The enhancement of blinking by increased laser power suggested that blinking might form part of the stress response mechanism in *E. coli*. We thus tested whether blinking was associated with cationic efflux, another important mechanism of stress response.

We observed surprising dynamics of a cationic membrane-permeable dye, tetramethyl rhodamine methyl-ester (TMRM), in blinking *E. coli*. As expected for this Nernstian voltage indicator (*16*), TMRM gradually accumulated in the cytoplasm over ~10 min. However, blinks in PROPS fluorescence coincided with precipitous stepwise drops in TMRM fluorescence that showed little or no recovery after the blink (Fig. 4A). The duration of the step in TMRM fluorescence coincided with the duration of the blink: At moderate intensities of red illumination (I = 10 W/cm^2) steps lasted less than 200 ms, whereas under little or no red illumination steps typically lasted several seconds (fig. S9). Stepwise disappearance of TMRM was also observed in cells without the PROPS plasmid, when only dim green illumination was used to image the TMRM (30 mW/cm^2; Fig. 4B and fig. S10). The duration of these steps was comparable to that of steps in cells with PROPS under dim red illumination (2 W/cm^2). The rapid disappearance of TMRM during a blink suggested an active-transport mechanism. Dissipation of V_m lowers the thermodynamic barrier to cationic efflux (Fig. 4C) (*6*). A concurrent dissipation of V_m and increase in membrane permeability would be sufficient to induce cationic efflux. PMF-dependent efflux of other cationic dyes has been observed in *E. coli* (*17*) in population-level assays that are insensitive to the dynamics of individual cells.

Bacterial electrophysiology is likely to differ in several key regards from its eukaryotic version due to the comparatively small surface area, yet high surface-to-volume ratio found in bacteria. With a typical membrane capacitance between 10^{-14} and 10^{-13} F, a single ion channel with a conductivity of 100 pS can alter the membrane potential with a time constant of 10^{-3} to 10^{-4} s. In contrast, neurons only fire through the concerted action of a large number of ion channels. Thus, bacterial electrophysiology is likely to be dominated by stochastic opening of individual ion channels and pores. Additionally, the ionic composition of bacteria is less robust than that of eukaryotes. A bacterium with a volume of 1 fl and a cytoplasmic Na^+ concentration of 10 mM contains only ~10^7 ions of Na^+. A single ion channel passing a current of 2 pA can deplete this store in less than 1 s. These simple estimates show that some of the tenets of neuronal electrophysiology may need rethinking in the context of bacteria.

References and Notes

1. B. Martinac, Y. Saimi, C. Kung, *Physiol. Rev.* **88**, 1449 (2008).
2. O. Béjà, E. N. Spudich, J. L. Spudich, M. Leclerc, E. F. DeLong, *Nature* **411**, 786 (2001).
3. O. Béjà *et al.*, *Science* **289**, 1902 (2000).
4. J. L. Spudich, C. S. Yang, K. H. Jung, E. N. Spudich, *Annu. Rev. Cell Dev. Biol.* **16**, 365 (2000).
5. P. Kolodner, E. P. Lukashev, Y. C. Ching, D. L. Rousseau, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11618 (1996).
6. Materials and methods are available as supporting material on *Science* Online.
7. A. K. Dioumaev, J. M. Wang, Z. Bálint, G. Váró, J. K. Lanyi, *Biochemistry* **42**, 6582 (2003).
8. G. Pucihar, T. Kotnik, D. Miklavčič, *J. Vis. Exp.* **33**, 1659 (2009).
9. G. Miesenböck, D. A. De Angelis, J. E. Rothman, *Nature* **394**, 192 (1998).
10. D. C. Fung, H. C. Berg, *Nature* **375**, 809 (1995).
11. D. Novo, N. G. Perlmutter, R. H. Hunt, H. M. Shapiro, *Cytometry* **35**, 55 (1999).
12. G. D. Sprott, W. G. Martin, H. Schneider, *Photochem. Photobiol.* **24**, 21 (1976).
13. T. Noumi, M. Maeda, M. Futai, *FEBS Lett.* **213**, 381 (1987).
14. A. K. Joshi, S. Ahmed, G. Ferro-Luzzi Ames, *J. Biol. Chem.* **264**, 2126 (1989).
15. W. J. Ingledew, R. K. Poole, *Microbiol. Rev.* **48**, 222 (1984).
16. C. J. Lo, M. C. Leake, T. Pilizota, R. M. Berry, *Biophys. J.* **93**, 294 (2007).
17. J. A. Bohnert, B. Karamian, H. Nikaido, *Antimicrob. Agents Chemother.* **54**, 3770 (2010).

Acknowledgments: We thank K. Rothschild for helpful discussions and for the GPR plasmid. We also thank H. Berg, R. Losick, J. Yuan, H. Inada, A. Garner, and D. Kahne for helpful discussions. G. Miesenbock provided the pHluorin plasmid. A. Fields, H. Bayraktar, V. Venkatachalam, and L. Bane helped with measurements. This work was supported by the Harvard Center for Brain Science, an Intelligence Community Postdoctoral Fellowship (J.M.K.), a National Science Foundation Graduate Fellowship (D.R.H.), a Helen Hay Whitney Postdoctoral Fellowship (A.D.D.), and a Charles A. King Trust Postdoctoral Fellowship (A.D.D). J.M.K., A.D.D., and A.E.C. have applied for a patent on PROPS as a sensor of membrane potential.

Supporting Online Material

www.sciencemag.org/cgi/content/full/333/6040/345/DC1
Materials and Methods
SOM Text
Figs. S1 to S11
Tables S1 to S6
References (*18*–*33*)
Movies S1 to S7

24 February 2011; accepted 31 May 2011
10.1126/science.1204763

Downloaded from www.sciencemag.org on July 15, 2011

Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement

Farren J. Isaacs,[1]*†‡ Peter A. Carr,[2,3]*‡§ Harris H. Wang,[1,4,5,6]* Marc J. Lajoie,[1,7] Bram Sterling,[2,3] Laurens Kraal,[1] Andrew C. Tolonen,[1] Tara A. Gianoulis,[1,6] Daniel B. Goodman,[1,5] Nikos B. Reppas,[8] Christopher J. Emig,[9] Duhee Bang,[10] Samuel J. Hwang,[11] Michael C. Jewett,[1,12] Joseph M. Jacobson,[2,3] George M. Church[1,6]

We present genome engineering technologies that are capable of fundamentally reengineering genomes from the nucleotide to the megabase scale. We used multiplex automated genome engineering (MAGE) to site-specifically replace all 314 TAG stop codons with synonymous TAA codons in parallel across 32 *Escherichia coli* strains. This approach allowed us to measure individual recombination frequencies, confirm viability for each modification, and identify associated phenotypes. We developed hierarchical conjugative assembly genome engineering (CAGE) to merge these sets of codon modifications into genomes with 80 precise changes, which demonstrate that these synonymous codon substitutions can be combined into higher-order strains without synthetic lethal effects. Our methods treat the chromosome as both an editable and an evolvable template, permitting the exploration of vast genetic landscapes.

The conservation of the genetic code, with minor exceptions (*1*), enables exchange of gene function among species, with viruses, and across ecosystems. Experiments involving fundamental changes to the genetic code could substantially enhance our understanding of the origins of the canonical code and could reveal new subtleties of how genetic information is encoded and exchanged (*1*, *2*). Modifying the canonical genetic code could also lead to orthogonal biological systems with new properties. For instance, a new genetic code could prevent the correct translation of exogenous genetic material and lead to the creation of virus-resistant organisms. Additionally, a recoded genome could enhance the incorporation of unnatural amino acids into proteins, because existing suppressor systems must compete with native translation factors (*3*–*5*).

The construction of a new genetic code requires methods to manipulate living organisms at a whole-genome scale. Such methods are only now becoming attainable through the advent of advanced tools for synthesizing, manipulating, and recombining DNA (*6*). This has led to a number of impressive genome-scale studies, which include removing transposable elements (*7*), refactoring phage genomes (*8*), genome merging (*9*), whole-genome synthesis (*10*), and transplantation (*11*). Whole-genome de novo synthesis offers the ability to create new genomes without a physical template. Its main limitations are the cost of accurate in vitro DNA assembly and introduction of synthetic DNA into organisms (*12*). For

this reason, de novo synthesis is chosen when trying to create a small number of new DNA constructs of modest size (<10 kb) and high fidelity (*8*, *10*, *12*, *13*). Notably, however, the digital template used in de novo synthesis currently originates almost exclusively from sequences found in nature or minor variants thereof.

Redesigned genomes require approaches that reconcile the desired biological behavior with challenges inherent to biological complexity. Engineering biological systems can be unpredictable, as a single misplaced or misdesigned allele can be lethal. To address these challenges, we have developed approaches that integrate synthetic DNA and recombination methods to introduce genome-wide changes dynamically in living cells, thereby engineering the genome through viable intermediates. In recent work, we developed multiplex automated genome engineering (MAGE), which rapidly generates genetic diversity for strain and pathway engineering (*14*). To augment MAGE's ability to introduce nucleotide-scale mutations across the genome, a complementary method was required to assemble modified chromosomes in vivo.

Here, we report the development of a hierarchical conjugative assembly genome engineering (CAGE) method and its integration with MAGE toward reengineering the canonical genetic code of *E. coli* (Fig. 1)—an organism with broad utility in basic and applied research. The *E. coli* genetic code contains three stop codons (UAG, UAA, and UGA) whose translation termination is mediated by two release factors, RF1 and RF2. RF1 recognizes the termination codons UAA and UAG, whereas RF2 recognizes UAA and UGA. We hypothesized that replacing all TAG codons with synonymous TAA codons would abolish genetic dependence on RF1 and permit the newly reassigned TAA codons to be recognized by RF2. This will enable us to test and leverage the redundancy of the genetic code by deleting RF1 (Δ*prfA*), providing a blank TAG codon that could be cleanly reassigned to new function. Given that codon utilization bias has been shown to affect translation efficiency (*15*, *16*) and viral infectivity (*13*), we sought to determine whether *E. coli* could maintain viability with the systematic replacement of the 314 TAG codons.

On the basis of the MG1655 genome annotation, we identified 314 *E. coli* genes that contain the TAG stop codon (fig. S1 and table S1). We focused initially on reassigning all 314 stop codons (TAG) to the synonymous stop codon (TAA) in a modified *E. coli* MG1655 strain (EcNR2: *E. coli* MG1655 Δ*mutS*::*cat* Δ(*ybhB-bioAB*)::[*λcI857* Δ(*cro-ea59*)::*tetR-bla*]). A mismatch repair–deficient (Δ*mutS*) strain was used to achieve high-frequency allelic replacement (*17*). We used MAGE to simultaneously introduce subsets of the TAG-to-TAA codon changes into 32 separate strains (Fig. 1). Specifically, the EcNR2 genome was divided into 32 regions; 31 of these contained 10 targets, and the other contained the remaining four targets. This division was pursued for four reasons. First, pilot experiments (fig. S2) and associated computational predictions showed that the use of pools of 10 or more oligonucleotides (oligos) for MAGE (*14*) achieves highly efficient allelic replacement. Second, limiting the number of MAGE cycles for codon conversions minimizes the total number of cell divisions (six to eight per cycle) in the presence of λ red proteins (which promote recombination and are mutagenic) and deficient mismatch repair (MMR) (*18*). This reduces the number of undesired secondary mutations. Third, the use of smaller oligo pools enabled rapid accumulation of the desired codon conversions in parallel and quantitative measurements of conversion frequencies. Finally, we anticipated that certain codons might be recalcitrant to codon conversion or cause an aberrant phenotype, so it was advantageous to test mutations in small subsets. Candidates included 43 essential genes (*19*) that are terminated by TAG (fig. S1) and 39 genes in which the TAG stop codon overlaps a second reading frame (table S2). Thus, parallel allelic replacement across the 32 regions in

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [2]Center for Bits and Atoms, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Program in Biophysics, Harvard University, Cambridge, MA 02138, USA. [5]Program in Medical Engineering and Medical Physics, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA. [6]Wyss Institute, Harvard University, Cambridge, MA 02115, USA. [7]Program in Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [8]Joule Unlimited, Cambridge, MA 02139, USA. [9]Department of Bioengineering, Stanford University, Palo Alto, CA 94305, USA. [10]Department of Chemistry, Yonsei University, Shinchon 134, Seoul 120-749, Korea. [11]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [12]Department of Chemical and Biological Engineering and Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208, USA.

*These authors contributed equally to this work.
†Present address: Department of Molecular, Cellular and Developmental Biology, Systems Biology Institute, Yale University, New Haven, CT 06520, USA.
‡To whom correspondence should be addressed. E-mail: farren.isaacs@yale.edu (F.J.I.); carr@mit.edu (P.A.C.)
§Present address: Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420, USA.

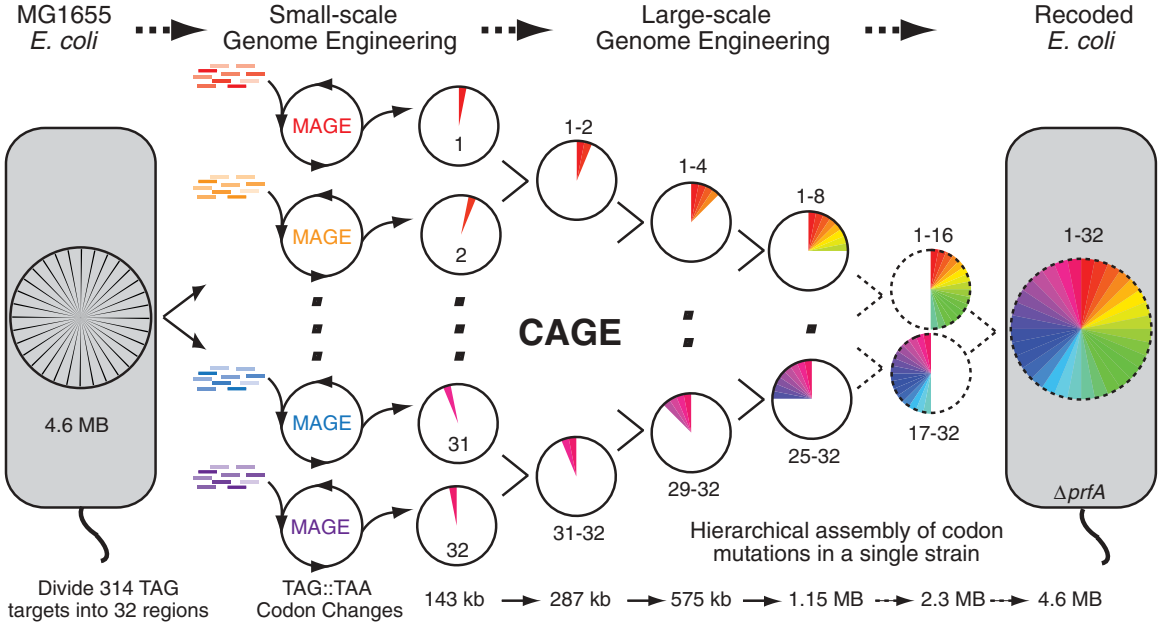


Fig. 1. Strategy for reassigning all 314 TAG codons to TAA in *E. coli*. First, the genome was split into 32 regions each containing 10 TAG stop codons. In parallel, MAGE was used to execute all 10 TAG::TAA codon modifications in a single strain for each genomic region. These partially recoded strains were paired such that a targeted genomic region of one strain (donor) was strategically transferred into a second strain (recipient), permitting the hierarchical consolidation of modified genomic regions using CAGE (see Fig. 4A). This five-stage process transfers genomic fragments ranging in size from ~154 kb to ~2.3 Mb in a controlled manner until a single recoded strain is constructed that lacks the TAG stop codon throughout. Thus far, 28 of 31 conjugations have been completed, where the dotted arrows denote outstanding conjugation steps and dotted genomes represent half- and full-genome strains that have not yet been completed. Once all TAG codons have been converted to TAA, the *prfA* gene will be deleted to inactivate TAG translational termination.

Downloaded from www.sciencemag.org on July 15, 2011

separate strains would enable rapid identification of potentially troublesome alleles.

The 314 oligonucleotides encoding the specified TAG-to-TAA codon mutations (table S3) were computationally designed by means of a software tool (optMAGE, http://arep.med.harvard.edu/optMAGE) on the basis of prior MAGE optimization experiments (*14*). These oligos were repeatedly applied over 18 MAGE cycles to introduce the codon replacements across 32 cultures (10 targets per culture). We developed two methods based on mismatch amplification mutation assay polymerase chain reaction (MAMA-PCR) (*20*) to quickly assay target codons. Multiplex allele-specific colony quantitative PCR (MASC-qPCR) (*21*) (fig. S3) was used to identify clones that contain the greatest number of codon conversions, and multiplex allele-specific colony PCR (MASC-PCR) (*21*) (fig. S4) was used to measure frequencies of allele replacement at each targeted position. MASC-PCR permitted simultaneous single–base pair (1-bp) measurements (TAG versus TAA) at 10 chromosomal sites per clone (fig. S4).

After 18 MAGE cycles, allelic replacement frequencies were analyzed for all 314 TAG-to-TAA mutations (Fig. 2) in 1504 clones (47 clones for each of the 32 recoded segments). Allelic replacement frequencies exhibited a high degree of variability among the targets (Fig. 2, outer ring; table S4). The average allelic replacement frequency observed was 37 ± 19% after 18 cycles, and 42% of the population was unconverted after 18 cycles; we observed 1 to 10 converted alleles per clone across the remaining population (Fig. 3A). These measurements suggest the evolution of two types of cells in our mixed cultures: one that appears largely resistant to allelic replacements, and another that readily permits them. With this knowledge, future MAGE methodology could be modified to select highly recombinogenic clones after fewer cycles (e.g., 5 versus 18). Notably, comparable distributions of allelic replacement frequencies were observed for TAG codons present in essential genes, codons overlapping a second reading frame, and codons distributed at various positions throughout both replicating arms (table S5). Moreover, allelic replacement frequencies did not correlate with distance from the origin of replication (oriC), nor with recombination hotspots [e.g., Chi-sites, DNA motifs (5′-GCTGGTGG-3′) in the genome where homologous recombination could be enhanced] or direction and level of transcription.

All individual TAG-to-TAA conversions were observed, indicating that no TAG stop codon in *E. coli* is required for survival or robust growth. Of 314 codon targets, 298 could be assayed using MASC-PCR, whereas the remaining modifica-

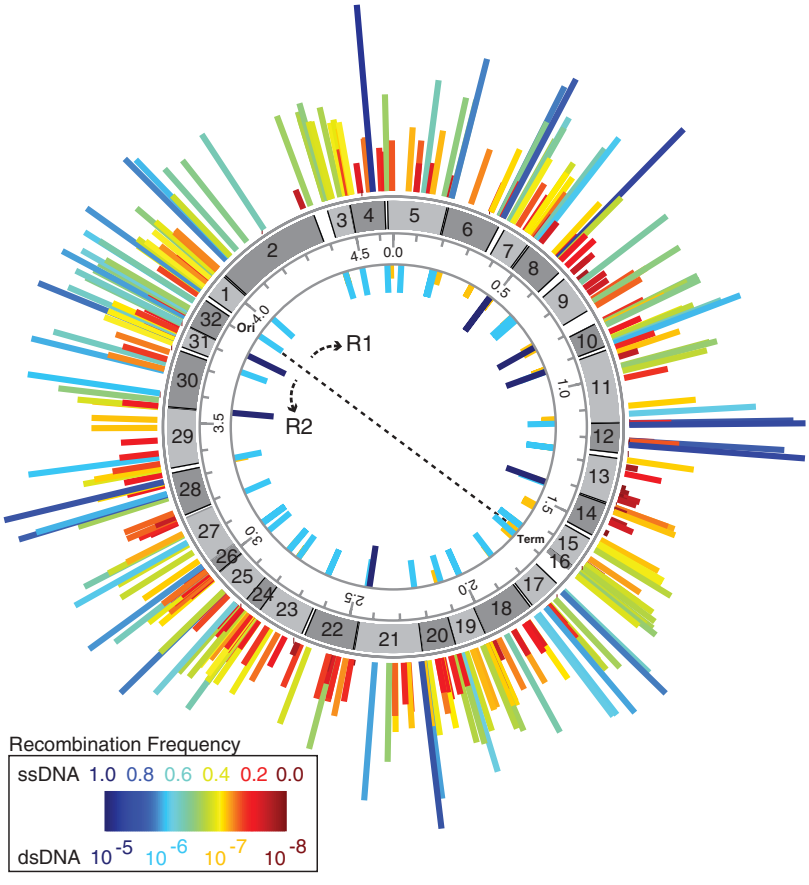


Fig. 2. Frequency map of oligo-mediated TAG::TAA codon replacements and genetic marker integrations across the *E. coli* genome at each replacement position. Circular map illustrates (from inner circle outward): (i) frequency of dsDNA selectable marker integrations; (ii) genome coordinates (in Mb): position of origin (Ori) and terminus (Term) and direction of the two replication forks (R1 and R2); (iii) location of the 32 targeted chromosomal segments; and (iv) frequency of TAG::TAA replacements across all TAG codons—after 18 MAGE cycles—denoted by height- and color-coded bars (scale bar indicates integration frequency).

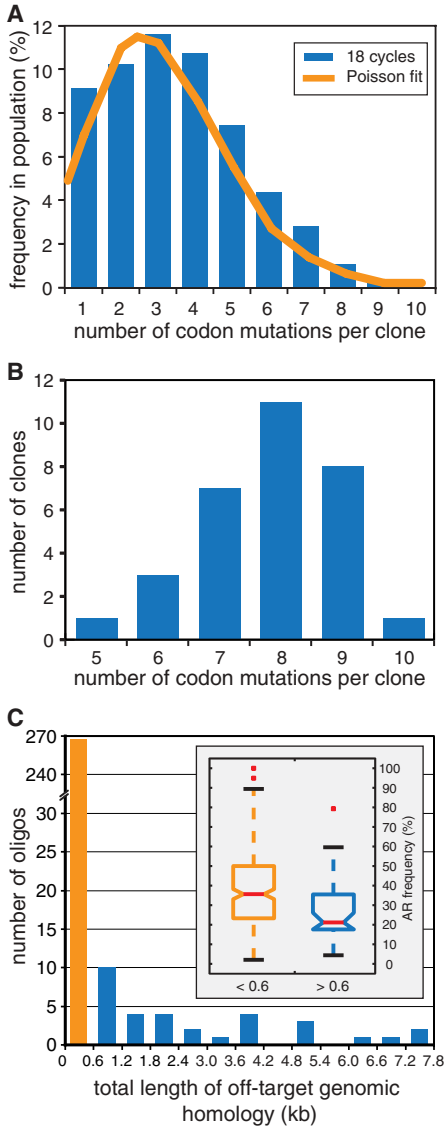


Fig. 3. Clonal rate and distribution of genome modifications after 18 cycles of MAGE. (**A**) Histogram of the frequencies of clones containing 1 to 10 conversions found among 1504 clones screened. A Poisson fit is shown (solid yellow line) for a subpopulation that excludes the zero-conversion group. (**B**) Distribution of modifications among the group of top clones (one for each of the 31 groups of 10 targeted modifications; one additional strain not shown had conversion at all four codon sites). (**C**) Distribution of the 314 90-mer oligos by their extent of total secondary sequence similarity to the *E. coli* genome. Inset box plot: Oligos with a mistarget score of more than 600 bp show, on average, a 32% decrease in allelic replacement (AR) frequency relative to oligos scoring less than 600 bp (25.6% versus 37.6%, $P < 0.003$).

Downloaded from www.sciencemag.org on July 15, 2011

tions were confirmed by direct DNA sequencing. By screening to identify maximally modified cells (Fig. 3B, 5 to 10 modifications per clone with a median of 8) and minimizing aberrant phenotypes (i.e., auxotrophy, decreased fitness) across 1504 clones, we isolated the top clone from each of the 32 populations after 18 MAGE cycles. These clones collectively accumulated 246 of 314 (78%) desired mutations after 18 MAGE cycles. Clones that did not contain all of the codon changes were subjected to an additional 6 to 15 MAGE cycles to convert the remaining TAG codons.

Given that λ Red facilitates highly efficient recombination using short regions of complementary sequences, it was important to assess the potential effects of oligonucleotide hybridization to other (unintended) regions of the genome. We found that 90-mer oligos that have multiple regions of high sequence similarity throughout the genome have a reduced recombination frequency (Fig. 3C) but that these oligos rarely cause mutagenesis at those other locations (see below). To estimate this effect quantitatively, we performed BLAST alignments of each oligo against the entire genome. To compute a mistarget score, we summed the lengths of the BLAST matches for each oligo sequence against the rest of the *E. coli* genome (blastn, word size = 11, expectation value = 10). Although the majority of oligos (~270) showed only minor sequence similarity to the genome (mistarget score < 600), we found that the score strongly correlated with the frequency of allelic replacement (Fig. 3C). Recombination frequencies were decreased by more than 30% for oligos having many regions of high sequence similarity in the genome (mistarget scores > 600 bp; $P < 0.003$). This information will be useful as a predictor of allele replacement frequency for future oligo designs and can be incorporated into automated design software such as optMAGE.

To directly verify the presence of codon conversions and to obtain a snapshot of secondary mutations accumulated during the MAGE process, we performed Sanger sequencing of ~300 bp surrounding each modified TAG replacement site (96 kb overall, ~3 kb in each of the 32 top strains). Sequencing confirmed the accuracy of our MASC-PCR method and verified the 16 TAG-to-TAA conversions not detected by this assay. Background mutations outside of the 90-bp regions targeted by MAGE oligos consisted of 6 substitutions, 0 insertions, and 3 deletions; in contrast, mutations within the targeted regions included 4 substitutions, 1 insertion, and 28 deletions (fig. S5). The use of a MMR-deficient (Δ*mutS*) strain rendered the expected bias toward substitution mutations in the nontargeted regions. Deletion mutations are probably enriched in oligo-targeted regions because internal deletions are common errors in many oligonucleotide chemical synthesis processes (*22*, *23*). We have developed strategies to minimize both sources of error: (i) optimized oligo synthesis to reduce deletions (e.g., fig. S7), and (ii) the use of chemically modified oligos that are not recognized by MMR to achieve efficient allelic replacement in the presence of a functional MMR pathway.

Because we initially performed codon changes in small subsets, we could easily identify candidate

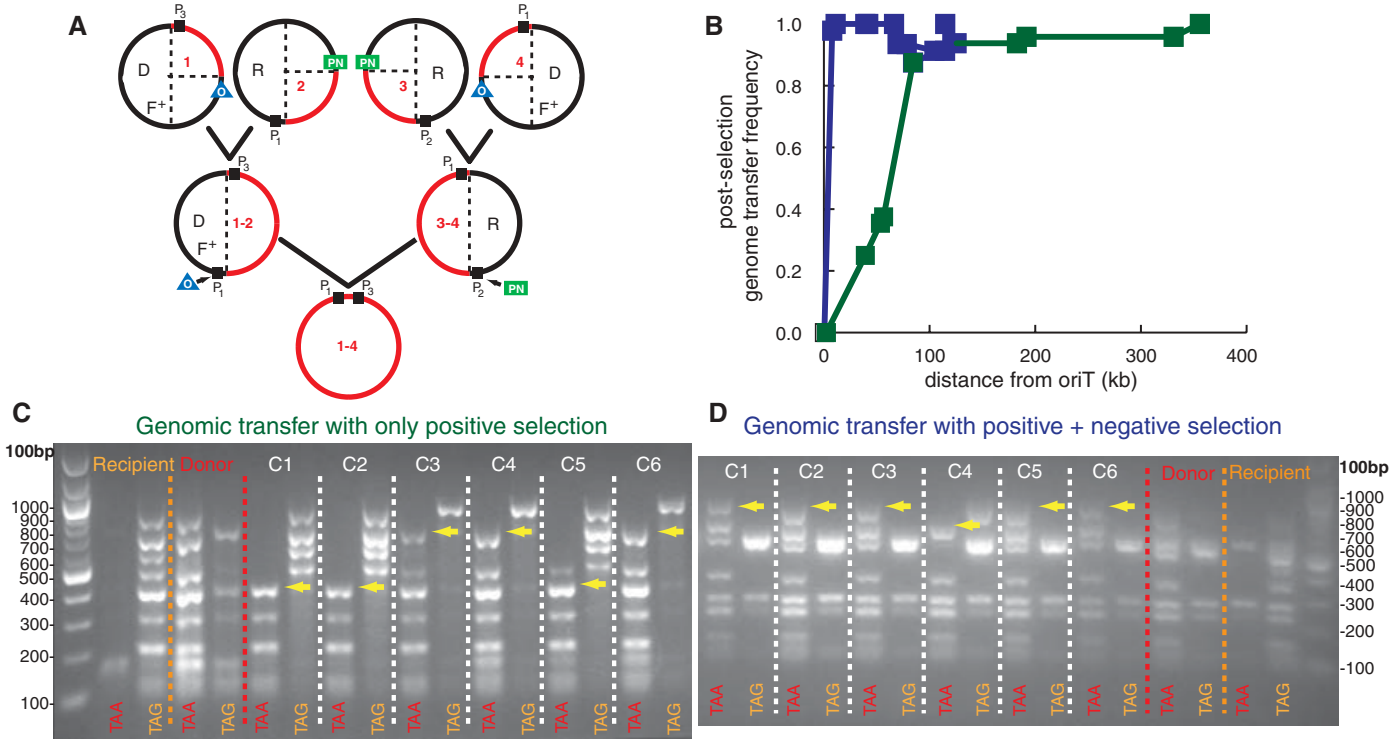


Fig. 4. Hierarchical CAGE methodology for controlled genome transfer. (**A**) Two pairs of strains illustrate the design and methodology of CAGE, with recoded genomic regions in red. Partially recoded strains are split into conjugation pairs. The donor strain (D) contains the following: *oriT-kanamycin* cassette (O, blue triangle); positive selection gene (P_n, n = 1, 2, or 3, black rectangle); and F plasmid (F^+). The recipient strain (R) contains the following: positive-negative selection gene (PN) and P_n, flanking its recoded region. DNA transfer is initiated at O in the donor genome, ensuring transfer of the desired codon mutations and downstream P_n. After conjugation, a specific set of three simultaneous selections is applied to yield a recombinant strain that contains the recoded genomic fragment from the donor strain while retaining the other recoded region in the recipient genome. Placement of the PN marker downstream of the *oriT* sequence in the recipient genome ensures that the entire desired region of the donor genome is inherited in the recombinant strain. All conjugation factors are maintained episomally on F′, so only a ~2-kb *oriT* sequence must be inserted onto the genome to generate a highly controllable *Hfr* donor strain. Because there is no scar between the two recoded regions from the conjugation parents, only one recombination is required to insert the O (donor) or PN (recipient) directly into P_n for the next round of conjugation. This conjugation assembly-selection strategy is implemented in five stages to merge the genomes of 32 recoded strains into a single strain (see Fig. 1). (**B**) Genome transfer frequency as a function of the distance from O. Plots of two conjugations of genome segments illustrating the transfer of ~120 kb with positive and negative selection (blue) and ~360 kb with only positive selection (green) to assemble recoded genomic DNA from donor and recipient genomes. (**C** and **D**) MASC-PCR images of TAA alleles transferred under positive selection alone (C) and positive and positive-negative selection (D). Yellow arrows indicate the genomic point of transfer, which illustrate the inheritance of the donor TAA alleles in the conjugated strain.

Downloaded from www.sciencemag.org on July 15, 2011

mutations that lead to aberrant phenotypes. Growth rates across all 32 top strains (fig. S6 and table S6, average of 47 min per doubling) showed modest deviations from the growth rate of the ancestral strain (42 min per doubling). These changes in growth could be attributed to either the codon changes or the accumulation of secondary mutations in our MMR-deficient strain. Additional phenotypic assays showed a sustained high recombination frequency and a 2.8% frequency of auxotrophy on minimal M9–glucose minimal medium after ~366 generations (table S6). These values compare favorably to previous studies (*24*) in which serial passage of a Δ*mutS* strain resulted in 9% frequency of auxotrophy after ~250 generations.

After converting all TAG codons to TAA across 32 *E. coli* strains, we initiated a five-stage hierarchical assembly (Fig. 1 and table S7) of the modified chromosomal segments into a single strain (Fig. 4). To accomplish this, we developed the hierarchical CAGE method, which is rooted in conjugation, a key mechanism for gene transfer in bacteria (*25*, *26*). In contrast to natural mechanisms of conjugal DNA transfer where the *oriT* sequence and conjugal factors act as a contiguous genetic construct, our approach physically decouples the episomally expressed conjugal factors from the chromosomally integrated *oriT* sequence. The *oriT* sequence is fused with a kanamycin resistance gene (*oriT-kan*) so that it can be easily integrated into any permissible locus across the *E. coli* genome via λ Red–mediated double-stranded DNA (dsDNA) recombination (*27*). Thus, we can precisely control the genomic position at which conjugal transfer is initiated (Fig. 4A). This strategy allowed us to use a tractable ~2-kb cassette in place of a cumbersome 30-kb *Hfr* fragment for consecutive manipulations throughout the genome.

Before conjugation, we converted the 32 strains that collectively contain all TAG-to-TAA modifications into 16 pairs of strains primed for large-scale genome transfer (Fig. 1). Within each conjugation pair, a donor strain transfers its recoded genomic region to a recipient strain, which inherits the donor genome and retains its recoded genomic region. Genome transfer is controlled by the precise placement of positive and positive-negative selectable markers integrated with an engineered conjugation strategy to obtain the desired recombinant genomes. Precise placement of these markers into "safe insertion regions" (SIRs: intergenic regions that are not annotated for any coding or regulatory function) by dsDNA recombination (*27*) was intended to maintain genomic integrity and to attain the desired combination of recoded donor and recipient genomes in the recombinant strain (Fig. 4A). In total, two genetic markers were inserted into each of the donor and recipient strains, yielding a total of 64 markers across the 32 modified strains. In the donor strain, the recoded region was flanked by an upstream *oriT-Kan* cassette and a downstream positive selectable marker (P_1, e.g., zeo^R, $spec^R$, $gent^R$). In the recipient strain, the recoded region was flanked with a different positive selectable marker (P_2) and a positive-negative selectable marker (PN) such as *tolC* (*28*) or *galK* (*29*). The frequencies of integration among selectable marker cassettes exhibited a high degree of site-specific variability (Fig. 2, inner ring). On average, 59 clones (~10^{-6} frequency) were observed per recombination. However, dsDNA recombination frequencies spanned >3 logs across 81 integration sites tested. Twelve intergenic sites yielded no observable recombinants despite repeated (three or more) attempts. The remaining 69 sites performed as follows: 23 sites at ~10^{-7}, 38 sites at ~10^{-6}, and 8 sites at ~10^{-5} recombination frequencies.

Placement of complementary selectable markers across all 32 strains served as anchor points that enabled hierarchical assembly of recoded genomic fragments. By design, this permits the use of modular *oriT-kan* and *tolC* cassettes throughout the assembly process. Rather than having to prepare a cassette for each SIR, three *oriT-kan* cassettes and three *tolC* cassettes that insert directly into the three positive markers (zeo^R, $spec^R$, and $gent^R$ genes) are sufficient to guide the remaining four stages of hierarchical assembly. Because *oriT-kan* and *tolC* are not inherited by the recipient strain, each strain can be prepared for subsequent conjugations by simply inserting an *oriT-kan* (donor) or a PN (recipient) directly into one of the strain's inherited positive markers (Fig. 4A).

In the first stage of the hierarchical conjugation strategy (Fig. 1 and table S7), 32 strains each containing 10 codon modifications were merged to produce 16 strains with 20 modifications. Transfer of 1/32 of the genome (~150 kb) occurs at a frequency of ~10^{-4} (Fig. 4B), 2 logs greater than half-genome transfer (*21*). This result supports prior findings that the probability of transferring a specific marker decreases exponentially with its distance from *oriT* (*26*). The relationship between genome transfer efficiency and the distance from *oriT* revealed useful parameters for designing our engineered conjugation scheme. In the absence of a positive-negative marker in the recipient strain, MASC-PCR analysis showed reduced transfer frequency from loci that are in close proximity (<10 kb) to *oriT*, resulting in the uncontrolled transfer of the donor genome—specifically, the loss of mutated TAA codons from the donor genome and the retention of one to four TAG codons from the recipient genome (Fig. 4C). Upon the inclusion of a positive-negative marker [e.g., *tolC* (*28*), *galK* (*29*)] in the recipient genome, desired postconjugal strains were readily selected; that is, full transfer of mutated TAA codons from the donor genome was achieved by selecting for the loss of *tolC* or *galK* placed among the TAG codons in the recipient genome (Fig. 4D). Together, these results demonstrate the requirement for robust positive and positive-negative selectable markers that strategically flank the recoded genomes in the donor and recipient strains (Fig. 4A). Moreover, MASC-PCR analysis across all codon loci shows that conjugation efficiency is sustained throughout the region of transfer, indicating contiguous transfer of the donor genomic fragment.

Using CAGE, we then consolidated the 32 original strains into eight recoded strains, each with 1/8 of the genome recoded. Two of these eight strains exhibited a dysfunctional *tolC* phenotype (i.e., they simultaneously passed both positive and negative *tolC* selections). Although mutations conferring simultaneous novobiocin sensitivity and colicin E1 resistance have been identified (*28*), there is no literature precedence for the phenotype that we observed. We have discovered two routes to this phenotype. In one strain, the causative allele was present in *tolC*, and we corrected the phenotype by replacing the dysfunctional copy with a functional one. In the other strain, the causative allele appears to be outside of *tolC*. Indeed, *tolC* works in concert with a number of other genes (e.g., *btuB*, *tolA*, *tolQ*, and *tolR*) that have been implicated in dysfunctional negative selection (*30*). Recognizing that the ancestral strains also carried the dysfunctional allele, we reconstructed this 1/8th strain using MAGE, and used it to complete the full set of 1/4 genomes (28 of 31 conjugations). These four strains, which contain up to 80 modifications per genome, can be combined to complete the assembly of a fully recoded strain containing all 314 TAG-to-TAA codon conversions.

In light of the challenges arising from spontaneous point mutations, we sought to assess the effects of MAGE and CAGE on genome stability. Therefore, we performed whole-genome sequencing for the two dysfunctional strains and an additional functional control (figs. S8 and S9 and tables S8 to S11). These strains have 110, 102, and 128 secondary mutations, respectively [total number of single-nucleotide polymorphisms (SNPs) and indels; fig. S8]. After ~960 cell divisions (table S11), the majority of SNPs were transition mutations (98.4% transitions and 1.6% transversions; table S9) and the overall background mutation rate was 2.5×10^{-8} per bp per replication (1 error per genome per ~9 replications; table S10). These results are consistent with a ΔmutS phenotype (*31*). Our measured error rate was lower than we expected, given that the cumulative potential mutations include contributions from a MMR-deficient strain, repeated exposure to induction of the λ Red recombination system, and conjugation-based genomic manipulations. A mechanistic hypothesis for the lower error rate is that the conjugation process acts as a backcross and removes deleterious secondary mutations through the isolation of clones that maximize fitness. To examine this idea further, we explored the potential functional consequences of these SNPs as indicated by the COG category of the gene or regulatory region associated with the SNP (*32*–*34*). We used a hypergeometric distribution to determine the enrichment

Downloaded from www.sciencemag.org on July 15, 2011

level of the three main COG categories across all three strains. Both SNPs associated with metabolism (117 SNPs, $P < 4 \times 10^{-4}$) and SNPs associated with information storage and processing (29 SNPs, $P < 0.05$) were shown to be significantly enriched, whereas SNPs associated with cell signaling and transduction (98 SNPs, $P > 0.05$) were not (fig. S9). Future work will be needed to sequence additional strains throughout the ancestral conjugation tree to characterize the frequency, inheritance patterns, and functional bias of such mutations.

This study, which integrates in vivo genome engineering from the nucleotide to the megabase scale, demonstrates the successful replacement of all genomic occurrences of the TAG stop codon in the *E. coli* genome. We found that cells can incorporate all individual TAG-to-TAA codon changes, and that these changes can be assembled into genomes with up to 80 modifications with mild phenotypic consequences. The scarless introduction of codon changes via MAGE enabled the first genome-wide allelic replacement frequency map using single-stranded DNA oligos in *E. coli* (Fig. 2). In addition, our engineered conjugation experiments produced a complementary recombination frequency map of intergenic dsDNA integration sites across the genome (Fig. 2). Together, these experiments revealed both highly accessible and recalcitrant sites for both small- and large-scale chromosomal modifications. These data could serve as valuable resources for future genome engineering efforts. Moreover, synthetic approaches such as the one pursued here may help to refine the existing genome annotation by revealing unannotated functional genetic loci, such as short peptides (*35*) or minigenes (*36*).

Introducing genome-wide changes dynamically in a living cell permits engineering in the cell's native biological context. In contrast to in vitro genome synthesis (*10*) and transplantation methods (*12*) that introduce discrete and abrupt changes in a single genome, our genome engineering technologies treat the chromosome as an editable and evolvable template and generate targeted and combinatorial modifications across many (~10^9) genomes in vivo (*14*). MAGE is optimal for introducing small modifications in sequence design space, whereas CAGE is designed for taking bigger leaps via large-scale assembly of many modified genomes. Together, these genome editing methods are advantageous when the designed genomes share >90% sequence similarity to existing templates or when many targeted mutations dispersed across the chromosome are desired (e.g., genome recoding).

