ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

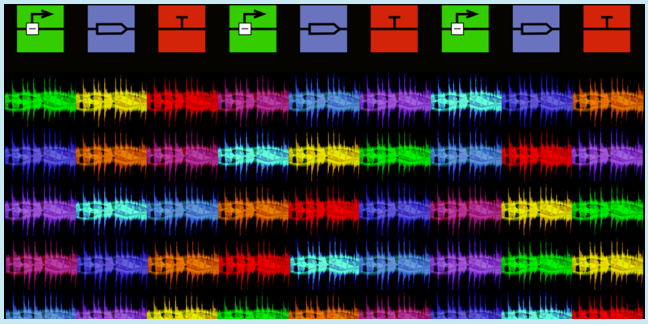
Randomized BioBrick Assembly: A Novel DNA Assembly Method for Randomizing and Optimizing Genetic Circuits and Metabolic Pathways

Sean C. Sleight* and Herbert M. Sauro

Department of Bioengineering, University of Washington, Seattle, Washington 98195, United States

Supporting Information

ABSTRACT: The optimization of genetic circuits and metabolic pathways often involves constructing various iterations of the same construct or using directed evolution to achieve the desired function. Alternatively, a method that randomizes individual parts in the same assembly reaction could be used for optimization by allowing for the ability to screen large numbers of individual clones expressing randomized circuits or pathways for optimal function. Here we describe a new assembly method to randomize genetic circuits and metabolic pathways from modular DNA fragments derived from PCR-amplified BioBricks. As a proof-of-principle for this method, we successfully assembled CMY (Cyan-Magenta-Yellow) three-gene circuits using Gibson Assembly that express CFP, RFP, and YFP with independently randomized promoters, ribosome binding sites, transcriptional terminators, and all parts randomized simultaneously. Sequencing results from 24 CMY circuits with various parts randomized show that 20/24 circuits are distinct and expression varies over a 200-fold range above background levels. We then adapted this method to randomize the same parts with enzyme coding sequences from the lycopene biosynthesis pathway instead of fluorescent proteins, designed to independently express each enzyme in the pathway from a different promoter. Lycopene production is improved using this randomization method by about 30% relative to the highest polycistronic-expressing pathway. These results demonstrate the potential of generating nearly 20,000 unique circuit or pathway combinations when three parts are permutated at each position in a three-gene circuit or pathway, and the methodology can likely be adapted to other circuits and pathways to maximize products of interest.



KEYWORDS: *synthetic biology, DNA assembly, BioBricks, genetic circuits, combinatorial, randomized*

DNA assembly methods[1−11] are essential for most synthetic biology research and applications due to the high cost of DNA synthesis for constructing genetic circuits and metabolic pathways. The optimization of circuits and pathways often requires constructing various iterations of the same construct or using directed evolution to achieve the desired function.[12−17] Alternatively, a combinatorial assembly method that randomizes individual parts, instead of individual nucleotides as in directed evolution, could be used for optimization. Combinatorial assembly of parts requires that each part of a particular type (*e.g.*, promoters, coding sequences, transcriptional terminators) has the same compatible or modular overlap sequences for *independent* assembly. If multiple part types with modular overlaps are mixed together in the same assembly reaction, this allows for randomized assembly since there is some probability that each part will be assembled into a unique circuit or pathway. Randomized assembly methods must produce multiple randomized constructs that can be screened or selected for simultaneously, whereas combinatorial assembly methods may or may not produce multiple constructs simultaneously (often one unique construct is produced for each assembly reaction). The benefit of a randomized assembly method is that the experimenter could either decide an exact circuit to assemble from individual parts (*e.g.*, promoter A, ribosome binding site A, coding sequence A, and terminator A), randomize a particular part type (*e.g.*, promoters A, B, and C), or randomize all parts at the same time. Such a randomization method gives the experimenter access to a parts library as individual purified PCR products that exist in the freezer without the need for processing (hence making BioBricks into something more closely resembling "DNA Legos") that can be assembled randomly. Hence, a major motivation for this assembly method is to break down circuits and pathways into individual units and be able to reassemble these parts in a modular way, both quickly, and affordably.

The synthetic biology combinatorial assembly methods developed to date[1,3,7,9,11,18] use traditional cloning methods (*e.g.*, restriction enzymes and DNA ligase) to create compatible overlapping sequences between DNA fragments to assemble. The minimal overlap size of restriction enzyme-based assembly

Received: May 10, 2013
Published: July 10, 2013

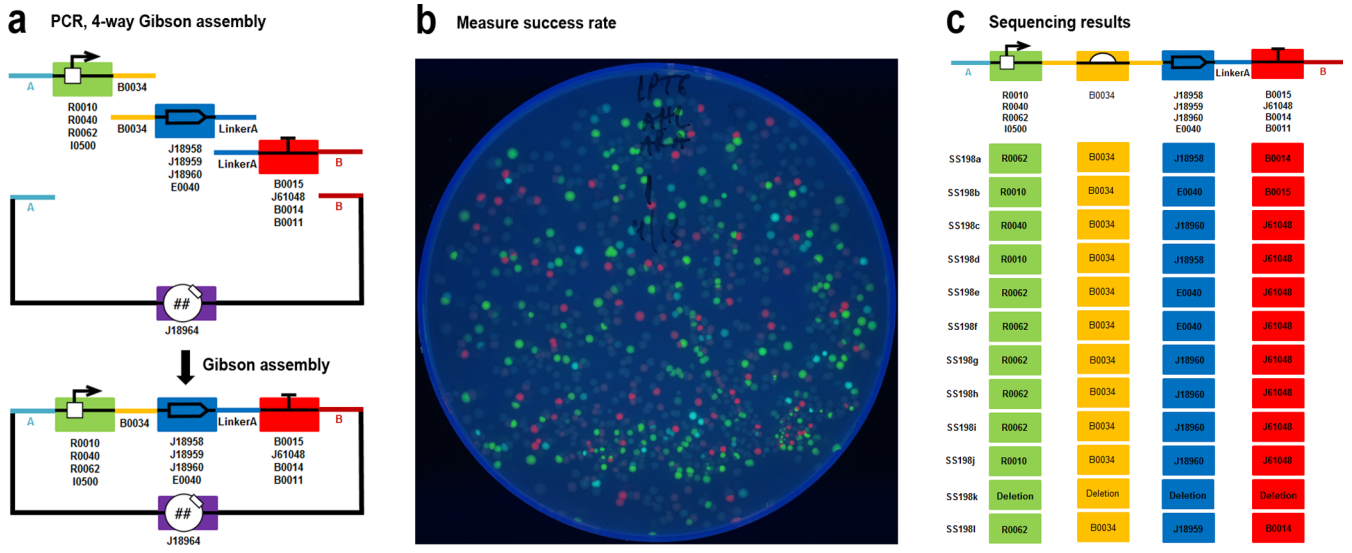


Figure 1. One-gene randomization methodology and results. (a) Four promoters (green box with regulation symbol and bent arrow), four coding sequences (blue right-facing arrow), and four transcriptional terminators (red "T" symbol) are randomized simultaneously into a vector (purple circle with ## and box symbol) to assemble randomized one-gene circuits. The colored lines on the side of each part illustrate standardized overlap sequences in Supplementary Table 1, where lines of the same color represent complementary, overlapping sequences. BioBrick part numbers (BBa_) from the Registry of Standard Biological Parts are shown for each part. Each promoter (pLacZYA = R0010, pTetR = R0040, pLuxR = R0062, pBAD = I0500) is PCR-amplified with an "A" prefix for a left-overlap and B0034 RBS for a right-overlap. Each coding sequence (J18958 = eYFP, J18959 = maxRFP, J18960 = eCFP, and E0040 = GFPmut3b) has the complementary B0034 RBS for a left-overlap and "LinkerA" for the right-overlap. Each terminator (B0015, J61048, B0014, B0011) has the complementary "LinkerA" for the left-overlap and "B" suffix for the right-overlap. The J18964 = pGA3K4 vector was amplified with the complementary "A" prefix and "B" suffix for left- and right-overlaps, respectively. The 13 purified PCR products were added to the same Gibson assembly reaction to generate randomized one-gene circuits. (b) The qualitative success rate (number of fluorescing colonies/total number of colonies) is measured by plating the transformants on selective media and visualizing the colonies under UV light. (c) Sequencing results show that 3/4 promoters, 4/4 coding sequences, and 3/4 terminators added to the assembly reaction are identified among 12 sequenced clones (SS198A-L), generating 8/11 distinct circuits (SS198K has a deletion).

is a distinct advantage for combinatorial methods due to the critical spacing needed between individual parts, such as that between the ribosome binding site (RBS) and coding sequence. However, there are disadvantages to restriction enzyme-based assembly such as the removal of restriction sites, gel extractions (in some cases), and relatively low assembly success rates. In contrast, recombination-based assembly methods such as Gibson Assembly,[6] In-Fusion,[8,19] and SLIC[20] join DNA fragments with complementary overlap sequences and have the advantage of not requiring removal of restriction sites before assembly, can generate scarless assembly between parts (if desired), and have relatively high assembly success rates in our hands. However, recombination-based assembly methods have the disadvantage of requiring at least a 15 bp overlap between DNA fragments, and there is normally the need to design primers for every unique construct to assemble. Thus, we wanted to create a recombination-based randomized assembly method that (1) creates multiple unique constructs per assembly reaction, (2) randomizes circuits and pathways without disrupting the critical spacing between parts, (3) enables plug-and-play assembly (*i.e.*, no processing is required once the part is engineered), and (4) can be easily adapted to other parts without any complicated rules (since each part type has the same exact standardized overlap sequences, it would be a matter of attaching these overlap sequences to parts of interest and going straight to assembly).

