Although the transformation based method used in our testing pipeline optimizes for the delivery of recoded DNA with minimal opportunity for recombination, the protocol (*5*) is not directly amenable to the assembly of multiple segments. A challenge is the size of the recoded DNA regions that can be delivered. Electroporation can only accommodate sequences of up to several hundred kb (*7*). While it is advantageous to individually test 50 kb segments for effects of recoding on host strain fitness, assembling a genome sequentially from these will be time-consuming. Therefore, a method such as CAGE is ideal because it is amenable to hierarchical assembly (*32*). The testing pipeline also does not easily support the repetitive use of dual selective markers, which are needed to iteratively add additional segments to the same strain at isogenic loci. Techniques described herein will make it possible to extend our testing protocol to larger segment assemblies and, if necessary, to combine the segment assembly and testing protocols.

Here we introduce a next generation protocol, termed CAGE 2.0, for integrating recoded segments into a single *E. coli* strain. In place of the CAGE protocol that uses RecA-mediated recombination for hierarchical assembly, CAGE 2.0 employs λ integrase-mediated site-specific recombination for the targeted assembly of recoded segments. Bypassing the need for the large homology regions that are required for homologous recombination, as well as the need for a recombinogenic background, mitigates against recombination events that can lead to loss of recoded regions during the assembly process. Thus, CAGE 2.0 provides routes to further eliminate undesired recombination, replace segments isogenically, and can be employed to introduce multiple recoded segments in a single strain (Fig. 1).

Results

To expand on the capabilities of CAGE in genome engineering, we devised CAGE 2.0 to deliver recoded segments by site-specific recombination using phage attachment sites that are flexibly shuffled with minimal recombineering to enable genome assembly. Recoded-DNA segments are delivered using vectors that harbor an F plasmid origin, which in *E. coli* has the capacity to replicate extremely large DNA segments ranging from 600 kb to 2.5 Mb (*7, 33, 34*).

In the CAGE 2.0 protocol (Fig. 1), low copy F origin plasmids containing recoded segments are directly conjugated into recipient *E. coli* and site-specifically recombined onto the genome by λ integrase-mediated recombination between the vector *attP* and chromosomal *attB* sites, which are inserted at defined loci by recombineering. Next, the corresponding chromosomal segment is eliminated by λ Red-directed exchange using a chloramphenicol resistance cassette. The resultant strain is then modified for subsequent use as a recipient by removal of residual vector sequences and *att* sites flanking the recoded segment in sequential steps using MAGE (*10*) and *E. coli* thymidine kinase *(tdk)* as a dual selectable marker (*35*). Finally, the host is set for introduction of the next segment by introduction of a new *tdk-attB* cassette in the locus where the recoded segment is to be integrated (Fig. 1). A novelty of the protocol is the use of *tdk*, a marker that can be used repeatedly in the same strain to assist in either adding or removing elements from the genome, allowing for segment assembly.

While developing the testing protocol, we initially attempted to integrate recoded segments both close to and far away from their natural location in *E. coli* K-12 BW38028 (*36*). In this recombinogenic ($RecA^+$) strain, we found recombination between natural and recoded segments sequences was problematic at the isogenic locus. Accordingly, we chose two precautions in our initial CAGE 2.0 experiment. We elected to adjacently assemble two recoded segments (segments 43R and 47R (*5*)) in genomic locations greater than 1-Mb away from their natural locations in order to minimize recombination. We also used the *recA1* host *E. coli* K-12 DH10B-MAGE for subsequent CAGE 2.0 experiments to minimize recombination (*5, 37, 38*). By expressing λ integrase during the conjugation process, we were able to recover single-copy recoded segment integrants from the resulting population of cells. Proceeding directly with these clones, we reset the strain by eliminating residual vector sequences and the *attL* and *attR* sites created by integration (Fig. 1).

