# Marionette: *E. coli* containing 12 highly-optimized small molecule sensors

Adam J. Meyer[1], Thomas H. Segall-Shapiro[1], and Christopher A. Voigt[1,*]

[1] Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Correspondence and request for materials should be addressed to C.A.V. (cavoigt@gmail.com)

Key words: *Synthetic biology, directed evolution, biosensor, gene expression*

**Cellular processes are carried out by many interacting genes and their study and optimization requires multiple levers by which they can be independently controlled. The most common method is via a genetically-encoded sensor that responds to a small molecule (an "inducible system"). However, these sensors are often suboptimal, exhibiting high background expression and low dynamic range. Further, using multiple sensors in one cell is limited by cross-talk and the taxing of cellular resources. Here, we have developed a directed evolution strategy to simultaneously select for less background, high dynamic range, increased sensitivity, and low crosstalk. Libraries of the regulatory protein and output promoter are built based on random and rationally-guided mutations. This is applied to generate a set of 12 high-performance sensors, which exhibit >100-fold induction with low background and cross-reactivity. These are combined to build a single "sensor array" and inserted into the genomes of *E. coli* MG1655 (wild-type), DH10B (cloning), and BL21 (protein expression). These "Marionette" strains allow for the independent control of gene expression using 2,4-diacetylphophloroglucinol (DAPG), cuminic acid (Cuma), 3-oxohexanoyl-homoserine lactone (OC6), vanillic acid (Van), isopropyl β-D-1-thiogalactopyranoside (IPTG), anhydrotetracycline (aTc), L-arabinose (Ara), choline chloride (Cho), naringenin (Nar), 3,4-dihydroxybenzoic acid (DHBA), sodium salicylate (Sal), and 3-hydroxytetradecanoyl-homoserine lactone (OHC14).**

Advances in biology are often tied to new methods that use external stimuli to control the levels of gene expression[1-3]. Pioneered in the early 1980s, so-called inducible systems were developed that allow genes to be turned on by adding a small molecule inducer to the growth media[4]. These consist of a protein transcription factor (*e.g.*, LacI) whose binding to a DNA operator in a promoter is controlled by the inducer (*e.g.*, IPTG). Initially co-opted from natural regulatory networks, over the years many versions were designed to improve performance. In the 1990s, additional systems were developed that responded to other inducers, notably arabinose and aTc, which became common tools in the field. In 1997, Lutz and Bujard published a seminal paper that combined three (IPTG, arabinose, aTc) that could be easily interchanged on a two-plasmid system[5]. Its organizational simplicity, compatibility, and quantified response functions were revolutionary. Beyond providing a new tool to biologists to control multiple genes with independent "strings," it facilitated researchers with quantitative backgrounds to enter biology[6-7]. Armed with the new ability to control two genes with precision, physicists and engineers built the first synthetic genetic circuits, performed single molecule experiments inside cells, deconstructed the origins of noise in gene expression, determined how enzyme balancing impacts metabolic flux, elucidated rules underlying the assembly of molecular machines, and built synthetic symbiotic microbial communities, just to highlight a few[8-19].

Sensor performance is quantified by its response function; in other words, how the concentration of inducer changes the activity of the output promoter (Figure 1a). Often, this promoter retains a residual activity in the absence of inducer ("leakiness"). This hampers the ability to explore low expression levels or keep a gene in the off state, particularly needed for toxic proteins[20-21]. Another important parameter is the dynamic range, defined as the ratio of promoter activity in the on and off states. When this is large, both high and low expression can be explored as well as many intermediate states. The sensitivity is the concentration of inducer that turns a sensor on (defined as 50% activation). A lower sensitivity reduces the amount of a chemical that must be added to the media. Further, when multiple sensors are combined into one cell, they can interfere with each other's response functions (Figure 1b). Some small molecules bind non-cognate regulators and this can lead to off-target activation (cross reactivity) or competitive inhibition with the cognate small molecule (antagonism)[22-27]. Finally, each sensor requires cellular resources (*e.g.*, ribosomes) to function and the activation of one sensor can influence another indirectly due to resource competition[28]. Each sensor also requires ~1-2 kb of DNA and this becomes increasingly difficult to carry on plasmids. These challenges limit the number of sensors that can be put in a single cell and the maximum reported to date is four[29].

