Multiplex genome editing by natural transformation

Ankur B. Dalia[a], EmilyKate McDonough[a], and Andrew Camilli[a,b,1]

[a]Department of Molecular Biology and Microbiology and [b]Howard Hughes Medical Institute, Tufts University School of Medicine, Boston, MA 02111

Edited by Roy Curtiss III, Arizona State University, Tempe, AZ, and approved May 7, 2014 (received for review April 8, 2014)

Editing bacterial genomes is an essential tool in research and synthetic biology applications. Here, we describe multiplex genome editing by natural transformation (MuGENT), a method for accelerated evolution based on the cotransformation of unlinked genetic markers in naturally competent microorganisms. We found that natural cotransformation allows scarless genome editing at unprecedented frequencies of ∼50%. Using DNA substrates with randomized nucleotides, we found no evidence for bias during natural cotransformation, indicating that this method can be used for directed evolution studies. Furthermore, we found that natural cotransformation is an effective method for multiplex genome editing. Because MuGENT does not require selection at edited loci *in cis*, output mutant pools are highly complex, and strains may have any number and combination of the multiplexed genome edits. We demonstrate the utility of this technique in metabolic and phenotypic engineering by optimizing natural transformation in *Vibrio cholerae*. This was accomplished by combinatorially editing the genome via gene deletions and promoter replacements and by tuning translation initiation of five genes involved in the process of natural competence and transformation. MuGENT allowed for the generation of a complex mutant pool in 1 wk and resulted in the selection of a genetically edited strain with a 30-fold improvement in natural transformation. We also demonstrate the efficacy of this technique in *Streptococcus pneumoniae* and highlight the potential for MuGENT to be used in multiplex genetic interaction analysis. Thus, MuGENT is a broadly applicable platform for accelerated evolution and genetic interaction studies in diverse naturally competent species.

Tools for multiplexed genome editing are limited in number and currently are developed only for use in model bacteria. The method known as "multiplexed automated genome engineering" or MAGE was developed in *Escherichia coli* and has been widely successful in "accelerated evolution" of this species, which has been exploited for metabolic and phenotypic engineering applications (1–3). This technique also was critical for "recoding" the *E. coli* genome, in which all UAG stop codons were replaced with synonymous UAA codons (4, 5). MAGE relies on highly efficient recombineering with ssDNA oligonucleotides. Mechanistically this method relies on annealing of ssDNA oligos to the lagging strand during DNA replication and can introduce point mutations or small insertions and deletions into the genome at efficiencies of up to ∼20% (2, 3, 6, 7). A key feature of this technique is the absence of selection for mutations *in cis*, which allows multiplexed mutations to be distributed randomly in output mutant pools, where individual cells in this population have any number and combination of genome edits. MAGE demonstrates the utility of methods for multiplexed genome editing in microbial systems; however, because of the requirement for highly efficient recombineering, this method is not easily adapted to nonmodel microorganisms.

Recently, the Cas-9 endonuclease derived from the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated proteins (Cas) system has been exploited for targeted genome engineering in nonmodel bacterial microorganisms (8). However, this method requires Cas9 selection at edited genomic loci. Therefore CRISPR/Cas-mediated genome editing cannot produce complex mutant pools as described above for MAGE, so the use of this technique for accelerated evolution of phenotypes in microbial systems is limited.

Natural competence and transformation is a trait shared by diverse microbial species (9). It involves the uptake of DNA from the extracellular environment followed by the integration of this DNA into the genome by homologous recombination. During natural transformation, only a fraction of cells in the population become competent and are transformed (9). It has been demonstrated previously that unlinked markers in naturally competent bacteria can be cotransformed, indicating that each competent cell has the ability to take up multiple DNA molecules (10). However, the use of cotransformation for multiplex genome editing applications has not been explored previously. Here we optimize natural cotransformation and demonstrate its use as a method for multiplex genome editing in naturally competent *Vibrio cholerae* and *Streptococcus pneumoniae*.

Results and Discussion

Optimization of Natural Cotransformation. As a first step, we optimized cotransformation of two unlinked markers in *V. cholerae*, selecting for one marker and screening for integration of the other. We used a PCR product to replace a neutral gene (VC1807, a transposase pseudogene with an authentic frameshift) with an antibiotic resistance (Ab^R) marker (selected) and a PCR product to introduce a nonsense point mutation into the β-galactosidase gene (*lacZ*) (unselected) (Fig. 1*A*). We found that the highest rates of cotransformation (∼50–65%) were obtained when the unselected marker had ≥2-kb arms of homology and was present at high concentrations (3 μg/mL) (Fig. 1 *B* and *C*). There were fewer constraints on the selected marker: As expected, increasing the length of homology or the amount of the selected marker increased the number of transformants

Significance

The ability to generate mutants is essential in microbiology research. Although methods have been developed for making defined single mutations in bacterial genomes, methods for simultaneously generating multiple defined mutations (i.e., multiplex genome editing) have been limited to model species such as *Escherichia coli*. Diverse microbial species have the ability to take up exogenous DNA naturally and integrate it into their genome, a process known as "natural transformation." Although natural transformation has been exploited for making single mutations, it has not been used previously for multiplex genome editing. Here, we describe multiplex genome editing by natural transformation and demonstrate the utility of this method in two naturally competent microbial species. This method should be broadly applicable for diverse research applications.