References and Notes

1. A. Ambrogelly, S. Palioura, D. Söll, *Nat. Chem. Biol.* **3**, 29 (2007).
2. R. D. Knight, S. J. Freeland, L. F. Landweber, *Nat. Rev. Genet.* **2**, 49 (2001).
3. T. S. Young, P. G. Schultz, *J. Biol. Chem.* **285**, 11039 (2010).
4. H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **464**, 441 (2010).
5. T. Mukai *et al.*, *Nucleic Acids Res.* **38**, 8188 (2010).
6. P. A. Carr, G. M. Church, *Nat. Biotechnol.* **27**, 1151 (2009).
7. G. Pósfai *et al.*, *Science* **312**, 1044 (2006); 10.1126/science.1126439.
8. L. Y. Chan, S. Kosuri, D. Endy, *Mol. Syst. Biol.* **1**, 0018 (2005).
9. M. Itaya, K. Tsuge, M. Koizumi, K. Fujita, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15971 (2005).
10. D. G. Gibson *et al.*, *Science* **319**, 1215 (2008); 10.1126/science.1151721.
11. C. Lartigue *et al.*, *Science* **317**, 632 (2007).
12. D. G. Gibson *et al.*, *Science* **329**, 52 (2010); 10.1126/science.1190719.
13. J. R. Coleman *et al.*, *Science* **320**, 1784 (2008).
14. H. H. Wang *et al.*, *Nature* **460**, 894 (2009).
15. B. Irwin, J. D. Heck, G. W. Hatfield, *J. Biol. Chem.* **270**, 22801 (1995).
16. G. A. Gutman, G. W. Hatfield, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3699 (1989).
17. H. M. Ellis, D. Yu, T. DiTizio, D. L. Court, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6742 (2001).
18. N. Costantino, D. L. Court, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15748 (2003).
19. S. Y. Gerdes *et al.*, *J. Bacteriol.* **185**, 5673 (2003).
20. R. S. Cha, H. Zarbl, P. Keohavong, W. G. Thilly, *PCR Methods Appl.* **2**, 14 (1992).
21. See supporting material on *Science* Online.
22. P. A. Carr *et al.*, *Nucleic Acids Res.* **32**, e162 (2004).
23. A. B. Oppenheim, A. J. Rattray, M. Bubunenko, L. C. Thomason, D. L. Court, *Virology* **319**, 185 (2004).
24. P. Funchain *et al.*, *Genetics* **154**, 959 (2000).
25. H. Ochman, J. G. Lawrence, E. A. Groisman, *Nature* **405**, 299 (2000).
26. G. R. Smith, *Cell* **64**, 19 (1991).
27. D. Yu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5978 (2000).
28. J. A. DeVito, *Nucleic Acids Res.* **36**, e4 (2008).
29. S. Warming, N. Costantino, D. L. Court, N. A. Jenkins, N. G. Copeland, *Nucleic Acids Res.* **33**, e36 (2005).
30. M. Masi, P. Vuong, M. Humbard, K. Malone, R. Misra, *J. Bacteriol.* **189**, 2667 (2007).
31. R. M. Schaaper, R. L. Dunn, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6220 (1987).
32. H. Li, R. Durbin, *Bioinformatics* **25**, 1754 (2009).
33. H. Li *et al.*, 1000 Genome Project Data Processing Subgroup, *Bioinformatics* **25**, 2078 (2009).
34. W. McLaren *et al.*, *Bioinformatics* **26**, 2069 (2010).
35. M. R. Hemm *et al.*, *J. Bacteriol.* **192**, 46 (2010).
36. E. Jacinto-Loeza, S. Vivanco-Domínguez, G. Guarneros, J. Hernández-Sánchez, *Nucleic Acids Res.* **36**, 4233 (2008).

Acknowledgments: We thank S. Zhang for helpful discussions and extensive use of laboratory resources, R. Kolter for the JC411 strain, and members of the Church and Jacobson labs for helpful discussions. Supported by the NSF (SynBERC, Center for Bits and Atoms, and Genes and Genomes Systems Cluster), the U.S. Department of Energy, an NSF graduate fellowship (H.H.W.), a National Defense Science and Engineering Graduate Fellowship (M.J.L.), and an NIH K99/R00 award (M.C.J.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/333/6040/348/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S13
References

27 April 2010; accepted 20 May 2011
10.1126/science.1205822

Downloaded from www.sciencemag.org on July 15, 2011

Linking Context with Reward: A Functional Circuit from Hippocampal CA3 to Ventral Tegmental Area

Alice H. Luo,[1]* Pouya Tahsili-Fahadan,[2] Roy A. Wise,[1] Carl R. Lupica,[3] Gary Aston-Jones[2]

Reward-motivated behavior is strongly influenced by the learned significance of contextual stimuli in the environment. However, the neural pathways that mediate context-reward relations are not well understood. We have identified a circuit from area CA3 of dorsal hippocampus to ventral tegmental area (VTA) that uses lateral septum (LS) as a relay. Theta frequency stimulation of CA3 excited VTA dopamine (DA) neurons and inhibited non-DA neurons. DA neuron excitation was likely mediated by disinhibition because local antagonism of γ-aminobutyric acid receptors blocked responses to CA3 stimulation. Inactivating components of the CA3-LS-VTA pathway blocked evoked responses in VTA and also reinstatement of cocaine-seeking by contextual stimuli. This transsynaptic link between hippocampus and VTA appears to be an important substrate by which environmental context regulates goal-directed behavior.

Efficient reward-seeking requires that environmental stimuli be properly interpreted, to predict when and where reward can be expected. The ventral tegmental area (VTA) and its dopaminergic projections are critical components of a reward circuit. Although the influence of temporal cues on this system has been well studied (*1*), it is unknown how the VTA system relates to other contextual information, such as where reward can be expected. The hippocampus organizes aspects of context into a relational memory network (*2*). Interactions between hippocampus and VTA are important for context-reward associations (*3*). However, the circuitry by which the two interact remains to be elucidated.

We used pseudo-rabies virus (PRV), a retrograde, transsynaptic tracer (*4*), to delineate circuit projections to VTA. Beginning 48 hours after unilateral injection into VTA, PRV (+) neurons

[1]Behavioral Neuroscience Section, Behavioral Neuroscience Research Branch, National Institute on Drug Abuse Intramural Research Program, National Institutes of Health, Baltimore, MD 21224, USA. [2]Department of Neurosciences, Medical University of South Carolina, Charleston, SC 29425, USA. [3]Electrophysiology Research Section, Cellular Neurobiology Branch, National Institute on Drug Abuse Intramural Research Program, National Institutes of Health, Baltimore, MD 21224, USA.

*To whom correspondence should be addressed. E-mail: alice_luo@hotmail.com

www.sciencemag.org/cgi/content/full/333/6040/348/DC1

Supporting Online Material for

Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement

Farren J. Isaacs,* Peter A. Carr, Harris H. Wang, Marc J. Lajoie, Bram Sterling, Laurens Kraal, Andrew C. Tolonen, Tara A. Gianoulis, Daniel B. Goodman, Nikos B. Reppas, Christopher J. Emig, Duhee Bang, Samuel J. Hwang, Michael C. Jewett, Joseph M. Jacobson, George M. Church*

*To whom correspondence should be addressed. E-mail: farren.isaacs@yale.edu (F.J.I.); gmc@harvard.edu (G.M.C.)

Published 15 July 2011, *Science* **333**, 348 (2011)
DOI: 10.1126/science.1205822

This PDF file includes:

Supporting Online Material:

Table of Contents:

Materials & Methods

Strains and Culture Conditions

The λ-prophage was obtained from strain DY330 (*1*), modified to include the *bla* gene and introduced into wild-type MG1655 *E.coli* by P1 Transduction at the *bioA*/*bioB* gene locus and selected on ampicillin to yield the strain EcNR1 (λ-Red^{+}). Replacement of *mutS* with the chloramphenicol resistance gene (*chloramphenicol acetyl transferase, cmR*) in EcNR1 produced EcNR2 (mutS^{-}, λ-Red^{+}). EcNR2 was grown in low salt LB-min medium (10 g tryptone, 5 g yeast extract, 5 g NaCl in 1 L dH_2O) for optimal electroporation efficiency and compatibility with zeocin selection. EcNR2 was used as the ancestral strain for all recoded strains reported in this manuscript.

Oligonucleotides

All oligonucleotides were obtained from Integrated DNA Technologies. Oligonucleotides (Table S3) used in the MAGE process were designed according to the following specifications: 1) 90 nucleotides in length, 2) contain a single mutation to effect the TAG to TAA codon conversion, 3) two phosphorothioate linkages at both the 5' and 3' ends to attenuate exonuclease activity and to increase half life, 4) minimize secondary structure (ΔG threshold values, self-folding energy), 5) target lagging strand at the replication fork. No additional purification was used following oligonucleotide synthesis. Primers were purchased from IDT for the multiplex PCR assays and loci sequencing reactions (see description below and Tables S12 and S13).

MAGE-generated Codon Conversions

A single clone of the EcNR2 strain was grown in liquid cell culture, which was used to inoculate 32 separate cultures for parallel modification of all TAG codons. Modification of these codons was achieved through continuous MAGE (*2*) cycling. Each culture was grown at 30°C to mid-logarithmic growth (*i.e.,* OD_{600} of ~0.7) in a rotor drum at 200 RPM. To induce expression of the λ Red recombination proteins (Exo, Beta and Gam), cell cultures were shifted to a 42°C water bath with

vigorous shaking for 15 min and then immediately chilled on ice. In a 4°C environment, 1 mL of cell culture was centrifuged at 16,000x g for 30 seconds. Supernatant media was removed and cells were re-suspended in 1 mL dH_2O (Gibco cat# 15230). This wash process was repeated. Supernatant water was removed the pellet was resuspended in the appropriate pool of 10 oligos (1 uM per MAGE oligo in 50 uL dH_2O). The resuspended oligos/cell mixture was transferred to a pre-chilled 96-well, 2 mm gap electroporation plate (BTX, USA) and electroporated with a BTX electroporation system using the following parameters: 2.5 kV, 200 Ω, and 25 μF. The electroporated cells were immediately transferred to 3 mL of LB-min media for recovery. Recovery cultures were grown at 30 °C in a rotator drum for 2-2.5 hours. Once cells reached mid-logarithmic growth they proceeded to the next MAGE cycle. This approach introduces genomic modifications while allowing cells to evolve and adapt to those changes. Moreover, this approach is designed to explore extensive genotype and phenotype landscapes by creating combinatorial genomic variants that leverage the size of the cell population. After 18 MAGE cycles, cells from each population were isolated on LB-min agar plates. Forty-seven clones from each of the 32 cycled populations were selected and subjected to genotype and phenotype analyses. From each population the clone with the greatest number of modifications (an average of 8 modifications per clone) and minimal aberrant phenotypes (*i.e.,* auxotrophy, decreased fitness)was selected. Further MAGE cycles were employed (typically 6 cycles, but in some cases up to 15) to yield strains with complete sets of 10 targeted modifications.

Genotype Analyses

TAG-to-TAA codon conversions were analyzed using three main methods: 1) Multiplex allele-specific colony PCR (MASC-PCR), 2) Multiplex allele-specific colony quantitative PCR (MASC-qPCR) and 3) Sanger DNA sequencing.

Multiplex Allele Specific Colony PCR (MASC-PCR)

Based on previously described allele-specific PCR techniques, we developed the MASC-PCR method to test for TAG-to-TAA codon conversion in our recoded strains (the ancestral EcNR2 strain was the negative control). Three primers were designed for each locus: 1) a forward primer for the

TAG sequence, 2) a forward primer for the TAA sequence and 3) a reverse primer compatible with both forward primers (Table S12). Primers were designed for a target T_m of 62° C. The two forward primers were identical except that the most 3' nucleotide hybridized to produce either a GC base pair for the wildtype (TAG) codon or an AT base pair for the mutant (TAA) codon. Thus, every clone from each of the 32 populations was interrogated via two MASC-PCR reactions, in which each reaction assayed 10 different loci (with one set assaying four loci). One reaction assayed the wild type (TAG) sequence and a second reaction assayed the mutant (TAA) sequence, yielding two binary reactions that revealed the sequences of the targeted codons (Fig. S4). A clone containing the mutant allele generated PCR products only using the mutant allele primers and not the WT primers and vice versa for a clone with the wild-type allele. To minimize nonspecific amplification of MASC-PCR primers, a gradient PCR was performed to experimentally determine the optimal annealing temperature for each MASC-PCR primer pool (typically between 64 - 67° C). Multiple loci were queried in a single PCR reaction using the multiplex PCR mast2er mix kit from Qiagen. Each MASC-PCR primer set produced amplicon lengths of 100, 150, 200, 250, 300, 400, 500, 600, 700, or 850 bps, corresponding to up to 10 different genomic loci. We found that using a 1:100 dilution of saturated clonal culture in water as template generated the best MASC-PCR specificity. Typical 20 uL MASC-PCR reactions included 10 uL 2x Qiagen multiplex PCR master mix, 0.2 uM of each primer, and 1 uL of template. MASC-PCR cycles were conducted as follows: polymerase heat activation and cell lysis for 15 min at 95° C, denaturing for 30 sec at 94° C, annealing for 30 sec at experimentally determined optimal temperature (64-67° C), extension for 80 sec at 72°C, repeated cycling 26 times, and final extension for 5 min at 72°C. Gel electrophoresis on a 1.5% agarose gel (0.5x TBE) produced the best separation for a 10-plex MASC-PCR reaction. (See Fig. S4 for representative gel picture of MASC-PCR reaction.)

Mulitplex allele-specific quantitative colony PCR (MASC-qPCR)

In complement to MASC-PCR analyses, we also developed a highly multiplexed quantitative PCR screen to rapidly identify highly modified clones (Fig. S3). Typical multiplexed qPCR reactions

employ multiple fluorescence and distinct detection events to assess multiple PCR reactions in one sample, and are generally limited by the available optics and fluorescence to 4 channels. Instead, we needed a robust, economical test that employed many different non-optimized primers, did not require more expensive fluorescently labeled oligos, and would work for 10-plex reactions. We accomplished these goals with SYBR Green I detection, which gauges the total amount of DNA produced in the reaction. Two qPCR reactions were compared for each clone evaluated, one with 10 pairs of primers matched to the unmodified TAG genes, and the other with 10 primer pairs matched to the intended TAA modifications. The TAG reactions were expected to proceed most efficiently with a wild-type template, and the TAA reactions most efficiently with a fully modified template. Intermediate values between these extremes also provided an effective, though nonlinear gauge of the extent of modification for each clone (Fig. S3 A-C).

Each colony was used as template for a pair of qPCR reactions comparing the amplification efficiency when matched to primers terminating in wild-type or targeted mutant sequence. The experimental measurement for a given clone is then compared to the equivalent values measured for the unmodified starting (negative control) strain. This reference value is subtracted from each ΔC_t to yield a $\Delta\Delta C_t$, with unmodified clones scoring close to zero (as with the negative control colonies). The largest $\Delta\Delta C_t$ values were expected to indicate the most modified clones, which we confirmed by genotyping clones with varying $\Delta\Delta C_t$ values (Fig. S3C) Large numbers of clones could be quickly assessed using this approach (up to 190 per 384-well plate, plus 2 negative controls). A typical assessment of MAGE-cycled clones comprised of 4 groups per plate, i.e. for each culture targeting 10 modifications, 2-4 control colonies and 44-46 queried colonies. After identification of the most promising clones, site-specific qPCR genotyping (Fig. S3D) was used to identify which specific sites had been modified, selecting the best clones for further modification.

Individual bacterial colonies were picked into 0.5 mL sterile distilled deionized water, with 5 μL of this suspension used as template in 20 μL qPCR reactions containing 1x NovaTaq buffer, 0.5 U NovaTaq Hotstart DNA Polymerase (EMD Biosciences), 250 μM each dNTP, 0.5x SYBR Green I

(Invitrogen), and 5% DMSO. Primer concentrations were 50 nM for each primer (i.*e*. 500 nM total for 10 forward primers and 500 nM total for 10 reverse primers). A typical qPCR program included a 10 minute hot start at 95° C, followed by 40 cycles (95° C for 30 seconds, 60° C for 30 seconds, 72° C for 30 seconds) finishing with a melt curve analysis. All reactions were performed in a 7900 HT system (Applied Biosystems, Inc.). PCR primers for all sites were designed to have a melting temperature estimated at 62° C. Reverse primers were chosen to yield amplicons in the size range of 200-225 bp. No optimization was needed for qPCR primer sequences or for multiplex/singleplex reaction conditions.

Sanger Sequencing of 314 TAG to TAA loci

DNA sequencing was employed to confirm the results of the above PCR assays and to determine genotypes for 16 sites that gave ambiguous results by MASC-PCR. Amplicons 200-300 bp in length surrounding each of the 314 TAG sites were sequenced from the top-scoring clones by colony PCR as above. Sanger sequencing to confirm allelic replacements was performed by Agencourt Bioscience Corporation and the Biopolymer Facility in the Department of Genetics at Harvard Medical School. Mutations were identified by sequence alignment to the reference MG1655 genome.

Phenotype Analyses

To ensure that the codon replacements did not introduce any significant aberrant phenotypes, we conducted a number of experiments that assessed the fitness of the recoded strains. These experiments included measurements of: 1) strain growth rates, 2) auxotrophic rates and 3) frequency of recombination. Growth rate measurements were obtained by growing replicates of the recoded strains in LB-min media in 96-well plates at 30°C and obtaining OD_{600} measurements using Molecular Devices plate readers (M5 and SpectraMax Plus). Auxotrophic rates were obtained by spotting all clonal isolates (1504) from the MAGE-cycled experiments on M9 minimal media plates (200 mL 5x M9 medium, 1 mL 1 M $MgSO_4$, 5 mL 40% glucose, 100 μL 0.5% vitamin B1 (thiamine), 1 mL D-biotin (0.25 mg/mL), up to 1 L water + 15g Agar). The recombination frequency of each isolate was obtained by

performing the allelic replacement protocol using a lacZ 90-mer oligo that produced a premature stop codon in the chromosomal *lacZ* gene. In general, 250-500 cells were plated on LB-min+Xgal/IPTG (USB Biochemicals) agar plates. Frequency of allelic replacement was calculated by dividing the number of white colonies by the total number of colonies on plates. All phenotypic results are reported in Table S6.

Hierarchical Conjugation Assembly Genome Engineering (CAGE)

Donor and recipient strains were grown in 3 mL LB-min containing the appropriate positive selectable antibiotics. Once cells reached logarithmic-saturated growth, 2 mL samples of each culture were transferred to 2 mL Eppendorf tubes. Cells were washed three times in order to remove antibiotics present in the growth cultures. The washing procedure consisted of centrifuging samples at 5000 rpm for 2 minutes at room temperature, removing the supernatant, and re-suspending the cell pellet in fresh LB-min containing no antibiotics. After the final wash, the donor and recipient pellets were concentrated in 100 μL LB-min in order to enhance cell-cell contact during conjugation. Conjugation was initiated by combining 80 μL of ~20x concentrated donor culture and 20 μL of ~20x concentrated recipient culture. In order to minimize F pilus shearing, cells were gently mixed by pipetting. In order to minimize turbulence that can disrupt cell-cell contact during conjugation and to maximize genome transfer, the entire 100 μL donor-recipient mixture was transferred as a series of 2 x 20 μL and 6 x 10 μL spots onto an LB-min agarose plate lacking antibiotics. This conjugation plate was incubated at 32° C for 0.5-2 hours, then the cells were re-suspended directly off of the plate using 1.5 mL LB-min and concentrated into a final volume of 250 μL. Desired recombinant genomes were selected by inoculating 5 μL of the concentrated post-conjugation culture into LB-min containing the correct combination of positive selection antibiotics (e.g., 10 μg/mL zeocin, 95 μg/mL spectinomycin, and 7.5 μg/mL gentamycin). The conjugated cells that populated the selected culture were then subjected to a negative selection using either *tolC* or *galK* to ensure proper DNA transfer of TAA codons at critical junction points between donor and recipient cells (see Fig. 4).

This engineered conjugation method was tested for the first (1/32 genome, ~143 kb) and last (1/2 genome, ~2.3 Mb) chromosomal transfer steps in the hierarchical assembly experiment (Fig. 1). By selecting for different combinations of markers across the donor and recipient genomes and subsequent screening of specific genomic loci, recombinant clones were isolated that contained the transfer of half or full (otherwise unmodified) genomes at a frequency of ~2.5×10^6 (from a population of 10^9-10^{10} cells), indicating the successful DNA transfer from an integrated *oriT* with episomal expression of conjugal factors. Equivalent frequencies were observed for full genome transfers.

Upon completion of the conjugation process, we also observed the anticipated loss of the *oriT-kan* cassette in the recombinant strain. This observation yields a subtle, yet very useful feature of our engineered conjugation system. By not inheriting the *oriT* sequence, the strains are positioned to proceed to a subsequent conjugation by a one-step integration of the *oriT-Kan* cassette in a new, targeted chromosomal locus (Fig. 4A).

Illumina Whole Genome Sequencing

We prepared a paired-end Illumina sequencing library for three 1/8 genome strains C21 (regions 17-20), C22 (21-24), and C23 (25-28) using barcoded Illumina adapters. The barcoded library was sequenced on one lane using an Illumina GAII.

1. Genome prep (Qiagen Genome Prep kit)
2. Sheared 5 µg of gDNA to target size = 200 bp using covaris (estimated median band size 250 bp)
3. PCR purified DNA (QIAquick PCR purification kit)
4. End repair (Epicentre End-it™ DNA End-Repair kit)

Component	Volume (1x)
DNA sample	35
10x End-Repair Buffer	10
1 mM dNTPs	10
End-Repair Enzyme mix	5

dH2O	40
Total (µL)	100

Incubate at 25 C for 30 minutes

5. PCR purified DNA (QIAquick PCR purification kit)
6. A-tailing (NEB Klenow Fragment (3'->5' exo-))

Component	Volume (1x)
DNA sample	32
Klenow buffer	5
1 mM dATP	10
Klenow (3'->5' exo-)	3
Total (µL)	50

Incubate for 30 minutes at 37 C

7. PCR purified DNA (QIAquick PCR purification kit)
8. Adapter ligation (adapters complements of Morten Sommer)
    a. C21: TopPE-1 barcode = AGC
    b. C22.DO:T: TopPE-3 barcode = CTA
    c. C23: TopPE-4 barcode = TCT

Component	Volume (1x)
DNA sample	31
Rapid ligase buffer (2x)	35
PE adapter (50 µM)	2
Enzymatics Rapid (T4) ligase	2

Incubated at 20 C for exactly 10 minutes in a thermocycler, then immediately added PBI for the PCR purification

9. PCR purification (Qiagen MinElute PCR purification kit)
10. Gel purified adapter-ligated sequencing libraries on 2% agarose gel in 0.5x TBE (cut 2 mm bands corresponding to approximately 225 bp) (Qiagen Gel Purification kit)
11. PCR amplified sequencing libraries (2x KAPA HiFi Ready Mix; 11 cycles)
    a. Standard Illumina PE PCR Primers

Component	Volume (1x)
2X KAPA HiFi Ready Mix	25
PE_PCR-f	1
PE_PCR-r	1
dH2O	13
Template	4

Step	Temp	Time (min)
1	95	5:00
2	98	0:20
3	62	0:15
4	72	1:15
5	Go to step 2	11x
6	72	3:00
7	4	Forev

12. PCR purified (QIAquick PCR purification kit)
13. Validated sequencing library
    a. Cloned Illumina libraries using Invitrogen TOPO ZeroBlunt II
    b. Transformed into OneShot Top 10 electrocompetent cells
    c. Genewiz sequenced insert (Sanger sequencing; Seq. primer = M13 forward (Invitrogen): GTAAAACGACGGCCAG)
14. Size-selected sequencing libraries for ~225 bp bands (E-Gel® SizeSelect™ gels)
15. PCR purified Illumina libraries (Qiagen MinElute)
16. Quantitated contents of C21, C22.DO:T, and C23 libraries
    a. PAGE, Low DNA Mass Ladder (Invitrogen), and SYBR gold staining
    b. Densitometry
17. Prepared sequencing library by adding all 3 components to a final concentration of 10 nM
18. Sample QC, Clustering, and sequencing performed by BPF
    a. Standard Illumina PE Sequencing Primers

Genome Assembly and Sequence Analysis

Read Sorting and Processing

The raw Illumina reads in FASTQ format were preprocessed and sorted using the 6-bp barcodes in the paired end adaptors. Reads that contained anomalous barcodes were discarded. Reads containing any bases with a quality score of 2, also called the *Read Segment Quality Control Indicator (*based on *Illumina Quality Scores* by Tobia Man*)*, were discarded at this step, but all other reads were kept. After preprocessing, all reads were exactly 34 base pairs long.

Reference-based Assembly

The expected FASTA sequence of the EcNR2 parent strain was assembled by manually modifying the FASTA sequence of E. coli K-12 strain MG1655 to reflect the removal of mutS and the insertion of the lambda prophage genome into the bioAB operon. Next, the preprocessed reads were sorted into separate files by pair group and the *Burrows-Wheeler Aligner* program (*BWA*) (*3*) was used to separately align the paired reads from each of the three strains to the expected EcNR2 FASTA sequence. The *sample* algorithm was used to align the reads. The distribution of insert sizes was inferred at runtime. During the read alignment step of *BWA*, (the *aln* command), a value of 10 was used for the suboptimal alignment cutoff.

Indel and SNP Filtering

After alignment, the *SAMtools* package (*4*) was used to create and sort BAM files for the assemblies. From these BAM files we generated a set of raw SNPs and short indels with respect to this reference assembly. These were then filtered using several criteria. First, using the *varFilter* script within *SAMtools*, we removed SNPs where the root mean squared mapping quality was less than 10, and indels where the root mean squared mapping quality was less than 25. We fitted the read coverage of each assembly to a gamma distribution and used the 99.95th and 0.05th percentile cutoffs for minimum and maximum read depth, beyond which SNPs and indels were discarded. We also discarded SNPs within 3 base pairs of a gap, and SNPs that occurred more densely than three within one 10 base pair window.

Region Masking

We used custom scripts to further filter SNPs and indels by masking regions of poor assembly. We masked regions containing many truncated reads, many incorrect read pairings, many non-unique alignments, and regions with motifs known to be problematic in Illumina sequencing (GGCnG). We defined truncated read regions as those containing multiple incompletely mapped reads, separated by less than one read length, containing at least 4 truncated reads and having a number of truncated reads totaling at least one half of the length of the contiguous region in which they were found.

Regions with incorrect read pairings were defined using the following method. We found read pairs whose insert size was outside of the 99.9th and 0.1th percentile of a fitted normal distribution of mate pair distance. These reads were counted in a 34-bp rolling window. As a thresholding step we chose contiguous regions where 10 or more of these reads were found in one window length. Additionally included were contiguous regions where only one read in a pair could be mapped, and these were thresholded with a rolling window in a similar fashion, using a 6-read cutoff. As a final masking step, we removed SNPs stemming from the replacement of amber stop codons as well as SNPs and indels where surrounding context was *GGCnC*, as these regions are known to be hotspots for Illumina sequencing errors.

Annotation

After removing SNPs and indels in the masked regions as described above, we attempted to associate the remaining SNPs and indels with functional consequences. We used a modified version of *Ensembl's SNP Effect Predictor* software (*5*), and the *Ensembl Bacteria* database to find SNPs that occurred within genes. We further categorized these by synonymous and non-synonymous coding changes, frameshift mutations, premature stop mutations, mutations in the 5' and 3' UTRs, and mutations less than 100 base pairs upstream of a transcript start site (Fig. S8). Coordinates were lifted over from ECNR2 to MG1655 to permit annotation of the SNPs and indels. This resulted in C21, C22, and C23 having 4,5, and 5 mutations respectively having no corresponding liftover coordinates in ECNR2. These are referred to as the "unmappable" in Figure S8.

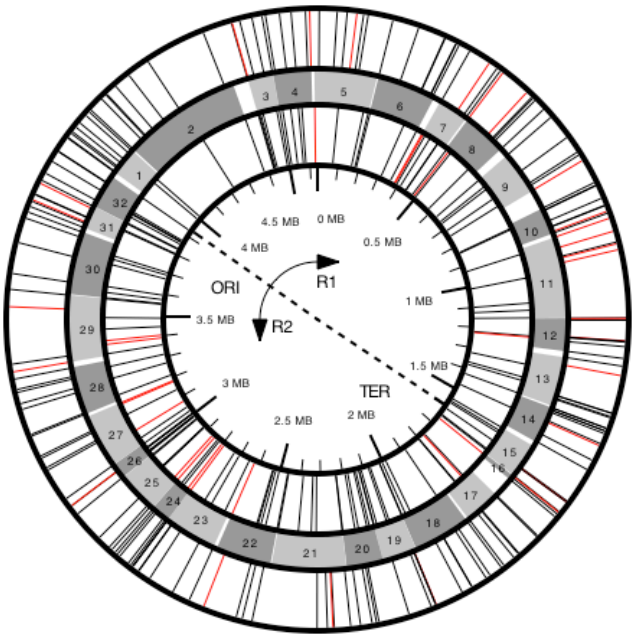


Figure S1. Circular representation of the *E. coli* genome depicting the coordinates and orientation of all TAG codons. TAG codons found in essential genes(*6*) are shown in red. The outer ring plots all clockwise transcribed TAG codons on the + DNA strand whereas the inner ring plots all counterclockwise transcribed TAG codons on the – DNA strand. The middle ring depicts the 32 sections of the genome targeted for TAG-to-TAA conversion. The inner circle plots the genomic coordinates, origin of replication (ORI), terminus (TER) and replichores 1 (R1) and 2 (R2).

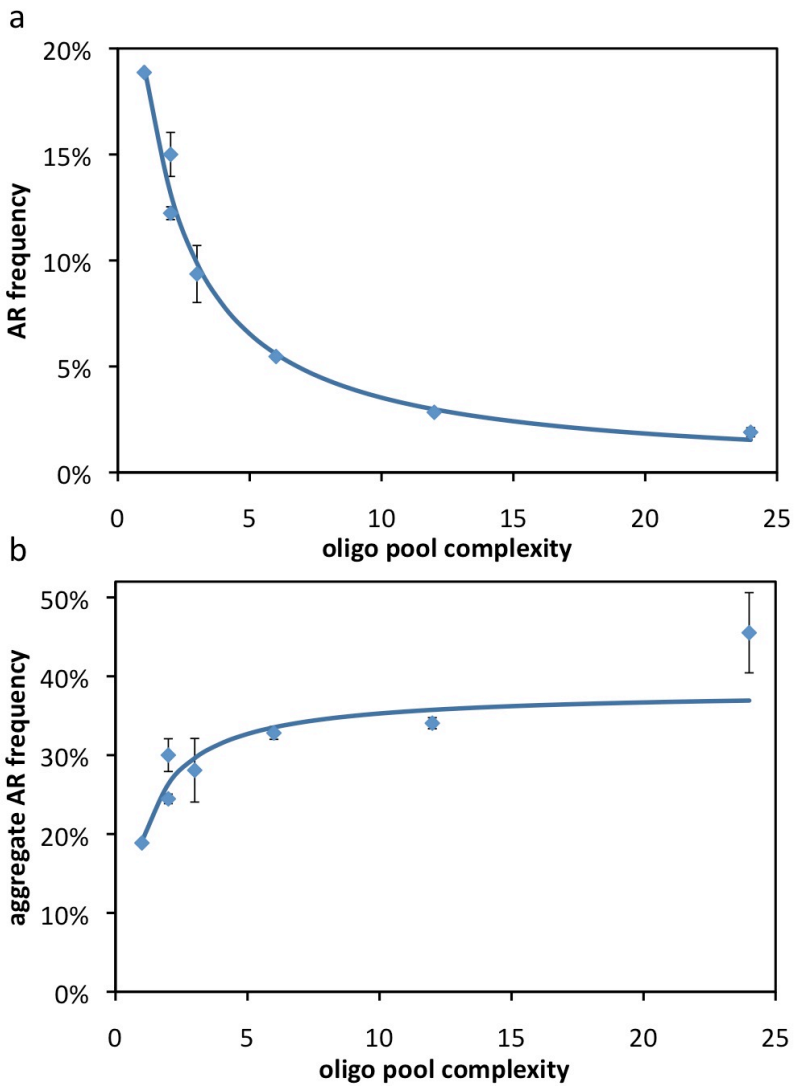


Figure S2. Multiplexing in MAGE leads to higher aggregate allele replacement (AR) frequencies. Multiplex oligo recombineering experiments were performed with equimolar oligo pools ranging in complexity from one to 24 oligos. AR frequencies were quantified for one conversion site corresponding to one oligo present in all pools. While individual AR frequencies (a) decrease as a function of higher complexity, the overall aggregate frequency (estimated as the product of individual frequency and pool complexity) (b) increases. Allele frequencies were measured using MASC-qPCR and curves are fit to the formula $y=a(1-e^{-b/x})$ for plot a and $y=ax(1-e^{-b/x})$ for plot b. Error bars indicate standard error (n=2).

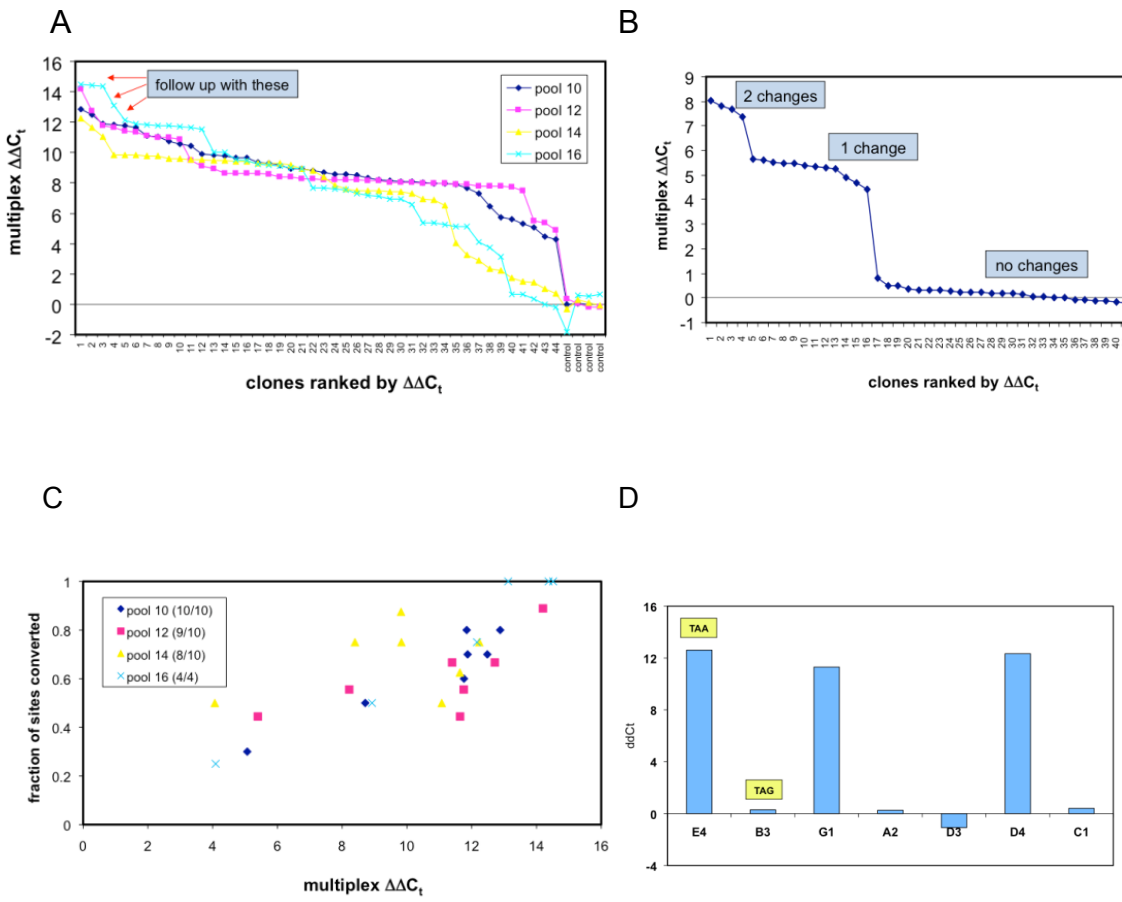


Figure S3. Multiplex allele-specific colony quantitative PCR (MASC-qPCR) to rapidly screen for the most modified clones. Multiplex reactions were used to determine which clones in the complex pool had been most highly modified. **A**. Clones from cultures receiving 18 cycles of MAGE processing (pools of 10 oligos) were sorted by their mutiplex $\Delta\Delta C_t$ scores. Small numbers of top-scoring clones (typically 3-5) were then assessed at each TAG site of interest. **B**. When only two modifying oligos are used for allele replacement, mutiplex $\Delta\Delta C_t$ values are more visibly clustered into groups representing 0, 1, or 2 modifications. **C**. Correlation between mutiplex $\Delta\Delta C_t$ scores from (A) and the number of specific modifications achieved. The top mutiplex $\Delta\Delta C_t$-scoring clone was found to have the most allele conversions roughly 70% of the time. The legend indicates the number of modifications observed in the top-scoring clones. **D**. Singleplex reactions were used to genotype the most promising clones. Shown are 7 clones assayed at the tfaS stop codon, with singleplex $\Delta\Delta C_t$ values of 0.0±0.7 for wild-type TAG and 12.1±0.7 for modified TAA.

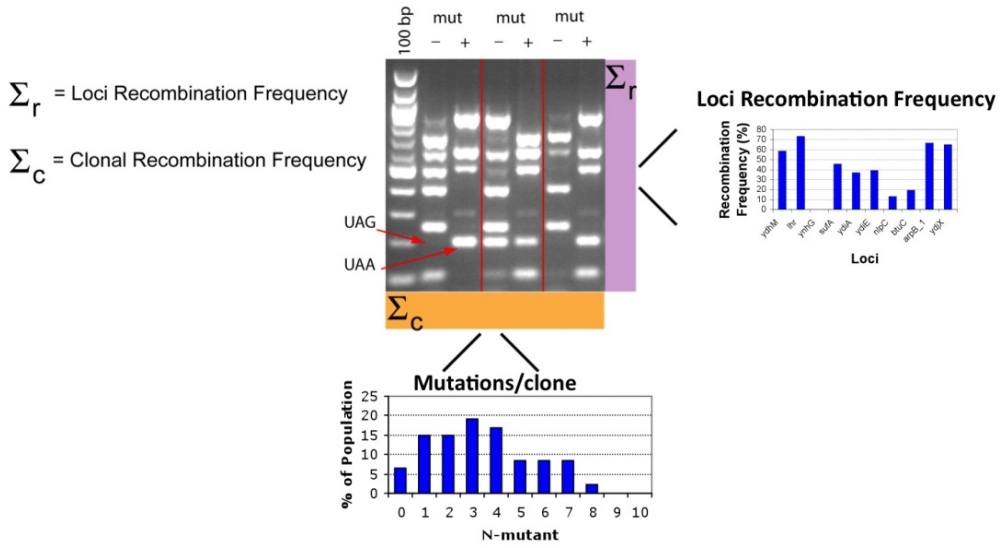


Figure S4. Multiplex allele-specific colony PCR (MASC-PCR) for the detection of codon conversions in each clone. The gel illustrates results from 3 MAGE-cycled clones. Ten sites are investigated in each lane where two MASC-PCRs are conducted for each clone: one reaction interrogates the TAG loci (-) and another reaction interrogates the TAA loci (+). Each reaction provides a binary output through the presence or absence of an amplicon band. Together, both TAG and TAA reactions provide sufficient information to determine the conversion status of a given codon. Summation of the rows of the 46 clones provides loci frequency data for each codon (plotted in Figure 2 in the main text). Summation of the columns of each clone provides a histogram of the mutations per clone (plotted in Fig. 3 in the main text).

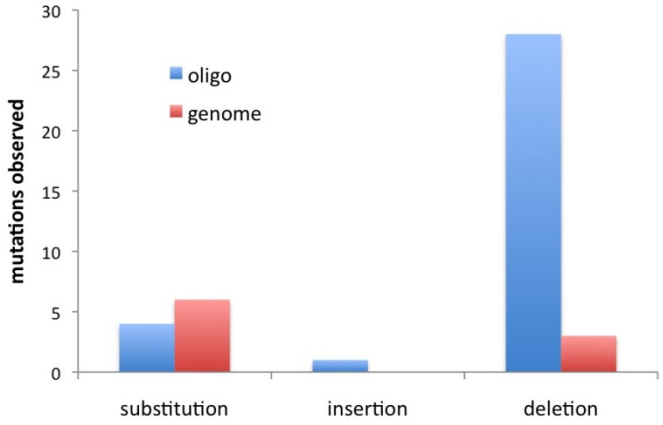


Figure S5. Off-target mutation in modified strains. From 96.3 kb of total Sanger sequencing, the majority of unwanted mutations were observed in regions corresponding to the annealing sites for the 90-mer oligonucleotides (blue). Sequenced mutations falling outside these 90 bp regions are shown in red. The principal error is a deletion, mostly single base deletions. These errors correspond to common defects arising from oligo synthesis.