Performing this process repeatedly and efficiently required the use of a robust dual selectable marker. Our tests with *E. coli tdk* showed that this alternative marker allowed for the maintenance of a stable selective advantage (analogous to *tolC* (*39*)) over multiple rounds of selection using commercially available reagents in a single strain (Fig. 2A, B, and C). Using positive and negative selection conditions optimized for use in GROs (*1*), we found that *E. coli tdk* was more robust than herpes simplex virus thymidine kinase (*hsv-tk*) (*40*) over multiple rounds of selection. When optimizing the negative selection for *tdk*, we circumvented the use of dP because it is a potent mutagenic reagent (*40*), and instead used AZT, a chain-terminator that results in failure of cell division at high doses

(*41*) (Fig. 2A and B; see the Supplement for details.) *E. coli tdk* is also attractive because it can be used in a wide range of organisms (*42, 43*), and can maintain utility over multiple rounds without requiring extensive strain modifications, as was required to optimize for *tolc* selection (39).

By using *tdk*, we developed a protocol to reset the strain after addition of the recoded segment 47R (*5*) (Fig. 1). Initially, we experienced difficulty eliminating the integrated pYES1L vector sequences containing the F replication origin. By using MAGE and negative selection of the adjacent *tdk* marker, we were able to delete the plasmid backbone. The region containing the other attachment site was easily removed by swapping with *tdk* and introduction of a new *attB* site. We subsequently integrated segment 43R at the new *attB* site adjacent to 47R.

When we tested growth of the strain containing the adjacent 43R and 47R segments, we found the normalized generation time was significantly slower than strains containing either sequence alone (Fig. 3A and B). This result motivated us to consider the importance of context in strain development and re-examine alternative routes to assembly of the recoded segments. As a test case, we chose segment 16R (*5*) for isogenic replacement in a *recA1* host. Using the CAGE 2.0 protocol, we first integrated segment 16R on the genome adjacent to the natural locus. Attempts to eliminate the corresponding wild type region together with *tdk* and *attL* using MAGE and negative selection resulted in strains exhibiting recombination between the natural and recoded segment. As an alternative, we first removed the wild type segment by λ Red-directed

exchange for a chloramphenicol resistance cassette and then deleted *tdk* together with vector and *attL* sequences using MAGE. Subsequent removal of *attR* site with *tdk* selection enabled the isogenic replacement to be completed.

Scarless replacements are potentially important for preserving regulatory features in adjacent loci that could affect the expression or function of genes within recoded segments that are required to maintain strain fitness. Analysis of the growth curve of the segment 16R isogenic replacement revealed a growth rate similar to the parental wild-type strain, suggesting isogenic replacements that preserve the genome context of the segment are useful intermediates for subsequent assemblies (Fig. 3C). The integration of recoded segments at natural loci is a key consideration for optimizing strain phenotypes and will eventually be a prerequisite for fully isogenic construction of *rE.coli-57*.

Discussion

The generation of RROs on practical timescales requires improved technologies for assembling modified genomes from large synthetic DNA sequences. These methods would ideally offer the ability to easily transfer large segments of recoded DNA into host organisms and efficiently integrate them into the genome at their corresponding locus. A primary obstacle with many current techniques in *E. coli* is the occurrence of genetic recombination between wild type and recoded sequences after delivery into host strains. As we demonstrated in our testing pipeline, one way to circumvent this problem is to retain the recoded segment episomally in a *recA1* strain, while deleting the corresponding segment from the host chromosome prior to integration. Combining this order of operations with CAGE 2.0 may provide a facile route to generate isogenic replacements. As shown previously, this approach requires an additional Cas9 step to eliminate residual free delivery vector following integration of the recoded segment (*5*) and protocols to subsequently remove the Cas9 plasmid (see Supplemental Methods).

