ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

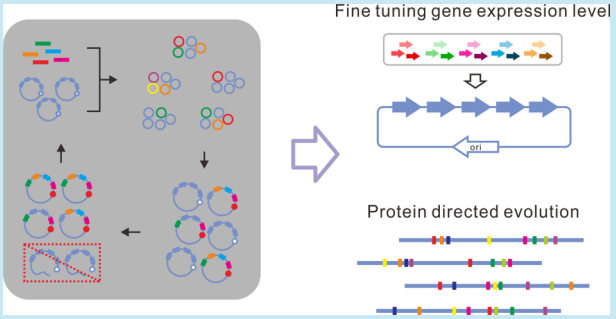
Multiplex Iterative Plasmid Engineering for Combinatorial Optimization of Metabolic Pathways and Diversification of Protein Coding Sequences

Yifan Li,† Qun Gu,† Zhenquan Lin, Zhiwen Wang, Tao Chen,* and Xueming Zhao

Key Laboratory of Systems Bioengineering, Ministry of Education, and Department of Biochemical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, People's Republic of China

Supporting Information

ABSTRACT: Engineering complex biological systems typically requires combinatorial optimization to achieve the desired functionality. Here, we present Multiplex Iterative Plasmid Engineering (MIPE), which is a highly efficient and customized method for combinatorial diversification of plasmid sequences. MIPE exploits ssDNA mediated λ Red recombineering for the introduction of mutations, allowing it to target several sites simultaneously and generate libraries of up to 10^7 sequences in one reaction. We also describe "restriction digestion mediated co-selection (RD CoS)", which enables MIPE to produce enhanced recombineering efficiencies with greatly simplified coselection procedures. To demonstrate this approach, we applied MIPE to fine-tune gene expression level in the 5-gene riboflavin biosynthetic pathway and successfully isolated a clone with 2.67-fold improved production in less than a week. We further demonstrated the ability of MIPE for highly multiplexed diversification of protein coding sequence by simultaneously targeting 23 codons scattered along the 750 bp sequence. We anticipate this method to benefit the optimization of diverse biological systems in synthetic biology and metabolic engineering.



KEYWORDS: *ssDNA recombineering, plasmid library, combinatorial optimization, protein directed evolution, synthetic biology, metabolic engineering*

A major challenge in metabolic engineering and synthetic biology is to engineer complex artificial circuits with sophisticated functions. Due to the inherent complexity of biological systems, it is often difficult to rationally design every component in a synthetic gene network to arrive at optimal performance. One solution to this problem is to engineer libraries of designs in parallel and screen for those that exhibit the desired functions.[1−7] Traditionally, genomic diversities can be obtained by spontaneous or induced mutagenesis, which introduces mutations randomly across the entire chromosomes. Recent advances in genetic engineering and DNA synthesis techniques allow to generate more focused variations in catalytic enzymes,[8] promoters,[4,9] ribosome binding sites (RBSs),[2,10] intergenic regions,[11] and transcription factors[12] to search for improved biological functions.

Engineering collections of synthetic biology designs in parallel depends heavily on the ability to construct large plasmid libraries with targeted variations. Restriction digestion/ligation based methods have been widely used but are laborious, limited by unique restriction sites, and incapable of parallel assembly. Several efficient and advanced techniques for the assembly of multigene constructs have been developed in the past several years, including the *in vitro* sequence and ligation-independent cloning (SLIC),[13] the Gibson isothermal assembly,[14] and the DNA assembler,[15] and the circular polymerase extension cloning (CPEC).[16] These techniques have the common advantages that they are relatively sequence independent and allow cloning multiple inserts in a single step. Some techniques among them have demonstrated the ability of creating combinatorial synthetic gene libraries. For example, the Gibson assembly method has been used to construct a library that consists of ∼10^2 two-gene constructs.[17] Based on DNA assembler, an approach named "customized optimization of metabolic pathways by combinatorial transcriptional engineering (COMPACTER)" has been developed for rapid tuning gene expression in synthetic pathways.[4] CPEC is based on polymerase overlap extension and has also been described for cloning combinatorial DNA libraries with up to 5 separately amplified pieces.[16] However, all of these methods require PCR amplification of DNA sequences, which has intrinsic limitations such as high mutation rate and inconvenience in cloning very long (or short) sequences. In addition, the library size generated by these methods is relatively small (up to 10^3 when assembling ≤3 pieces, except CPEC reported to reach 10^5) and usually drops dramatically when assembling >3 pieces in parallel,[5] limiting their applications in high-throughput

Received: April 29, 2013
Published: August 28, 2013

biological optimizations. A unique method "Reiterative Recombination" has been reported to construct large mock libraries of at least 10^4 biosynthetic pathways in the chromosome of *Saccharomyces cerevisiae* but requires multiple elongation steps to construct multigene pathways and depends on the efficient recombination systems of the host cells.[5]

Multiplex Automatable Genome Engineering (MAGE) is a powerful technique to generate libraries of targeted mutations in a population of *Escherichia coli* (*E. coli*) chromosomes.[2] It has shown the ability to combinatorially optimize the RBS sequences of 20 genes in the lycopene biosynthesis pathway, creating 4.3 billion genomic variants per day. This technique is based on the oligo mediated λ Red recombineering (OMR), in which Redβ recombinase binds to the introduced oligonucleotides (oligos), facilitating their annealing to the lagging strand at the replication fork.[18] By removing the endogenous mismatch repair (MMR) machinery and optimizing introduced synthetic oligos, point mutations can be introduced into *E. coli* genome at frequencies up to 30%.[2,19] Co-selection (CoS) strategy has also been described to greatly enhance MAGE efficiency by ~4-fold, in which an oligo was used to repair a broken selectable marker (e.g., antibiotic resistance gene) to enhance allele replacement (AR) frequency of nearby nonselectable alleles presumably by selecting for cells that take up oligos and that have a permissive replication fork in the desired region of the genome.[20] Despite these great improvements, MAGE has currently been limited to the modification of *E. coli* genomes. Hence, there is a pressing need to expand the applicability of this highly multiplexed engineering idea to the diversification of plasmid sequences and the optimization of genetic components that function in other host cells.

Here, we present MIPE, which uses OMR for highly efficient and multiplexed plasmid sequence diversification. In this approach, plasmids and multiple synthetic oligos are coelectroporated into cells with λ Red functions, introducing combinations of mutations in plasmid sequences and generating libraries of up to 10^7 sequences. Restriction digestion mediated co-selection (RD CoS) strategy is also developed, which allows MIPE to use a single restriction site region as CoS marker for several cycles and directly select mutant plasmid molecules. Finally, we show that MIPE can be used to combinatorially optimize synthetic metabolic pathways and to diversify protein coding sequence for directed evolution.

RESULTS AND DISCUSSION

Characterization and Optimization of Oligo Meditated Plasmid Recombineering. Previously, OMR has mainly been used to the engineering of single-copy replicons such as bacterial chromosomes and BACs;[18] very few efforts have attempted to modify plasmid DNA with recombineering.[21] Plasmids are multicopy replicons, which may complicate their OMR in the following aspects. First, a recombination event always generates mixed populations of mutant and parental plasmid molecules (Supporting Information, Discussion), which requires further efforts to isolate pure clones and confounds the calculation of AR frequencies (Figure 1). Second, we speculate that different plasmid molecules within a single cell may compete with each other for the limited recombineering recourses (oligos and Redβ proteins), possibly reducing the average probability of recombination events on each plasmid molecule.