The Randomized BioBrick Assembly method described here requires each PCR product of a particular type (*e.g.*, promoters, coding sequences, and transcriptional terminators) to have a unique and sufficiently long overlap (18−28 bp) on either side of the part for modular, recombination-based assembly. In other words, there needs to be a standard overlapping "scar sequence" or "linker" in between parts so that each DNA fragment can be assembled independently of its functional sequence and in the correct order. By contrast, scarless-based assembly methods[18] have more flexibility over the sequences in between parts but normally require the generation of new parts for the assembly of every unique construct and do not encourage reuse of the same parts. The first step in developing this method was to assemble circuits with standardized scar sequences that function properly and have a steady-state expression level that is not affected by scar sequences. We first attempted to minimize the overlap between fragments to 8−12 bp to not change the crucial spacing between parts but determined that the success rate was too low and that the minimum overlap size needed to be at least 15 bp for routine use. Then we attempted to increase the length of the scar sequence that exists between genetic circuits assembled with Standard BioBrick assembly.[2,4,5] We found that increasing the scar sequence length to 15 or 20 bp between the promoter-RBS, and RBS-coding sequence junctions caused severely decreased steady-state expression, but a 20-bp scar sequence in between the coding sequence and terminator (using multiple versions) did not significantly affect expression levels. Therefore, we decided to use the existing Standard BioBrick Assembly scar sequences between the promoter-RBS and RBS-coding sequence to maintain proper spacing and create a new scar/overlap sequence for the coding sequence-terminator junction. Since the spacing between the RBS and coding sequence is short, we decided to use the RBS itself as the overlap sequence.

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506−518

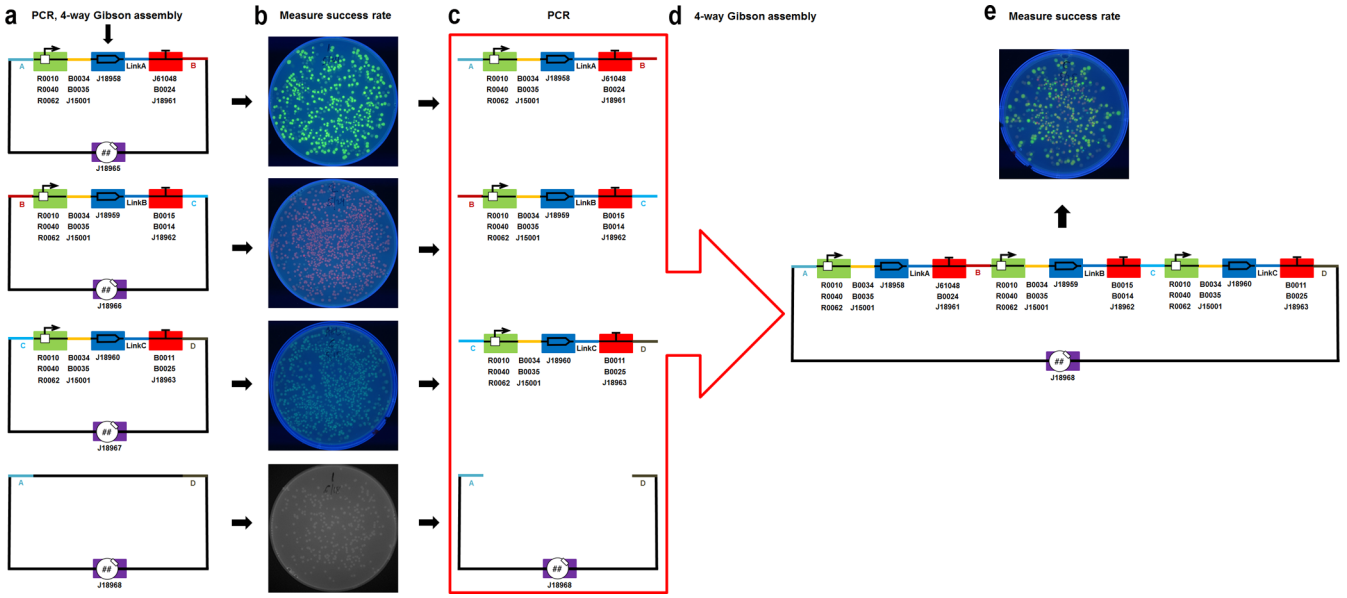


Figure 2. Three-gene randomization methodology. (a) Independent assembly of randomized eYFP, maxRFP, and eCFP one-gene circuits. In this example, promoters, RBSs, and terminators are simultaneously randomized. Each randomized J18958 = eYFP, J18959 = maxRFP, and J18960 = eCFP circuit is randomized with the same promoters (R0010, R0040, R0062) and RBSs (B0034, B0035, J15001) but have different terminators (J61048, B0024, J18961 for eYFP; B0015, B0014, J18962 for maxRFP; B0011, B0025, J18963 for eCFP). The randomized eYFP, maxRFP, and eCFP circuits are constructed using the J18965 = pSS3K1, J18966 = pSS3K2, and J18967 = pSS3K3 vectors, respectively, having unique prefix and suffix sequences. The J18968 = pSS3C1 vector without a functional insert is the vector used for three-gene randomization. (b) The qualitative success rate of randomized eYFP, maxRFP, and eCFP circuits is measured by plating the transformants on selective media and visualizing the colonies under UV light. As a negative control, colonies transformed with the pSS3C1 vector lacking a functional insert (used for three-gene randomization) are visualized under blue light. (c) Randomized eYFP, maxRFP, and eCFP circuits are selectively PCR-amplified from a diluted Gibson assembly reaction. The pSS3C1 vector is also PCR-amplified to create the vector for three-gene randomization. (d) A second 4-way Gibson assembly reaction is performed to combine the pool of randomized one-gene circuits with the pSS3C1 vector to construct randomized three-gene circuits that express eYFP, maxRFP, and eCFP. (e) The qualitative success rate of the assembly reaction is visualized under UV light, and multiple colonies are grown in selective media to measure the success rate (number of colonies expressing three fluorescent proteins/total number of colonies).

After these design considerations, particular parts were chosen to test this assembly method, and the assembly method was initially performed to randomize one-gene circuits (Figure 1) and later expanded (Figure 2) to randomize both three-gene circuits (Figures 3 and 4) and the lycopene biosynthesis pathway (Figure 5). In theory, if three promoters, RBSs, and terminators are simultaneously randomized at each position in a three-gene circuit or pathway, this allows for the potential of generating nearly 20,000 circuit combinations ($3^3 \times 3^3 \times 3^3$ = 19,683). Thus, our method is novel compared to other existing methods in the following three ways: (1) at the one-gene circuit level, randomized circuits are generated from a single assembly reaction (this requires optimization to add multiple part types per assembly reaction) and combined to generate randomized three-gene circuits; (2) promoters, RBSs, and terminators can be simultaneously randomized (coding sequences can also be randomized); and (3) Gibson assembly is used for this method instead of restriction enzymes and ligase. We expect that this randomization method will be useful for quickly producing circuits and pathways with high efficiency, optimizing metabolic flux to maximize products of interest (*e.g.*, biofuels) and is likely adaptable to other circuits and pathways in various organisms.

RESULTS AND DISCUSSION

Generation of Randomized One-Gene Circuits. An example of a one-gene randomized assembly, where four promoters, four coding sequences, and four transcriptional terminators were used to construct randomized one-gene circuits using Gibson assembly,[6] is shown in Figure 1. First, individual parts (promoters, coding sequences, and transcriptional terminators) were PCR-amplified with a standardized overlap that joins parts during the assembly reaction (Figure 1a, Table 1, Supplementary Tables 1−3). The promoter has a left-overlap consisting of a prefix on the vector and a RBS as the right-overlap. Likewise, the coding sequence has a left-overlap of the complementary RBS sequence and right-overlap of a unique 22-bp "linker" sequence. The terminator has a left-overlap of the complementary "linker" sequence and right-overlap of the suffix on the vector. Finally, the vector has the complementary prefix and suffix sequences as the left- and right-overlaps, respectively. In order to achieve randomized assembly, instead of adding one part type per assembly reaction (one promoter, one coding sequence, one terminator, and one vector), multiple fragments of the same part type are mixed together in the same assembly reaction (in this example 13 parts were mixed). The assembly reaction was transformed into competent cells, the transformants were plated out, and the colonies were visualized under UV light to measure the qualitative success rate (Figure 1b). The qualitative success rate of this 4-way assembly reaction using 13 parts is about 40% (352/888), based on the number of fluorescing colonies divided by the total number of colonies since the vector is lacking a functional insert. After functional screening, 11 clones expressing either GFP, YFP, CFP, and RFP and one non-functional clone (strain ID numbers SS198A−L) were sequenced to quantify the part distribution and success rate.