As we move towards generation of *rE.coli-57*, hierarchical methods of assembly will dramatically reduce the time to final strain construction compared to sequential addition of 87 segments. For hierarchical assembly using CAGE 2.0, the protocol must support the excision, maintenance, transfer, and integration of increasingly large recoded genome segments. Currently, genomic excision experiments support use of this protocol for at least 8 segments (see Supplemental Methods), indicating this approach has the potential to drastically reduce the number of steps required for hierarchical assembly.

An alternative to the use of the large vectors required for CAGE 2.0 would be the development of a conjugation-based RMCE method, where very large recoded genome regions flanked by orthogonal *attP* sites are transferred from donor to recipient strains and integrated at isogenic loci. This method takes advantage of well-characterized attachment sites commercially utilized in Gateway products, which are capable of unidirectional integration. With CAGE 2.0, many tools for the task have been prototyped, such as *tdk*-mediated insertion and removal of recombinase sites, integration that is simultaneous with conjugation, and strong markers that select for the delivery of conjugated DNA into the recipient strain. A potential hurdle to overcome in developing this combined approach is the possibility of undesired recombination observed previously using RMCE in recombinogenic (*recA*$^+$) *E. coli* (*28*), and which potentially can occur due to reversions from *recA1* to *recA*$^+$. An ideal host strain for RMCE + CAGE 2.0 would exhibit a low frequency of unwanted recombination, favoring predictable and selectable isogenic replacements.

We have shown that CAGE 2.0 in *E. coli* can be facilitated by using *recA1* host strains that reduce undesired recombination events. Performing this protocol in other organisms where recombination is essential or where each modification is time consuming will provide a greater challenge. In these cases, achieving isogenic replacements will likely require deletion of the corresponding wild type locus before recoded segments are integrated. If double crossovers can be successfully minimized, then backgrounds that enable recombination could offer advantages for organisms that are normally resilient to modification. A potential approach would be to partially delete the wild type sequences

(before segment integration), leaving the flanking homology arms intact, allowing for subsequent integration screening using *tdk* or fluorescent marker expression. In this case, the presence of cellular recombination pathways may offer an endogenous route to strain reset. This would be desirable in hard to modify strains such as mammalian cells, where advance deletion of one chromosomal fragment using CRISPR (*44, 45*) (i.e., before the new sequence in introduced into the cell) may not be disruptive to the cell (due to the presence of multiple chromosomes) and where bacterial artificial chromosome (BAC) libraries provide an engineerable substrate that contains large homologous regions.

As one begins to modify BACs for genome project write (GP-Write) (*31*) and deliver them to mammalian cells or other organisms, the CAGE 2.0 protocol may be directly useful in a number of ways. Conjugal integration provides a route to reduce BAC vectors previously utilized for genome sequencing to single copy and proceed with either MAGE modification or replacement of synthesized regions assisted by positive and negative *tdk* selections, within a strain suitable for BAC maintenance (*38*). After modification, these vectors can be excised or potentially transferred to other strains as genome integrants using RP4 conjugation (*15*). If necessary, several segments can be assembled together using the same protocol prior to transfer to another strain. The ability to modify and transfer BAC vectors and genomes into mammalian cells and other organisms from biocontained *E. coli* (*1,3*) is even more advantageous, offering a method that maximizes biosafety.

The CAGE 2.0 protocol described here further extends the technologies available for constructing RROs. Our results highlight potential opportunities and challenges for each in recoded genome assembly (Fig. 4). Both for these protocols and as one moves towards conjugation and integrase assisted genome engineering in *E. coli*, it will be desirable to further minimize recombination proficiency, which may allow for automation of the protocol. Recently, we have utilized Cas9 to eliminate strains that exhibit undesired mutations, which could assist in minimizing *recA1* reversion (*46*). With CAGE 2.0, we have generated a new set of tools that expands current capabilities in genome construction, and are now developing RMCE + CAGE 2.0 to facilitate large-scale hierarchical approaches to RRO assembly.

