

Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA

Ryan R Gallagher^{1,2,4}, Zhe Li^{1,2,4}, Aaron O Lewis^{1,3} & Farren J Isaacs^{1,2}

¹Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, USA. ²Systems Biology Institute, Yale University, West Haven, Connecticut, USA. ³Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, USA. ⁴These authors contributed equally to this work. Correspondence should be addressed to F.J.I. (farren.isaacs@yale.edu).

Published online 4 September 2014; doi:10.1038/nprot.2014.082

Multiplex automated genome engineering (MAGE) is a powerful technology for *in vivo* genome editing that uses synthetic single-stranded DNA (ssDNA) to introduce targeted modifications directly into the *Escherichia coli* chromosome. MAGE is a cyclical process that involves transformation of ssDNA (by electroporation) followed by outgrowth, during which bacteriophage homologous recombination proteins mediate annealing of ssDNAs to their genomic targets. By iteratively introducing libraries of mutagenic ssDNAs targeting multiple sites, MAGE can generate combinatorial genetic diversity in a cell population. Alternatively, MAGE can introduce precise mutant alleles at many loci for genome-wide editing or for recoding projects that are not possible with other methods. In recent technological advances, MAGE has been improved by strain modifications and selection techniques that enhance allelic replacement. This protocol describes the manual execution of MAGE wherein each cycle takes ~2.5 h, which, if carried out by two people, allows ~10 continuous cycles of MAGE-based mutagenesis per day.

INTRODUCTION

The growing availability of whole-genome sequence data sets generated by massively parallel sequencing permits detection of potential associations between genotype and phenotype¹. These advances motivate the development of high-throughput genome editing technologies to systematically elucidate the causative mutations underlying important phenotypes. Similarly, expanding protein and RNA structure databases provide valuable resources for researchers seeking to engineer macromolecules with new binding specificities². This application demands tools permitting targeted mutagenesis of the genetic loci that contribute to the structure of a ligand's binding pocket. Finally, as complex metabolic networks controlling the flow of biomolecules through cells are elucidated, efforts to alter native metabolism for engineered biosynthesis of desired compounds are rapidly growing in basic research and industrial biotechnology³. To maximize the efficiency of engineered biosynthesis, it will be necessary to recalibrate flow through each branch of the engineered pathway by targeted, multisite mutagenesis. In all three cases, there is demand for tools that leverage existing knowledge by targeting mutations to many genetic loci and by tuning mutagenesis to permit certain modifications while excluding others. Such tools would permit analysis of the effects of many mutations in one gene, one genetic pathway or across many sites in the genome.

MAGE allows an investigator to generate many independent or combinatorial genome modifications in a population of live cells⁴. Chemically synthesized ssDNAs targeting different chromosomal loci can be combined into a complex oligonucleotide (oligo) pool for simultaneous recombination into the genome (Fig. 1). The resulting population contains cells carrying different combinations of independent single-site and multisite mutations. The ability to rapidly generate combinatorial mutations and to specify either discrete or degenerate mutations at each target site is a feature that distinguishes MAGE from alternative methods. For instance, although genome editing technologies based on targeted nucleases are functional in many species, they have limited multiplexing capability and require provision of separate repair

templates for discrete mutations⁵. Directed enzyme evolution methods provide the useful feature of incorporating both random and targeted mutations, but they cannot be used to mutagenize many sites in a large genetic network⁶.

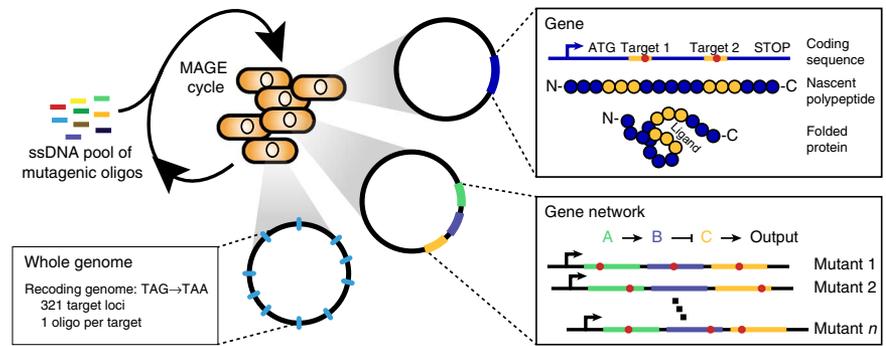
Overview of MAGE

MAGE harnesses phage homologous recombination proteins to create targeted, rapid, scarless modifications of bacterial chromosomes across many genomic loci^{7,8}. The recombination proteins used in MAGE act on synthetic ssDNA (oligos) introduced to a cell population by electroporation. These oligos are designed with 5'- and 3'-terminal homology arms that are complementary to target sequences in the genome. Once in the cell, oligos are proposed to anneal to their lagging strand targets on the bacterial chromosome as it separates into leading and lagging strands during replication, and then they are stably inherited after another round of replication⁹ (Fig. 2a). 5' and 3' homology arms flank the sequence corresponding to the specified mutation. In this region, the synthetic ssDNA can skip bases, mispair or add bases with respect to the target region, causing a deletion, mismatch or insertion, respectively (Fig. 2b).

In general, MAGE experiments can be divided into three classes, characterized by varying degrees of scale and complexity: (i) many target sites, single genetic mutations; (ii) single target site, many genetic mutations; and (iii) many target sites, many genetic mutations. In the first class, MAGE has been used to recode all 321 instances of the TAG stop codon for the synonymous TAA codon using 321 discrete ssDNAs. This project yielded a strain of *E. coli* with only 63 'active' codons and a 64th 'blank' codon available for site-specific incorporation of nonstandard amino acids^{10,11}. In the second class, MAGE could be used to explore the effects of all possible amino acid substitutions at a single target locus. In such an experiment, it would be possible to purchase a single degenerate ssDNA containing the NNN triplet at its center—such an oligonucleotide could give all possible amino acid substitutions. In the third class, MAGE has been used to construct diverse cell populations

PROTOCOL

Figure 1 | Multiplex automated genome engineering (MAGE) processes and applications. MAGE can be used to generate many mutations in genes (targets in orange), gene networks (targets in green, orange and purple) or full genomes (targets in blue). Depending on the number of ssDNAs targeted to each locus, MAGE can be used to edit the bacterial chromosome (one oligo per locus) or to generate combinatorial diversity (>1 oligo per locus). Combinatorial diversity can be explored in a population of cells that have incorporated different mutations (red) at different targets. Populations undergoing MAGE mutagenesis can be cycled iteratively until the desired level of diversification or recoding is achieved.



containing combinations of alleles across many loci involved in the biosynthesis of lycopene⁴ or aromatic amino acids¹². In this implementation, discrete oligos designed to knockout competing pathways by deletion can be mixed with degenerate oligos designed to randomize target positions in the coding sequence or regulatory regions of key pathway enzymes (Fig. 2). The highly diverse population resulting from a MAGE experiment can be used downstream to screen or select for mutants with a prescribed phenotype (e.g., overproduction of a metabolite or small molecule).

Although an automated device has been built to enable continuous mutagenesis sustaining ~10 cycles per day⁴, this protocol describes the manual process for scientists who seek to adopt the MAGE technology in their day-to-day experiments.

Development of MAGE

Although MAGE was originally developed with the goal of performing targeted codon reassignments for whole-genome recoding^{10,11}, early experiments revealed the ease of generating diversity across cell populations by mutagenesis targeting multiple loci for degenerate mutations. As such, the first demonstration of MAGE highlighted its ability to drive diversity across many genetic loci by evolving the deoxyxylulose-5-phosphate (DXP) biosynthesis pathway to enhance lycopene production⁴. Several clones capable of enhanced lycopene production were isolated from >10¹⁰ genomic variants generated in 4 d of continuous MAGE cycling (35 cycles), a striking improvement over conventional metabolic engineering approaches that take months to years⁴.

MAGE was inspired and has been enabled by three main research fields, each of which has progressed immensely over recent years.

First, extraordinary advances in *de novo* DNA synthesis mean that relatively long ssDNAs (~90 nt) can be produced at high accuracy and low cost^{13–15}. High-fidelity synthesis ensures that the mutations introduced are the ones specified during oligo design, rather than errors inherited during ssDNA chemical synthesis. Low-cost synthesis means that >25 loci can be targeted for single or degenerate mutations—potentially engendering >>10¹² unique genotypes—for <\$1,000 worth of DNA (\$37.5 per 90-mer phosphorothioated oligo at a 100-μmol synthesis scale). Second, MAGE technology is based on decades of research into the mechanisms of homologous recombination in yeast¹⁶, bacteria¹⁷ and bacteriophage¹⁸. As a result of this research, the process of homologous recombination is well understood at genetic, molecular and biochemical levels. Such knowledge has enabled rational efforts to improve the efficiency of the MAGE process by changing the target strain's genetic background or modifying synthetic ssDNAs to avoid degradation. Third, the field of directed evolution has shed light on how mutagenesis and selection or screening technologies can be combined to isolate biomolecules with enhanced activity^{19–23}. MAGE extends the work of that field by applying its lessons to the chromosomes of living bacteria, permitting parallel or continuous directed evolution of gene networks or genomes.

Workflow of a typical MAGE project

MAGE projects involve six stages (Fig. 3):

- Describe the desired phenotype or genotype. A prescribed phenotype might relate to the performance of an enzyme or full

Figure 2 | Proposed mechanism and categories of MAGE-based mutagenesis. (a) The mutagenesis oligo is bound by β-protein and anneals to the lagging strand of the replication fork, between discontinuous Okazaki fragments. At the end of replication, one wild-type (WT) chromosome and one chimeric chromosome exist. During the second round of replication, two strands of the chimeric chromosome segregate into a fully WT and a fully mutant (mut) chromosome. (b) MAGE is capable of introducing insertion, deletion or mismatch mutations. By targeting coding, regulatory or intergenic regions, these mutations can be used to modify transcription rates, translation rates, mRNA stability, enzyme activity and so on. RBS, ribosome binding sites.