Author contributions: A.B.D. designed research; A.B.D. and E.M. performed research; A.B.D. contributed new reagents/analytic tools; A.B.D., E.M., and A.C. analyzed data; and A.B.D., E.M., and A.C. wrote the paper.

Conflict of interest statement: The authors have filed a patent application on this work.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: andrew.camilli@tufts.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1406478111/-/DCSupplemental.

MICROBIOLOGY

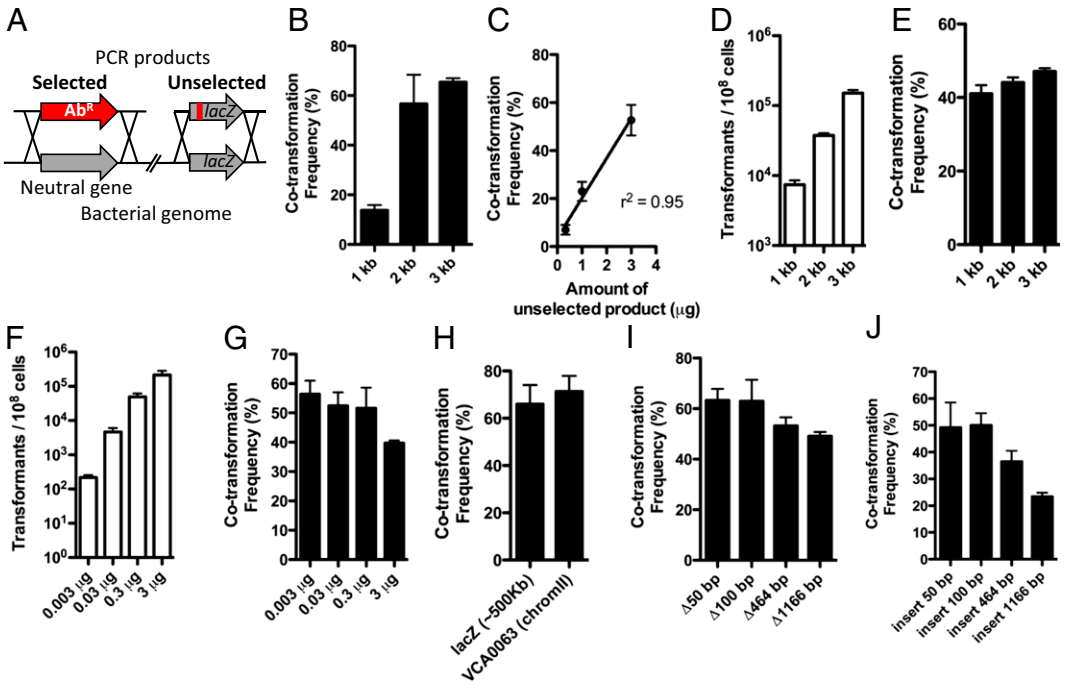

Fig. 1. Optimization of cotransformation in *V. cholerae*. (*A*) Diagram of the approach used to study cotransformation of two unlinked markers. The neutral locus targeted for replacement with an Ab^R marker (also known as the "selected product") is VC1807, a transposase pseudogene containing an authentic frameshift. The unselected PCR product for *B*–*G* contains a transversion point mutation that introduces a premature stop codon into the *lacZ* gene. *B*–*G* represent data from cotransformation assays in which one variable is systematically altered. Unless otherwise noted, reactions in *B*–*J* contained PCR products with 3-kb arms of homology, the selected product at 30 ng/mL, and the unselected product at 3 μg/mL (*B*) Cotransformation assays in which the size of homology in the unselected product was varied. (*C*) Cotransformation assays in which the concentration of the unselected PCR product was varied. *D* and *E* show transformation efficiency (*D*) and cotransformation frequency (*E*) when varying the size of homology in the selected product. The selected product was at 300 ng/mL in these assays. *F* and *G* show transformation efficiency (*F*) and cotransformation frequency (*G*) when varying the concentration of the selected product. (*H*) Cotransformation assays using two distinct unselected genetic markers, one in *lacZ*, which is ~500 kb from the selected marker on the genome, and the other upstream of VCA0063, which is on a different chromosome from the selected marker. (*I*) Cotransformation assays using unselected products to generate deletions of the indicated size in the *lacZ* gene. (*J*) Cotransformation frequency of insertion mutations as measured by reverting strains with deletions of the indicated size in *lacZ* back to WT. All data are from of at least two biological replicates and are shown as the mean ± SD.

obtained (Fig. 1 *D* and *F*); however, this increase did not substantially alter cotransformation frequency, even when the selected product was present at equimolar concentration to the unselected product (Fig. 1 *E* and *G*). This result indicated that the increased concentration of unselected DNA in the transformation reaction, not the ratio of selected:unselected DNA, is critical for optimal cotransformation. Also, distance between the selected and unselected markers on the genome did not alter cotransformation frequency, because an unselected marker on a distinct chromosome exhibited a cotransformation frequency similar to that of the *lacZ* marker (Fig. 1*H*). Genetic edits were not limited to point mutations, because deletions and insertions of 50–1,166 bp were obtained with cotransformation frequencies of ~60–25%, respectively (Fig. 1 *I* and *J*).