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506−518

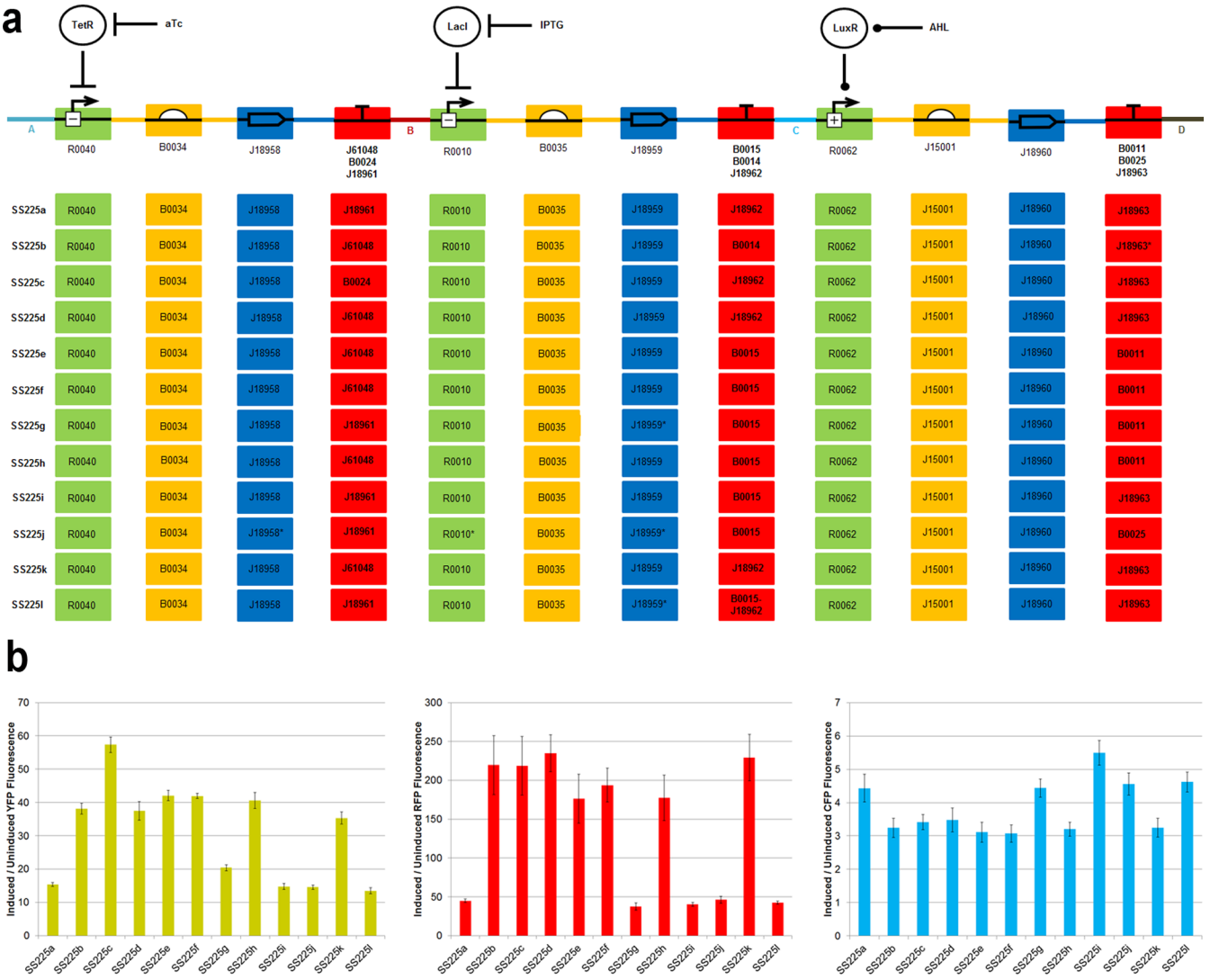


Figure 3. Sequencing results and characterization of CMY circuits with randomized terminators. (a) Sequencing results of 12 randomized CMY circuits with terminators randomized (strains SS225A–L) on the pSS3K4 vector. Nine of the 12 SS225 circuits are unique (SS225E = F = H, and SS225D = K). The regulation for each promoter is shown (see Methods), where regulation lines ending in a perpendicular line indicate inhibition and lines ending in a solid circle indicate activation. The terminators that were randomized are bolded and listed below each part symbol for a particular position in the circuit. Hybrid parts (two parts with homology that join together during the assembly reaction) are shown with two part names separated by a dash (*e.g.*, B0015-J18962), and starred parts indicate mutated sequences. See sequences of hybrid and mutated parts in Supplementary Table 4. (b) Characterization results of 12 randomized CMY circuits with terminators randomized. The left, center, and right panels show the induced/uninduced fluorescence for eYFP, maxRFP, and eCFP, respectively, in each of the 12 circuits (SS225A–L). The height of each bar represents the mean of eight experiments ± SD. See Methods for detailed information on circuit characterization.

Sequencing results identified 3/4 promoters, 4/4 coding sequences, and 3/4 terminators among these circuits, with 8/11 distinct circuits, demonstrating that multiple unique circuits can be generated using a single assembly reaction (Figure 1c).

There is, however, a bias of certain parts to assemble more frequently than others in this small sample size, and randomization is not 100% effective (a success rate of 72.7% in this example). Sequencing more plasmids from individual clones may have identified all 12 parts used in the assembly reaction, and using a larger sample size would have allowed us to rigorously determine whether there is a statistical bias for certain parts to be more frequently assembled than others. Using the small number of randomized circuits, we performed Chi-square tests to test whether the observed distribution of parts at each position in the circuit (promoters, coding sequences, and terminators) were consistent with the expected distribution and determined there was a significant bias for promoters and terminators (Table 2). The non-functional clone sequenced (and several others later sequenced) have a dominant mutation that involves amplification of prefix and suffix restriction sites located on the pGA3K3 vector, generating an inverted repeat that caused self-ligation of the vector (see Methods and Supplementary Table 4 for the exact sequence). It was therefore decided to create new vectors that lacked any restriction sites in the prefix/suffix to maximize success rates for expanding the method to three-gene randomization (see Methods).

Optimizing One-Gene and Three-Gene Assembly Reactions. In contrast to Figure 1 where promoters, coding sequences, and terminators were randomized simultaneously, we ultimately wanted to simultaneously randomize promoters, RBSs, and terminators in three-gene circuits, while keeping the

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506–518

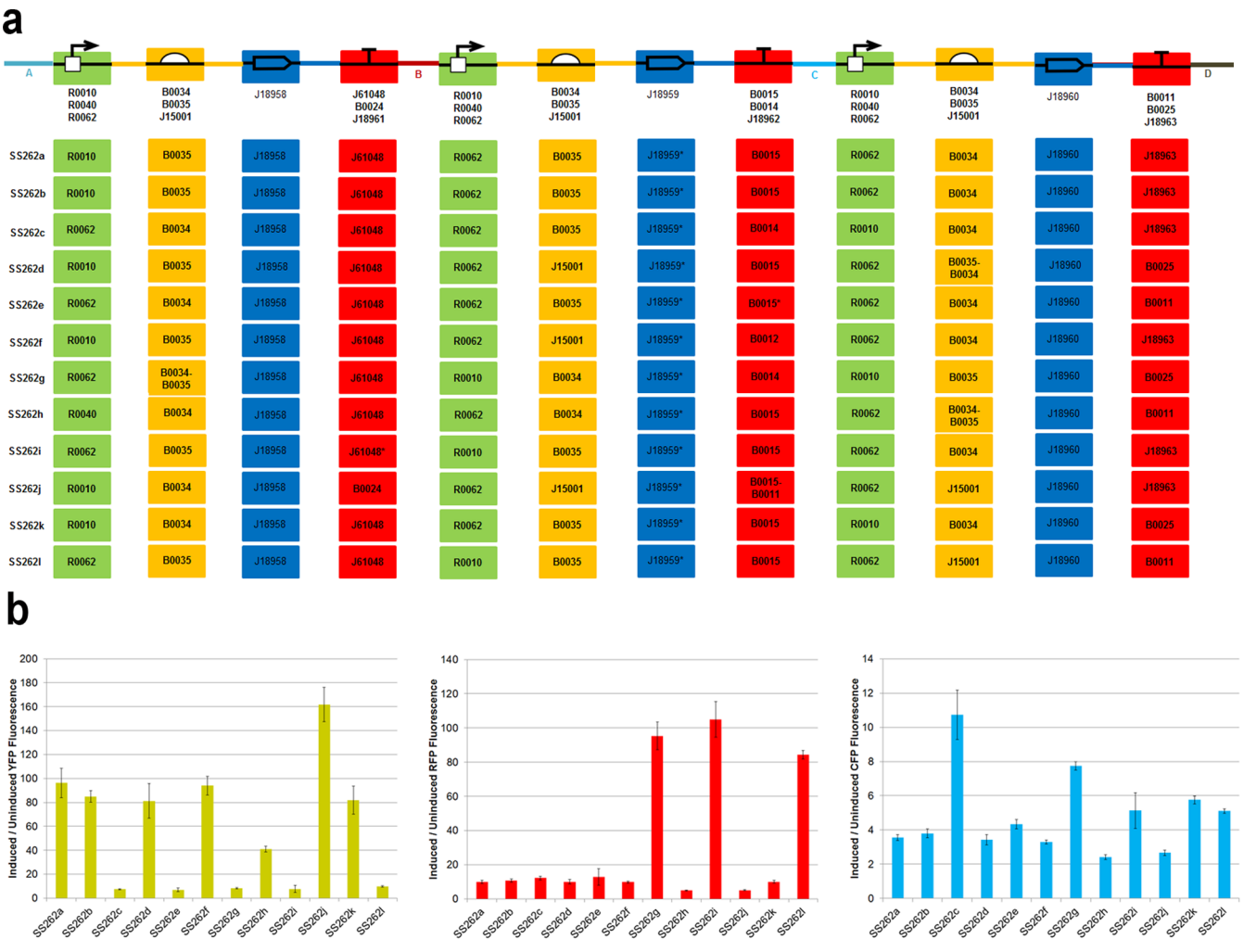


Figure 4. Sequencing results and characterization of CMY circuits with randomized promoters, RBSs, and terminators. (a) Sequencing results of 12 randomized CMY circuits with the promoters, RBSs, and terminators simultaneously randomized (strains SS262A–L) on the pSS3C1 vector. Eleven of the 12 SS262 circuits are unique (SS262A = B). The parts that were randomized are bolded and listed below each part symbol for a particular position in the circuit. Hybrid parts (two parts with homology that join together during the assembly reaction) are shown with two part names separated by a dash (*e.g.*, B0034-B0035), and starred parts indicate mutated sequences. See sequences of hybrid and mutated parts in Supplementary Table 4. (b) Characterization results of 12 randomized CMY circuits. The left, center, and right panels show the induced/uninduced fluorescence for eYFP, maxRFP, and eCFP, respectively, in each of the 12 circuits (SS262A–L). The height of each bar represents the mean of eight experiments ± SD. See Methods for detailed information on circuit characterization.

