Precise manipulation of bacterial chromosomes by conjugative assembly genome engineering

Natalie J Ma[1,2], Daniel W Moonan[2,3] & Farren J Isaacs[1,2]

[1]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA. [2]Systems Biology Institute, Yale University, West Haven, Connecticut, USA. [3]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, USA. Correspondence should be addressed to F.J.I. (farren.isaacs@yale.edu).

Published online 4 September 2014; doi:10.1038/nprot.2014.081

Conjugative assembly genome engineering (CAGE) is a precise method of genome assembly using conjugation to hierarchically combine distinct genotypes from multiple *Escherichia coli* strains into a single chimeric genome. CAGE permits large-scale transfer of specified genomic regions between strains without constraints imposed by *in vitro* manipulations. Strains are assembled in a pairwise manner by establishing a donor strain that harbors conjugation machinery and a recipient strain that receives DNA from the donor. Within strain pairs, targeted placement of a conjugal origin of transfer and selectable markers in donor and recipient genomes enables the controlled transfer and selection of desired donor-recipient chimeric genomes. By design, selectable markers act as genomic anchor points, and they are recycled in subsequent rounds of hierarchical genome transfer. A single round of CAGE can be completed in a week, thus enabling four rounds (hierarchical assembly of 16 strains) of CAGE to be completed in roughly 1 month.

INTRODUCTION

CAGE is a method that permits the construction of chimeric genomes from many distinct strains. CAGE is rooted in conjugation, a direct cell-to-cell contact that drives transfer of DNA and acts as an important mechanism for horizontal gene transfer in bacteria[1–3]. The machinery required for conjugation is often encoded on self-transmissible conjugative plasmids (F plasmids) that use conjugation as a means for self-dissemination, and it includes the proteins for recognition and capture of recipient strains (F pilus), formation of cell-to-cell contact for DNA transfer (mating apparatus) and transfer of plasmid DNA[4,5]. When a cell containing the conjugative plasmid (F^+, donor cell) recognizes a recipient cell that lacks the conjugative plasmid (F^-), it binds the recipient cell using the F pilus to form the mating apparatus[6] (**Fig. 1a**). The conjugative plasmid in the donor strain is nicked at the origin of transfer site (*oriT*), and a single strand of the plasmid is then transferred to the recipient strain through the mating apparatus in a direction-dependent manner[5]. The *oriT* site possesses a sequence-dependent directionality of transfer, resulting in the transfer of the same strand starting at the 5′ end during each conjugal event[7,8]. Upon entry into the cell, the complement strand is synthesized and the plasmid recircularizes, thus converting the recipient bacterium into an F^+ donor capable of transmitting the conjugative plasmid to other recipient strains that lack the same conjugative plasmid[4,6]. Although conjugative plasmids are often independent genetic elements, high-frequency recombination (Hfr) strains exist wherein conjugal elements are integrated into the *E. coli* genome[2,6]. These Hfr strains possess the same machinery as strains containing conjugative plasmids, but during conjugation donors transfer a strand of their genome in a direction-dependent manner to the recipient cell beginning at a genomic *oriT* site[2,9] (**Fig. 1b**). Once the complement of this strand is synthesized, the transferred genomic DNA from the donor strain can undergo recBCD-mediated recombination with the recipient genome, replacing part or all of the recipient genome[2].

CAGE differs from natural plasmid and Hfr conjugation in two key aspects: first, it decouples the transfer of genetic information from the conjugative machinery, and second it uses selectable markers to precisely determine the composition of the final chimeric genome. CAGE decouples the transfer of genetic information from the conjugal machinery through site-specific placement of the *oriT* in the genome, whereas the conjugative machinery is encoded on a separate plasmid (pRK24), enabling rapid conversion of strains between donor and recipient status[4,10]. During conjugation, genomic transfer initiates at the genomic *oriT* and transfers a strand of the donor genome independent of the conjugative machinery located on pRK24 (**Fig. 1c**). It is hypothesized that the complement is then synthesized upon entry into the recipient, producing a double-stranded copy of donor DNA that undergoes recBCD-mediated recombination with the recipient genome[2]. The result is a recipient cell containing a chimeric genome and the pRK24 conjugative plasmid (**Fig. 1c**). If pRK24 is not actively maintained through selection with antibiotics, it is lost from most cells in the population after outgrowth and selection for recombinant genomes (see below), effectively enabling easy conversion of chimeric strains to recipient status in subsequent rounds of CAGE.

Although conjugation enables the transfer of genetic information, isolation of chimeric genomes possessing the desired portions of donor and recipient genomes is possible in CAGE only through the precise placement of selectable markers, which is achieved by λ-Red recombination of double-stranded DNA[11,12]. Choice of marker placement is important for three main reasons. First, site-specific placement of markers establishes boundaries on genomic regions to be transferred from the donor and the genomic regions to be retained in the recipient after conjugation (**Fig. 2**). Second, the choice of marker placement is guided by successive rounds of CAGE because the selectable markers in one round serve as anchor positions for insertion of *oriT* and other markers in subsequent rounds of CAGE (**Figs. 2b** and **3**). Third,

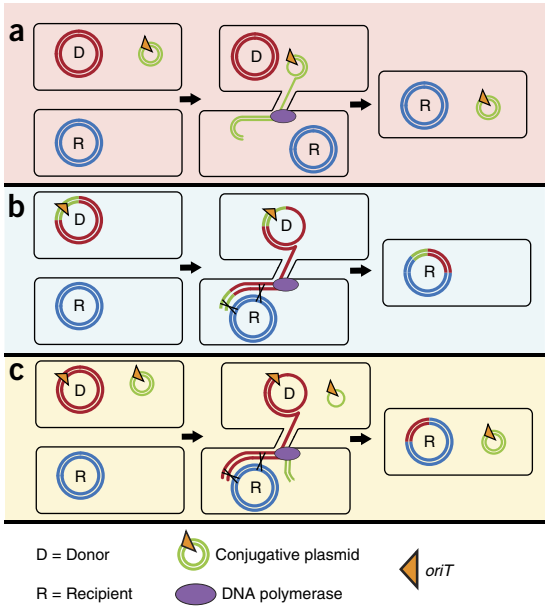

Figure 1 | Comparison of conjugation in plasmids, Hfr strains and CAGE. CAGE differs from natural mechanisms of conjugation (F plasmids and Hfr strains) in the decoupling of the origin of transfer (*oriT*) from the conjugal machinery. (**a**) In plasmid-based systems, the conjugative machinery transfers itself to recipients in the form of a plasmid carrying the *oriT*. (**b**) Hfr strains contain all conjugal elements and *oriT* on the genome, which permits transfer of a part of the genome or the whole genome. (**c**) In CAGE, site-specific integration of *oriT* in the donor genome enables transfer of genomic information independent of the conjugative machinery. The donor genome is transferred with the *oriT* as the transfer initiation point, whereas placement of selectable markers in donor and recipient genomes determines the genetic composition of the chimeric genome (**Figs. 2** and **3**).

because chimeric strains will inherit both positive markers (P_1 and P_2; **Fig. 2b**), markers should be stably integrated into regions that will not reduce cellular fitness. We have identified several safe insertion regions (SIRs) within *E. coli* MG1655 that do not contain any known genes or regulatory regions (**Supplementary Table 1**). To calculate recombination efficiencies at each locus, we performed recombineering into SIRs with selectable markers, counted the total number of colony-forming units with antibiotic resistance after recovery and then divided by the estimated total number of cells in the initial population[10]. The *oriT-kan* and positive-negative selectable marker (P/N in **Fig. 2b**) may also be placed at SIRs, but because these markers are not inherited by the chimeric strain they may also be placed in non-essential genes that do not cause fitness impairment in laboratory conditions.

To control the direction and start site of conjugal transfer from the donor, the *oriT* sequence is integrated into the donor strains at a minimum of 3 kb upstream of the desired fragment in the genome (**Fig. 3**). The *oriT* is fused with a kanamycin gene resistance (*kan*) cassette to permit site-specific placement of the *oriT* at any SIR in the genome. When located on the genome, the transfer of a 100-kb region occurs at an efficiency of 10^{-2} to 10^{-3}, whereas the transfer of 2.3 Mb (half of the genome) can occur at an efficiency of 10^{-6}, enabling isolation of multiple chimeric genomes from a diverse population of $\geq 10^9$ cells present in culture[10]. To maximize this efficiency, the *oriT-kan* cassette should be oriented away from the desired recipient fragment to provide the maximum transfer distance between donor and recipient fragments in which recombination can occur (**Fig. 3**). Because the transfer start site within the *oriT-kan* is downstream of the *kan* cassette, the *kan* resistance gene is not transferred during conjugation. The *oriT* sequence that is transferred at the start of the donor fragment is lost during recombination with the recipient genome owing to lack of homology (**Fig. 2b**). Loss of the *oriT-kan* in the resulting chimeric genome facilitates the immediate re-introduction of the *oriT-kan* in place of a positive marker (described below) to position the strain for the next round of CAGE.