To address the first issue, we used a different method to estimate AR frequencies for plasmid recombineering compared

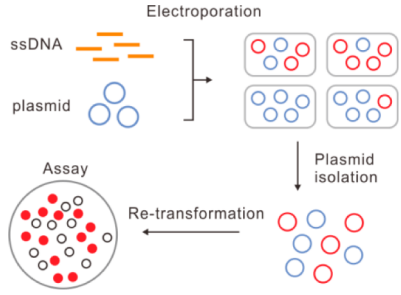


Figure 1. Method for assaying plasmid recombineering AR frequencies. Cells after recombineering usually contain mixtures of parental and mutant plasmids, which confounds the calculation of AR frequencies. To solve this problem, plasmid mixtures are amplified, isolated, and retransformed to determine the AR frequencies at single plasmid molecular level. This diagram shows the case of plasmid coelectroporation, but this assaying method was also utilized for targeting resident plasmids.

with that used for genome recombineering (Figure 1). After electroporation, cell mixtures were not directly plated to calculate frequencies. Rather, they were incubated overnight with antibiotics to amplify the plasmids, which were then isolated and retransformed to calculate AR frequencies. Low concentrations (<1 ng/reaction) of plasmids were added in the retransformation reactions, ensuring most competent cells take in only one plasmid molecule and allowing the estimation of AR frequencies at single plasmid molecular level.

To address the second issue, we tested the AR frequencies under three different conditions. In condition A, 0.5 μM oligos were used in each recombineering reaction, which was previously reported to be saturating to achieve optimal AR frequency for genome recombineering.[2,22] In condition B, 5 μM oligos were added in 10 times diluted electrocompetent cells, resulting in 100 times higher oligo/cell ratios than condition A, in case that multicopy plasmids may require condensed oligos to achieve optimal efficiencies. In condition C, 1 ng plasmid to be engineered is coelectroporated together with the 0.5 μM oligos, providing an instant low plasmid concentration within the cells when recombineering. Co-electroporation has been used to engineer plasmids but was reported to have no effect on AR frequencies.[21,23] However, these experiments were mainly performed in strains with active MMR systems, which yield very low AR frequencies. In addition, the frequencies were estimated by directly plating the cell mixtures after recombineering and therefore could not tell the recombination frequencies at the plasmid molecular level.

We constructed a series of testing plasmids to characterize OMR with four different plasmid replication origins (pSC101, p15A, pBR322, and pUC18). Each plasmid contained a replication origin with a certain orientation, a *bla* gene for plasmid selection and an inactivated *kan* gene with two consecutive stop codons in its ORF for screening mutants. Oligo kan_on was designed to restore the Kan resistant phenotype and the AR frequencies were determined by the ratio of the number of colonies formed on Kan plates to that on Amp plates.

As shown in Figure 2a, even with oligos corresponding to the lagging strand, targeting resident plasmids yielded much decreased AR frequencies (1.2%, 1.2%, 0.6% and 0.2% for pSC101, p15A, pBR322, and pUC18, respectively) compared with the frequencies targeting the *E. coli* chromosomes under similar conditions.[2,19] In addition, the AR frequencies for high

dx.doi.org/10.1021/sb400051t | ACS Synth. Biol. 2013, 2, 651−661

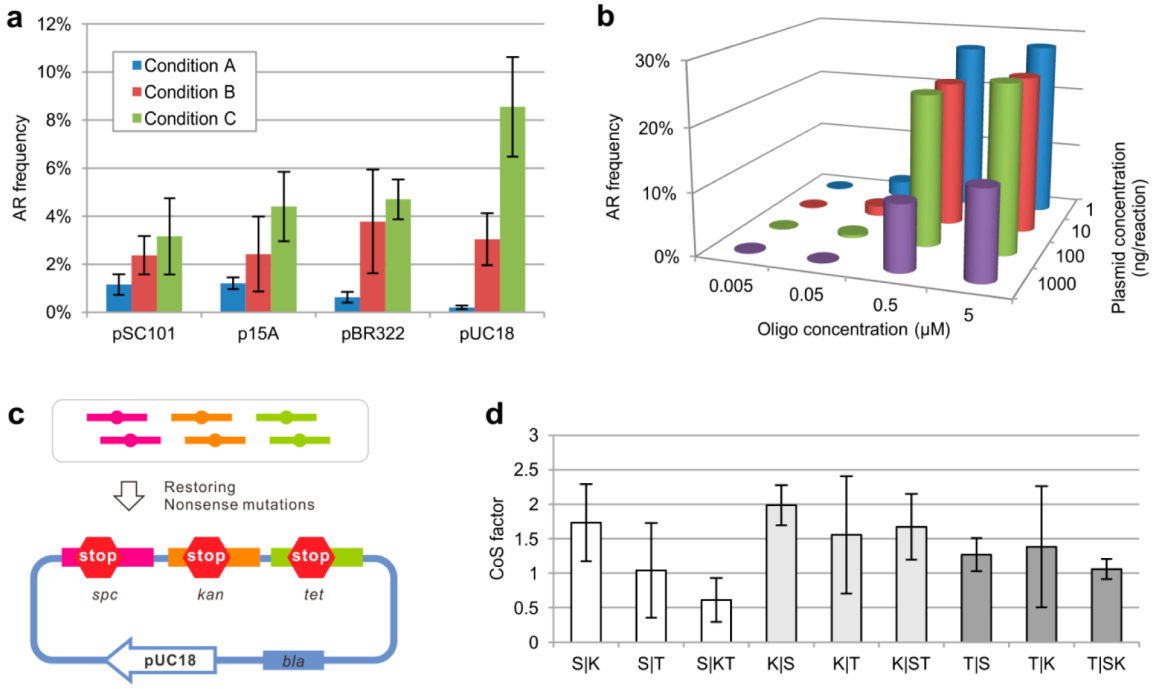


Figure 2. Characterization and optimization of plasmid recombineering. (a) AR frequencies for targeting the lagging strand of four different types of plasmids under three different conditions. In condition A, 0.5 μM oligos were electroporated into 50 μL competent cells harboring the targeted plasmids. In condition B, 5 μM oligos were added in 10 times diluted competent cells harboring the targeted plasmids for electroporation, resulting in 100 times higher oligo/cell ratios than that for condition A. In condition C, 0.5 μM oligos and 1 ng plasmid were coelectroporated. Oligo kan_on was introduced to restore the Kan resistant phenotype and the AR frequencies were determined by the ratio of colony numbers on Kan plates to that on Amp plates. Error bars represent standard deviation; $n = 3$. (b) AR frequencies with varying concentrations of oligos and plasmids. Oligo rfp_off inactivating the *rfp* gene and pRFP07 were coelectroporated and the AR frequencies were calculated by the ratio of white versus total colonies. Data are the mean of triplicate experiments. (c) Method for charactering coselection effect. pLX15 containing three inactivated antibiotic resistance markers was coelectroporated with three oligos designed to restore each of the antibiotic resistant phenotypes. 1 μM assaying oligo, 0.05 μM CoS oligos, and 1 ng plasmid were introduced in each reaction. AR frequencies were determined by plating the cells on plates with different combinations of antibiotics, and from which the CoS factors were calculated. (d) CoS factors when each of Spc^+, Kan^+, and Tet^+ serving as the assaying phenotype. "S|K" indicates that Spc^+ is coselected by Kan^+, and similar for the others. Error bars represent standard deviation; $n = 2$.

copy number plasmids were much lower than that for medium and low copy number plasmids. However, the AR frequencies could be greatly improved by increasing the oligo/cell ratios or by coelectroporation. For all types of plasmids tested, coelectroporation always yielded the highest AR frequencies. These observations reinforce our speculation that the multicopy nature of plasmids decreases the AR efficiencies and suggest that coelectroporation is absolutely necessary to arrive at optimal AR frequencies, especially for high copy number plasmids.

At the same time, we found that higher copy number plasmids had higher frequencies when plasmids were coelectroporated with their lagging strand targeted. A possible reason for this phenomenon is that plasmids with higher copy number have higher frequency to initiate replication and thus generates more permissible replication forks for oligo to anneal to. Targeting the lagging strand of the plasmid always produced higher AR frequencies than targeting the leading strand, which was in agreement with the observations in chromosomes recombineering (Supporting Information (SI) Figure S1).