Acknowledgments

This research was supported by Department of Energy grant DE-FG02-02ER63445 to G.M.C., Synthetic Biology Engineering Research Center from the National Science Foundation SA5283-11210 to G.M.C., DARPA contract HR0011-13-1-0002 to G.M.C., NSF grant number 106394 to B.L.W., Summer Honors Undergraduate Research Program Grant to E.A., and Origins Summer Undergraduate Research Grant to M.M. We thank Timothy Henion for writing assistance on this manuscript

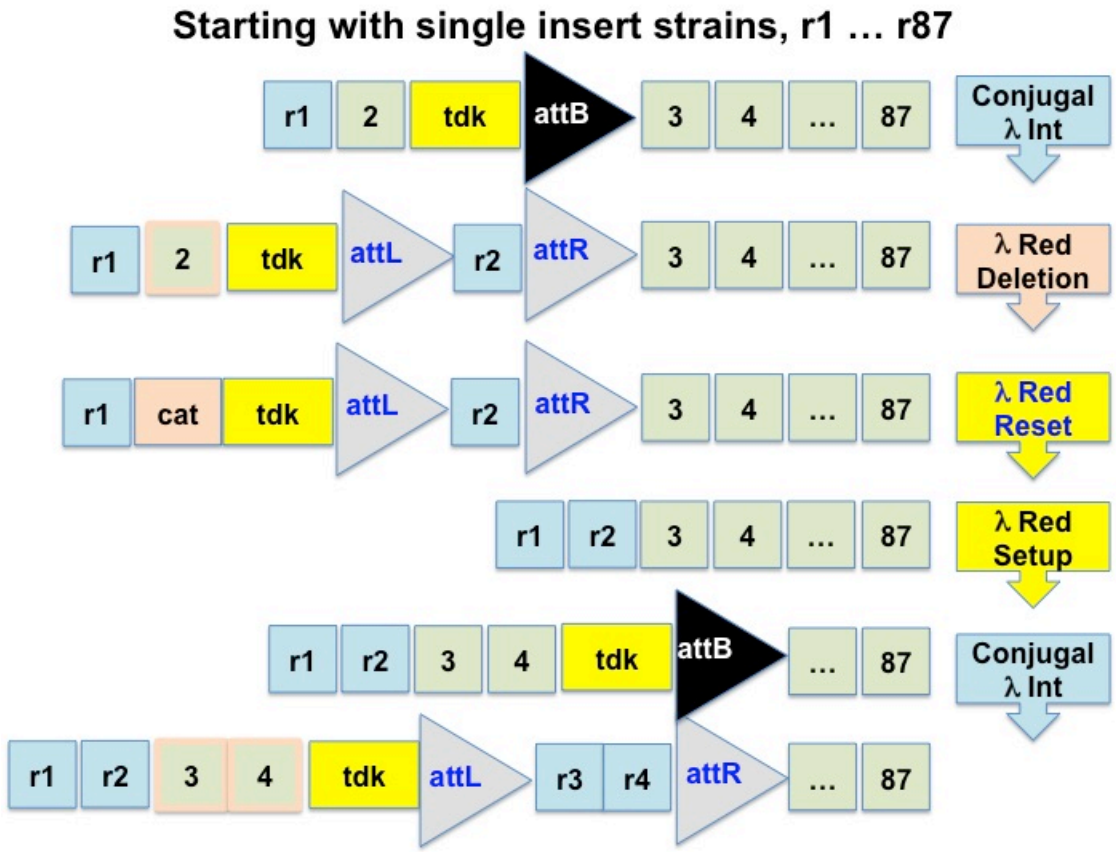

Figure 1. Strategy for CAGE 2.0. **Conjugal λ Integration** of the vector into the recipient strain at single copy. The recipient strain contains the dual selective marker *E. coli* thymidine kinase (*tdk*) and λ integrase *attB* site to the right of the chromosomal sequence that will be replaced. **λ Red Deletion** of chromosomal sequence corresponding to the recoded segment using a chloramphenicol cassette (*cat*). **λ Red Reset Step:** The regions surrounding the recoded segment are removed using *tdk* cassettes and MAGE oligos. **λ Red Setup Step:** To prepare for addition of an additional segment, a cassette containing *tdk* and *attB* is added to the chromosome.