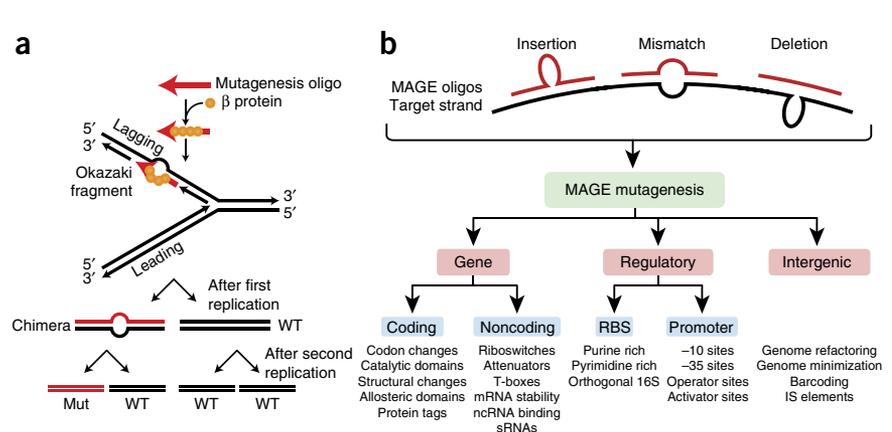
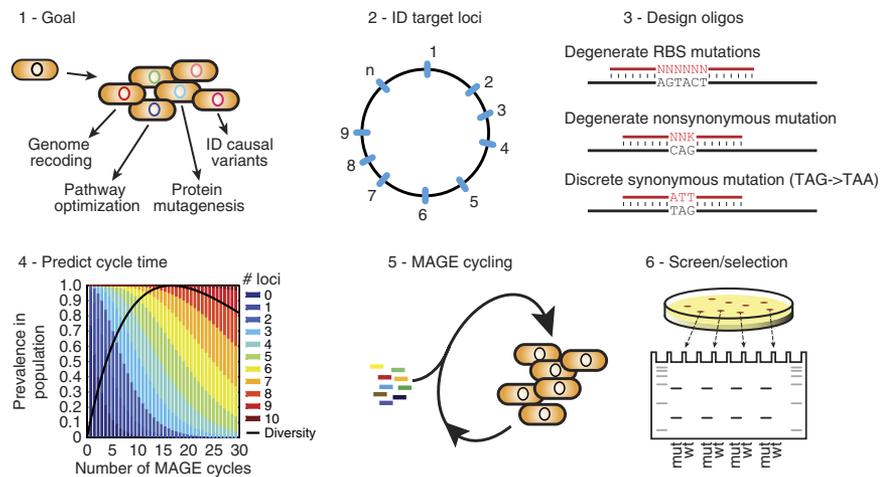


Figure 3 | The MAGE workflow. Each MAGE project involves six steps: (1) determining genome editing goals; (2) identifying target loci for mutagenesis; (3) designing mutagenic MAGE oligos; (4) predicting the number of MAGE cycles to accomplish genome editing goals; (5) performing MAGE cycling; and (6) screening or selecting for the desired genotype or phenotype.



gene network in converting the substrate to product. Moreover, a designed genome could require 10^2 – 10^{3+} specific mutations (e.g., genome recoding in which all TAG stop codons are reassigned to TAA codons^{10,11}).

• *Choose target loci.* For example, before using MAGE to optimize biomolecule production, all genes participating in a pathway should be identified. Existing computational algorithms such as Flux Balance Analysis²⁴ or Optknock²⁵ can help identify relevant loci. For each of these, transcriptional and translational control sites can be tuned up or down to modulate gene expression. Moreover, regulatory, catalytic and substrate-binding sites of a protein can be targeted to modify enzyme activity (Fig. 2). For a recoding project, all instances of the sequence to be modified can be located using bioinformatic scripts and an annotated reference genome file.

• *Design ssDNAs to modify target sites.* Genes can be knocked out or recoded by discrete deletion or substitution oligos. Alternatively, all possible substitutions can be made using degenerate oligos that target promoters¹², ribosome-binding sites²⁶, enzyme allosteric feedback sites²⁷ and so on (Fig. 2).

• *Predict the required cycle time.* More cycles increase the prevalence of allelic replacements and population diversity. Mathematical models can guide the investigator's initial choice of cycle number (Box 1 and Fig. 4), and samples of the mutagenized population can be withdrawn and assayed between cycles to track the progress of mutagenesis.

Box 1 | Modeling a population undergoing multiplexed genome engineering

A population undergoing MAGE can be treated as a collection of binary events, representing allelic replacements at the n genomic loci targeted for mutagenesis³. Neglecting off-target recombination, linkages between loci and effects on fitness, allelic replacements can be assumed to occur independently and with a fixed ARF per cycle. ARFs can be estimated for MAGE oligos using empirically determined formulae⁴⁶ or measured directly, and then used in a combinatorial model to predict the population's evolution as a function of cycle number N (Fig. 4); after N cycles, replacement with a given ARF will have occurred with probability $p = 1 - (1 - \text{ARF})^N$.

For example, in the simplest case, all n loci are targeted by nondegenerate oligos, all with the same ARF, allowing the population's evolution to be completely described by the distribution of the number of allelic replacements per clone k , which is a binomial distribution $f(k, n, p) = \binom{n}{k} (1 - \text{ARF})^k (1 - \text{ARF})^{n-k}$ (Fig. 4b; for $n = 10$ or 30 , $\text{ARF} = 3\%$, and $N = 5$ through 90). From that distribution, values important to particular experiments can be obtained, such as the cumulative distribution function $F(k; n, p) = \sum_{i=0}^k f(i, n, p)$. In terms of these common distributions, the prevalence of clones with at least k mutations is then $1 - F(k; n, p) + f(k, n, p)$ (Fig. 4a; for $k = 1$). In experiments targeting a particular genotype (e.g., recoding), the goal is to reliably obtain a clone with all loci modified ($k = n$), the prevalence of which is just $f(k = n, n, p) = (1 - (1 - \text{ARF})^N)^n$ (Fig. 4e; solid line, for $n = 10$; $\text{ARF} = 3\%$, red, or 15% , blue). A screen of s clones will find a clone that has prevalence f in the population with probability r if $f > 1 - (1 - r)^{\frac{1}{s}}$. Because f increases with cycle number, this condition determines the minimum number of cycles after which a screen would reliably (i.e., with 95% confidence given a 96-colony screen) render the desired genotype (Fig. 4e; dashed lines). Coselection can significantly speed up the evolution of this top clone (Fig. 4e; shown for a fivefold increase in ARF frequency, from 3 to 15% per locus).

For experiments targeting a particular phenotype, a criterion for cycle time is not as easily defined; each additional cycle is as valuable as the adaptive diversity that it generates. In our binomially distributed example, we can calculate the variance in the number of replacements per clone $np(1 - p)$ as a measure of overall diversity (Fig. 4c,d,f, black line). Diversity attains its maximum at $p = \frac{1}{2}$ (Figs. 4c,d, for $\text{ARF} = 3$ and 15%), but diversification experiments typically use degenerate pools that include oligos encoding the WT, thus preventing the decrease in diversity that would accompany vanishing WT alleles. The simplest degenerate case would have each locus targeted by an oligo encoding the WT allele with ARF_w and oligos encoding mutants with ARF_m ; in this case, the distribution of the number of allelic replacements per clone is still binomial, $f(k, n, p)$, but with $p = 1 - \sum_{i=0}^{\lfloor N/2 \rfloor} \binom{N}{2i} (\text{ARF}_w)^{2i} (1 - \text{ARF}_m)^{N-2i}$ (Fig. 4d, for $\text{ARF}_w = \text{ARF}_m = 3\%$ and $n = 10$). Thus, the diversity generated after many cycles by a degenerate pool has its asymptote above zero and may even approach its maximum.

PROTOCOL

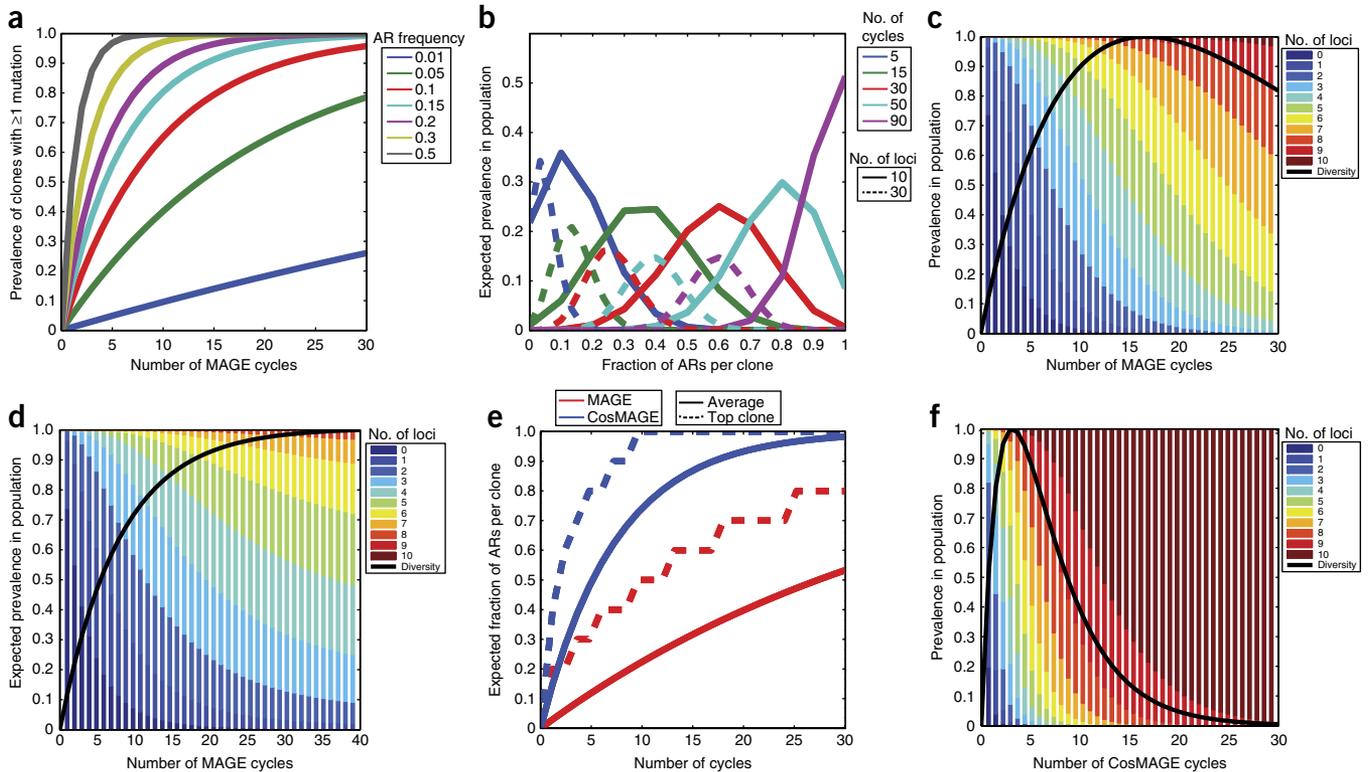


Figure 4 | Model-guided MAGE experimental design. (a) As more MAGE cycles are performed, the prevalence of clones containing any allelic replacement (AR) rises more steeply for higher ARFs (colored lines), regardless of the number of loci targeted (curves shown are for ten loci). (b) As more MAGE cycles are performed with a pool targeting many loci, the mean fraction of all targeted ARs that have been acquired per clone increases. For example, half the cells in a population will carry at least $0.6 \times 10 = 6$ replacements after 30 cycles, if targeted with a 3% ARF at each of ten loci (solid red line). (c,d) After many cycles, population diversity (here, variance in number of replacements per clone; black line) will decrease as wild-type alleles are lost from the population (c), unless the oligo pool is degenerate (d) and includes oligos encoding the wild-type allele (plots for 3% ARF at each of ten loci). (e,f) Compared with MAGE (c), cosMAGE (e,f) reduces the number of cycles required to achieve multisite mutations. For example, it would take 15 cycles to replace five out of ten targeted alleles in half the population at an ARF of 3% each, but only five cycles if that frequency were increased fivefold by coselection, to 15%. (f) By using cosMAGE, multisite mutations accumulate more rapidly than without (compare with c). However, diversity (black line) falls faster as the population rapidly converts to all mutant alleles. See **Box 1** for model description and predictions.