Assessing Bias During Natural Cotransformation. Next, we performed cotransformation experiments with PCR products that had either six (N6) or 30 (N30) nucleotides randomized in the *lacZ* gene. To increase the complexity of mutations at the *lacZ* locus, we performed multiple cycles of cotransformation with the N6 and N30 unselected products by using selected products that alter the Ab^R marker at the neutral locus at each cycle (Fig. 2 *A* and *B*). Based on deep sequencing of the input PCR product and output transformant pools, we found no increase in cotransformation frequency for sequences closer to the WT in either the N6 or N30 samples (Fig. 2 *C* and *D*). Furthermore, we found a significant correlation between the abundance of N6 mers in the input PCR pool and in the output transformant pool, further indicating that there is little to no bias in the N6 mers recombined into the genome during cotransformation (Fig. 2*E*). Thus these data suggest that natural cotransformation can be used for unbiased directed evolution at a single genetic locus.

Multiplexed Genome Editing by Natural Transformation Optimizes Natural Transformation in *V. cholerae*. Editing genomes in multiplex in the absence of selection can be used for accelerated evolution to optimize metabolic pathways and phenotypes (1–3). Thus we assessed whether natural cotransformation can be used for multiplex genome editing. Because genome edits do not require selection, output transformants can have any number of edits, and by using multiple cycles of cotransformation we can increase the complexity of gene edits in the final transformant pool (Fig. 3*A*).

As a proof-of-concept, we optimized the phenotype of natural transformation in *V. cholerae*, because many of the genes involved in natural transformation and their regulation are well characterized (11–14). In our approach, we targeted genetic loci that would impact distinct steps of natural transformation, including uptake of transforming DNA (tDNA) into the periplasm (*tfoX*), transport across the inner membrane (*tfoX* and *hapR*), protection of cytoplasmic single-stranded tDNA (*dprA*), and homology searching/integration of tDNA (*recA*) (Fig. 3*B*). The *tfoX*, *hapR*, and *recA* genes were targeted for promoter replacement (promoter construct = LacI-inducible P_{tac} and *rrnB* antiterminator) and ribosome-binding site (RBS) tuning; *dprA* was targeted for RBS tuning alone, because this gene is within an operon. RBS tuning was accomplished by semirandomized mutagenesis of two key positions within the RBSs of these four

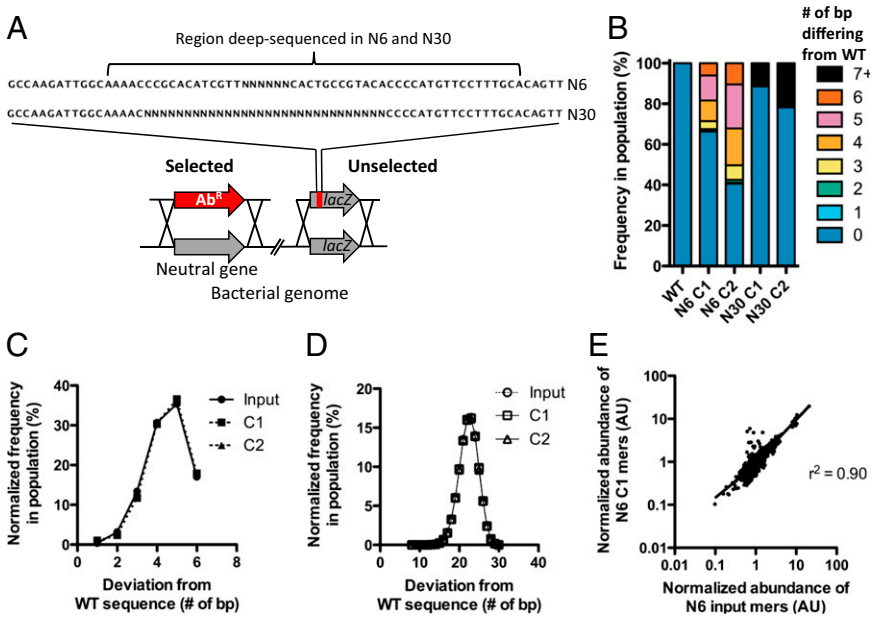

Fig. 2. Assessing bias during natural cotransformation in *V. cholerae*. (*A*) Diagram of the approach used to randomize six (N6) or 30 (N30) base pairs in the *lacZ* gene by cotransformation and the deep-sequenced region. (*B*) Frequency of randomized bases in the *lacZ* gene following two cycles (C1 and C2) of cotransformation with the N6 and N30 PCR products. (*C* and *D*) The composition of the N6 and N30 regions in the input PCR product and output cotransformant pools for N6 (*C*) and N30 (*D*) samples as measured by divergence of sequences from the WT consensus. (*E*) Linear regression of the abundance of all 4,096 N6 mers, excluding the WT sequence, in the input PCR product and output cotransformant pool for the N6 C1 sample.