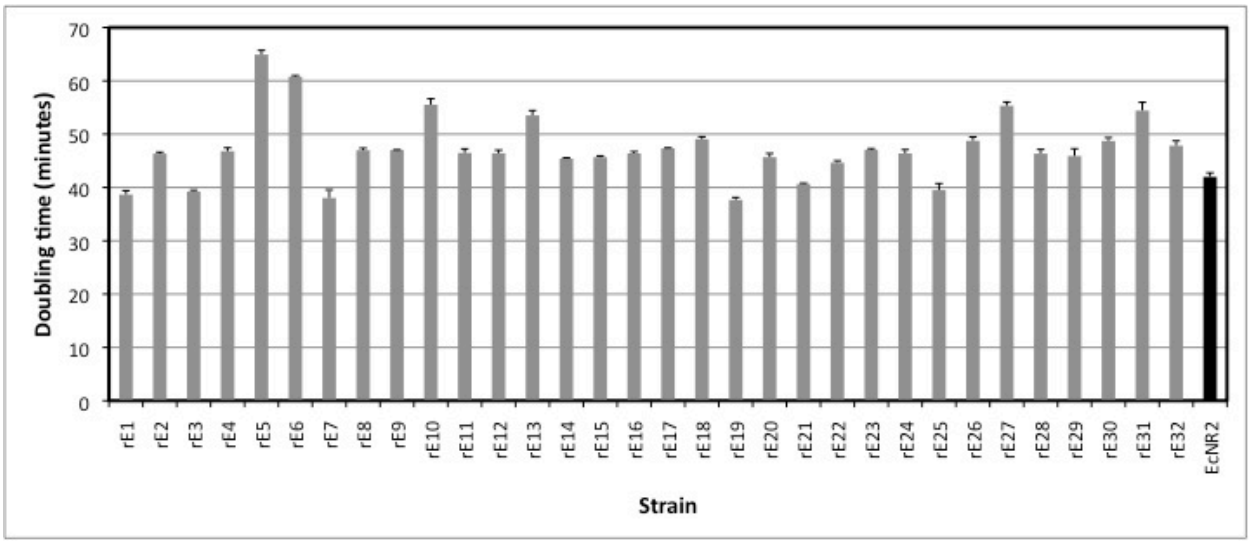


Figure S6. Individual growth rates of all 32 "top" recoded strains (10 out of 10 changes each, grey, rE1-rE32) following the successful replacement of all TAG-to-TAA codons versus that of the ancestral strain (black, EcNR2). A mix of increased and decreased growth rates was observed across the 32 strains with an average of 47 minutes/division. This is a mild decrease versus the growth rate (42 minutes/division) of the ancestral strain. Our parallelized MAGE approach across 32 strains allows us to easily identify strains with notable growth phenotypes (e.g., rE5, rE6). These strains can be investigated further to determine if these growth impediments are due to the codon changes or whether they arise from secondary mutations elsewhere in the genome

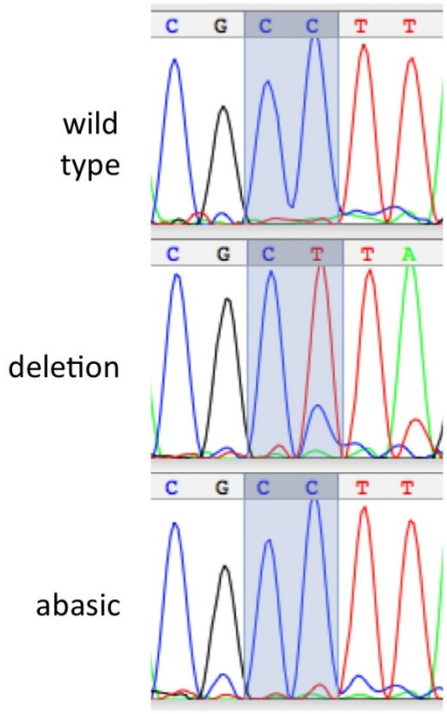
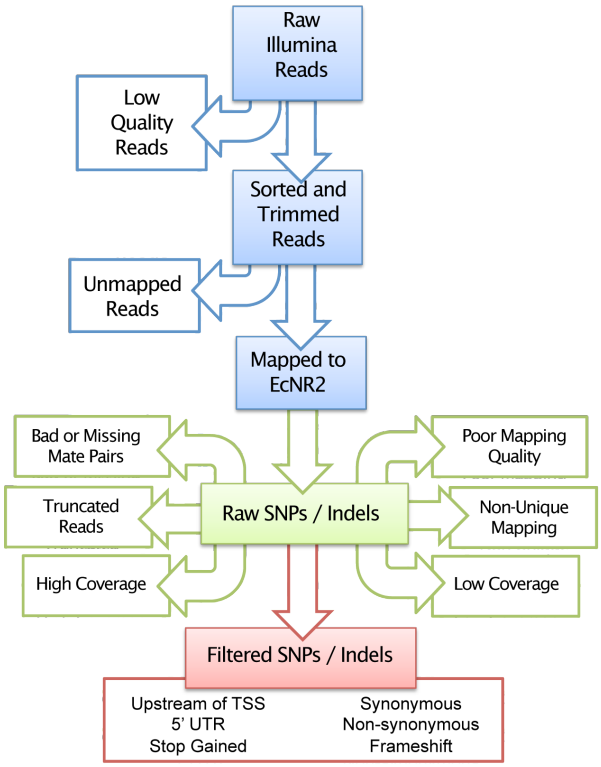


Figure S7. Some oligo defects can be fixed by host repair mechanisms. Oligos containing internal deletions are likely to result in equivalent mutations in the genome, but oligo synthesis chemistry can be optimized to minimize such deletions. In such cases, measures such as more aggresive deprotection and coupling conditions can then give rise to damaged oligos containing abasic sites. However, this second type of defect is readily repaired in the host. Three similar purified oligos were used to modify the selectable chromosomal tolC gene in separate cultures. An upstream modification (not shown) in each oligo creates a stop codon in tolC—selection against the tolC protein ensures only cells that have incorporated this oligo survive. PCR amplification of the resulting population and sequencing of this potentially heterogeneous product allows assessment of the effect of modifications at a second site. Top: only the initial stop codon was employed for this oligo, leaving the wild-type sequence C97-T102 of the tolC gene. Middle: this oligo coded for a deletion of C99, effectively shifting the subsequent peaks left one base position. However, a notable fraction (less than one-third) has not been shifted, indicating possible repair events (this fraction is very unlikely to arise from a defect in the oligo). Bottom: this oligo contained an abasic site at position 99, but the resulting population is almost completely wild-type, indicating likely cellular repair. These experiments were performed in strain EcNR2, which includes a deletion of the mutS mismatch repair gene.



	C21	C22	C23	Total
Raw Reads	-	-	-	50,050,024
Low Quality Reads	-	-	-	10,049,324
Sorted and Trimmed	11,753,698	16,353,342	11,893,660	40,000,700
Unmapped Reads	448,348	506,335	413,368	1,368,051
Mapped to EcNR2	11,305,350 (96.19%)	15,847,007 (96.90%)	11,480,292 (96.52%)	38,632,64 (96.58%)
Raw SNPs / Indels	269	323	227	819
Poor Mapping Quality	108	148	75	331
Bad/Missing Mates	11	11	10	32
Truncated Reads	4	10	7	21
High Coverage	2	0	2	4
Low Coverage	3	5	3	11
Non-Unique Mapping	7	10	10	27
Filtered SNPs / Indels	110	128	102	340
Synonymous	29	33	24	86
Non-synonymous	50	46	41	137
Frameshift	11	11	10	32
Stop gained	2	4	2	8
5' UTR	2	3	3	8
Upstream of Transcript	2	4	1	7
Within Noncoding	2	8	4	14
Unmappable	4	5	5	14
No annotation	8	14	12	34

Figure S8. Bioinformatics sequencing analysis process and secondary mutation breakdown. We sequenced the entire genomes of three 1/8 recoded strains and identified off-target SNPs and short (1 bp) indels using BWA sequencing alignment (*3*). On-target TAG conversions are not included in this analysis. We found an average of 113 mutations/genome after each strain went through approximately 960 doublings, multiple lambda red inductions, and several conjugations. This corresponds to 2E-8 mutations/bp/doubling, which is consistent with the predicted basal mutation rate of the ancestral strain (EcNR2). These results indicate that MAGE and CAGE do not significantly compromise genome stability. Also see Fig. S9 and Tables S8-S11 for supporting data and information.

.

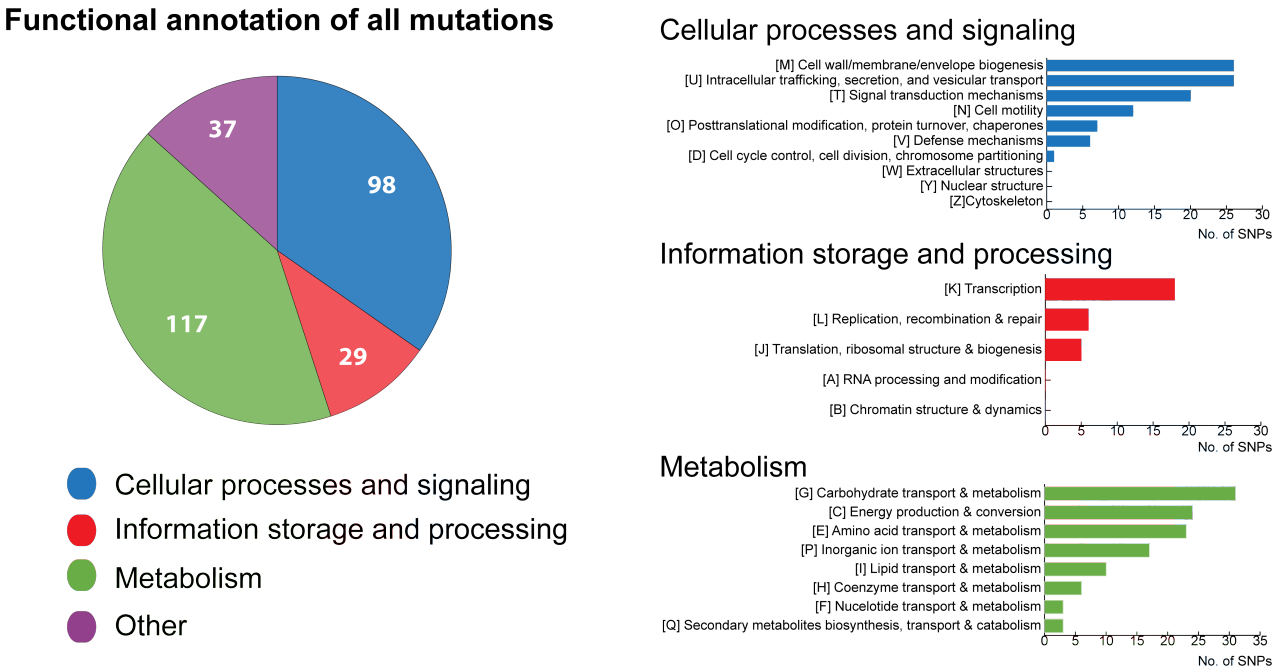


Figure S9 Functional annotation of all mutations as indicated by the COG category of the gene or regulatory region associated with the SNP or indel. Functional annotation of all mutations are summed across the three sequenced genomes. See Table S8 for complete list of secondary mutations.

Table S1. List of genes containing TAG codon and TAG genomic coordinate.

Genome Size = 4639675

Set		Gene	Essential	Direction	Gene.up	Gene.lo	Region.Size	TAG.coordinate
SIR.32.1					**3920623**	**3921005**	**383**	
	1	rbsB	0	forward	3934301	3935191	**891**	3935191
	1	rbsR	0	forward	3936250	3937242	**993**	3937242
	1	ilvL	0	forward	3948345	3948443	**99**	3948443
	1	ilvA	0	forward	3953354	3954898	**1545**	3954898
	1	wzzE	0	forward	3967054	3968100	**1047**	3968100
	1	aslB	0	forward	3980981	3982216	**1236**	3982216
	1	hemY	0	complemer	3984709	3985905	**1197**	3984709
	1	yigG	0	complemer	4000836	4001216	**381**	4000836
	1	recQ	0	forward	4003887	4005716	**1830**	4005716
	1	rmuC	0	forward	4015356	4016783	**1428**	4016783
SIR.1.2					**4019889**	**4019968**	**80**	
	2	tatD	0	forward	4021577	4022359	**783**	4022359
	2	yihL	0	forward	4058470	4059180	**711**	4059180
	2	yihW	0	forward	4072692	4073477	**786**	4073477
	2	yiiD	0	forward	4075472	4076461	**990**	4076461
	2	cpxP	0	forward	4103843	4104343	**501**	4104343
	2	frwD	0	forward	4145161	4145502	**342**	4145502
	2	zraR	0	forward	4201343	4202668	**1326**	4202668
	2	yjaB	0	complemer	4211703	4212146	**444**	4211703
	2	yjdK	0	forward	4350607	4350903	**297**	4350903
	2	blc	0	complemer	4375212	4375745	**534**	4375212
SIR.2.3					**4417723**	**4417928**	**206**	
	3	ulaF	0	forward	4421723	4422409	**687**	4422409
	3	yjfY	0	complemer	4422539	4422814	**276**	4422539
	3	priB	1	forward	4423543	4423857	**315**	4423857
	3	ytfB	0	complemer	4426102	4426740	**639**	4426102
	3	ytfG	0	complemer	4431187	4432047	**861**	4431187
	3	msrA	0	complemer	4439561	4440199	**639**	4439561
	3	ytfN	0	forward	4442135	4445914	**3780**	4445914
	3	b4250	0	phantom g	4472250	4472414	**165**	4472414
	3	idnR	0	complemer	4488164	4489162	**999**	4488164
	3	b4273	0	forward, ins	4496618	4497523	**906**	4497523
SIR.3.4					**4497523**	**4497625**	**103**	
	4	yjgX	0	complemer	4498066	4498512	**447**	4498066
	4	b4283	0	forward, ins	4505220	4505486	**267**	4505486
	4	fecE	0	complemer	4508713	4509480	**768**	4508713
	4	nanC	0	complemer	4536808	4537524	**717**	4536808
	4	fimB	0	forward	4538980	4539582	**603**	4539582
	4	yjiN	0	complemer	4563989	4565269	**1281**	4563989
	4	yjiP	0	forward, yji	4567021	4567332	**312**	4567332
	4	mcrB	0	complemer	4575981	4577360	**1380**	4575981

	4	yjjQ	0	forward	4601500	4602225	**726**	4602225
	4	yjjV	1	forward	4611504	4612283	**780**	4612283
SIR.4.5					**4612358**	**4612628**	**271**	
	5	lplA	1	complemer	4621124	4622140	**1017**	4621124
	5	creC	0	forward	4634719	4636143	**1425**	4636143
	5	kefF	0	forward	47246	47776	**531**	47776
	5	yabI	0	forward	71351	72115	**765**	72115
	5	thiQ	0	complemer	72229	72927	**699**	72229
	5	murF	1	forward	94650	96008	**1359**	96008
	5	mutT	0	forward	111044	111433	**390**	111433
	5	pdhR	0	forward	122092	122856	**765**	122856
	5	sfsA	0	complemer	160782	161486	**705**	160782
	5	btuF	0	complemer	177662	178462	**801**	177662
SIR.5.6					**182308**	**182445**	**138**	
	6	cdaR	0	forward	182463	183620	**1158**	183620
	6	yaeI	0	complemer	184257	185069	**813**	184257
	6	yafL	0	forward	246712	247461	**750**	247461
	6	yagY	0	complemer	308582	309250	**669**	308582
	6	eaeH	0	forward	313581	314468	**888**	314468
	6	b0299	0	forward, ins	314811	315677	**867**	315677
	6	yahL	0	forward	343400	344215	**816**	344215
	6	yahM	0	forward	344628	344873	**246**	344873
	6	prpE	0	forward	351930	353816	**1887**	353816
	6	cynR	0	complemer	357015	357914	**900**	357015
SIR.6.7					**374180**	**374608**	**429**	
	7	frmR	1	complemer	378830	379105	**276**	378830
	7	b0361	0	forward, ins	380898	381803	**906**	381803
	7	yaiS	1	complemer	383283	383840	**558**	383283
	7	tra5_1	0	complemer	390963	391829	**867**	390963
	7	araJ	0	complemer	410521	411705	**1185**	410521
	7	pgpA	1	forward	435813	436331	**519**	436331
	7	yajL	0	complemer	442275	442865	**591**	442275
	7	panE	0	complemer	442828	443739	**912**	442828
	7	lon	0	forward	458112	460466	**2355**	460466
	7	mdlA	0	forward	468095	469867	**1773**	469867
SIR.7.8					**474460**	**474528**	**69**	
	8	ybaA	1	forward	475896	476249	**354**	476249
	8	ybaJ	0	complemer	479558	479932	**375**	479558
	8	priC	1	complemer	489509	490036	**528**	489509
	8	aes	0	complemer	498238	499197	**960**	498238
	8	ybcF	1	forward	549662	550555	**894**	550555
	8	tra5_2	0	forward, ins	566361	567227	**867**	567227
	8	ybcL	0	forward	570116	570667	**552**	570667
	8	rzpD	0	forward	577330	577791	**462**	577791
	8	ybcY	0	complemer	581375	582029	**655**	581375
	8	hokE	0	forward	607059	607211	**153**	607211
SIR.8.9					**608475**	**608607**	**133**	
	9	ybdL	0	forward	632809	633969	**1161**	633969

	9	ybdM	0	complemer	633970	634599	**630**	633970
	9	rlpA	0	complemer	663325	664413	**1089**	663325
	9	ybeQ	0	complemer	674793	675770	**978**	674793
	9	ubiF	0	forward	694324	695499	**1176**	695499
	9	ybfH	0	compleme	715611	715820	**210**	715611
	9	ybfD	0	forward	737315	738076	**762**	738076
	9	nei	0	forward	745158	745949	**792**	745949
	9	abrB	0	complemer	745946	746992	**1047**	745946
	9	sucB	1	forward	760745	761962	**1218**	761962
SIR.9.10					**769909**	**770318**	**410**	
	10	bioD	0	forward	811493	812170	**678**	812170
	10	moaE	0	forward	818518	818970	**453**	818970
	10	ybhM	0	forward	820016	820729	**714**	820729
	10	ybhS	0	complemer	825342	826475	**1134**	825342
	10	ybiH	0	complemer	829195	829866	**672**	829195
	10	ybiR	0	forward	852870	853988	**1119**	853988
	10	yliD	0	forward, gs	871113	872024	**912**	872024
	10	yliE	0	forward	872202	874550	**2349**	874550
	10	ybjK	1	forward	886646	887182	**537**	887182
	10	rimK	0	forward	891190	892092	**903**	892092
SIR.10.11					**892731**	**892782**	**52**	
	11	ybjR	1	forward	904136	904966	**831**	904966
	11	cspD	0	complemer	921589	921813	**225**	921589
	11	lolA	1	forward	936595	937206	**612**	937206
	11	ycaI	0	forward	963543	965807	**2265**	965807
	11	lpxK	1	forward	967589	968575	**987**	968575
	11	ycbK	1	forward	982298	982846	**549**	982846
	11	ycbX	0	complemer	1005714	1006823	**1110**	1005714
	11	tra5_3	0	complemer	1093498	1094364	**867**	1093498
	11	yceO	0	complemer	1118530	1118670	**141**	1118530
	11	yceF	0	complemer	1145234	1145818	**585**	1145234
SIR.11.12					**1146763**	**1146844**	**82**	
	12	plsX	0	forward	1146844	1147914	**1071**	1147914
	12	fabH	1	forward	1147982	1148935	**954**	1148935
	12	pabC	0	forward	1152523	1153332	**810**	1153332
	12	b1146	1	forward, yn	1202247	1202447	**201**	1202447
	12	ymfM	0	forward	1203045	1203383	**339**	1203383
	12	ymfR	0	forward	1204772	1204954	**183**	1204954
	12	tfaE	0	complemer	1207740	1208342	**603**	1207740
	12	ycgX	1	complemer	1211926	1212330	**405**	1211926
	12	ymgC	0	forward	1215971	1216219	**249**	1216219
	12	ycgY	0	forward	1244383	1244823	**441**	1244823
SIR.12.13					**1255250**	**1255869**	**620**	
	13	hemA	0	forward	1262937	1264193	**1257**	1264193
	13	b1228	1	forward, yc	1285932	1286207	**276**	1286207
	13	kch	0	complemer	1307040	1308293	**1254**	1307040
	13	sapB	0	complemer	1352529	1353494	**966**	1352529
	13	pspF	0	complemer	1364959	1365936	**978**	1364959

	13 ycjU	0 forward	1378172	1378831	**660**	1378831
	13 ycjV	0 forward	1378845	1379926	**1082**	1379926
	13 ycjG	0 forward	1386954	1387919	**966**	1387919
	13 ycjZ	0 forward	1390015	1390914	**900**	1390914
	13 intR	0 complemer	1410024	1411259	**1236**	1410024
SIR.13.14			**1416328**	**1416497**	**170**	
	14 sieB	0 forward	1416695	1417183	**489**	1417183
	14 b1354	0 forward (nc	1417192	1417368	**177**	1417368
	14 ydaY	0 forward	1423645	1424004	**360**	1424004
	14 b1367	0 forward	1417192	1417368	**177**	1417368
	14 tfaR	0 forward	1430435	1431010	**576**	1431010
	14 paaG	0 forward	1456288	1457076	**789**	1457076
	14 paaX	1 forward	1461563	1462513	**951**	1462513
	14 b1402	0 complemer	1465945	1466850	**906**	1465945
	14 ydcL	0 forward	1500481	1501149	**669**	1501149
	14 rhsE	0 forward	1525914	1527962	**2049**	1527962
SIR.14.15			**1529675**	**1529765**	**91**	
	15 yddJ	0 complemer	1542408	1542743	**336**	1542408
	15 ddpF	0 complemer	1555136	1556062	**927**	1555136
	15 yddV	0 complemer	1563782	1565164	**1383**	1563782
	15 ydeV	0 complemer	1596641	1598233	**1593**	1596641
	15 ydeA	0 forward	1615052	1616242	**1191**	1616242
	15 marA	1 forward	1617598	1617981	**384**	1617981
	15 marB	0 forward	1618013	1618231	**219**	1618231
	15 tfaQ	0 complemer	1632334	1632909	**576**	1632334
	15 ydfP	0 complemer	1637054	1637551	**498**	1637054
	15 rzpQ	0 forward	1646847	1647065	**219**	1647065
SIR.15.16			**1647140**	**1647331**	**192**	
	16 b1578	1 InsD, forwa	1648905	1649561	**657**	1649561
	16 ynfA	1 complemer	1653371	1653697	**327**	1653371
	16 speG	0 forward	1654208	1654768	**561**	1654768
	16 dmsD	0 forward	1662530	1663144	**615**	1663144
SIR.16.17			**1676002**	**1676376**	**375**	
	17 ydhM	0 forward	1724047	1724646	**600**	1724646
	17 lhr	0 forward	1727111	1731727	**4617**	1731727
	17 ynhG	0 complemer	1755745	1756749	**1005**	1755745
	17 sufA	1 complemer	1762042	1762410	**369**	1762042
	17 ydiA	0 forward	1785469	1786302	**834**	1786302
	17 ydiE	0 forward	1787637	1787828	**192**	1787828
	17 nlpC	0 complemer	1790291	1790755	**465**	1790291
	17 btuC	0 complemer	1792196	1793176	**981**	1792196
	17 arpB_1	0 forward	1801118	1801591	**474**	1801591
	17 ydjX	0 forward	1831425	1832135	**711**	1832135
SIR.17.18			**1839427**	**1839513**	**87**	
	18 nudG	0 forward	1839514	1839921	**408**	1839921
	18 b1788	0 complemer	1872102	1872206	**105**	1872102
	18 yeaL	0 forward	1872376	1872822	**447**	1872822
	18 yeaX	0 forward	1883869	1884834	**966**	1884834

	18	yobB	0	forward	1923464	1924120	**657**	1924120
	18	exoX	0	forward	1924144	1924806	**663**	1924806
	18	yecN	0	forward	1950290	1950685	**396**	1950685
	18	otsA	0	complemer	1978212	1979636	**1425**	1978212
	18	dcyD	0	complemer	1996518	1997504	**987**	1996518
	18	yedM	0	complemer	2010025	2010375	**351**	2010025
SIR.18.19					**2010450**	**2010649**	**200**	
	19		0	complemer	2010724	2011038	**315**	2010724
	19	fliN	1	forward	2019112	2019525	**414**	2019525
	19	fliP	0	forward	2019893	2020630	**738**	2020630
	19	fliQ	0	forward	2020640	2020909	**270**	2020909
	19	yedS_1	0	forward	2032075	2032560	**486**	2032560
	19	cbl	0	complemer	2057988	2058938	**951**	2057988
	19	b1996	0	complemer	2066976	2067881	**906**	2066976
	19	hisL	0	forward	2088020	2088070	**51**	2088070
	19	wbbJ	0	complemer	2102518	2103108	**591**	2102518
	19	wcaM	0	complemer	2112526	2113920	**1395**	2112526
SIR.19.20					**2113920**	**2113931**	**12**	
	20	wcaL	0	complemer	2113931	2115151	**1221**	2113931
	20	wcaC	0	complemer	2128877	2130094	**1218**	2128877
	20	asmA	0	complemer	2137783	2139636	**1854**	2137783
	20	baeR	0	forward	2162300	2163022	**723**	2163022
	20	gatR_1	0	complemer	2169419	2169751	**333**	2169419
	20	tra5_4	0	forward, In	2168556	2169422	**867**	2169422
	20	yegV	0	forward	2179118	2180083	**966**	2180083
	20	yegW	0	complemer	2180057	2180803	**747**	2180057
	20	yehQ	0	forward	2207122	2209122	**2001**	2209122
	20	yohC	0	complemer	2223066	2223653	**588**	2223066
SIR.20.21					**2223728**	**2223748**	**21**	
	21	yohF	0	complemer	2224531	2225292	**762**	2224531
	21	sanA	0	forward	2230900	2231619	**720**	2231619
	21	yejA	0	forward	2270386	2272200	**1815**	2272200
	21	yejE	1	forward	2273295	2274320	**1026**	2274320
	21	b2191	0	forward (nc	2286927	2287049	**123**	2287049
	21	rcsD	0	forward	2311510	2314182	**2673**	2314182
	21	rcsC	0	complemer	2315049	2317898	**2850**	2315049
	21	yfaT	0	complemer	2332358	2332981	**624**	2332358
	21	menF	0	complemer	2377370	2378665	**1296**	2377370
	21	yfcO	0	complemer	2447250	2448071	**822**	2447250
SIR.21.22					**2450358**	**2450378**	**21**	
	22	yfcU	0	complemer	2450378	2453023	**2646**	2450378
	22	tfaS	0	forward	2468837	2469127	**291**	2469127
	22	ypdI	0	forward	2492720	2492995	**276**	2492995
	22	yfdY	0	complemer	2493072	2493314	**243**	2493072
	22	yfeO	0	forward	2507652	2508908	**1257**	2508908
	22	mntH	0	complemer	2509490	2510728	**1239**	2509490
	22	xapR	0	complemer	2519615	2520499	**885**	2519615
	22	yfeR	0	complemer	2523952	2524878	**927**	2523952

	22	yffB	1	forward	2589269	2589625	**357**	2589625
	22	hda	1	complemer	2616097	2616843	**747**	2616097
SIR.22.23					**2627035**	**2627237**	**203**	
	23	yfgG	0	forward	2627312	2627503	**192**	2627503
	23	pbpC	0	complemer	2643035	2645347	**2313**	2643035
	23	yphA	0	forward	2671368	2671790	**423**	2671790
	23	yfhB	0	complemer	2695937	2696572	**636**	2695937
	23	kgtP	0	complemer	2722470	2723768	**1299**	2722470
	23	yfiA	0	forward, rai	2735176	2735517	**342**	2735517
	23	yfjQ	0	forward	2766687	2767508	**822**	2767508
	23	yfjR	0	forward	2767725	2768426	**702**	2768426
	23	ypjC	1	complemer	2781660	2783033	**1374**	2781660
	23	ygaQ	0	forward	2784419	2786671	**2253**	2786671
SIR.23.24					**2784751**	**2784770**	**20**	
	24	ygaR	0	forward	2784770	2785456	**687**	2785456
	24	yqaC	0	forward	2785664	2786260	**597**	2786260
	24	gabT	0	forward	2790757	2792037	**1281**	2792037
	24	ygaU	1	complemer	2794359	2794808	**450**	2794359
	24	ygaM	0	forward	2798156	2798497	**342**	2798497
	24	luxS	0	complemer	2812240	2812755	**516**	2812240
	24	mltB	0	complemer	2822513	2823598	**1086**	2822513
	24	srlE	0	forward	2824414	2825373	**960**	2825373
	24	norW	0	forward	2831934	2833067	**1134**	2833067
	24	ascB	0	forward	2839012	2840436	**1425**	2840436
SIR.24.25					**2840436**	**2840499**	**64**	
	25	hycI	1	complemer	2840595	2841065	**471**	2840595
	25	hypB	0	forward	2849023	2849895	**873**	2849895
	25	ygbA	1	complemer	2854475	2854828	**354**	2854475
	25	pphB	0	forward	2857782	2858438	**657**	2858438
	25	ygcN	0	forward	2890679	2891950	**1272**	2891950
	25	mazF	0	complemer	2908778	2909113	**336**	2908778
	25	relA	0	complemer	2909439	2911673	**2235**	2909439
	25	rumA	0	complemer	2911721	2913022	**1302**	2911721
	25	ppdC	0	complemer	2960463	2960786	**324**	2960463
	25	ptsP	0	complemer	2964210	2966456	**2247**	2964210
SIR.25.26					**2967074**	**2967277**	**204**	
	26	mutH	0	forward	2967684	2968373	**690**	2968373
	26	ygeA	0	complemer	2977965	2978657	**693**	2977965
	26	kduI	0	complemer	2981310	2982146	**837**	2981310
	26	yqeF	0	complemer	2982433	2983614	**1182**	2982433
	26	ygeI	0	forward	2991660	2991878	**219**	2991878
	26	pbl	1	forward	2991961	2992463	**503**	2992463
	26	ygeM	0	complemer	2993336	2993767	**432**	2993336
	26	b2860	0	complemer	2994394	2995299	**906**	2994394
	26	ygeP	0	complemer	2995711	2996010	**300**	2995711
	26	yqeC	0	complemer	3012309	3013079	**771**	3012309
SIR.26.27					**3013091**	**3013107**	**17**	
	27	ygfJ	0	forward	3013182	3013760	**579**	3013760

	27 recJ	0 complemer	3034395	3036128	**1734**	3034395
	27 argO	1 complemer	3066195	3066830	**636**	3066195
	27 yggU	0 forward	3094405	3094695	**291**	3094695
	27 mutY	0 forward	3101035	3102087	**1053**	3102087
	27 glcC	0 forward	3126294	3127058	**765**	3127058
	27 yghQ	0 complemer	3129363	3130430	**1068**	3129363
	27 yghT	0 forward	3132153	3132845	**693**	3132845
	27 ygiZ	1 complemer	3169901	3170233	**333**	3169901
	27 yqiB	1 complemer	3174880	3175302	**423**	3174880
SIR.27.28			**3181420**	**3181760**	**341**	
	28 b3045	0 forward, Ins	3184532	3185437	**906**	3185437
	28 yqiI	0 forward	3188654	3189718	**1065**	3189718
	28 ygiP	0 complemer	3203346	3204278	**933**	3203346
	28 yhaK	0 forward	3252341	3253042	**702**	3253042
	28 tdcA	0 complemer	3264149	3265087	**939**	3264149
	28 yhaC	0 forward	3266437	3267624	**1188**	3267624
	28 ecfH	0 forward, yra	3294431	3295006	**576**	3295006
	28 yhbW	0 forward	3301470	3302477	**1008**	3302477
	28 nlpI	0 complemer	3306062	3306946	**885**	3306062
	28 dacB	0 forward	3326985	3328418	**1434**	3328418
SIR.28.29			**3331548**	**3331657**	**110**	
	29 ptsN	1 forward	3344600	3345091	**492**	3345091
	29 yhcC	0 complemer	3351143	3352072	**930**	3351143
	29 yhcE	0 forward	3363207	3364951	**1745**	3364951
	29 yhcO	1 complemer	3383879	3384151	**273**	3383879
	29 mreC	1 complemer	3396897	3398000	**1104**	3396897
	29 yhdN	0 complemer	3437163	3437531	**369**	3437163
	29 kefB	0 complemer	3476824	3478629	**1806**	3476824
	29 frlD	1 forward	3501189	3501974	**786**	3501974
	29 bioH	0 complemer	3542096	3542866	**771**	3542096
	29 gntX	0 forward	3542904	3543587	**684**	3543587
SIR.29.30			**3544296**	**3544352**	**57**	
	30 malQ	0 complemer	3546008	3548092	**2085**	3546008
	30 glgA	0 complemer	3564623	3566056	**1434**	3564623
	30 yrhA	0 forward	3581064	3581477	**414**	3581477
	30 ugpQ	0 complemer	3585393	3586136	**744**	3585393
	30 gadE	0 forward	3656389	3656916	**528**	3656916
	30 gadX	0 complemer	3663009	3663833	**825**	3663009
	30 yhjK	0 complemer	3681653	3683602	**1950**	3681653
	30 yhjR	0 complemer	3694020	3694208	**189**	3694020
	30 yiaG	0 forward	3717501	3717791	**291**	3717791
	30 xylR	0 forward	3733002	3734180	**1179**	3734180
SIR.30.31			**3735275**	**3735445**	**171**	
	31 sgbU	0 forward	3747255	3748115	**861**	3748115
	31 rhsA	0 forward	3760206	3764339	**4134**	3764339
	31 yibA	1 forward	3764360	3765202	**843**	3765202
	31 lctD	0 forward, lld	3777850	3779040	**1191**	3779040
	31 yibK	0 forward	3779238	3779711	**474**	3779711

	31	yibD	0	complemer	3787070	3788104	**1035**	3787070
	31	htrL	0	complemer	3790849	3791706	**858**	3790849
	31	coaD	1	forward	3807848	3808327	**480**	3808327
	31	rph	0	complemer	3813857	3814572	**716**	3813857
	31	dinD	0	forward	3815783	3816607	**825**	3816607
SIR.31.32					**3816682**	**3816822**	**141**	
	32	ligB	0	complemer	3817511	3819193	**1683**	3817511
	32	setC	0	forward	3834976	3836160	**1185**	3836160
	32	yicO	0	complemer	3840478	3841812	**1335**	3840478
	32	ivbL	0	complemer	3850913	3851011	**99**	3850913
	32	yidH	0	complemer	3853983	3854330	**348**	3853983
	32	glvC	0	complemer	3860010	3861626	**1617**	3860010
	32	yidS	0	forward, cb	3867400	3868464	**1065**	3868464
	32	yieH	0	forward	3894797	3895462	**666**	3895462
	32	yieJ	0	forward, cb	3896045	3896632	**588**	3896632
	32	atpE	0	complemer	3918973	3919212	**240**	3918973
SIR.32.1					**3920623**	**3921005**	**383**	

Table S2. List of genes containing TAG codons that overlap the open reading frame of a second gene.

TAG gene	overlaps the	of gene	coding change
abrB	tail	nei	silent*
aes	tail	hemH	silent
B1996	tail	yoeA	silent
B2860	head	ygeO	G6R
B3045	head	yqiG	G6R
blc	tail	sugE	silent
exoX	tail	ptrB	silent
frwD	tail	yijO	T280I
hypB	head	hypC	G4S
kgtP	tail	yfiM	silent
ligB	tail	yicG	silent
mdlA	head	mdlB	S3N
murF	head	mraY	V3I
nei	tail	abrB	silent*
nudG	tail	ynjH	T80I
paaX	head	paaY	D7N
panE	head	yajL	S13N
rbsR	tail	hsrA	S465F
sapB	head	sapC	S5N
sgbU	head	sgbE	E3K
tatD	tail	rfaH	silent
tfaE	tail	ymfS	P129L
tfaS	tail	yfdK	P138L
ugpQ	tail	yhhA	T143I
wbbJ	head	wbbK	V6I
wcaC	head	wcaD	S9N
ybjR	tail	ybjS	silent
ycbX	tail	ycbW	silent
ycjG	tail	mpaA	P255L
yeaL	tail	yeaM	T260I
yegV	tail	yegW	silent*
yegW	tail	yegV	silent*
ygcN	head	ygcO	A4T
ygeA	tail	lysR	A308V
yibD	tail	yibQ	A316V
yidH	head	yidG	S4N
yidS	tail	dgoT	silent
ytfB	tail	ytfA	S104F
zraR	tail	purD	silent

* two TAG-terminated genes overlap tails with terminal CTAG sequences
(converted to TTAA, all silent coding changes)

Table S3. List of oligonucleotide sequences for every TAG-to-TAA mutation

Region	Gene	Sequence of 90mer mutator oligonucleotides
1	rbsB	a*t*gacgtatttatgtcaccatcaggtcatacaacctgattaaaaTtactgcttaacaaccagtttcagatcaaccggatacttagcct*g*a
1	rbsR	g*a*aaacaaccgatggtaataatttgatcaaaagacagcgtaaatTtaagccgaaccgcgttccatcagaatcggagtaagttgtaatc*g*t
1	ilvL	c*g*gacctttcggtgcgggggtcttagttcgttaaggcttgatctTtaagcctttcctcgtccaagtgcagccccgcacggtgggataa*t*a
1	ilvA	c*a*caaatgacgttgtcgcgcgggtaggcctgataagcgaagcgctatcaggcatttttccTtaacccgccaaaaagaacctgaacgcc*g*g
1	wzzE	c*t*ttcacatcgattctcttcgaataagcggcgagcgcctttgcgctcaccgcagcagtgttgTtatttcgagcaacggcgggttaatg*c*g
1	aslB	g*g*cgtttacgccgcatccggcaatcaaccgcaggcggccgccgatttTtacttactcaccaccagcaaatgcgcatgcataatgtcgc*t*g
1	hemY	a*g*ctgcagctatgcgtcgcgacggtttgatgttaacgttgcagaataacccgccacagtaAttccttctcacccggaggcaagcacct*c*c
1	yigG	t*t*cagtatgctaatgcttgcacctactgattcatcaaaacaataAttatggagtatttatgagtaaggaatatatgaacgatggttca*t*t
1	recQ	a*a*cgtttacggagtcttcatactggcacttttttatgctgctgaTtactcttcgtcatcgccatcaacatgcgcacgaatcagcgcca*t*a
1	rmuC	g*t*cgatgccctagatttctacccggcttaactactcccaatgggTtagcgcgactgctgattatattcatcatcgcgttgataagctt*c*a
2	tatD	a*a*agagattaagcacagtgtgaagaataccgagttccgcaaactTtaaaacgcaatcccaaacagtgttttgacattagcatccgtgg*t*g
2	yihL	g*t*tgtagaatctttcttgcgttatttattgtcaccataaaacacTtacagattaatgatgaggtctacggcctcaaatcgccaatatt*c*a
2	yihW	g*g*ctatgtggcaaggtatgaagccaggttaacggtcacaaaaatTtacgcgtcttcctggggaagaatgatttcaatttggctgtctt*g*t
2	yiiD	c*c*aaggactaatataaatgacctgaatgagagaggcagcaccgcccctctccatcagTtactcttcttcgttcccgccctcttcatac*g*g
2	cpxP	g*g*ggaagacagggatggtgtctatggcaaggaaaacagggtttaTtactgggaacgtgagttgctactactcaatagcttcaacgatg*a*a
2	frwD	g*g*aatatcgccggcagtaccacagccaactgacagaaaaaccgaTtactccagaataagatgggtttgctgcggcgcagaaagcactt*t*g
2	zraR	t*a*tccactgggacgactgcttctgccggaaagatatcggctggcgcgctatcgaacgcgagcagaaTtaacgcgacagttttgccaat*a*g
2	yjaB	g*g*gaaaccgtatccgttgctgaatctggcgtatgtgggggcgtaAgcgatttttctcatcccggtgggctgaacttcaggtttagta*c*c
2	yjdK	t*t*agagaatagtgccatttttatgttcttcttatccttcctggTtacttgtaaaactgacggaaatagtcatcaactgtcataatgt*t*g
2	blc	t*t*tgatgtcagtaaatttatttgggtacagcagcctggtagttaAtgagtgctgagtttcagaccaataatccccaatacgattaacg*c*c
3	ulaF	a*g*accagttatctcccgaggaaggaaatttccgcagcgcgtgttTtacttctgcccgtaataagcgttaggaccgtgtttacgcatga*a*g
3	yjfY	g*g*gagtaacggcaatatgcacgccagcgcggatatttatcgctaAcaccaggtataaccaacgaaacattgccatagtttgctttgcc*c*c
3	priB	t*g*aaacggcagaacttgcgacgacggaaataacgtgccatatggTtagtctccagaatctatcaattcaatctgctcggcatgcaaaa*c*c
3	ytfB	c*t*gttcacccgccagccggatggcagttttattcgtgcgcggtaAatgaagggtcaaccaaaagccagctaaggattgtctctttgat*g*a
3	ytfG	a*c*gttagccgaaagcgtaagccatctttttaatgttaataactaAttaattaaagtggcatcctcccgcatcctctctgataatgacg*g*g
3	msrA	t*t*actgtggaattggcggaattggcgtctgtctgccgccggaagcataAcgttacgggtacaaatgtagtttgttgataaagtgctct*t*t
3	ytfN	g*c*cttgtttgtggcgtaaactgccgtagacaaatattcgcattgTtaaaactcgaactgatagagcaaatccagtgcctggtctacac*c*a
3	b4250	a*t*atctgcgcgccgcttcttttaatacctcaaccaggcactcgcTtaccggacgatcatcacgaagaatatcggcaagcggaatagct*t*c
3	idnR	a*c*tgttgacctcgggtatcagatttatcacggcaacacgctttaAgttattttctcttcattacacgtattctctttgccggacgttg*c*c
3	b4273	g*t*ataaaatatccgtattcatatcagcacaaggtggatttgcccTtatatttccagacatctgttatcacttaacccattacaagccc*g*c