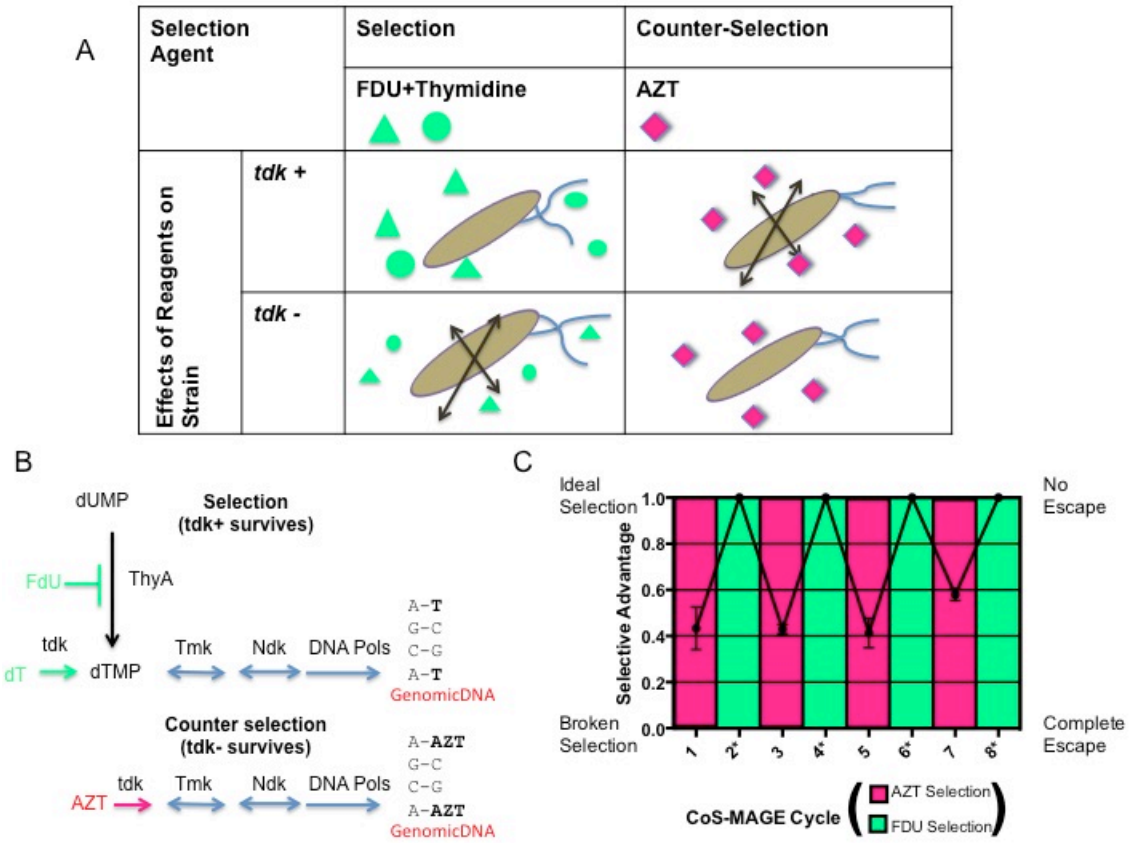

Figure 2. Repetitive *tdk* selection and counterselection maintains a stable selective advantage over multiple cycles (A, B, and C). A) Using a selection reagent FDU utilized for herpes simplex virus thymidine kinase (*HSV-tk*) (*40*), but replacing the mutagenic counterselection reagent dP (6-(β-D-2-deoxyribofuranosyl)-3,4-dihydro-8*H*-pyramido[4,5-c][1,2]oxazin-7-one) with the chain terminator AZT, we developed selection and counterselection conditions for *tdk* using commercially available reagents. B) Selection Mechanism of Action: *Tdk* rescues strains in the presence of FDU, which blocks thymidine salvage. Counterselection Mechanism of Action: The presence of *tdk* allows AZT, a chain terminator that inhibits DNA synthesis, to be incorporated into DNA after phosphorylation with *tdk* (*47, 48*). C) To test the robustness of reuse of the *E. coli tdk* gene over multiple cycles in the same strain, we tested the maintenance of the strain

to selective advantage during co-transformation with a set of MAGE oligos designed to implement modifications over 10 genomic locations.

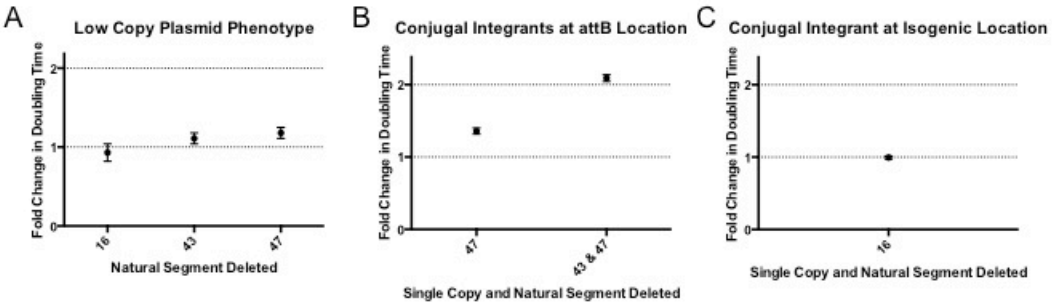

Figure 3. Phenotypic analysis of adjacent and isogenic recoded segments. A) Phenotype established by transformation of recoded segment vector and deletion of the corresponding wild type region (*5*). B) Phenotype established by iterative assembly of two segments at a non-natural location. C) Phenotype established for an isogenic segment replacement.

Method	Scale of Transfer	Advantages and Limitations
CAGE (*32*)	-A genome scale technology -Conjugation removes segment size limit restriction of electroporation	-Validated method for hierarchical genome assembly -Allows scarless recoded genome replacements -Restricted to RecA+ strains, which increases risk of double crossovers