While *oriT* determines the start of conjugal transfer, two positive selectable markers (e.g., $spec^R$, zeo^R or $gent^R$; represented by P_1 and P_2 in **Figs. 2b** and **3**) ensure the inheritance of both desired donor and recipient regions in chimeric genomes. In the donor, the desired genomic region for transfer is flanked by an *oriT* at the 5′ end and a positive selectable marker at the 3′ end to select for transfer of the full fragment. The recipient strain contains a second positive marker positioned at the start of the desired recipient genomic fragment. Selection for this marker enriches for chimeric strains that have retained the desired recipient portion of the genome after conjugal transfer with the donor.

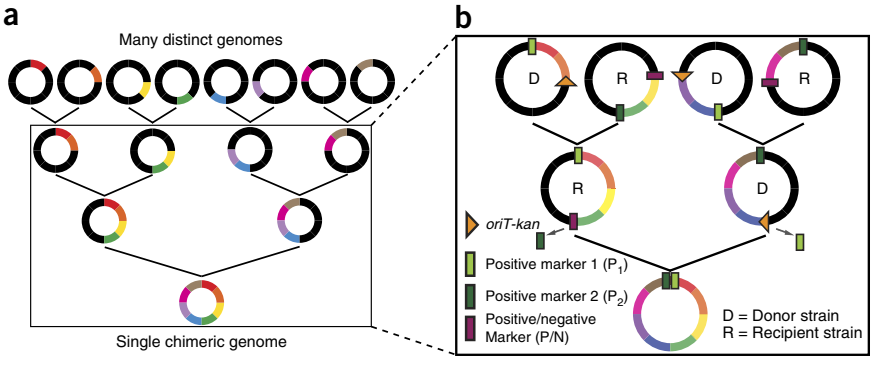

Figure 2 | Overview of CAGE. CAGE assembles regions of many distinct genomes into a single chimeric genome, combining many mutations in multiple *E. coli* strains into a single strain. (**a**) As an example, genome fragments across eight strains are assembled into a single strain through three iterative conjugations of donor-recipient pairs. CAGE genome assemblies are performed in a pairwise manner to permit hierarchical assembly. The number of rounds of CAGE (n) required to assemble G genomes together is derived from the equation $G = 2^n$. For example, if there are fragments in eight genomes to assemble, three rounds of CAGE are required. (**b**) Shown are the marker placements that enable precise manipulation of genomic fragment assembly for the final two conjugation events. In the donor, the *oriT* appended to kanamycin resistance (*oriT-kan*) is located upstream of the desired fragment and positive marker 1 (P_1) is located downstream of the desired fragment. In the recipient, positive marker 2 (P_2) is located upstream of the desired fragment and the P/N marker is located ~3 kb downstream of the site of *oriT-kan* in the donor genome. When the strains are conjugated, the *oriT-kan* is lost and the P/N is overwritten, permitting facile recycling of these markers in preparation for the next round of CAGE.

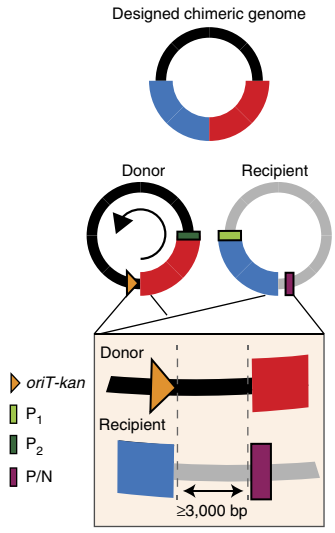

Figure 3 | Placement of *oriT-kan* and P/N markers to ensure integrity of the desired genomic regions. To increase the frequency of desired chimeric genomes, the *oriT-kan* should be placed in the direction that maximizes the distance between the end of the donor fragment and start of the recipient fragment. Transfer thus proceeds away from the junction where the donor and recipient fragments meet (arrow within donor genome), maximizing the distance in which recombination can occur. The *oriT-kan* should also optimally be placed in the donor strain at a minimum of 3 kb from the start of the desired donor fragment. In the recipient strain, the P/N marker should be placed at least 3 kb downstream of the *oriT-kan* site in the donor genome, so it will be overwritten by the donor genome during conjugation and recombination in CAGE assembly.

Thus, selection for both markers ensures that recombination has occurred between the two selectable markers in surviving cells, preserving both the desired donor and recipient fragments in the chimeric genome (**Fig. 2b**). By using these two markers, we find that chimeric genomes can be assembled from fragments with as little as 3 kb of distance between them, which provides sufficient distance for homologous recombination to occur . However, resultant chimeric genomes may show loss of early portions of the donor fragment (i.e., genomic region immediately 3′ to *oriT* sequence fragment), suggesting that recombination between the transferred donor genome fragment and recipient genome may occur over a large region following the *oriT*.

To select for recombinant cells where the donor genomic fragment has recombined in close proximity (roughly 3–10 kb) of the *oriT*, a P/N selectable marker (such as *tolC*, *galK* or *thyA*; represented by P/N in **Figs. 2b** and **3**) is introduced in the recipient genome at the genomic locus immediately 3′ to the *oriT* locus in the donor genome. Negative selection against the P/N selectable marker after conjugation ensures that the remaining chimeric genomes possess the full-length donor fragment (**Fig. 3**). In contrast to positive selections, which use antibiotics to select for incorporation of an antibiotic resistance marker, negative selections select for loss of function of a gene in cells through introduction of an agent that targets the gene product and causes cell death. This protocol has been optimized for use with the *tolC* gene[10], which is placed in the recipient at a minimum of 3 kb downstream from where the *oriT* is located in the donor. During conjugation, chimeric genomes lose *tolC* if recombination occurs between the *oriT* and the start of the desired donor fragment, thus possessing the entire desired donor fragment, but they retain *tolC* if recombination occurs within the donor fragment region, resulting in partial loss of desired donor fragment (**Fig. 3**). After conjugation, positive selection for P_1 and P_2 enriches for genomes that have incorporated the donor fragment without overwriting any of the recipient fragment. The subsequent negative selection against the presence of functional tolC using colicin E1 eliminates all cells that have retained *tolC* owing to

Box 1 | Negative selection of P/N marker *tolC* ● TIMING 18 h

This CAGE protocol has been optimized for the use of P/N marker *tolC*. TolC is a transmembrane protein that confers resistance to SDS but also permits cell entry of colicin E1, a small toxic peptide produced by *E. coli* strain JC411 [ColE1] to reduce intraspecific competition[13,14]. Deletion of the *tolC* gene prevents import of colicin E1 and confers resistance to the colicin E1 toxin. However, because loss of function can result not only from full deletion of *tolC* but also from nonsense and frameshift point mutations within the gene, negative selections are time-dependent, requiring the user to enrich the population of cells that have lost the gene before 'escape' mutants containing point mutations within the gene overtake the population. The following procedure is used for *tolC* negative selection. Protocols for other P/N markers (e.g., *galK* and *thyA*) exist[22,23].

1. Prepare colicin E1 according to refs. 13,14, or 20.
2. After the deletion of *tolC* by recombineering or CAGE, grow the resulting strain for at least 6 h to overnight to eliminate any residual tolC protein in the cell membrane. Simultaneously, start a culture of its ancestral strain containing functional *tolC* as a control. Grow both cultures in an incubator at 30–34 °C with shaking at 300 r.p.m.
3. Dilute both experiment and control cultures to 3 µl in 3 ml of fresh LB in a 14-ml culture tube.
4. Grow the cultures to early log stage (0.1 or 0.2 OD at 600 nm).
5. Remove colicin E1 from the –80 °C freezer and place it on ice to thaw.
6. For each of the experiment and control strains, mix 1.5 µl of colicin E1, 150 µl of LB and 1.5 µl of culture in separate wells of a 96-well flat-bottom plate.
7. Place the plate in a shaking plate reader on the kinetic setting for 6–10 h with reads every 10 min under 282 c.p.m. orbital shaking to monitor the cell growth. If a plate reader is unavailable, place the plate in a shaking incubator under the conditions from step 2 and check for growth every hour.
8. When the *tolC*-deleted cultures reach early mid-log, as indicated by the OD_{600} value, but the control culture that contains *tolC* has not yet increased in OD_{600} values, remove the plate from the plate reader and plate serial dilutions on agar to isolate single colonies for screening.

recombination within the donor fragment[13,14] (**Fig. 2b** and **Box 1**). This ensures that the remaining chimeric genomes have recombined upstream of the start of the desired donor fragment and thus possess the entire desired donor fragment. Furthermore, loss of *tolC* in the chimeric genome enables immediate re-introduction into recipient strains of the next CAGE round in place of one of the positive selectable markers.