We further determined the impact of plasmid and oligo concentrations on AR frequency for coelectroporation. Oligo rfp_off was designed to introduce an inactivating nonsense mutation in the *rfp* gene on a pUC18 derived plasmid pRFP07. rfp_off and pRFP07 were coelectroporated and the AR frequencies were calculated by the ratio of white versus total colonies. In general, AR frequency increased as the oligo concentration increased and the plasmid concentration decreased (Figure 2b). A plateau of frequencies (more than 20%) could be achieved with oligo concentration no less than 0.5 μM and plasmid concentration no more than 100 ng per reaction. Decreasing plasmid concentration to less than 1 ng per reaction did not further increase recombineering frequencies (data not shown). On this plateau, varying oligo or plasmid concentrations had very slight effect on AR frequencies. However, reducing oligo concentration to 0.05 μM dramatically reduced AR frequencies even with very low amount of plasmid. Increasing plasmid concentration to 1000 ng/reaction also decreased frequencies even with high oligo concentrations (10.6% for 0.5 μM oligos and 14.2% for 5 μM oligos). This observation is again in line with our speculation that competition between different plasmid molecules reduces AR frequencies.

To assess the effect of coselection on plasmid recombineering, we constructed a testing plasmid pLX15 containing three inactivated antibiotic resistance genes (*spc*, *kan*, and *tet*), each encoding a reversible nonsense mutation (Figure 2c). pLX15 was coelectroporated with three different oligos (CoS_Spc_on, CoS_Spc_on, and CoS_Kan_on), each restoring one of the inactivated genes. In an individual recombineering reaction, one oligo was selected as the assaying oligo (for assaying AR frequency, not for CoS) while another oligo (or both of the other two oligos) was used as the CoS oligo. Assaying oligo was introduced at 1 μM concentration whereas CoS oligo at 0.05

dx.doi.org/10.1021/sb400051t | *ACS Synth. Biol.* 2013, 2, 651−661

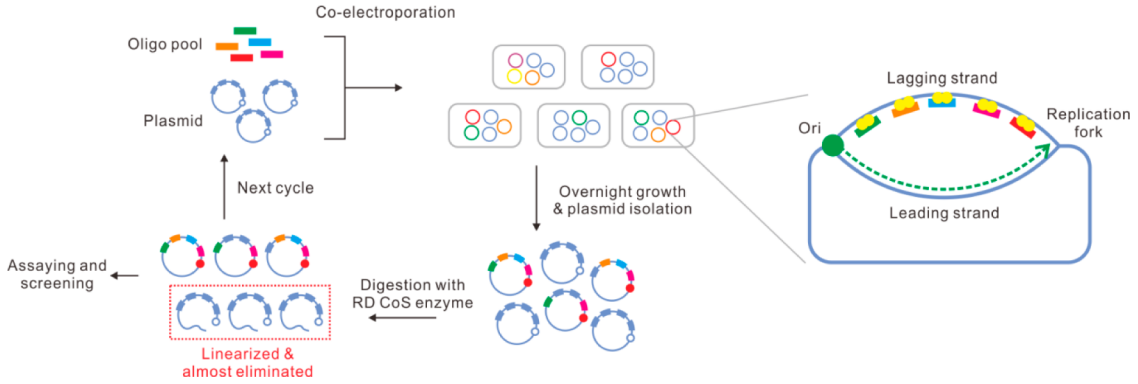


Figure 3. Scheme and mechanism of MIPE. In each cycle, plasmids (pure plasmids to initiate the first cycle of MIPE or digested plasmid mixtures from the previous cycle) are coelectroporated with oligo mixtures. An RD CoS oligo is also introduced to switch a unique restriction site (white circle) to another site (red circle), which is previously absent from the plasmid. During plasmid replication, several oligos are very likely to incorporate into the plasmid simultaneously, often inherited together by the same daughter plasmid.[20] (The dividing plasmid is schematized, with the five oligos binding the Redβ recombinases (yellow circles) and simultaneously annealing to the lagging strand at the replication fork.) The resulting cells are incubated overnight with antibiotics for plasmid amplification. In the next day, the isolated plasmid mixture is digested with the enzyme cutting the native RD CoS site (white circle) followed by coelectroporation to initial the next cycle. These procedures enable WT plasmids to be linearized and almost eliminated, and the mutants to be enriched. The switched RD CoS site (red circle) can serve as native RD CoS marker in the next cycle. MIPE requires a single day for each cycle including 0.5 h for plasmid isolation, 0.6 h for digestion, 0.4 h for electroporation, and an overnight for plasmid amplification. Given the small size of plasmid molecules and the large abundant of commercial restriction enzymes, it is usually easy to find sites for RD CoS. If not, a new site can be introduced into a nonessential region on the plasmid by the oligo recombineering method we are introducing here or site directed mutagenesis.

μM.[20] The effect of CoS could be assessed by CoS factor, which was calculated by the ratio of AR frequencies with CoS to that without CoS. As Figure 2d shows, CoS factors were larger than one in most cases, indicating a positive effect of CoS on plasmid recombineering. However, we observed the CoS factors were smaller than two in most cases, which was different from previous observations that CoS factors were around 4 when modifying *E. coli* genomes.[20] At the same time, CoS by two markers did not statistically yield higher CoS factors than CoS by a single marker, which appears to contradict previous observation that double CoS yielded higher AR frequencies.[20] Several factors may have contributed to these different CoS effects (SI Discussion).

In summary, to achieve optimal AR frequency for plasmid recombineering, it is recommended to (1) use coelectroporation to target the lagging strand of the plasmid, (2) introduce no more than 100 ng plasmid DNA and no less than 0.5 μM oligos in each reaction, and (3) use Co-Selection.

Construction and Testing of MIPE. Although MAGE has demonstrated its powerful ability to engineer *E. coli* genomes in a highly multiplexed manner, the original MAGE cycle of electroporation follow by cell recovery and growth may not be suitable to engineer plasmid DNA. To this end, we have designed MIPE for more efficient and convenient engineering of plasmid DNA via OMR (Figure 3). MIPE contains several *in vitro* plasmid manipulation steps, which allows the strategy of "restriction digestion mediated co-selection (RD CoS)". In The first round of MIPE, plasmids, and oligo mixtures are coelectroporated into recombineering cells, allowing several sites on the plasmid to be modified simultaneously. An RD CoS oligo is also included in the mixtures to switch a unique restriction site (Site A) on the plasmid to a second site initially absent from the plasmid (Site B). The resulting cells are incubated overnight with antibiotics for plasmid amplification. In the next day, the plasmid mixture is isolated, digested with the enzyme cutting Site A, and coelectroporated with oligo mixtures to start the next cycle. During digestion and retransformation, WT plasmids (plasmids with Site A) are linearized and thus have greatly reduced transformation efficiencies (SI Discussion), whereas mutant plasmids (plasmids with Site B) are resistant to digestion, remain circular, and transform efficiently. Therefore, WT plasmids are almost totally eliminated while the mutants are highly enriched and dominate the plasmid population. In the next cycle, Site B can be utilized as RD CoS marker and switched to a third site (Site C) and so on. (We do not recommend switching site B directly back to Site A, because if so, incomplete digestion may cause the WT plasmids to be exponentially enriched in the second round of MIPE. In addition, considering the small size of plasmids and the large number of available restriction enzymes, it is usually very easy to find multiple sites for RD CoS.) MIPE offers several advantages in this manner. First, it allows plasmid coelectroporation, which is critical to achieve optimal AR efficiencies. Second, it enables a 6bp restriction site region to be used as CoS marker iteratively for several cycles. Third, it allows directly selecting the recombinant plasmid molecules by removing the WT clones by restriction digestion. Fourth, it separates the diversification and the screening processes so that gene circuits can be diversified in MMR defective *E. coli* but function not only in *E. coli* but also in host cells deficient of sophisticated recombination systems.