Electroporation-based segment delivery and λ-integrase-mediated single site recombination (5) extended by the strategies for strain reset and setup shown in this work.	-Vectors containing segments and segment assemblies of up to 6x50 kb in size -Validated for single 50 Kb segments -Extending beyond single segments would require dual-selective markers and techniques utilized in this work -Restricted by practical size limit for electroporation (200-300 Kb).	-Efficient method for genomic addition of single segments -Ideal as a "testing pipeline" for assessing function of recoded genes -Does not require removal of strain markers after segment extraction -Preventing recombination is advantageous, possibly limiting strain choice -Size limitations make genome scale recoding time consuming
CAGE 2.0 (excision of segments as vectors and delivery by conjugation and single site λ recombination) as described in this work	-Vectors containing segments and segment assemblies of at least 8x50 kb in size -Validated for two 50 Kb segments -Excision tested for up to 8 segment genomic region -F vectors may greatly increase capacity (up to 2.6 Mb are reported)	-Validated method for scarless and isogenic segment replacements -Allows for recoded segment testing in natural context -Expandable to direct capture of multiple segments, limited only by requirement to excise and maintain large vector sequences -Conjugation and maintenance of multiple segments may require addition of strong selection markers -Preventing recombination is advantageous, possibly limiting strain choice
RMCE + CAGE 2.0 (delivery of conjugated genome and two site recombination)	-Theoretically a genome-scale technology, though single site methods may be more streamlined for smaller segments -Unrestricted by vector capacity	-Theoretically enables hierarchical assembly and rapid strain construction -Two site unidirectional λ integration validated by Gateway® technology -High frequency of unwanted RecA-mediated artifacts in early tests -An ideal option for larger segment assemblies if recombination problems can be solved

Figure 4. Survey of methodologies for assembling radically recoded genomes using 50 kb segments.