- **Conduct MAGE cycling.** The MAGE cycle (**Fig. 5**) involves inducing competence, electroporation to transform cells with mutagenic ssDNA and then outgrowth to permit recombination of ssDNAs with the chromosome.
- **Identify desirable clones.** For recoding projects, the desired genotype is known ahead of time and either sequencing or PCR-based methods can be used to determine whether the desired genotype has been achieved. For phenotype-oriented projects, desirable clones must be isolated from the diverse population generated by MAGE before their genotype is determined. Recombinant cells arising from MAGE mutagenesis can be assayed by the use of selective media, by visual screening or by PCR screening. Selection allows for extremely high-throughput analysis—in principle, the entire population can be assayed—but requires a selectable phenotype. A visual screen allows analysis of up to $\sim 10^4$ – 10^6 independent clones but requires a visible phenotype.

Figure 5 | The MAGE cycle. Each cycle begins with a cell population growing at mid-log phase. Cells are heat shocked to induce recombination proteins, washed to remove salts and then mixed with ssDNA and electroporated to transform. An outgrowth period allows cells to recover and mutations to be fixed. After outgrowth, the mutagenized population can be diluted, grown to mid-log phase and then fed into another round of MAGE, or it can be subjected to a screen or selection to isolate desirable genotypes or phenotypes.

A PCR screen based on primers complementary to mutant or wild-type (WT) alleles can be designed to interrogate any mutation in up to $\sim 10^2$ independent clones.

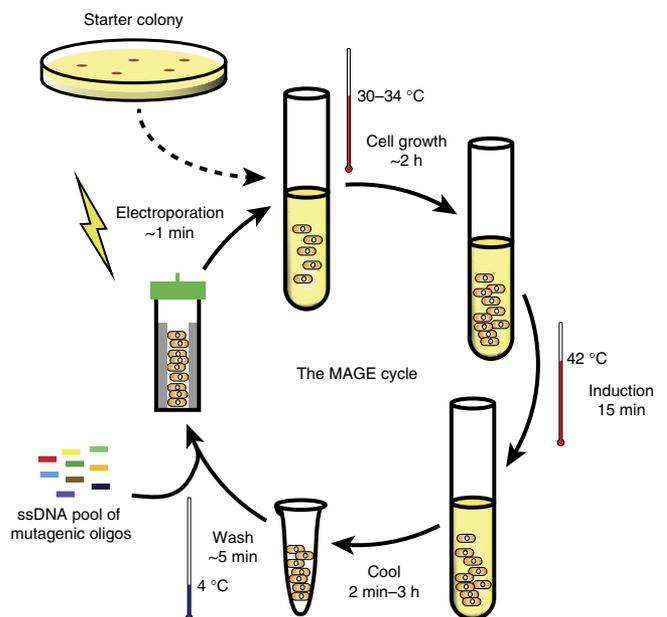
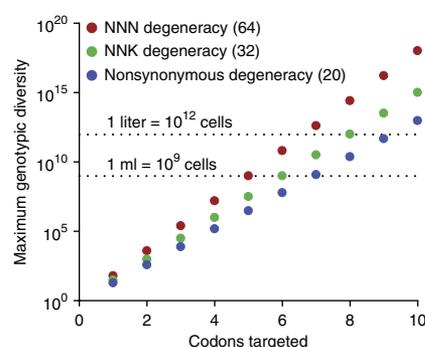


Figure 6 | Mathematical calculations for functional diversity of mutagenized protein-coding sequences. Depending on its design, a MAGE oligo can introduce all possible substitutions at a target codon (NNN degeneracy, 64 possibilities), all nonsynonymous and some synonymous substitutions (NNK degeneracy, 32 possibilities) or only nonsynonymous substitutions (nonsynonymous degeneracy, 20 possibilities). Nonsynonymous mutations are more efficient, in that fewer cells are needed to explore all possible mutations. MAGE enables complete randomization of amino acid sequence while restricting the genetic diversity to 20 codons rather than 64 codons. Therefore, targeting seven codons for all possible substitutions (red) results in at least 1,000× more possible genotypes (diversity) than if only nonsynonymous mutations are made (green). Greater genotypic diversity is inefficient because it requires a larger population to sample. Dotted horizontal lines show the maximum achievable population complexity for 1-ml and 1-liter cultures.



Advantages of MAGE

Five main features contribute to the utility of MAGE:

- **Fast and efficient.** Each round of MAGE requires ~2.5 h, after which the frequency of allelic replacement at a specific site can exceed 30% (refs. 4,10,28). Repeated MAGE cycles—wherein the population from one round proceeds into a subsequent round of oligo-based mutagenesis—yield more population diversity and lead to accumulation of genetic mutants.
- **Highly scalable.** The space of possible sequences represented in a population increases rapidly as more bases are targeted for mutagenesis (~4ⁿ where *n* is the number of targeted bases; Fig. 6). Compared with traditional targeted mutagenesis techniques, this gives the investigator access to more functional diversity for downstream analysis without requiring more time at the bench. MAGE affords a great deal of control over the extent of mutagenesis and the diversity of the cell population. For instance, an oligo pool designed to introduce only nonsynonymous mutations would allow researchers to exhaustively sample the full diversity of amino acid sequence space while minimizing the coding DNA sequence space. This feature of MAGE permits a greater number of residues to be targeted for complete randomization without increasing the size of the cell population needed to fully sample sequence diversity (Fig. 6). Similarly, a MAGE oligo pool could be designed to site-specifically introduce substitutions for a specific class of amino acids (e.g., all nonpolar residues).

- **Targeted and multiplexed.** Mutagenesis technologies generally enable targeted mutagenesis at single sites or random multisite mutations, but not both. MAGE is unique in its ability to introduce degenerate mutations at multiple genetic loci.
- **Cheap.** No expensive enzymes or buffers are required; the only capital costs are for electroporation equipment and incubators (standard laboratory equipment), and the only recurring costs are for growth media and synthetic DNA, whose price is dropping precipitously¹⁴. Moreover, DNA microchips provide an immediate source of inexpensive mutagenic oligos¹⁵.
- **Operates continuously in vivo.** This allows for MAGE experiments to proceed in the context of the normal evolutionary pressures that maintain cell viability, thereby preventing the accumulation of deleterious mutations.

Limitations of MAGE

The MAGE technology possesses four main limitations:

- **Limited portability.** Except for a recent report of similar technology in *Corynebacterium glutamicum*²⁷, MAGE has been described only in *E. coli*. Given that the proposed mechanisms underlying MAGE (i.e., mutagenic oligos that hybridize at the lagging strand during DNA replication) are conserved throughout all domains of life, and given recent reports that identified *bet* homologs that are functional as ssDNA recombinases²⁹ in both prokaryotes and eukaryotes³⁰,

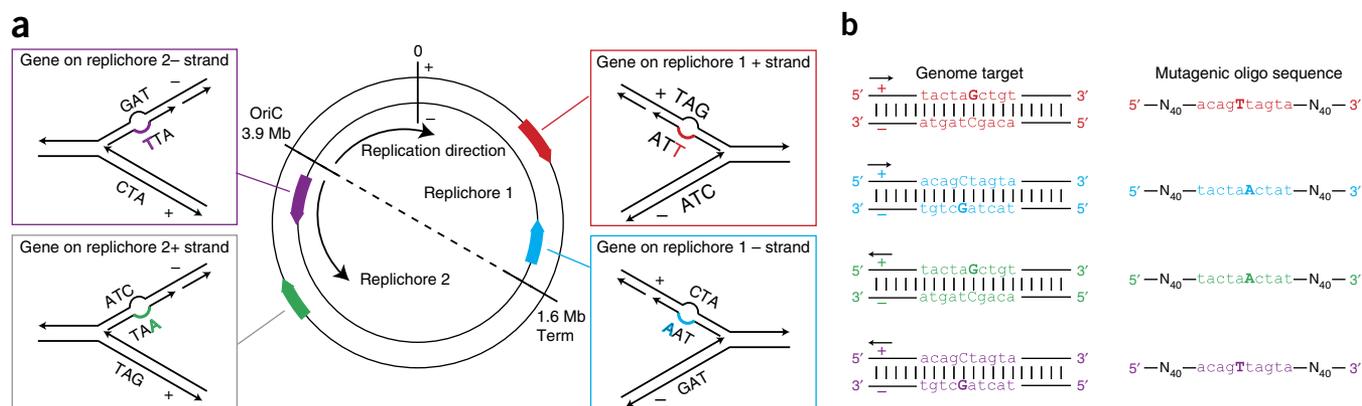


Figure 7 | Oligo design guidelines for target loci on different strands and different replichores (chromosome halves that are bidirectionally replicated). (a) MAGE allelic replacement frequency is highest when oligos anneal to the lagging strand of the replication fork as it passes through the target loci. To mutate genes on the replichore 1 (+) strand or replichore 2 (–) strand, the oligo sequence should be the reverse complement of the target gene sequence. To mutate genes on the replichore 1 (–) strand or replichore 2 (+) strand, the oligo sequence should be the same as the gene sequence. (b) Target genome sequences alongside mutagenic oligos provide explicit examples of four common scenarios depicted in a to guide the design of mutagenic oligos at the lagging strand of DNA synthesis. Bold letters indicate targeted bases.



PROTOCOL

Figure 8 | Oligo design parameters affecting allelic replacement frequency. (a–f) Length (a); folding energy (b); phosphorothioate backbone modifications (c); concentration (d); insertions and mismatches (e); and deletions (f). Each parameter is plotted on the x axis, and ARF—scored either by a selectable or a screenable marker—is plotted on the y axis. In general, we use 90 mer oligos delivered at 1 μM (singleplex) or 0.5–10 μM (multiplex) final concentration with two phosphorothioate (PO, indicated by *) bonds at the 5' end and with a minimum of predicted folding energy (≥ -12 kcal/mol). Error bars indicate s.d.

genome editing via MAGE could be developed in diverse organisms.

- **Undersampling.** The capacity of MAGE to generate cell populations containing vast genetic diversity outstrips the ability of current methods to examine the genotype or phenotype of clones. As a result, much of the diversity generated in a MAGE experiment is never sampled, and new high-throughput ways of analyzing complex cell populations are required to fully explore MAGE-diversified populations.