Currently, CAGE has been paired with multiplex automated genome engineering (MAGE)[15] for *in vivo* genome recoding of an *E. coli* strain lacking TAG codons[10,16]. CAGE enabled parallel assembly of 32 *E. coli* genome fragments wherein all TAG codons had been converted to TAA using MAGE. Because of its hierarchical nature, assembly of these 32 fragments after MAGE occurred in parallel with five steps of CAGE[16]. In contrast to *de novo* genome synthesis, the *in vivo* nature of MAGE and CAGE provides component testing before final assembly and removal of lethal genotypes from the population. In addition, CAGE can be used to overwrite deleterious genotypes resulting from spontaneous mutations that appear in Δ*mutS* strains during MAGE[16]. Together, CAGE and MAGE offer an *in vivo* alternative for genome editing of many chromosomes instead of *de novo* synthesis of a single genome. We also envision that CAGE can be used to combine and randomize pathway and genome components distributed across multiple genomic regions.

Although CAGE is currently limited to *E. coli*, we hypothesize that similar methods of genome assembly may be possible owing to the broad host range and transfer capabilities of IncPα conjugative plasmids[17–19] (the group to which pRK24 belongs) and other functionally related plasmids, provided that the following requirements are met: (i) conjugal machinery is functional in the donor; (ii) genome recombination mechanisms in the recipient are similar to the recombination mechanisms of *E. coli*; and (iii) functional selectable markers are available. In addition, negative selection steps with *tolC* and other P/N markers have a higher rate of false positives than positive selections, as secondary mutations can render the P/N marker nonfunctional and enable cell survival. Ultimately, increased stringency of P/N markers could reduce or eliminate genetic screening between steps of CAGE to verify correct sequence integrity of chimeric genomes. A recent paper has already demonstrated a negative selection of *tolC* with improved stringency[20]. Such advances could allow for CAGE to be performed exclusively in liquid medium to reduce each stage of the protocol, eliminate genetic screens and facilitate the automation of the entire procedure.

MATERIALS

REAGENTS

- *E. coli* strain ECNR2 (Addgene strain no. 26931; see Isaacs *et al.*[10], Sharan *et al.*[12] and Gregg *et al.*[20])
- *E. coli* strain containing the pRK24 plasmid (Addgene strain no. 51950)
- *E. coli* strain JC411 (ColE1) for colicin E1 preparation (Addgene strain no. 51949)
- *E. coli* strain *oriT-kan* template (Addgene strain no. 51946)
- *E. coli* strain *spec*^{R} template (Addgene strain no. 51948)
- *E. coli* strain *tolC* template (Addgene strain no. 51029)
- Kanamycin (American Bioanalytical, cat. no. AB01100)
- Carbenicillin (American Bioanalytical, cat. no. AB00285)
- Tetracycline (American Bioanalytical, cat. no. AB02024)
- Zeocin (Life Technologies, cat. no. R25001)
- Spectinomycin (MP Biomedicals, cat. no. 0215206701)
- Gentamicin (MP Biomedicals, cat. no. 0219005705)
- SDS (American Bioanalytical, cat. no. AB01922)
- Colicin E1 (see DeVito[13], Schwartz and Helinski[14] and Gregg *et al.*[20] for preparation)
- Sterile H_2O
- Tryptone (American Bioanalytical, cat. no. AB02031)
- Yeast extract (American Bioanalytical, cat. no. AB01208)
- Sodium chloride (NaCl; American Bioanalytical, cat. no. AB01915)
- Agar (American Bioanalytical, cat. no. AB01185)
- Sodium hydroxide, 1 N (1 N NaOH) made from NaOH pellets (JT Baker, cat. no. 3722) **! CAUTION** It is corrosive; use it only with proper personal protective equipment.
- Glycerol (American Bioanalytical, cat. no. AB00751)
- 2G HiFi HotStart ReadyMix (high-fidelity PCR mix; Kapa Biosystems, cat. no. KK2602)
- 2G HotStart ReadyMix with Loading Dye (Kapa Biosystems, cat. no. KK5609)
- 2G multiplex PCR mix (Kapa Biosystems, cat. no. KK5802)
- DpnI (New England BioLabs, cat. no. R0176S)
- PCR purification kit (Qiagen, cat. no. 28104)
- Gel extraction kit (Qiagen, cat. no. 28604)
- Agarose (American Bioanalytical, cat. no. AB00972)
- TBE buffer, 10× (American Bioanalytical, cat. no. AB14024)
- Nonmutagenic DNA stain such as GelGreen (Biotium, cat. no. 41004) or SYBR Safe (Life Technologies, cat. no. S33102)
- DNA ladder (New England BioLabs, cat. no. N3200S)
- Chimeric primers for recombineering (Steps 14 and 15)
- DNeasy blood and tissue kit for genomic DNA preparation (Qiagen, cat. no. 69504)

EQUIPMENT

- ApE (A Plasmid Editor, http://biologylabs.utah.edu/jorgensen/wayned/ape/) or any other DNA sequence manipulation software
- Thermal cycler (Bio-Rad, cat. no. 185-1196)
- 0.2-ml PCR tubes (Bio-Rad, cat. no. TLS0801)

TABLE 1 | Final concentration of selection agents.

Selection agent[a]	Final concentration (μg/ml unless stated otherwise)[b]	Solvent
Ampicillin	50	Water
Carbenicillin	50	Ethanol, 50% (vol/vol)
Chloramphenicol	20	Ethanol
Kanamycin	30	Water
Spectinomycin	95	Water
Tetracycline	12	Ethanol, 70% (vol/vol)
Zeocin	10	Water
Gentamicin	7.5	Water
SDS	0.005% (wt/vol)	Water

[a]All selection agents except SDS are antibiotics. [b]Stock concentrations are made at 1,000× the final concentration.

- Flat caps for PCR tubes (Bio-Rad, cat. no. TCS0803)
- 96-well PCR plates (Thermo Scientific, cat. no. AB-0800-L)
- Plastic films for 96-well PCR plates (Bio-Rad, cat. no. MSB-1001)
- Gel electrophoresis apparatus (Bio-Rad)
- Blue light transilluminator for imaging DNA gels (Clare Chemical Research, cat. no. DR89X)
- Gel dock for imaging large gels (Bio-Rad, cat. no. 1708195)
- Blue light transilluminator for gel dock (Bio-Rad, cat. no. 1708182)
- Stationary incubator set at 30–34 °C (Thermo Scientific, cat. no. SHKE6000)
- Shaking incubator set at 30–34 °C (Thermo Scientific, cat. no. SHKE6000)
- Shaking water bath set at 42 °C (Thermo Scientific, cat. no. SHKE7000)
- Microcentrifuge kept at 4 °C (Eppendorf, cat. no. 5415C)
- Microcentrifuge kept at room temperature (~20–25 °C; Eppendorf, cat. no. 5424)
- Electroporation cuvettes with 1-mm gap (Bio-Rad, cat. no. 165-2089)
- Electroporator (Bio-Rad, cat. no. 1652662 and Harvard BTX, cat. no. 45-2008)
- Sterile 5-ml serological pipettes (Corning, cat. no. 356543)
- Sterile filtration unit (Thermo Scientific, cat. no. 568-0020)
- Cryotubes, 2 ml (Thermo Scientific, cat. no. 5000-1020)
- Sterile Petri dishes, 150 × 15 mm (Fisher Scientific, cat. no. 351058)
- Conical tubes, 50 ml (Corning, cat. no. 430829)
- Sterile 1.7-ml microcentrifuge tubes (GeneMate, cat. no. C-3371-1)
- Sterile 14-ml culture tubes (Corning, cat. no. 352059)
- Sterile 2.0-ml microcentrifuge tubes (Fisher Scientific, cat. no. 02-681-271)
- Spectrophotometer (Biochrom, cat. no. 80-5000-00) and cuvettes (BrandTech Scientific, cat. no. 759086D)
- Sterile 96-well flat-bottom plates with lids (Corning, cat. no. 3595)
- V-bottom 96-well plates (Nunc, cat. no. 249662)
- V-bottom 96-well plate lids (Nunc, cat. no. 264122)
- Plate reader for negative selections (optional; Bio-Tek, cat. no. 11-120-531)

REAGENT SETUP

LB broth, Lennox (LB) Combine 10 g of tryptone, 5 g of yeast extract and 5 g of NaCl in 1 liter of water, adjust the pH to 7.5 by adding 2 ml of 1 M NaOH and autoclave it. Store the medium at room temperature. If it is kept sterile, the medium will last for several months. If selecting agents are required in the medium, filter-sterilize the solutions of selection agents and add them to aliquots of sterile medium at the time of use. Keep the solutions of selection agents at −20 °C indefinitely (**Table 1**).