As a proof-of-principle demonstration, we first used MIPE to introduce a stop codon in *rfp* gene coselected by changing a nearby restriction site. RD CoS oligo rfp_CoS_*Eco*RI was designed to change the *Pst*I sequence located directly after the TAA stop codon of the *rfp* gene to *Eco*RI and was coelectroporated with 1 μM rfp_off and 1 ng pRFP07 into recombineering cells. Plasmid mixtures isolated after recombineering were digested with *Pst*I and then retransformed. The AR frequencies were calculated by the ratio of white versus total colonies. AR frequencies without RD CoS were measured in a similar way except that the plasmids were not digested with *Pst*I. As shown in Figure 4, MIPE yielded AR frequencies up to 60% for introducing a point mutation in a single round. RD

dx.doi.org/10.1021/sb400051t | *ACS Synth. Biol.* 2013, 2, 651−661

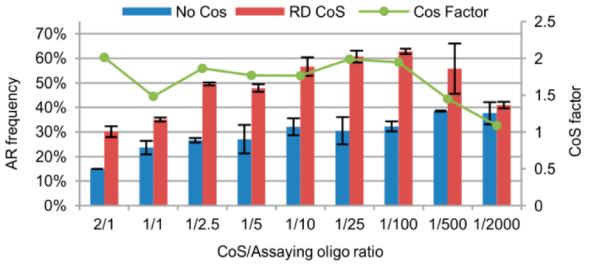


Figure 4. MIPE for introducing a point mutation. Varying concentrations of RD CoS oligo switching the *Pst*I site downstream of *rfp* to *Eco*RI, 1 μM rfp_off introducing a stop codon in the *rfp* gene, and 1 ng pRFP07 were coelectroporated into electrocompetent cells. The resulting plasmids were digested with *Pst*I and retransformed to determine the AR frequencies with RD CoS. Frequencies without CoS were determined by retransformation of the collected plasmid without digestion. CoS factors were calculated by the ratio of AR frequencies with CoS to that without CoS. Error bars represent standard deviation; $n = 2$.

CoS strategy generally increased AR frequencies by a factor of 2, which was consistent with the CoS experiments with pLX15.

The AR frequencies with RD CoS were robust within a wide range of RD CoS oligo concentration. Increasing RD CoS oligo concentration to 1:1 of the assaying oligo concentration however reduced AR frequencies both with and without RD CoS, possibly due to the competition between oligos for getting into the cells or annealing to the Redβ proteins.[20] Diluting RD CoS oligo to 1/500 and 1/2000 of the assaying oligo concentration reduced CoS factor to 1.45 and 1.09 respectively. We speculate the reason for the drop of CoS factors is that restriction digestion cannot provide very strict selections as antibiotic resistance markers; the procedure of digestion and retransformation can only reduce but not eliminate parental plasmids. We also tested if two very adjacent oligos would interfere with each other when annealing to the replicating plasmid. Surprisingly, RD CoS oligo that was even overlapping with the assaying oligo for 15bp did not yield decreased CoS factor (SI Figure S2).

MIPE for Combinatorial Optimization of Metabolic Pathways. To challenge MIPE for combinatorial diversification of plasmid sequence, we applied this method to optimize the 5-gene riboflavin biosynthetic pathway (Figure 5a). Riboflavin (vitaminB2) is required for a wide variety of cellular processes and is supplemented for feed and food fortification purposes in humans and animals to maintain health.[24] It is a yellow pigment, so its production can be easily screened by the color intensities of the colonies. The pathway was constructed by cloning the five genes (*ribC*, *ribE*, *ribD*, *ribB*, and *ribA*) from *E. coli* MG1655 genome into an operon under the control of a Trc promoter on p15A plasmid, resulting in a ~8.4k plasmid pRib23 (Figure 5a).

Unlike previous approaches using the same degenerate canonical Shine–Dalgarno sequence to enhanced translation efficiency of every gene in the pathway,[2] our strategy was to carefully design customized degenerate oligos with RBS Calculator[10] to fine-tune gene expression level (SI Tables S1 and S2). For a particular degenerate sequence, the designing principle was to produce a library of RBSs with strengths spanning a wide range ($>10^3$) and exhibiting near uniform distribution (SI Figure S3). Generally, a core region of 6–10 nucleotides in a certain RBS was targeted for mutagenesis, which could generate a library of 12–24 different sequences. Different RBS regions could be combinatorially modified, giving rise to a total library complexity of $\sim 10^6$ different pathways ($16 \times 16 \times 24 \times 24 \times 12 = 1.8 \times 10^6$) .

Oligo mixture including the five oligos targeting the RBSs and an RD CoS oligo were introduced together with pRib23 into recombineering cells to initiate MIPE. The RD CoS oligo for the first round was designed to change an *Xho*I site downstream of the replication origin to *Eco*RI. Plasmid mixture isolated in the next day was digested with *Xho*I and coelectroporated to initiate the second cycle. The *Eco*RI site was switched to *BamH*I and then to *Afl*II for RD CoS in

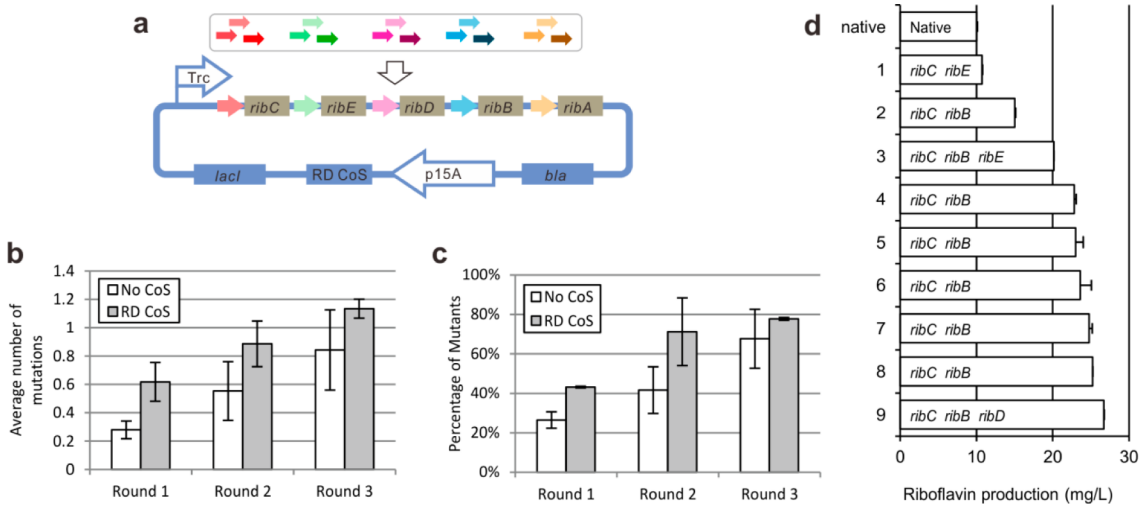


Figure 5. MIPE for combinatorial optimization of riboflavin biosynthetic pathway. (a) The five genes of the pathways were cloned into an operon under the control of a Trc promoter on p15A plasmid. To fine tune gene expression, libraries of oligos with various RBS strengths were introduced, simultaneously targeting all of the five RBS regions. A 6 bp restriction site region downstream of the replication origin served as RD CoS marker. (b) Average number of mutations and (c) percentage of mutant clones generated after different cycles of MIPE. Samples without CoS were prepared in an identical way to that with RD CoS except that plasmid mixtures were not digested to linearize the plasmid. Error bars represent standard deviation; $n = 2$. (d) Riboflavin production of the overproducing clones. Error bars represent standard deviation; $n = 2$. Labels in the bar indicate the genes with mutant RBS sequences. Notably, none of the clones are identical although they share a very similar mutation pattern.