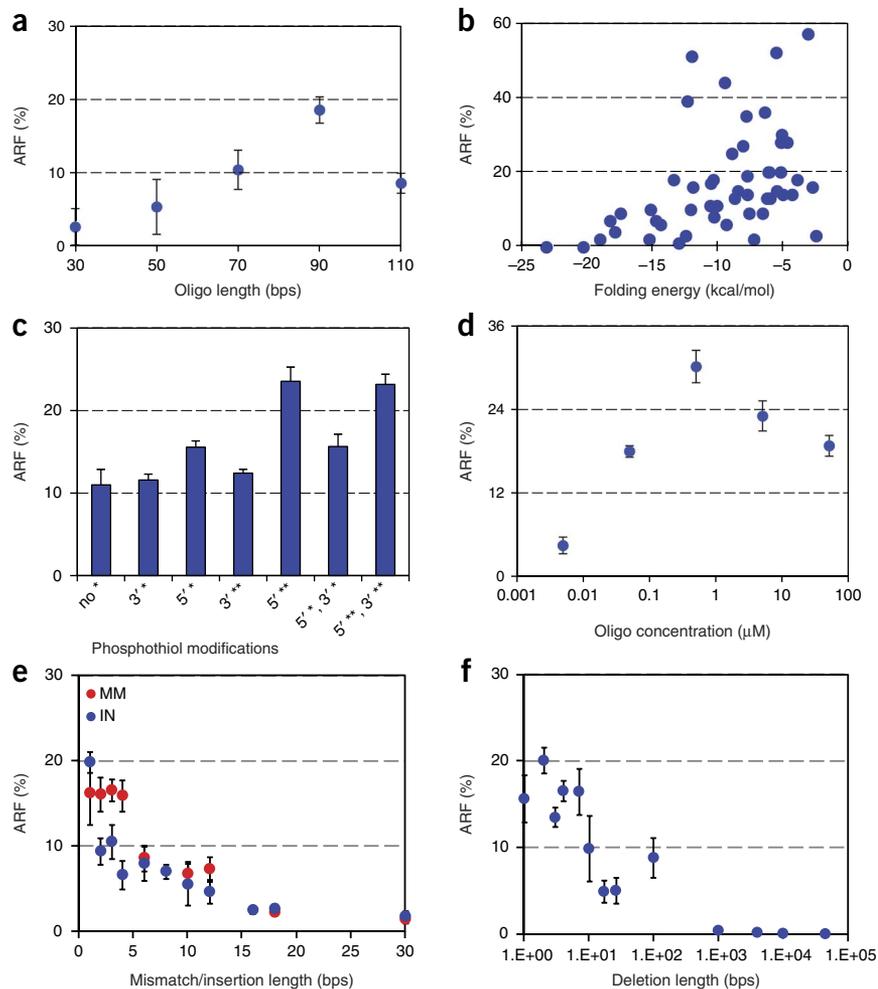
- **Requirement for detailed prior knowledge.** Target selection for MAGE mutagenesis is dependent on *a priori* knowledge of biological systems. Therefore, it is difficult to design MAGE oligos for processes that

are poorly understood. In such instances, strategies that combine MAGE with random mutagenesis methods could expedite isolation of organisms with prescribed phenotypes.

- **Low frequency of multisite mutants.** Although the MAGE allelic replacement frequency (ARF) for single loci is high, that frequency is shared by multiple loci targeted by a complex pool of oligos. For large pools targeting many sites, the replacement frequency for any single site can be low. Strategies that increase the concentration of mutagenic oligos in the cell, facilitate hybridization of oligos to targets or avoid the cell's error correction mechanisms all could improve aggregate ARF. This would help increase the rate at which a population undergoing MAGE accumulates multisite mutations.

Experimental design

Target selection. Although MAGE enables the construction of extremely diverse cell populations, targeting too many sites can engender a genetic space far larger than downstream assays could handle. For instance, full degeneracy mutation at ten codons gives $>10^{18}$ possible genotypes (i.e., $(4^3)^{10}$) or $>10^6$ liters of confluent *E. coli* (10^9 cells per ml $\times 10^3$ ml/liter $\times 10^6$ liters) if each genotype was represented by a single cell. To reduce the genetic space, it is necessary to prioritize important sites during experimental design. For MAGE-based diversification of a single protein whose crystal structure is known, mutagenic oligos can be targeted to bases encoding amino acids that are



most closely associated with the enzyme's structure or function. For diversification of a gene network, computational models of gene regulation can help identify loci that have a major role in the network's function.

MAGE oligo design. The mechanism hypothesized to underlie MAGE involves hybridization of ssDNA at the lagging strand of the replication fork (Fig. 2a). Supporting this model, oligos that anneal to the leading strand display markedly reduced ARF³¹. Moreover, strain modifications that increase the distance between Okazaki fragments on the lagging strand increase ARF³². Therefore, it is important to consider strandedness when designing oligos. As replication of the *E. coli* chromosome starts from a single origin and proceeds bidirectionally, it is possible to identify leading or lagging strands at each locus. Figure 7 shows how the target strand changes according to the genome coordinate and therefore guides how the oligo's strandedness should be chosen. Other important parameters in oligo design (Fig. 8) include:

- **Length.** Optimal mutagenic oligo length is 90 nt (ref. 4; Table 1). Below 40 nt, the ARF falls rapidly probably because short ssDNAs do not bind well to the β -protein⁹ and possess reduced homology to the chromosome. The frequency increases up to 90 nt and then decreases, either owing to errors in synthetic DNAs or increased likelihood of secondary structure formation that prevents oligo binding at its genomic target (Fig. 8a).
- **Structure.** ssDNAs with higher predicted ΔG score³³ recombine

TABLE 1 | Optimal range of MAGE oligo design parameters.

Parameter	Optimum
Length	90 nt
Concentration	0.5–1 μM
Phosphorothioate bonds	2 at 5' terminus
Folding energy	≥–12 kcal/mol
Mistarget score	≤600 nt

at higher frequencies likely because unstructured ssDNAs can easily hybridize with target genomic loci⁴ (Fig. 8b).

- **Backbone modifications.** Phosphorothioate bonds connecting the two 5'-most bases increase ARF, probably by preventing exonuclease-mediated degradation of ssDNA⁴ (Fig. 8c).
- **Mismatch position.** ARF is highest when the modification to be introduced is >10 nt from the oligo's 5' and 3' termini. Nearer the center of the ssDNA, the mutation is most likely to be stabilized by homology arms and least susceptible to exonucleases, provided that the oligonucleotide has been designed with limited secondary structure.
- **Off-target score.** ssDNAs with significant homology to multiple genetic loci will recombine with lower frequency, presumably because off-target sites compete with target sites to anneal the ssDNA. Specifically, oligos with >600 bp of off-target homology—which are determined by summing the lengths of all homologous regions identified by a BLAST search against the *E. coli* genome—showed a 32% decrease in ARF relative to those with <600 bp of off-target homology¹⁰.

- **Mismatch identities.** The mismatch repair pathways are highly sensitive to certain mismatches and relatively insensitive to others. For example, the T-T, A-G and C-C mismatches are bound poorly by MutS, whereas the G-G mismatch is readily recognized³⁴. In a methyl mismatch repair (MMR⁺) strain, repair can be avoided by selecting weakly recognized mismatches. Repair can also be avoided by including a string of synonymous mismatches near the desired nonsynonymous mismatch³⁵.
- **Base modifications.** ssDNAs containing modified bases can cause nearby mismatches to escape detection in MMR⁺ strains. For instance, 5-methyl-C and fluoro-uridine can increase ARF when placed within 5 bp of a mismatch position, although the cost of such modifications could be prohibitive³⁶.

The MAGE cycle. Given that the MAGE process is cyclical in nature, it can be applied iteratively to increase the diversity of a cell population or to increase the frequency of desired modifications. Each cycle of MAGE entails induction of recombination proteins, electroporation to introduce ssDNA and outgrowth to allow recovery and segregation of the mutation (Fig. 5; Steps 1–12 of the PROCEDURE). The cycle time is roughly 2.5 h, with additional time required at the beginning and end of cycling to grow isogenic starter cultures and to characterize mutations, respectively. It is possible for one person to perform five MAGE cycles in a typical workday with incubation between cycles leaving time for other tasks. Two people can perform ten continuous cycles of MAGE in a full day. Because the hands-on time for each cycle is a small fraction of total cycle time, independent MAGE experiments can be synchronized to enable parallel cycling. Negative (no ssDNA added) and positive (addition of a confirmed oligo with a screenable phenotype, e.g., lacZ_{off} from Table 2) controls can be included to aid in troubleshooting.

TABLE 2 | Some commonly used mutagenesis oligos.

Name	Sequence (5'–3')
bla _{on}	gcgaaaactctcaaggatcttaccgctgttgagatccagttcgat G taaccactcgtgcaccaactgatcttcagcatctttacttt
bla _{off}	gcgaaaactctcaaggatcttaccgctgttgagatccagttcgat T taaccactcgtgcaccaactgatcttcagcatctttacttt
cat _{on}	gcatcgtaaagaacattttgaggcatttcagtcagttgctcaatg T acctataaccagaccggttcagctggatattacggccttttaaa
cat _{off}	gcatcgtaaagaacattttgaggcatttcagtcagttgctcaatg A acctataaccagaccggttcagctggatattacggccttttaaa
kan _{on}	cgcgattaaattccaacatggatgctgatttatatgggtataaat G ggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgct
kan _{off}	cgcgattaaattccaacatggatgctgatttatatgggtataaat A ggctcgcgataatgtcgggcaatcaggtgcgacaatctatcgct
tolC _{on}	tgggttcagttcgttgagccagccgagaacctgatgcaagtta T cagcaagcacgccttagtaaccgggaattgcgtaagtctgccgc
tolC _{off}	tgggttcagttcgttgagccagccgagaacctgatgcaagtta A cagcaagcacgccttagtaaccgggaattgcgtaagtctgccgc
lacZ _{on}	atgattacggattcactggcctgcttttacaacgctgctgactg G gaaaacctggcgttacccaactaatcgcttcagcacatccc
lacZ _{off}	atgattacggattcactggcctgcttttacaacgctgctgactg A gaaaacctggcgttacccaactaatcgcttcagcacatccc
mutS _{on}	atcacaccccttaataatcagggaaaccggacataacccat G agtgaatagaaaatttcgacgccatacggccatgatgcagcagta
mutS _{off}	atcacaccccttaataatcagggaaaccggacataacccat C agtgaatagaaaatttcgacgccatacggccatgatgcagcagta

The lacZ_{off} oligo can be used as a positive control to verify proper execution of the MAGE protocol; the mutS_{on} oligo can be used to restore MMR in ΔmutS strains at the end of MAGE mutagenesis. Capitalized, boldface letters indicate mutation sites.



TABLE 3 | Strains used in MAGE.