LB agar plates Combine 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl and 15 g of agar in 1 liter of water; adjust the pH to 7.5 by adding 2 ml of 1 M NaOH and autoclave the medium. Pour molten medium into sterile Petri dishes. If selection agents are required in the medium, filter-sterilize them first and then add them after autoclaved medium has cooled to 50 °C. Agar plates should be stored at 4 °C for up to 3 months.

Glycerol, 80% (vol/vol) Combine 80 ml of glycerol with 20 ml of water, mix it and filter it through a sterile filtration unit. Store it at room temperature. If it is kept sterile, glycerol will last for several months. ▲ **CRITICAL** Glycerol should be kept sterile to maintain frozen stocks of bacterial strains.

Colicin E1 preparation Prepare colicin E1 according to DeVito[13], Schwartz and Helinski[18] or Gregg *et al.*[20]. Store it at −80 °C, and thaw it on ice as needed.

PROCEDURE

Determination of genome assembly order ● TIMING 15 min

1| Draw a schematic of the final chimeric genome.

2| Divide the chimeric regions into two and draw a pair of half-chimeric genomes above the fully chimeric genome.

3| Repeat Step 2 until chimeric regions match regions in currently existing strains. The schematic should resemble **Figure 2a**.

4| In each genome pair, designate one genome as the donor and one genome as the recipient.

5| In donor strains, draw *oriT-kan* at the start of the donor fragment facing a direction that maximizes transfer distance between donor and recipient fragments to maximize CAGE efficiency. Draw the P_1 marker at the end of donor fragment (**Figs. 2b** and **3**).

6| In recipient strains, draw the P/N marker at the start of the region that will be overwritten by the donor strain, outside of the region to retain in each recipient. Draw the P_2 marker on the other side of the recipient fragment to keep (**Figs. 2b** and **3**).

Design of selection marker cassettes and screening primers ● TIMING 1 h

7| Choose an insertion site for the selection marker. See **Figure 3** for guidance on placing markers and **Supplementary Table 1** for SIRs in the *E. coli* MG1655 genome.

8| Use the National Center for Biotechnology Information (NCBI) genome database (http://www.ncbi.nlm.nih.gov/genome) to query the *E. coli* MG1655 genome (tax id no. 511145) for the sequence 500 bp upstream and downstream of the insertion site.

9| Highlight and copy the sequence into the ApE software.

10| Insert an 'N' in the middle of the sequence where a marker will be integrated.

11| Highlight 50 bp upstream of the 'N' and label them as 'Homology Arm'. Repeat this downstream of the 'N'.

TABLE 2 | Universal primers for amplifying selectable markers.

Cassette name	Selection type	Length (bp)	Template	Forward primer (preceded 5′ by a 50-bp-long homology arm)	Reverse primer (preceded 5′ by a 50-bp-long homology arm)
Kanamycin-*oriT*	Positive	1,949	*E. coli* strain rEc2 (refs. 10,16)	5′-CCTGTGACGGAAGATCACTTCG-3′	5′-AGCGCTTTTCCGCTGCA-3′
oriT	NA	784	pRK24	5′-GGCGCTCGGTCTTGCCTT-3′	5′-AGCGCTTTTCCGCTGCA-3′
Kanamycin	Positive	1,165	PBK*KS*plus	5′-CCTGTGACGGAAGATCACTTC-3′	5′-AACCAGCAATAGACATAAGCGG-3′
Chloramphenicol	Positive	1,015	PBC*KS*plus	5′-CCTGTGACGGAAGATCACTTCG-3′	5′-AACCAGCAATAGACATAAGCGG-3′
Gentamicin	Positive	819	PBBR1MCS-5	5′-CCGGGAAGCCGATCTCG-3′	5′-ACGCACACCGTGGAAAC-3′
Spectinomycin	Positive	1,201	Addgene strain no. 51948	5′-CAGCCAGGACAGAAATGC-3′	5′-TGCAGAAATAAAAAGGCCTGC-3′
Zeocin	Positive	762	PTEF1/Zeo	5′-GGTGTTGACAATTAATCATCGGC-3′	5′-AGCTTGCAAATTAAAGCCTTCG-3′
TolC	Positive/negative	1,746	pZE21-tolC	5′-TTGAGGCACATTAACGCC-3′	5′-TCTAGGGCGGCGGATT-3′
GalK	Positive/negative	1,270	*E. coli* strain rEc25 (refs. 10,16)	5′-CCTGTTGACAATTAATCATCGGC-3′	5′-AAAAAAAACCCCGCCCTGT-3′

NA, not applicable.

12| Replace 'N' with a marker cassette sequence. See **Supplementary Data** for markers and *oriT* sequences.
▲ **CRITICAL STEP** The *oriT-kan* cassette should only be integrated into donor strains.

13| Annotate the universal parts of the forward and reverse primers for the marker in the sequence (**Table 2**). The universal parts will always be in the 3′ end of the primer.

14| Highlight both the universal forward primer and its adjacent homology arm sequence, and annotate this as 'F primer full'. Order this sequence as a single-stranded oligonucleotide with desalting.

15| Highlight both the universal reverse primer and its adjacent homology arm sequence and annotate as 'R primer full'. Order the reverse-complement of the sequence as a single-stranded oligonucleotide with desalting.

16| At >150 bp upstream (5′) of the selection gene cassette, highlight a DNA sequence that has a melting temperature of ~59 °C and annotate this as the 'F screening primer.' Order this sequence as a single-stranded oligonucleotide with desalting.

17| At >150 bp downstream (3′) of the cassette, highlight a DNA sequence that has a melting temperature of ~59 °C, and annotate this as the 'R screening primer'. Order the reverse complement of this sequence as a single-stranded oligonucleotide with desalting.

18| Repeat Steps 7–17 for each marker to be integrated.

Amplification and isolation of selection marker DNA ● TIMING 2–3 h

19| Prepare 200 μM storage stocks of the primers and dilute 20 μM aliquots to serve as working stocks. Store both storage and working stocks at –20 °C indefinitely.

20| To generate double-stranded selection marker DNA for recombineering, prepare the DNA template as follows: if the template is plasmid DNA, dilute it to 1 ng/μl in sterile water. If the template is bacterial cells, dilute 1 μl of a confluent culture in 19 μl of sterile water.

21| Assemble the following reaction (optimized for using the Kapa 2G HiFi HotStart ReadyMix) in a 0.2-ml tube:

Compound	Amount per sample	Final concentration
DNA template	1 µl	NA
F primer (20 µM)	1 µl	0.4 µM
R primer (20 µM)	1 µl	0.4 µM
2× PCR mix	25 µl	1×
Water	Up to 50 µl	NA

NA, not applicable.

▲ **CRITICAL STEP** Never vortex the PCR mix, as this may inactivate the polymerase.

22| Cap the tube and perform the PCR with the following conditions (optimized for using the Kapa 2G HiFi PCR mix):

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 3 min			
2–26	98 °C, 20 s	59 °C, 15 s	72 °C, 30 s/kb of desired product	
27			72 °C, 5 min	
28				12 °C

23| Remove the tube from the thermocycler and add 1 µl of DpnI to the reaction to degrade any template DNA. Pipette it up and down to mix the reaction.

24| Return the tube to the thermocycler and run the following program: incubation at 37 °C for 60 min; inactivation of the enzyme at 80 °C for 20 min; and holding at 4 °C.
▲ **CRITICAL STEP** If DpnI digestion is not performed, the presence of residual template DNA may result in false positives during recombineering.

25| Visualize 4 µl of the PCR product on agarose gel, as described by Isaacs *et al.*[10].
? TROUBLESHOOTING

26| If the PCR produced multiple bands, run all of the PCR product on a gel with nonmutagenic DNA stain, cut the desired band out of the gel and purify the PCR product using the gel extraction kit according to the manufacturer's instructions. If the PCR shows only one clear band in Step 25, purify the PCR product directly, to increase yield, using the PCR purification kit according to the manufacturer's instructions. Elute the DNA into sterile water.
▲ **CRITICAL STEP** Never expose DNA intended for recombineering or sequencing to UV light, as this will mutagenize the DNA.
■ **PAUSE POINT** Store DNA for recombineering at 4 °C overnight or at –20 °C indefinitely.