coding sequences constant. Thus, non-randomized eYFP, maxRFP, and eCFP one-gene circuits with "standard" promoters, RBSs, and terminators (exact parts described in Supplementary Table 5) were first rationally constructed to not have repeated parts when combined together into a three-gene circuit (by contrast randomized circuits may or may not have repeated parts). Non-randomized assembly thus requires adding four parts per assembly reaction (one promoter, one coding sequence, one terminator, and one vector), whereas randomized assembly requires adding additional parts of a particular part type, increasing the number of parts per assembly to greater than four. Randomizing promoter and terminator parts only requires adding additional part types to the same assembly reaction, but since the RBS is on the overlap sequence itself and cannot be assembled independently due to its short size, the RBS was randomized by adding different promoters and coding sequence parts to the assembly reaction, each with different RBSs attached on the overlap. In theory, only promoter and coding sequence fragments with the same RBS should assemble together. Because standard one-gene circuits were constructed, randomization success rates could be quantified relative to non-randomized assembly. One-gene randomization assembly reactions required much optimization, especially the simultaneous randomization of promoters, RBSs, and terminators due to the excessive amount of DNA added to the reaction when equimolar ratios are used (Supplementary Tables 5–8). After optimization, using between 10 and 40 ng total DNA for each part instead of equimolar ratios, the success rate of simultaneous promoter-RBS-terminator randomization for YFP, RFP, and CFP circuits on average improved from 30% to 67.6% (Supplementary Tables 5–8). On average, non-randomized assembly has a higher success rate than randomized assembly (89.0% *vs* 67.6%), not surprisingly, and the mean success rate decreases with the total number of parts used in the assembly reaction (Supplementary Figure 1, Supplementary Table 8).

After optimization of one-gene circuit randomization, two approaches were considered for three-gene randomized

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506–518

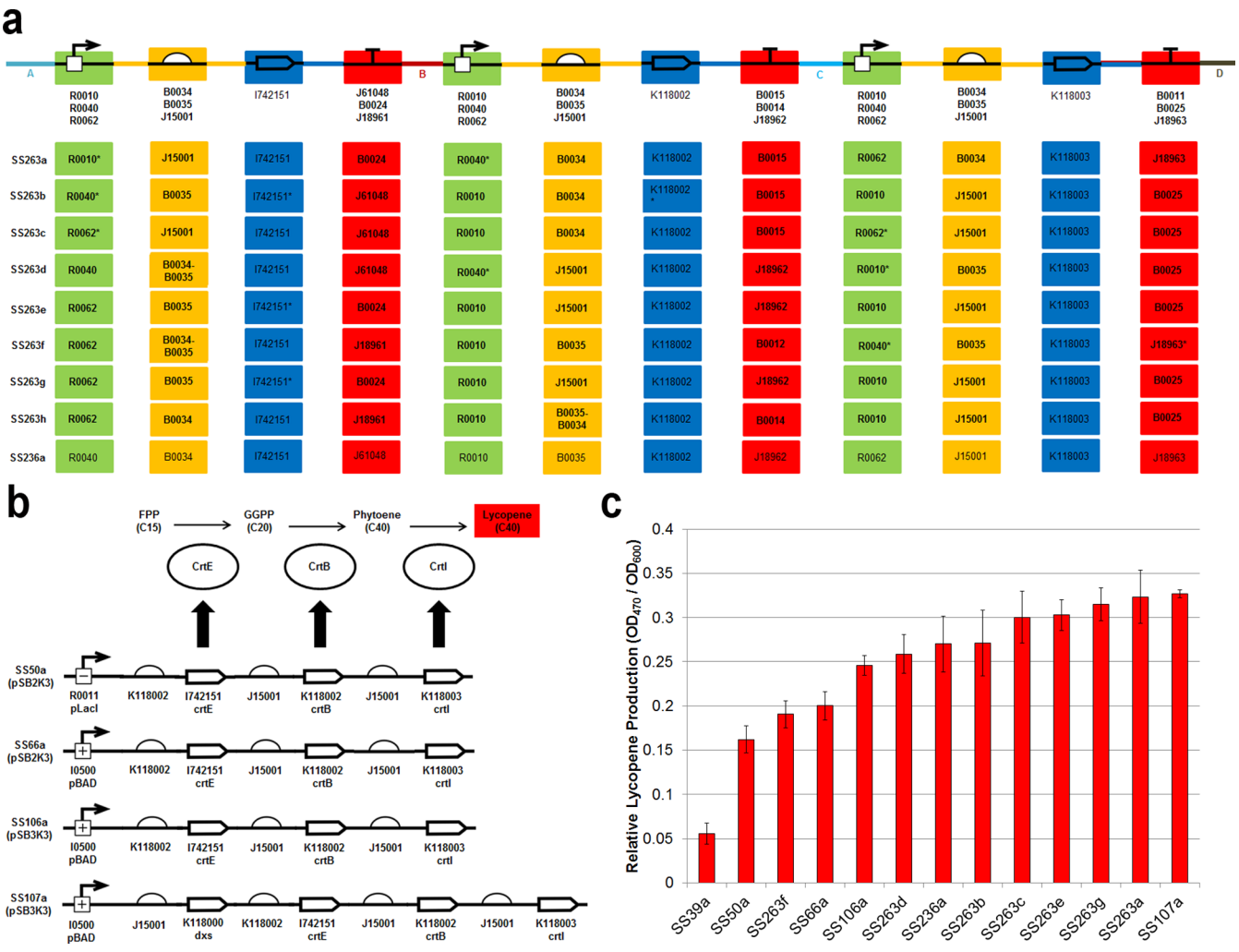


Figure 5. Randomized lycopene biosynthesis pathway sequencing results and characterization. (a) Sequencing results of eight randomized lycopene biosynthesis pathways (strains SS263A–H) with the promoters, RBSs, and terminators simultaneously randomized, and one non-randomized pathway (SS236A). Seven of the 8 SS263 circuits are unique (SS263E = G). The parts that were randomized are bolded and listed below each part symbol for a particular position. Hybrid parts (two parts with homology that join together during the assembly reaction) are shown with two part names separated by a dash (*e.g.*, B0034-B0035), and mutated parts are starred. See sequences of hybrid parts and mutated parts in Supplementary Table 4. (b) Rationally designed, lycopene biosynthesis pathways under polycistronic expression. The CrtE, CrtB, and CrtI gene products catalyze the reactions that metabolize FPP to lycopene. SS50A controls lycopene production from a pLacI promoter on an inducible-copy vector (pSB2K3, see Methods). SS66A controls lycopene production from a pBAD promoter on pSB2K3. SS106A is the same as SS66A, except that a medium-copy vector (pSB3K3) is used instead. SS107A is the same as SS106A, except that the *dxs* gene with a RBS is inserted upsteam of *crtE*, *crtB*, and *crtI* in the pathway. (c) Lycopene production in eight randomized pathways (SS263A–G; H was not used for characterization due to poor growth), one non-randomized pathway (SS236A), four non-randomized polycistronic pathways (SS50A, SS66A, SS106A, SS107A), and one negative control (SS39A) that lacks the *crtE*, *crtB*, and *crtI* genes. The height of each bar represents the mean of three experiments ± SD. See Methods for detailed information on inducer experiments and the lycopene extraction procedure.

assembly. The first approach involved an attempt to engineer an entire three-gene circuit from 10 individual parts in the same assembly reaction (PromoterA-CodingSequenceA-TerminatorA-PromoterB-CodingSequenceB-TerminatorB-PromoterC-CodingSequenceC-TerminatorC-Vector), but despite multiple attempts the assembly reaction failed, likely because there is a limit to the number of parts to join in a single reaction with these particular part sizes and overlap lengths. Instead, it was decided to first randomize one-gene circuits that express either YFP, CFP, or RFP in one assembly reaction, then selectively amplify these randomized circuits, and perform a second assembly reaction to combine the circuits together into randomized three-gene circuits.