dx.doi.org/10.1021/sb400051t | *ACS Synth. Biol.* 2013, 2, 651–661

Table 1. Mutant RBS Sequences[a] of *ribC* and *ribB*, Their Predicted Strengths[b], and Their Relative Rankings[c] in the Synthetic Libraries

	ribC			*ribB*		
clones	sequence	strength (a.u.)	ranking (%)	sequence	strength (a.u.)	ranking (%)
native	TTTCAGAAGTCGTAAGT	145	100.0[d]	AGGAGCTCTGTTTTACC	112	100.0[d]
1	TT---GGAGTTAAAAGT	2271	65.6	AGGAGCTCTGTTTTACC	112	100.0
2	TTTCAGGAGGTACAAGT	20607	37.5	ACTGAGGAAGTTTTACC	2883	62.5
3	TTTCAGGAGTTAAAAGT	8377	50.0	ACTAAGGAAGTTTTACC	16381	25.0
4	TTTAAGAAGGTAAAAGT	12561	43.8	ACTAAGGAAGTTTTACC	16381	25.0
5	TTTAAGGAGTTAAAAGT	53744	18.8	AGTAAGGAAGTTTTACC	10258	29.2
6	TTTAAGGAGGTACAAGT	138912	12.5	AGTAAGGATGTTTTACC	3578	58.3
7	TTTCAGGAGGTACAAGT	20607	37.5	AGTAAGGAGGTTTTACC	66704	8.3
8	TTTAAGGAGGTACAAGT	138912	12.5	AGTAAGGATGTTTTACC	3578	58.3
9	TTTCAGGAGTTACAAGT	3562	62.5	AGTAAGGAAGTTTTACC	10258	29.2

[a]17 bp sequences ahead of ATG are shown. [b]The strengths are calculated with RBS calculator. [c]For example, 37.5% means that this RBS sequences is predicted to be among the top 37.5% strongest in the population. [d]The native RBS is not in the library, but their predicted strength are weaker than any RBS in the mutant library.

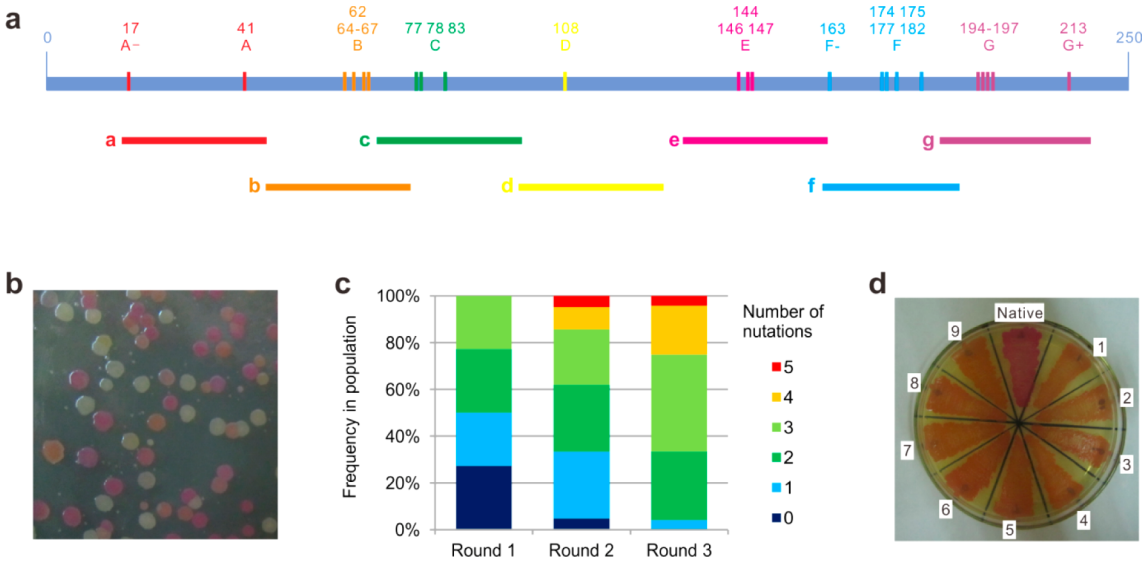


Figure 6. MIPE for diversification of *rfp* sequence. (a) Twenty three codons (small bars on the blue line with their relative positions depicted) in the 750 bp *rfp* sequence were simultaneously targeted by seven oligos (colored lines e to g showing their covered regions on the sequence). The 23 codons are grouped into 10 regions designated by the capital letters above the regions. The numbers above the capital letters were the exact positions of the targeted codons. (b) Photograph of the colonies from the library generated after the 2 rounds of MIPE. Photo was taken by an ordinary digital camera. (c) Frequency distribution of mutations in the libraries generated after different cycles of MIPE. All modified nucleotides covered by the same oligo were regarded as one mutation. Data was collected from two independently prepared samples and $n > 20$ for each sample. (d) Nine clones with orange colors could be isolated and analyzed to elucidate the genetic basis for the shifted spectrum.

subsequent cycles. We introduced 1 ng plasmid DNA in the first cycle and could generate more than 10^5 clones. To maintain the same library size in subsequent rounds, 200 ng DNA was required because only a small proportion of DNA molecules in the digested mixture were circular, resulting in greatly reduced transformation efficiencies. We performed MIPE in an iterative manner for three cycles while always maintaining large libraries with ~10^5 clones.

To determine the effect of sequence diversification, plasmid mixtures from each round of MIPE were transformed into MG1655 and their genotype were assayed by MASC-PCR.[25] MIPE could generate an average of 0.63 mutations and 43% mutant plasmids (plasmid with at least one mutation) in the population in a single round. These numbers increased to 1.13 mutations and 78% mutants after three rounds (Figure 5b and c). With RD CoS, the average number of mutations generated in the first round was about two times larger than that without RD CoS, which was again in agreement with the CoS factor of 2 observed in previous experiments.

Diversified pRib23 plasmids from the all three rounds of MIPE were transformed into MG1655 to screen for clones with improved riboflavin production. By screening 10^4 colonies for increased production of the yellow pigment by visual inspection, we isolated nine overproducing clones for further phenotyping (charactering riboflavin production) and genotyping (sequencing RBS regions). Among the nine clones, seven of them exhibited more than 2-fold improved riboflavin production (Figure 5d). The best one showed a 2.67-fold improvement compared with the parental clone. Sequencing of these nine clones revealed that none of them were identical; yet the best eight of them collectively had mutations in the RBS regions of both *ribC* and *ribB* (Figure 5d). Further analysis of the sequences by RBS calculator revealed that they were all relatively strong RBS sequences (Table 1), suggesting that the

dx.doi.org/10.1021/sb400051t | *ACS Synth. Biol.* 2013, 2, 651−661

Table 2. AR Frequencies (%) for Diversifying *rfp* Sequence

	regions[a]										
	A−	A	B	C	D	E	F−	F	G	G+	avg.
round 1[b]	0.0	27.3	4.5	45.5	36.4	13.6	0.0	0.0	18.2	9.1	14.8
round 2	0.0	20.0	10.0	45.0	65.0	35.0	0.0	10.0	45.0	15.0	26.1
round 3	4.2	37.5	16.7	54.2	83.3	54.2	0.0	0.0	45.8	20.8	31.3

[a]The 41 targeted nucleotides were grouped into 10 different regions, since mutations within the same region were very likely to be introduced simultaneously.[33] The regions covered by the same oligo are distinguished by adding a "+" or "−". [b]Two independent replicates and at least 20 clones were sequenced for each sample.

overexpression of *ribC* and *ribB* is critical to riboflavin overproduction and the effect of these two genes may be synergistic. However, for the variant performances of riboflavin production among the overproducing clones, no general relationship can be identified between the production and the predicted strengths of the mutant RBSs (Table 1). In addition, the mutant RBSs are not necessarily among the strongest in the synthetic libraries to achieve the best selected performance. This case reinforces the importance of exploiting library based methods to fine-tune synthetic gene networks to achieve optimal performances.