Name	Genotype	Description
MG1655	F ⁻ λ - <i>ilvG-rfb-50 rph-1</i>	Unmodified ancestor strain, approximates wild type
EcNR1	MG1655 Δ (<i>ybhB-bioAB</i>)::[λ cI857 N(<i>cro-ea59</i>):: <i>tetR-bla</i>]	Insertion of mutant prophage with recombination genes
EcNR2	EcNR1 Δ <i>mutS::cat</i>	Deletion of <i>mutS</i> with <i>cat</i> marker impairs MMR
Nuc5 ⁻ (ref. 40)	EcNR2 Δ [<i>xonA, recJ, xseA, exoX, redα</i>]	Markerless nuclease deletions prevent oligo degradation
EcNR3 (ref. 32)	Nuc5- <i>dnaG.Q576A</i>	Primase allele increases Okazaki fragment size
EcM2.1 (ref. 43)	EcNR3 1255700:: <i>tolQRA</i>	Nuclease deletions, primase allele and extra <i>tolQRA</i> operon to prevent dual colicin E1 and SDS resistance
EctMMR ³⁹	MG1655 <i>mutS</i> (A134V) <i>mutL</i> (G62S)	Temperature-sensitive MMR alleles to interrupt mismatch repair only during MAGE cycling

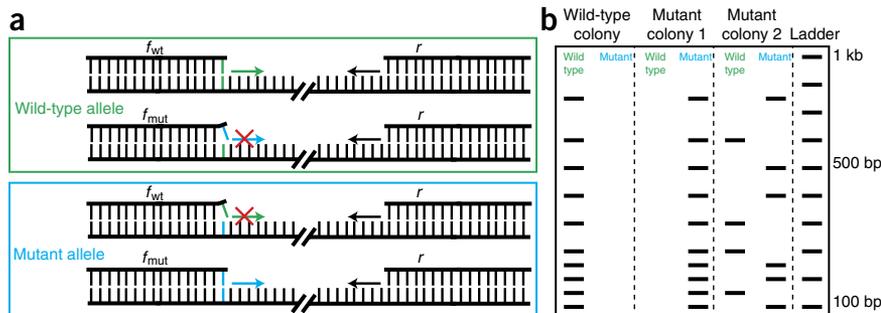
Preparation of MAGE-competent cells. Before mutagenic ssDNA is transformed, the target cell population must be made competent. First, recombination proteins—which are too toxic to be expressed constitutively—must be transiently induced (Steps 1–5 of PROCEDURE). The endogenous repair proteins of many bacteria possess some activity toward ssDNA³⁷, but this basal capacity is too low for MAGE. Instead, MAGE strains overexpress *exo*, *bet* and *gam* genes from an integrated and heat shock-inducible λ -prophage^{4,10,38}. The WT λ -phage uses these genes to facilitate replication and genetic exchange with its host. For MAGE, *bet* is the most important of the three—it is a *recA* homolog with ssDNA-binding and recombinase functions. The *bet* recombinase coats ssDNAs introduced by electroporation to protect them from degradation, prevent self-folding and facilitate target sequence hybridization at the replication fork¹⁸. The 5' exonuclease *exo* and the RecBCD inhibitor *gam* are required for dsDNA recombination, but they have been shown to be dispensable in ssDNA recombination³⁸. These three genes are under control of the phage *pL* promoter, which in turn is regulated by the temperature-sensitive mutant repressor *cI857*. This repressor is destabilized at 42 °C, losing its ability to bind *pL* operator sites and allowing strong transcription of the *exo-bet-gam* operon. A 15-min heat shock in a 42 °C shaking water bath results in a pulse of *bet* expression sufficient for high-frequency allelic replacement. Second, cells must be washed twice with chilled distilled water to remove electrolytes (Steps 6–8 of PROCEDURE). Electroporation is an extremely efficient method of DNA delivery, but it requires

low-conductivity cell slurry (otherwise, electrical arcing will occur and cause cell death). Third, because the MAGE process is replication dependent, efficient allelic replacement requires that cells be collected for electroporation at mid-log phase (OD₆₀₀ ~0.5) when cells are growing fastest with active replication forks. In early growth, cell density is too low, and in later stages of growth cells enter a stationary phase during which replication slows down markedly.

Strain modifications to enhance MAGE. After a mutagenic oligo hybridizes to its chromosomal target, a mismatch exists at the target site. The mutation will not be inherited if the cell's MMR machinery is allowed to correct it before a subsequent round of replication. To increase MAGE ARF by avoiding MMR, a *mutS*-deficient strain is commonly used (Table 3). This strain does not produce the protein that binds mismatched positions to identify them for repair. In an MMR-deficient strain, the improvement in oligo-mediated allelic replacement is roughly 100-fold (refs. 4,38). The background mutation rate—i.e., at positions not targeted by ssDNAs—is increased ~50-fold (refs. 10,11). A recently described strain with temperature-sensitive expression of MMR machinery permits high-efficiency MAGE with a reduced rate of off-target mutation³⁹. Alternatively, the reduced ARF seen in MMR⁺ cells can be overcome by careful selection of modifications in the mutagenic oligo, as MMR detects some mismatches less effectively³⁴. Other strain modifications increase the frequency of recombination without influencing the stability of mismatches. Knockouts

Figure 9 | Genotype assays by MASC-PCR.

(a) *f*_{mut} and *f*_{WT} primers share a common *r* primer and are identical, except for the 3'-terminal base—which can either anneal to the mutant or the wild-type allele (respectively). At the optimal *T*_m, extension occurs only if the forward primer is an exact match for its target site. (b) Examples of binary MASC-PCR results. Each screened colony is interrogated by two multiplex reactions: one with *f*_{WT}/*r* pairs for ten sites, and the other with *f*_{mut}/*r* pairs for ten sites. Screened colonies are unmodified at all ten target sites (wild-type colony), modified at all ten target sites (mutant colony 1) or modified at some but not other target sites (mutant colony 2). Each band corresponds to a primer set that interrogates a specific locus and that is designed to produce amplicons of a characteristic length.



Box 2 | Optimizing PCR conditions for MASC-PCR ● TIMING 1 d

1. Run a temperature gradient PCR to confirm the T_m for MASC-PCR primers. Each set of three primers requires a gradient of eight reactions each for f_{mut-r} and f_{WT-r} pairs. For a single 20- μ l reaction, mix the following (for multiple PCRs, prepare a master mix and dispense aliquots, accordingly):

Component	Volume (μ l)	Final
2 \times KAPA 2G fast multiplex readymix	10	1 \times
Forward primer (WT or mut) (20 μ M)	1	0.2 μ M
Reverse primer (20 μ M)	1	0.2 μ M
1:20 DI H ₂ O dilution of an ancestor strain (WT at all targeted loci)	1	Trace amount
Sterile DI H ₂ O	7	

2. Prepare eight reactions for each f_{mut-r} and f_{WT-r} pair, and then run the following PCR program:

Cycle	Denature	Anneal	Extend	Hold
1	95 °C for 5 min			
2–25	95 °C for 30s	gradient (e.g., 62–67 °C) for 30s	72 °C for 30s/kb–60s/kb	
26		72 °C for 10 min	4 °C	

3. Visualize 4 μ l of each PCR product on a 1.5% (wt/vol) agarose gel. At optimal T_m , amplicons will appear at the expected length only in the f_{WT-r} reaction. Amplicons should be absent from the f_{mut-r} reaction at optimal T_m , although at lower temperatures nonspecific amplification is expected. Confirm that the optimal T_m for each primer set in the pool is sufficiently close (± 3 °C) to allow multiplexing.

▲ **CRITICAL STEP** Failure to determine optimal T_m will prevent reliable genotyping of isolated colonies.

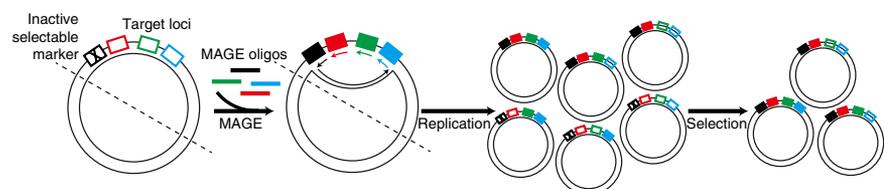
of nucleases can increase the ARF by preventing breakdown of oligos that have entered the cells⁴⁰. Modifications to the primosome machinery can increase ARF by exposing longer stretches of the lagging strand for hybridization with mutagenic oligos³².

Selection and screening. Recombinant cells from MAGE or coselection MAGE (cosMAGE; see ‘coselection MAGE’ below) experiments targeting selectable markers or a visible phenotype can be identified by plating them on appropriate selective or differential media^{4,12}. However, because most mutations do not present screenable or selectable phenotypes, multiplex allele-specific colony PCR (MASC-PCR) can be used to simultaneously interrogate the genotypes of many mutagenized loci^{10,11} (Fig. 9 and Box 2). MASC-PCR affords single base-pair resolution and permits the detection of single-nucleotide polymorphisms. For each targeted locus, three primers are designed: (i) a forward primer specific to the WT (f_{WT}) sequence, (ii) a forward primer specific to the mutant (f_{mut}) and (iii) a reverse primer (r) common to both. The two forward primers only differ at their 3'-terminal bases, allowing for discrimination of SNPs with an allele-specific PCR. Two MASC-PCRs are required to screen each colony: one to assay the WT genotype using f_{WT} and r primers, and the other to assay the mutant genotype using f_{mut} and r primers. If the colony contains the mutant allele, an amplicon will only be produced by the f_{mut} and r primers but not the f_{WT} and r primers.

For the WT allele, only f_{WT} and r primers will produce an amplicon. At least ten loci can be interrogated in a single reaction by designing each primer set to produce amplicons of different length (e.g., 100, 150, 200, 250, 300, 400, 500, 600, 700 and 850 bp). All the primers are designed with the same T_m (melting temperature, ~ 62 °C); however, the optimal T_m for each MASC-PCR primer pool should be empirically determined by a gradient PCR (Box 2). At the pool's precise T_m , MASC-PCR provides a binary result allowing unambiguous identification of mutant or WT clones. However, at temperatures below the T_m , amplicons can appear in both mutant and WT reactions. Conversely, above the T_m , amplicons are occasionally absent from both reactions.

Coselection MAGE. When a single or a few loci are targeted, the ARF by MAGE is sufficiently high to screen for the desired genotype. However, the aggregate replacement frequency is distributed across many genomic target sites when a complex pool of oligos is used. For example, when targeting three loci, one would expect only 0.1% triple mutants after one round of MAGE (assume 30% aggregate frequency for the pool = 10% per target site; assume target site independence, 10% frequency at each site: $0.1^3 = 0.1\%$). The generation of a population that samples combinatorial diversity across many sites requires either many cycles of MAGE or a higher rate of incorporation per locus. CosMAGE overcomes this challenge by including an oligo that activates a silenced selectable marker (safe insertion regions for cosMAGE marker integration

Figure 10 | Mechanism for coselection MAGE. Target-site oligos (red, green and blue) and the selectable marker oligo (black) are mixed into a pool and introduced into the cell population. Cells that incorporate the selectable marker oligo are likely to incorporate oligos used to mutagenize nearby target sites. Selection eliminates unmodified cells, leaving only cells enriched for multisite mutations. Inactive selectable marker is shown as a black rectangle with an 'X' inside; different target loci are shown as red, green and blue empty rectangles. Mutated sites are shown as solid rectangles in corresponding colors.