Recombineering selection markers into the genomes of donor and recipient strains ● TIMING 3 d
▲ **CRITICAL** Work in sterile conditions to prevent contamination of strains.

27| Place a 50-ml conical tube filled with sterile water in ice to chill.

28| Use a serological pipette to add 3 ml of LB into a 14-ml sterile culture tube. Prepare one culture tube for each strain to recombineer.

29| Start the cultures of strains to recombineer using one of the following options, depending on the current growth medium of the strain: for colonies growing on agar plates, use Step 29A; for bacteria growing in liquid culture, use Step 29B.
(A) If bacteria are growing as colonies on solid medium
(i) Use a sterile pipette tip (mounted on a pipette) to pick up a colony from the plate.

(ii) Dip the pipette tip in the 3 ml of LB in the culture tube and pipette up and down to dislodge cells from the tip.

(B) If bacteria are growing in liquid culture

(i) Dilute the culture by adding 3 µl of confluent culture into the 3 ml of freshly prepared LB.

30| Grow the cultures in an incubator at 30–34 °C with shaking at 300 r.p.m. until the OD reaches mid-log phase (0.4–0.6 OD at 600 nm).

31| While the cultures grow, prepare the following for each strain: label two empty 1.7-ml microcentrifuge tubes as 'experiment' and 'negative control' and place them on ice; add 150 ng of selection marker DNA from Step 26 to sterile water to a final volume of 50 µl in a 1.7-ml tube and place it on ice; add 50 µl of sterile water to a 1.7-ml microcentrifuge tube as a negative control and place it on ice; label two electroporation cuvettes as 'experiment' and 'negative control' and place them on ice; label two 14-ml culture tubes for bacterial recovery as 'experiment' and 'negative control', pipette 1 ml of LB into each and leave them at room temperature.

32| When the cultures reach mid-log phase, transfer the culture tubes to a shaking water bath at 42 °C for 15 min with shaking at 250 r.p.m. to induce the expression of the λ-Red proteins[10,12], and then immediately transfer them to ice.
▲ **CRITICAL STEP** Do not allow the cells to remain at 42 °C longer than 15 min, as λ-Red proteins are toxic and will compromise cell viability.

33| Transfer 1 ml of the culture to each prechilled empty 1.7-ml tube.

34| Centrifuge the tubes at 13,500*g* for 1 min at 4 °C to pellet the cells.

35| Remove the supernatant from the cell pellet.
▲ **CRITICAL STEP** If any supernatant remains, salts from the medium may reduce the electroporation efficiency or cause arcing, thus reducing recombineering efficiency.

36| Resuspend the cells in 1 ml of chilled sterile water.

37| Repeat Steps 34–36 once.

38| Repeat Steps 34 and 35, and then resuspend the cells in the 50 µl of the prechilled DNA mix prepared in Step 31 and transfer them to the prechilled electrocuvette.

39| Electroporate the cells using the following settings: voltage—1,800 V; capacitance—25 µF; resistance—200 Ω; and cuvette—1 mm.
? TROUBLESHOOTING

40| Immediately aspirate the 1 ml of LB from each 14-ml recovery culture tube and transfer it to an electrocuvette to dilute the cells for recovery.

41| Transfer the 1 ml of LB + cells back to the culture tubes.

42| Repeat Steps 33–41 for the negative control sample, resuspending instead in sterile prechilled water in Step 38.

43| Place the recovery cultures into the shaking incubator at 30–34 °C for 2 h.
? TROUBLESHOOTING

44| Prewarm and label the agar plates containing selection agents to select for the recombineered marker.

45| Plate 50 µl of recovery culture from the 'experiment' culture tube onto an agar plate.

46| Centrifuge the remaining recovery culture in a 1.7-ml tube at 13,500*g* for 1 min at room temperature.

47| Remove the supernatant until ~50 µl remains.

48| Resuspend the cells in the remaining ~50 µl of supernatant and plate them on an agar plate containing the selection antibiotic.

49| Repeat Steps 45–48 for negative control sample.

50| Incubate all agar plates (Steps 45, 48 and 49) overnight at 30–34 °C.
? TROUBLESHOOTING

51| Under sterile conditions, pick >10 colonies into 150 µl of LB containing antibiotic to select for the recombineered marker in a 96-well flat-bottom plate. Pick also one colony from the ancestral plate or inoculate 1 µl of ancestral culture (Step 29) into a designated well containing LB without antibiotic as a control.

52| Grow the cultures at 30–34 °C to late log phase or overnight.

53| Recombineered strains can be stored indefinitely at –80 °C in 11% (vol/vol) glycerol. For that, mix 2 ml of 80% (vol/vol) glycerol with 5 ml of LB and add 70 µl into each well of a 96-well plate. Next, add 50 µl of culture into each well and pipette it up and down to mix. Seal it with foil film for storage.
▲ **CRITICAL STEP** To access clones at a later date, remove the film entirely. Do not stab through the foil, as this can lead to contamination of the frozen stock.

Screening for successfully recombineered clones ● TIMING 3 h
54| Add 19 µl of sterile water into each well of a V-bottom 96-well plate.

55| Dilute the cultures from Step 51 for PCR screening by adding 1 µl of culture from each well of a 96-well flat-bottom plate to the corresponding well of the 96-well V-bottom plate.

56| Prepare 200 µM storage stocks of the screening primers designed in Steps 16 and 17 and dilute them to make 20 µM working stocks. Store both stocks at –20 °C indefinitely.

57| Prepare a PCR master mix by combining the following for each clone to be screened:

Compound	Amount per sample	Final concentration
F primer (20 µM)	0.5 µl	0.5 µM
R primer (20 µM)	0.5 µl	0.5 µM
2× 2G Fast HotStart ReadyMix with loading dye	10 µl	1×
Water	up to 20 µl	NA

NA, not applicable.

58| Add 19 µl of master mix into each well of a 96-well PCR plate.

59| Add 1 µl of the diluted culture to each well of the 96-well PCR plate.

60| Place the sealing film on the 96-well PCR plate and place it in the thermocycler using the following program (optimal for Kapa 2G Fast HotStart ReadyMix with loading dye):

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 3 min			
2–26	95 °C, 15 s	58 °C, 15 s	72 °C, 15 s/kb of desired product	
27			72 °C, 5 min	
28				12 °C

61| Visualize 4 µl of PCR product on an agarose gel, as described by Isaacs *et al.*[10].

Box 2 | Determination of optimal selection agent concentrations ● TIMING 20 h

Because resistance markers are genomic, reduced copy number of each gene could reduce the resistance at high selection agent concentrations. **Table 1** lists the selection agent concentrations used to select for genomically integrated selectable markers.

Growing strains in the presence of multiple different selection agents often requires reduction in the amount of each antibiotic used. To determine the optimal range of selection agent concentrations to work with, use the following steps (example selection agents used are spectinomycin and Zeocin).

1. Grow the four strains to late log phase or overnight: a strain containing both $spec^R$ and zeo^R markers; a strain containing only $spec^R$ marker; a strain containing only zeo^R marker; and a strain containing neither marker.
2. Mix 600 µl of the following LB-antibiotic concentrations and dispense in 150-µl aliquots into wells of a 96-well plate: 95 µg/ml spec and 10 µg/ml zeo (100% of the concentration in **Table 1**), 71.2 µg/ml spec and 7.5 µg/ml zeo (75% of the concentration in **Table 1**), 47.5 µg/ml spec and 5 µg/ml zeo (50% of the concentration in **Table 1**), 23.8 µg/ml spec and 2.5 µg/ml zeo (25% of the concentration in **Table 1**), and no selection agents.
3. Inoculate one well of each concentration with the strain containing both $spec^R$ and zeo^R markers, with only $spec^R$ marker, with only zeo^R marker or with no markers.
4. Grow the plate in a shaking plate reader on the kinetic setting for 16–20 h, with 282 c.p.m. orbital shaking and OD_{600} reads every 10 min.
5. Determine the concentration at which the strain containing both $spec^R$ and zeo^R markers grows but none of the other strains grow. Use this concentration of spec and zeo in experiments when they will be used in combination.

▲ CRITICAL STEP Because *tolC* is an efflux pump implicated in the export of several toxic small molecules, it may also influence the fitness of strains in the presence of some selection agents or small molecules (e.g., chloramphenicol). Specifically, we have found that chloramphenicol cannot be used in $tolC^-$ strains, even in the presence of the gene encoding chloramphenicol acetyltransferase.