Figure 2 shows a summary of the methodology to simultaneously randomize promoters, RBSs, and terminators to generate randomized three-gene circuits. The first step in three-gene randomization is to PCR-amplify all the parts of interest to randomize and independently assemble randomized YFP, RFP, and CFP one-gene circuits (Figure 2a). A detailed explanation of the parts required is given in the Supplementary Information, and Supplementary Table 9 provides a comprehensive list of all PCR products used in this study. Next, each YFP, RFP, and CFP assembly reaction is transformed into competent cells, and then transformants are plated on selective media and visualized under UV light to measure the qualitative success rate (Figure 2b). If success rates are sufficiently high, the third step is to create a pool of randomized one-gene circuits expressing either YFP, RFP, or CFP by selectively PCR-amplifying these circuits from the assembly reaction mixture itself (Figure 2c). Since each randomized one-gene circuit is

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506–518

Table 1. Parts and Vectors Used for Randomized BioBrick Assembly

part ID[a]	short ID[b]	part type[c]	description[d]
R0010	pLacZYA	promoter	pLacZYA promoter, repressed by LacI, inducible by IPTG
R0040	pTetR	promoter	pTetR promoter, repressed by TetR, inducible by aTc
R0062	pLuxR	promoter	pLuxR promoter, activated by LuxR when AHL is present
I0500	pBAD	promoter	pBAD promoter, activated by arabinose when AraC is present
B0034	RBS-A	RBS	RBS based on Elowitz repressilator
B0035	RBS-B	RBS	derivative of B0030 RBS based on Weiss thesis
J15001	RBS-C	RBS	synthetic RBS with SacI site
J18958	eYFP	coding sequence	enhanced Yellow Fluorescent Protein (eYFP)
J18959	maxRFP	coding sequence	maximum-expressing Red Fluorescent Protein (maxRFP)
J18960	eCFP	coding sequence	enhanced Cyan Fluorescent Protein (eCFP)
E0040	GFPmut3b	coding sequence	Green Fluorescent Protein with extra mutations from wildtype GFP (GFPmut3b)
I742151	crtE	coding sequence	*crtE* coding sequence (converts FPP to GGPP)
K118002	crtB	coding sequence	*crtB* coding sequence (converts GGPP to phytoene)
K118003	crtI	coding sequence	*crtI* coding sequence (converts phytoene to lycopene)
J61048	T1	transcriptional terminator	derived from the T1 terminator from *rnpB* gene of *E. coli* MG1655
B0024	T2	transcriptional terminator	reverse sequence of B0014
J18961	T3	transcriptional terminator	*bla* terminator from DNA2.0 plasmid
J18962	T4	transcriptional terminator	*rrnB1* terminator from DNA2.0 plasmid
B0015	T5	transcriptional terminator	consists of B0010 (T1 from *E. coli rrnB*) and B0012 (transcription terminator for the *E. coli* RNA polymerase)
B0014	T6	transcriptional terminator	consists of B0012 (transcription terminator for the *E. coli* RNA polymerase) and B0011
J18963	T7	transcriptional terminator	*rpn* terminator from DNA2.0 plasmid
B0025	T8	transcriptional terminator	reverse sequence of B0015
B0011	T9	transcriptional terminator	derived from *luxICDABEG* operon terminator of *Vibrio fischeri*
J18964	pGA3K4	vector	vector with pGA prefix/suffix, confers kanamycin resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18965	pSS3K1	vector	vector with pSS prefix A and suffix B, confers kanamycin resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18966	pSS3K2	vector	vector with pSS prefix B and suffix C, confers kanamycin resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18967	pSS3K3	vector	vector with pSS prefix C and suffix D, confers kanamycin resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18976	pSS3K4	vector	vector with pSS prefix A and suffix D, confers kanamycin resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18968	pSS3C1	vector	vector with pSS prefix A and suffix D, confers chloramphenicol resistance, p15a origin of replication, constitutively expresses LuxR and AraC
J18969	LinkerA	linker	coding sequence-terminator linker A
J18970	LinkerB	linker	coding sequence-terminator linker B
J18971	LinkerC	linker	coding sequence-terminator linker C
J18972	A	linker	prefix/suffix A
J18973	B	linker	prefix/suffix B
J18974	C	linker	prefix/suffix C
J18975	D	linker	prefix/suffix D

[a]The ID associated with parts in the Registry of Standard Biological Parts, where the sequence info can be found. [b]The common name used for each part. [c]Part type denotes whether the part is a promoter, ribosome binding site (RBS), coding sequence, transcriptional terminator, vector (plasmid backbone), or linker (overlap sequence). [d]Each part is described as synthetic or with respect to known genes in *E. coli* or *Vibrio fischeri*.

amplified from a vector with unique prefix and suffix sequences, this allows for these circuits to be combined in a second assembly reaction (Figure 2d) to create randomized three-gene circuits expressing YFP, RFP, and CFP. Note that letters designated on the prefix or suffix are designed to be complementary overlaps (*e.g.*, the "A" prefix on the YFP circuit will join with the "A" prefix on the pSS3C1 vector, the "B" suffix on the YFP circuit will join with the "B" prefix on the RFP circuit, *etc.*). Each randomized three-gene circuit thus has the possibility of different promoters, ribosome binding sites, and terminators at each position in the circuit. The final step is to transform the assembly reaction into competent cells and plate out the cells on selective media to visualize colonies under UV light (Figure 2e). Since CFP, RFP, and YFP produce Cyan, Magenta, and Yellow colors, respectively, under UV light, circuits expressing these three fluorescent proteins were termed CMY circuits. The wide variety of different colored colonies demonstrates the successful joining of different randomized one-gene circuits into multigene circuits, but not all

Table 2. Chi-Square Tests for Bias in Randomized Circuits[a]

	expected no.	observed nos.	χ^2 value	df	p-value	result
One-Gene Randomization (Figure 1)						
promoters	2.75	7, 3, 1, 0	10.455	3	0.0151	bias
coding sequences	2.75	2, 3, 5, 1	3.182	3	0.3644	no bias
terminators	2.75	1, 8, 2, 0	14.091	3	0.028	bias
Three-Gene Randomization						
CMY terminator randomization (Figure 3)						
terminators, YFP	4	6, 1, 5	3.5	2	0.1738	no bias
terminators, RFP	3.667	6, 1, 4	3.455	2	0.1777	no bias
terminators, CFP	4	4, 1, 7	4.5	2	0.1054	no bias
CMY promoter, RBS, terminator randomization (Figure 4)						
promoters, YFP	4	6, 1, 5	3.5	2	0.1738	no bias
RBSs, YFP	3.667	5, 6, 0	5.636	2	0.0597	no bias
terminators, YFP	4	11, 1, 0	18.5	2	0.0001	bias
promoters, RFP	4	3, 0, 9	10.5	2	0.0052	bias
RBSs, RFP	4	2, 7, 3	3.5	2	0.1738	no bias
terminators, RFP	3.333	8, 2, 0	10.4	2	0.0055	bias
promoters, CFP	4	3, 0, 9	10.5	2	0.0052	bias
RBSs, CFP	3.333	7, 1, 2	6.2	2	0.045	bias
terminators, CFP	4	3, 3, 6	1.5	2	0.4724	no bias
promoters, all	12	12, 1, 23	20.167	2	0.0001	bias
RBSs, all	11	14, 14, 5	4.909	2	0.0859	no bias
lycopene promoter, RBS, terminator randomization (Figure 5)						
promoters, YFP	2.667	1, 2, 5	3.25	2	0.1969	no bias
RBSs, YFP	2	1, 3, 2	1	2	0.6065	no bias
terminators, YFP	2.667	3, 3, 2	0.25	2	0.8825	no bias
promoters, RFP	2.667	6, 2, 0	7	2	0.0302	bias
RBSs, RFP	2.333	3, 1, 3	1.143	2	0.5647	no bias
terminators, RFP	2.333	3, 1, 3	1.143	2	0.5647	no bias
promoters, CFP	2.667	5, 1, 2	3.250	2	0.1969	no bias
RBSs, CFP	2.667	1, 2, 5	3.250	2	0.1969	no bias
terminators, CFP	2.667	0, 6, 2	7	2	0.0302	bias
promoters, all	8	12, 5, 7	3.25	2	0.1969	no bias
RBSs, all	7	5, 6, 10	2	2	0.3679	no bias
Three-Gene Promoter, RBS, and Terminator Randomization (Pooled) (Figures 4 and 5)						
promoters	20	24, 6, 30	15.6	2	0.0004	bias
RBSs	18	19, 20, 15	0.778	2	0.6778	no bias
terminators, YFP	10.667	20, 5, 7	12.438	2	0.002	bias
terminators, RFP	9.333	17, 4, 7	9.929	2	0.007	bias
terminators, CFP	10.667	7, 10, 15	3.063	2	0.2163	no bias

[a]For each position in a particular randomized assembly reaction, the expected no. (total number of sequenced circuits/parts mixed in the assembly reaction), observed nos. (frequency of a particular part occurring at a particular position), χ^2 value (Chi-square value), df (degrees of freedom), p-value, and result of the test are shown. Hybrid parts (Supplementary Table 4) were not included in the analysis since the Chi-square test relies on integers (*e.g.*, there cannot be a frequency of 0.5 for a hybrid part). The Chi-square test tests the null hypothesis that the observed frequency distribution of each randomized part at a particular position is consistent with the expected theoretical distribution. If the p-value is $\leq$0.05, this indicates that there is a significant bias for certain parts to be assembled at a higher frequency than other parts, whereas a p-value of >0.05 indicates the null hypothesis is true, an indication of successful randomization (no bias). The statistical results show that for each library of randomized circuits (Figures 1, 3, 4, and 5), in 10/28 (35.7%) cases there is a significant bias for certain parts at a particular position in the circuit, and in 18/28 (64.3%) the observed parts are uniformly distributed without bias. However, Chi-square calculations are only reliable when all expected values are 5 or greater. Unfortunately, most of the time, we did not achieve this result due to the small sample size of sequenced circuits. Therefore, we pooled all randomizations that involved three promoters, three RBSs, and nine terminators (three separate randomizations using three sets of different terminators were used) and find that all but one of these randomizations meet the requirement where all expected values are 5 or greater (one randomization had only 4). The results show that there is significant bias with regards to promoter randomization (likely due to the repeated operator sequences in R0040) and certain terminators combinations, but successful randomization with RBSs and other terminator combinations.

randomized circuits express three fluorescent proteins (see below).