References

1. D. J. Mandell *et al.*, Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55-60 (2015).
2. N. J. Ma, F. J. Isaacs, Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell Syst.* **3**, 199-207 (2016).
3. M. J. Lajoie *et al.*, Genomically recoded organisms expand biological functions. *Science* **342**, 357-360 (2013).
4. M. J. Hammerling *et al.*, Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. *Nat. Chem. Biol.* **10**, 178-180 (2014).
5. N. Ostrov *et al.*, Design, Synthesis and Testing Towards a 57-Codon Genome. *Science* **353**, 819-822 (2016).
6. M. O'Connor, M. Peifer, W. Bender, Construction of large DNA segments in *Escherichia coli*. *Science* **244**, 1307-1312 (1989).
7. H. Shizuya *et al.*, Cloning and stable maintenance of 300-kilobasae-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc. Natl. Acad. Sci. USA* **89**, 8794-8797 (1992).
8. Y. Sheng, V. Mancino, B. Birren, Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Res.* **23**, 1990-1996 (1995).
9. J. Song, J. M. Bradeen, S. K. Naess, J. P. Helgeson, J. Jiang, BIBAC and TAC clones containing potato genomic DNA fragments larger than 100 kb are not stable in Agrobacterium. *Theor. Appl. Genet.* **107**, 958-964 (2003).

10. H. H. Wang *et al.*, Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).

11. H. H. Wang *et al.*, Genome-scale promoter engineering by coselection MAGE. *Nat. Methods* **9**, 591-593 (2012).

12. P. A. Carr *et al.*, Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Res.* **40**, e132 (2012).

13. F. van Vliet, B. Silva, M. van Montagu, J. Schell, Transfer of RP4::mu plasmids to Agrobacterium tumefaciens. *Plasmid* **1**, 446-455 (1978).

14. J. M. Alonso, A. N. Stepanova, Arabidopsis transformation with large bacterial artificial chromosomes. *Methods Mol. Biol.* **1062**, 271-283 (2014).

15. V. L. Waters, Conjugation between bacterial and mammalian cells. *Nat. Genet.* **29**, 375-376 (2001).

16. T. Kunik *et al.*, Genetic transformation of HeLa cells by Agrobacterium. *Proc. Natl. Acad. Sci. USA* **98**, 1871-1876 (2001).

17. M. Kalkum, R. Eisenbrandt, R. Lurz, E. Lanka, Tying rings for sex. *Trends Microbiol.* **10**, 382-387 (2002).

18. C. Lartigue *et al.*, Genome transplantation in bacteria: changing one species to another. *Science* **317**, 632-638 (2007).

19. D. G. Gibson *et al.*, Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science* **319**, 1215-1220 (2008).

20. M. A. Algire *et al.*, New selectable marker for manipulating the simple genomes of Mycoplasma species. *Antimicrob. Agents Ch.* **53**, 4429-4432 (2009).

21. C. Lartigue *et al.*, Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**, 1693-1696 (2009).

22. G. A. Benders *et al.*, Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.* **38**, 2558-2569 (2010).

23. D. G. Gibson *et al.*, Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52-56 (2010).

24. H. H. Wang, Synthetic genomes for synthetic biology. *J. Mol. Cell Biol.* **2**, 178-179 (2010).

25. C. Tagwerker *et al.*, Sequence analysis of a complete 1.66 Mb Prochlorococcus marinus MED4 genome cloned in yeast. *Nucleic Acids Res.* **40**, 10375-10383 (2012).

26. B. J. Karas *et al.*, Direct transfer of whole genomes from bacteria to yeast. *Nat. Methods* **10**, 410-412 (2013).