genes (Fig. S1) (15, 16). The mismatch repair (MMR) system can prevent or correct genetic edits subsequent to integration (17, 18). Therefore, we also targeted *mutS*, a critical component of MMR, for inactivation. In total, there are 1,000 possible combinations for these genome edits.

First, we used cotransformation to introduce genome edits into a population of cells in multiplex. PCR products for each mutation were mixed at equimolar concentrations with a selectable marker in transformation reactions. To perform multiple cycles of multiplexed genome editing by natural transformation (MuGENT), different Ab^R markers were swapped at the neutral locus at each cycle (Fig. 3*A*). Transformants were screened by multiplex allele-specific colony (MASC) PCR (3), and after a single cycle of coselection (C1) we found that ∼50% of the population had at least one genetic edit (Fig. 3*C*). After a second cycle of coselection (C2/R0), ∼90% of the population contained at least one edit, and ∼4% had edits at all five loci. Both cycles of coselection were accomplished in less than 1 wk. Thus MuGENT is a feasible and highly effective strategy for generating complex mutant pools within a defined set of loci.

Next, we wanted to select and characterize edited strains with the phenotype of improved natural transformation. Therefore we subjected the C2/R0 mutant pool to two additional rounds of natural transformation using only a selected marker to enrich for strains with a phenotype of increased natural transformability (R1 and R2). After these two additional rounds of enrichment, edits at *tfoX* and *recA* were found in ∼100% and ∼90% of the population, respectively, suggesting that these edits enhance natural transformation (Fig. 3*D*). Indeed, when we tested defined edited strains with the WT RBS, we found that the transformation efficiencies of the *tfoX*, *recA*, and *tfoX recA* strains are greater than that of the parent strain (Fig. 3*E*). Next, we isolated seven randomly chosen colonies from the final enriched pool and found that they all had transformation efficiencies higher than the parent strain, and many were improved compared with any defined singly and doubly edited strains (Fig. 3*E*, *Bottom*). Thus, MuGENT allowed for the rapid isolation of multiply edited strains with improved natural transformation phenotypes, representing an increase of up to ∼30-fold over the parent strain and ∼sixfold over any singly edited strain, likely due to the combinatorial effect of these RBS-optimized genome edits. Assessing the combinatorial space explored in these experiments in a sequential manner using classic techniques would take an inordinate amount of time and effort. Thus, these experiments demonstrate that MuGENT is an excellent platform for accelerated evolution in naturally competent microbes.

MuGENT Rapidly Generates All Possible Mutant Combinations of a Defined Gene Family in *S. pneumoniae*. Genetic redundancy can hinder uncovering phenotypes in organisms. Using MuGENT, we can reveal redundancies by generating pools of defined mutant combinations. To test this ability and to demonstrate MuGENT in another species, we targeted the four pneumococcal histidine triad (*pht*) genes in *S. pneumoniae* for inactivation. These genes have been characterized previously as redundant zinc-binding proteins (19, 20). Using MuGENT, we introduced premature tandem stop codons into *phtA*, *phtB*, *phtD,* and *phtE* in a combinatorial fashion. Cotransformation frequency was lower in *S. pneumoniae* than in *V. cholerae*. Nonetheless, after five cycles of MuGENT, which took 1 wk to perform, we obtained all 16 possible combinations for these genome edits (Fig. 4 *A* and *C*). We hypothesized that the difference in editing frequency between *V. cholerae* and *S. pneumoniae* may be caused by differences in the efficacy of MMR in these bacteria. To test this notion, we repeated the combinatorial *pht* gene inactivation experiment in a strain lacking MMR and found that editing frequencies were improved dramatically (Fig. 4*B*). In contrast, MMR showed a minimal effect when tested in *V. cholerae* (Fig. S2). The basis for this differential effect is currently unknown. Thus, use of MMR-deficient *S. pneumoniae* will increase the speed of MuGENT but also may increase the frequency of off-target mutations in the genome. Indeed, this effect is observed during MAGE, which commonly is performed in MMR-deficient strains (4). Recently, it was demonstrated that