PROTOCOL

listed in **Supplementary Table 1**) among the pool of oligos that mutate target loci²⁸. The selection oligo enhances the incorporation of the mutagenic oligos in two important ways. First, selection enriches for cells that have undergone an oligo-mediated recombination event. Second, recent work showed that selected clones are highly enriched for incorporation of mutagenic oligos within 500 kb of the selectable marker locus²⁸. Multiple oligos in close proximity cosegregate at high frequency during cosMAGE, which increases the probability of integration into the same daughter genome (**Fig. 10**). To enhance multiple incorporation of

mutagenic oligos, only a low concentration of the selectable marker oligo (i.e., 1% of the total MAGE oligo pool) is required.

CosMAGE greatly enhances the ARF of multiplex loci. To target eight sites by conventional MAGE, only one colony in 1.5×10^7 would contain six or more mutations, given an oligo pool aggregate replacement frequency of 30% (3.75% per target site)²⁸. CosMAGE improves the ARF of each target site by around four-fold (to 15.6%; ref. 28), so that one colony in 10^4 can achieve the same result. This enhancement substantially reduces the screening required to isolate a desired genotype. Furthermore, by minimizing

Box 3 | CosMAGE cycling ● TIMING 2 d to 2 weeks

To use *tolC* as a dual selectable marker for cosMAGE, start from a strain with the active *tolC* allele and whose genotype has been optimized for repeated on- and off-selection⁴³. The starting strain's *tolC* allele can be assessed by growth in SDS: resistance is conferred only by the active (WT) allele. A control strain with active TolC can also be used to verify stocks of colicin E1, as this strain should be sensitive to the protein. The oligo pool used during odd MAGE cycles (1, 3, 5...) will include an oligo designed to introduce a nonsense mutation in *tolC*. Conversely, the oligo pool used during even MAGE cycles (2, 4, 6...) will include an oligo designed to revert the nonsense mutation introduced during odd cycles.

1. Follow Steps 1–12 of the main PROCEDURE. For odd-numbered cycles of MAGE, supplement the cell/DNA slurry at Step 8 with 0.05 μM final concentration of *tolC_off* oligo. Perform negative selection (option A) to enrich cells that have acquired the *tolC* nonsense mutation. For even-numbered cycles of MAGE, supplement the cell/DNA slurry at Step 8 with 0.05 μM final concentration of *tolC_on* oligo. Perform positive selection (option B) to enrich cells that have reverted to the active *tolC* allele.

▲ **CRITICAL STEP** Do not change the normal 0.5–10 μM final concentration of target loci oligos at Step 8.

? TROUBLESHOOTING

(A) Negative selection for odd cycles of MAGE ● TIMING 1 d to 1 week

(i) Outgrow cells in a shaking incubator at 300 r.p.m. between 30 and 34 °C for at least 6 h for complete turnover of the TolC outer membrane protein. Prepare cell culture of the ancestor strain with functional TolC as negative control.

▲ **CRITICAL STEP** Insufficient outgrowth will not eliminate residual TolC protein in mutant cells leading to failure of negative selection.

(ii) Dilute the cells 1:100 and grow them to early-log phase ($\text{OD}_{600} \sim 0.2$). Thaw colicin E1 on ice before use.

(iii) Mix 30 μl of colicin E1, 30 μl of recovered culture and 3 ml of LB for each strain; mix well by pipetting, and then add 150- μl volumes into a 96-well culture plate (~20 wells per sample).

(iv) Incubate the plate in a shaking plate reader for 6–12 h at 30–34 °C with 300 r.p.m. orbital shaking and OD_{600} measurements every 10 min to monitor cell growth. Stop the experiment when samples mutagenized by the *tolC_off* oligo grow to mid-log (OD_{600} 0.4–0.6), but the control strain (containing active *tolC* allele) remains suppressed. The selected samples are subjected to the next MAGE cycle (Step 3 of the main PROCEDURE) or mutants are isolated (step 1A(v) of this box).

▲ **CRITICAL STEP** Longer selection time will cause degradation of colicin E1, loss of selection and a high false-positive rate.

Once this begins to happen, control wells containing isogenic active TolC cells will start to grow despite the presence of colicin E1.

(v) Make 1:10 serial dilutions in LB and spread 50 μl of each diluted culture on both LB and LB-SDS plates.

(vi) Incubate the plates at 30–34 °C overnight until colonies appear.

(vii) Count the colonies from each plate to determine negative selection efficiency (number of colonies on the LB-SDS plate divided by the number of colonies on the LB plate).

(viii) Pick several colonies to grow each in both LB and LB-SDS liquid media in a shaking incubator at 300 r.p.m. between 30 and 34 °C for 6–12 h. These cultures can be frozen, used for Sanger sequencing or used for downstream analysis.

(B) Positive selection for even cycles of MAGE ● TIMING 1 d to 1 week

(i) Recover cells in a shaking incubator at 300 r.p.m. between 30 and 34 °C for at least 3 h to permit full expression of TolC.

(ii) Inoculate 30 μl of each mutagenized sample into 3 ml of LB for a 1:100 dilution. As a negative control, inoculate an isogenic culture with the inactive *tolC* allele 1:100 into 3 ml of LB. Supplement each tube with 0.005% SDS (final concentration).

(iii) Allow the cells to grow for 3–12 h in a shaking incubator at 30–34 °C. Stop the experiment once it is clear that mutagenized cultures have grown to mid-log phase (OD 0.4–0.6), but the negative control strain remains suppressed. The selected samples are subjected to the next MAGE cycle (Step 3 of the main PROCEDURE) or mutants are isolated (next step).

▲ **CRITICAL STEP** Longer selection time will cause degradation of SDS, loss of selection and a high false-positive rate.

(iv) Make several 1:10 serial dilutions in LB liquid medium and spread 50 μl of each diluted culture on LB and LB-SDS plates.

(v) Incubate the plates at 30–34 °C overnight until colonies appear.

(vi) Count viable cells from each plate to determine positive selection efficiency (number of cells on LB-SDS plate/number of cells on LB plate).

(vii) Pick several colonies to grow each in both LB and LB-SDS liquid media in a shaking incubator at 300 r.p.m. between 30 and 34 °C for 6–12 h. These cultures can be frozen, used for Sanger sequencing or used for downstream analysis.

the number of cycles required to obtain highly mutagenized cells, cosMAGE limits the accumulation of background mutations in *ΔmutS* strains.

Through the use of dual-selectable markers (e.g., *tolC*⁴¹ and *galK*⁴²), cosMAGE can be extended to continuous cycling, further enhancing the ability to accumulate many mutations. For example, the outer membrane porin *tolC* confers resistance to SDS but sensitivity to colicin E1 (a small toxic peptide). Strains with a nonsense allele of *tolC* lose SDS resistance but gain colicin E1 resistance. In cycled cosMAGE, the *tolC_off* oligo (Table 2) is used to introduce a nonsense mutation, and then in a subsequent cycle the *tolC_on* oligo is used to revert *tolC* to the WT (Box 3). This cycle can be repeated for as many

rounds as are necessary to accumulate several target-site mutations in a single chromosome. In a proof-of-principle experiment, *tolC*-based dual selection was used to isolate a colony with 80 target loci mutations after only 18 cosMAGE cycles²⁸. Long-term positive/negative cycling for coselection carries the risk of accumulating mutations that render the selectable gene nonfunctional. For instance, during *tolC* coselection, strains can acquire mutations that cause simultaneous SDS and colicin E1 resistance. In recent advances, we show that such problems can be avoided by strain and protocol modifications⁴³ that decrease the frequency of SDS and colicin E1-resistant clones, thereby preserving the function of the *tolC* gene for more cycles of coselection.

MATERIALS

REAGENTS

- Appropriate *E. coli* strain (see Reagent Setup and Table 3)
- User-defined synthetic ssDNA oligos for MAGE; see Table 2 for some commonly used mutagenesis oligos (standard purification; Integrated DNA Technologies)
- User-defined PCR primers for MASC-PCR (standard purification; Integrated DNA Technologies)
- Sterile distilled H₂O, chilled on ice
- Tryptone (American Bioanalytical, cat. no. AB02031)
- Yeast extract (American Bioanalytical, cat. no. AB01208)
- Sodium chloride (NaCl; American Bioanalytical, cat. no. AB01915)
- Agar (American Bioanalytical, cat. no. AB01185)
- NaOH, 10 M, made from NaOH pellets (JT Baker, cat. no. 3722)
- 2× Kapa 2G fast multiplex readymix (KAPA Biosystems, cat. no. KK2602)
- Agarose (American Bioanalytical, cat. no. AB00972)
- DNA ladder (New England Biolabs, cat. no. N3200S)
- Kanamycin (American Bioanalytical, cat. no. AB01100)
- Carbenicillin (American Bioanalytical, cat. no. AB00285)
- Tetracycline (American Bioanalytical, cat. no. AB02024)
- Zeocin (Life Technologies, cat. no. R25001)
- Spectinomycin (MP Biomedicals, cat. no. 0215206701)
- Gentamicin (MP Biomedicals, cat. no. 0219005705)
- SDS (American Bioanalytical, cat. no. AB01922)
- Colicin E1 (see previous papers for preparation)^{44,45}
- Glycerol (American Bioanalytical, cat. no. AB00751)
- Nucleic acid stains: ethidium bromide (Invitrogen, cat. no. 15585011), SYBR Gold (Invitrogen, cat. no. S-11494) or GelGreen (Biotium, cat. no. 41004)

EQUIPMENT

- Stationary incubator set at 30–34 °C (Thermo Scientific, cat. no. 4359)
- Shaking incubator set at 30–34 °C (Thermo Scientific, cat. no. 4359)
- Shaking water bath set at 42 °C (Thermo Scientific, cat. no. 4303)
- Stationary water bath set at 50 °C (Thermo Scientific, cat. no. 18002AQ)
- Insulated ice bucket (Fisher Scientific, cat. no. 1167673)
- Microcentrifuge kept at 4 °C (Eppendorf, cat. no. 5415C)
- Microcentrifuge kept at room temperature (~25 °C; Eppendorf, cat. no. 5424)
- Electroporator (Bio-Rad, cat. no. 1652662, or Harvard BTX, cat. no. 45–2006)
- Electroporation cuvettes with 1 mm (Bio-Rad, cat. no. 165-2089) or 2 mm (Bio-Rad, cat. no. 165-2086) gaps
- Thermal cycler (Bio-Rad, cat. no. 185-1196)
- PCR tubes, 0.2 ml (Bio-Rad, cat. no. TLS0901)
- Cap strips for PCR tubes (Bio-Rad, cat. no. TCS-0801)
- 96-well PCR plates (Thermo Scientific, cat. no. AB-0849)
- Plastic films for 96-well PCR plates (Bio-Rad, cat. no. MSB-1001)
- Single-channel pipettes for 0.5–1,000-μl volumes (Rainin) and sterile filter tips
- Multichannel pipettes for 0.5–200-μl volumes (Rainin)
- Agarose gel electrophoresis apparatus (Bio-Rad)

- Gel documentation system (Bio-Rad, cat. no. 1708195)
- Sterile 1.7-ml microcentrifuge tubes (GeneMate, cat. no. C-3620-4)
- Sterile 14-ml culture tubes (BD Falcon, cat. no. 352059)
- Sterile Petri dishes (Corning, cat. no. 351058)
- Sterile syringe filter units (Millipore, cat. no. SLGP033RS)
- Sterile 10-ml syringes (BD, cat. no. 309604)
- Sterile 50-ml vacuum filter units (Millipore, cat. no. SCGP00525)
- Sterile 2-ml cryotubes (Corning, cat. no. 420917)
- Toothpicks, nonsterile (VWR, cat. no. 500029-808)
- Glass beads, nonsterile (Sigma, cat. no. Z265942)
- Sterile 96-well flat-bottom culture plates (Corning, cat. no. 3596)
- Sterile reservoirs (Corning, cat. no. 4870)
- Spectrophotometer (Biochrom, cat. no. 80-5000-00)
- Spectrophotometry cuvettes (Brand-Tech Scientific, cat. no. 759086D)
- Plate reader for negative selections (optional; Bio-Tek, cat. no. 11120531)

REAGENT SETUP

***E. coli* strain** MAGE is performed in strain EcNR2, an *E. coli* MG1655 derivative that contains a modified λ-Red prophage to provide recombination proteins and a knockout mutation in *mutS* to prevent mismatch repair⁴. Strains related to EcNR2 with modifications in replication or nuclease genes can improve MAGE ARF (Table 3). To isolate single colonies, streak the bacterial strain that will be used for MAGE. Use a Luria-Bertani (LB, Lennox) plate supplemented with antibiotics selective for the streaked strain to avoid contamination.