MIPE for Diversification of Protein Coding Sequence. Generating high quality libraries is the initial and critical step for successful protein directed evolution efforts. Previously, protein coding sequence diversification depends heavily on error-prone polymerase chain reaction (epPCR) or DNA shuffling.[26] Targeted mutagenesis (including saturation mutagenesis) has also been performed to introduce well designed mutations using overlap extension PCR (OE-PCR)[27,28] and QuikChange.[29] However, OE-PCR are inconvenient to introduce two mutations that are ∼100 bp apart or to introduce multiple mutations within a short region (<500 bp) and are usually limited to target less than 3 sites in a single reaction. QuikChange is independent of PCR reactions but has several disadvantages such as low efficiency and accuracy.[30] A recently developed methods, "Multichange Isothermal Mutagenesis", overcomes several technical problems of QuikChange but still suffers from the inherent limitations of PCR based methods.[30]

Here, we challenged MIPE to simultaneously targeted 41 nucleotides (23 codons)[31] (SI Table S3) scattered along the 750 bp *rfp* gene sequence on pRFP07, which was extremely difficult to be implemented with OE-PCR or QuikChange. To this end, seven oligos, which together covered 85% of the total *rfp* sequence, were design to introduce degenerate or defined mutations on the selected targets (SI Table S4). We performed three rounds of MIPE while always keeping ∼10^6 library size by introducing 10 ng pRFP07 in the first cycle and 200 ng DNA mixture in the next two cycles. Plasmid libraries generated from each cycle were transformed into MG1655 to assess the effect of mutagenesis.

MIPE could very effectively alter the functionality of *rfp*, as reflected by the diversified colors and intensities of the colonies (Figure 6b). MIPE introduced 1.5, 2.3, and 2.9 mutations (all modified nucleotides covered by the same oligo were regarded as one mutation) per clone and yielded on average 15%, 26%, and 31% AR frequency for the targeted codons after the first, second and third round, respectively (Table 2). Highly broadened AR frequency distributions were also generated at that same time (Figure 6c). Among a total of 60 sequenced mutants, 54 clones were unique, indicating good diversities. (The identical mutant clones were those that have a single modified position with no degeneracy or low degeneracy.) We isolated and sequenced nine clones with very similar blue-shifted orange colors to explore the genetic bases of the modified spectrum (Figure 6d). Surprisingly, none of the nine clones contained mutations around Q66, which was reported to be the chromophore position and could alter the spectrum in fluorescent proteins derived from *rfp* or *gfp*.[31,32] Instead, 8 of 9 clones contain mutations on position C (SI Table S5), suggesting the amino acids in this region may have contributed to the blue-shifted fluorescence.

This case demonstrated the ability of MIPE for simultaneous mutagenesis of up to 10 positions in 750 bp protein coding sequences. The uniqueness of MIPE for protein directed evolution lies in that it is a semirational approach, which is more focused than epPCR and more multiplexed than targeted mutagenesis. On one hand, it avoids mutations on extremely conserved or unrelated regions, thus creating more customized and higher quality libraries. On the other hand, it simultaneously targets many sites (10 in our case, but we see no limitations for more and can be up to 20 in genome engineering according to Lajoie[34]), which allows greatly enhanced ability for combinatorially exploring protein sequences.

AR frequencies exhibited a high degree of variability among targets (Table 2). Although introduced by the same oligo, mutations on position A− and G+ were introduced at decreased frequencies than those on A and G, which was likely due to the oligo chew-back phenomenon observed by Wang.[33] Oligo F exhibited extremely low AR frequency possibly due to the complex interactions between different oligos in the large oligo pool.[35] The AR efficiencies for diversifying pRFP07 (introducing 1.5 mutations per round) were at the same level to that of multiplexed genome engineering (introducing 1−2 mutations per round),[20,34] whereas that for pRib23 (introducing only 0.63 mutations per round) were much lower than genome engineering. Two factors may have contributed to the reduced efficiencies for pRib23 engineering. First, pRib23 has p15A replication origin, which showed lower AR frequencies than pUC18 plasmid under similar conditions. Second, the oligos diversifying pRib23 were designed to have more mismatches on average than that for diversifying *rfp* or replacing stop codons. (7.8 mismatches for pRib23, 5.4 for pRFP07, and 1 for replacing stop codons).[20]

We have attempted to perform more than three cycles for the diversification of both pRib23 and *rfp* sequences but could not maintain good library quality or high complexity due to the formation of "tough" plasmids (SI Discussion). However, we successfully solved this problem and performed MIPE for seven cycles by using multiple 6 bp regions as RD CoS markers instead of a single site (SI Figure S4).

Unique Features of MIPE. Compared with approaches based on DNA assembly techniques, MIPE offers several unique aspects for the construction of plasmid libraries. First,

MIPE is based on OMR and is therefore free of PCR amplifications. PCR is required for most DNA assembly techniques and may produce several concerns or limitations in many conditions. To begin with, PCR introduces unwanted mutations at much higher rate than *in vivo* DNA replication, which becomes an especial concern when engineering multigene synthetic pathways on large constructs. Although, to achieve optimal efficiency, OMR was usually performed in MMR-deficient strains, which have higher replication mutation rate than strains with active MMR systems, MIPE still allows much lower mutation rate (2.5×10^{-8} for MMR-deficient strains according to Isaacs[25]) than PCR even with the highest fidelity DNA polymerases (4.4×10^{-7} for Phusion DNA Polymerase). However, another source of mutations can arise at sufficient frequency from chemical synthesis of long (>80mer) oligos, which could be a major source of mutations for MIPE in our experience. Furthermore, introducing combinations of mutations by PCR-based methods is often constrained by the upper limit of primer synthesis and the lower limit of reasonable PCR product length. MIPE, on the other hand, is highly flexible and customized for introducing combinations of mutations; two mutations within 40 bp can be introduced by a single primer while those that are 40 bp apart can be introduced by separate primers. Finally, unlike DNA assembly methods that require several pairs of primers to PCR amplify individual parts, MIPE requires only a single primer for each mutation (or several adjacent mutations).

In addition, MIPE introduces mutations on existing plasmids and avoids the reconstruction of the entire plasmid molecules, whereas DNA assembly methods for creating plasmid libraries are intrinsically *de novo* plasmid construction procedures and thus suffer from low efficiency and accuracy. The library's size generated with assembly methods often decreases dramatically ($<10^2$ colonies per reaction) when assembling more than three pieces, although the size may be acceptable when constructing small molecules with fewer pieces.[5] MIPE, on the other hand, could theoretically construct libraries with more than 10^7 clones under optimal conditions. (Introducing a 100 ng plasmid with optimal efficiency of 10^6 cfu/ng plasmid DNA gives 10^8 clones, but we have observed ~10 fold decreased transformation efficiency with high oligo concentration.) MIPE also avoids mis-assembly errors, which are always present in the libraries generated by assembly methods.[40] A specific concern of MIPE is that plasmid multimers may form during recombineering process.[21] We have identified four multimeric plasmid molecules among the 70 isolated clones in the pRFP07 libraries, which may not be a problem given the large library size generated.

Another unique feature of MIPE is that it generates diversified sequences with combinations of wild type and mutant alleles. In the population generated by MIPE, the fraction of mutations on a certain site is determined by its AR frequency per cycle and the number of cycles performed. This feature actually allows MIPE to control the average divergence of the resulting population and thus to construct more customized plasmid libraries. For example, in experiments when many sites are tested for a certain function, more multiplexed oligos (up to 20 according to Lajoie[34]) can be introduced with less MIPE cycles, allowing low average AR frequency on each site and thus avoiding the abolishment of the function resulted from too many combined mutations on a single clone. When only few sites (e.g., less than 4) are tested for combined effects, high AR frequency on each site can be achieved by introducing less multiplexed oligos with increased MIPE cycles. However, if only the final clone with all targeted mutations is required, assembly methods are more favorable than MIPE.

A major concern of MIPE is that increased length of introduced mutation would decrease the recombineering frequency.[2] However, this length is sufficient to cover the core region many regulatory elements in bacterial such as promoters,[36] short protein tags,[37] 5′UTR,[2,38] intergenic regions,[11] and various kinds of RNA devices.[39] In addition, the AR efficiencies may be further increased by decreasing oligo degradation through nuclease inactivation and increasing the Okazaki fragment length.[34,35] Another disadvantage is that the AR frequencies may vary greatly among targets. Further experiments should be performed in detail to elucidate the factors affecting the recombination frequency.