Sensors can be improved using biophysical models, rational engineering and directed evolution[5, 23, 26, 30-36]. Improving performance requires that screens or selections be performed in the presence and absence of inducer[37]. Such dual selections have been accomplished by sorting cells based low and high fluorescence, utilizing proteins that can be both toxic and selective (*e.g.,* TetA or HSV-TK), or by deploying separate positive and negative selections[26, 35, 37-44]. This has been applied to improving the response functions and eliminating cross-reactivity between pairs of regulators[23, 26, 35, 45].

Here, we have developed a selection methodology that allows us to intervene at multiple steps in order to simultaneously select for improved response functions and decreased crosstalk (Figure 1c). First, we built 12 sensors that respond to different small molecules and made directed changes in order to improve function. These sensors were then subjected to multiple rounds of negative-positive selection. Negative selection is facilitated by a promiscuous A294G mutant of the phenylalanine aminoacyl tRNA-synthetase (PheS)[46]. In the absence of inducer, leaky transcription of PheS leads to the charging of phenylalanyl-tRNA with the non-canonical amino acid 4-Chloro-DL-phenylalanine (Cl-Phe). Adding more Cl-Phe increases the stringency of the selection by making PheS transcription more toxic[43, 47]. Positive selection is facilitated by a thermostable DNA polymerase (DNAP), either KOD DNAP from the archaea *Thermococcus kodakarensis*[48] (for stringent replication) or the engineered PK6 DNAP[49] (to introduce mutations). In the presence of inducer, the level of DNAP can be quantified by PCR amplification of the

sensor library after the cells are emulsified with primers[50-51]. Multiple properties of the response function can be simultaneously improved during one cycle of negative-positive selection and stringency altered by changing the concentrations of Cl-Phe and inducer. Further, cross reactions can be selected against by adding the inducer or regulator from a problematic system. The dual selection is applied over multiple rounds to create a highly-optimized set of sensors. From these, we identify 12 that can be used together in a single strain, and these are combined and integrated into the genomes of *E. coli* MG1655, DH10B, and BL21 in order to create the Marionette family of strains. Genome integration increases stability and reduces the cellular resources required to maintain regulator expression. It also simplifies the use of the inducible systems, where only the output promoters need to be incorporated into a design (*e.g.*, for the expression of multiple proteins from a plasmid).

An initial set of 17 putative sensors were designed. Each sensor consists of a weak constitutive promoter ($P_{LacI}$, Supplementary Table 1) driving the expression of the regulatory gene and an output promoter that is acted on by the regulator. The regulatory genes were either codon optimized and synthesized or cloned (Methods and Supplementary Table 4). The output promoters were either obtained from the literature or, in the case of $P_{Van}$, rationally designed by inserting cognate operator sequences into unregulated promoters (Supplementary Table 1). The activity of the output promoter was measured through the expression of yellow fluorescent protein (YFP) using flow cytometry and reported in relative promoter units (RPUs) (Supplementary Figure 1 and Methods). Each complete sensor was cloned into the same sensor plasmid architecture (Supplementary Figure 2). Some sensors require additional genes, which are encoded as an operon with the regulatory protein. For the Ara-inducible system, the transporter *araE* was included in order to produce a graded response[52]. For the Ery-inducible system, the ribosome methylase $ery^R$ was included to confer resistance to Ery[53].

Rational mutations were made to improve some sensors prior to performing the direction evolution experiments. Multiple versions of each sensor, each with different a promoter and RBS used to drive the expression of the regulator, were tested (Supplementary Figure 3). The version with the largest dynamic range was chosen for further optimization. Then, a number of potential improvements gleaned from the literature or rationally designed were evaluated, the results of which are shown in Supplementary Figure 4. After this step, we reduced the set of sensors to 14, removing copper, glucaric acid and paraquat inducible systems because their responses were too small for subsequent optimization. The response functions of the initial sensors are shown in Figure 2 (grey curves). Each function was obtained by fitting the experimental data to the equation

$$y = y_{min} + (y_{max} - y_{min})\frac{x^n}{K^n + x^n} \quad , \qquad (1)$$

where $y$ is the promoter activity in RPU, $x$ is the concentration of the small molecule, $y_{min}$ is the leakiness, $y_{max}/y_{min}$ is the dynamic range, $K$ is the sensitivity, and $n$ is the cooperativity. The raw data points, including error bars, used for this fit are shown in Supplementary Figure 5. While they all show some response, the high leakiness, low dynamic range and low sensitivity are apparent for many. Of the 14, the aTc- and OC6-sensors produced a good enough response to not require additional optimization (Table 1).