4	yjgX	a*t*gatgggttctcatccaaatccgtgtgacagccttaactcctaAccgaataattacctggagcagtatccccgaaaaattgcctgtt*a*c
4	b4283	c*c*gaaagagataattgaaaggggtgttgcattgacagattgaatTtacagtagccttttaatatttcattctccatttcaatccgtt*g*t
4	fecE	c*a*ccccgagccggtatctggcaggccgatgtgcctaatgaggtaAattgcacaggccgtaagaaccaaaccacgactgaatgaaactg*g*a
4	nanC	t*t*atcggaaaacagttatcgcattggtgtgtcatttaaactgtaAtagacaggagacagtcacaatgaataaaacaataacggcgctt*g*c
4	fimB	t*a*agaataatgtagttttaacaccatccctggtatctcaactatctTtataaaacagcgtgacgctgtcgtcctctggctctatccc*a*g
4	yjiN	t*t*gctgtcgcagctcccggccttgttccccctcggcaatttttaAaaatccatgaaaaactaagtggctaaccgcgagagagatcaaa*t*a
4	yjiP	c*c*gtgataaaatagcatcgggatgaccaacggtagcggctggcgtttatcatgctctatatggcgTtacatcaccgccatggaatagc*g*c
4	mcrB	t*a*taaacaacagaaatggaccaacaaattattaggggactcataAtggaacagcccgtgatacctgtccgtaatatctattacatgct*t*a
4	yjjQ	t*c*tgaataaactctcaataccgatactactcatgacgcatttttTtatgagtgcgacatttctcttcttaattcggctgtgttccatt*c*t
4	yjjV	g*c*attgtgcgccaactgccggatgcggcgtgaacgccttatccggcTtacggcacgttaaacaacgtatacgtgttattaagcaacgc*t*t
5	lplA	a*g*gaaaaagagctacgggagttatcggcatggatggcgggggctgtaaggtaAttacccgcccatgcgggcaactttctcttcgattt*g*c
5	creC	g*c*aatggcagcaggatacgaagactatgtgggaagaatttgaagTtatgtgaagtgacggtgaagtcgaagcgaggccagcacgccac*c*t
5	kefF	a*t*gggtacaatcagcgctgccgaaccgagataaatcagcgcctgaatcagcgtatggTtatccatgatgggcctcctgccattccagc*a*g
5	yabI	g*a*cgcggcaagcgtcgcatcaggcaacccgcacaagactatctcTtaaaccccaaccactttacgcaaaatatcgatatacaccggca*t*c
5	thiQ	t*g*accaatgagttgttgagcggtaaggcgagtgcttcggcactattggggattacgggttaAtatgcactgtcggatgcggcgtgaac*g*c
5	murF	g*t*taaagccggaataatatttgaccaaatgttcggccagccaaaTtaacatgtcccattctcctgtaaagcgcgtactacctcttcca*t*g
5	mutT	c*a*tcaggcatcgtgtgcgaatgtcggatgcggcgaaaacgccttatctgacTtacagacgtttaagcttcgcaattaccggttcattg*g*c
5	pdhR	c*g*cttgatttacaacatcttctggataatttttaccagaaaaatcaTtaattctttcgttgctccagacgacgcagagaacgctcacg*g*c
5	sfsA	g*c*tgaaggcatggctcttaaaaaatcactgccggttacattgtaAtaaagtaagtaactggttaatttacattctggtcgcgtgcgca*a*a
5	btuF	t*a*tcctcgctgcacaacagctctgtaatgcgctttcacaggtagattaAcgcctgacctttcaggtgctcagtggggattgaacaat*g*c
6	cdaR	g*c*tgaggggaatgagccagcgccgcctcaaccttaacgcataacTtaccgctcttcatccagttgtaacgccacatacagcagcaacc*t*g
6	yaeI	c*g*tctgaactgccgcccggaagtaacgatgctggaactggtgtaAatcgtgaattttatcaccagaataaagcaggtaatttttgat*t*t
6	yafL	a*c*ctgcgtacaactgccggatgcggcgtaaacgccttatccgtcTtacaaaatcgtctcttccgtcaaaatcctgcgcgcgcccaaaa*a*a
6	yagY	c*g*ggttgcactttggcagggtgataagttcattcccgtgaaataAatagctgtgcagatggataacgacaatgcctttacgacggttc*t*c
6	eaeH	c*a*tgtgagtcacctctgactgagagtttactcacttagccgcgtgtccaTtattgctgggtaagaggcactggcgtataggtcacctc*g*g
6	b0299	a*c*ccgaggactttatgatcctacccacgtaatatggacacaggcTtaagcgaggttcttgttttcaaattgttccggactgaggccgc*c*a
6	yahL	a*c*tctcaggtggaataacaacttgcaggtggaacccttctgaaTtaaagagtgtagcggattttaacgtgcttatttacaggataat*a*t
6	yahM	g*a*tttacgaaggggtgacgcagcggtgagcagataatttactctTtacgtaatcaacctgatttgccgaatcacggccatcggcgatg*a*t
6	prpE	c*a*tcggcacacagtgccggatgcaccgcgcgacgcatccgacaataacTtactcttccatcgcctggcggatctgatccaacgacgca*g*g
6	cynR	t*g*gcgttggataaatgcgcggttgttggcggaaatgaatcacggtaAccgctgacacaggtctgcggctggaattgcctgatgcgcta*c*g
7	frmR	t*c*cgttgacgacactattgaactggttcgtgcctatcttaaataActgaatctattaccatattgaggaagagcgagagatgaaatca*c*g
7	b0361	g*t*agtgattggcgcaaactgcctgataggtaattggatttgcccTtatatttccagacatctgttatcacttaacccattacaagccc*g*c
7	yaiS	t*a*tttcacagtcaaacttgctgcattaaaaaaacataaaagctaAgaacgacgcgattatatgcgtcatgatcgcctgcgtgcagttg*c*a
7	tra5_1	g*g*cggcctcagtccggaacaatttgaaaacaagaacctcgcttaAgcctgtgtccatattacgtgggtaggatcaatcagccccccac*c*c
7	araJ	g*t*cgctataagcgccagcaagcggcggatactccggtgctggcgaaaccactggggtaAgttatagtctcggtgctcattacttattg*c*c

7	pgpA	g*t*cgtaggatttaaataagagtccaggcctgatgagacgtgacaagcgtcacatcaggcatcggtgcacaaTtacgacagaataccca*g*c
7	yajL	g*g*cgtgaaaaagcccatgaagtggcatcacaactggtgatggcggcagggatttataattattacgagtaAtgtcggatgcggcaaac*g*t
7	panE	a*g*tgaatatgagcgcatcggcactggtttgcctcgcccctggtaAtgaagagactgaagccgtcaccactatcgatctgctggttcgc*g*g
7	lon	a*g*cctgccagccctgtttttattagtgcattttgcgcgaggtcaTtattttgcagtcacaacctgcataccagacggttcattttgca*g*c
7	mdlA	g*g*cgaaccgtacgctaacaggcgcttgagagtcggccacagttggctaaaaTtacgcatcgacggcctcctcgcgattttccggagcg*t*c
8	ybaA	t*t*tcttacgctgggccgagcaacatttgtgaatttttgggctgacggcacgccgcacagcccgtcacgctacTtattcgtcgatgatt*g*a
8	ybaJ	t*t*tgtcaatgcgacgaaagagaatcctgcgagtttatcttgttaAaattattacaaccataggtagaagtatgtccgaaaaaccttta*a*c
8	priC	g*c*gctggaaaaaatcgaaaacaggttagcgcgtttaacccgctaAcaatggagagaatatgtcactggaaaatgcccctgacgatgtc*a*a
8	aes	g*a*ggctcttcgcgacggcgctcagttctttaccgctcagctttaAcgatacgcggcaacaagattagccatcatttcgatatgttccg*g*c
8	ybcF	a*c*gcgtcaagcgtcgcatcaggcacaaatgtctaatgcctacgaTtacagcgaaatacaggtccccgcttcgcccgccagcgtctctt*c*a
8	tra5_2	t*g*atttatgctggttgatcctacccacgtaatatggacacaggcTtaagcgaggttcttgttttcaaattgttccggactgaggccgc*c*a
8	ybcL	t*c*gatctttaagttagtcttacctcttgccagcatactctcaccTtactttatctcataaactggtgttatctcagcggttgcgattt*t*a
8	rzpD	t*t*gctcaaaataattgcatgagttgcccatcgatatggtcagctTtatctgcactgctcattaatatacttctgggttccttccagtt*g*t
8	ybcY	a*a*aaccaaggttcaaggtactgtagtaatgttttccgcttcaggaaaaaaataAcatccaaccgcagcatgttcttgcttaagacgtg*c*t
8	hokE	t*a*aatcatatcaggcgtaataccacaacccttaagttagcgcttatgggattactccccgcacaacggTtacttcttcggttcgtaag*c*g
9	ybdL	c*t*gaagcaaattaacggcctgcaagagttgtttgccgaccgtcaatattcccgcgcctggacggttaaatagTtaaagctggcgcagg*c*g
9	ybdM	t*g*aagcaaattaacggcctgcaagagttgtttgccgaccgtcaatattcccgcgcctggacggttaaataActaaagctggcgcaggc*g*t
9	rlpA	c*a*aaccgaagcccaattacagtcatttattactaccgcgcagtaAcgtaaagcaggcatctgatgtgtcaatttctgtaaatgacgtt*a*a
9	ybeQ	c*a*tccggaaccagagatgtcagtcaggaaattcgtgagactgaagatttactgtaAagataaaaacaagggccggggaatttcattcg*g*c
9	ubiF	c*g*tcttatccgacctacggttggcacgcatccggcaatgttgtaaggTtacaaccctaacgcatatttcagcgcctgacgtttcaaca*c*g
9	ybfH	c*a*ccgatgggtgcgcttgtctatggctccggaggtctgctgattgataatggttaAttacgcatagcaggatctggtcacccccggtt*g*c
9	ybfD	c*g*caaggactgacgtcaggtagtttctaaccatagctgccctttTtattcaaacgtgaaatcaataagctcatcagggacatcgcaaa*c*a
9	nei	t*g*ccattgcgcggtttatttcaacctatgcgccgaagcggtcggcTtagtgctggcagccagggcaccagtaaaacgggcgagatgac*a*g
9	abrB	t*g*acagggcctgccattgcgcggtttatttcaacctatgcgccgaagcggtcggcctaAtgctggcagccagggcaccagtaaaacgg*g*c
9	sucB	g*a*taatgccttatccggtctacagtgcaggtgaaacttaaactaTtacacgtccagcagcagacgcgtcggatcttccagcaactctt*t*g
10	bioD	t*c*gaagacgcgatctcgctcgcaatttaaccaaatacagaatggTtacaacaaggcaaggtttatgtactttccggttgccgcatttt*c*t
10	moaE	c*g*taaacgtatgtactgagcggtgaaattgccggacgcagcggtgccttatccggctaacaaaaaaTtaccagcgttttgccgcctgc*t*g
10	ybhM	g*c*gatgtgaagtttagttaagttctttagtatgtgcatttacggTtaatgaaaaaaacgcgtatgcctttgccagacaagcgttatag*c*t
10	ybhS	t*t*tatcggcctgacgtggctgaaaaccaaacgtcggctggattaAggagaagagcatgtttcatcgcttatggacgttaatccgcaaa*g*a
10	ybiH	c*a*tatcgacctgattttgcaaggattatcgcaaaggagtttgtaAtgatgaaaaaacctgtcgtgatcggattggcggtagtggtact*t*g
10	ybiR	t*c*tgaattaatcttcaaaacttaaagcaaaaggcggactcataatccgccttttttatttgccagaccTtagttggccgggagtataa*c*t
10	yliD	t*t*tcctgtgaggtgattacccttcaagcaatattcaaacgtaaTtatcctttaattttcggatccagcgcatcgcgtaaaccatcgc*c*c
10	yliE	g*a*ctgactgtaagtacgaacttattgattctggacatacgtaaaTtactcttttactaattttccacttttatcccaggcggagaatg*g*c
10	ybjK	t*c*ggttcaaggttgatgggtttttgttatctaaaacttatctaTtaccctgcaaccctctcaaccatcctcaaaatctcctcgcgcg*a*t
10	rimK	c*g*caaaaagcgcaggcaaaaccatgatcagtaatgtgattgcgaTtaaccacccgttttcaggcaatattctgtcgtagcgtggcgtt*c*g

11	ybjR	a*t*cgagaaaactgccgcctggctacgcgaccacggaaaactgccgcgTtaatcctgcccgtatttctccagcaatgcttcggcaatcg*c*c
11	cspD	g*c*cagtgttattgtgcccgtcgaagtagaagcggcagtcgcataActcttctgtctcattgtgtacatcctaaaggcaaaatgccagc*c*c
11	lolA	a*a*agtattatccgaaaaatcgagcgacagattgctcactcaggtgcctTtacttacgttgatcatctaccgtgacgccttgcggcggg*g*t
11	ycaI	t*t*atgcattcaaaaaaccagcatttgttgaaatagccgcatattTtacccgttatccactggcacgccaaaccactgatgataccaac*g*a
11	lpxK	a*g*cgacattcatgactccatcaatcgaacgctgccgcggcgtaaTtagttgccagaagccagcaaggttagttgcgtaagcagtttcg*c*t
11	ycbK	a*a*taatacgatagttcatactgcccctgtttcgttaagcaattaTtaccagtgccgtgctggcccggtatcaatatgcacaaagttac*t*a
11	ycbX	t*g*ctgtagctgtgtaccgaagactgcacttaagttggcgcgttaAactgcctgttcgaggctgaagctgtcaacattaacctgcggtt*t*c
11	tra5_3	g*g*cggcctcagtccggaacaatttgaaaacaagaacctcgcttaAgcctgtgtccatattacgtgggtaggatcaaaatgccgcttta*t*g
11	yceO	g*g*attgttggcaattatggatctgatctatgtgttttatttttaAttttccgcactgacataaactgacaaatatttaatctccctttc*c*a
11	yceF	g*c*taatcgcgctgtgtcagatgttacgtcgggaagggaaaaacccgctgatgggataAtttttgcgccggatgtggttacgtcgcat*c*c
12	plsX	a*t*ataccgtcacttgcaaactgcgagttcgctggcagcgtcctgTtaccgcagagttccgcttttgccaccgtccagcagctcaaaac*c*a
12	fabH	c*c*agggaacacaaatgcaaattgcgtcatgttttaatccttatcTtagaaacgaaccagcgcggagccccaggtgaatccaccgccaa*a*g
12	pabC	t*a*ccaccagcaataacaagattatcaataacacttttttcatgaTtaattcgggcgctcacaaagtggggctaaatattcatataacg*t*t
12	b1146	c*g*attagcctcatcgttctgtggttaaaaattgaaagtgttctgTtaatctttcggatagatatccggtcttaagtcagatttcgtaa*t*t
12	ymfM	c*c*atcccatacatagccggttaatttagtgctcatgaccgacctTtatttccctgaatttacgacgaaactgttcgagcgggctgaag*c*a
12	ymfR	t*c*accggtgcgtcacttttcgttgaaataatcccgaaaagaacaTtatttacctccgccgacagacgactgtgtacggtcgagatat*c*g
12	tfaE	c*a*gataccagcaccgcgccggatgtcatctggcctgaactgccggaggcgtaAgccattcaatatctggcgcactggatgtatcaacc*a*g
12	ycgX	c*t*tttatatattgaagatgtagaaaataataaaccattaatctaActaaagttggatacttaagaaatgcttcataattcagtaaggc*a*t
12	ymgC	c*a*tggtgatacagctgatgtttattctaaaaccttactcaagttTtaagagagcacggattccctgtcattaataatgatatccgtta*g*t
12	ycgY	t*t*ctccggatggcgttgtcggcagataacaccatagctatggtttttTtacacagacggttttcaagtcctggcgggaagtattgc*g*c
13	hemA	c*a*taggcgtaaatgcaccctgtaaaaaaagaaaatgatgtactgTtactccagcccgaggctgtcgcgcagaatattcaggcgttcgt*t*a
13	b1228	a*a*aagtccctgaactctcaagcgaatccgcaatcaaatattctgcccatgcggggaaggatgagaagcttcgaTtaaggttcggctcg*a*g
13	kch	g*a*actagaaacgaaagctgacagtaaagaatcggcgcaaaaataAagtttaaagacggaagagtaaacttattgggcaccgcaatggt*g*c
13	sapB	c*t*gatattttgggtgccatggctaaccctctgaaacataaggaatggtatgccttacgataAcgtatacagcgaaaagcgcccgccgg*g*c
13	pspF	a*c*ttatcatcagttccgcgcgttgttgaaaaagcaccagatttaAcgaagttgagcaatacaacttgccggatgcggcgtgaatgctt*t*a
13	ycjU	t*t*tgaatatgttgtaacgaaagctgagccatgttgattcctttgTtatacgttttgccagaaggccgataaccgcggccaggtgagtg*a*t
13	ycjV	a*t*tttctcctgcatgataatcatttaacgcaccagcacggaccaTtaactcgtgcccccaaccgtggtgtagagcataaattccgcac*c*g
13	ycjG	g*t*ttccgcccatctcctccgacgaagccagcgaaaaatatctctttgcgatggcaaacttgctgcgctggcatcTtaaagatgcaatt*c*g
13	ycjZ	a*a*tacgttaagagaagaattaaattaatgtatcagcaagtagtgTtaaacagcacctttatacttcagcctatctataaataatgaga*a*t
13	intR	a*t*gaacgacgaacagatagccatgttgaatgcgcggttatcgtaAttgcaaagtttgcccccaatttgccccatttagtaccagagaa*c*t
14	sieB	a*c*cgaaactggtgcgcgccgatggttggagaagcacacagtaagTtaacgattaaaacgtctactcctgctgttccagaataacttca*t*a
14	b1354	g*g*ctttacgcagcaacgactgggatctctggcaggcagaagtcgTtatacgtagcaacaagatgtgagattcgcaaaatgcaaaaaat*c*g
14	ydaY	t*t*ccccctttctcccccgaactccgataaccaccgctgggcagTtatggctttttatcgtatcctgcttctgctgctctttaccct*g*g
14	b1367	g*a*tttaattaaagacaccatcacatctctgtaacagagtgtacgtaattaacaaTtacacacactgctcctgaaaatactggtcatcc*a*g
14	tfaR	t*a*aatggaaagatacatctaattatagaatttatatgttttaccTtacggcagtgctggccattcaatatcctgtgcagttgacgtat*c*a

14	paaG	t*g*ccgctcccaatcactgccacagtttgcacatttatcatcatgTtatttccccgtgaactgcggtgagcgtttagccaggaacgcac*t*g
14	paaX	c*g*gatggacaaaactctcttctggcacaaccggagtcagaccgtTtatctgataaattggcataacgcctcctgttcaatattcaagc*c*g
14	b1402	c*g*ggcttgtaatgggttaagtgataacagatgtctggaaatataAgggcaaatccagggttcttgtcccatgatgttcaaaactgtca*a*t
14	ydcL	g*g*atggggttttacggcgcaccggataaaaagcgccatctaacgatggcgcgtctggaaTtacttttgttaacgtcaaacatggtgg*c*a
14	rhsE	c*a*acaaataaacaaagggtaattaaaaaaatgacaagtttgccttaataatttgaagcatctgaactcccttgctttgcttttattg*g*a
15	yddJ	g*a*ccttctggacagatcgttttcctccagttcatggccttttaAtgaataaatcatcactaattctataagtgtctggtcttgttga*t*a
15	ddpF	g*c*aatcattagcaataagggaggacggacgtgagctccggtgctggcgggcgctgtaAagatgctaacaacgcatcaggcattcacgc*g*c
15	yddV	a*g*acgaggtagaaaccgtgttgaactctggaaagccagtctttaAatgcgccaggatgcagaggtaatcatgaagctaaccgatgcgg*a*t
15	ydeV	g*a*ataactatactcaaaaactaacagccacggtcatcatgatgtggctgtcaatgaaaTtataacccaggcgctttccataacgacgt*c*g
15	ydeA	g*c*cgcgttaatttggtcaatcattatatttcgccgctggccagtgacactcgaagaacagacgcaataAttgaaaggcccattcgggc*c*t
15	marA	g*g*cgaatcgcgctttttacatccattaaatcattacaacagctaAttgaaaacgtgacaacgtcactgaggcaatcatgaaaccactt*t*c
15	marB	a*t*gggtactggcagtgataagtcggatgcgctcggcgtgccctattataatcaacacgctatgtaAtttgttctggccccgacatctc*g*g
15	tfaQ	t*a*aatggaaagatacatctaattatagaatttatatgttttaccTtacggcagtgctggccattcaatatcctgtgcagttgacgtat*c*a
15	ydfP	a*a*attttacattccgcgatgtaaatgttccgatttaatattaccTtacatttgatgctttttatctcttaaagattcatagatctgtt*g*a
15	rzpQ	g*c*agcggcattaagcattgccagaatgattacatacccgcgttaAtagctaatcaacaaagctaaggttagtaattaaggagttctcc*a*c
16	b1578	c*g*ggcttgtaatgggttaagtgataacagatgtctggaaatataAgggcaaatccacaggtgaactacgctcctctgtttgttacgca*a*t
16	ynfA	g*c*tgacacttcctcttatcttattgataaaatggatttatgttcTtacgtgcgcccccagcccgcaacaatgatcaacatgccgcaaa*g*c
16	speG	c*a*caaaacaccgggtcagactctcctgaagccgaccgcacaataAcattaataataatcgatcgtatttttgatggtgtaaacccgtt*c*g
16	dmsD	a*t*cacaactgttaattcctgtcgcggttaaaccgctgtttcgataAgatcaccgtaatattgccggatggtgacgattacgcgcaatc*c*g
17	ydhM	g*c*gctggcccatgtaaaaaacattattgcgacgcctgccgtttaAcaggcatttttatcaccagacgaccgggagcctttatgtcat*c*t
17	lhr	a*a*cgccgatgtttacgctgctgcgtgaggcgggattttcaagttcgccacaagggctggattggggataAagaaaggactgacggatg*c*c
17	ynhG	a*a*tggcgcacattgtgcgccatttttttgtctgccgtttaccgTtactgcgtcacgcgtaacatattcccttgctctggttcaccat*t*c
17	sufA	t*t*tgacatcgtcagttgcttcagtattacgagacatagtaccgcTtataccccaaagctttcgccacagccacattcattctgggctt*t*a
17	ydiA	a*c*caagatcctcgatatcatgggccttagtcgccgaatgtactaAagaactagtgcattagcttatttttttgttatcatgctaacca*c*c
17	ydiE	c*a*tgacgggcaagaatatctgctccgtaaaactcaggctggcaagctgctgttgaccaagtaAcctttaactcgagcagctgacttcc*a*g
17	nlpC	c*t*actgtattcaggagtaaatggacggcgatgtatcatcgccgtccatttttcgacataTtagattcgtctcgcctgccagaaatttt*t*c
17	btuC	t*c*agtaaaaatcttagcttttaatcatagaccgtctttttgcggTtaacgtcctgcttttaacaataaccagataaacaccggcgcac*c*t
17	arpB_1	c*a*tagagaaaaaaactatacacccaacagtaaatctcaatggtaAtgcatgttttttgcaatctccttctgacgctatttttttgtcgc*c*a
17	ydjX	g*c*ttcgcgccgcagcccacttactcacccaaaaaatgaaggataAaacgatgttgcaacattattcagtgtcatggaaaaaaggactg*g*c
18	nudG	t*g*ctgacattccattattagaggcgtttatggctttacgcgccgccagaccagcggattaAtgctaagggttttgtcatcacgctggc*a*t
18	b1788	g*a*aaaccgcggattatttgcgcattgcgcaaccagaactggtcaatttccgccgcactgcgtaaTtagccggtccgttcgataagaag*a*a
18	yeaL	g*g*cgtaccggttgggccgcttattgccgcgggtctggtttcgctgattgtggggaaacagtaAttaatctcgcaatatatcggcctgg*c*g
18	yeaX	a*a*agtatgttgatctgttgttcgcgtgcgaagggtaaacgcctggtgttggatttgtaAtttgccgaagccggatgtggcgctgagcg*c*g
18	yobB	c*g*ttcatcacagggttggcaaggcgatatcattccattacgctaAgctttttcggcctggagcatgccatgttgcgcattatcgatac*a*g
18	exoX	c*c*ggagctgcgtttaacactgaaacattatctggaaaatacttaAtccgcaggcgtagcgggtaatgttccctgcgccagcgcgacca*g*a

18	yecN	a*a*tctttggtatatgccctgggagttggttttctccctgcgttaAcgcacaatacgccactttctttttcccggatttttacgttatg*t*c
18	otsA	a*g*gtcgatgtgctgttagttccacttacgggagattaaccgctcTtacgcaagctttggaaaggtagcaactttatcgcgctgctggc*t*t
18	dcyD	a*t*caataaccagttgtatactttcttgcatgagcgtttccggagTtaaacgtggggatgataggcgaacagcgcaggcgcaccgccgg*t*a
18	yedM	c*g*taaagttgtctagaaactctgcaatattagtcgatataaactTtaatttatgccgaatgggtgggttaccagttccctgatttttt*c*t
19	fliE	c*c*cacgcggtgagatcggatgtaattaacaacaggttacaaaacTtacacctgcatgctcatcacttcctgatacgccgccaccagct*t*a
19	fliN	g*a*tatcattactccgtctgagcgaatgcgccgcctgagccgttaAtgatgaataaccacgctactgtgcaatcttccgcgccggtttc*t*g
19	fliP	t*g*gcaattgctggtcggttcgctggcgcagagcttttacagctaAagaggcaaaatgacacctgaatcggtcatgatgatggggactg*a*a
19	fliQ	a*c*cgtggatgctcaatctgttgctggattacgtccgcaccttgttcactaacctgccgtatatcatcgggtaAccgtactatgttgca*g*g
19	yedS_1	g*c*taatggtgttgcaacctaccgtaacaatggcttctttggttaAgtcgatggcttgaactttgctttgcaatatcagggtaataatg*a*a
19	cbl	g*g*agccaaattcctgaaaagcccgcttttatagcgggatttttgTtatatctgataatcaatttcctcttcactgctttccatcacct*g*c
19	b1996	t*t*ggcaggccgtaaacaaatccgcattcgtggctggatttgcccTtatatttccagacatctgttatcacttaacccattacaagccc*g*c
19	hisL	g*a*cacgcgttcaatttaaacaccaccatcatcaccatcatcctgactaAtctttcaggcgatgtgtgctggaagacattcagatcttc*c*a
19	wbbJ	g*g*taaatggaccgccagtggtaaaattgaccgcagaaacaacgaTtatgctttttcccataatttggtctcatgattgtatttcttta*t*g
19	wcaM	t*t*gttcctaagtatgactccatttttccaggaatggtcgcaaatTtactccctcagttccggcaatctaaagttaatcttctccacat*t*a
20	wcaL	g*g*aacgtgcgtcgggagagtttttaaatggcatgcgttctcctTtataaagcctgcagcaagctggcgagttctcgattgatcacct*g*c
20	wcaC	t*t*ttgacgttaaccagcagataaatcagcggcagcaacaggtagTtacagattctgatagaagttgacatactcctccagcatctgtt*g*t
20	asmA	c*t*ttacgcgtcttatcaggcctacaggactggtcacctggtaacTtacatcttctccagcaacttcttcacatctttgccattacggg*a*a
20	baeR	c*c*gcgccgtttatggcgtcggttaccgctgggaagccgacgcctgccgcatcgtttaAttttagcgacattattttgttagccggaga*t*g
20	gatR_1	t*g*ccggatgcggtgtaaacaccttatccggcatacagaacaataTtacaccgtaattaattttacccgcgtcttctgcaatgccagtt*t*a
20	tra5_4	g*g*cggcctcagtccggaacaatttgaaaacaagaacctcgcttaAgcctgtgtccatattacgtgggtaggatcagtcatcccttcgt*g*g
20	yegV	g*c*gccaacgcgcgaggaactactcctcgcacacaaaaacgtataAatcgctgcgacagtggctaatgctgtactcaataggccgttgc*t*g
20	yegW	g*c*gtcgtgggttgtcgggcatcggggcggtgattgtgcgccaacgcgcgaggaaTtactcctcgcacacaaaaacgtatagatcgctg*c*g
20	yehQ	a*a*cggctggcgctatcaggaacagagtgctattggtcaggcataAttgctggcgaatgacctcaccgcggtgggtttttatcggctgg*c*a
20	yohC	c*g*tcaaaggagaatcgtgaggaaatgctgcattgctgacatttacgccagcaatgcaacgtcaaagaactttTtagaacaaacgataa*c*c
21	yohF	c*g*accaaaatattacccaatgaaatatccacgcacctcactgcgTtattctgggttgaactgtggattcgccaacataaagccgccat*c*c
21	sanA	a*c*acccgaacagttacttgaattacaaaagaaacaaggaaagtaAttatggatgtacagcagttttttgtcgttgccgttttttcct*t*a
21	yejA	t*g*gtatgacgtcaataaagcggccaaactgccgtccgccagcaaacagggagagtaAatgggcgcttacctgattcgccgtctgttgc*t*g
21	yejE	g*g*tgaagccgtccgcgacgcatttgatcctaataaggcggtgtaAcatgacgcaaactctgttagcgattgaaaatttgtcggtgggt*t*t
21	b2191	g*a*gccgatttttctcccgtaaatgccttgaatcagcctatttaAaccgtttcttcgccatttaaggcgttatccccagttttagtg*a*g
21	rcsD	a*a*atacatcagcgacattgacagttatgtcaagagcttgctgtaAcaaggtagcctattacatgaacaatatgaacgtaattattgcc*g*a
21	rcsC	g*g*catttgcactgaatgccggatgcggcgtaaacgccttatccgtcTtacgaatcccgcgatttcctgaccctctcggcatataacgt*c*a
21	yfaT	g*g*aagcgaatcatcagcattagcaagccctgttcccacaagagacagggacaagcacgctgTtatcgcgccagaaaatttaaacgata*a*a
21	menF	a*a*aaatagattattgatatgaatcggtaatgatgcgactcattaTtattccatttgtaataaagtacgcagccctgccgctttgttgt*c*g
21	yfcO	t*a*cgggtaatggcaggggcgaatgcccctgcctggtgtatctaagggTtacaactcattcagtatgtactggaaccggcaagtagaat*c*g
22	yfcU	t*a*gctaacggtactgttgtcgccccacggcatactgagcgaaatTtacatgcccttatccgcccggttgtcatactcgtagcggtagc*c*a

22	tfaS	c*c*agcaccgcgccggatgtcatctggcctgaactgccggaggagtaAgccattcaatatctggcgcactggaagtatcgaccagctcc*a*g
22	ypdI	c*c*ttttccgacaaaggcgaattatcttattcatcctgtgcgataAcgtctttcacatggcaggtgaacaatcatgtctgatgctgatg*t*g
22	yfdY	t*g*atgctgatgtgaacacctgcccaaatgagccatcctgacccaTtaccgcctgccataccccaccataaagcgaatcatcaagccac*a*a
22	yfeO	a*t*ggtcaatcgtccgaagcaacagccacccccacgataacgtttaAcaaaaaatgctctcctttatgatttaagagttatggcgatgat*a*c
22	mntH	t*t*acgccgcatccgccaatagtgccagatgcgacgctcattcaaTtacaatcccagcgccgtccccaccaacaaccagatattcagcg*c*c
22	xapR	t*t*cacttagcagactactatttccggcaattcctgtctcctcacTtactgtgtcaatgcagccaacagcttaaccatcgcgggcgtca*c*c
22	yfeR	g*c*gataatctactccatcaaatttaacttctgattgattactgccggggctcccggcagtatgaattcactTtatatctgatacaacg*g*a
22	yffB	g*g*tttcagtgattccagttatcagcaatttttccatgaggtgtaAtctatgtcgtgcccggttattgagctgacacaacagcttattc*g*c
22	hda	a*g*ttcggataaggcgttcgcgccgcatccgacaataaacaccttatTtacaacttcagaatttctttcacaaacggaatggtcagctt*a*c
23	yfgG	g*c*tcagcaatccacactctccgtcgaatcaccggtacaacgttaAcggttaccttctccactttcacagaacataacggcacttcgct*g*t
23	pbpC	t*t*taaaataaaatgagttttagcaacaatcagggacaaaacagaTtattgcatgacaaatttcactgtcgcgatttgccccacgtcat*c*c
23	yphA	g*t*cagtatcgctggcgcgttcttgctattggcaattaccgggccgggggcaatttctctcgatcggcgttaAgaaagatgccggatgc*g*g
23	yfhB	a*a*ctccggttataaaagcgcggcgcattatacacggacgctatgctttactTtattccagttgctggagttcaccgcgcggggttact*c*g
23	kgtP	g*a*tgtcgccggtgcaagcaccggctataccgtctggcaactgacccgtcaTtaaagacgcatccccttcccttgcgatgtagcatca*a*a
23	yfiA	c*g*caacatcggtgaaagacgccaacttcgtcgaagaagttgaagaagagtaAtcctttatattgagtgtatcgccaacgcgccttcgg*g*c
23	yfjQ	c*c*gtgcgctgtgggtgatggcagaaacactgctcacgcaactgcagtaAacgtttcatgttgccacgttgttaatatcggacaccacc*t*g
23	yfjR	g*c*cttaccccattttcgctttatccagcattcactgcttacgtttgttccggccgacagcccaccggaataAcgcaggcgttgttatc*a*a
23	ypjC	a*a*accaacatgaaaatcacttttttcttatattatcgaaaaggttTtattcatttcttttagcgcattcaaaaaactgatcggcattat*t*t
23	ygaQ	c*g*tagcattacagcgccagacaggataggagaaagaaaaacataAgtttatcaacgagatacaatgattgaaagaggtgataacagcg*g*t
24	ygaR	g*c*gaagatcagtaaagatatagaaggtggtatccctggctattaAcaaggtcaggttttgattccattcattaaagatccagtaacaa*a*a
24	yqaC	a*t*taaaaattatgatgggtccacgcgtgtcggcggtgaggcgtaActtaataaaggttgctctacctatcagcagctctacaatgaat*t*c
24	gabT	t*c*accattgaagacgctcagatccgtcagggtctggagatcatcagccagtgttttgatgaggcgaagcagtaAcgccgctcctatgc*c*g
24	ygaU	t*g*acgccaattcccattatccagcaggcgatggctggcaattaaTtactcttccggaatacgcaacacttgccccggataaatttttat*c*c
24	ygaM	g*t*aggtatttttatcggcgcactgttaagcatgcgcaaatcgtaAtgcaaaaatgataataaatacgcgtctttgaccccgaagcctg*t*c
24	luxS	t*t*tgaactggctttttcaattaattgtgaagatagtttactgaTtagatgtgcagttcctgcaacttctctttcggcagtgccagtt*c*t
24	mltB	a*a*ttttacgaggaggattcagaaaaaagctgattagccagagggaagctcacgccccctcttgtaaatagTtactgtactcgcgcca*g*c
24	srlE	a*c*tgtactgatcgcctggtttgtctccggttttatctatcaataAaggctgaaacatgaccgttatttatcagaccaccatcacccgt*a*t
24	norW	a*t*cggatgaaagaggcatttggattgttgaaaacattgccgatgtaAgtgggctactgtgcctaaaatgtcggatgcgacgctggcgc*g*t
24	ascB	a*t*cattctggtggtataaaaaagtgattgccagtaatggggaagatttagagtaAgtaacagtgccggatgcggcgtgaacgccttat*c*c
25	hycI	t*c*accccatcaagaacatccctgtcctgattccttaatgaaaaaTtactcttcttccaccgctaactgcgcgaagccgccatttcctt*c*c
25	hypB	a*a*gggatggaccagtggctgaactggctggagacacagcgatgtgcataAgcgttcccggccagatccgcaccattgacggcaaccag*g*c
25	ygbA	t*g*atttaactgcaaattgccggacagatctgcctgtccggcataTtattcatgaggtttttcggacgatatttttccggcagttctg*g*c
25	pphB	g*c*taaccagatttatattgataccggatcgccgaacagcgggcggctgtcattttataaaataaagtaAtctcatgcttcttctgtga*a*g
25	ygcN	t*g*gcatctggcgggcgatgttatcaggagtctgcgatgtctgtaAcccgtaatctctggcgcgttgctgatgcgccgcacattgttcc*g*g
25	mazF	g*t*caggtggaaacctgtgaccagaatagaagtgagttagtaacaTtacccaatcagtacgttaattttggctttaatgagttgtaatt*c*c

25	relA	a*c*agtatatatcaatctacattgtagatacgagcaaatttcggcTtaactcccgtgcaaccgacgcgcgtcgataacatccggcacct*g*g
25	rumA	c*c*tctcctttagggaccagacctgccgaaatcggcaaatcgcaaTtatttaacgcgcgagaaaagtaccatcgattccagatgtcccg*t*g
25	ppdC	t*c*cagacgattggaatggtagacccttaacatagcggctcctgaTtactgacgattcgggcaatgcaggcgggtcatctcgccttctc*t*g
25	ptsP	g*g*ctggttgccggataccgttaaaagatatgtatatgatccgcgTtataaccctccgcgaatcagcccgcccatgccgcgacgctcca*t*a
26	mutH	a*a*tttcaccagtgcactactggcccgtcattttctgatccagtaAccatcgctttgacctgccgctttccgggcatataattaccgct*t*c
26	ygeA	a*g*gcgtttagtgggcatttacaagcggggttaccgaaactggtcacttctcttgacgctattttgtcgtcagTtacgacagcataaaa*g*c
26	kduI	t*g*cacttaaaatcataacaataccttattcgtttatgcccacaaTtagcgcaaatctttaacggccacatggtccatatcatcaaaga*c*c
26	yqeF	g*a*ttaatgacggcagatatttatcggtaaggaggaaatgcagagTtattcgtcacgttcaatggtcaatgccacaccctggcccccgc*c*g
26	ygeI	t*c*agaagataatcctatctcaccttactataaaattgagtggtaAacgattgaccaaataaatcataacgagatacttttataatagt*a*t
26	pbl	t*a*tgccccaaaatatatattgtatataccaggcttaatgaactaAacaatcgtaaggcttcagcaaaatgagcctaaagcctcttttt*t*t
26	ygeM	t*t*acaattagattatattcaaatcattaaacttgagccagggagTtataagtcctcagggcttgccaaaacagaattattcatatcca*a*a
26	b2860	a*c*agatcaatggaagtaaattccaattgtttattggatttgcccTtatatttccagacatctgttatcacttaacccattacaagccc*g*c
26	ygeP	c*t*tatttaagtgatattggttgtctggagattcaggggccagtTtaaacttgctcttttcttctggtgtttatctttgaaagatttg*t*a
26	yqeC	t*g*tgaaaatattcataccaacctcaatggatatcctttcagtaacccggaatacccgggccgTtaattcacaaatctgcgcgcgattg*c*a
27	ygfJ	c*c*ggacgactttattacagcgaaggaaaggtatactgaaatttaAaaaacgtagttaaacgattgcgttcaaatatttaatccttccg*g*c
27	recJ	g*g*gattgtacccaatccacgctctttttatagagaagatgacgTtaaattggccagatattgtcgatgataatttgcaggctgcggt*t*g
27	argO	c*t*ctggaggcaagcttagcgcctctgttttatttttccatcagatagcgcTtaactgaacaaggcttgtgcatgagcaataccgtctc*t*c
27	yggU	a*a*tccgcaacaaatcccgccagaaatcgcggcgttaattaattaAgtatcctatgcaaaaagttgtcctcgcaaccggcaatgtcggt*a*a
27	mutY	g*t*ggagcgtttgttacagcagttacgcactggcgcgccggtttaAcgcgtgagtcgataaagaggatgatttatgagcagaacgattt*t*t
27	glcC	g*c*caccatttgattcgctcggcggtgccgctggagatgaacctgagttaActggtattaaatctgcttttcatacaatcggtaacgct*t*g
27	yghQ	a*c*tgagtcagccgagaagaatttccccgcttattcgcaccttccTtaaatcaggtcatacgcttcgagatacttaacgccaaacacca*g*c
27	yghT	t*g*gttgatgcagaaaaagcgattacggattttatgaccgcgcgtggttatcactaAtcaaaaatggaaatgcccgatcgccaggaccg*g*g
27	ygiZ	t*t*ctctgtctatgagagccgttaaaacgactctcatagattttaTtaatagcaaaatataaaccgtccccaaaaaagccaccaaccac*a*a
27	yqiB	a*g*ggttaacaggctttccaaatggtgtccttaggtttcacgacgTtaataaaccggaatcgccatcgctccatgtgctaaacagtatc*g*c
28	b3045	c*g*ggcttgtaatgggttaagtgataacagatgtctggaaatataAgggcaaatccagttatcatcattggttgtgcgtcagcttatgc*c*g
28	yqiI	t*t*gttgagggaccgtttagtaaagacgtaatcctgaaaatcacctataactaAtatctaatacaaacactaaaacgggccatcaggcc*c*g
28	ygiP	a*t*tcaaaagaaagcgcgaaaaggatcccgagtgatttctggtggatTtacatgacctgatagccttcatcgggcttgcccagccgttg*c*t
28	yhaK	g*c*cgattccccactgcgcgctttgctgatagatttgcctgtctaAttgtttttacagggagatgatgatgagtaaaaaattggccaaa*a*a
28	tdcA	a*a*gacagacaggtggattatttatgtttaaataaaaacaaataaTtaaccaacttctattaattgccttcgtctacacccattataag*a*t
28	yhaC	a*g*agaggactatacgtatgattacgacctttaaatgcaatataAacttaataacttcttacaaagtctattattaataccgtagata*t*t
28	ecfH	g*c*cgggtgagcggcgtgaagcgggtaactacggcgtttacgtttattaaataAcagccccttgtaatgcctgatgcgacgcttgccgc*g*t
28	yhbW	t*c*gtttgagctggcgatggatgttaaggaagagttgttgggataAtgtgtcttaacgcgggaagccttatccgagctggcaacgctgt*c*c
28	nlpI	c*g*ttaaggtgatggcaatcaaaaaagattacgggctgatgtgtacgtcagTtattgctggtccgattctgccaggtcatcttggtcct*g*g
28	dacB	c*g*ccgtattccgttagtgcgttttgaaagccgtttgtataaagatatttatcagaacaattaAtcaaaaagaaaccccggcacatggc*c*g
29	ptsN	c*c*agagcgatgaagagctgtatcaaatcattacggataccgaaggtactccggatgaagcgtaAttattcggtaatgtctcttttaga*c*g