It is worth noting that the improved clones can be further optimized with MIPE for both diversifying synthetic pathways and protein coding sequences. The recombineering oligos can be redesigned based on the knowledge obtained from previous optimization stages. In this manner, cycles of diversification, screening, and characterization process can be performed for deeply understanding the genotype-phenotype relationship of certain biological functions.

An important reason for us to choose the riboflavin pathway and the *rfp* gene as the target for MIPE is that their mutation effects can be easily detected and screened. However, MIPE has the ability to produce high quality libraries and thus does not necessarily require any high-throughput screening technologies for the exploration of the libraries (similar to what is required for chromosome random mutagenesis or epPCR). However, high-throughput screen technologies are developing very fast with the emergence of some novel and widely applicable strategies.[40−42] Further considering the ability of MIPE to generate very large size libraries, we believe the combination of MIPE with high-throughput screen technologies will be very powerful for creating and optimizing novel biological functions.

Overall, MIPE is a highly efficient and flexible method for combinatorial mutagenesis of plasmid sequences. Its unique features such generating large library size and avoiding PCR amplifications could benefit many applications which are previously challenging to perform using DNA assembly or site directed mutagenesis methods. In addition, this approach is accessible and cost-effective for all laboratories because it requires only the recombineering strains and synthetic primers. We anticipate MIPE to be widely used for the optimization of diverse synthetic biological designs.

METHODS

Methods for plasmid construction and MASC-PCR for assaying the genotype of pRib23 libraries can be found in the Supporting Information

Strains, Media, and Reagents. All recombineering experiments were performed in HME68, which had a genotype of W3110 Δ(*argF-lac*)*U169 galK*$_{tyr145UAG}$ *mutS*<>*cat* [λ cI857 Δ(*cro-bioA*)] and was a gift from Professor Court.[21] *E. coli* MG1655 from CGSC was used for phenotyping of the diversified plasmids for both pathway optimization and protein evolution. *E. coli* DH5α was used for molecule cloning. LB media was used for cell growth in most cases. Ampicillin (Amp), Kanamycin (kan), Tetracycline (Tet), and Spectinomycin (Spc) were added at concentrations of 100 mg/mL, 10 mg/mL, 30 mg/mL, 50 mg/mL, respectively.

Primers. SI Table S6 presents a full list of primers used in this study. All oligos were obtained from Sangon Biotech. Oligos for AR were usually 80−89 bp in length and contained four phosphorothioate linkages at the 5′ terminus unless designated otherwise.[2] The degenerate sequences in the oligos modifying the RBS regions were designed with the "Genome Editing" function of "RBS Calculator".[10] The Max Length of Mutation Box was usually set to 10. The Minimum and Maximum Translation Initiation Rates were set according to the strength of the native RBS sequences to produce RBS strengths over a range of more than 10^3. An estimated library size was usually set to 13 to 26 sequences. Primers for CPEC were designed with J5.[43]

Recombineering Methodology. Recombineering was done according to previously described procedures.[18] Briefly, strains containing the defective λ prophage (HME68 in this work) were grown at 32 °C to midlog (OD_{600} = 0.4−0.6). The cultures were shifted to 42 °C for 15 min to induce λ Red and then quick-chilled in ice water slurry for 10 min. For electroporation preparation, the cells were washed twice with cold-sterile dH_2O and finally concentrated 100-fold in 10% glycerol (final OD_{600} = 40). To modify plasmids with coelectroporation, plasmids and oligos were mixed in 50 μL electrocompetent cells for electroporation. For targeting resident plasmids, each plasmid was transformed into the HME68 cells prior to recombineering. The preparation of electrocompetent cells harboring plasmids was identical to that without plasmids except the addition of antibiotics during overnight growth for plasmid selection. An Eppendorf electroporator was used for electroporation (0.1 cm cuvette, 1.80 kV). After electroporation, cells were allowed to recover for 2 h in LB before adding antibiotics for plasmid selection or plating for the estimation of library size. Cells containing λ prophage were always incubated at 32 °C except when inducing λ Red. The concentrations of plasmids and oligos used in recombineering varied between different experiments and are described in details in the main text.

Calculation of CoS Factors with pLX15. To characterize the coselection effect, primers CoS_Spc_on, CoS_Kan_on, and CoS_Tet_on were coelectroporated with pLX15. The harvested plasmids were retransformed, and the resulting cell mixtures were plated with different combinations of antibiotics for the calculation of AR frequencies. The event of restoring the Spc, Kan and Tet resistant phenotype by recombineering is denoted as Spc^+, Kan^+, and Tet^+, respectively. Thus, the AR efficiency of restoring both Spc and Kan, for example, can be denoted as $p(Spc^+Kan^+)$ and can be calculated by the ratio of the colony numbers on plate containing Spc and Kan to that containing Amp. CoS factors were calculated by the ratio of AR frequencies with Cos to that without CoS. For example, the effect of Kan^+ coselection on Spc^+ is calculated as CoS Factor($Spc^+|Kan^+$) = $p(Spc^+|Kan^+)/p(Spc^+)$, where $p(Spc^+|Kan^+)$ is the conditional probability of Spc^+ in the case of Kan^+ and can be calculated as $p(Spc^+|Kan^+)=p(Spc^+Kan^+)/p(kan^+)$. If both Kan^+ and Tet^+ are coselected, the conditional AR efficiency is $p(Spc^+|Kan^+Tet^+) = p(Spc^+Kan^+Tet^+)/p(Kan^+Tet^+)$ and the CoS Factor($Spc^+|Kan^+Tet^+$) = $p(Spc^+|Kan^+Tet^+)/ p(Spc^+)$.

Performing MIPE. In each MIPE cycle, plasmid DNA, which can be either pure plasmid to initial the first cycle or digested plasmid mixture obtained from previous cycles, was coelectroporated with oligo mixtures into electrocompetent recombineering cells. More than 10 ng plasmid in the first cycle and more than 200 ng DNA in subsequent cycles were introduced, which could generate a library size of ∼10^5 for diversifying pRib23 and ∼10^6 for pRFP07. 0.5 μM of each recombineering oligo were used while RD CoS oligos were added to 1/20 of the total oligo concentration. The resulting cell mixtures after electroporation were incubated in 5 mL LB for 2 h, followed by overnight growth with appropriate antibiotics and plating to estimate the library size. Plasmid mixtures were isolated with Plasmid Miniprep Kit from Sangon Biotech and the concentrations were measured by absorbance at 260 nm. For restriction digestion, Thermo Scientific FastDigest Enzymes were used with digestion systems set up according to the manufacture's protocols. Typically, less than 1 μg DNA was digested in 40 μL systems for 0.5 h to ensure complete digestion. The digested mixtures were purified with DNA Purification Kit from Sangon Biotech and transformed to initiate the next MIPE cycle.

Screening and Characterizing Riboflavin Overproduction Strains. Diversified pRib23 plasmids were transformed into MG1655 for screening overproducing clones. The resulting cells mixtures were appropriately diluted and plated on LB plate with Amp and IPTG and incubated at 37 °C for 24 h before screening for improved clones. An overall of 10^4 colonies were screened for increased production of the yellow pigment by visual inspection. Nine overproducing clones were picked, which were then incubated in 3 mL Amp-LB media overnight for plasmid isolation and further characterization. To assess the productivity of the isolated clones, the overnight cultures were diluted to OD_{600} = 0.02 to inoculate 3 mL Amp-LB media supplemented with 0.2 mM IPTG in 15 mL test tubes. The final titer of riboflavin production was measured by the OD_{444} after 12 h of production. Isolated plasmids were sequenced to determine their genotype.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary methods, discussion, Figures S1−S4, and Tables S1−S6. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: chentao@tju.edu.cn.