Generation of Three-Gene Randomized Circuits. We first randomized only the terminators independently for YFP, CFP, and RFP circuits and then assembled these circuits together into randomized CMY circuits, as described in Figure 2. This example involved mixing one promoter, one coding sequence, three terminators, and one vector for each one-gene assembly reaction, then combining these one-gene circuits together to generate CMY circuits with randomized terminators. Sequencing results demonstrated that 9/12 CMY circuits (strain ID numbers SS225A–L) were distinct and that all nine terminators used for assembly were identified in at least one circuit (Figure 3a). We performed Chi-square tests on

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506–518

the 12 randomized circuits and determined in this particular case there was no statistical bias for particular terminators to assemble more frequently than others, using this small sample size (Table 2). Interestingly, there are assembly "errors" that occur when two terminators with homology recombine during the assembly reaction (*e.g.*, B0015-J18962 hybrid terminator in SS225L), but the generation of new, unexpected parts may be viewed as beneficial. Another fortuitous result was identifiying CMY circuits (SS225D and K) that happened to have the exact parts rationally designed for each of the "standard" one-gene circuits, now combined into a "standard" three-gene circuit (a 1/27 chance that a particular colony would have this combination of parts since three terminators were randomized at three different positions in the circuit). Expression data of these 12 circuits show that expression levels vary from about 3- to 235-fold above background levels (Figure 3b, arbitrary expression levels are shown in Supplementary Figure 6a). These 12 circuits have similar YFP, RFP, and CFP expression levels, with the exception of circuits with a particular terminator (J18961) that causes decreased steady-state expression of YFP. Circuits with this same terminator also have increased background (uninduced) RFP background, presumably due to read-through transcription from the upstream promoter, causing the induced/uninduced expression ratio to be low relative to other circuits without this terminator. As for the success rate, despite not alternating antibiotic selection (see Methods) going from one-gene to three-gene randomized assembly (Figure 2d), the success rate for the assembly reaction that generated these circuits was 13.5% (= 13 successful clones expressing three fluorescent proteins/96 total clones tested). Using optimization strategies (see below, and Methods and Supplementary Information) can increase the three-gene randomization success rate.

Next, we independently randomized the promoters, RBSs, and terminators simultaneously in CMY circuits exactly as described in Figure 2. This example involved mixing nine promoters (three different promoters each with three different RBSs), three coding sequences (the same coding sequence with three different RBSs), and three terminators for each one-gene assembly reaction, then combining these one-gene circuits together into randomized CMY circuits. Sequencing results show that 11/12 CMY circuits (strain ID numbers SS262A–L) are distinct (Figure 4a), a slight improvement over randomizing only terminators which produced 9/12 unique circuits. All three promoters, all three RBSs, and 7/9 terminators were identified among these circuits. In this small sample size, there is a statistical bias for parts at certain positions to assemble more successfully than others (*e.g.*, R0040 is identified less frequently than R0010 or R0062) (Table 2). On average, there is a statistical bias for promoter and terminator randomization, whereas there is no statistical bias for RBS randomization (Table 2). Also, as with the case of terminators, since some RBSs have homology (*e.g.*, B0034 and B0035), hybrid RBSs can be created as an artifact of the assembly process (noted as either B0034-B0035 or B0035-B0034 with exact sequences listed in Supplementary Table 4). Other hybrid parts were created as well, including another hybrid terminator that was created between B0015 and B0011 due to shared homology. Figure 4b shows the relative expression levels in these circuits that vary from 2- to 160-fold above background levels (arbitrary expression levels are shown in Supplementary Figure 6b). Unlike the CMY circuits with randomized terminators, these 12 circuits show wider variation in expression levels between individual circuits, likely because randomizing different promoters and RBSs has a greater effect on gene expression relative to transcriptional terminators. From the expression results, there is useful information that can be extracted. For instance, there is a clear pattern between the expression levels in circuits with R0010 and R0062 promoters. R0010 have higher induced/uninduced expression on average compared to R0062. This result can be explained because although R0062 is the strongest promoter, it is also the most leaky promoter, so it has a higher uninduced expression level, causing lowered induced/uninduced expression. This pattern of R0010 having higher induced/uninduced expression is more pronounced with YFP and RFP expression and less clear with CFP expression, probably due to the fact that CFP expression has higher background levels in general.

The success rates of the reactions that generated these particular randomized circuits was initially only 7.3% (= 7/96) using alternating antibiotic selection and random screening of fluorescing colonies, as opposed to 20.8% (= 10/48) with non-randomized three-gene assembly. After specific screening methods were developed further (described in Methods), the success rate of randomized three-gene assembly was improved to 34.4% (= 11/32). Thus, the ability to identify potentially hundreds of positive clones per assembly reaction negates the fact that three-gene randomization assembly reactions are low in general. Also, by sequencing various clones that generated particular colors, it was revealed that two-gene circuits are commonly generated due to the use of parts with significant homology that can hybridize during the assembly reaction (parts with homology hybridize to each other instead of their designated linker sequences). This phenomenon is especially a problem when randomizing the same three promoters and three RBSs in YFP, RFP, and CFP circuits. Taken together, these results suggest that in order to maximize success rates for three-gene randomization, (1) parts with significant homology should not be used, (2) alternating use of antibiotic selection should be used at each assembly step, and (3) efficient screening methods must be developed to identify positive clones.

Generation of Randomized Lycopene Biosynthesis Pathways. In order to adapt this randomization methodology for applied purposes, we then independently randomized the same promoters, RBSs, and terminators in the lycopene biosynthesis pathway. The difference between parts used for this randomization *versus* parts used in Figure 2 is that instead of YFP, RFP, and CFP coding sequences, the *crtE*, *crtB*, and *crtI* coding sequences were used in their place. After selecting 12 colonies that produced a red pigment on selective media, sequencing results demonstrate that 7/8 randomized pathways (strain ID numbers SS263A–H) were distinct (Figure 5a) and four pathways had deletions (possibly due to the fact that the metabolic load may have been too high in these pathways, reducing growth rate, and allowing for mutants to outcompete cells with functional pathways). Interestingly, the part spectrum from these randomized pathways is very different than for the randomized circuits shown in Figure 4a (see Supplementary Figure 2 for a direct comparison). Interestingly, on average randomized lycopene biosynthesis pathways are about twice as likely to have mutations relative to randomized CMY circuits (Supplementary Table 4). There are also mutations found in promoters that were not seen in circuits expressing fluorescent proteins. These sequencing results highlight the difference between expressing fluorescent proteins and enzymes used in

metabolic pathways; the latter is actually doing work in the cell and more likely to disrupt cellular physiology. Apparently the strong promoters and RBSs used for randomization may have overwhelmed cells and caused stunted growth, resulting in the recovery of successful lycopene producers, albeit with mutations. Although the choice of parts also had an effect on the success rate of this assembly reaction (which was not quantified, but certainly less than 5%), all of the randomized pathways were generated with a single assembly reaction, and the majority of identified red colonies did produce a satisfactory red color when grown in test tubes and spun down. Taken together, these results suggest that the parts used for lycopene biosynthesis pathway randomization may not have been ideal and that use of other parts may have resulted in selection of clones with less mutated parts and higher assembly success rates.

In order to measure lycopene production in these randomized pathways relative to other rationally designed monocistronic- and polycistronic-expressing pathways, the inducer concentrations were first optimized to maximize lycopene production. The different pathways used included a non-randomized, monocistronic lycopene pathway that was rationally designed, eight randomized monocistronic pathways, three polycistronic-expressing pathways, and a polycistronic-expressing pathway with the addition of *dxs* to increase metabolic flux (Figure 5b, Supplementary Figures 3 and 4). Figure 5c shows the lycopene production for each uniquely tuned pathway. The pathway that produced the highest amount of lycopene on average (SS107A) is a rationally designed polycistronic pathway that in addition to *crtE*, *crtB*, and *crtI* produces *dxs*, a gene known to increase metabolic flux through this pathway.[14,21−23] However, this pathway should be treated as a special case since it is the only pathway tested with the *dxs* gene and also because the lycopene production in this pathway cannot be well-controlled (*i.e.*, the maximum level of lycopene is achieved without induction; with induction the cells grow poorly). The highest-expressing randomized pathway (SS263A) after proper tuning produces about the same amount of lycopene as SS107A and is not statistically different (unpaired two-tailed t test, p-value = 0.9305). The SS263A pathway produces on average about 30% more lycopene than the highest polycistronic-expressing pathway (SS106A) and about 20% more lycopene than the "standard" non-randomized SS236A pathway. Further improvements of lycopene production may be achieved by using a different set of parts to change *crtE*, *crtB*, and *crtI* expression or with the addition of other genes known to increase metabolic flux through the lycopene biosynthesis pathway such as *dxs* or *idi*, or the combination of both approaches.