62| Compare PCR bands from recombineered clones with the PCR band from the ancestral strain; successfully recombineered clones should show a shift in size equal to the size of the integrated marker.
? TROUBLESHOOTING

63| Choose one successfully recombineered clone, and while working in sterile conditions dilute 3 µl of culture from the well of the chosen clone in the 96-well plate (Step 51) into 3 ml of fresh LB and antibiotic for marker. Positive selection combined with PCR screening generally ensure clone genotype.

64| Grow the culture to late log phase or overnight at 30–34 °C with shaking.
■ PAUSE POINT Recombineered strains can be stored indefinitely at –80 °C as 1 ml of confluent culture mixed with 200 µl of 80% (vol/vol) sterile glycerol in 2-ml cryotubes.

Transfer of the conjugative plasmid (pRK24) to the donor strain ● TIMING 6–10 h

65| Determine the concentrations of the selective agents to be used according to **Box 2**.

66| Prepare agar plates containing no selection agents and two selection agents—one for the CAGE donor strain and the other for pRK24.

67| Start cultures of the strain containing the conjugative plasmid pRK24 and of the desired CAGE donor strain (Step 63) in separate 14-ml culture tubes with 2 ml of LB supplemented with antibiotics to select for each strain and to maintain conjugative plasmid. Start the cultures according to Step 29.
▲ CRITICAL STEP Only CAGE donor strains should receive the pRK24 plasmid in this step. CAGE recipients will not be used until Step 87.

68| Prewarm the agar plates without antibiotics for conjugation. Use one plate for each conjugation strain mixture, as well as one plate for the CAGE strain only and one plate for the pRK24 strain only.

69| When cultures reach late log phase (OD_{600} = 1.0), transfer them to 2-ml microcentrifuge tubes (**Fig. 4**).

70| Centrifuge the tubes at 13,500*g* for 1 min at room temperature.

71| Remove the supernatant without disturbing the cell pellets.

72| Resuspend the cells in 2 ml of fresh antibiotic-free LB.

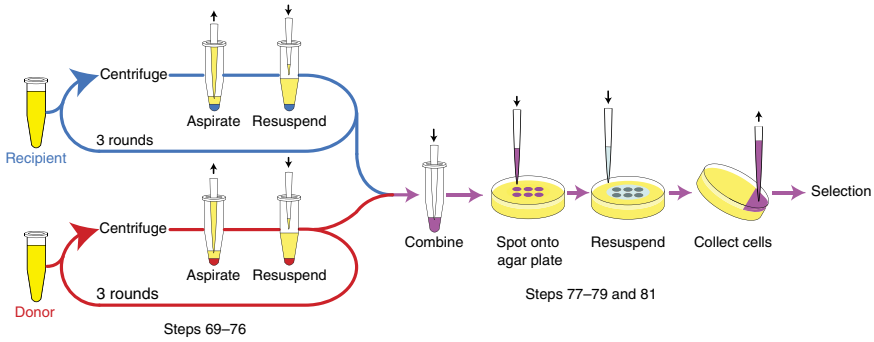

Figure 4 | Overview of the conjugation procedure in CAGE. Donor and recipient strains are independently grown to late-log phase, rinsed to remove antibiotics, and concentrated and mixed in a 4:1 ratio of donor to recipient cells. Cells are then spotted on an agar plate to conjugate on a solid surface, incubated for 1 h and rinsed off the plate with LB in preparation for selection.

73| Repeat Steps 70–72 once.

74| Repeat Steps 70 and 71 once and resuspend the cell pellets in 100 µl of fresh antibiotic-free LB.

75| Normalize the cell density in culture for each strain by diluting 10 µl of culture in 990 µl of fresh medium in a spectrophotometer cuvette. Check the OD at 600 nm on the spectrophotometer and calculate the amount of medium to add with the following equation: medium to add in µl = $(OD_0 \times V_0)/(OD_F)-V_0$, where OD_0 is the current OD_{600} value, V_0 is the current volume that the cells are suspended in and OD_F is the desired OD_{600} value.

76| Add the medium until the diluted OD_{600} of strains is uniform and between 0.15 and 0.25, which indicates an OD_{600} of 15–25 in cell density.

77| Combine 80 µl of strain containing pRK24 and 20 µl of CAGE donor strain (ratio of 4:1) in a microcentrifuge tube, pipetting gently to mix.
▲ **CRITICAL STEP** Pipetting roughly or vortexing may reduce the conjugation efficiency by shearing the F pilus.

78| Pipette the cell mixture onto a prewarmed agar plate lacking antibiotics as two 20-µl spots and six 10-µl spots.

79| Pipette the remaining cells of each strain separately onto prewarmed agar plates lacking antibiotics as controls.

80| Transfer all agar plates to 30–34 °C and incubate them for 1 h.

81| Remove the cells from each plate by rinsing 750 µl of fresh medium over the plate multiple times, as seen in **Figure 4**. Aspirate the medium from the plate and place it in a microcentrifuge tube. Repeat with another 750 µl of fresh medium and add it to the same tube.

82| Centrifuge the tubes at 13,500*g* for 1 min at room temperature to pellet the cells.

83| Remove the supernatant and resuspend the cells in 250 µl of fresh medium.

84| Serially dilute the cells 1:10 in LB repeatedly to 10^{-6} and plate dilutions 10^{-4} through 10^{-6} on agar plates containing selection agents for both the desired CAGE donor and the pRK24 conjugative plasmid.

85| Incubate the plates overnight at 30–34 °C.
? TROUBLESHOOTING

86| Pick one colony from each plate prepared in Step 84 and place it into 3 ml of LB containing antibiotics to select for both the CAGE donor strain and the pRK24 conjugative plasmid. Grow the cells to late log phase or overnight.
▲ **CRITICAL STEP** Conjugation is highly efficient; if a strain grows on a plate containing the antibiotics to select for both the donor and recipient strains, whereas the controls of donor only and recipient only do not grow, it is extremely likely that the picked colony is correctly conjugated. To ensure this, though, always retain growing conjugated strains in medium supplemented with the appropriate antibiotics.
■ **PAUSE POINT** CAGE donor strains containing pRK24 can be stored indefinitely at –80 °C as 1 ml of confluent culture mixed with 200 µl of 80% (vol/vol) sterile glycerol in 2-ml cryotubes.

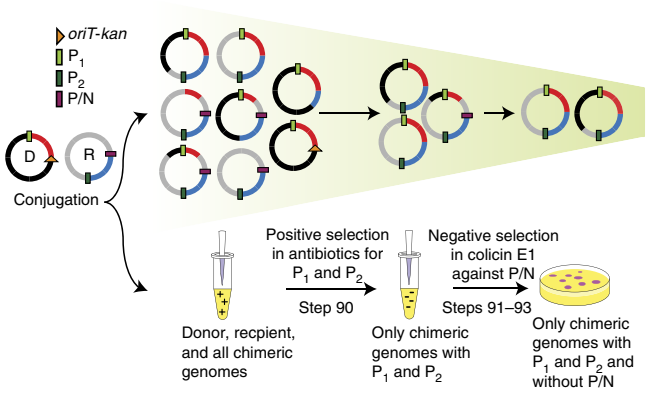

Figure 5 | Overview of the selection procedure in CAGE. After conjugation, cells are exposed to serial rounds of positive and negative selection to select for chimeric genomes of a desired genotype. Donor, recipient and all chimeric genomes are first exposed to positive selection with antibiotics, eliminating all genomes that do not possess both positive selectable markers (P_1 and P_2) in a single genome. The resulting chimeric strains are then challenged with negative selection in colicin E1 to eliminate any genomes that retain the P/N selectable marker, resulting in only chimeric genomes that contain the entire desired donor and recipient fragments.

Conjugation of strains ● **TIMING 6–10 h**

87| Start the cultures of CAGE donor and recipient strains in separate 14-ml culture tubes with 2 ml of fresh LB supplemented with antibiotics to select for each strain, as well as to maintain the conjugative plasmid (in the donor strains only). Start the cultures according to Step 29.

▲ **CRITICAL STEP** Determine the concentrations of selective agents to be used according to **Box 2**.

88| Prepare the agar plates containing selective agents for markers P_1 and P_2.

89| Repeat Steps 68–83 with the CAGE donor and recipient strains, resulting in a donor-recipient cell mixture in 250 µl of LB after conjugation (**Fig. 4**). If the amount of genome to be transferred from donor to recipient exceeds 2.3 Mb (roughly half of the *E. coli* genome), increase the conjugation time in Step 80 to 2 h.