29	yhcC	t*a*cgttataaattgtgtgagaaatgcggaattcccctcaacgacTtactccgttggaggtagccacggacgtcccagcgccgatccct*g*t
29	yhcE	g*a*gccgatttttctcccgtaaatgccttgaatcagcctatttaAaccgtttcttcgccatttaaggcgttatccccagttttagtg*a*g
29	yhcO	g*g*ctgaactgtacaaataaaccctcatcgtcttgtccgacgatattgcccccggttcgggggctttttttgcgTtaatgacgaacat*t*a
29	mreC	a*g*caacagcgcaatgaggaaagagagccagattacccagcgtccctggctacgatagctcgccaTtattgccctcccggcgcacgcgc*a*g
29	yhdN	c*c*gccatttttgccagctcaccaatgcgatacatacatactccaTtagttatcgttgattttgtccaacaacttgtcagcataatcgc*c*g
29	kefB	t*t*tgcgcccgcaataaaacgttttcggattgccatctttaccctTtactcaaattcatcccagccgtccagctggcgtcgttcttgtt*g*c
29	frlD	a*t*gacattaccgcaggcgatagcgcagggaacggcgtgcgcggcgaaaaccattcagtaccacggtgcctggtaAgtataacgttggc*g*t
29	bioH	c*t*ttaccgggctgtcagaagaagagttggctgcggcgagtcagtaaagttctgtctcgccatttcaaaagccacTtacaccctctgct*t*c
29	gntX	g*g*tgcggcggctgtccaggtctggtgcctttgtcgaaccttgtaAagcctcgatgatgggcgtattataaccaactaaaatagtcaac*t*a
30	malQ	c*c*ggataagacgcgtcaagcatcgcatccggcatcaaccgcactTtacttcttcttcgctgcagctctgcgccgtctgtccaaatcct*t*c
30	glgA	g*a*tgaatatgtaaacggagcattcatataggcgtttcctgaaaaTtatttcaagcgatagtaaagctcacggtacgacttcgccgcga*c*c
30	yrhA	t*t*tacttatgacattaaaggtgatgctgccaacttactgatttaAtgtatgatggtgtttttgaggtgctccagtggcttctgtttct*a*t
30	ugpQ	c*t*gccggagaggagaaacggcgacatgctgaatcagcccagcacgccgcagcctgacattccgttgaaaaTtattgggccgtaaagtt*c*g
30	gadE	c*g*tatgagcgacatcgtcaccctgggtatcacatcttatttttaAtcaggacataagcaactgaaattgatggctggcatgacgaggg*a*t
30	gadX	t*c*ccggtcccctatgccgggtttttttttatgtctgagtaaaactTtataatcttattccttccgcagaacggtcagtgccgtaaaaat*t*c
30	yhjK	c*a*caaaccagcacttttaaagttttgtaatcagtttggggtagTtacttttcttccaggtaactctcttcgaagatttcaatagga*g*t
30	yhjR	t*c*cccgcacccccctgctatcccagtacggccattcaccaatcccTtacttttgttgcgcaaactctgccagcaacggccagcgtttta*a*t
30	yiaG	c*g*tttgattcaagccaacccggcattaagtaagcagttgatggaataAacttttatccactttattgctgtttacggtcctgatgaca*g*g
30	xylR	t*a*tgacacgacgccaaaagagtatcgcgatgtaaatagcgaggtcatgttgtaAcgatgatgagaattgtcggcgtcacatcaggtaa*c*g
31	sgbU	g*g*agattattcaggcgcggcgttggattgaagcgcgtatgcaggaggctggatttatatgttaAagcaactgaaagccgacgtgctgg*c*g
31	rhsA	t*g*gaattgtcgtcactccagacagtcaaacgataaaaaaagataAcccttgtggaggttcctgcaatgtcaaatacataccagaaaag*a*a
31	yibA	t*g*acaatgaaattataacttccgctattgataagctgaagcgttcataAcgtgagttgcctatgcacagtgggggattcccgccggca*c*g
31	lctD	t*a*aagagttgcctgcggcactggctcccatggcgaaagggaatgcggcataAtcgtttgcgccccctcaccctaaccctctccctcag*g*g
31	yibK	g*g*tgtatgaagcctggcggcagttggggtatccgggagcggtattgagagattaAttactgtaggccggatagatgattacatcgcat*c*c
31	yibD	c*a*cgctgatcgtttacttccagcatcagtggcaaggctggggcaaacagcccgaagcggcgaagtttaacgTtagcgcaaattaaagg*c*g
31	htrL	t*a*tctcagttggtatcaaaacgctacccactaatggagggaaacTtattttcttgaaaggaaaattctcagcgcctgcatttttgccc*c*t
31	coaD	a*t*cagggcgatgtcacccatttcctgccggagaatgtccatcaggcgctgatggcgaagttagcgtaAcgtttatgccggatggtatg*c*c
31	rph	c*a*gtcgccttaaaaatcagtttgccagcgccgccttctgcgtcgTtacaatggattcgattcccctcgggccagagccaacaagatga*g*t
31	dinD	a*t*caagcagttggaaaacagtgtaaaaattacagagaagaaataAcagaaactaaaaactctttttgttgattgagacacccgatgcg*t*a
32	ligB	c*t*tactggcgctacgccttaagctgggattaccggtcttttactacagccacgaaggacaTtaaggttcaaaacctgtgatctgctgg*g*c
32	setC	t*t*gggtattgcgatgatttgcctgctgtttattaaagatatttaActcaccagtctgggctaagaaaaatcggcgtggcgcaaacaga*g*c
32	yicO	c*t*gcgtaatcagaactcatgatatctggaaaccctcgccagttaTtaatccaccagaataatcttcagtgcaaacagagctgcgacca*c*c
32	ivbL	g*g*tttgcgccggcggggttttggaatcgtgtgttgttccagtccTtacggcgcattgccgacgaccaccaccacacgcacgacgacca*c*t
32	yidH	g*c*tccggctgcagtcctgggtcggcaatgcgtctggctttgcggTtatccggcatacaacaccagtcccataacgatcaccgccacga*c*c
32	glvC	t*c*cgcgtggttgctgaacatcgggatccagttttgcggtaaaacTtagtcaatcagacctccgcccatgttgcccaccacaccaaaga*g*a

32	yidS	g*a*ttatgcgcagtggtgtggcacatattccacagttgaaggattaAccaacgcgcttcacatcgcccaccagcaggatataagagagc*g*c
32	yieH	t*t*acctgaactgtggaaagcgcgtggttgggatattacggcataAttcttcacactcccttcacttaccccgcttaaattggcgctca*a*a
32	yieJ	t*g*tctccactgcggcaagctgagactgtggggtgatttttcgtaAttatttaaataatgagaacaggccggagcgtaattcacacatc*c*g
32	atpE	t*a*gttaacgttctgatattgctctttaaataaaagcaacgcttaTtacgcgacagcgaacatcacgtacagacccagacctacagcga*t*c

Table S4. Frequency of TAG codon conversions at each locus

Region	Gene	Frequency of Codon Conversions
1	rbsB	39.1
1	rbsR	71.1
1	ilvL	41.3
1	ilvA	37.2
1	wzzE	29.8
1	aslB	20.5
1	hemY	37.2
1	yigG	51.2
1	recQ	39.5
1	rmuC	28.9
2	tatD	29.3
2	yihL	72.5
2	yihW	68.3
2	yiiD	42.5
2	cpxP	55.0
2	frwD	52.5
2	zraR	
2	yjaB	57.5
2	yjdK	10.3
2	blc	48.6
3	ulaF	43.5
3	yjfY	47.2
3	priB	45.5
3	ytfB	45.7
3	ytfG	51.1
3	msrA	40.9
3	ytfN	40.5
3	b4250	46.5
3	idnR	37.8
3	b4273	39.0
4	yjgX	23.8
4	b4283	34.1
4	fecE	37.2
4	nanC	15.4
4	fimB	
4	yjiN	27.3

4	yjiP	27.9
4	mcrB	94.9
4	yjjQ	22.0
4	yjjV	
5	lplA	48.9
5	creC	25.6
5	kefF	32.5
5	yabI	26.2
5	thiQ	14.6
5	murF	57.5
5	mutT	15.6
5	pdhR	32.6
5	sfsA	51.2
5	btuF	11.4
6	cdaR	51.2
6	yaeI	72.7
6	yafL	29.3
6	yagY	
6	eaeH	15.9
6	b0299	34.9
6	yahL	73.3
6	yahM	53.3
6	prpE	20.0
6	cynR	79.3
7	frmR	47.4
7	b0361	60.5
7	yaiS	54.8
7	tra5_1	37.8
7	araJ	25.5
7	pgpA	21.3
7	yajL	42.2
7	panE	36.2
7	lon	35.9
7	mdlA	27.7
8	ybaA	14.3
8	ybaJ	57.4
8	priC	66.0
8	aes	13.3
8	ybcF	34.8
8	tra5_2	15.2
8	ybcL	90.9
8	rzpD	44.2
8	ybcY	18.2

8	hokE	19.6
9	ybdL	14.3
9	ybdM	11.1
9	rlpA	6.8
9	ybeQ	55.3
9	ubiF	13.3
9	ybfH	12.2
9	ybfD	25.6
9	nei	31.9
9	abrB	51.5
9	sucB	26.7
10	bioD	22.7
10	moaE	52.5
10	ybhM	46.5
10	ybhS	33.3
10	ybiH	52.6
10	ybiR	17.4
10	yliD	38.9
10	yliE	44.4
10	ybjK	58.5
10	rimK	68.3
11	ybjR	21.1
11	cspD	33.3
11	lolA	10.3
11	ycaI	46.3
11	lpxK	50.0
11	ycbK	43.9
11	ycbX	22.5
11	tra5_3	34.2
11	yceO	64.3
11	yceF	
12	plsX	83.0
12	fabH	21.3
12	pabC	89.1
12	b1146	12.8
12	ymfM	78.7
12	ymfR	25.5
12	tfaE	
12	ycgX	
12	ymgC	89.4
12	ycgY	
13	hemA	34.1
13	b1228	6.1

13	kch	18.9
13	sapB	2.2
13	pspF	4.4
13	ycjU	10.9
13	ycjV	2.3
13	ycjG	4.4
13	ycjZ	9.8
13	intR	
14	sieB	31.9
14	b1354	28.3
14	ydaY	6.4
14	b1367	
14	tfaR	33.3
14	paaG	8.7
14	paaX	7.1
14	b1402	13.0
14	ydcL	2.1
14	rhsE	33.3
15	yddJ	37.8
15	ddpF	37.8
15	yddV	45.0
15	ydeV	31.0
15	ydeA	10.9
15	marA	33.3
15	marB	41.9
15	tfaQ	45.5
15	ydfP	
15	rzpQ	46.7
16	b1578	45.9
16	ynfA	42.6
16	speG	50.0
16	dmsD	40.4
17	ydhM	60.0
17	lhr	73.3
17	ynhG	
17	sufA	47.7
17	ydiA	37.8
17	ydiE	43.9
17	nlpC	13.3
17	btuC	22.0
17	arpB_1	69.8
17	ydjX	64.9
18	nudG	64.3

18	b1788	35.6
18	yeaL	15.6
18	yeaX	17.8
18	yobB	31.1
18	exoX	14.7
18	yecN	56.5
18	otsA	28.9
18	dcyD	10.9
18	yedM	32.6
19	fliE	34.8
19	fliN	25.5
19	fliP	47.8
19	fliQ	34.1
19	yedS_1	45.5
19	cbl	31.9
19	b1996	32.6
19	hisL	42.2
19	wbbJ	62.8
19	wcaM	47.4
20	wcaL	17.0
20	wcaC	45.7
20	asmA	20.0
20	baeR	21.7
20	gatR_1	36.2
20	tra5_4	17.8
20	yegV	28.3
20	yegW	17.4
20	yehQ	38.3
20	yohC	70.2
21	yohF	33.3
21	sanA	85.2
21	yejA	25.0
21	yejE	33.3
21	b2191	21.1
21	rcsD	34.9
21	rcsC	26.8
21	yfaT	50.0
21	menF	70.0
21	yfcO	2.9
22	yfcU	21.7
22	tfaS	29.8
22	ypdI	23.4
22	yfdY	21.7

22	yfeO	48.9
22	mntH	14.9
22	xapR	8.7
22	yfeR	23.4
22	yffB	41.3
22	hda	23.4
23	yfgG	6.8
23	pbpC	
23	yphA	17.0
23	yfhB	27.7
23	kgtP	38.3
23	yfiA	22.7
23	yfjQ	38.6
23	yfjR	17.0
23	ypjC	26.5
23	ygaQ	40.9
24	ygaR	57.1
24	yqaC	40.0
24	gabT	22.0
24	ygaU	37.2
24	ygaM	20.9
24	luxS	30.2
24	mltB	36.4
24	srlE	43.9
24	norW	14.6
24	ascB	
25	hycI	19.4
25	hypB	45.7
25	ygbA	15.9
25	pphB	61.4
25	ygcN	24.4
25	mazF	73.3
25	relA	19.1
25	rumA	44.7
25	ppdC	35.1
25	ptsP	14.7
26	mutH	27.7
26	ygeA	23.4
26	kduI	39.1
26	yqeF	17.8
26	ygeI	39.1
26	pbl	23.9
26	ygeM	58.5

26	b2860	23.9
26	ygeP	27.7
26	yqeC	12.8
27	ygfJ	71.1
27	recJ	34.0
27	argO	43.2
27	yggU	41.3
27	mutY	65.2
27	glcC	56.5
27	yghQ	42.6
27	yghT	31.9
27	ygiZ	15.2
27	yqiB	21.1
28	b3045	26.7
28	yqiI	10.9
28	ygiP	26.7
28	yhaK	47.8
28	tdcA	78.3
28	yhaC	68.9
28	ecfH	19.6
28	yhbW	34.8
28	nlpI	82.9
28	dacB	19.6
29	ptsN	40.0
29	yhcC	34.1
29	yhcE	24.4
29	yhcO	66.7
29	mreC	20.0
29	yhdN	18.6
29	kefB	33.3
29	frlD	33.3
29	bioH	33.3
29	gntX	40.9
30	malQ	18.2
30	glgA	51.1
30	yrhA	0.0
30	ugpQ	67.4
30	gadE	69.0
30	gadX	29.5
30	yhjK	59.1
30	yhjR	23.4
30	yiaG	28.6
30	xylR	

31	sgbU	37.0
31	rhsA	74.3
31	yibA	36.4
31	lctD	17.0
31	yibK	59.6
31	yibD	35.7
31	htrL	23.3
31	coaD	19.6
31	rph	53.5
31	dinD	
32	ligB	59.6
32	setC	51.6
32	yicO	70.2
32	ivbL	21.4
32	yidH	51.1
32	glvC	61.7
32	yidS	48.9
32	yieH	65.2
32	yieJ	56.5
32	atpE	24.2

Table S5. Frequency of TAG codon conversions.

TAG Gene Subset	# Codons Assayed	Allelic Replacement Frequency
All TAG genes	298/314	36.6% +/-19%
TAG transcript in + transcription direction	169/177	36.5% +/-19%
TAG transcript in - transcription direction	129/137	36.6% +/-19%
Essential genes with TAG	41/43	35.3% +/-19%
TAG codons overlapping 2nd reading frame	37/39	33.7% +/-19%

Frequency of allelic replacement is tabulated for: 1) all TAG codons, 2) TAG codons broken down by direction (+/-) of transcription, 3) essential genes that contain TAG codons, 4) TAG codons overlapping the reading frame of a second gene. The number of codons assayed is represented as a fraction of the number of successfucl mascPCR reactions over the total number of genes.

Table S6. Properties of MAGE-cycled Strains

Straiin Measurement	Ancestral EcNR2 Strain	Average of 32 TAG-to-TAA recoded strains
Growth Rates (30C)	42'	47'(+/-6)
Recombination Efficiencies	23%	21.6%(+/-2.5%)
Auxotrophic Rates	-	2.8%

Table S7. Five-stage hierarchical conjugation process

Stage	No. of Strains	No. of Conjugations	Average Size of Genomic Transfer	Strain Numbers
1	32	16	143 kb	rEc1-rEc32
2	16	8	267 kb	C1-C16
3	8	4	575 kb	C17-C24
4	4	2	1.15 MB	C25-C28
5	2	1	2.3 MB	C29-C31

Table S8. List and annotation of all secondary mutations

Strain	Feature	Start	End	WT	Mutant	Gene	Consequence	AA Change	COG	Type	Product
C21	snp	2055	2055	C	T	thrA	SYNONYMOUS_CODING	H	COG0527	E	fused aspartokinase I and homoserine dehydrogenase I
C21	snp	66039	66039	T	C	araD	NON_SYNONYMOUS_CODING	H/R	COG0235	G	L-ribulose-5-phosphate 4-epimerase
C21	snp	71695	71695	C	T	yabI	SYNONYMOUS_CODING	V	COG0586	S	inner membrane protein, SNARE_assoc family
C21	snp	76786	76786	A	G	sgrR	NON_SYNONYMOUS_CODING	W/R	COG4533	R	transcriptional DNA-binding transcriptional activator of sgrS sRNA
C21	snp	151352	151352	A	G	yadK	NON_SYNONYMOUS_CODING	V/A	COG3539	NU	predicted fimbrial-like adhesin protein
C21	snp	160383	160383	G	A	dksA	SYNONYMOUS_CODING	D	COG1734	T	DNA-binding transcriptional regulator of rRNA transcription, DnaK suppressor protein
C21	snp	195795	195795	C	T	cdsA	NON_SYNONYMOUS_CODING	A/V	COG0575	I	CDP-diglyceride synthase
C21	snp	332913	332913	C	T	yahB	NON_SYNONYMOUS_CODING	V/I	COG0583	K	predicted DNA-bindng transcriptional regulator
C21	indel	342758	342758	G	-	yahK	FRAMESHIFT_CODING	-	COG1064	R	predicted oxidoreductase, Zn-dependent and NAD(P)-binding
C21	snp	444898	444898	C	T	yajR	SYNONYMOUS_CODING	S	COG2814	G	predicted transporter
C21	snp	547694	547694	A	G	ylbE	SYNONYMOUS_CODING	E	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	indel	547831	547832	-	G	ylbE	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	658193	658193	A	G	tatE	SYNONYMOUS_CODING	K	-	-	TatABCE protein translocation system subunit
C21	snp	760107	760107	C	T	sucA	NON_SYNONYMOUS_CODING	P/S	COG0567	C	2-oxoglutarate decarboxylase, thiamin-requiring
C21	snp	804520	804520	G	A	ybhJ	NON_SYNONYMOUS_CODING	A/T	COG1048	C	predicted hydratase
C21	snp	822574	822574	T	C	ybhO	NON_SYNONYMOUS_CODING	N/S	COG1502	I	cardiolipin synthase 2
C21	snp	827752	827752	A	G	ybhF	SYNONYMOUS_CODING	C	COG1131	V	fused predicted transporter subunits of ABC superfamily: ATP-binding components
C21	snp	853211	853211	A	G	ybiR	SYNONYMOUS_CODING	R	COG0471	P	predicted transporter

C21	indel	883564	883564	G	-	cmr	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	987651	987651	G	A	asnS	SYNONYMOUS_CODING	N	COG0017	J	asparaginyl tRNA synthetase
C21	snp	995889	995889	A	G	ssuA	SYNONYMOUS_CODING	A	COG0715	P	putative aliphatic sulfonate binding protein
C21	snp	1005546	1005546	C	T	ycbW	SYNONYMOUS_CODING	V	-	-	predicted protein
C21	snp	1056395	1056395	A	G	torT	SYNONYMOUS_CODING	P	COG1879	G	periplasmic sensory protein associated with the TorRS two-component regulatory system
C21	indel	1093083	1093084	-	G	ycdT	FRAMESHIFT_CODING	-	COG3706	T	diguanylate cyclase
C21	snp	1110049	1110049	A	G	mdoG	NON_SYNONYMOUS_CODING	T/A	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	1146795	1146795	A	G	plsX	UPSTREAM (-49)	-	COG0416	I	probable phosphate acyltransferase
C21	snp	1241046	1241046	A	G	cvrA	SYNONYMOUS_CODING	S	COG3263	P	predicted cation/proton antiporter
C21	snp	1284157	1284157	A	G	narH	NON_SYNONYMOUS_CODING	Y/C	COG1140	C	nitrate reductase 1, beta (Fe-S) subunit
C21	snp	1456641	1456641	A	G	paaG	SYNONYMOUS_CODING	A	COG1024	I	predicted enoyl-CoA hydratase
C21	snp	1493717	1493717	A	G	ydcJ	NON_SYNONYMOUS_CODING	I/V	COG5383	S	predicted metalloenzyme
C21	snp	1513952	1513952	T	C	ydcW	SYNONYMOUS_CODING	C	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	1537774	1537774	G	A	narZ	SYNONYMOUS_CODING	T	COG5013	C	nitrate reductase 2 (NRZ), alpha subunit
C21	snp	1562401	1562401	T	C	dos	NON_SYNONYMOUS_CODING	T/A	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	1657766	1657766	G	A	ynfE	STOP_GAINED	W/*	COG0243	C	probable selenate reductase, periplasmic
C21	snp	1693853	1693853	G	A	uidA	SYNONYMOUS_CODING	G	COG3250	G	beta-D-glucuronidase
C21	snp	1753006	1753006	C	T	ydhZ	NON_SYNONYMOUS_CODING	E/K	-	-	predicted protein
C21	snp	1837592	1837592	G	A	ynjE	SYNONYMOUS_CODING	Q	COG2897	P	predicted thiosulfate sulfur transferase
C21	indel	1866698	1866698	C	-	yeaG	FRAMESHIFT_CODING	-	COG2766	T	protein kinase, function unknown; autokinase
C21	indel	1950710	1950710	C	-	cmoA	UPSTREAM (-16)	-	COG2226	H	tRNA cmo(5)U34 methyltransferase, SAM-dependent
C21	snp	1959694	1959694	T	C	argS	NON_SYNONYMOUS_CODING	C/R	COG0018	J	arginyl-tRNA synthetase
C21	snp	1971125	1971125	A	G	cheW	NON_SYNONYMOUS_CODING	V/A	COG0835	NT	purine-binding chemotaxis protein

C21	snp	1972133	1972133	A	G	cheA	NON_SYNONYMOUS_CODING	S/P	COG0643	NT	fused chemotactic sensory histidine kinase in two-component regulatory system with CheB and CheY: sensory histidine kinase/signal sensing protein
C21	indel	2016276	2016277	-	A	fliJ	FRAMESHIFT_CODING	-	COG2882	NUO	flagellar protein
C21	snp	2037966	2037966	G	A	yedY	SYNONYMOUS_CODING	P	COG2041	R	membrane-anchored, periplasmic TMAO, DMSO reductase
C21	snp	2044469	2044469	C	T	yeeJ	NON_SYNONYMOUS_CODING	P/L	-	-	probable adhesin
C21	snp	2081468	2081468	T	C	sbcB	NON_SYNONYMOUS_CODING	V/A	COG2925	L	exonuclease I
C21	snp	2180057	2180057	C	T	yegW	SYNONYMOUS_CODING	*	COG2188	K	predicted DNA-binding transcriptional regulator
C21	snp	2208976	2208976	A	T	yehQ	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	2247576	2247576	T	C	yeiE	NON_SYNONYMOUS_CODING	Q/R	COG0583	K	predicted DNA-binding transcriptional regulator
C21	snp	2277333	2277333	T	C	bcr	SYNONYMOUS_CODING	A	COG2814	G	bicyclomycin/multidrug efflux system
C21	snp	2360599	2360599	T	C	yfaY	NON_SYNONYMOUS_CODING	S/G	COG1058	R	conserved protein
C21	snp	2407281	2407281	T	C	yfbR	NON_SYNONYMOUS_CODING	V/A	COG1896	R	5'-nucleotidase
C21	indel	2483331	2483332	-	C	evgS	FRAMESHIFT_CODING	-	COG0642	T	hybrid sensory histidine kinase in two-component regulatory system with EvgA
C21	snp	2520915	2520915	T	C	xapB	NON_SYNONYMOUS_CODING	S/G	COG2211	G	xanthosine transporter
C21	snp	2554980	2554980	A	G	eutC	NON_SYNONYMOUS_CODING	W/R	COG4302	E	ethanolamine ammonia-lyase, small subunit (light chain)
C21	snp	2556896	2556896	T	C	intZ	NON_SYNONYMOUS_CODING	L/P	COG0582	L	CPZ-55 prophage; predicted integrase
C21	snp	2573112	2573112	G	A	eutP	NON_SYNONYMOUS_CODING	P/L	COG4917	E	conserved protein with nucleoside triphosphate hydrolase domain
C21	snp	2609439	2609439	A	G	hyfI	NON_SYNONYMOUS_CODING	S/G	COG3260	C	hydrogenase 4, Fe-S subunit
C21	snp	2641755	2641755	T	C	rlmN	NON_SYNONYMOUS_CODING	N/S	COG0820	R	23S rRNA m(2)A2503 methyltransferase, SAM-dependen
C21	snp	2853870	2853870	G	A	fhlA	NON_SYNONYMOUS_CODING	R/H	COG3604	KT	DNA-binding transcriptional activator

C21	snp	2935824	2935824	A	G	fucK	NON_SYNONYMOUS_CODING	D/G	COG1070	G	L-fuculokinase
C21	snp	3119818	3119818	T	C	glcB	SYNONYMOUS_CODING	Q	COG2225	C	malate synthase G
C21	snp	3122613	3122613	G	A	glcF	NON_SYNONYMOUS_CODING	P/L	COG0247	C	glycolate oxidase iron-sulfur subunit
C21	snp	3154895	3154895	G	A	dkgA	NON_SYNONYMOUS_CODING	R/H	COG0656	R	2,5-diketo-D-gluconate reductase A
C21	snp	3176187	3176187	T	C	tolC	SYNONYMOUS_CODING	S	COG1538	MU	transport channel
C21	snp	3176753	3176753	A	G	tolC	NON_SYNONYMOUS_CODING	Q/R	COG1538	MU	transport channel
C21	snp	3177340	3177340	T	C	tolC	NON_SYNONYMOUS_CODING	Y/H	COG1538	MU	transport channel
C21	snp	3193359	3193359	G	A	hldE	STOP_GAINED	Q/*	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	3194503	3194503	G	A	hldE	SYNONYMOUS_CODING	V	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	3211874	3211874	T	C	rpoD	NON_SYNONYMOUS_CODING	L/P	COG0568	K	RNA polymerase, sigma 70 (sigma D) factor
C21	snp	3233809	3233809	G	A	rlmG	NON_SYNONYMOUS_CODING	A/V	COG2813	J	23S rRNA mG1835 methyltransferase, SAM-dependent
C21	snp	3246194	3246194	T	C	yqjA	NON_SYNONYMOUS_CODING	F/L	COG0586	S	required, with yghB, for membrane integrity; inner membrane protein
C21	snp	3322912	3322912	T	C	folP	NON_SYNONYMOUS_CODING	T/A	COG0294	H	7,8-dihydropteroate synthase
C21	snp	3356518	3356518	G	A	gltB	NON_SYNONYMOUS_CODING	D/N	COG0069	E	glutamate synthase, large subunit
C21	snp	3364975	3364975	G	A	yhcF	NON_SYNONYMOUS_CODING	G/R	-	-	predicted transcriptional regulator
C21	snp	3369748	3369748	G	A	nanT	SYNONYMOUS_CODING	V	COG2814	G	sialic acid transporter
C21	snp	3462098	3462098	T	C	gspL	NON_SYNONYMOUS_CODING	V/A	COG3297	U	general secretory pathway component, cryptic
C21	snp	3542361	3542361	G	A	bioH	NON_SYNONYMOUS_CODING	P/L	COG0596	R	carboxylesterase of pimeloyl-CoA synthesis
C21	snp	3753588	3753588	T	C	aldB	NON_SYNONYMOUS_CODING	Q/R	COG1012	C	aldehyde dehydrogenase B
C21	snp	3794765	3794765	A	G	rfaC	NON_SYNONYMOUS_CODING	K/R	COG0859	M	ADP-heptose:LPS heptosyl transferase I
C21	snp	3831732	3831732	A	G	yicI	NON_SYNONYMOUS_CODING	F/L	COG1501	G	predicted alpha-glucosidase
C21	snp	3867170	3867170	G	A	yidR	NON_SYNONYMOUS_CODING	R/C	COG0823	U	conserved protein
C21	snp	3923701	3923701	G	A	mnmG	5PRIME_UTR	-	COG0445	D	5-methylaminomethyl-2-thiouridine modification at tRNA U34

C21	snp	4051916	4051916	C	T	glnG	SYNONYMOUS_CODING	K	COG2204	T	fused DNA-binding response regulator in two-component regulatory system with GlnL: response regulator/sigma54 interaction protein
C21	snp	4054114	4054114	C	T	glnL	SYNONYMOUS_CODING	L	COG3852	T	sensory histidine kinase in two-component regulatory system with GlnG
C21	snp	4085593	4085593	A	G	yiiG	NON_SYNONYMOUS_CODING	H/R	-	-	conserved lipoprotein
C21	snp	4111725	4111725	T	C	uspD	NON_SYNONYMOUS_CODING	V/A	COG0589	T	stress-induced protein
C21	indel	4159177	4159178	-	C	fabR	FRAMESHIFT_CODING	-	COG1309	K	DNA-binding transcriptional repressor
C21	snp	4161425	4161425	G	A	btuB	5PRIME_UTR	-	COG4206	H	vitamin B12/cobalamin outer membrane transporter
C21	indel	4271828	4271828	G	-	uvrA	FRAMESHIFT_CODING	-	COG0178	L	ATPase and DNA damage recognition protein of nucleotide excision repair excinuclease UvrABC
C21	snp	4407528	4407528	A	G	rlmB	SYNONYMOUS_CODING	G	COG0566	J	23S rRNA (Gm2251)-methyltransferase
C21	indel	4441815	4441815	G	-	ytfM	FRAMESHIFT_CODING	-	COG0729	M	predicted outer membrane protein and surface antigen
C21	indel	4473579	4473580	-	G	yjgL	FRAMESHIFT_CODING	-	-	-	predicted protein
C21	snp	4490470	4490470	C	T	idnT	NON_SYNONYMOUS_CODING	A/T	COG2610	GE	L-idonate and D-gluconate transporter
C21	snp	4517532	4517532	A	G	yjhU	SYNONYMOUS_CODING	I	COG2390	K	KpLE2 phage-like element; predicted DNA-binding transcriptional regulator
C21	snp	4526189	4526189	A	G	sgcQ	NON_SYNONYMOUS_CODING	V/A	COG0434	R	KpLE2 phage-like element; predicted nucleoside triphosphatase
C21	snp	4567683	4567683	C	T	yjiP	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C21	snp	4589938	4589938	G	A	tsr	NON_SYNONYMOUS_CODING	V/I	COG0840	NT	methyl-accepting chemotaxis protein I, serine sensor receptor

C22	snp	21442	21442	G	A	ribF	SYNONYMOUS_CODING	Q	COG0196	H	bifunctional riboflavin kinase/FAD synthetase
C22	snp	48106	48106	A	G	kefC	NON_SYNONYMOUS_CODING	Q/R	COG0475	P	potassium:proton antiporter
C22	snp	64216	64216	C	T	polB	NON_SYNONYMOUS_CODING	R/H	COG0417	L	DNA polymerase II
C22	snp	76664	76664	C	T	sgrR	SYNONYMOUS_CODING	P	COG4533	R	transcriptional DNA-binding transcriptional activator of sgrS sRNA
C22	snp	81083	81083	G	A	leuB	SYNONYMOUS_CODING	G	COG0473	CE	3-isopropylmalate dehydrogenase
C22	snp	95687	95687	A	G	murF	SYNONYMOUS_CODING	V	COG0770	M	UDP-N-acetylmuramoyl-tripeptide:D-alanyl-D-alanine ligase
C22	snp	97551	97551	C	T	murD	SYNONYMOUS_CODING	Y	COG0771	M	UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase
C22	snp	114829	114829	A	G	hofC	NON_SYNONYMOUS_CODING	L/P	COG1459	NU	assembly protein in type IV pilin biogenesis, transmembrane protein
C22	snp	178904	178904	A	G	mtnN	SYNONYMOUS_CODING	L	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	329644	329644	A	G	betT	NON_SYNONYMOUS_CODING	T/A	COG1292	M	choline transporter of high affinity
C22	snp	352488	352488	C	T	prpE	NON_SYNONYMOUS_CODING	H/Y	COG0365	I	propionate--CoA ligase
C22	snp	385577	385577	T	C	tauB	SYNONYMOUS_CODING	G	COG4525	P	taurine transporter subunit
C22	indel	389540	389541	-	T	yaiT	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	392842	392842	G	A	yaiT	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	401199	401199	G	A	phoA	NON_SYNONYMOUS_CODING	D/N	COG1785	P	bacterial alkaline phosphatase
C22	snp	495410	495410	T	C	htpG	NON_SYNONYMOUS_CODING	V/A	COG0326	O	molecular chaperone HSP90 family
C22	snp	500572	500572	A	G	gsk	SYNONYMOUS_CODING	V	COG0524	G	inosine/guanosine kinase
C22	snp	510896	510896	T	C	glsA1	NON_SYNONYMOUS_CODING	V/A	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	517157	517157	A	G	ybbN	NON_SYNONYMOUS_CODING	L/P	COG3118	O	DnaK co-chaperone, thioredoxin-like protein
C22	snp	517804	517804	A	G	ybbO	SYNONYMOUS_CODING	G	COG0300	R	predicted oxidoreductase with NAD(P)-binding Rossmann-fold domain
C22	snp	524652	524652	A	G	rhsD	NON_SYNONYMOUS_CODING	Q/R	COG3209	M	rhsD element protein
C22	snp	547694	547694	A	G	ylbE	SYNONYMOUS_CODING	E	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	indel	547831	547832	-	G	ylbE	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	549600	549600	A	G	ylbF	NON_SYNONYMOUS_CODING	T/A	-	-	conserved protein

C22	indel	560913	560914	-	G	sfmD	FRAMESHIFT_CODING	-	COG3188	NU	predicted outer membrane export usher protein
C22	snp	609015	609015	T	C	entD	SYNONYMOUS_CODING	L	COG2977	Q	phosphopantetheinyltransferase component of enterobactin synthase multienzyme complex
C22	snp	626595	626595	A	G	entE	NON_SYNONYMOUS_CODING	I/V	COG1021	Q	2,3-dihydroxybenzoate-AMP ligase component of enterobactin synthase multienzyme complex
C22	snp	686392	686392	G	A	gltI	SYNONYMOUS_CODING	T	COG0834	ET	glutamate and aspartate transporter subunit
C22	indel	867638	867638	C	-	gsiA	FRAMESHIFT_CODING	-	COG1123	R	glutathione transporter ATP-binding protein, ABC superfamily
C22	snp	877125	877125	G	T	rimO	NON_SYNONYMOUS_CODING	A/E	COG0621	J	ribosomal protein S12 methylthiotransferase; radical SAM superfamily
C22	snp	894684	894684	T	C	potG	SYNONYMOUS_CODING	R	COG3842	E	putrescine transporter subunit: ATP-binding component of ABC superfamily
C22	indel	913399	913400	-	A	ybjE	FRAMESHIFT_CODING	-	COG2431	S	predicted transporter
C22	snp	921252	921252	A	G	macB	SYNONYMOUS_CODING	V	COG1136	V	fused macrolide transporter subunits of ABC superfamily: ATP-binding component/membrane component
C22	snp	966954	966954	A	G	msbA	NON_SYNONYMOUS_CODING	T/A	COG1132	V	fused lipid transporter subunits of ABC superfamily: membrane component/ATP-binding component
C22	snp	1015393	1015393	C	T	fabA	NON_SYNONYMOUS_CODING	E/K	COG0764	I	beta-hydroxydecanoyl thioester dehydrase
C22	indel	1026029	1026030	-	C	mgsA	FRAMESHIFT_CODING	-	COG1803	G	methylglyoxal synthase
C22	snp	1059816	1059816	C	T	torA	SYNONYMOUS_CODING	S	COG0243	C	trimethylamine N-oxide (TMAO) reductase I, catalytic subunit
C22	snp	1093365	1093365	C	T	ycdT	STOP_GAINED	R/*	COG3706	T	diguanylate cyclase

C22	snp	1194738	1194738	A	G	icd	SYNONYMOUS_CODING	V	COG0538	C	e14 prophage; isocitrate dehydrogenase, specific for NADP+
C22	indel	1293726	1293726	G	-	insZ	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	1298991	1298991	G	A	oppB	5PRIME_UTR	-	COG0601	EP	oligopeptide transporter subunit
C22	snp	1356669	1356669	T	C	puuP	SYNONYMOUS_CODING	G	COG0531	E	putrescine importer
C22	indel	1377889	1377890	-	G	ycjT	FRAMESHIFT_CODING	-	COG1554	G	predicted hydrolase
C22	indel	1435247	1435247	T	-	MicC	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	1456641	1456641	A	G	paaG	SYNONYMOUS_CODING	A	COG1024	I	predicted enoyl-CoA hydratase
C22	snp	1549184	1549184	C	T	fdnH	NON_SYNONYMOUS_CODING	P/S	COG0437	C	formate dehydrogenase-N, Fe-S (beta) subunit, nitrate-inducible
C22	indel	1592272	1592273	-	C	yneO	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	1645619	1645619	G	A	ydfX	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	1731204	1731204	G	A	lhr	NON_SYNONYMOUS_CODING	G/D	COG1201	R	predicted ATP-dependent helicase
C22	snp	1873938	1873938	G	A	yeaN	NON_SYNONYMOUS_CODING	S/N	COG2807	P	predicted transporter
C22	snp	1913971	1913971	T	C	yebR	NON_SYNONYMOUS_CODING	Q/R	COG1956	T	free methionine-(R)-sulfoxide reductase
C22	snp	1951551	1951551	A	G	cmoB	NON_SYNONYMOUS_CODING	N/S	COG2227	H	tRNA mo(5)U34 methyltransferase, SAM-dependent
C22	snp	1964524	1964524	G	A	cheZ	NON_SYNONYMOUS_CODING	R/C	COG3143	NT	chemotaxis regulator, protein phosphatase for CheY
C22	snp	2007291	2007291	G	A	yedE	NON_SYNONYMOUS_CODING	V/M	COG2391	R	predicted inner membrane protein
C22	snp	2014597	2014597	G	A	fliI	NON_SYNONYMOUS_CODING	R/H	COG1157	NU	flagellum-specific ATP synthase
C22	snp	2034663	2034663	T	C	hchA	SYNONYMOUS_CODING	L	COG0693	R	Hsp31 molecular chaperone
C22	snp	2272188	2272188	A	G	yejA	SYNONYMOUS_CODING	K	COG4166	E	microcin C transporter YejABEF, periplasmic binding protein; ABC family
C22	snp	2356185	2356185	G	A	yfaU	NON_SYNONYMOUS_CODING	A/V	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	2414786	2414786	T	C	pta	NON_SYNONYMOUS_CODING	L/P	COG0857	R	phosphate acetyltransferase

C22	snp	2492972	2492972	T	C	ypdI	NON_SYNONYMOUS_CODING	Y/H	-	-	predicted lipoprotein involved in colanic acid biosynthesis
C22	snp	2532053	2532053	T	C	ptsI	5PRIME_UTR	-	COG1080	G	PEP-protein phosphotransferase of PTS system (enzyme I)
C22	snp	2593249	2593249	T	C	ypfI	SYNONYMOUS_CODING	Q	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	2682993	2682993	G	A	glyA	SYNONYMOUS_CODING	G	COG0112	E	serine hydroxymethyltransferase
C22	snp	2683583	2683583	T	C	glyA	5PRIME_UTR	-	COG0112	E	serine hydroxymethyltransferase
C22	snp	2794168	2794168	C	T	ygaE	NON_SYNONYMOUS_CODING	A/V	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	2794938	2794938	T	C	yqaE	NON_SYNONYMOUS_CODING	Y/C	COG0401	S	regulated by cyaR sRNA, UPF0057 family; predicted membrane protein
C22	indel	2822518	2822518	G	-	mltB	FRAMESHIFT_CODING	-	COG2951	M	membrane-bound lytic murein transglycosylase B
C22	snp	2846778	2846778	A	G	hycC	SYNONYMOUS_CODING	T	COG0651	CP	hydrogenase 3, membrane subunit
C22	snp	2892481	2892481	T	C	ygcP	SYNONYMOUS_CODING	S	COG1954	K	predicted anti-terminator regulatory protein
C22	snp	2924262	2924262	G	A	ygdH	UPSTREAM (-68)	-	COG1611	R	conserved protein, UPF0717 family
C22	snp	2930004	2930004	T	C	fucO	SYNONYMOUS_CODING	V	COG1454	C	L-1,2-propanediol oxidoreductase
C22	snp	2935824	2935824	A	G	fucK	NON_SYNONYMOUS_CODING	D/G	COG1070	G	L-fuculokinase
C22	snp	2978919	2978919	C	T	araE	STOP_GAINED	W/*	COG2814	G	arabinose transporter
C22	snp	2991561	2991561	A	G	ygeI	UPSTREAM (-99)	-	-	-	predicted protein
C22	snp	3048078	3048078	C	T	gcvT	SYNONYMOUS_CODING	A	COG0404	E	aminomethyltransferase, tetrahydrofolate-dependent, subunit (T protein) of glycine cleavage complex
C22	indel	3100356	3100357	-	C	trmB	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	3126227	3126227	A	G	glcC	UPSTREAM (-1)	-	COG2186	K	DNA-binding transcriptional dual regulator, glycolate-binding
C22	snp	3145438	3145438	T	C	yghX	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND

C22	snp	3168182	3168182	G	A	qseB	SYNONYMOUS_CODING	R	COG0745	TK	quorum sensing DNA-binding response regulator in two-component regulatory system with QseC
C22	snp	3176270	3176270	A	G	tolC	NON_SYNONYMOUS_CODING	D/G	COG1538	MU	transport channel
C22	snp	3176601	3176601	C	T	tolC	SYNONYMOUS_CODING	I	COG1538	MU	transport channel
C22	snp	3176753	3176753	A	G	tolC	NON_SYNONYMOUS_CODING	Q/R	COG1538	MU	transport channel
C22	snp	3176813	3176813	A	G	tolC	NON_SYNONYMOUS_CODING	D/G	COG1538	MU	transport channel
C22	snp	3177340	3177340	T	C	tolC	NON_SYNONYMOUS_CODING	Y/H	COG1538	MU	transport channel
C22	snp	3177366	3177366	T	C	tolC	SYNONYMOUS_CODING	N	COG1538	MU	transport channel
C22	snp	3206772	3206772	A	G	ttdT	NON_SYNONYMOUS_CODING	S/G	COG0471	P	L-tartrate/succinate antiporte
C22	snp	3233951	3233951	T	C	ygjP	UPSTREAM (-31)	-	COG1451	R	predicted metal dependent hydrolase
C22	snp	3244782	3244782	T	C	exuR	NON_SYNONYMOUS_CODING	F/L	COG2186	K	DNA-binding transcriptional repressor
C22	indel	3268503	3268504	-	C	RNaseP_bact_a	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	3339659	3339659	A	G	kdsD	SYNONYMOUS_CODING	V	COG0794	M	D-arabinose 5-phosphate isomerase
C22	snp	3364975	3364975	G	A	yhcF	NON_SYNONYMOUS_CODING	G/R	-	-	predicted transcriptional regulator
C22	snp	3476899	3476899	G	A	kefB	SYNONYMOUS_CODING	R	COG0475	P	potassium:proton antiporter
C22	snp	3533758	3533758	G	A	envZ	STOP_GAINED	Q/*	COG0642	T	sensory histidine kinase in two-component regulatory system with OmpR
C22	snp	3555218	3555218	T	C	rtcB	NON_SYNONYMOUS_CODING	Y/C	COG1690	S	conserved protein
C22	snp	3830564	3830564	T	C	yicI	NON_SYNONYMOUS_CODING	N/S	COG1501	G	predicted alpha-glucosidase
C22	snp	3880833	3880833	A	G	dnaA	NON_SYNONYMOUS_CODING	M/T	COG0593	L	chromosomal replication initiator protein DnaA, DNA-binding transcriptional dual regulator
C22	snp	3908724	3908724	C	T	pstS	STOP_GAINED	W/*	COG0226	P	periplasmic phosphate binding protein, high-affinity

C22	snp	3947340	3947340	G	A	yifB	NON_SYNONYMOUS_CODING	P/L	COG0606	O	predicted bifunctional enzyme and transcriptional regulator
C22	indel	3968441	3968443	ACG	-	wecB	NON_SYNONYMOUS_CODING	T/-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	snp	3978613	3978613	C	T	wecG	NON_SYNONYMOUS_CODING	T/M	NOT_FOUND	NOT_FOUND	NOT_FOUND
C22	indel	4122303	4122303	C	-	cytR	FRAMESHIFT_CODING	-	COG1609	K	DNA-binding transcriptional dual regulator
C22	indel	4159177	4159178	-	C	fabR	FRAMESHIFT_CODING	-	COG1309	K	DNA-binding transcriptional repressor
C22	snp	4222412	4222412	A	G	metH	NON_SYNONYMOUS_CODING	N/D	COG1410	E	homocysteine-N5-methyltetrahydrofolate transmethylase, B12-dependent
C22	indel	4364734	4364735	-	C	dcuA	FRAMESHIFT_CODING	-	COG2704	R	C4-dicarboxylate antiporter
C22	snp	4370483	4370483	A	G	groL	NON_SYNONYMOUS_CODING	N/S	COG0459	O	Cpn60 chaperonin GroEL, large subunit of GroESL
C22	snp	4420466	4420466	A	G	ulaD	SYNONYMOUS_CODING	V	COG0269	G	3-keto-L-gulonate 6-phosphate decarboxylase
C22	snp	4432875	4432875	C	T	cpdB	NON_SYNONYMOUS_CODING	A/T	COG0737	F	2':3'-cyclic-nucleotide 2'-phosphodiesterase
C22	snp	4510965	4510965	A	G	fecC	SYNONYMOUS_CODING	G	COG0609	P	KpLE2 phage-like element; iron-dicitrate transporter subunit
C22	snp	4576825	4576825	A	G	mcrB	NON_SYNONYMOUS_CODING	I/T	COG1401	V	5-methylcytosine-specific restriction enzyme McrBC, subunit McrB
C23	snp	31174	31174	G	A	carB	NON_SYNONYMOUS_CODING	A/T	COG0458	EF	carbamoyl-phosphate synthase large subunit
C23	snp	86391	86391	A	G	ilvI	SYNONYMOUS_CODING	A	COG0028	EH	acetolactate synthase III, large subunit
C23	snp	109201	109201	A	G	secA	NON_SYNONYMOUS_CODING	H/R	COG0653	U	preprotein translocase subunit, ATPase
C23	snp	155935	155935	T	C	ecpD	SYNONYMOUS_CODING	K	COG3121	NU	predicted periplasmic pilin chaperone
C23	snp	255356	255356	C	T	pepD	NON_SYNONYMOUS_CODING	A/T	COG2195	E	aminoacyl-histidine dipeptidase (peptidase D)
C23	snp	348383	348383	G	A	prpB	NON_SYNONYMOUS_CODING	D/N	COG2513	G	2-methylisocitrate lyase
C23	snp	402393	402393	G	T	psiF	5PRIME_UTR	-	-	-	conserved protein, PsiF family, pho regulon
C23	indel	410116	410117	-	G	mak	FRAMESHIFT_CODING	-	COG1940	KG	manno(fructo)kinase

C23	snp	544887	544887	A	G	allD	NON_SYNONYMOUS_CODING	V/A	COG2055	C	ureidoglycolate dehydrogenase
C23	snp	547694	547694	A	G	ylbE	SYNONYMOUS_CODING	E	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	indel	547831	547832	-	G	ylbE	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	701381	701381	C	T	nagA	SYNONYMOUS_CODING	A	COG1820	G	N-acetylglucosamine-6-phosphate deacetylase
C23	snp	839948	839948	G	A	fiu	SYNONYMOUS_CODING	I	COG4774	P	catecholate siderophore receptor Fiu
C23	snp	856139	856139	C	T	ybiT	SYNONYMOUS_CODING	N	COG0488	R	fused predicted transporter subunits of ABC superfamily: ATP-binding components
C23	snp	878192	878192	C	T	yliI	SYNONYMOUS_CODING	D	COG2133	G	soluble aldose sugar dehydrogenase
C23	snp	900391	900391	G	A	artM	STOP_GAINED	Q/*	COG4160	E	arginine transporter subunit
C23	snp	940791	940791	T	C	dmsA	NON_SYNONYMOUS_CODING	Y/H	COG0243	C	dimethyl sulfoxide reductase, anaerobic, subunit A
C23	snp	940992	940992	G	A	dmsA	NON_SYNONYMOUS_CODING	D/N	COG0243	C	dimethyl sulfoxide reductase, anaerobic, subunit A
C23	snp	971270	971270	C	T	ycbJ	NON_SYNONYMOUS_CODING	S/L	-	-	conserved protein
C23	snp	1028847	1028847	T	C	rlmI	NON_SYNONYMOUS_CODING	T/A	COG1092	R	23S rRNA m(5)C1962 methyltransferase, SAM-dependent
C23	indel	1083798	1083799	-	G	ycdB	FRAMESHIFT_CODING	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	indel	1094834	1094835	-	A	ymdE	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	1134404	1134404	C	T	flgG	SYNONYMOUS_CODING	I	COG4786	N	flagellar component of cell-distal portion of basal-body rod
C23	snp	1214078	1214078	C	T	ycgF	SYNONYMOUS_CODING	G	COG2200	T	anti-repressor for YcgE, blue light-responsive; FAD-binding; has c-di-GMP phosphodiesterase-like EAL domain, but does not degrade c-di-GMP
C23	indel	1418811	1418812	-	G	ydaT	FRAMESHIFT_CODING	-	-	-	Rac prophage; predicted protein

C23	snp	1438612	1438612	G	A	ydbK	NON_SYNONYMOUS_CODING	T/I	COG0674	C	fused predicted pyruvate-flavodoxin oxidoreductase: conserved protein/conserved protein/FeS binding protein
C23	snp	1456641	1456641	A	G	paaG	SYNONYMOUS_CODING	A	COG1024	I	predicted enoyl-CoA hydratase
C23	snp	1470771	1470771	C	T	ydbA	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	1478597	1478597	C	T	ynbC	NON_SYNONYMOUS_CODING	P/L	COG2267	I	predicted hydrolase
C23	indel	1506241	1506241	G	-	ydcP	FRAMESHIFT_CODING	-	COG0826	O	predicted peptidase
C23	indel	1590763	1590764	-	C	yneO	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	1605564	1605564	G	A	tam	SYNONYMOUS_CODING	P	COG4106	R	trans-aconitate methyltransferase
C23	snp	1858728	1858728	T	C	ydjL	NON_SYNONYMOUS_CODING	D/G	COG1063	ER	predicted oxidoreductase, Zn-dependent and NAD(P)-binding
C23	indel	1887631	1887632	-	C	fadD	FRAMESHIFT_CODING	-	COG0318	IQ	acyl-CoA synthetase (long-chain-fatty-acid--CoA ligase)
C23	snp	1934359	1934359	T	C	zwf	5PRIME_UTR	-	COG0364	G	glucose-6-phosphate 1-dehydrogenase
C23	snp	1934660	1934660	T	C	hexR	UPSTREAM (-16)	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	1946339	1946339	C	T	nudB	SYNONYMOUS_CODING	A	COG0494	LR	dihydroneopterin triphosphate pyrophosphatase
C23	snp	1964142	1964142	T	C	flhB	NON_SYNONYMOUS_CODING	Q/R	COG1377	NU	flagellin export apparatus, substrate specificity protein
C23	snp	2031261	2031261	G	A	yedR	NON_SYNONYMOUS_CODING	A/V	-	-	inner membrane protein
C23	snp	2041770	2041770	G	A	mtfA	SYNONYMOUS_CODING	T	COG3228	S	anti-repressor for DgsA(Mlc)
C23	snp	2120281	2120281	G	A	manB	SYNONYMOUS_CODING	G	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	2125083	2125083	G	A	fcl	SYNONYMOUS_CODING	S	COG0451	MG	bifunctional GDP-fucose synthetase: GDP-4-dehydro-6-deoxy-D-mannose epimerase/ GDP-4-dehydro-6-L-deoxygalactose reductase
C23	snp	2157936	2157936	A	G	mdtC	SYNONYMOUS_CODING	V	COG0841	V	multidrug efflux system, subunit C

C23	snp	2235267	2235267	G	A	mglC	NON_SYNONYMOUS_CODING	T/I	COG4211	G	methyl-galactoside transporter subunit
C23	snp	2277333	2277333	T	C	bcr	SYNONYMOUS_CODING	A	COG2814	G	bicyclomycin/multidrug efflux system
C23	snp	2334052	2334052	G	A	yfaA	SYNONYMOUS_CODING	D	COG4685	S	conserved protein, DUF2138 family
C23	snp	2354609	2354609	A	G	glpC	NON_SYNONYMOUS_CODING	Q/R	COG0247	C	sn-glycerol-3-phosphate dehydrogenase (anaerobic), small subunit
C23	snp	2412038	2412038	G	A	ackA	NON_SYNONYMOUS_CODING	G/S	COG0282	C	acetate kinase A and propionate kinase 2
C23	snp	2475477	2475477	T	C	dsdC	NON_SYNONYMOUS_CODING	I/V	COG0583	K	DNA-binding transcriptional dual regulator
C23	snp	2480643	2480643	A	G	emrK	NON_SYNONYMOUS_CODING	V/A	COG1566	V	EmrKY-TolC multidrug resistance efflux pump, membrane fusion protein component
C23	snp	2578169	2578169	C	T	tktB	NON_SYNONYMOUS_CODING	T/M	COG0021	G	transketolase 2, thiamin-binding
C23	snp	2603047	2603047	T	C	hyfD	NON_SYNONYMOUS_CODING	L/P	COG1009	CP	hydrogenase 4, membrane subunit
C23	snp	2614239	2614239	A	G	yfgC	NON_SYNONYMOUS_CODING	S/G	COG4783	R	predicted peptidase, has TPR repeats
C23	indel	2651708	2651709	-	C	IS128	WITHIN_NON_CODING_GENE	-	NOT_FOUND	NOT_FOUND	NOT_FOUND
C23	snp	2708059	2708059	A	G	rpoE	5PRIME_UTR	-	COG1595	K	RNA polymerase, sigma 24 (sigma E) factor
C23	snp	2717738	2717738	G	A	yfiP	NON_SYNONYMOUS_CODING	R/Q	COG3148	S	conserved protein, DTW domain
C23	snp	2776890	2776890	T	C	ypjA	NON_SYNONYMOUS_CODING	T/A	COG3468	MU	adhesin-like autotransporter
C23	indel	2891983	2891985	CAC	-	ygcO	NON_SYNONYMOUS_CODING	H/-	COG2440	C	predicted 4Fe-4S cluster-containing protein
C23	snp	2934953	2934953	T	C	fucI	NON_SYNONYMOUS_CODING	F/L	COG2407	G	L-fucose isomerase
C23	snp	2935824	2935824	A	G	fucK	NON_SYNONYMOUS_CODING	D/G	COG1070	G	L-fuculokinase
C23	snp	3080060	3080060	T	C	yggG	SYNONYMOUS_CODING	D	COG0501	O	heat shock protein binding to Era protein; predicted peptidase
C23	indel	3132833	3132833	T	-	yghT	FRAMESHIFT_CODING	-	COG0125	F	predicted ATP-binding protein
C23	snp	3176187	3176187	T	C	tolC	SYNONYMOUS_CODING	S	COG1538	MU	transport channel
C23	snp	3176541	3176541	T	C	tolC	SYNONYMOUS_CODING	Y	COG1538	MU	transport channel

C23	snp	3176564	3176564	A	G	tolC	NON_SYNONYMOUS_CODING	D/G	COG1538	MU	transport channel
C23	snp	3176582	3176582	A	G	tolC	NON_SYNONYMOUS_CODING	Q/R	COG1538	MU	transport channel
C23	snp	3176753	3176753	A	G	tolC	NON_SYNONYMOUS_CODING	Q/R	COG1538	MU	transport channel
C23	snp	3177340	3177340	T	C	tolC	NON_SYNONYMOUS_CODING	Y/H	COG1538	MU	transport channel
C23	snp	3212411	3212411	G	A	rpoD	NON_SYNONYMOUS_CODING	R/H	COG0568	K	RNA polymerase, sigma 70 (sigma D) factor
C23	snp	3218212	3218212	G	A	patA	NON_SYNONYMOUS_CODING	A/T	COG4992	E	putrescine:2-oxoglutaric acid aminotransferase, PLP-dependent
C23	snp	3306890	3306890	G	A	nlpI	SYNONYMOUS_CODING	C	COG4785	R	lipoprotein involved in osmotic sensitivity and filamentation
C23	indel	3328375	3328376	-	T	dacB	FRAMESHIFT_CODING	-	COG2027	M	D-alanyl-D-alanine carboxypeptidase
C23	snp	3357140	3357140	G	A	gltB	NON_SYNONYMOUS_CODING	S/N	COG0069	E	glutamate synthase, large subunit
C23	snp	3364975	3364975	G	A	yhcF	NON_SYNONYMOUS_CODING	G/R	-	-	predicted transcriptional regulator
C23	snp	3389175	3389175	T	C	tldD	SYNONYMOUS_CODING	E	COG0312	R	predicted peptidase
C23	snp	3530385	3530385	T	C	yhgE	NON_SYNONYMOUS_CODING	Y/C	-	-	predicted inner membrane protein
C23	indel	3533587	3533587	C	-	envZ	FRAMESHIFT_CODING	-	COG0642	T	sensory histidine kinase in two-component regulatory system with OmpR
C23	snp	3560885	3560885	A	G	glpD	NON_SYNONYMOUS_CODING	I/V	COG0578	C	sn-glycerol-3-phosphate dehydrogenase, aerobic, FAD/NAD(P)-binding
C23	snp	3896903	3896903	G	A	yieK	NON_SYNONYMOUS_CODING	P/S	COG0363	G	predicted 6-phosphogluconolactonase
C23	snp	4019567	4019567	A	G	ubiB	NON_SYNONYMOUS_CODING	Q/R	COG0661	R	2-octaprenylphenol hydroxylase
C23	indel	4159177	4159178	-	C	fabR	FRAMESHIFT_CODING	-	COG1309	K	DNA-binding transcriptional repressor
C23	snp	4200557	4200557	T	C	zraS	SYNONYMOUS_CODING	I	COG0642	T	sensory histidine kinase in two-component regulatory system with ZraR

C23	snp	4245058	4245058	C	A	malK	STOP_GAINED	Y/*	COG3839	G	fused maltose transport subunit, ATP-binding component of ABC superfamily/regulatory protein
C23	snp	4366062	4366062	C	T	aspA	NON_SYNONYMOUS_CODING	V/I	COG1027	E	aspartate ammonia-lyase
C23	snp	4509897	4509897	T	C	fecD	NON_SYNONYMOUS_CODING	S/G	COG0609	P	KpLE2 phage-like element; iron-dicitrate transporter subunit

Table S9. Total number and type of SNPs observed across the three genomes

SNP: Ref/Var	Fraction of SNPs	Total SNPs	C21	C22	C23	STOP_CODONS
A/G	30.1%	75	26	33	16	0
G/A	25.7%	64	23	20	21	4
C/T	16.5%	41	14	15	12	3
T/C	26.1%	65	22	23	20	0
A/T	0.4%	1	1	0	0	0
G/T	0.8%	2	0	1	1	0
C/A	0.4%	1	0	0	1	1
TRANSITIONS	98.4%	245	85	91	69	7
TRANSVERSIONS	1.6%	4	1	1	2	1

Table S10. Calculation of secondary mutation rates of recoded genomes based on results from Illumina sequencing of three genomes. Three genomes - C21, C22, C23 - have been subjected to ~25-30 rounds of MAGE cycling, dsDNA recombinations to place selectable markers in targeted positions across the genome and two conjugation protocols.

Metric	Average	Total	C21	C22	C23
Secondary mutation rate (INDELS + SNPs + unanno SNPs)	**2.52E-08**	2.52E-08	2.48E-08	2.86E-08	2.23E-08
Secondary mutation rate (INDELs + SNPs)	2.17E-08	2.17E-08	2.19E-08	2.43E-08	1.90E-08
SNP mutation rate	1.85E-08	1.85E-08	1.92E-08	2.05E-08	1.58E-08
INDEL mutation rate	3.20E-09	3.20E-09	2.68E-09	3.79E-09	3.12E-09
Size of E. coli genome (bp)	4.65E+06	*1.39E+07*	4.65E+06	4.65E+06	4.65E+06
All Mutations per population per division (INDELS + SNPs + unanno SNPs)	0.12	0.352	0.115	0.133	0.104
Mutations per population per division (INDELS + SNPs)	0.10	0.303	0.102	0.113	0.088
SNPs per population per division	0.09	0.258	0.089	0.095	0.074
INDELS per population per division	0.01	0.045	0.012	0.018	0.015
Calculated No. of cell divisions	964.0	964	964	964	964
Total Mutations (INDELs, SNPs + unanno SNPs) per genome	113.0	339	111	128	100
Total Mutations (INDELs + SNPs) per genome	97.3	292	98	109	85
SNPs per genome	83.0	249	86	92	71
INDELs per genome	14.3	43	12	17	14
Unannotated	15.7	47	13	19	15

Table S11. Estimates for total number of cell divsions: ~25-30 rounds of MAGE cycling, dsDNA recombinations to place selectable markers in targeted positions across the genome and two conjugation protocols.

	Calculated # cell divisions
Total	**964**
MAGE	**366**
Recombinations	**258**
Conjugation 1	**127**
Intermediate recombination	**86**
Conjugation 2	**127**

MAGE

Step	divisions per	No.	# cell divisions
MAGE cycles	6	25	150
o/n growths	15	6	90
Re-dilution	5	6	30
Colony/Plating	30	2	60
Outgrowth	12	3	36
MAGE total			**366**

Selectable marker dsDNA recombinations

Step	Divsions per X
Colony/plating	30
Outgrowth, mid-log	10
Induce @ 42C, 15 min	0
Electroporation	0
Recover 1 hour	0
Plate o/n	30
Colony outgrowth to mid-log	10
Dilute, re-grow, freeze	6
Divisions/recombination	**86**
No. of recombinations	3
Total	**258**

Conjugations

Step	Divsions per X
Colony/plating	30
Outgrowth, log-stationary	12
Conjugate (2-15 hrs)	10
Positive Selection	15
Dilute 1/100, outgrowth mid-log	6
Dilute 1/100 in ColE1	6
Colony/plating	30
Outgrowth, log-stationary	12
Dilute, re-grow, freeze	6
Total	**127**

Table S12. Sequences of MASC-PCR primers

Gene	MASC-PCR_forward_TAG_primer	MASC-PCR_forward_TAA_primer	MASC-PCR_reverse_primer
rbsB	ATCTGAAACTGGTTGTTAAGCAGTAG	ATCTGAAACTGGTTGTTAAGCAGTAA	CACGCTAGCCCATACACCA
rbsR	GAACGCGGTTCGGCTTAG	GAACGCGGTTCGGCTTAA	TTACTGCTTCTCTATATTCTCCATGCA
ilvL	ACTTGGACGAGGAAAGGCTTAG	ACTTGGACGAGGAAAGGCTTAA	AGCAGGCAACGTCGAATG
ilvA	AGGTTCTTTTTGGCGGGTTAG	AGGTTCTTTTTGGCGGGTTAA	TGGCGTTTTCGATGCTGG
wzzE	CCGCCGTTGCTCGAAATAG	CCGCCGTTGCTCGAAATAA	TGATAAAACGCCGCCAGG
aslB	CATTTGCTGGTGGTGAGTAAGTAG	CATTTGCTGGTGGTGAGTAAGTAA	AAGCGACTCCATCGGCG
hemY	CAGAATAACCCGCCACAGTAG	CAGAATAACCCGCCACAGTAA	TTCGCCTTAGTGCCTCATAAAC
yigG	TGCACCTACTGATTCATCAAAACAATAG	TGCACCTACTGATTCATCAAAACAATAA	TTGAATAGTTAAGCGTTGTCTTACTTTTAATG
recQ	TGATGGCGATGACGAAGAGTAG	TGATGGCGATGACGAAGAGTAA	TCGGGACCGGGGCT
rmuC	GAATATAATCAGCAGTCGCGCTAG	GAATATAATCAGCAGTCGCGCTAA	AAAAATTGTTCCAGAAGTGTAACAGATTG
tatD	CACTGTTTGGGATTGCGTTTTAG	CACTGTTTGGGATTGCGTTTTAA	GAGTAGCCTTTAGGGTTTTCGAAC
yihL	CCGTAGACCTCATCATTAATCTGTAG	CCGTAGACCTCATCATTAATCTGTAA	TACAGGCTTTATGACCCAGGC
yihW	CCCCAGGAAGACGCGTAG	CCCCAGGAAGACGCGTAA	CCGTATCGCTGGGTGAAAAAC
yiiD	GCGGGAACGAAGAAGAGTAG	GCGGGAACGAAGAAGAGTAA	GGGAAAGTTGTTTTTCTAAATTAACAACAGA
cpxP	TAGCAACTCACGTTCCCAGTAG	TAGCAACTCACGTTCCCAGTAA	TTACCGTGACCAAACGAGTGA
frwD	GCAAACCCATCTTATTCTGGAGTAG	GCAAACCCATCTTATTCTGGAGTAA	CGAATCTGTTGCACAGGCG
zraR	TGGCAAAACTGTCGCGTTAG	TGGCAAAACTGTCGCGTTAA	GGCCTGCCGCTGGA
yjaB	GCGTATGTGGGGGCGTAG	GCGTATGTGGGGGCGTAA	CGCATTCTCCATGAAGTGGTC
yjdK	ATGACTATTTCCGTCAGTTTTACAAGTAG	ATGACTATTTCCGTCAGTTTTACAAGTAA	TATGTGTTTTCTCATGTGAGATAAACCAG
blc	GGTACAGCAGCCTGGTAGTTAG	GGTACAGCAGCCTGGTAGTTAA	TGAGTCCGCTAACCCGATG
ulaF	CGCTTATTACGGGCAGAAGTAG	CGCTTATTACGGGCAGAAGTAA	CTGCACCAGTGATGGCAG
yjfY	CAGCGCGGATATTTATCGCTAG	CAGCGCGGATATTTATCGCTAA	TACGAAACCATGAAAGCGGC
priB	CAGATTGAATTGATAGATTCTGGAGACTAG	CAGATTGAATTGATAGATTCTGGAGACTAA	GATGTCGCGAGTACCGATG
ytfB	AGTTTTATTCGTGCGCGGTAG	AGTTTTATTCGTGCGCGGTAA	CTTCCACGCCTTACACCGTA
ytfG	GCGTAAGCCATCTTTTTAATGTTAATAACTAG	GCGTAAGCCATCTTTTTAATGTTAATAACTAA	TTGAATGGCACTGGCAAGG
msrA	GCCGCCGGAAGCATAG	GCCGCCGGAAGCATAA	AAGTTGATTAATTTGTGTGAAATACACATTGA
ytfN	GGATTTGCTCTATCAGTTCGAGTTTTAG	GGATTTGCTCTATCAGTTCGAGTTTTAA	CGCCCCTGGTGCG
b4250	TCGTGATGATCGTCCGGTAG	TCGTGATGATCGTCCGGTAA	GATGCCGCCTGGCG
idnR	ATCACGGCAACACGCTTTAG	ATCACGGCAACACGCTTTAA	AGTAAAAACAGTTCTGTGTCCGG
b4273	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GTCGCTTGTACTGAAAAGTATTCTGTA
yjgX	CGTGTGACAGCCTTAACTCCTAG	CGTGTGACAGCCTTAACTCCTAA	ACACGCCAGTTCGCCC

b4283	TGGAGAATGAAATATTAAAAAGGCTACTGTAG	TGGAGAATGAAATATTAAAAAGGCTACTGTAA	CGAATATCACCTGAGACAATTTATAAAACG
fecE	CCGATGTGCCTAATGAGGTAG	CCGATGTGCCTAATGAGGTAA	GATCCATCAACTGGCCTGTATAAC
nanC	GCATTGGTGTGTCATTTAAACTGTAG	GCATTGGTGTGTCATTTAAACTGTAA	TTGAAGATATTCTGGTTAACACCACC
fimB	ACAGCGTCACGCTGTTTTATAG	ACAGCGTCACGCTGTTTTATAA	GAGTGATGTGGCTATCAGTCAATTAC
yjiN	GTTCCCCCTCGGCAATTTTTAG	GTTCCCCCTCGGCAATTTTTAA	CAACCGCACCGGCTG
yjiP	CTATTCCATGGCGGTGATGTAG	CTATTCCATGGCGGTGATGTAA	GATCAACTCCAACAGGGCG
mcrB	GACCAACAAATTATTAGGGGACTCATAG	GACCAACAAATTATTAGGGGACTCATAA	AGCCCTCGGCGTGAAA
yjjQ	ATTAAGAAGAGAAATGTCGCACTCATAG	ATTAAGAAGAGAAATGTCGCACTCATAA	CGAGATCCGCGACATGG
yjjV	ATACGTTGTTTAACGTGCCGTAG	ATACGTTGTTTAACGTGCCGTAA	GCATGATAAGACGCGTTAAGCG
lplA	GCGGGGGCTGTAAGGTAG	GCGGGGGCTGTAAGGTAA	GATTACTGCACAACACCCTGC
creC	GACTTCACCGTCACTTCACATAG	GACTTCACCGTCACTTCACATAA	CCGGTACCGCTTAAATTCAGG
kefF	CAGGAGGCCCATCATGGATAG	CAGGAGGCCCATCATGGATAA	ACCACCAGCACCAGCG
yabI	GCGTAAAGTGGTTGGGGTTTAG	GCGTAAAGTGGTTGGGGTTTAA	CTATCGCCCGCCAGATGG
thiQ	CACTATTGGGGATTACGGGTTAG	CACTATTGGGGATTACGGGTTAA	ATTCCTGCCGGAATGCAGA
murF	CTTTACAGGAGAATGGGACATGTTAG	CTTTACAGGAGAATGGGACATGTTAA	GACGGGTAAGCCCACAGC
mutT	TTGCGAAGCTTAAACGTCTGTAG	TTGCGAAGCTTAAACGTCTGTAA	CCAAAGCAGTGACATTTGTCTGA
pdhR	CGTCTGGAGCAACGAAAGAATTAG	CGTCTGGAGCAACGAAAGAATTAA	GGAAACGTTCTGACATGGGTTATT
sfsA	AAATCACTGCCGGTTACATTGTAG	AAATCACTGCCGGTTACATTGTAA	AGGCCGCTATAAATAGCAGATG
btuF	AATGCGCTTTCACAGGTAGATTAG	AATGCGCTTTCACAGGTAGATTAA	TACCGCTGTGCCGACTATATC
cdaR	TACAACTGGATGAAGAGCGGTAG	TACAACTGGATGAAGAGCGGTAA	GCCATACTCACGGCGG
yaeI	ACGATGCTGGAACTGGTGTAG	ACGATGCTGGAACTGGTGTAA	GTTGCTGCTGTATGTGGCG
yafL	TTTGACGGAAGAGACGATTTTGTAG	TTTGACGGAAGAGACGATTTTGTAA	ATGGCCAATCACTCACTTGC
yagY	GGTGATAAGTTCATTCCCGTGAAATAG	GGTGATAAGTTCATTCCCGTGAAATAA	GCGCTTCGCGCTTGA
eaeH	AGTGCCTCTTACCCAGCAATAG	AGTGCCTCTTACCCAGCAATAA	CTCAGCCTGATGTTTTTCAATAAAGAC
b0299	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	TATTTTTGGCAACTCATTTAAATATGACATTGA
yahL	ACGTTAAAATCCGCTACACTCTTTAG	ACGTTAAAATCCGCTACACTCTTTAA	AACACATTCACAAAACATCTGCATG
yahM	CGGCAAATCAGGTTGATTACGTAG	CGGCAAATCAGGTTGATTACGTAA	CGCATCAATTATCTGGCGAGTT
prpE	GCCAGGCGATGGAAGAGTAG	GCCAGGCGATGGAAGAGTAA	GTGTCGGATGCGACGC
cynR	GGCGGAAATGAATCACGGTAG	GGCGGAAATGAATCACGGTAA	GCCTACGTAGAGCACTGAACT
frmR	GAACTGGTTCGTGCCTATCTTAAATAG	GAACTGGTTCGTGCCTATCTTAAATAA	GCATTCAAAGGTATGGTCGATACC
b0361	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GCTGCATCGTATACCCACAG
yaiS	CTTGCTGCATTAAAAAAACATAAAAGCTAG	CTTGCTGCATTAAAAAAACATAAAAGCTAA	AGTAGAGGGAATGCCGCC
tra5_1	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	CTGATGATAGCCTCAACGTTGATG
araJ	GCGAAACCACTGGGGTAG	GCGAAACCACTGGGGTAA	GACCGTTTATATCAAACGGTTGGG
pgpA	CCGCTGGGTATTCTGTCGTAG	CCGCTGGGTATTCTGTCGTAA	CGCAGAGCGGTTAACAGG

yajL	CGGCAGGGATTTATAATTATTACGAGTAG	CGGCAGGGATTTATAATTATTACGAGTAA	GGGATTGATGTGGTTTCTCTGC
panE	GCCTCGCCCCTGGTAG	GCCTCGCCCCTGGTAA	ACCAGGCAGCACGATCA
lon	TATGCAGGTTGTGACTGCAAAATAG	TATGCAGGTTGTGACTGCAAAATAA	CCTTACAACACCCCTTAATAAGGG
mdlA	AGGCCGTCGATGCGTAG	AGGCCGTCGATGCGTAA	TCAGGACCGCAATCCCC
ybaA	GATTCGAGTCAATCATCGACGAATAG	GATTCGAGTCAATCATCGACGAATAA	GCAGCGATGTTTGTGCTG
ybaJ	GAGAATCCTGCGAGTTTATCTTGTTAG	GAGAATCCTGCGAGTTTATCTTGTTAA	GCAATCGAAGTTGGCGAAAAAAT
priC	GCGCGTTTAACCCGCTAG	GCGCGTTTAACCCGCTAA	GTATGTGCGTGAACTGCGT
aes	CAGTTCTTTACCGCTCAGCTTTAG	CAGTTCTTTACCGCTCAGCTTTAA	TACGCCGATAACCACGATTACA
ybcF	GGGACCTGTATTTCGCTGTAG	GGGACCTGTATTTCGCTGTAA	TGAAACTGCCGCTGGTG
tra5_2	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	CTACAGTTTGAACTCAACTCACCTG
ybcL	CTGAGATAACACCAGTTTATGAGATAAAGTAG	CTGAGATAACACCAGTTTATGAGATAAAGTAA	TAATTCTCAATATAAAATTCATAACCACAAGGAAATAATAT
rzpD	CAGAAGTATATTAATGAGCAGTGCAGATAG	CAGAAGTATATTAATGAGCAGTGCAGATAA	CGATGCAGCTAAAATTTGTGGC
ybcY	TAATGTTTTCCGCTTCAGGAAAAAAATAG	TAATGTTTTCCGCTTCAGGAAAAAAATAA	GCTGAACCGTGTTGATACATCAA
hokE	CGCTTACGAACCGAAGAAGTAG	CGCTTACGAACCGAAGAAGTAA	ATCGGCACGTTAAATCATATCAGG
ybdL	GCCTGCGCCAGCTTTAG	GCCTGCGCCAGCTTTAA	TATTTGCACTGGTATCAGAAAAAAGGC
ybdM	GCGCCTGGACGGTTAAATAG	GCGCCTGGACGGTTAAATAA	GCGAGCGCACCAGAC
rlpA	GTCATTTATTACTACCGCGCAGTAG	GTCATTTATTACTACCGCGCAGTAA	GCGACACAAGCATCGTTACC
ybeQ	AAATTCGTGAGACTGAAGATTTACTGTAG	AAATTCGTGAGACTGAAGATTTACTGTAA	GCACTGGTGGAACAGGC
ubiF	GCTGAAATATGCGTTAGGGTTGTAG	GCTGAAATATGCGTTAGGGTTGTAA	ATCGAATATGTTCGGCAAATTCAAAAC
ybfH	GAGGTCTGCTGATTGATAATGGTTAG	GAGGTCTGCTGATTGATAATGGTTAA	GTACATTTTCATAGAAGGTATCTAAATCACCG
ybfD	GATGAGCTTATTGATTTCACGTTTGAATAG	GATGAGCTTATTGATTTCACGTTTGAATAA	CTTTCAGTCGATGTAACTCATGAAGG
nei	CCTGGCTGCCAGCACTAG	CCTGGCTGCCAGCACTAA	TTTATGACTGCCTACCTCGCC
abrB	CGAAGCGGTCGGCCTAG	CGAAGCGGTCGGCCTAA	ATAAGCATCATGGGGCGCT
sucB	TCTGCTGCTGGACGTGTAG	TCTGCTGCTGGACGTGTAA	GTCAGCGTCGCATCAGG
bioD	AGTACATAAACCTTGCCTTGTTGTAG	AGTACATAAACCTTGCCTTGTTGTAA	CGGGAAGAACTCTTTCATTTCGC
moaE	GCGGCAAAACGCTGGTAG	GCGGCAAAACGCTGGTAA	CGTCAATCCGACAAAGACAATCA
ybhM	AAGGCATACGCGTTTTTTTCATTAG	AAGGCATACGCGTTTTTTTCATTAA	TTACTGGCAGGGATTATCTTTACCG
ybhS	CCAAACGTCGGCTGGATTAG	CCAAACGTCGGCTGGATTAA	CTGTTGTTAGGTTTCGGTTTTCCT
ybiH	AAGGATTATCGCAAAGGAGTTTGTAG	AAGGATTATCGCAAAGGAGTTTGTAA	GTCATAGGCGGCTTGCG
ybiR	TTAGTTATACTCCCGGCCAACTAG	TTAGTTATACTCCCGGCCAACTAA	ATGAGCCGGTAAAAGCGAC
yliD	CGCTGGATCCGAAAATTAAAGGATAG	CGCTGGATCCGAAAATTAAAGGATAA	AATAAAATTATCAGCCTTATCTTTATCTTTTCGTATAAA
yliE	TGGGATAAAAGTGGAAAATTAGTAAAAGAGTAG	TGGGATAAAAGTGGAAAATTAGTAAAAGAGTAA	CAGCAATATTTGCCACCGCA
ybjK	TTGAGAGGGTTGCAGGGTAG	TTGAGAGGGTTGCAGGGTAA	AACTTTTCCGCAGGGCATC
rimK	GCCTGAAAACGGGTGGTTAG	GCCTGAAAACGGGTGGTTAA	TACAACCTCTTTCGATAAAAAGACCG
ybjR	TGGAGAAATACGGGCAGGATTAG	TGGAGAAATACGGGCAGGATTAA	GTGATGCTCGCGGGC

cspD	GAAGCGGCAGTCGCATAG	GAAGCGGCAGTCGCATAA	CAGTAAAGTGCTGCGTGTCTG
lolA	TCACGGTAGATGATCAACGTAAGTAG	TCACGGTAGATGATCAACGTAAGTAA	CTGACCACCATAGCCACGG
ycaI	TGCCAGTGGATAACGGGTAG	TGCCAGTGGATAACGGGTAA	ACGCACAACAGTAATCAGTGC
lpxK	GCTGGCTTCTGGCAACTAG	GCTGGCTTCTGGCAACTAA	TTTGTATTTCCAGCCCATTTTTTCAG
ycbK	CAGCACGGCACTGGTAG	CAGCACGGCACTGGTAA	CAGAACTCATCTTCTTTTTCCGGG
ycbX	GCACTTAAGTTGGCGCGTTAG	GCACTTAAGTTGGCGCGTTAA	TGTGGGTGCCCATGCC
tra5_3	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	TGTGACTCAGGAAAATGCGTTAAA
yceO	GCAATTATGGATCTGATCTATGTGTTTTATTTTTAG	GCAATTATGGATCTGATCTATGTGTTTTATTTTTAA	TTCAGCATCAGAGATGTCTGGT
yceF	GAAAAACCCGCTGATGGGATAG	GAAAAACCCGCTGATGGGATAA	CCTGATAAGGCGTTTACGCC
plsX	AGCGGAACTCTGCGGTAG	AGCGGAACTCTGCGGTAA	CGCCGTTGCACTGATAATACG
fabH	CGCGCTGGTTCGTTTCTAG	CGCGCTGGTTCGTTTCTAA	GTTAAAGGTGATTTTCGCTAATTCTACTG
pabC	TTTGTGAGCGCCCGAATTAG	TTTGTGAGCGCCCGAATTAA	CAGGCGCTATCGACCG
b1146	AGACCGGATATCTATCCGAAAGATTAG	AGACCGGATATCTATCCGAAAGATTAA	TCGAAATTGAGAGATACAGCTCGT
ymfM	AGTTTCGTCGTAAATTCAGGGAAATAG	AGTTTCGTCGTAAATTCAGGGAAATAA	CGGCGGTGGTAATGGC
ymfR	TCTGTCGGCGGAGGTAAATAG	TCTGTCGGCGGAGGTAAATAA	TATAGCCATTGGGATGCGTGG
tfaE	AACTGCCGGAGGCGTAG	AACTGCCGGAGGCGTAA	TCCTTCTATTAAAGAGGTTATCCGTACTC
ycgX	CTTTTATATATTGAAGATGTAGAAAATAATAAACCATTAATCTAG	CTTTTATATATTGAAGATGTAGAAAATAATAAACCATTAATCTAA	CAAATATCAAGTTGATAGTATTAGTCTGGTGAT
ymgC	AGGGAATCCGTGCTCTCTTAG	AGGGAATCCGTGCTCTCTTAA	GAGAGCGAGAACTCTCCATTCT
ycgY	GGACTTGAAAAACCGTCTGTGTAG	GGACTTGAAAAACCGTCTGTGTAA	AGCGCTTACTCCGACAACG
hemA	GCCTCGGGCTGGAGTAG	GCCTCGGGCTGGAGTAA	ATGTTGTGAACGTTCGTCCTG
b1228	GCTCGAGCCGAACCTTAG	GCTCGAGCCGAACCTTAA	CGCTTCCCGATAAGGGAGC
kch	CAGTAAAGAATCGGCGCAAAAATAG	CAGTAAAGAATCGGCGCAAAAATAA	CAGCGCGGCGTTTGC
sapB	CATAAGGAATGGTATGCCTTACGATAG	CATAAGGAATGGTATGCCTTACGATAA	ACATAGCGTGGGACAAACTC
pspF	CGTTGTTGAAAAAGCACCAGATTTAG	CGTTGTTGAAAAAGCACCAGATTTAA	AAAGAAAGCGTTCCGGCC
ycjU	GGCCTTCTGGCAAAACGTATAG	GGCCTTCTGGCAAAACGTATAA	GATACAACGCGTAGTTCTGGAAC
ycjV	GTTGGGGGCACGAGTTAG	GTTGGGGGCACGAGTTAA	GTTCCTTTCCTCGGCCTGT
ycjG	ACGGGCGAATTGCATCTTTAG	ACGGGCGAATTGCATCTTTAA	GGCTCAGGCGTTTGAATTG
ycjZ	GGCTGAAGTATAAAGGTGCTGTTTAG	GGCTGAAGTATAAAGGTGCTGTTTAA	CCCGTTTTCGGTATCATAATCAGC
intR	TTGAATGCGCGGTTATCGTAG	TTGAATGCGCGGTTATCGTAA	TCTTTGTAATTTGTTGATTTTCTTGCATTATTTC
sieB	GCAGGAGTAGACGTTTTAATCGTTAG	GCAGGAGTAGACGTTTTAATCGTTAA	AAAAGTGAACTTTTAGATATAAACAACCTTTATCAA
b1354	GAATCTCACATCTTGTTGCTACGTATAG	GAATCTCACATCTTGTTGCTACGTATAA	TACCCCATACCAACCCCCA
ydaY	AGCAGGATACGATAAAAAAGCCATAG	AGCAGGATACGATAAAAAAGCCATAA	GCCTGCGTCATTTTGACTAATGA
b1367	GTATTTTCAGGAGCAGTGTGTGTAG	GTATTTTCAGGAGCAGTGTGTGTAA	TTTCATTAAGTTTACGGACAAGAAATTCATATTC
tfaR	GCCAGCACTGCCGTAG	GCCAGCACTGCCGTAA	GCATTGACCTGACCAGTCCA
paaG	CGCAGTTCACGGGGAAATAG	CGCAGTTCACGGGGAAATAA	AGACGCCGCTTCAATGAC