General Conclusions and Future Directions. Overall, the Randomized BioBrick Assembly methodology can be used to simultaneously randomize different parts to produce a library of circuits and pathways with varying expression levels. Once individual parts are ready for assembly, three-gene circuits and pathways can be generated in only 3 days (Methods). On average there is a statistical bias for certain parts to be randomized more efficiently than others (Table 2, pooled Chi-square results), but 90% (18/20) of circuits and pathways sequenced with promoters, RBSs, and terminators simultaneously randomized are unique (Figures 4 and 5). Much optimization was required to get this assembly to work with acceptable success rates (Methods, Supplementary Tables 5−8, Supplementary Figure 5), and despite optimization, improvements on this methodology can be suggested for applied use. For generating randomized pathways to properly balance enzyme levels, a wide variety of parts with varying strengths should be used for randomization due to the metabolic load associated with expressing enzymes. Another suggested improvement is to only use unique parts that do not have significant homology when randomizing different parts simultaneously in order to maximize assembly success rates. In general, we expect that this methodology can be easily adapted to other parts and other organisms besides *E. coli* since it does not have complicated rules (*i.e.*, attach standardized overlaps shown in Supplementary Table 1 to parts of interest and assemble). We hope that this assembly method will help to develop new complex circuits and pathways in various organisms, be used in combination with other engineering methods,[22] and be amenable to robotic automation.

METHODS

Parts, Vectors, and Plasmid Engineering. All standard biological parts (BioBricks) were obtained from the Registry of Standard Biological Parts (partsregistry.org) except for select parts obtained from DNA2.0 (J18961, J18962, and J18963 terminators), the Klavins Lab (maxRFP coding sequence, equivalent to the E1010 Monomeric Red Fluorescent Protein, except that the first few codons were re-engineered to increase expression levels), or the Elowitz Lab (eYFP and eCFP coding sequences). All parts and vectors used in this study are described in Table 1, and all strains produced using the Randomized BioBrick Assembly method are described in Supplementary Table 3. All vectors except the inducible-copy vector (pSB2K3, described on partsregistry.org) are derived from the pSB3K3 plasmid (described on partsregistry.org), a medium copy number plasmid (20−30 plasmids/cell) with a p15A pMR101-derived replication origin and kanamycin resistance gene.[24] The pSB3K3 plasmid has a repeated NotI restriction site in the prefix and suffix, causing vector self-ligation during Gibson assembly reactions, as determined by the 2011 University of Washington iGEM team. The team re-engineered this vector without NotI sites and thereby greatly improved 2-way Gibson assembly success rates with their new vector named pGA3K3. In this study, the pGA3K3 vector was re-engineered to express *luxR* and *araC* from a constitutive promoter on I0500 (modified without the pBAD promoter) upstream of the prefix in the reverse direction to make pGA3K4. After it was determined that pGA3K4 4-way Gibson assembly reaction failures were caused by a different type of self-ligation (main text and Supplementary Table 4), the existing prefix and suffix were removed entirely and replaced with new prefixes and suffixes that lacked restriction sites or inverted repeats. These pSS3K1, pSS3K2, pSS3K3, and pSS3K4 vectors used for one-gene randomization are indistinguishable from pGA3K4 except for having different prefixes and suffixes that were designed to have a T_m of 57−59 °C (Table 1, Supplementary Table 1). The pSS3C1 vector used for three-gene randomization was also engineered from pGA3K4, except that the kanamycin expression cassette was replaced with a chloramphenicol expression cassette from pGA1C3 (another vector constructed by the 2011 UW iGEM team).

PCR and Randomized Assembly Reactions. PCR was performed using Phusion High-Fidelity PCR Master Mix according to the manufacturer's instructions, with 3% DMSO, and the primer annealing step was performed at 57 °C for 30 s. Primers to PCR-amplify promoters, coding sequences, tran-

scriptional terminators, and vectors were obtained from IDT and designed to have a T_m of 57−59 °C on the 3′ end specific to the template DNA. The standardized overlap sequences on the 5′ end are shown in Supplementary Table 1. PCR products were purified using the Qiagen PCR Purification Kit (without gel extraction) and assembled using Gibson assembly,[6] as described, but with some exceptions. For one-gene randomized assembly, instead of using equimolar ratios of products for assembly, various optimization strategies were used. After optimization, it was determined that using the same total nanograms of DNA (10−40 ng/part) for each part in the assembly reaction gave higher success rates than using various insert:vector molar ratios (Supplementary Tables 5−8). A possible explanation for these results is given in Supplementary Information. For PCR reactions requiring >5 μL total DNA, assembly reaction volumes were doubled (*e.g.*, instead of 15 μL Gibson assembly master mix + 5 μL total DNA and water, 30 μL Gibson assembly master mix + 10 μL total DNA and water was used). For three-gene randomized assembly, each of the three one-gene randomization assembly reactions were diluted in water 1:100 (1 μL of assembly reaction into 99 μL of water), and then 0.5 μL of the diluted reaction was added as template DNA to a new 25 μL PCR reaction. Primers to selectively amplify the pool of randomized one-gene circuits are listed and described in Supplementary Table 2. After these PCR products (normally having a "smear" of different size products due to the different parts used) are purified using the Qiagen PCR Purification kit, another Gibson assembly reaction is performed using equimolar ratios (using the largest possible randomized circuit size in bp for calculations) to assemble randomized three-gene circuits. Important to the success rate of three-gene randomized assembly is the ability to independently assemble YFP, CFP, and RFP circuits on different vectors with standardized prefix and suffix sequences so that the randomized pools of circuits can be selectively amplified from the first assembly reaction and combined into three-gene circuits in the second assembly reaction. Alternating antibiotic selection is used at each step of this process, going from individual parts PCR-amplified from ampicillin resistance-conferring vectors to randomized one-gene circuits on vectors conferring kanamycin resistance, to randomized three-gene circuits on chloramphenicol-conferring vectors.

Transformations. Assembly reactions were transformed into Stellar (Clontech) or NEB Turbo chemically competent *E. coli* cells that overexpress LacI ($lacI^q$). Typically 2 μL of the completed assembly reaction was incubated with 30 μL of competent cells in an Eppendorf tube for 30 min on an ice block, heat shocked at 42 °C for 30 s, put back on the ice block for 1 min, and then grown with 200 μL of SOC media at 37 °C shaking at 250 rpm for 1 h. After incubation, 200 μL of the transformed cell culture was plated on LB agar plates with appropriate antibiotics (50 μg/mL kanamycin or 34 μg/mL chloramphenicol) and inducers (100 μM IPTG and 100 nM AHL), then incubated overnight at 37 °C, or allowed to incubate an extra day at room temperature to allow for colonies to develop stronger colors for visualization. NEB Turbo cells produce colonies with stronger colors compared to Stellar cells.

Functional Screening of Positive Clones and Assembly Success Rates. In general, at least three colonies were screened for the correct insert using functional screening or colony PCR with primers specific to the vector and/or insert. Functional screening to identify positive clones after Gibson assembly and transformation, or to determine qualitative assembly success rates, was determined by visualizing colonies on a blue light transilluminator (Clare Chemical) with a Fotodyne FOTO/Analyst Express System or on a UV transilluminator (Fotodyne, 312 nm wavelength, low intensity) with a Fotodyne Apprentice System equipped with a color digital camera (after removal of the UV filter). Screening for positive clones in one-gene randomized circuits only involved visualization of fluorescing colonies, whereas a number of approaches were used to screen for positive clones in three-gene randomized circuits due to variations in colony color under UV light (cells were shielded against UV light to prevent UV-induced damage). Under UV light, colonies expressing all three fluorescent proteins range from dark orange to dark purple to bright green. The approach that maximized success rate was to identify colonies that were orangeish under normal light conditions (both eYFP- and eCFP-expressing colonies appear yellow and maxRFP-expressing colonies appear red, thus producing an orangeish color when these colors are expressed in combination). Lycopene-producing transformant colonies produced visible a red pigment and were visualized in normal light conditions. Qualitative assembly success rates for one-gene randomization were measured by counting the total number of fluorescing colonies and dividing this number by the total number of colonies. Quantitative assembly success rates for three-gene randomization was determined by the counting the number of clones expressing eYFP, maxRFP, and eCFP significantly above background levels (see next section for details) and dividing this number by the total number of clones tested, or by DNA sequencing.

Circuit and Pathway Characterization. The R0010 (pLacZYA with CAP binding site, regulated by LacI), R0040 (pTetR), and R0062 (pLuxR) promoters can be induced by isopropyl-β-D-thiogalactopyranoside (IPTG), anhydrotetracycline (aTc), and 3-oxohexanoyl-homoserine lactone (AHL), respectively (the NEB Turbo cells do not express TetR and thus do not require aTc for induction). Positive clones were grown in LB broth plus 50 μg/mL kanamycin or 34 μg/mL chloramphenicol, 100 μM IPTG, and 100 nM AHL at 37 °C shaking at 250 rpm for 24 h. Steady-state fluorescence and optical density (OD_{600}) were quantified by thoroughly mixing all cultures and pipetting 200 μL into a black, clear-bottom 96-well microplate (Costar) and taking measurements in a Tecan M200 Pro monochromator plate reader using the appropriate excitation/emission wavelengths (eCFP: 439 ex/476 em, GFPmut3b: 485 ex/516 em, eYFP: 514 ex/548 em, maxRFP: 584 ex/620 em). Clones that expressed eYFP, maxRFP, and eCFP significantly above background levels were grown in a test tube with 8 mL of LB and appropriate antibiotics at 37 °C shaking at 250 rpm overnight and saved for long-term storage at −80 °C in 15% glycerol, and then plasmids were extracted using the Qiagen Miniprep Kit and sequenced using primers specific to the vector and/or insert. After plasmid sequencing, correctly constructed plasmids were transformed into MG1655 $Z1^{3,25-27}$ (SS39A), which constitutively overexpresses LacI and TetR from the chromosome. Eight transformant colonies of each randomized CMY circuit were grown in LB and appropriate antibiotics at 37 °C shaking at 250 rpm overnight and then saved for long-term storage at −80 °C in 15% glycerol. Circuits were characterized by first streaking out freezer stocks on LB agar plates with appropriate antibiotics (50 μg/mL kanamycin or 34 μg/mL chloramphenicol) and incubation at 37 °C overnight, and then colonies were grown for 24 h in 1.5 mL of LB + 50 μg/mL kanamycin or 34 μg/mL

dx.doi.org/10.1021/sb4000542 | *ACS Synth. Biol.* 2013, 2, 506−518

chloramphenicol at 37 °C shaking at 250 rpm in an Eppendorf deep 96-well plate sealed with a Thermo Scientific gas-permeable membrane for maximum oxygen diffusion. The cultures were then diluted 1:1000 into 1.5 mL of LB + 50 μg/mL kanamycin or 34 μg/mL chloramphenicol, with and without inducers (100 μM IPTG, 100 nM AHL, and 1 μg/mL aTc) and grown at 37 °C shaking at 250 rpm for 24 h, and then steady-state fluorescence and OD_{600} was measured (Supplementary Figure 6). To calculate the induced/uninduced ratio (a metric to measure circuit on−off behavior), induced and uninduced fluorescence readings are first divided by OD_{600} to normalize for cell density, and then the induced fluorescence/OD_{600} value is divided by uninduced fluorescence/OD_{600} value for each replicate (Figures 3b and 4b).