Selecting for desired chimeras ● **TIMING 2 d**

90| Select for chimeric genomes containing both positive markers by diluting 5 µl of donor-recipient mixture from Step 89 in 3 ml of fresh LB containing antibiotics to select for P_1 and P_2 (**Fig. 5**), and grow the culture overnight in a shaking incubator at 30–34 °C.

? TROUBLESHOOTING

91| Select for chimeric genomes that have lost the positive/negative marker (**Fig. 5**) by performing a negative selection on the overnight culture from Step 90 in colicin E1 according to **Box 1**.

? TROUBLESHOOTING

92| Prewarm the agar plates containing positive marker antibiotics.

93| Serially dilute the cultures 1:10 from the negative selection (Step 91) and plate them on agar plates containing selection agents for positive markers P_1 and P_2. Incubate the plates overnight at 30–34 °C.

94| Pick colonies into wells of a 96-well flat-bottom plate each containing 150 µl of LB with antibiotic to select for positive markers P_1 and P_2. Pick the ancestral donor and recipient colonies or inoculate 1 µl of the ancestral donor and recipient culture (Step 87) into designated wells with LB lacking antibiotics as a control.

95| Grow the cultures in flat-bottom plates at 30–34 °C to late log phase or overnight. For storing post-conjugation strains, see storage methods in Step 53.

■ **PAUSE POINT** Postconjugation strains can be stored in LB-glycerol at –80 °C indefinitely.

Genetic screening of chimeric genomes ● **TIMING 3 h**

96| Repeat Steps 54–62 for screening for P_1, P_2, P/N and *oriT-kan* markers. Correct chimeric genomes should only possess P_1 and P_2 markers at sites identical to their position in the donor and recipient genomes.

97| After verification of markers, chimeric genomes can be screened for the correct genotypes using one of the following methods, on the basis of the nature of mutations or genotypes unique to donor and recipient fragments. Choose option A for screening for insertions or deletions larger than 3 bp long by sequencing; option B for screening

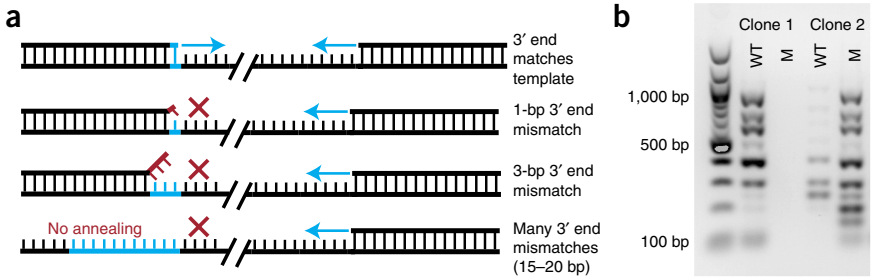

Figure 6 | Genotyping by MASC-PCR. (**a**) MASC-PCR interrogates genotype by 3′ primer annealing. If the primer is complementary at the 3′ end, amplification occurs to produce a fragment. However, mismatch between the template and one primer at the 3′ end produces less PCR product and a weaker or absent gel band. This effect increases with increasing number of mismatches. (**b**) The genotype for each clone can be determined by looking at the bands in the wild-type (WT) and mutant (M) lanes. For example, clone 1 only contains bands in the WT lane, indicating that it is wild type at all loci. In contrast, clone 2 contains bands in the M lane, indicating that it is mutant at all loci. The appearance of faint bands in the WT reaction of clone 2 can be resolved with temperature gradient PCR (**Box 3**).

for mutations that are 3 bp long or less, by multiplex allele-specific colony-PCR (MASC-PCR; **Fig. 6**); or option C for screening for multiple mutations of varying types, or when screening the final chimeric genome, by whole-genome sequencing.

(A) If mutations are sequence insertions or deletions >3 bp long, verify the genotype by sequencing

(i) Design primers to amplify the regions to sequence by following Steps 16 and 17.

(ii) Amplify DNA fragments for sequencing according to Steps 54–62.

(iii) If the PCR produces multiple bands, cut the desired band out of the gel and purify the PCR product using the gel extraction kit according to the manufacturer's instructions. If the PCR shows only one clear band, purify the PCR product directly, to increase yield, using the PCR purification kit according to the manufacturer's instructions.

(iv) Elute the DNA into sterile water and submit it for sequencing.

? TROUBLESHOOTING

(B) If mutations are <3 bp long, screen by MASC-PCR

(i) Design MASC-PCR primers according to **Box 3** (see also **Fig. 6**).

(ii) Prepare 200 µM stocks of each primer with sterile water and dilute an aliquot to 20 µM for working stocks.

(iii) Prepare a wild-type primer mix by combining all the wild-type–specific primers with the general primer in water such that each primer is at 1 µM concentration; for example, add 50 µl of each 20 µM primer into a final volume of 1 ml of primer mix.

(iv) Create a separate mutant primer mix by combining all the mutant-specific primers with the general primer in water such that each primer is at 1 µM concentration; for example, add 50 µl of each 20 µM primer into a final volume of 1 ml of primer mix.

(v) Prepare the clones to screen by following Steps 54 and 55.

(vi) Assemble a PCR master mix for the wild-type screen and one for the mutant screen by combining the following

Box 3 | MASC-PCR primer design

MASC-PCR is a multiplexed version of allele-specific colony PCR, and it allows simultaneous screening of SNPs or mutations of 3 bp or fewer at multiple loci[10,16]. In allele-specific colony PCR, three primers must be designed for each locus to screen: a general primer either upstream or downstream of the locus, a wild-type primer with its 3′ end complementary to the wild-type sequence at that locus and a mutant primer with its 3′ end complementary to the mutant sequence at that locus. Either the general primer or both the wild-type and mutant primers must be reverse-complemented to yield a PCR product. Two PCRs are performed on each clone or colony to screen: one with the wild-type primer and general primer, and the other with the mutant primer and the general primer. When the clones are run on a gel side by side, clones that are wild type show a band in PCR containing the wild-type primer, whereas clones that are mutant show a band in the PCR containing the mutant primer (**Fig. 6**).

To multiplex these reactions, each primer set must generate a different DNA fragment size (between 100 and 1,200 bp) when run on a gel. For example, ten loci can be screened with primer sets that generate the following fragment sizes: 100, 150, 200, 250, 300, 400, 500, 600, 700 and 800 bp. To ensure that all primers can be used in the same reaction, the melting temperature for every primer should be within a 1.5 °C range using Kun's Oligonucleotide Calculator (http://arep.med.harvard.edu/kzhang/cgi-bin/myOligoTm.cgi)[24]. Wild-type or mutant primer pools are then created by combining all wild-type and general primers or all mutant and general primers, respectively. To determine the optimal annealing temperature for the PCR, pools of primers should be tested on a gradient PCR with a control of known genotype before use. The gradient PCR should then be run on a gel. The annealing temperature at which each band appears in only the wild-type or the mutant lane, but not both, is the optimal temperature (**Fig. 6**).

for each clone (optimized for Kapa 2G multiplex mix):

Compound	Amount per sample	Final concentration
Primer mix	4 µl	Depends on the number of primer sets
2G multiplex mix	10 µl	1×
Water	Up to 20 µl	NA

NA, not applicable.

(vii) Add 19 µl of master mix into each well of the 96-well PCR plate.
(viii) Add 1 µl of the diluted culture to each well of the 96-well PCR plate.
(ix) Place sealing film on the 96-well PCR plate and place the plate in a thermocycler using the following program (optimal for Kapa 2G Fast HotStart ReadyMix with loading dye):

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 3 min			
2–26	95 °C, 15 s	See **Box 3**, 15 s	72 °C, 60 s	
27			72 °C, 5 min	
28				12 °C

(x) Visualize 4 µl of PCR product on agarose gel, as described by Isaacs *et al.*[10].
▲ **CRITICAL STEP** When agarose gel is loaded with DNA samples, load wild-type reactions in odd wells and mutant reactions in adjacent even wells (**Fig. 6b**). This enables easy visualization of each clone's genotype.
? TROUBLESHOOTING

(C) Multiple mutations of varying types

(i) If there are multiple regions requiring verification via sequencing or if this is the final chimeric genome, prepare chosen clones for whole-genome sequencing using a commercially available kit protocol. To rule out the possibility of large-scale genomic rearrangements in final strains, pulsed-field gel electrophoresis may also be performed on the chosen clones as described in Smith *et al.*[21].

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 3**.

TABLE 3 | Troubleshooting table.