paaX	AGGCGTTATGCCAATTTATCAGATAG	AGGCGTTATGCCAATTTATCAGATAA	ACCAATATGTCCATCTTCTCCTACAA
b1402	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GTATATCTGAATGGAGATACGACAATCAG
ydcL	CACCATGTTTGACGTTAACAAAAAGTAG	CACCATGTTTGACGTTAACAAAAAGTAA	TACATTGAACAACAGCAAACACCT
rhsE	AGGGAGTTCAGATGCTTCAAATTATTAG	AGGGAGTTCAGATGCTTCAAATTATTAA	TTTGAATTGTTATCAAACCCAAGGTAAAC
yddJ	CCTCCAGTTCATGGCCTTTTTAG	CCTCCAGTTCATGGCCTTTTTAA	ATTAATCCCAAGAGCGCTCCC
ddpF	TGGCGGGCGCTGTAG	TGGCGGGCGCTGTAA	GGGATTGTGATTGGTATGATCCG
yddV	GAACTCTGGAAAGCCAGTCTTTAG	GAACTCTGGAAAGCCAGTCTTTAA	GAGTGTATCGGGCTGCATAC
ydeV	GGAAAGCGCCTGGGTTATAG	GGAAAGCGCCTGGGTTATAA	GACATTAAAAATGAAACTTATTAAATTGTCAGAGG
ydeA	CACTCGAAGAACAGACGCAATAG	CACTCGAAGAACAGACGCAATAA	AAAGCAAGTCGGAAGAGCTG
marA	TTTTTACATCCATTAAATCATTACAACAGCTAG	TTTTTACATCCATTAAATCATTACAACAGCTAA	GTCGGGGCCAGAACAAAC
marB	GCCCTATTATAATCAACACGCTATGTAG	GCCCTATTATAATCAACACGCTATGTAA	GGACGCTATGAAACCTGGC
tfaQ	GCCAGCACTGCCGTAG	GCCAGCACTGCCGTAA	AGTATAAGTGCCATTGCCCG
ydfP	TGAATCTTTAAGAGATAAAAAGCATCAAATGTAG	TGAATCTTTAAGAGATAAAAAGCATCAAATGTAA	TTCTGATTAAGACCTTTTATCTTACCCTTAAG
rzpQ	CAGAATGATTACATACCCGCGTTAG	CAGAATGATTACATACCCGCGTTAA	CGTGTCCGGCGCAC
b1578	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	AACGGTCTTTTGCTAAATTTCTCCT
ynfA	TGGGGGCGCACGTAG	TGGGGGCGCACGTAA	CTGTGACCGGTGGTATGGC
speG	GAAGCCGACCGCACAATAG	GAAGCCGACCGCACAATAA	AATTACTGGATCTGGAAAATAATACCGTG
dmsD	CGGTTAAACCGCTGTTTCGATAG	CGGTTAAACCGCTGTTTCGATAA	TCCGCCGAGCGCC
ydhM	CGACGCCTGCCGTTTAG	CGACGCCTGCCGTTTAA	ATGCCGCGTTTACCCAGT
lhr	AAGGGCTGGATTGGGGATAG	AAGGGCTGGATTGGGGATAA	AACTTGATGATTCTAATGCATAGAAGAAGAG
ynhG	GTTACGCGTGACGCAGTAG	GTTACGCGTGACGCAGTAA	GAAGCGCCCTGATGGG
sufA	GGCGAAAGCTTTGGGGTATAG	GGCGAAAGCTTTGGGGTATAA	TTCGCGATAAGTAGTGGCAAC
ydiA	CCTTAGTCGCCGAATGTACTAG	CCTTAGTCGCCGAATGTACTAA	ACGGGTGGCGTACTCC
ydiE	CAAGCTGCTGTTGACCAAGTAG	CAAGCTGCTGTTGACCAAGTAA	CATTTGATGACTGGTTTGCCC
nlpC	GGCAGGCGAGACGAATCTAG	GGCAGGCGAGACGAATCTAA	GAAGGATGTTTTTCAGTATCCGGC
btuC	TGGTTATTGTTAAAAGCAGGACGTTAG	TGGTTATTGTTAAAAGCAGGACGTTAA	ACACTTTGAGGCGACATTGAC
arpB_1	CACCCAACAGTAAATCTCAATGGTAG	CACCCAACAGTAAATCTCAATGGTAA	AACGGAGGAGTAATGTTTAATCAGGT
ydjX	ACTTACTCACCCAAAAAATGAAGGATAG	ACTTACTCACCCAAAAAATGAAGGATAA	CCAGCAACAGCCAGTAAACA
nudG	GCCAGACCAGCGGATTAG	GCCAGACCAGCGGATTAA	GTCAGGGCATTTTTGAAGGTTTG
b1788	ATCGAACGGACCGGCTAG	ATCGAACGGACCGGCTAA	TATCGCGCCGCAGAAAGA
yeaL	GCTGATTGTGGGGAAACAGTAG	GCTGATTGTGGGGAAACAGTAA	GGTTTGGATACCGGGCG
yeaX	CGCCTGGTGTTGGATTTGTAG	CGCCTGGTGTTGGATTTGTAA	GTGAACTTTGTCGTGCGTGA
yobB	GGCGATATCATTCCATTACGCTAG	GGCGATATCATTCCATTACGCTAA	TATAACGCCATATTGCTGTACTTGATC
exoX	TTTAACACTGAAACATTATCTGGAAAATACTTAG	TTTAACACTGAAACATTATCTGGAAAATACTTAA	TATGACAACGTCACCGCACA
yecN	GTTGGTTTTCTCCCTGCGTTAG	GTTGGTTTTCTCCCTGCGTTAA	AGATCGTAAACCTGCGTACCA

otsA	ACCTTTCCAAAGCTTGCGTAG	ACCTTTCCAAAGCTTGCGTAA	CCTACGGTGAGTTAAGCGAAC
dcyD	GCCTATCATCCCCACGTTTAG	GCCTATCATCCCCACGTTTAA	TAAACCCGAGCAGTAAGCCA
yedM	CCCACCCATTCGGCATAAATTAG	CCCACCCATTCGGCATAAATTAA	GAAACCTTTGCAAAGTCTGGAATC
fliE	GTGATGAGCATGCAGGTGTAG	GTGATGAGCATGCAGGTGTAA	TTTTTTTTGAAGTCATTCGATGCGC
fliN	GCCGCCTGAGCCGTTAG	GCCGCCTGAGCCGTTAA	CGATCACCGGTGACATAATAAAAAAG
fliP	GCGCAGAGCTTTTACAGCTAG	GCGCAGAGCTTTTACAGCTAA	GCGGTTCGCACAGCG
fliQ	CCTGCCGTATATCATCGGGTAG	CCTGCCGTATATCATCGGGTAA	CAGGGTGTGAAAGGTATCGAC
yedS_1	CGTAACAATGGCTTCTTTGGTTAG	CGTAACAATGGCTTCTTTGGTTAA	TCACCATTATCAAACTCGCCAC
cbl	GCAGTGAAGAGGAAATTGATTATCAGATATAG	GCAGTGAAGAGGAAATTGATTATCAGATATAA	ATAATGAGGGAGAATCAGCAAAGC
b1996	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	AAAATCAGCAACGAAGAGACGC
hisL	CCATCATCACCATCATCCTGACTAG	CCATCATCACCATCATCCTGACTAA	AGCTATGCGTAAACGAGTGTTG
wbbJ	TGAGACCAAATTATGGGAAAAAGCATAG	TGAGACCAAATTATGGGAAAAAGCATAA	CTCTTTAGCAGAATGGACTAATGCG
wcaM	GCCGGAACTGAGGGAGTAG	GCCGGAACTGAGGGAGTAA	GTGATGTTACGCCGCTTCG
wcaL	AGCTTGCTGCAGGCTTTATAG	AGCTTGCTGCAGGCTTTATAA	ACGAATAACAAAATGTTTACCATTTTCAACA
wcaC	GAGTATGTCAACTTCTATCAGAATCTGTAG	GAGTATGTCAACTTCTATCAGAATCTGTAA	TTGAAATGCCCACTCCAGC
asmA	GAAGAAGTTGCTGGAGAAGATGTAG	GAAGAAGTTGCTGGAGAAGATGTAA	AAGTGGAAGAGATCGACGCC
baeR	GCCTGCCGCATCGTTTAG	GCCTGCCGCATCGTTTAA	TAACGGAGGCAATTCCGGT
gatR_1	CGCGGGTAAAATTAATTACGGTGTAG	CGCGGGTAAAATTAATTACGGTGTAA	AGGCCGGTGGCGG
tra5_4	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	CGTGCTTGTTCAGGATCTGG
yegV	TCCTCGCACACAAAAACGTATAG	TCCTCGCACACAAAAACGTATAA	TAATTCATGATGTTGATGCCATCGG
yegW	CGTTTTTGTGTGCGAGGAGTAG	CGTTTTTGTGTGCGAGGAGTAA	TCCGGCATTTCCAACGC
yehQ	AGAGTGCTATTGGTCAGGCATAG	AGAGTGCTATTGGTCAGGCATAA	ATTGTTCGCAAAGCAAAACACC
yohC	GGCTACGGTTATCGTTTGTTCTAG	GGCTACGGTTATCGTTTGTTCTAA	CGATCGTCAAAGGAGAATCGTG
yohF	GAATCCACAGTTCAACCCAGAATAG	GAATCCACAGTTCAACCCAGAATAA	TTGTGATTATTTTTTGTTCGCCCATG
sanA	TTACTTGAATTACAAAAGAAACAAGGAAAGTAG	TTACTTGAATTACAAAAGAAACAAGGAAAGTAA	GGCACACTCTGGCACAAAC
yejA	GCCAGCAAACAGGGAGAGTAG	GCCAGCAAACAGGGAGAGTAA	GAAGAAGACAATCAGCAGGATGG
yejE	CATTTGATCCTAATAAGGCGGTGTAG	CATTTGATCCTAATAAGGCGGTGTAA	GAGAGCTGATGCGGATAATCTG
b2191	GTAAATGCCTTGAATCAGCCTATTTAG	GTAAATGCCTTGAATCAGCCTATTTAA	TTTGTCTCAGCCGATGCC
rcsD	CAGTTATGTCAAGAGCTTGCTGTAG	CAGTTATGTCAAGAGCTTGCTGTAA	TCAGAACAATGATCGACAGGCT
rcsC	GGAAATCGCGGGATTCGTAG	GGAAATCGCGGGATTCGTAA	AGTGACCTTAAGTCCGGCAG
yfaT	CGTTTAAATTTTCTGGCGCGATAG	CGTTTAAATTTTCTGGCGCGATAA	CCGATAATCACGCCCCG
menF	GCTGCGTACTTTATTACAAATGGAATAG	GCTGCGTACTTTATTACAAATGGAATAA	CCGCCCAGCGTCGG
yfcO	GTTCCAGTACATACTGAATGAGTTGTAG	GTTCCAGTACATACTGAATGAGTTGTAA	GAGGAGTTGAGCGCGTTG
yfcU	GGGCGGATAAGGGCATGTAG	GGGCGGATAAGGGCATGTAA	TCATCAGGAACAACCGGCG
tfaS	CTGAACTGCCGGAGGAGTAG	CTGAACTGCCGGAGGAGTAA	CCGGTTTAATTATTCAATGGGGGC

ypdI	CGAATTATCTTATTCATCCTGTGCGATAG	CGAATTATCTTATTCATCCTGTGCGATAA	TAAAGAGTGTAAAGACTAGTCAGGACATAA
yfdY	GGGTATGGCAGGCGGTAG	GGGTATGGCAGGCGGTAA	CGCAGAGAACTATTCTTATATACCACAAA
yfeO	CCACCCCACGATAACGTTTAG	CCACCCCACGATAACGTTTAA	CACGATAGCCACGCTCATAC
mntH	ACGGCGCTGGGATTGTAG	ACGGCGCTGGGATTGTAA	CAAAAACCGTTGGTGGCGT
xapR	TTGGCTGCATTGACACAGTAG	TTGGCTGCATTGACACAGTAA	TCTGCGATACTCAAAGATTCGAAC
yfeR	AAGGGGATCCGTTGTATCAGATATAG	AAGGGGATCCGTTGTATCAGATATAA	GAATGGCGCAGTTGCACT
yffB	TATCAGCAATTTTTCCATGAGGTGTAG	TATCAGCAATTTTTCCATGAGGTGTAA	AACGGTAAAACCGATCGCC
hda	CGTTTGTGAAAGAAATTCTGAAGTTGTAG	CGTTTGTGAAAGAAATTCTGAAGTTGTAA	CGAATACTATGGATTAACTTCGCGTAG
yfgG	CGAATCACCGGTACAACGTTAG	CGAATCACCGGTACAACGTTAA	TAGCTCTTTTCAGGACATTTGGTATTT
pbpC	CGACAGTGAAATTTGTCATGCAATAG	CGACAGTGAAATTTGTCATGCAATAA	AACGCCCCGGTGAGC
yphA	CAATTTCTCTCGATCGGCGTTAG	CAATTTCTCTCGATCGGCGTTAA	GGGTAAGTCAGGCGTTTCATTTAA
yfhB	GTGAACTCCAGCAACTGGAATAG	GTGAACTCCAGCAACTGGAATAA	CTCATCCGCCAGTATTCCTTC
kgtP	GGAAGGGGATGCGTCTTTAG	GGAAGGGGATGCGTCTTTAA	AGCCGCATCCTGTAATCACA
yfiA	CGTCGAAGAAGTTGAAGAAGAGTAG	CGTCGAAGAAGTTGAAGAAGAGTAA	GCCAGTAACGGGTTTTCCG
yfjQ	GCTCACGCAACTGCAGTAG	GCTCACGCAACTGCAGTAA	TTCTGCCTGTGTCATGAGAGAT
yfjR	CGACAGCCCACCGGAATAG	CGACAGCCCACCGGAATAA	ACTGGCAGCACCGGC
ypjC	GTTTTTTGAATGCGCTAAAAGAAATGAATAG	GTTTTTTGAATGCGCTAAAAGAAATGAATAA	TATGAATTACGTGATTCCATACAGACATTT
ygaQ	CAGACAGGATAGGAGAAAGAAAAACATAG	CAGACAGGATAGGAGAAAGAAAAACATAA	TGTTCAGCACGACCATACTGATA
ygaR	AAGGTGGTATCCCTGGCTATTAG	AAGGTGGTATCCCTGGCTATTAA	TAGGTAGAGCAACCTTTATTAAGCTACG
yqaC	CGGCGGTGAGGCGTAG	CGGCGGTGAGGCGTAA	TAAAAATATCTACATTTCTGAAAAATGCGCA
gabT	TTTTGATGAGGCGAAGCAGTAG	TTTTGATGAGGCGAAGCAGTAA	GCGGCGATGTTGGCTT
ygaU	GTTGCGTATTCCGGAAGAGTAG	GTTGCGTATTCCGGAAGAGTAA	AGGGTATCGGGTGGCG
ygaM	GTTAAGCATGCGCAAATCGTAG	GTTAAGCATGCGCAAATCGTAA	CGCAACGCTTCTGCCG
luxS	GTTGCAGGAACTGCACATCTAG	GTTGCAGGAACTGCACATCTAA	ATGCCCAGGCGATGTACA
mltB	GCTGGCGCGAGTACAGTAG	GCTGGCGCGAGTACAGTAA	AGACTCGGCAGTTGTTACGG
srlE	GGTTTGTCTCCGGTTTTATCTATCAATAG	GGTTTGTCTCCGGTTTTATCTATCAATAA	GGATGGAGTGCACCTTTCAAC
norW	GATTGTTGAAAACATTGCCGATGTAG	GATTGTTGAAAACATTGCCGATGTAA	GTGTTGCATTTGGACACCATTG
ascB	CCAGTAATGGGGAAGATTTAGAGTAG	CCAGTAATGGGGAAGATTTAGAGTAA	CGCTTATCGGGCCTTCATG
hycI	AGTTAGCGGTGGAAGAAGAGTAG	AGTTAGCGGTGGAAGAAGAGTAA	TGGGCGGCGCTGG
hypB	GAGACACAGCGATGTGCATAG	GAGACACAGCGATGTGCATAA	TTTGCGGGTTGGATTCTCCT
ygbA	TATCGTCCGAAAAAACCTCATGAATAG	TATCGTCCGAAAAAACCTCATGAATAA	GTGCGTCTAATCGCCGC
pphB	GGCGGCTGTCATTTTATAAAATAAAGTAG	GGCGGCTGTCATTTTATAAAATAAAGTAA	ACTGACTGTGGTCACAAACGA
ygcN	AGGAGTCTGCGATGTCTGTAG	AGGAGTCTGCGATGTCTGTAA	CGTACAACACGGAAATAAATTGGC
mazF	GCCAAAATTAACGTACTGATTGGGTAG	GCCAAAATTAACGTACTGATTGGGTAA	AACACCACCTGGAATAGCAGAT
relA	TCGGTTGCACGGGAGTTAG	TCGGTTGCACGGGAGTTAA	CGGGCTCTTTACGCACTG

rumA	GTACTTTTCTCGCGCGTTAAATAG	GTACTTTTCTCGCGCGTTAAATAA	CAGACAATACGCCCAGGTTTC
ppdC	CATTGCCCGAATCGTCAGTAG	CATTGCCCGAATCGTCAGTAA	ACTTTGCACCAGAATCATCTCTG
ptsP	CTGATTCGCGGAGGGTTATAG	CTGATTCGCGGAGGGTTATAA	GCTCCAAAGGTGCGAATAATAGC
mutH	CCCGTCATTTTCTGATCCAGTAG	CCCGTCATTTTCTGATCCAGTAA	GTAAAAGGAGTGAGTACGCCCT
ygeA	TGTCGCTTTTATGCTGTCGTAG	TGTCGCTTTTATGCTGTCGTAA	TCAGCGCCGCAGAAAGT
kduI	GCCGTTAAAGATTTGCGCTAG	GCCGTTAAAGATTTGCGCTAA	TAACATTAATGTTGTGTTTAGCCCATTC
yqeF	TTGACCATTGAACGTGACGAATAG	TTGACCATTGAACGTGACGAATAA	CAACTTCCCCGCCAACG
ygeI	CTATCTCACCTTACTATAAAATTGAGTGGTAG	CTATCTCACCTTACTATAAAATTGAGTGGTAA	GTTCATCTTTTATTATTCCACGTTTTTTTAACAG
pbl	CAAAATATATATTGTATATACCAGGCTTAATGAACTAG	CAAAATATATATTGTATATACCAGGCTTAATGAACTAA	CTTCCAGGTATTCCTCTGATAATCAAC
ygeM	GGCAAGCCCTGAGGACTTATAG	GGCAAGCCCTGAGGACTTATAA	TTCCCATATATTCATCCTGAGTGTAAATG
b2860	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	ATCGTGCCTTTTCATTTATTTCCAG
ygeP	ACACCAGAAGAAAAGAGCAAGTTTAG	ACACCAGAAGAAAAGAGCAAGTTTAA	CCGTCTGCGTTTCTCCG
yqeC	GCGCGCAGATTTGTGAATTAG	GCGCGCAGATTTGTGAATTAA	TCGAGTTTTGCAGCCTCTGT
ygfJ	AGCGAAGGAAAGGTATACTGAAATTTAG	AGCGAAGGAAAGGTATACTGAAATTTAA	GATGAACTGTTGCATCGGCG
recJ	TCATCGACAATATCTGGCCAATTTAG	TCATCGACAATATCTGGCCAATTTAA	CTGTACGCAGCCAGCC
argO	TGCACAAGCCTTGTTCAGTTAG	TGCACAAGCCTTGTTCAGTTAA	AATCGCTGCCTTACGCG
yggU	CAGAAATCGCGGCGTTAATTAATTAG	CAGAAATCGCGGCGTTAATTAATTAA	TAACCAAAGCCACCAGTGC
mutY	GGCGCGCCGGTTTAG	GGCGCGCCGGTTTAA	CGCGAGATATTTTTTCATCATTCCG
glcC	GCTGGAGATGAACCTGAGTTAG	GCTGGAGATGAACCTGAGTTAA	GGGCAAAATTGCTGTGGC
yghQ	CTCGAAGCGTATGACCTGATTTAG	CTCGAAGCGTATGACCTGATTTAA	ACCAACTGGCGATGTTATTCAC
yghT	CGCGCGTGGTTATCACTAG	CGCGCGTGGTTATCACTAA	GACGATGGTGGTGGACGG
ygiZ	TGGGGACGGTTTATATTTTGCTATTAG	TGGGGACGGTTTATATTTTGCTATTAA	ATCGCCAAATTGCATGGCA
yqiB	CGATGGCGATTCCGGTTTATTAG	CGATGGCGATTCCGGTTTATTAA	AAAATCCTGACTCTGGCCTCA
b3045	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	CGTTTGTGCGACACCCG
yqiI	GACGTAATCCTGAAAATCACCTATAACTAG	GACGTAATCCTGAAAATCACCTATAACTAA	ATGACCTGTGGTTATAGAATCATAATGG
ygiP	GATGAAGGCTATCAGGTCATGTAG	GATGAAGGCTATCAGGTCATGTAA	TCCAGTGCCATTCTGGTTTG
yhaK	CTTTGCTGATAGATTTGCCTGTCTAG	CTTTGCTGATAGATTTGCCTGTCTAA	AATGGCGCTGGATGATACC
tdcA	ACGAAGGCAATTAATAGAAGTTGGTTAG	ACGAAGGCAATTAATAGAAGTTGGTTAA	GGGAAACCCCTTGCGC
yhaC	CGTATGATTACGACCTTTTAAATGCAATATAG	CGTATGATTACGACCTTTTAAATGCAATATAA	CGTAATAAACCGTTCCGGCG
ecfH	GTAACTACGGCGTTTACGTTTATTAAATAG	GTAACTACGGCGTTTACGTTTATTAAATAA	CACTGATTATGCTGCGCAAAG
yhbW	GATGTTAAGGAAGAGTTGTTGGGATAG	GATGTTAAGGAAGAGTTGTTGGGATAA	GCAACGCACTGGTGCATAT
nlpI	GCAGAATCGGACCAGCAATAG	GCAGAATCGGACCAGCAATAA	TGAAATTGATCAGCAATTTTCATTGAAAAG
dacB	GCCGTTTGTATAAAGATATTTATCAGAACAATTAG	GCCGTTTGTATAAAGATATTTATCAGAACAATTAA	GCATAAACAAAGCGCACTTTGT
ptsN	GGTACTCCGGATGAAGCGTAG	GGTACTCCGGATGAAGCGTAA	CCAGTTGCTCTGCAATATACACC
yhcC	GGCTACCTCCAACGGAGTAG	GGCTACCTCCAACGGAGTAA	CAGTTGTTTTTCGCTTTTTCCGG

yhcE	GTAAATGCCTTGAATCAGCCTATTTAG	GTAAATGCCTTGAATCAGCCTATTTAA	CACCCAGCTCGACCAAGA
yhcO	CATTTGCGTTTTAATGTTCGTCATTAG	CATTTGCGTTTTAATGTTCGTCATTAA	CTTTTTTGTTATTAAAATTCACATTTTTAACACTTAGTATC
mreC	CGCCGGGAGGGCAATAG	CGCCGGGAGGGCAATAA	GACGTTAATCACTAAAAACTCTGCCC
yhdN	GTTGTTGGACAAAATCAACGATAACTAG	GTTGTTGGACAAAATCAACGATAACTAA	CCTGCACAATGCCTTTTGACT
kefB	CGGCTGGGATGAATTTGAGTAG	CGGCTGGGATGAATTTGAGTAA	AAAACTCGGCGCATATCGC
frlD	ACCACGGTGCCTGGTAG	ACCACGGTGCCTGGTAA	AACCCCCTGCGCGATAT
bioH	CGTTGAAGCAGAGGGTGTAG	CGTTGAAGCAGAGGGTGTAA	CTGCATATAGCCAAAATAATGCTTGAATC
gntX	GTGCCTTTGTCGAACCTTGTAG	GTGCCTTTGTCGAACCTTGTAA	GCATCGGAAATACGGATCATAGTAAT
malQ	GCTGCAGCGAAGAAGAAGTAG	GCTGCAGCGAAGAAGAAGTAA	TCGATAAGCCAATTCCGGAAG
glgA	CGTGAGCTTTACTATCGCTTGAAATAG	CGTGAGCTTTACTATCGCTTGAAATAA	TCTGATCGTAAGCGACTCCC
yrhA	TGATGCTGCCAACTTACTGATTTAG	TGATGCTGCCAACTTACTGATTTAA	CTGATTACGTGCAGCTTTCCC
ugpQ	CCGAACTTTACGGCCCAATAG	CCGAACTTTACGGCCCAATAA	TTTTGTCGGGAAGCGAGTAAAC
gadE	CACCCTGGGTATCACATCTTATTTTTAG	CACCCTGGGTATCACATCTTATTTTTAA	ACAACGGTATTAACAGCTTTCTTCTT
gadX	CGTTCTGCGGAAGGAATAAGATTATAG	CGTTCTGCGGAAGGAATAAGATTATAA	TGGTAGGACCAAAATTTATCCACAAAG
yhjK	GAAGAGAGTTACCTGGAAGAAAAGTAG	GAAGAGAGTTACCTGGAAGAAAAGTAA	AGGACCTGAAAGTAAAGGCTTTTAAA
yhjR	CAGAGTTTGCGCAACAAAAGTAG	CAGAGTTTGCGCAACAAAAGTAA	GACGGTGGGTAAAATCAACATTAAATG
yiaG	GCATTAAGTAAGCAGTTGATGGAATAG	GCATTAAGTAAGCAGTTGATGGAATAA	GTTGATGTGCATTAAGCCACG
xylR	GATGTAAATAGCGAGGTCATGTTGTAG	GATGTAAATAGCGAGGTCATGTTGTAA	GGATAAGGCGTTCACGCC
sgbU	GCAGGAGGCTGGATTTATATGTTAG	GCAGGAGGCTGGATTTATATGTTAA	CGGTTTAATGACGTTGGCTATTTC
rhsA	CCAGACAGTCAAACGATAAAAAAAGATAG	CCAGACAGTCAAACGATAAAAAAAGATAA	GGAAAGTCCGATTATTGCTTCAATAC
yibA	GCTATTGATAAGCTGAAGCGTTCATAG	GCTATTGATAAGCTGAAGCGTTCATAA	TGAAAACGCCTGGGATATTAATCTC
lctD	CGAAAGGGAATGCGGCATAG	CGAAAGGGAATGCGGCATAA	CGCCATCCTGATAGCTTACG
yibK	CGGGAGCGGTATTGAGAGATTAG	CGGGAGCGGTATTGAGAGATTAA	AACGGCGGTGATTGAAAACG
yibD	TAGCGCCTTTAATTTGCGCTAG	TAGCGCCTTTAATTTGCGCTAA	TAACCTGCCGCCAGACATTA
htrL	CGCTGAGAATTTTCCTTTCAAGAAAATAG	CGCTGAGAATTTTCCTTTCAAGAAAATAA	CGCGACGAGATAACGCG
coaD	CTGATGGCGAAGTTAGCGTAG	CTGATGGCGAAGTTAGCGTAA	GATGGTAAACCGGGCTATTTCG
rph	GAGGGGAATCGAATCCATTGTAG	GAGGGGAATCGAATCCATTGTAA	AAATTCAATAAACTGGCGCTGATATG
dinD	GGAAAACAGTGTAAAAATTACAGAGAAGAAATAG	GGAAAACAGTGTAAAAATTACAGAGAAGAAATAA	CACTGCACCTGGACCTTG
ligB	GCAGATCACAGGTTTTGAACCTTAG	GCAGATCACAGGTTTTGAACCTTAA	AAATATTGCACCTTTCTTTTCCCCC
setC	ATGATTTGCCTGCTGTTTATTAAAGATATTTAG	ATGATTTGCCTGCTGTTTATTAAAGATATTTAA	AACGTGCCATAGCGGGC
yicO	GCACTGAAGATTATTCTGGTGGATTAG	GCACTGAAGATTATTCTGGTGGATTAA	TGCGGCGTGAACGTCT
ivbL	CGGCAATGCGCCGTAG	CGGCAATGCGCCGTAA	TGTCCACTTCCTGGAAGGC
yidH	CTGGTGTTGTATGCCGGATAG	CTGGTGTTGTATGCCGGATAA	CTTCAGCATCAATGACCTGACTTC
glvC	GGCGGAGGTCTGATTGACTAG	GGCGGAGGTCTGATTGACTAA	ATTGCCGACCCCGCC
yidS	GCACATATTCCACAGTTGAAGGATTAG	GCACATATTCCACAGTTGAAGGATTAA	TTACCTGGTCGCTGGTTTCT

yieH	GTGGTTGGGATATTACGGCATAG	GTGGTTGGGATATTACGGCATAA	CAGCGGTTCTTTTGCGAAGA
yieJ	GACTGTGGGGTGATTTTTCGTAG	GACTGTGGGGTGATTTTTCGTAA	CAGCTGCATCCTTCGCTG
atpE	GATGTTCGCTGTCGCGTAG	GATGTTCGCTGTCGCGTAA	ATTGTTGCGTTAAGATTCACAGCA

Table S13. Sequences of MASC-qPCR primers

Gene	MASC-qPCR_forward_TAG_primer	MASC-qPCR_forward_TAA_primer	MASC-qPCR_reverse_primer
rbsB	ATCTGAAACTGGTTGTTAAGCAGTAG	ATCTGAAACTGGTTGTTAAGCAGTAA	GGTCAGCATTAATGCTGCCA
rbsR	GAACGCGGTTCGGCTTAG	GAACGCGGTTCGGCTTAA	TTTAGGCGTTGCTGTAAGTGC
ilvL	ACTTGGACGAGGAAAGGCTTAG	ACTTGGACGAGGAAAGGCTTAA	TGCCCGCAACGCATG
ilvA	AGGTTCTTTTTGGCGGGTTAG	AGGTTCTTTTTGGCGGGTTAA	ACCGCGTGATGATTTTAGAGC
wzzE	CCGCCGTTGCTCGAAATAG	CCGCCGTTGCTCGAAATAA	CATCTCCCGATGCTGCG
aslB	CATTTGCTGGTGGTGAGTAAGTAG	CATTTGCTGGTGGTGAGTAAGTAA	CAGATTAAATCTGACTAAGCCGGC
hemY	CAGAATAACCCGCCACAGTAG	CAGAATAACCCGCCACAGTAA	GCTTATCGTCCACTGACAGATG
yigG	TGCACCTACTGATTCATCAAAACAATAG	TGCACCTACTGATTCATCAAAACAATAA	TTGAATAGTTAAGCGTTGTCTTACTTTTAATG
recQ	TGATGGCGATGACGAAGAGTAG	TGATGGCGATGACGAAGAGTAA	TCGCTTCTTTACGGGAACGA
rmuC	GAATATAATCAGCAGTCGCGCTAG	GAATATAATCAGCAGTCGCGCTAA	CGTGGGCGACCATATCCG
tatD	CACTGTTTGGGATTGCGTTTTAG	CACTGTTTGGGATTGCGTTTTAA	CATTGTCGATCCGGCAACC
yihL	CCGTAGACCTCATCATTAATCTGTAG	CCGTAGACCTCATCATTAATCTGTAA	ACATGGATTTTGACGCCACG
yihW	CCCCAGGAAGACGCGTAG	CCCCAGGAAGACGCGTAA	ACGCGTTAAATCGCTCCAG
yiiD	GCGGGAACGAAGAAGAGTAG	GCGGGAACGAAGAAGAGTAA	GATGATGAATTAAGCATTATCTGCGATAATT
cpxP	TAGCAACTCACGTTCCCAGTAG	TAGCAACTCACGTTCCCAGTAA	ACTGACCAGCCGTCCATAAG
frwD	GCAAACCCATCTTATTCTGGAGTAG	GCAAACCCATCTTATTCTGGAGTAA	CGACTGGAGCACGCTAAGA
zraR	TGGCAAAACTGTCGCGTTAG	TGGCAAAACTGTCGCGTTAA	AAAACTGGCGGATGACGAG
yjaB	GCGTATGTGGGGGCGTAG	GCGTATGTGGGGGCGTAA	CGCATTCTCCATGAAGTGGTC
yjdK	ATGACTATTTCCGTCAGTTTTACAAGTAG	ATGACTATTTCCGTCAGTTTTACAAGTAA	CAGACTCATTGGCCAGGTTATTC
blc-	GGTACAGCAGCCTGGTAGTTAG	GGTACAGCAGCCTGGTAGTTAA	TTGCCTGGGCGATGAAATC
ulaF	CGCTTATTACGGGCAGAAGTAG	CGCTTATTACGGGCAGAAGTAA	ACCGGGGAGTAACGGC
yjfY	CAGCGCGGATATTTATCGCTAG	CAGCGCGGATATTTATCGCTAA	ACACTTCATGCGTAAACACGG
priB	CAGATTGAATTGATAGATTCTGGAGACTAG	CAGATTGAATTGATAGATTCTGGAGACTAA	CCAGCTGACGCTGGTATTTTG
ytfB	AGTTTTATTCGTGCGCGGTAG	AGTTTTATTCGTGCGCGGTAA	AGATATCGACCCATTGCAGGT
ytfG	GCGTAAGCCATCTTTTTAATGTTAATAACTAG	GCGTAAGCCATCTTTTTAATGTTAATAACTAA	CGACAAAAATAAGCCCCCACA
msrA	GCCGCCGGAAGCATAG	GCCGCCGGAAGCATAA	CAGAGGGTTAAAATTCAGACAGCTG
ytfN	GGATTTGCTCTATCAGTTCGAGTTTTAG	GGATTTGCTCTATCAGTTCGAGTTTTAA	CGATAAACTTCACCGTGTACCG
b425	TCGTGATGATCGTCCGGTAG	TCGTGATGATCGTCCGGTAA	TCAGTCACACATTGTGCGC
idnR	ATCACGGCAACACGCTTTAG	ATCACGGCAACACGCTTTAA	AGTTTTTGCAGCGTAACGGT
b427	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GGTTTTTAAGGGAGAGAGGAATATACC
yjgX	CGTGTGACAGCCTTAACTCCTAG	CGTGTGACAGCCTTAACTCCTAA	AACAGGATTGGCACTGTCG
b428	TGGAGAATGAAATATTAAAAAGGCTACTGTAG	TGGAGAATGAAATATTAAAAAGGCTACTGTAA	AAATGGGCTAATTCGGCAGTAC
fecE	CCGATGTGCCTAATGAGGTAG	CCGATGTGCCTAATGAGGTAA	CAAGCCTGTACAGCTCTGGT
nanC	GCATTGGTGTGTCATTTAAACTGTAG	GCATTGGTGTGTCATTTAAACTGTAA	CATGCCGTACCTGCGC
fimB	ACAGCGTCACGCTGTTTTATAG	ACAGCGTCACGCTGTTTTATAA	TTATTTATTTTGAAGAACATAAACCATTGTATTTTGTATG
yjiN	GTTCCCCCTCGGCAATTTTTAG	GTTCCCCCTCGGCAATTTTTAA	GCCTGTCTGGCGAGCT
yjiP	CTATTCCATGGCGGTGATGTAG	CTATTCCATGGCGGTGATGTAA	GCACGACAGTAACATCCACC

mcrB	GACCAACAAATTATTAGGGGACTCATAG	GACCAACAAATTATTAGGGGACTCATAA	AGCCCTCGGCGTGAAA
yjjQ	ATTAAGAAGAGAAATGTCGCACTCATAG	ATTAAGAAGAGAAATGTCGCACTCATAA	TCTCTGCGCTCACTTCTCAT
yjjV	ATACGTTGTTTAACGTGCCGTAG	ATACGTTGTTTAACGTGCCGTAA	CATGATAAGACGCGACAAGCG
lplA	GCGGGGGCTGTAAGGTAG	GCGGGGGCTGTAAGGTAA	AGGTGAAGTTTGATCTTAACGTCAG
creC	GACTTCACCGTCACTTCACATAG	GACTTCACCGTCACTTCACATAA	ATCGCTACGGTAATCAGCACG
kefF	CAGGAGGCCCATCATGGATAG	CAGGAGGCCCATCATGGATAA	CCCCAATCTCGGCAAAGTG
yabI	GCGTAAAGTGGTTGGGGTTTAG	GCGTAAAGTGGTTGGGGTTTAA	GTATGACCAATGAGTTGTTGAGCG
thiQ	CACTATTGGGGATTACGGGTTAG	CACTATTGGGGATTACGGGTTAA	GTTAATTCGCCACCCGTTGAT
murF	CTTTACAGGAGAATGGGACATGTTAG	CTTTACAGGAGAATGGGACATGTTAA	CGTCGTTACGCACCACC
mutT	TTGCGAAGCTTAAACGTCTGTAG	TTGCGAAGCTTAAACGTCTGTAA	GACAGCGTCGCATCAGG
pdhR	CGTCTGGAGCAACGAAAGAATTAG	CGTCTGGAGCAACGAAAGAATTAA	GGAAACGTTCTGACATGGGTTATT
sfsA	AAATCACTGCCGGTTACATTGTAG	AAATCACTGCCGGTTACATTGTAA	TTGCTTCTCCTTAACACGCAC
btuF	AATGCGCTTTCACAGGTAGATTAG	AATGCGCTTTCACAGGTAGATTAA	GAATTGTCCCGCCGCC
cdaR	TACAACTGGATGAAGAGCGGTAG	TACAACTGGATGAAGAGCGGTAA	TCCTCGACGACTTACGTCAC
yaeI	ACGATGCTGGAACTGGTGTAG	ACGATGCTGGAACTGGTGTAA	TTTCAGATTGTCGTACATAGCGTTAC
yafL	TTTGACGGAAGAGACGATTTTGTAG	TTTGACGGAAGAGACGATTTTGTAA	GGCGACTCCTTTTTCTCAACG
yagY	GGTGATAAGTTCATTCCCGTGAAATAG	GGTGATAAGTTCATTCCCGTGAAATAA	GCGCCTGACTGAAGGC
eaeH	AGTGCCTCTTACCCAGCAATAG	AGTGCCTCTTACCCAGCAATAA	TGATTCATACAGGCTGAGTTCAC
b029	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	AATTTTAAGACGGATATCTATTTCGTTTCCA
yahL	ACGTTAAAATCCGCTACACTCTTTAG	ACGTTAAAATCCGCTACACTCTTTAA	CTGCGAAACCACTACTGGTG
yahM	CGGCAAATCAGGTTGATTACGTAG	CGGCAAATCAGGTTGATTACGTAA	CGCATCAATTATCTGGCGAGTT
prpE	GCCAGGCGATGGAAGAGTAG	GCCAGGCGATGGAAGAGTAA	TGGGTAGGCCTGATAAGACG
cynR	GGCGGAAATGAATCACGGTAG	GGCGGAAATGAATCACGGTAA	AACAACCGCGCATTCCTC
frmR	GAACTGGTTCGTGCCTATCTTAAATAG	GAACTGGTTCGTGCCTATCTTAAATAA	GGGTAAATGCGTCGGTATGG
b036	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GAAGCTTAATGTTAAAGAGAAAAATATTCAGATTG
yaiS	CTTGCTGCATTAAAAAAACATAAAAGCTAG	CTTGCTGCATTAAAAAAACATAAAAGCTAA	CATAGCAACATTGCTATGGACATACA
tra5	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	CTAAATTAACTTCAGATACTATCGATTTAACTAACG
araJ	GCGAAACCACTGGGGTAG	GCGAAACCACTGGGGTAA	AAAAGGTGCCAGAACCGTAG
pgpA	CCGCTGGGTATTCTGTCGTAG	CCGCTGGGTATTCTGTCGTAA	ATCGGAACTTCGCGCGA
yajL	CGGCAGGGATTTATAATTATTACGAGTAG	CGGCAGGGATTTATAATTATTACGAGTAA	CGATCTCGACCAGAACAAAACC
panE	GCCTCGCCCCTGGTAG	GCCTCGCCCCTGGTAA	ACCAGGCAGCACGATCA
lon-	TATGCAGGTTGTGACTGCAAAATAG	TATGCAGGTTGTGACTGCAAAATAA	TGAATCCTTCAAGGTACGAACG
mdlA	AGGCCGTCGATGCGTAG	AGGCCGTCGATGCGTAA	ACCACTTTCAACGGCAGGT
ybaA	GATTCGAGTCAATCATCGACGAATAG	GATTCGAGTCAATCATCGACGAATAA	AGCGGAGGGAATCGAAACC
ybaJ	GAGAATCCTGCGAGTTTATCTTGTTAG	GAGAATCCTGCGAGTTTATCTTGTTAA	GGTGATCTGCGGCTGAGTA
priC	GCGCGTTTAACCCGCTAG	GCGCGTTTAACCCGCTAA	TTATTTCTCACTTTTCCGCCTCATC
aes-	CAGTTCTTTACCGCTCAGCTTTAG	CAGTTCTTTACCGCTCAGCTTTAA	ATTCAGGTGATGTGCCCGG
ybcF	GGGACCTGTATTTCGCTGTAG	GGGACCTGTATTTCGCTGTAA	GTTCGGTTAACTGGTGCTCTATTC
tra5	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	TCCTTAAAAATCAGATTCGTGCTCAC
ybcL	CTGAGATAACACCAGTTTATGAGATAAAGTAG	CTGAGATAACACCAGTTTATGAGATAAAGTAA	ATTATCAACAGTTAAAGTATGCGTCTCC
rzpD	CAGAAGTATATTAATGAGCAGTGCAGATAG	CAGAAGTATATTAATGAGCAGTGCAGATAA	CGATGCAGCTAAAATTTGTGGC
ybcY	TAATGTTTTCCGCTTCAGGAAAAAAATAG	TAATGTTTTCCGCTTCAGGAAAAAAATAA	ACGGAGGAAGAAACCTCATTG