CMY Color Visualization and Inducer Experiments. Media was prepared in an Eppendorf deep 96-well plate with 1.5 mL of LB + 50 μg/mL kanamycin or 34 μg/mL chloramphenicol, 100 μM IPTG, 100 nM AHL, and 1 μg/mL aTc (for full induction), and then a serial dilution was performed to dilute this media 2-fold successively with each column on the plate (see Supplementary Figure 3). The top column has full induction and each column is diluted 1:2 successively until a 1:1024 dilution is achieved, and the bottom column has no inducers. Uninduced cultures were diluted 1:1000 into the media and grown at 37 °C shaking at 250 rpm for 24 h, and then steady-state fluorescence and OD_{600} were measured. For color visualization, the cultures were centrifuged at 2500 rpm for 10 min at 4 °C, the supernatant was removed, then the cells were washed with 500 μL of water, centrifuged again at 2500 rpm for 10 min at 4 °C, the supernatant was again removed, and then the cells were resuspended in 100 μL of water and transferred into a black, clear-bottom 96-well microplate (Costar). The plate was incubated for an additional 24 h at room temperature and then visualized with a white light (lycopene pathways) or UV transilluminator (CMY circuits).

Lycopene Extractions. Strains with lycopene biosynthesis pathways and a negative control that does not produce lycopene (SS39A) were grown without inducers in test tubes in 5 mL of LB + 50 μg/mL kanamycin or 34 μg/mL chloramphenicol at 37 °C shaking at 250 rpm for 24 h. Overnight cultures were diluted 1:1000 into test tubes with 10 mL of LB + 50 μg/mL kanamycin or 34 μg/mL chloramphenicol with IPTG, AHL and aTc inducer concentrations that maximize lycopene production (Supplementary Table 10) and grown at 30 °C shaking at 250 rpm for 48 h in the dark (this temperature and incubation time combination was visually determined to produce the most lycopene when comparing test strains at 30 °C *vs* 37 °C and 24 h *vs* 48 h, using each of the four combinations). Each replicate culture was first transferred (200 μL) into a black, clear bottom 96-well microplate, and optical density was measured at OD_{600} to normalize lycopene production with cell density. The remaining culture was centrifuged at 3000*g* in a Sorvall 23R Legend centrifuge for 10 min at 4 °C. The supernatant was removed, and the cell pellet was resuspended in 400 μL of acetone and then transferred to a black Eppendorf tube (lycopene is light-sensitive). Tubes were vortexed continuously on a microtube holder on a high setting for 15 min and then heated on a heat block at 55 °C for 15 min. The tubes were vortexed again for 5 s and then centrifuged at 17,000*g* in a tabletop centrifuge for 3 min. The supernatant (100 μL) was mixed with 100 μL of water, and then the entire volume was transferred to a microplate (the supernatant was diluted in water since acetone can degrade the plastic in the microplate). Relative lycopene production was measured by absorbance at OD_{470}/OD_{600} for each replicate.

Strain Availability. All strains and sequence information will be made available via AddGene (addgene.org) using the strain IDs in Supplementary Table 3. The sequences for all basic parts are available on the Registry of Standard Biological Parts (partsregistry.org).

■ ASSOCIATED CONTENT

Supporting Information

This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sleight@uw.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank BEACON: An NSF Center for the Study of Evolution in Action for funding this research, the Registry of Standard Biological Parts, DNA2.0, the Klavins and Elowitz laboratories for parts, the SBOL Visual V0.0.0 team for part symbols, the 2011 University of Washington iGEM team, Sauro lab (Wilbert Copeland, Bryan Bartley, Kyung Kim, and Michal Galdzicki), and Klavins lab (Rob Egbert) members for useful materials and discussions.

■ REFERENCES

(1) Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S. (2002) Combinatorial synthesis of genetic networks. *Science 296*, 1466−1470.

(2) Knight, T. (2003) Idempotent Vector Design for Standard Assembly of Biobricks, *DSpace* http://hdl. handle. net/1721.1/21168.

(3) Cox, R. S., Surette, M. G., and Elowitz, M. B. (2007) Programming gene expression with combinatorial promoters. *Mol. Syst. Biol. 3*, 145.

(4) Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol. 26*, 787−793.

(5) Shetty, R. P., Endy, D., and Knight, T. F., Jr. (2008) Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng. 2*, 5.

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

(7) Peisajovich, S. G., Garbarino, J. E., Wei, P., and Lim, W. (2010) Rapid diversification of cell signaling phenotypes by modular domain recombination. *Science 328*, 368−372.

(8) Sleight, S. C., Bartley, B. A., Lieviant, J. A., and Sauro, H. M. (2010) In-Fusion BioBrick assembly and re-engineering. *Nucleic Acids Res. 38*, 2624−2636.

(9) Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., and Peccoud, J. (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS One 6*, e16765.

(10) Litcofsky, K. D., Afeyan, R. B., Krom, R. J., Khalil, A. S., and Collins, J. J. (2012) Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nat. Methods 9*, 1077−1080.

(11) Sarrion-Perdigones, A., Falconi, E. E., Zandalinas, S. I., Juárez, P., Fernández-del-Carmen, A., Granell, A., and Orzaez, D. (2011) GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PloS One 6*, e21622.

(12) Yokobayashi, Y., Weiss, R., and Arnold, F. H. (2002) Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16587−91.
(13) Umeno, D., Tobias, A. V., and Arnold, F. H. (2005) Diversifying carotenoid biosynthetic pathways by directed evolution. *Microbiol. Mol. Biol. Rev. 69*, 51−78.
(14) Tao, L., Jackson, R. E., and Cheng, Q. (2005) Directed evolution of copy number of a broad host range plasmid for metabolic engineering. *Metab. Eng. 7*, 10−7.
(15) Haseltine, E. L., and Arnold, F. H. (2007) Synthetic gene circuits: design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct. 36*, 1−19.
(16) Esvelt, K. M., Carlson, J. C., and Liu, D. R. (2011) A system for the continuous directed evolution of biomolecules. *Nature 472*, 499−503.
(17) 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.
(18) Hillson, N. J., Rosengarten, R. D., and Keasling, J. D. (2012) j5 DNA assembly design automation software. *ACS Synth. Biol. 1*, 14−21.
(19) Zhu, B., Cai, G., Hall, E. O., and Freeman, G. J. (2007) In-fusion assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *BioTechniques 43*, 354−359.
(20) Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods 4*, 251−256.
(21) Kang, M.-J., Yoon, S.-H., Lee, Y.-M., Lee, S.-H., Kim, J.-E., Jung, K.-H., Shin, Y.-C., and Kim, S.-W. (2005) Enhancement of lycopene production in Escherichia coli by optimization of the lycopene synthetic pathway. *J. Microbiol. Biotechnol. 15*, 880−886.
(22) 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.
(23) Yoon, S.-H., Park, H.-M., Kim, J.-E., Lee, S.-H., Choi, M.-S., Kim, J.-Y., Oh, D.-K., Keasling, J. D., and Kim, S.-W. (2007) Increased beta-carotene production in recombinant Escherichia coli harboring an engineered isoprenoid precursor pathway with mevalonate addition. *Biotechnol. Prog. 23*, 599−605.
(24) Lutz, R., and Bujard, H. (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res. 25*, 1203−1210.
(25) Sleight, S. C., Bartley, B. A., Lieviant, J. A., and Sauro, H. M. (2010) Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng. 4*, 12.
(26) Sleight, S. C., and Sauro, H. M. (2012) Design and construction of a prototype CMY (Cyan-Magenta-Yellow) genetic circuit as a mutational readout device to measure evolutionary stability dynamics and determine design principles for robust synthetic systems. *Artificial Life 13*, 481−488.
(27) Yang, S., Sleight, S. C., and Sauro, H. M. (2013) Rationally designed bidirectional promoter improves the evolutionary stability of synthetic genetic circuits. *Nucleic Acids Res. 41*, e33.

■ NOTE ADDED AFTER ASAP PUBLICATION

This article was published ASAP on July 10, 2013. Figure 5 has been replaced. Throughout the paper, the relative lycopene production is now reported, not absolute values. The correct version was published on September 9, 2013.