Step	Problem	Possible reason	Solution
25, 62, 97A(iv), 97B(x)	DNA bands are of incorrect size or there are multiple bands	Primers anneal incorrectly or are nonspecific	Run primers on NCBI Primer Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and determine specificity; if primers are nonspecific, extend their length to increase specificity Check the melting temperature of the primers and increase the annealing temperature of PCR
25, 62, 97A(iv), 97B(x)	No DNA bands in the expected product size	One primer is not in reverse-complement orientation, or it does not anneal to the template	Check the original design to ensure that primers are present in the template sequence, check that only one primer is reverse-complement orientation
	No band at the correct size, but there is a large band at the bottom of the gel at ~100 bp	Primers anneal to each other (primer dimers)	Check the complementarity of primers to each other and themselves. Extend or shorten the primers to decrease complementarity

(continued)

TABLE 3 | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
39	Arcing during electroporation	Residual salt in either the cell pellet or the DNA	Remove extra medium from the cell pellet during the washes in Steps 33–38, or wash the cell pellet once more with chilled sterile water
43	No cell outgrowth during recovery	Cultures left too long at 42 °C or too many cells were killed during electroporation	Restart cultures and ensure they remain at 42 °C for a maximum of 15 min and a minimum of 10 min, wash cell pellet once more (Step 37) with chilled sterile water
50	No colony growth on the negative control or experiment plates, but growth is observed on the positive control	Recombineering failed	Check the arms of homology to ensure they are correct, and minimize secondary structure using NCBI's Primer Blast or other primer design software. If necessary, redesign the arms of homology or move the integration site to identify better SIRs
50, 85	No colonies on any of the plates or colonies on all the plates	Antibiotics are too strong or too weak	Create new agar plates with antibiotic and plate the recombineering culture again See **Box 2** to determine correct concentrations of selection agents when using multiple selection agents at once
85	No colony growth on the negative control or experiment plates, but growth is observed on the positive control	Conjugation failed	Re-do the conjugation, handling cells more gently to prevent damage to F pili Ensure that the wash steps to prepare strains for conjugation (Steps 70–74) are performed in LB without antibiotic
62	All PCR bands show wild-type (no marker present at locus), but strains are resistant to the marker antibiotic	Selectable marker is integrated elsewhere in the genome	Check the arms of homology to ensure direction and specificity to a single region in the genome If the template for recombineering DNA was a plasmid (Step 20), screen clones for presence of the plasmid. Ensure that the DpnI digest (Steps 23 and 24) was performed properly
90	No cell growth during selection in liquid medium	Conjugation failed	Ensure that wash steps to prepare strains for conjugation (Steps 70–74) are performed in LB without antibiotic Check that *oriT-kan* is facing the correct direction and re-do the conjugation, handling cells more gently Increase the conjugation time to 2 h to increase the likelihood and amount of genetic information transferred If conjugation fails repeatedly, reposition the markers to allow a larger region between the P_1 and P_2 markers or between the *oriT* site and the P/N marker
91	Both experiment and negative control grow	Colicin E1 has degraded	Use a fresh tube of colicin E1 to set up another negative selection. Dilute fewer cells into each well

● TIMING

Steps 1–6, determination of genome assembly order: 15 min
Steps 7–18, design of selection marker cassettes and screening primers: 1 h
Steps 19–26, amplification and isolation of selection marker DNA: 2–3 h
Steps 27–53, recombineering selection markers into the genome of donor and recipient strains: 3 d
Steps 54–64, screening for successfully recombineered clones: 3 h
Steps 65–86, transfer of conjugative plasmid to the donor strain: 6–10 h
Steps 87–89, conjugation of strains: 6–10 h
Steps 90–95, selecting for desired chimeras: 2 d

Steps 96 and 97, genetic screening of chimeric genomes: 3 h
Box 1, negative selection of P/N marker *tolC*: 18 h
Box 2, determination of optimal antibiotic concentrations: 20 h

ANTICIPATED RESULTS

In the context of genome recoding, each round of CAGE decreases the number of strains by 50%, with the resultant genomes possessing twice the amount of recoded genome. These new strains are paired and CAGE is repeated until a single strain with all the desired genetic information remains (**Fig. 2**). Two important observations can determine the success of a single CAGE experiment. The outcome of selections after the conjugation step is the first important observation. During the positive liquid selections after conjugation (Step 90), a confluent postconjugal culture alongside no growth observed in the ancestral donor and recipient negative control cultures is a strong indicator of a successful CAGE experiment. Subsequent negative selections in colicin E1 (Step 91) of these cultures should also show growth, in contrast to no growth in the recipient strains, which contain the *tolC* gene. The other crucial observation that determines the success of a CAGE experiment is the genetic analysis by MASC-PCR or genome sequencing that shows the presence of desired genetic loci from both the ancestral donor and recipient strains in the chimeric strain after conjugation and selection.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS The authors acknowledge support from the U.S. Department of Energy (DE-FG02-02ER63445), Defense Advanced Research Projects Agency (N66001-12-C-4020, N66001-12-C-4211), Arnold and Mabel Beckman Foundation (F.J.I.), Gruber Foundation (N.J.M.), the National Institute of Health Cellular and Molecular Biology Training Grant (N.J.M.), the National Institute of Health Genetics Training Grant (N.J.M.), and the National Institute of Health Biophysics Training Grant (D.W.M.) and the Yale Chemical Biology Institute Training Grant (D.W.M.).

AUTHOR CONTRIBUTIONS F.J.I. developed the CAGE methodology in collaboration with A. Tolonen, M. Lajoie and G. Church. N.J.M. finalized the CAGE protocol, wrote the manuscript and created figures. D.W.M. provided feedback on manuscript and created figures.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

1. Ochman, H., Lawrence, J.G. & Groisman, E.A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).
2. Smith, G.R. Conjugational recombination in *E. coli*: myths and mechanisms. *Cell* **64**, 19–27 (1991).
3. Lederberg, J. & Tatum, E. Gene recombination in *Escherichia coli*. *Nature* **53**, 673–684 (1946).
4. Pansegrau, W. *et al.* Complete nucleotide sequence of Birmingham IncPα plasmids: compilation and comparative analysis. *J. Mol. Biol.* **239**, 623–663 (1994).
5. Lanka, E. & Wilkins, B.M. DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64**, 141–169 (1995).
6. Curtiss, R. Bacterial conjugation. *Annu. Rev. Microbiol.* **23**, 69–136 (1969).
7. Fürste, J. & Pansegrau, W. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. USA* **86**, 1771–1775 (1989).
8. Guiney, D. & Yakobson, E. Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. *Proc. Natl. Acad. Sci. USA* **80**, 3595–3598 (1983).
9. Wollman, E.-L., Jacob, F. & Hayes, W. Conjugation and genetic recombination in *Escherichia coli* K-12. *Cold Spring Harb. Symp. Quant. Biol.* **21**, 141–162 (1956).
10. Isaacs, F.J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
11. Court, D.L., Sawitzke, J.A. & Thomason, L.C. Genetic engineering using homologous recombination. *Annu. Rev. Genet.* **36**, 361–388 (2002).
12. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009).
13. DeVito, J.A. Recombineering with *tolC* as a selectable/counter-selectable marker: remodeling the rRNA operons of *Escherichia coli*. *Nucleic Acids Res.* **36**, e4 (2008).
14. Schwartz, S.A. & Helinski, D.R. Purification and characterization of colicin E1. *J. Biol. Chem.* **246**, 6318–6327 (1971).
15. Gallagher, R.R., Li, Z., Lewis, A.O. & Isaacs, F.J. *Nat. Protoc.* **9**, 2301–2316 (2014).
16. Lajoie, M.J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
17. Bates, S., Cashmore, A.M. & Wilkins, B.M. IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: involvement of the Tra2 mating system. *J. Bacteriol.* **180**, 6538–6543 (1998).
18. Thomas, C.M. & Smith, C.A. Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. *Annu. Rev. Microbiol.* **41**, 77–101 (1987).
19. Waters, V.L. Conjugation between bacterial and mammalian cells. *Nat. Genet.* **231**, 375–376 (2001).
20. Gregg, C.J. *et al.* Rational optimization of *tolC* as a powerful dual selectable marker for genome engineering. *Nucleic Acids Res.* **42**, 4779–4790 (2014).
21. Smith, C., Econome, J. & Schutt, A. A physical map of the *Escherichia coli* K 12 genome. *Science* **236**, 1448–1453 (1987).
22. Warming, S., Costantino, N., Court, D.L., Jenkins, N.A. & Copeland, N.G. Simple and highly efficient BAC recombineering using *galK* selection. *Nucleic Acids Res.* **33**, e36 (2005).
23. Wong, Q.N.Y. *et al.* Efficient and seamless DNA recombineering using a thymidylate synthase A selection system in *Escherichia coli*. *Nucleic Acids Res.* **33**, e59 (2005).
24. SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA* **95**, 1460–1465 (1998).