hokE	CGCTTACGAACCGAAGAAGTAG	CGCTTACGAACCGAAGAAGTAA	GCGGCGGGTTAGAGC
ybdL	GCCTGCGCCAGCTTTAG	GCCTGCGCCAGCTTTAA	CGCCATCAAATCACCGCC
ybdM	GCGCCTGGACGGTTAAATAG	GCGCCTGGACGGTTAAATAA	GCGGTTTCTACCCTGGATGA
rlpA	GTCATTTATTACTACCGCGCAGTAG	GTCATTTATTACTACCGCGCAGTAA	GGCGCTTCATGATACGAGC
ybeQ	AAATTCGTGAGACTGAAGATTTACTGTAG	AAATTCGTGAGACTGAAGATTTACTGTAA	GAGAAATGTCACTTCCATCTCCC
ubiF	GCTGAAATATGCGTTAGGGTTGTAG	GCTGAAATATGCGTTAGGGTTGTAA	TCGAATCCTCGTACCCCAG
ybfH	GAGGTCTGCTGATTGATAATGGTTAG	GAGGTCTGCTGATTGATAATGGTTAA	AAAATAATAAACGCAGCCAGTATCTTC
ybfD	GATGAGCTTATTGATTTCACGTTTGAATAG	GATGAGCTTATTGATTTCACGTTTGAATAA	TCATCAAATCGTTGTAGATTCATTACTGC
nei-	CCTGGCTGCCAGCACTAG	CCTGGCTGCCAGCACTAA	TTTATGACTGCCTACCTCGCC
abrB	CGAAGCGGTCGGCCTAG	CGAAGCGGTCGGCCTAA	ACTGGAGATTCCTCGATTTTCCT
sucB	TCTGCTGCTGGACGTGTAG	TCTGCTGCTGGACGTGTAA	GATAAGACACGTTAAACGTCACATCA
bioD	AGTACATAAACCTTGCCTTGTTGTAG	AGTACATAAACCTTGCCTTGTTGTAA	AGGAATAGTGGATAACTGTCTCCAG
moaE	GCGGCAAAACGCTGGTAG	GCGGCAAAACGCTGGTAA	CAGCCCGGGGTTGTACG
ybhM	AAGGCATACGCGTTTTTTTCATTAG	AAGGCATACGCGTTTTTTTCATTAA	CCAAAGGTACAATTATCGCCGC
ybhS	CCAAACGTCGGCTGGATTAG	CCAAACGTCGGCTGGATTAA	TTATCTTCATCGTAGATGGCGATG
ybiH	AAGGATTATCGCAAAGGAGTTTGTAG	AAGGATTATCGCAAAGGAGTTTGTAA	CCAGCGATTCAACGCGC
ybiR	TTAGTTATACTCCCGGCCAACTAG	TTAGTTATACTCCCGGCCAACTAA	TGTGCCAATTACCCTGACGA
yliD	CGCTGGATCCGAAAATTAAAGGATAG	CGCTGGATCCGAAAATTAAAGGATAA	AATACATTTAACAAAATAACTATCTGAACGCTG
yliE	TGGGATAAAAGTGGAAAATTAGTAAAAGAGTAG	TGGGATAAAAGTGGAAAATTAGTAAAAGAGTAA	CAGCAATATTTGCCACCGCA
ybjK	TTGAGAGGGTTGCAGGGTAG	TTGAGAGGGTTGCAGGGTAA	GGATAAAACTGAAGTTGCCCTGAA
rimK	GCCTGAAAACGGGTGGTTAG	GCCTGAAAACGGGTGGTTAA	AGCCTGACAGTTATCACATTCAAAAA
ybjR	TGGAGAAATACGGGCAGGATTAG	TGGAGAAATACGGGCAGGATTAA	CCGCAAGTCAGCAAAAGAGC
cspD	GAAGCGGCAGTCGCATAG	GAAGCGGCAGTCGCATAA	TGGGAGATTGGTTTTTCACCG
lolA	TCACGGTAGATGATCAACGTAAGTAG	TCACGGTAGATGATCAACGTAAGTAA	AGGATCATAGAATGTAAATGCCCG
ycaI	TGCCAGTGGATAACGGGTAG	TGCCAGTGGATAACGGGTAA	AAGGAGCGATAACATGAAGGTATC
lpxK	GCTGGCTTCTGGCAACTAG	GCTGGCTTCTGGCAACTAA	TGGTATCGATTTGCAGCAAGG
ycbK	CAGCACGGCACTGGTAG	CAGCACGGCACTGGTAA	GGATCTGCATCAGTGTCAGG
ycbX	GCACTTAAGTTGGCGCGTTAG	GCACTTAAGTTGGCGCGTTAA	CACGAGCAGGTCAATTTGCT
tra5	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	TGTGACTCAGGAAAATGCGTTAAA
yceO	GCAATTATGGATCTGATCTATGTGTTTTATTTTTAG	GCAATTATGGATCTGATCTATGTGTTTTATTTTTAA	TTTCAATAATTTTGAATTAATAATGAAAGAATCTTCATAAGT
yceF	GAAAAACCCGCTGATGGGATAG	GAAAAACCCGCTGATGGGATAA	GCGTATCGAAGCTCCGATG
plsX	AGCGGAACTCTGCGGTAG	AGCGGAACTCTGCGGTAA	TACCGGTACGAGTGACAATCC
fabH	CGCGCTGGTTCGTTTCTAG	CGCGCTGGTTCGTTTCTAA	GTTCTTCAGCTGGCCCCT
pabC	TTTGTGAGCGCCCGAATTAG	TTTGTGAGCGCCCGAATTAA	CGGCATAAAGCTGTTCACCG
b114	AGACCGGATATCTATCCGAAAGATTAG	AGACCGGATATCTATCCGAAAGATTAA	AAGCGTTTTCTGTTACATCGAGC
ymfM	AGTTTCGTCGTAAATTCAGGGAAATAG	AGTTTCGTCGTAAATTCAGGGAAATAA	CTCATCCCCGCGCCA
ymfR	TCTGTCGGCGGAGGTAAATAG	TCTGTCGGCGGAGGTAAATAA	CCCGACCGACTCTGCC
tfaE	AACTGCCGGAGGCGTAG	AACTGCCGGAGGCGTAA	CGAACTTTCAAAAGCAGATGCTTATC
ycgX	CTTTTATATATTGAAGATGTAGAAAATAATAAACCATTAATCTAG	CTTTTATATATTGAAGATGTAGAAAATAATAAACCATTAA	CAAATATCAAGTTGATAGTATTAGTCTGGTGAT
ymgC	AGGGAATCCGTGCTCTCTTAG	AGGGAATCCGTGCTCTCTTAA	AATGTAGGCAGTACAAAATAAGTTAGGC
ycgY	GGACTTGAAAAACCGTCTGTGTAG	GGACTTGAAAAACCGTCTGTGTAA	CGAAAGAGCAACCGTGTGAC
hemA	GCCTCGGGCTGGAGTAG	GCCTCGGGCTGGAGTAA	CATCACTTAACTGCGCATATTCG

b122	GCTCGAGCCGAACCTTAG	GCTCGAGCCGAACCTTAA	TCGACCAAGGTTCGACTCG
kch-	CAGTAAAGAATCGGCGCAAAAATAG	CAGTAAAGAATCGGCGCAAAAATAA	CCGACGTAAATCGCGTTGG
sapB	CATAAGGAATGGTATGCCTTACGATAG	CATAAGGAATGGTATGCCTTACGATAA	ATTGATAACCGAGAAATTGCTGGTC
pspF	CGTTGTTGAAAAAGCACCAGATTTAG	CGTTGTTGAAAAAGCACCAGATTTAA	GGATGCGCAATTCGATGC
ycjU	GGCCTTCTGGCAAAACGTATAG	GGCCTTCTGGCAAAACGTATAA	CTCCTCAAGCCCGGCA
ycjV	GTTGGGGGCACGAGTTAG	GTTGGGGGCACGAGTTAA	CGGTACAGGGTAATAACTTTTTCATTATTG
ycjG	ACGGGCGAATTGCATCTTTAG	ACGGGCGAATTGCATCTTTAA	GGCTCAGGCGTTTGAATTG
ycjZ	GGCTGAAGTATAAAGGTGCTGTTTAG	GGCTGAAGTATAAAGGTGCTGTTTAA	GTAAAGCACATAGAACTCCGAACA
intR	TTGAATGCGCGGTTATCGTAG	TTGAATGCGCGGTTATCGTAA	ATGATGCGATCGCCTTCTTC
sieB	GCAGGAGTAGACGTTTTAATCGTTAG	GCAGGAGTAGACGTTTTAATCGTTAA	AACAAGATGTGAGATTCGCAAAATG
b135	GAATCTCACATCTTGTTGCTACGTATAG	GAATCTCACATCTTGTTGCTACGTATAA	GATTCAAATTTTGCAGTGCAGCA
ydaY	AGCAGGATACGATAAAAAAGCCATAG	AGCAGGATACGATAAAAAAGCCATAA	CAGCCATTCCAGCATAACGTC
b136	GTATTTTCAGGAGCAGTGTGTGTAG	GTATTTTCAGGAGCAGTGTGTGTAA	GCTCTGCCATTTATTAAAACTCCTGA
tfaR	GCCAGCACTGCCGTAG	GCCAGCACTGCCGTAA	AGTATAAGTGCCATTGCCCG
paaG	CGCAGTTCACGGGGAAATAG	CGCAGTTCACGGGGAAATAA	TTTCCCCGCGTCACGC
paaX	AGGCGTTATGCCAATTTATCAGATAG	AGGCGTTATGCCAATTTATCAGATAA	GCATAACGCAATTATCCTGAATGTTC
b140	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	GTATATCTGAATGGAGATACGACAATCAG
ydcL	CACCATGTTTGACGTTAACAAAAAGTAG	CACCATGTTTGACGTTAACAAAAAGTAA	GCTATTTTGCCAGTAGCTGCG
rhsE	AGGGAGTTCAGATGCTTCAAATTATTAG	AGGGAGTTCAGATGCTTCAAATTATTAA	TTATAAAATGACCATTCCCTCTACTAATTTTTTG
yddJ	CCTCCAGTTCATGGCCTTTTTAG	CCTCCAGTTCATGGCCTTTTTAA	GGGATTAAGTAGCCAGATATGTGTTAC
ddpF	TGGCGGGCGCTGTAG	TGGCGGGCGCTGTAA	GAGTGAAGTTGCGGTGCC
yddV	GAACTCTGGAAAGCCAGTCTTTAG	GAACTCTGGAAAGCCAGTCTTTAA	TCTTCACGTTTGTATCCCCAGA
ydeV	GGAAAGCGCCTGGGTTATAG	GGAAAGCGCCTGGGTTATAA	ATAAGGGTTGTGCTTAATGCTTTAAGA
ydeA	CACTCGAAGAACAGACGCAATAG	CACTCGAAGAACAGACGCAATAA	TGGGGAAGCCTACGCAG
marA	TTTTTACATCCATTAAATCATTACAACAGCTAG	TTTTTACATCCATTAAATCATTACAACAGCTAA	GTGGGTGTTCCTGCGATG
marB	GCCCTATTATAATCAACACGCTATGTAG	GCCCTATTATAATCAACACGCTATGTAA	AGTAGGACTGGCAAGTGCG
tfaQ	GCCAGCACTGCCGTAG	GCCAGCACTGCCGTAA	AGTATAAGTGCCATTGCCCG
ydfP	TGAATCTTTAAGAGATAAAAAGCATCAAATGTAG	TGAATCTTTAAGAGATAAAAAGCATCAAATGTAA	GAAAAACACAGCAATGAAAGAATAGCT
rzpQ	CAGAATGATTACATACCCGCGTTAG	CAGAATGATTACATACCCGCGTTAA	GCTATCTATGAACCAGCTACCTCG
b157	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	TCTTTCAGTGCCAGAAATTTTTCAG
ynfA	TGGGGGCGCACGTAG	TGGGGGCGCACGTAA	TGTCGTGATGCATGAAATCTACG
speG	GAAGCCGACCGCACAATAG	GAAGCCGACCGCACAATAA	AATTAGATTACACAGCGGTTACTTTACTG
dmsD	CGGTTAAACCGCTGTTTCGATAG	CGGTTAAACCGCTGTTTCGATAA	AGATGTGTAGATGATGCATGAGCA
ydhM	CGACGCCTGCCGTTTAG	CGACGCCTGCCGTTTAA	ACGTTGGCGATAGTATTCCG
lhr-	AAGGGCTGGATTGGGGATAG	AAGGGCTGGATTGGGGATAA	ACTGTGGGTTGACGGCG
ynhG	GTTACGCGTGACGCAGTAG	GTTACGCGTGACGCAGTAA	AGTTGACCAGCTGAGCAAC
sufA	GGCGAAAGCTTTGGGGTATAG	GGCGAAAGCTTTGGGGTATAA	CAGCATCCACTCCGGCT
ydiA	CCTTAGTCGCCGAATGTACTAG	CCTTAGTCGCCGAATGTACTAA	GGAGTTCGTCAGTTCTGTTCATTAC
ydiE	CAAGCTGCTGTTGACCAAGTAG	CAAGCTGCTGTTGACCAAGTAA	GGTATTGCGCAACTGGTTGG
nlpC	GGCAGGCGAGACGAATCTAG	GGCAGGCGAGACGAATCTAA	ACAATTACAGAAGGAGTAACTTTCATTTGT
btuC	TGGTTATTGTTAAAAGCAGGACGTTAG	TGGTTATTGTTAAAAGCAGGACGTTAA	ACACTTTGAGGCGACATTGAC
arpB	CACCCAACAGTAAATCTCAATGGTAG	CACCCAACAGTAAATCTCAATGGTAA	GCTCTGAACTCCGCGTC

ydjX	ACTTACTCACCCAAAAAATGAAGGATAG	ACTTACTCACCCAAAAAATGAAGGATAA	CGTATGGTTATTGGCATCGACA
nudG	GCCAGACCAGCGGATTAG	GCCAGACCAGCGGATTAA	CCCGCATTATCGTCCGGA
b178	ATCGAACGGACCGGCTAG	ATCGAACGGACCGGCTAA	GCGCGATCCCCCCTT
yeaL	GCTGATTGTGGGGAAACAGTAG	GCTGATTGTGGGGAAACAGTAA	TTGAGTTTTCGCCAGTGGC
yeaX	CGCCTGGTGTTGGATTTGTAG	CGCCTGGTGTTGGATTTGTAA	AACCAACTGCTGAACTGGC
yobB	GGCGATATCATTCCATTACGCTAG	GGCGATATCATTCCATTACGCTAA	ATGAATCGCCATCGCTTGTG
exoX	TTTAACACTGAAACATTATCTGGAAAATACTTAG	TTTAACACTGAAACATTATCTGGAAAATACTTAA	AATTGCGCGAGCTGAAAAC
yecN	GTTGGTTTTCTCCCTGCGTTAG	GTTGGTTTTCTCCCTGCGTTAA	AATCATGGAAATAATATTGGAATAGCCGG
otsA	ACCTTTCCAAAGCTTGCGTAG	ACCTTTCCAAAGCTTGCGTAA	CCTACGGTGAGTTAAGCGAAC
dcyD	GCCTATCATCCCCACGTTTAG	GCCTATCATCCCCACGTTTAA	CACGAAAAATAGAGATATAAAAACGCGC
yedM	CCCACCCATTCGGCATAAATTAG	CCCACCCATTCGGCATAAATTAA	TGTGGAGTTCAACACATTTAAACAATC
fliE	GTGATGAGCATGCAGGTGTAG	GTGATGAGCATGCAGGTGTAA	ACGTGGTGATGTGTCAAATGAG
fliN	GCCGCCTGAGCCGTTAG	GCCGCCTGAGCCGTTAA	AAGCACTGGCGCTAATTTTCA
fliP	GCGCAGAGCTTTTACAGCTAG	GCGCAGAGCTTTTACAGCTAA	GCGATGATTTTCGGAATAAACGAC
fliQ	CCTGCCGTATATCATCGGGTAG	CCTGCCGTATATCATCGGGTAA	TGATGGGGCAATGGCGA
yedS	CGTAACAATGGCTTCTTTGGTTAG	CGTAACAATGGCTTCTTTGGTTAA	TATGCAGCACCCAGACTTAAAC
cbl-	GCAGTGAAGAGGAAATTGATTATCAGATATAG	GCAGTGAAGAGGAAATTGATTATCAGATATAA	AAAAATCGGGTTTGACAAAAGATTTTTC
b199	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	CAACCTTTGGGAAAGCCCG
hisL	CCATCATCACCATCATCCTGACTAG	CCATCATCACCATCATCCTGACTAA	AGCTATGCGTAAACGAGTGTTG
wbbJ	TGAGACCAAATTATGGGAAAAAGCATAG	TGAGACCAAATTATGGGAAAAAGCATAA	TTAACCTCAGGAAACTCAATGAATTTAACC
wcaM	GCCGGAACTGAGGGAGTAG	GCCGGAACTGAGGGAGTAA	TCGTCATTATTTCATCCACCTGTAAAAT
wcaL	AGCTTGCTGCAGGCTTTATAG	AGCTTGCTGCAGGCTTTATAA	GCCCTTCGCTGAAGGC
wcaC	GAGTATGTCAACTTCTATCAGAATCTGTAG	GAGTATGTCAACTTCTATCAGAATCTGTAA	TGAGTCCCGCGCCAAT
asmA	GAAGAAGTTGCTGGAGAAGATGTAG	GAAGAAGTTGCTGGAGAAGATGTAA	AGATTGAGCCACCCTGGC
baeR	GCCTGCCGCATCGTTTAG	GCCTGCCGCATCGTTTAA	CTTGATTGAAATTGTGTATCGAAAACAGC
gatR	CGCGGGTAAAATTAATTACGGTGTAG	CGCGGGTAAAATTAATTACGGTGTAA	GGCGAACCGTTGAGTGG
tra5	CAATTTGAAAACAAGAACCTCGCTTAG	CAATTTGAAAACAAGAACCTCGCTTAA	ACAAAAAGGCGTGGTTACGC
yegV	TCCTCGCACACAAAAACGTATAG	TCCTCGCACACAAAAACGTATAA	AGCCAGCACATTTACCCACA
yegW	CGTTTTTGTGTGCGAGGAGTAG	CGTTTTTGTGTGCGAGGAGTAA	TCCGGCATTTCCAACGC
yehQ	AGAGTGCTATTGGTCAGGCATAG	AGAGTGCTATTGGTCAGGCATAA	CTCTTAGTAAAAAGCACTATTTCATCCTTG
yohC	GGCTACGGTTATCGTTTGTTCTAG	GGCTACGGTTATCGTTTGTTCTAA	CATGATGAGCATTCAGATAGTGGTT
yohF	GAATCCACAGTTCAACCCAGAATAG	GAATCCACAGTTCAACCCAGAATAA	GTGGCTGAAACGACGCG
sanA	TTACTTGAATTACAAAAGAAACAAGGAAAGTAG	TTACTTGAATTACAAAAGAAACAAGGAAAGTAA	GAAAAAGAGTCCGGGATTTTTTGC
yejA	GCCAGCAAACAGGGAGAGTAG	GCCAGCAAACAGGGAGAGTAA	GCACGAACACCTTCACCG
yejE	CATTTGATCCTAATAAGGCGGTGTAG	CATTTGATCCTAATAAGGCGGTGTAA	AGGCGTAAAATTGACAGCGC
b219	GTAAATGCCTTGAATCAGCCTATTTAG	GTAAATGCCTTGAATCAGCCTATTTAA	CAAGCGACAGTTCGGCTTC
rcsD	CAGTTATGTCAAGAGCTTGCTGTAG	CAGTTATGTCAAGAGCTTGCTGTAA	CAACACATGCGCATCCAGT
rcsC	GGAAATCGCGGGATTCGTAG	GGAAATCGCGGGATTCGTAA	GACGCGTTAGTGTCTTATCTGG
yfaT	CGTTTAAATTTTCTGGCGCGATAG	CGTTTAAATTTTCTGGCGCGATAA	CCGATAATCACGCCCCG
menF	GCTGCGTACTTTATTACAAATGGAATAG	GCTGCGTACTTTATTACAAATGGAATAA	GATACAGATGTGTCTGACGCC
yfcO	GTTCCAGTACATACTGAATGAGTTGTAG	GTTCCAGTACATACTGAATGAGTTGTAA	CATGTGATGGCCTTTCATCAGG
yfcU	GGGCGGATAAGGGCATGTAG	GGGCGGATAAGGGCATGTAA	ATCATGGCTGTAGTAGCCATCAA

tfaS	CTGAACTGCCGGAGGAGTAG	CTGAACTGCCGGAGGAGTAA	AACTGAAAAAGCAGCAATTGATTAATCA
ypdI	CGAATTATCTTATTCATCCTGTGCGATAG	CGAATTATCTTATTCATCCTGTGCGATAA	CTGGCGGTAATGAACTAGTGAC
yfdY	GGGTATGGCAGGCGGTAG	GGGTATGGCAGGCGGTAA	GATTTTGTTTAAGATGAGAGAAGATGCG
yfeO	CCACCCCACGATAACGTTTAG	CCACCCCACGATAACGTTTAA	CCTGCGTTTGCTGCAAG
mntH	ACGGCGCTGGGATTGTAG	ACGGCGCTGGGATTGTAA	GATCATCGCGGAAAAGGCC
xapR	TTGGCTGCATTGACACAGTAG	TTGGCTGCATTGACACAGTAA	CGAAACAACGGAGCCGC
yfeR	AAGGGGATCCGTTGTATCAGATATAG	AAGGGGATCCGTTGTATCAGATATAA	GAATGGCGCAGTTGCACT
yffB	TATCAGCAATTTTTCCATGAGGTGTAG	TATCAGCAATTTTTCCATGAGGTGTAA	CCCGGCAAAGGCTAACG
hda-	CGTTTGTGAAAGAAATTCTGAAGTTGTAG	CGTTTGTGAAAGAAATTCTGAAGTTGTAA	TCCACCGGAGCAGGTACT
yfgG	CGAATCACCGGTACAACGTTAG	CGAATCACCGGTACAACGTTAA	CATTACAACATATATATTAATTTTATATTTATTTCGCAGCC
pbpC	CGACAGTGAAATTTGTCATGCAATAG	CGACAGTGAAATTTGTCATGCAATAA	ATGGAAAAAGTACGTTCAATAGCCAT
yphA	CAATTTCTCTCGATCGGCGTTAG	CAATTTCTCTCGATCGGCGTTAA	TTTCTTCTGTCTGGAACCGATGA
yfhB	GTGAACTCCAGCAACTGGAATAG	GTGAACTCCAGCAACTGGAATAA	TACCCGATTGTTATGCACTAATACC
kgtP	GGAAGGGGATGCGTCTTTAG	GGAAGGGGATGCGTCTTTAA	GCATCAGGGGATGAGCCG
yfiA	CGTCGAAGAAGTTGAAGAAGAGTAG	CGTCGAAGAAGTTGAAGAAGAGTAA	TTCACACGACGAAACGCC
yfjQ	GCTCACGCAACTGCAGTAG	GCTCACGCAACTGCAGTAA	CCAGGAAAGTTAACTGACTGAAAATGA
yfjR	CGACAGCCCACCGGAATAG	CGACAGCCCACCGGAATAA	ACTGGCAGCACCGGC
ypjC	GTTTTTTGAATGCGCTAAAAGAAATGAATAG	GTTTTTTGAATGCGCTAAAAGAAATGAATAA	GGTGTTCCTTTGACTCTGACAAA
ygaQ	CAGACAGGATAGGAGAAAGAAAAACATAG	CAGACAGGATAGGAGAAAGAAAAACATAA	ATTATATACATCGGCTGCCAGCT
ygaR	AAGGTGGTATCCCTGGCTATTAG	AAGGTGGTATCCCTGGCTATTAA	AGGCATTCCATTTAAACGCTTTTC
yqaC	CGGCGGTGAGGCGTAG	CGGCGGTGAGGCGTAA	TATTTATTATAACTAGTCCATCATGAGGAGAGT
gabT	TTTTGATGAGGCGAAGCAGTAG	TTTTGATGAGGCGAAGCAGTAA	TCACGCGTATAATTGTTATTGATTCTTTTT
ygaU	GTTGCGTATTCCGGAAGAGTAG	GTTGCGTATTCCGGAAGAGTAA	ATTTATCTGGCTGCGGCG
ygaM	GTTAAGCATGCGCAAATCGTAG	GTTAAGCATGCGCAAATCGTAA	AGAGATACTAGATGTAGTTGAAAAAAGATTCAA
luxS	GTTGCAGGAACTGCACATCTAG	GTTGCAGGAACTGCACATCTAA	TTCCGCCAATGAGATCTTCTGG
mltB	GCTGGCGCGAGTACAGTAG	GCTGGCGCGAGTACAGTAA	CCTGTTCACTTAACTGCATCAGTT
srlE	GGTTTGTCTCCGGTTTTATCTATCAATAG	GGTTTGTCTCCGGTTTTATCTATCAATAA	GGATGGAGTGCACCTTTCAAC
norW	GATTGTTGAAAACATTGCCGATGTAG	GATTGTTGAAAACATTGCCGATGTAA	TTATTGCTGCGGCGCG
ascB	CCAGTAATGGGGAAGATTTAGAGTAG	CCAGTAATGGGGAAGATTTAGAGTAA	GGTGGAAGAAGAGTAGTTTTTCATTAAGG
hycI	AGTTAGCGGTGGAAGAAGAGTAG	AGTTAGCGGTGGAAGAAGAGTAA	TGGGGAAGATTTAGAGTAGGTAACAG
hypB	GAGACACAGCGATGTGCATAG	GAGACACAGCGATGTGCATAA	TTCGGCTTCATTAATTACGCTCAT
ygbA	TATCGTCCGAAAAAACCTCATGAATAG	TATCGTCCGAAAAAACCTCATGAATAA	TGATTGTGCGCGTGCTG
pphB	GGCGGCTGTCATTTTATAAAATAAAGTAG	GGCGGCTGTCATTTTATAAAATAAAGTAA	ATGGCTGGCGTTACCATTAAG
ygcN	AGGAGTCTGCGATGTCTGTAG	AGGAGTCTGCGATGTCTGTAA	CAGCAAACGGCAGGTGC
mazF	GCCAAAATTAACGTACTGATTGGGTAG	GCCAAAATTAACGTACTGATTGGGTAA	AAGGCGCAATGGTGGC
relA	TCGGTTGCACGGGAGTTAG	TCGGTTGCACGGGAGTTAA	ATCTTCACTTCATCATCAATATTCAGATTGA
rumA	GTACTTTTCTCGCGCGTTAAATAG	GTACTTTTCTCGCGCGTTAAATAA	CAGACAATACGCCCAGGTTTC
ppdC	CATTGCCCGAATCGTCAGTAG	CATTGCCCGAATCGTCAGTAA	ATACCAAACTTTTGCGACAGGG
ptsP	CTGATTCGCGGAGGGTTATAG	CTGATTCGCGGAGGGTTATAA	TCCGGAAAATGCAGATAGCTACT
mutH	CCCGTCATTTTCTGATCCAGTAG	CCCGTCATTTTCTGATCCAGTAA	CCACCAGAGAAAGGAAAATAATATTGTC
ygeA	TGTCGCTTTTATGCTGTCGTAG	TGTCGCTTTTATGCTGTCGTAA	CTAGATTATGCGGCAAGCGG
kduI	GCCGTTAAAGATTTGCGCTAG	GCCGTTAAAGATTTGCGCTAA	CTCGATGGTTTCAGTCGGTTC

yqeF	TTGACCATTGAACGTGACGAATAG	TTGACCATTGAACGTGACGAATAA	CCATGGGCTAGGGTTTTCAATTAA
ygeI	CTATCTCACCTTACTATAAAATTGAGTGGTAG	CTATCTCACCTTACTATAAAATTGAGTGGTAA	GAGCAAGAGTTGCTGGCG
pbl-	CAAAATATATATTGTATATACCAGGCTTAATGAACTAG	CAAAATATATATTGTATATACCAGGCTTAATGAACTAA	GCAGTTAAGCAGGAAAACAGTCG
ygeM	GGCAAGCCCTGAGGACTTATAG	GGCAAGCCCTGAGGACTTATAA	TCATCCCTTCCATTATTATTATAAGCAAAATCC
b286	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	ATCGTGCCTTTTCATTTATTTCCAG
ygeP	ACACCAGAAGAAAAGAGCAAGTTTAG	ACACCAGAAGAAAAGAGCAAGTTTAA	CCCGGTTCAAAGCTCTGC
yqeC	GCGCGCAGATTTGTGAATTAG	GCGCGCAGATTTGTGAATTAA	GATAGAACCATCGGCGCG
ygfJ	AGCGAAGGAAAGGTATACTGAAATTTAG	AGCGAAGGAAAGGTATACTGAAATTTAA	TTTGTGAATGAACTCACAGGTGC
recJ	TCATCGACAATATCTGGCCAATTTAG	TCATCGACAATATCTGGCCAATTTAA	TCGTAGTCAAAGATACCCCCTAAG
argO	TGCACAAGCCTTGTTCAGTTAG	TGCACAAGCCTTGTTCAGTTAA	TGCGGTCCATCCGGC
yggU	CAGAAATCGCGGCGTTAATTAATTAG	CAGAAATCGCGGCGTTAATTAATTAA	TGGCGCGCTTTCAGAATC
mutY	GGCGCGCCGGTTTAG	GGCGCGCCGGTTTAA	TTTTCATTAATCAGCATGGTTTGCTTG
glcC	GCTGGAGATGAACCTGAGTTAG	GCTGGAGATGAACCTGAGTTAA	GGTGATTGCGCTGTTATTGATTG
yghQ	CTCGAAGCGTATGACCTGATTTAG	CTCGAAGCGTATGACCTGATTTAA	ATGATACGTCAGTGGGAGAGATC
yghT	CGCGCGTGGTTATCACTAG	CGCGCGTGGTTATCACTAA	GACGATGGTGGTGGACGG
ygiZ	TGGGGACGGTTTATATTTTGCTATTAG	TGGGGACGGTTTATATTTTGCTATTAA	CCGGACAAACCGCTACC
yqiB	CGATGGCGATTCCGGTTTATTAG	CGATGGCGATTCCGGTTTATTAA	TGGCCGAATCGCCTCC
b304	GGTTAAGTGATAACAGATGTCTGGAAATATAG	GGTTAAGTGATAACAGATGTCTGGAAATATAA	TCTTATTATTAATCAGAGTGCTTAGTGAATATTTG
yqiI	GACGTAATCCTGAAAATCACCTATAACTAG	GACGTAATCCTGAAAATCACCTATAACTAA	GTCGCCCGGTCGATATTCT
ygiP	GATGAAGGCTATCAGGTCATGTAG	GATGAAGGCTATCAGGTCATGTAA	CGTTTTATGTCTGGTGGTTTAAGC
yhaK	CTTTGCTGATAGATTTGCCTGTCTAG	CTTTGCTGATAGATTTGCCTGTCTAA	TCATCCTCGGCTAAACGCG
tdcA	ACGAAGGCAATTAATAGAAGTTGGTTAG	ACGAAGGCAATTAATAGAAGTTGGTTAA	ATTCGCCCAGCCAGTCG
yhaC	CGTATGATTACGACCTTTTAAATGCAATATAG	CGTATGATTACGACCTTTTAAATGCAATATAA	CAATATATTGCAGAGATGGCGTAGG
ecfH	GTAACTACGGCGTTTACGTTTATTAAATAG	GTAACTACGGCGTTTACGTTTATTAAATAA	GGGGCGATGGTGGCA
yhbW	GATGTTAAGGAAGAGTTGTTGGGATAG	GATGTTAAGGAAGAGTTGTTGGGATAA	GCAACGCACTGGTGCATAT
nlpI	GCAGAATCGGACCAGCAATAG	GCAGAATCGGACCAGCAATAA	GTAGTACGTGTGCCTCAAAATTAATG
dacB	GCCGTTTGTATAAAGATATTTATCAGAACAATTAG	GCCGTTTGTATAAAGATATTTATCAGAACAATTAA	ATCAGCCTGATGCGCTACG
ptsN	GGTACTCCGGATGAAGCGTAG	GGTACTCCGGATGAAGCGTAA	GCCAGGGTTCGAGCCA
yhcC	GGCTACCTCCAACGGAGTAG	GGCTACCTCCAACGGAGTAA	GCGCCAGCAACATTGAGA
yhcE	GTAAATGCCTTGAATCAGCCTATTTAG	GTAAATGCCTTGAATCAGCCTATTTAA	CAAGCGACAGTTCGGCTTC
yhcO	CATTTGCGTTTTAATGTTCGTCATTAG	CATTTGCGTTTTAATGTTCGTCATTAA	AATGGATATGCGTGAAATGCTGAA
mreC	CGCCGGGAGGGCAATAG	CGCCGGGAGGGCAATAA	TCACAAAACCTGTGCCCAC
yhdN	GTTGTTGGACAAAATCAACGATAACTAG	GTTGTTGGACAAAATCAACGATAACTAA	GGCATGGCGGATAAATTTCAATC
kefB	CGGCTGGGATGAATTTGAGTAG	CGGCTGGGATGAATTTGAGTAA	CATCTTTGCGAACGTGATCG
frlD	ACCACGGTGCCTGGTAG	ACCACGGTGCCTGGTAA	AACCCCCTGCGCGATAT
bioH	CGTTGAAGCAGAGGGTGTAG	CGTTGAAGCAGAGGGTGTAA	TAATGACTATTGCGGAAAGATTGCG
gntX	GTGCCTTTGTCGAACCTTGTAG	GTGCCTTTGTCGAACCTTGTAA	GCCACATTCAGCGTTAGGC
malQ	GCTGCAGCGAAGAAGAAGTAG	GCTGCAGCGAAGAAGAAGTAA	TCGATGCTGGAAACCATTATCTC
glgA	CGTGAGCTTTACTATCGCTTGAAATAG	CGTGAGCTTTACTATCGCTTGAAATAA	CGATCGCGCACAGCAAATA
yrhA	TGATGCTGCCAACTTACTGATTTAG	TGATGCTGCCAACTTACTGATTTAA	TTGAGAAGCGGTGTAAGTGAAC
ugpQ	CCGAACTTTACGGCCCAATAG	CCGAACTTTACGGCCCAATAA	ATAATTCTCAGCGGGTGTTGC
gadE	CACCCTGGGTATCACATCTTATTTTTAG	CACCCTGGGTATCACATCTTATTTTTAA	TCTGCCAACAGTTCCTGCC

gadX	CGTTCTGCGGAAGGAATAAGATTATAG	CGTTCTGCGGAAGGAATAAGATTATAA	GTTTATAAAAAAATGGCTGATCTTATTTCCAGTA
yhjK	GAAGAGAGTTACCTGGAAGAAAAGTAG	GAAGAGAGTTACCTGGAAGAAAAGTAA	TTTAGTAAAACCACAGGTCTTGTAAGG
yhjR	CAGAGTTTGCGCAACAAAAGTAG	CAGAGTTTGCGCAACAAAAGTAA	GACGGTGGGTAAAATCAACATTAAATG
yiaG	GCATTAAGTAAGCAGTTGATGGAATAG	GCATTAAGTAAGCAGTTGATGGAATAA	TTGAAGGGACTTGCCTTACTACA
xylR	GATGTAAATAGCGAGGTCATGTTGTAG	GATGTAAATAGCGAGGTCATGTTGTAA	CGACTTGTCAGTCGGATAAGG
sgbU	GCAGGAGGCTGGATTTATATGTTAG	GCAGGAGGCTGGATTTATATGTTAA	ATCTCAACCACCACCATATCGT
rhsA	CCAGACAGTCAAACGATAAAAAAAGATAG	CCAGACAGTCAAACGATAAAAAAAGATAA	CTCTTAACTGTAATACTCTGGCAGATG
yibA	GCTATTGATAAGCTGAAGCGTTCATAG	GCTATTGATAAGCTGAAGCGTTCATAA	CGGCGTGTACACCGG
lctD	CGAAAGGGAATGCGGCATAG	CGAAAGGGAATGCGGCATAA	GGGGCAGGATAGCAGAAAATTG
yibK	CGGGAGCGGTATTGAGAGATTAG	CGGGAGCGGTATTGAGAGATTAA	ATACCACCGCCGCTGG
yibD	TAGCGCCTTTAATTTGCGCTAG	TAGCGCCTTTAATTTGCGCTAA	CCGCGTAATCCGTTCCG
htrL	CGCTGAGAATTTTCCTTTCAAGAAAATAG	CGCTGAGAATTTTCCTTTCAAGAAAATAA	TGTAACTTGATGTCTCTCAACACCTA
coaD	CTGATGGCGAAGTTAGCGTAG	CTGATGGCGAAGTTAGCGTAA	GATGGTAAACCGGGCTATTTCG
rph-	GAGGGGAATCGAATCCATTGTAG	GAGGGGAATCGAATCCATTGTAA	CCGGATTTCAGCGTAAACTCG
dinD	GGAAAACAGTGTAAAAATTACAGAGAAGAAATAG	GGAAAACAGTGTAAAAATTACAGAGAAGAAATAA	AGAGATGCATTCTTCTGACTTCTTG
ligB	GCAGATCACAGGTTTTGAACCTTAG	GCAGATCACAGGTTTTGAACCTTAA	AGTTGTATGCCGGTGTCTCA
setC	ATGATTTGCCTGCTGTTTATTAAAGATATTTAG	ATGATTTGCCTGCTGTTTATTAAAGATATTTAA	TCCCCACAACACGGCG
yicO	GCACTGAAGATTATTCTGGTGGATTAG	GCACTGAAGATTATTCTGGTGGATTAA	GTACGACATCGATGTTCAGCC
ivbL	CGGCAATGCGCCGTAG	CGGCAATGCGCCGTAA	CCTGCTGTTCCAGGAAATGAAC
yidH	CTGGTGTTGTATGCCGGATAG	CTGGTGTTGTATGCCGGATAA	GGATCAGCGCCACGATG
glvC	GGCGGAGGTCTGATTGACTAG	GGCGGAGGTCTGATTGACTAA	GCTTTTGAGTAGAGTTTCACTTCCG
yidS	GCACATATTCCACAGTTGAAGGATTAG	GCACATATTCCACAGTTGAAGGATTAA	TAACCGGCGGCGTGT
yieH	GTGGTTGGGATATTACGGCATAG	GTGGTTGGGATATTACGGCATAA	CAGCGGTTCTTTTGCGAAGA
yieJ	GACTGTGGGGTGATTTTTCGTAG	GACTGTGGGGTGATTTTTCGTAA	AAAATGTTCTGCAAGGTCCGG
atpE	GATGTTCGCTGTCGCGTAG	GATGTTCGCTGTCGCGTAA	AATTTCTTTTTGACGTTTTTCGATGGC

References (Supporting Online Material)

1. D. Yu *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5978 (May 23, 2000).
2. H. H. Wang *et al.*, *Nature* **460**, 894 (Aug 13, 2009).
3. H. Li, R. Durbin, *Bioinformatics* **25**, 1754 (Jul 15, 2009).
4. H. Li *et al.*, *Bioinformatics* **25**, 2078 (Aug 15, 2009).
5. W. McLaren *et al.*, *Bioinformatics* **26**, 2069 (Aug 15, 2010).
6. S. Y. Gerdes *et al.*, *J Bacteriol* **185**, 5673 (Oct, 2003).