MICROBIOLOGY

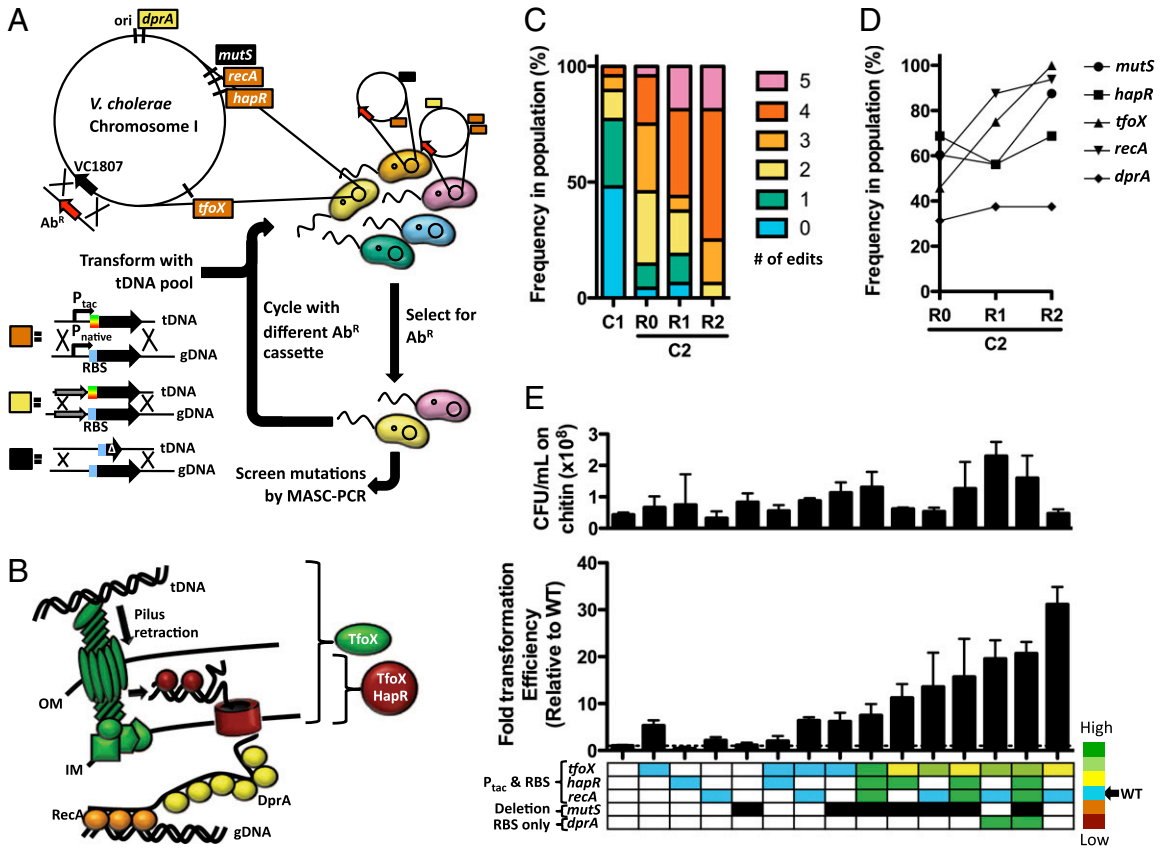

Fig. 3. MuGENT in *V. cholerae* optimizes natural transformation. (*A*) Strategy for generating complex heterogenic mutant populations using cotransformation and the genetic loci targeted in this experiment. (*B*) Diagram depicting the roles of targeted loci in *V. cholerae* natural transformation. TfoX and HapR are regulators that control the indicated processes. (*C*) Distribution of genome edits in the population following two cycles of MuGENT (C1 and C2) and two rounds of selection with only an Ab^R-conferring selected marker (R1 and R2). (*D*) Frequency of each genome edit following selection. (*E*) (*Top*) Final biomass on chitin and (*Middle*) transformation efficiencies from transformation assays. (*Bottom*) The grid under the *x* axis indicates the genotype of strains. A filled box indicates the presence of a genome edit, and the color indicates the strength of the RBS. Blue indicates that the native promoter was replaced with the P_{tac} promoter while maintaining the WT RBS. Black is used for *mutS*, because this gene was targeted for inactivation. Data (mean ± SD) are from four independent biological replicates.

use of temperature-sensitive MMR alleles allows efficient MAGE while limiting off-target mutations (21). Application of conditional MMR deficiency to *S. pneumoniae* also may allow efficient MuGENT while limiting off-target effects.

Concluding Remarks

We have demonstrated that MuGENT can be used for multiplex genome editing in two naturally transformable bacteria, Gram-negative *V. cholerae* and Gram-positive *S. pneumoniae*. Both these microorganisms are human pathogens, and MuGENT has the potential to uncover novel phenotypes and provide deep insight into how these bacteria interact with their mammalian hosts. Specifically, MuGENT provides the tools necessary to generate strains with large numbers of defined mutations rapidly and has the potential to uncover novel biology as a platform for genetic interaction studies.