27. C. N. Santos, D. D. Regitsky, Y. Yoshikuni, Implementation of stable and complex biological systems through recombinase-assisted genome engineering. *Nat. Commun.* **4**, 2503 (2013).

28. R. Krishnakumar *et al.*, Simultaneous non-contiguous deletions using large synthetic DNA and site-specific recombinases. *Nucleic Acids Res.* **42**, e111 (2014).

29. C. N. Santos, Y. Yoshikuni, Engineering complex biological systems in bacteria through recombinase-assisted genome engineering. *Nature protocols* **9**, 1320-1336 (2014).

30. M. C. Bassalo *et al.*, Rapid and Efficient One-Step Metabolic Pathway Integration in E. coli. *ACS Synth. Biol.* **5**, 561-568 (2016).

31. J. D. Boeke *et al.*, GENOME ENGINEERING. The Genome Project-Write. *Science* **353**, 126-127 (2016).

32. F. J. Isaacs *et al.*, Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333**, 348-353 (2011).

33. B. M. Barton, G. P. Harding, A. J. Zuccarelli, A general method for detecting and sizing large plasmids. *Anal. Biochem.* **226**, 235-240 (1995).

34. R. A. Meyers, *Molecular biology and biotechnology: a comprehensive desk reference. Physical maps of human chromosomes.* (John Wiley & Sons, 1995).

35. C. J. Gregg *et al.*, Rational optimization of tolC as a powerful dual selectable marker for genome engineering. *Nucleic Acids Res.* **42**, 4779-4790 (2014).

36. T. Conway *et al.*, Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *MBio* **5**, e01442-01414 (2014).

37. D. Yu *et al.*, An efficient recombination system for chromosome engineering in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **97**, 5978-5983 (2000).

38. E. C. Lee *et al.*, A Highly Efficient *Escherichia coli*-Based Chromosome Engineering System Adapted for Recombinogenic Targeting and Subcloning of BAC DNA. *Genomics* **73**, 56-65 (2001).

39. C. J. Gregg *et al.*, Rational optimization of tolC as a powerful dual selectable marker for genome engineering. *Nucleic Acids Res.* **42**, 4779-90 (2014).

40. Y. Tashiro, H. Fukutomi, K. Terakubo, K. Saito, D. Umeno, A nucleoside kinase as a dual selector for genetic switches and circuits. *Nucleic Acids Res.* **39**, e12 (2011).

41. L. P. Elwell *et al.*, Antibacterial activity and mechanism of action of 3'-azido-3'-deoxythymidine (BW A509U). *Antimicrob. Agents Ch.* **31**, 274-280 (1987).

42. V. de Berardinis *et al.*, A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1. *Mol. Syst. Biol.* **4**, 174 (2008).

43. S. Chakraborty, N. Christoforou, A. Fattahi, R. W. Herzog, K. W. Leong, A robust strategy for negative selection of Cre-loxP recombination-based excision of transgenes in induced pluripotent stem cells. *PLoS One* **8**, e64342 (2013).

44. S. M. Byrne, L. Ortiz, P. Mali, J. Aach, G. M. Church, Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.* **43**, e21 (2015).

45. J. F. Li *et al.*, Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688-691 (2013).

46. A. Chavez *et al.*, Precise Cas9 targeting enables genomic mutation prevention *BioRxiv* 058974, (2016).

47. L. P. Elwell, in *Molecular Genetics of Bacterial Pathogenesis,* V. L. Miller, J. B. Kaper, D. A. Portnoy, R. R. Isberg, Eds. (American Society for Microbiology, Washington, D.C., 1994), chap. 2, pp. 17-42.

48. D. Metzgar *et al.*, Acinetobacter sp. ADP1: an ideal model organism for genetic analysis and genome engineering. *Nucleic Acids Res.* **32**, 5780-5790 (2004).