TABLE 4 | Antibiotic concentrations to use for genome integrated selectable markers.

Antibiotic	Concentration (μg/ml)
Ampicillin	50
Carbenicillin	50
Chloramphenicol	20
Kanamycin	30
Spectinomycin	95
Tetracycline ^a	12
Zeocin ^{a,b}	10
Gentamicin	10
SDS	0.005% (vol/vol)

^aLight sensitive. ^bSalt and pH sensitive.

PROTOCOL

LB broth Combine 10 g of tryptone, 5 g of yeast extract and 5 g of NaCl in 1 liter of water, and adjust the pH to 7.5 by adding ~200 μl of 0.1 M NaOH. Store the medium at room temperature for up to 3 months. **▲ CRITICAL** LB medium must be autoclaved for sterility.

Antibiotic-supplemented LB broth If antibiotics are required in the medium, they should be suspended at 1,000 \times concentration in an appropriate solvent, filter-sterilized and then added to autoclaved medium after it has cooled to room temperature. Mixtures of LB medium with antibiotics should be freshly prepared, as they cannot be stored. See **Table 4** for antibiotic concentrations.

LB agar plates Combine 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl and 15 g of agar in 1 liter of water, and bring the pH to 7.5 by adding ~200 μl of 0.1 M NaOH. Autoclave the medium for sterility and to melt the agar, and then pour molten medium into sterile Petri dishes. If antibiotics are required in the medium, they must be filter-sterilized and then added to the autoclaved medium that has cooled to 50 $^{\circ}\text{C}$. LB agar solidifies below 50 $^{\circ}\text{C}$, but above 50 $^{\circ}\text{C}$ breakdown of antibiotics can occur.

Once the medium is solid, agar plates should be stored at 4 $^{\circ}\text{C}$ for up to 1 month.

Glycerol, 80% (vol/vol) Combine 80 ml of glycerol with 20 ml of water; mix the contents and then filter through a sterile vacuum filter unit. Store 80% (vol/vol) glycerol at room temperature for up to 1 year.

▲ CRITICAL Glycerol must be kept sterile to maintain frozen stocks of bacterial strains.

MAGE oligos Suspend MAGE oligos in deionized water (DI H₂O) at a DNA concentration of 50 μM . For single-site mutagenesis, the final concentration of the oligo is 1 μM (1 μl of oligo in a 50- μl total electroporation volume). For mutagenesis of more than one target site, a complex oligo pool is created by mixing the oligos that target separate loci. Generally, the final total oligo concentration in a complex pool is held at 0.5–10 μM , with each constituent contributing an equal amount. To create a 50- μl equimolar pool of n oligos at 10 μM aggregate concentration, (50/ n)- μl volumes of each 10 μM stock would be mixed. The oligo pool can be stored at –20 $^{\circ}\text{C}$ for at least 1 year; avoid frequent freeze/thaw cycles.

PROCEDURE

Preparation of electrocompetent cells and MAGE cycling ● TIMING 2 d to 1 week

1| For each MAGE experiment, set up a 3-ml culture of the appropriate bacterial strain in LB supplemented with appropriate antibiotics: 1 ml will be used for the MAGE experiment, 1 ml for the negative control and 1 ml for the (optional) positive control. Cultures can be established by inoculating the liquid medium with a single bacterial colony or by adding 30 μl of a confluent liquid culture (1:100 dilution). We suggest a ‘no DNA’ MAGE experiment as a negative control, and recombination with a lacZ_off oligo (**Table 2**) as a positive control. These controls allow MAGE ARF to be assayed in a lacZ⁺ strain by plating on X-Gal- and IPTG-containing media. In our hands, this oligo leads to ~30% white colonies, indicating a 30% ARF.

2| Grow the culture in a shaking incubator at 300 r.p.m. at 30–34 $^{\circ}\text{C}$ for ~2–3 h or until the cells reach mid-log phase, as measured by OD₆₀₀ of 0.4–0.6 (~10⁸ cells per ml).

▲ CRITICAL STEP Shaking is necessary to ensure adequate oxygenation and rapid growth.

▲ CRITICAL STEP The Red genes are under the control of a temperature-sensitive repressor; therefore, incubation should be performed at low temperature (<36 $^{\circ}\text{C}$) to avoid premature induction of recombination functions.

▲ CRITICAL STEP Media with high salt concentration may lead to low ARF or to failure of electroporation and are not compatible with certain antibiotics (e.g., zeocin).

▲ CRITICAL STEP Growth time varies with strains. Mid-log cultures contain a high proportion of actively replicating cells and are therefore conducive to high-frequency allelic replacement.

3| In preparation for Steps 6–11, chill 10 ml of DI H₂O on ice; thaw 50 μM oligo stocks (see Reagent Setup) and place them on ice and prechill two microcentrifuge tubes and a 0.1-cm electroporation cuvette for each sample and control. Each set of tubes should share a common label identifying the sample or control it will contain. For each sample or control, prepare oligo solution by adding 49 μl of sterile DI H₂O and 1 μl of the appropriate 50 μM oligo stock (lacZ_off oligo for the positive control or water for the negative control) to one of the prechilled labeled microcentrifuge tubes and chill it on ice. Also prepare a labeled culture tube containing 3 ml of LB for each sample or control (recovery tube).

4| Quickly transfer the culture tubes from Step 2 from the 30–34 $^{\circ}\text{C}$ incubator to the 42 $^{\circ}\text{C}$ shaking water bath. Induce the expression of λ -Red genes by heat shock for 15 min at 250 r.p.m.

▲ CRITICAL STEP Longer induction time results in cell death from toxicity of the Gam protein.

▲ CRITICAL STEP Vigorous shaking in a water bath is necessary to ensure adequate oxygenation, rapid heat transfer and complete induction.

5| After induction, quickly transfer the cultures from the shaking water bath to ice.

▲ CRITICAL STEP Keeping induced cells at low temperature until electroporation prevents breakdown of recombination proteins and maintains a high ARF.

▲ CRITICAL STEP Perform Steps 6–8 in a 4 $^{\circ}\text{C}$ cold room.

■ PAUSE POINT Cultures can be kept on ice for at most 3 h.

6| Transfer 1 ml of induced culture to each of the corresponding empty prechilled microcentrifuge tubes (from Step 3); there should be one tube for the experimental sample and one for each associated control. Centrifuge all tubes at maximum speed (13,000 r.p.m. or 14,000g) for 30s at 4 °C. Remove the supernatant.

7| Add 1 ml of ice-cold sterile DI H₂O to the cell pellet, and resuspend the cells with a pipette tip. Do not vortex. Centrifuge the mixture at maximum speed for (13,000 r.p.m. or 14,000g) for 30 s at 4 °C, and remove the supernatant. Repeat this step once.

8| After the second wash, remove the supernatant. Resuspend the cell pellet in 50 µl of oligo solution (from Step 3) with a pipette tip and transfer it to the corresponding prechilled electroporation cuvette. Keep it on ice until use. Competent cells should be electroporated within an hour of preparation.

▲ **CRITICAL STEP** Remove the supernatant carefully; otherwise, the cell pellet can be lost.

▲ **CRITICAL STEP** The supernatant contains electrolytes from LB and must be completely removed at each wash step to prevent arcing during electroporation.

9| Turn on the electroporator and set it to 1.8 kV, 25 µF capacitance and 200 ω resistance for a 0.1-cm gap cuvette. Settings for a 0.2-cm gap cuvette are as follows: 2.5 kV, 25 µF and 200 ω.

10| Wipe the outside of the cuvette dry. Deliver voltage to transform the cells with MAGE oligos.

▲ **CRITICAL STEP** After the electroporation shock has been delivered, ensure that the relaxation time (time constant) is at least 4.0 ms (this can vary for different electroporation instruments). Low time constants imply high electrolyte concentration (either from DNA-salt complex or improperly washed cells) and lead to increased lethality and low ARF. Well-washed cells with desalted DNA at an appropriate concentration exhibit time constants very close to the time constant for an equal volume of DI H₂O (~5.2 ms in the Bio-Rad instrument used throughout this protocol).

? TROUBLESHOOTING

11| Immediately after electroporation, add 1 ml of LB from the recovery tube (from Step 3) to the electroporation cuvette. Pipette it up and down to mix electroporated cells (~10% cells survive electroporation, i.e., 1 × 10⁷ cells per ml) with LB, and then transfer it back to the recovery tube to a final volume of 3 ml.

12| Incubate the recovery tube by shaking it at 300 r.p.m. at 30–34 °C. Different recovery times are appropriate for different applications: to perform a subsequent MAGE cycle (Fig. 5), follow option A; to pause between MAGE cycles, follow option B; to isolate recombinant clones for screening, follow option C; and to freeze a sample for later analysis, follow option D.

▲ **CRITICAL STEP** It is important to recover electroporated cultures long enough to allow at least two complete cycles of genome replication. If the recovery period is too short, cells will contain chimeric chromosomes in which oligo-genome mismatches still exist at target loci. Mutations by MAGE introduced on the lagging strand are completely inherited in cells only after two complete rounds of replication.

(A) To recover cells in preparation for a subsequent MAGE cycle

(i) Incubate the culture at 30–34 °C with shaking at 300 r.p.m. until the cells reach an OD₆₀₀ of 0.4–0.6.

(ii) Once it is at the appropriate density, port the recovered culture directly into an additional round of MAGE starting at Step 3. For EcNR2, appropriate cell density is achieved after ~2 h of recovery, permitting up to ten MAGE cycles per day.

(B) To pause between MAGE cycles

(i) Incubate the cells at 30–34 °C with shaking at 300 r.p.m. until they reach stationary phase (i.e., overnight).