MuGENT also may serve as a useful platform for accelerated evolution in nonpathogenic species of *Vibrio* and *Streptococci*. *Vibrio* species are naturally found in the aquatic environment. Chitin is a waste product of the food industry and is the most abundant biomolecule in aquatic environments, and *Vibrio* species naturally degrade chitin and use it as a source of carbon and nitrogen (22). Thus, these species could be exploited for biotechnology applications using chitin as an input carbon source. Additionally, some *Vibrio* species, namely *V. splendidus*, are capable of degrading and using alginate, further expanding the possible carbon sources that could be exploited for biotechnology applications (23). Currently, the use of these species has been limited by the lack of the genetic tools required for efficient metabolic and phenotypic engineering (23, 24). To date, natural competence and transformation have been demonstrated in a number of *Vibrio* species (25–27). Thus, MuGENT provides the genetic tools necessary for the development of *Vibrio* species for use in diverse biotechnology applications. The probiotic microbe *Streptococcus thermophilus* is commonly used in the dairy industry and is naturally competent (28). Thus, MuGENT may be used for metabolic engineering in *S. thermophilus* to alter or enhance its use in the dairy industry as well as enhance the probiotic activity of this species.

A large number of diverse species of microbes are known or are predicted based on bioinformatics to be naturally transformable and thus would be candidates for use of MuGENT. These include, but are not limited to, species of *Bacillus*, *Cyanobacterium*, *Lactococcus, Acinetobacter*, *Neisseria*, and *Haemophilus* (9).Therefore this method should be broadly applicable for diverse research and biotechnology applications.

Materials and Methods

Bacterial Strains and Culture Conditions. All *V. cholerae* and *S. pneumoniae* parent strains are described in Table S1. *V. cholerae* and *S. pneumoniae* were routinely grown exactly as described previously (29, 30). For *V. cholerae*, when appropriate, medium was supplemented with 50 μg/mL kanamycin,

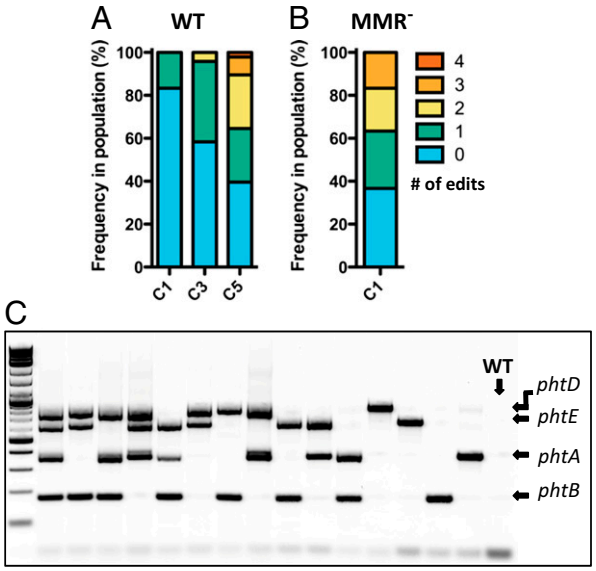

Fig. 4. MuGENT in *S. pneumoniae* rapidly generates all possible defined *pht* mutant strains. All four *pht* genes were targeted by MuGENT using PCR products that introduce tandem stop codons into each locus. (*A* and *B*) Distribution of genome edits in the population following cycles of MuGENT in a WT (*A*) or an MMR-deficient (*B*) pneumococcal strain. (*C*) MASC PCR of all 16 possible *pht* mutant strains made in the WT background. A band indicates the presence of a genome edit.

100 μg/mL spectinomycin, 100 μg/mL streptomycin, or 100 μg/mL ampicillin. For *S. pneumoniae*, when appropriate, medium was supplemented with 200 μg/mL spectinomycin, 4 μg/mL chloramphenicol, or 100 μg/mL streptomycin.

Generation of Mutant Constructs and Strains. Mutant constructs for selected and unselected PCR products throughout this study were generated via splicing by overlap extension (SOE) PCR exactly as previously described (31, 32) using Phusion polymerase (Thermo Scientific) because this enzyme has a lower error rate than other PCR polymerases. Briefly, two-piece SOE PCR was performed by amplifying an upstream region of homology with F1/R1 primers (UP arm) and a downstream region of homology with F2/R2 oligos (DOWN arm). The desired mutations were incorporated into the R1 and F2 primers used to amplify the UP and DOWN arms, respectively. Furthermore, the R1 and F2 primers were designed to share an overlapping sequence at their 5′ ends. Thus, splicing of these products was accomplished by mixing UP and DOWN arms as the template for PCR with F1 and R2 primers. For three-piece SOE PCRs, R1 and F2 primers contained sequences that overlap with a third PCR product (MIDDLE product) at their 5′ ends, and splicing was accomplished by mixing all three PCR products together as template for PCR with F1 and R2 primers. The primers used to generate all SOE products are listed in Table S2. In *V. cholerae*, the neutral locus targeted with the selected product was VC1807, a transposase pseudogene with an authentic frame-shift, which was replaced with a spectinomycin-, kanamycin-, or ampicillin-resistance marker. In *S. pneumoniae*, the selected product replaced SP_1051 with a chloramphenicol- or spectinomycin- resistance marker. The promoter construct consisting of P_{tac} and the *rrnB* antiterminator used during MuGENT in *V. cholerae* was derived from the end of a previously described Tn*10* transposon (31).