Author Contributions

†These authors contributed equally to this work.

Author Contributions

Y.L. designed the experiments; Y.L., Q.G., and Z.L. performed the experiments; Y.L. and T.C. wrote the manuscript; T.C. and X.Z. supervised the work; and all authors contributed to the discussion of the research.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Donald Court and Lynn Thomason for providing HME68 cells. We thank Peter Carr for helpful discussion on CoS MAGE. We thank Howard Salis for assistance with RBS calculator. This work was supported by National 973 Project [2011CBA00804, 2012CB725203]; National Natural Science Foundation of China [NSFC-21176182, NSFC-21206112]; and National High-tech R&D Program of China [2012AA02A702, 2012AA022103].

ABBREVIATIONS

MIPE, Multiplex Iterative Plasmid Engineering; MAGE, Multiplex Automatable Genome Engineering; AR, allelic replacement; OMR, oligo mediated λ Red recombineering; CoS, coselection; RD CoS, restriction digestion mediated coselection; WT, wild type; Amp, ampicillin; Kan, kanamycin; Tet, tetracycline; Spc, spectinomycin

REFERENCES

(1) Santos, C. N., and Stephanopoulos, G. (2008) Combinatorial engineering of microbes for optimizing cellular phenotype. *Curr. Opin. Chem. Biol. 12*, 168−176.

(2) Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., and Church, G. M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature 460*, 894−898.

(3) Yokobayashi, Y., Weiss, R., and Arnold, F. H. (2002) Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16587−16591.

(4) Du, J., Yuan, Y., Si, T., Lian, J., and Zhao, H. (2012) Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Res. 40*, e142.

(5) Wingler, L. M., and Cornish, V. W. (2011) Reiterative recombination for the *in vivo* assembly of libraries of multigene pathways. *Proc. Natl. Acad. Sci. U.S.A. 108*, 15135−15140.

(6) Ellis, T., Wang, X., and Collins, J. J. (2009) Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol. 27*, 465−471.

(7) Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science 330*, 70−74.

(8) Leonard, E., Ajikumar, P. K., Thayer, K., Xiao, W. H., Mo, J. D., Tidor, B., Stephanopoulos, G., and Prather, K. L. (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U.S.A. 107*, 13654−13659.

(9) Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U.S.A. 102*, 12678−12683.

(10) Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol. 27*, 946−950.

(11) Pfleger, B. F., Pitera, D. J., Smolke, C. D., and Keasling, J. D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol. 24*, 1027−1032.

(12) Alper, H., and Stephanopoulos, G. (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab. Eng. 9*, 258−267.

(13) Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat. Methods 4*, 251−256.

(14) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.

(15) Shao, Z., and Zhao, H. (2009) DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res. 37*, e16.

(16) Quan, J., and Tian, J. (2009) Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One 4*, e6441.

(17) Ramon, A., and Smith, H. O. (2011) Single-step linker-based combinatorial assembly of promoter and gene cassettes for pathway engineering. *Biotechnol. Lett. 33*, 549−555.

(18) Sharan, S. K., Thomason, L. C., Kuznetsov, S. G., and Court, D. L. (2009) Recombineering: A homologous recombination-based method of genetic engineering. *Nat. Protoc. 4*, 206−223.

(19) Costantino, N., and Court, D. L. (2003) Enhanced levels of λ Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A. 100*, 15748−15753.

(20) Carr, P. A., Wang, H. H., Sterling, B., Isaacs, F. J., Lajoie, M. J., Xu, G., Church, G. M., and Jacobson, J. M. (2012) Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Res. 40*, e132.

(21) Thomason, L. C., Costantino, N., Shaw, D. V., and Court, D. L. (2007) Multicopy plasmid modification with phage λ Red recombineering. *Plasmid 58*, 148−158.

(22) Sawitzke, J. A., Costantino, N., Li, X. T., Thomason, L. C., Bubunenko, M., Court, C., and Court, D. L. (2011) Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J. Mol. Biol. 407*, 45−59.

(23) Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 97*, 5978−5983.

(24) Shi, S., Chen, T., Zhang, Z., Chen, X., and Zhao, X. (2009) Transcriptome analysis guided metabolic engineering of *Bacillus subtilis* for riboflavin production. *Metab. Eng. 11*, 243−252.

(25) Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., Tolonen, A. C., Gianoulis, T. A., Goodman, D. B., Reppas, N. B., Emig, C. J., Bang, D., Hwang, S. J., Jewett, M. C., Jacobson, J. M., and Church, G. M. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science 333*, 348−353.

(26) Bloom, J. D., and Arnold, F. H. (2009) In the light of directed evolution: Pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U.S.A. 106* (Suppl 1), 9995−10000.

(27) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*, 51−59.

(28) Miyazaki, K., and Arnold, F. H. (1999) Exploring nonnatural evolutionary pathways by saturation mutagenesis: Rapid improvement of protein function. *J. Mol. Evol. 49*, 716−720.

(29) Hogrefe, H. H., Cline, J., Youngblood, G. L., and Allen, R. M. (2002) Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit. *Biotechniques 33*, 1158−1160, 1162, 1164−1165.

(30) Mitchell, L. A., Cai, Y., Taylor, M., Noronha, A. M., Chuang, J., Dai, L., and Boeke, J. D. (2013) Multichange isothermal mutagenesis: A new strategy for multiple site-directed mutations in plasmid DNA. *ACS Synth Biol. 2*, 473−477.

(31) Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004) Improved monomeric red, orange, and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol. 22*, 1567−1572.

(32) Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 91*, 12501−12504.

(33) Wang, H. H., Xu, G., Vonner, A. J., and Church, G. (2011) Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. *Nucleic Acids Res. 39*, 7336−7347.

(34) Lajoie, M. J., Gregg, C. J., Mosberg, J. A., Washington, G. C., and Church, G. M. (2012) Manipulating replisome dynamics to enhance λ Red-mediated multiplex genome engineering. *Nucleic Acids Res. 40*, e170.

(35) Mosberg, J. A., Gregg, C. J., Lajoie, M. J., Wang, H. H., and Church, G. M. (2012) Improving λ Red genome engineering in *Escherichia coli* via rational removal of endogenous nucleases. *PLoS One 7*, e44638.

(36) Wang, H. H., Kim, H., Cong, L., Jeong, J., Bang, D., and Church, G. M. (2012) Genome-scale promoter engineering by coselection MAGE. *Nat. Methods 9*, 591−593.

(37) Wang, H. H., Huang, P. Y., Xu, G., Haas, W., Marblestone, A., Li, J., Gygi, S., Forster, A. C., Jewett, M. C., and Church, G. M. (2012) Multiplexed *in vivo* His-tagging of enzyme pathways for *in vitro* single-pot multi-enzyme catalysis. *ACS Synth. Biol. 1*, 43−52.

(38) Egbert, R. G., and Klavins, E. (2012) Fine-tuning gene networks using simple sequence repeats. *Proc. Natl. Acad. Sci. U.S.A. 109*, 16817−16822.
(39) Isaacs, F. J., Dwyer, D. J., and Collins, J. J. (2006) RNA synthetic biology. *Nat. Biotechnol. 24*, 545−554.
(40) Peralta-Yahya, P., Carter, B. T., Lin, H., Tao, H., and Cornish, V. W. (2008) High-throughput selection for cellulase catalysts using chemical complementation. *J. Am. Chem. Soc. 130*, 17446−17452.
(41) Liang, J. C., Chang, A. L., Kennedy, A. B., and Smolke, C. D. (2012) A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res. 40*, e154.
(42) Sharma, V., Yamamura, A., and Yokobayashi, Y. (2011) Engineering artificial small RNAs for conditional gene silencing in *Escherichia coli*. *ACS Synth. Biol. 1*, 6−13.
(43) Hillson, N. J., Rosengarten, R. D., and Keasling, J. D. (2011) j5 DNA assembly design automation software. *ACS Synth. Biol. 1*, 14−21.