(ii) To start another round of MAGE the next day, proceed from Step 1.

(C) To isolate recombinant clones for screening

(i) Recover cells by incubating them at 30–34 °C with shaking at 300 r.p.m. for at least 3 h, and then proceed to Step 13 to prepare a tenfold serial dilution.

(D) To freeze the recovered MAGE culture for later analysis

(i) Mix 700 µl of the recovered culture with 300 µl of 80% (vol/vol) glycerol (final concentration 24%) in a 2-ml screw-top cryotube. Store it at –80 °C indefinitely. To recover the frozen sample, thaw it on ice and grow it in a large volume of culture medium (e.g., 1:100 dilution).

(ii) Follow Step 12A for an additional round of MAGE.

? TROUBLESHOOTING

PROTOCOL

Isolation of mutants engendered by MAGE ● TIMING 2 d to 1 week

13| Prepare serial tenfold dilutions (from undiluted to 10⁶-fold diluted) for each MAGE sample, positive and negative controls. Use at least 180 µl of water, LB or PBS as the diluent for each step. Change the pipette tips between dilution steps.

14| Spread 50 µl of each dilution on plates selective for the mutagenized strain. Incubate the plates at 30–34 °C overnight until colonies appear. Usually 10⁴–10⁶-fold dilutions give single colonies.

? TROUBLESHOOTING

15| By using a multichannel pipette, fill each well of a 96-well flat-bottom culture plate with ~150 µl of a mix of LB and an appropriate antibiotic.

16| Use sterile toothpicks to pick 48–96 colonies from each sample, and then transfer each colony to one well of the 96-well plate. Include at least one colony from the 'no oligo' negative control sample that did not undergo MAGE.

17| Grow the picked colonies in a shaking incubator at 300 r.p.m at 30–34 °C for ~2 h or until all wells have reached mid-log phase, as measured by the OD₆₀₀.

18| Set up two PCRs (step 1 of **Box 2**) for each isolated colony, one using *f*_{WT-r} and the other using *f*_{mut-r} primer sets. To supply template DNA, add 1 µl from each well of the 96-well plate to both WT and mutant reactions. The use of multichannel pipettes and sterile reservoirs greatly accelerates the process of aliquotting and adding template to PCRs.

19| Run the PCRs using the protocol specified in step 2 of **Box 2** and the optimal annealing temperature determined in step 3 of **Box 2**.

20| Visualize 4 µl of each PCR product on an ~1.5% (wt/vol) agarose gel to determine the genotype of each colony. If necessary, screen more colonies to find the desired mutants.

? TROUBLESHOOTING

21| Freeze and store the pure cultures (Step 12D) for the colonies that screened correctly, and then confirm genotype with Sanger sequencing.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 5**.

TABLE 5 | Troubleshooting table.

Step	Problem	Possible reason	Solution
10	Arcing during electroporation	Residual salt in either the cell pellet or the introduced DNA	Wash the cells one more time with chilled DI H ₂ O Drop-dialyze the DNA to reduce salt
12	No outgrowth during recovery	Cultures are left for a long time at 42 °C	Restart the cultures; place the cells on ice immediately after 15-min induction
		Too many cells are killed during electroporation	Wash the cells one more time with chilled DI H ₂ O
14	No colonies on either experiment or control plates	Inadequate outgrowth time	Allow the cells to grow longer before plating
		Poor MAGE ARF	Carry out more MAGE cycles; induce cells when OD ₆₀₀ reaches 0.4–0.6
		Antibiotics are too strong	Prepare new agar plates, supplementing appropriate amount of antibiotics as suggested in Table 4

(continued)

TABLE 5 | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
	Colonies on both experiment and negative control plates	Antibiotics are too weak	Prepare new agar plates, supplementing appropriate amount of antibiotics as suggested in Table 4
		Strain contamination	Re-streak strains; use sterile technique during MAGE
20	All PCR bands show the wild type	Poor MAGE allelic replacement frequency	Carry out more MAGE cycles; induce cells when OD ₆₀₀ reaches 0.4–0.6
		Inadequate screen	Screen more colonies
	Nonspecific results for MASC-PCRs: both f_{wt}/r and f_{mut}/r generate clear bands	T_m is not optimal	Optimal T_m should be determined by gradient PCR
		Picked colonies are not clonal or isogenic	Pick only easily resolved colonies; ensure sufficient recovery for segregation of mutation
Box 3	Both the experimental and negative control cultures grow	Colicin E1 has degraded	Use freshly prepared colicin E1 to set up another negative selection; dilute fewer cells in each well

● TIMING

Steps 1–12, preparation of electrocompetent cells and MAGE cycling: 2 d to 1 week

Steps 13–21, isolation of mutants engendered by MAGE: 2 d to 1 week

Box 2, optimizing PCR conditions for MASC-PCR: 1 d

Box 3, cosMAGE cycling: 2 d–2 weeks

ANTICIPATED RESULTS

MAGE is a valuable method to engineer bacterial genomes *in vivo* using ssDNAs. In general, up to ~30% of targeted alleles can be replaced in a single MAGE cycle in ~2.5 h. For >1 target site, expected ARFs and total population complexity can be predicted using the plots in **Figure 4**. CosMAGE further enhances the ARF by about fourfold, thus greatly improving the ARF for multiple target loci.

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

ACKNOWLEDGMENTS Funding was provided by the US Department of Energy (DE-FG02-02ER63445), Defense Advanced Research Projects Agency (N66001-12-C-4020, N66001-12-C-4211), and the Arnold and Mabel Beckman Foundation (F.J.I.).

AUTHOR CONTRIBUTIONS R.R.G., Z.L. and F.J.I. wrote the manuscript with contributions on mathematical modeling from A.O.L.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Metzker, M.L. Sequencing technologies—the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
- Ashworth, J. *et al.* Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441**, 656–659 (2006).
- Paddon, C.J. *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528–532 (2013).
- Wang, H.H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).

- Carroll, D. Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **83**, 409–439 (2014).
- Jackel, C., Kast, P. & Hilvert, D. Protein design by directed evolution. *Annu. Rev. Biophys.* **37**, 153–173 (2008).
- Zhang, Y., Buchholz, F., Muyrers, J.P. & Stewart, A.F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–128 (1998).
- Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**, 5978–5983 (2000).
- Ellis, H.M., Yu, D., DiTizio, T. & Court, D.L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. USA* **98**, 6742–6746 (2001).
- Isaacs, F.J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
- Lajoie, M.J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
- Wang, H.H. *et al.* Genome-scale promoter engineering by coselection MAGE. *Nat. Methods* **9**, 591–593 (2012).
- Caruthers, M.H. Gene synthesis machines: DNA chemistry and its uses. *Science* **230**, 281–285 (1985).
- Carr, P.A. & Church, G.M. Genome engineering. *Nat. Biotechnol.* **27**, 1151–1162 (2009).
- Kosuri, S. *et al.* Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat. Biotechnol.* **28**, 1295–1299 (2010).



16. Orr-Weaver, T.L., Szostak, J.W. & Rothstein, R.J. Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci. USA* **78**, 6354–6358 (1981).
17. Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. & Rehrauer, W.M. Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**, 401–465 (1994).
18. Muniyappa, K. & Radding, C.M. The homologous recombination system of phage λ . Pairing activities of β -protein. *J. Biol. Chem.* **261**, 7472–7478 (1986).
19. Haseltine, E.L. & Arnold, F.H. Synthetic gene circuits: design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 1–19 (2007).
20. Stemmer, W.P. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**, 10747–10751 (1994).
21. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage-T4 DNA polymerase. *Science* **249**, 505–510 (1990).
22. Ellington, A.D. & Szostak, J.W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
23. Soukup, G.A. & Breaker, R.R. Engineering precision RNA molecular switches. *Proc. Natl. Acad. Sci. USA* **96**, 3584–3589 (1999).
24. Orth, J.D., Thiele, I. & Palsson, B.O. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
25. Burgard, A.P., Pharkya, P. & Maranas, C.D. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657 (2003).
26. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
27. Binder, S., Siedler, S., Marienhagen, J., Bott, M. & Eggeling, L. Recombineering in *Corynebacterium glutamicum* combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res.* **41**, 6360–6369 (2013).
28. Carr, P.A. *et al.* Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Res.* **40**, e132 (2012).
29. Datta, S., Costantino, N., Zhou, X. & Court, D.L. Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl. Acad. Sci. USA* **105**, 1626–1631 (2008).
30. Zhang, Y., Muylers, J.P., Rientjes, J. & Stewart, A.F. Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol. Biol.* **4**, 1 (2003).
31. Li, X.T. *et al.* Identification of factors influencing strand bias in oligonucleotide-mediated recombination in *Escherichia coli*. *Nucleic Acids Res.* **31**, 6674–6687 (2003).
32. Lajoie, M.J., Gregg, C.J., Mosberg, J.A., Washington, G.C. & Church, G.M. Manipulating replisome dynamics to enhance λ Red-mediated multiplex genome engineering. *Nucleic Acids Res.* **40**, e170 (2012).
33. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
34. Babic, I., Andrew, S.E. & Jirik, F.R. MutS interaction with mismatch and alkylated base containing DNA molecules detected by optical biosensor. *Mutat. Res.* **372**, 87–96 (1996).
35. Sawitzke, J.A. *et al.* Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J. Mol. Biol.* **407**, 45–59 (2011).
36. Wang, H.H., Xu, G., Vonner, A.J. & Church, G. Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. *Nucleic Acids Res.* **39**, 7336–7347 (2011).
37. Swingle, B. *et al.* Oligonucleotide recombination in Gram-negative bacteria. *Mol. Microbiol.* **75**, 138–148 (2010).
38. Costantino, N. & Court, D.L. Enhanced levels of Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. USA* **100**, 15748–15753 (2003).
39. Nyerges, A. *et al.* Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic Acids Res.* **42**, e62 (2014).
40. Mosberg, J.A., Gregg, C.J., Lajoie, M.J., Wang, H.H. & Church, G.M. Improving Red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. *PLoS ONE* **7**, e44638 (2012).
41. DeVito, J.A. Recombineering with *tolC* as a selectable/counter-selectable marker: remodeling the rRNA operons of *Escherichia coli*. *Nucleic Acids Res.* **36**, e4 (2008).
42. Warming, S., Costantino, N., Court, D.L., Jenkins, N.A. & Copeland, N.G. Simple and highly efficient BAC recombineering using *galK* selection. *Nucleic Acids Res.* **33**, e36 (2005).
43. Gregg, C.J. *et al.* Rational optimization of *tolC* as a powerful dual-selectable marker for genome engineering. *Nucleic Acids Res.* **42**, 4779–4790 (2014).
44. Herschma, H.R. & Helinski, D.R. Comparative study of events associated with colicin induction. *J. Bacteriol.* **94**, 691 (1967).
45. Schwartz, S.A. & Helinski, D.R. Purification and characterization of colicin E1. *J. Biol. Chem.* **246**, 6318–6327 (1971).
46. Wang, H.H. & Church, G.M. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. *Methods Enzymol.* **498**, 409–426 (2011).