Natural Transformation and MuGENT in *V. cholerae*. Natural transformation of *V. cholerae* following growth on chitin from shrimp shells was done exactly as described previously (31). To assess transformation efficiencies and biomass on chitin, reactions were plated directly onto medium selective for the Ab^R marker (i.e., transformants) and onto medium lacking antibiotics to assess total viable cfus (i.e., total biomass on chitin). Transformation efficiency was defined as cfus of transformants/total viable cfus.

For cotransformation into *lacZ*, cells were plated on medium selective for the Ab^R marker and containing 40 μg/mL X-Gal (Sigma Aldrich) to assess cotransformation frequency.

For MuGENT, all PCR products, including the selected marker were added to transformation reactions at 3 μg/mL and had 3-kb arms of homology, because this concentration and length were found to be optimal for cotransformation. Under these condition, each cycle of MuGENT in a 1-mL reaction generated $\geq 10^5$ transformants. After reactions were incubated with DNA, samples were outgrown for 1 h in LB broth in the absence of antibiotics. A small aliquot (~1/10th of the reaction) was plated to assess transformation efficiency, and single colonies from selective plates were used for MASC PCR. The remainder of each transformation was inoculated into 50 mL of LB broth containing the appropriate antibiotic to select for transformants and was grown overnight at 37 °C with aeration. The following day, this culture was diluted 1:100 in medium lacking antibiotics and was grown to an OD_{600} of ~1.0. These cells then were washed, and ~10^8 cfus were placed onto chitin to repeat another cycle of MuGENT or to select for transformants from the mutant pool. After the first cycle of MuGENT, all subsequent transformations with this mutant pool were performed in the presence of 10 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma Aldrich) to induce expression of the P_{tac} promoter used in some genome edits. Growth in LB always was performed in the absence of IPTG, because we found that IPTG-induced expression of the edited gene *hapR* resulted in a growth defect.

Natural Transformation and MuGENT in *S. pneumoniae*. Natural transformation of *S. pneumoniae* was performed exactly as previously described (33). For MuGENT, 1.5 μg of each unselected product and 300 ng of the selected product were added to a 1-mL transformation reaction. All unselected products had 2.5- to 3-kb arms of homology, whereas the selected product had 1.5-kb arms of homology. After the addition of DNA, reactions were incubated at 37 °C in a 5% CO_2 incubator for 1 h. A small aliquot of each reaction (~1/10th) then was plated to assess transformation efficiency, and single colonies from selective plates were used for MASC PCR. The remainder of the transformation was plated for single colonies on medium selective for transformants. The following day, these plates were flooded with Todd Hewitt broth (BD Biosciences) supplemented with 5% (wt/vol) yeast extract (Fisher Scientific) (THY medium) to resuspend colonies. This bacterial slurry then was diluted in 10 mL of fresh THY medium to an OD_{600} of 0.05 and was grown to an OD_{600} of ~ 0.6. Cells then were washed, diluted, and retransformed to perform additional cycles of MuGENT.

MASC PCR. At each cycle of MuGENT, 24–48 single colonies were assessed for genome edits by MASC PCR essentially as described previously (3). All oligos used for MASC PCR are given in Table S2.

Analysis of High-Throughput Sequencing Data for Assessing Bias During Natural Cotransformation. After cotransformation of PCR products that randomized six (N6) and 30 (N30) bases in the *lacZ* gene of *V. cholerae*, we generated libraries for deep sequencing from genomic DNA purified from output transformant pools and from the input SOE PCR products. We did so by first PCR amplifying with ABD419 and ABD408. This PCR then was used as the template for a second round of PCR using ABD420 and a reverse primer that adds a unique 6-bp barcode sequence so we could distinguish samples run together on a single lane of the Illumina HiSeq. All primers used for preparing sequencing libraries are listed in Table S2.

After sequencing, data were analyzed on the Tufts University Galaxy server (34). First, we used the trim tool to remove the first six bases for N30 samples or 17 bases for N6 samples. Then, we used the clip tool to remove the constant sequence at the 3′ end of all molecules (N6 = 5′-CACTGCCGTA-CACCCCATGTTCCTTTGC-3′ and N30 = 5′-CCCCATGTTCCTTTGC-3′). We then used filter fastq to obtain reads of a length of six bases (N6) or 30 bases (N30) with a minimum quality score of 34 (on a scale of 0–41). To define the distribution of these reads in reference to their deviation from the WT consensus, we used the barcode splitter tool using the WT sequence as a reference and allowed any number (n = 1, 2, 3, … 30) of mismatches to define the distribution of sequences that were 1, 2, 3, to 30 bases different from the WT sequence. To define the exact abundance of each N6-mer in the input and output transformant pools, we used the barcode splitter tool using the sequence of each N6-mer as a reference and allowed for 0 mismatches.

ACKNOWLEDGMENTS. We thank Michael Malamy and Neil Greene for helpful discussions. This work was supported by National Institutes of Health Grants AI055058 and AI045746 (to A.C.). A.C. is a Howard Hughes Medical Institute Investigator.

1. Sandoval NR, et al. (2012) Strategy for directing combinatorial genome engineering in Escherichia coli. *Proc Natl Acad Sci USA* 109(26):10540–10545.
2. Wang HH, et al. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460(7257):894–898.
3. Wang HH, et al. (2012) Genome-scale promoter engineering by coselection MAGE. *Nat Methods* 9(6):591–593.
4. Isaacs FJ, et al. (2011) Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* 333(6040):348–353.
5. Lajoie MJ, et al. (2013) Genomically recoded organisms expand biological functions. *Science* 342(6156):357–360.
6. Ellis HM, Yu D, DiTizio T, Court DL (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci USA* 98(12):6742–6746.
7. Carr PA, et al. (2012) Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Res* 40(17):e132.
8. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31(3):233–239.
9. Lorenz MG, Wackernagel W (1994) Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* 58(3):563–602.
10. Erickson RJ, Copeland JC (1973) Congression of unlinked markers and genetic mapping in the transformation of *Bacillus subtilis* 168. *Genetics* 73(1):13–21.
11. Bergé M, Mortier-Barrière I, Martin B, Claverys JP (2003) Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol Microbiol* 50(2):527–536.
12. Chen I, Dubnau D (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol* 2(3):241–249.
13. Lo Scrudato M, Blokesch M (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet* 8(6):e1002778.
14. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science* 310(5755):1824–1827.
15. Salis HM (2011) The ribosome binding site calculator. *Methods Enzymol* 498:19–42.
16. Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27(10):946–950.
17. Tham KC, et al. (2013) Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol Cell* 51(3):326–337.
18. Claverys JP, Méjean V, Gasc AM, Sicard AM (1983) Mismatch repair in *Streptococcus pneumoniae*: Relationship between base mismatches and transformation efficiencies. *Proc Natl Acad Sci USA* 80(19):5956–5960.
19. Melin M, et al. (2010) Interaction of pneumococcal histidine triad proteins with human complement. *Infect Immun* 78(5):2089–2098.
20. Panina EM, Mironov AA, Gelfand MS (2003) Comparative genomics of bacterial zinc regulons: Enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins. *Proc Natl Acad Sci USA* 100(17):9912–9917.
21. Nyerges A, et al. (2014) Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic Acids Res* 42(8):e62.
22. Hunt DE, Gevers D, Vahora NM, Polz MF (2008) Conservation of the chitin utilization pathway in the *Vibrionaceae*. *Appl Environ Microbiol* 74(1):44–51.
23. Wargacki AJ, et al. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* 335(6066):308–313.
24. Alper H, Stephanopoulos G (2009) Engineering for biofuels: Exploiting innate microbial capacity or importing biosynthetic potential? *Nat Rev Microbiol* 7(10):715–723.
25. Chen Y, Dai J, Morris JG, Jr., Johnson JA (2010) Genetic analysis of the capsule polysaccharide (K antigen) and exopolysaccharide genes in pandemic *Vibrio parahaemolyticus* O3:K6. *BMC Microbiol* 10:274.
26. Gulig PA, Tucker MS, Thiaville PC, Joseph JL, Brown RN (2009) USER friendly cloning coupled with chitin-based natural transformation enables rapid mutagenesis of *Vibrio vulnificus*. *Appl Environ Microbiol* 75(15):4936–4949.
27. Pollack-Berti A, Wollenberg MS, Ruby EG (2010) Natural transformation of *Vibrio fischeri* requires *tfoX* and *tfoY*. *Environ Microbiol* 12(8):2302–2311.
28. Blomqvist T, Steinmoen H, Håvarstein LS (2006) Natural genetic transformation: A novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl Environ Microbiol* 72(10):6751–6756.
29. Dalia AB, Lazinski DW, Camilli A (2013) Characterization of undermethylated sites in *Vibrio cholerae*. *J Bacteriol* 195(10):2389–2399.
30. Shainheit MG, Mulé M, Camilli A (2014) The core promoter of the capsule operon of *Streptococcus pneumoniae* is necessary for colonization and invasive disease. *Infect Immun* 82(2):694–705.
31. Dalia AB, Lazinski DW, Camilli A (2013) Identification of a membrane-bound transcriptional regulator that links chitin and natural competence in *Vibrio cholerae*. *MBio* 5(1):e01028–e13.
32. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR (1989) Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. *Gene* 77(1):61–68.
33. Bricker AL, Camilli A (1999) Transformation of a type 4 encapsulated strain of *Streptococcus pneumoniae*. *FEMS Microbiol Lett* 172(2):131–135.
34. Giardine B, et al. (2005) Galaxy: A platform for interactive large-scale genome analysis. *Genome Res* 15(10):1451–1455.