The remaining 12 were then subjected to directed evolution using the dual selection (Figure 1c, Supplementary Figure 6). For each sensor, a library was constructed and cloned upstream of the operon containing PheS and DNAP (Figure 1a) on a dual selection plasmid. The initial library contained a mixture of random and rational mutations. For a subset of sensors, the output promoter was mutagenized: selected bases in the -10 box and -35 box were randomized. To control regulator expression, critical bases in the ribosome binding site (RBS) were randomized. For LacI and AraC, we partially mutagenized key amino acid residues based on prior work[23, 33-34]. Specifically, the genes were PCR amplified with primers of limited degeneracy (*e.g.,* VNA, WKK, or NDC) thus allowing LacI Q18, F161, W220, Q291, and L296 and AraC L133, E165, E169, and C280 to sample a subset of possible amino acids (Methods). The specific mutations made to the initial library for each sensor are shown in Supplementary Appendix 2.

Multiple rounds of the dual selection were performed with the initial library. Between 4 and 23 rounds were performed, depending on how many issues needed to be corrected for each sensor. Different interventions were performed during each round to bias solutions to address problems identified for each sensor. The conditions for each round are presented in detail in Supplementary Appendix 2. Typically, the stringency of the negative selection was increased at a particular round by increasing the concentration of Cl-Phe from 2 mM to 4 mM. This biases against leakiness in the absence of inducer. During positive selection, inducer was added to the surviving cells, leading to the expression of DNAP. During early rounds, the maximum amount of inducer was added. For some sensors, we sought to increase the sensitivity of the response by reducing the amount of inducer after each round; for example, BetI was induced with 1 mM Cho in the final round of selection, down from 5 mM Cho in the first round. After induction, the cells were encapsulated, lysed, and PCR amplified using the DNAP expressed by the sensor (Methods). In early rounds, additional random mutations throughout the sensor were introduced by using the PK6 DNAP during the positive selection (yielding an average of 1-2 mutations per kilobase). In later rounds, the stringent KOD DNAP was used to reduce the diversity in the population. After the amplification step of the positive selection, the constructs were recloned into the selection plasmid. Recloning allows the selection plasmid to be reset each round, thus preventing the accumulation of cheaters, and offers the opportunity to change the DNAP as needed. In some cases, in an effort to

combine multiple beneficial mutations into a single variant, gene shuffling[54] was used between rounds (Methods).

During the selection rounds, additional interventions were included in order to reduce crosstalk between systems. For example, crosstalk between the IPTG- and Ara- sensors is well known, where IPTG reduces the output of the AraC/$P_{BAD}$ system[23] (Supplementary Figure 7). To identify mutants that reduce this, 1000 μM IPTG was included during all rounds of selection for the Ara sensor. IPTG was also included in the negative selection to prevent the evolution of an IPTG-induced AraC mutant. Crosstalk between Van and DHBA was also observed and eliminated through the negative selection (Figure 2b, Supplementary Figure 8). Salycilate was also found to antagonize the Cuma sensor (Figure 2c, Supplementary Figure 7) and cross react with the DAPG sensor (Figure 2b, Supplementary Figure 8) and was thus added to both the positive and negative selections.

After all of the rounds of selection are complete, the library was assembled into a YFP screening plasmid (Supplementary Figure 6). Several clones were picked and assayed for output expression in the presence and absence of inducer (Supplementary Appendix 2). In addition, when selecting against crosstalk, the clones were screened for induction by these molecules. The mutants showing highest improvement (leakiness, sensitivity, dynamic range, and orthogonality) were identified and sequenced (Supplementary Appendix 2). Based on the mutations observed, one or more consensus sequences were constructed and re-screened to identify the best variant. The final sequences of the evolved sensors, including the mutations identified are provided in Supplementary Table 4. Mutations were identified throughout the sensors, including the promoter/RBS controlling regulator expression, synonymous and non-synonymous mutations throughout the regulator genes, and mutations/substitutions in the output promoters (Figure 2a). On average, about eight cumulative mutations were made to each sensor as a result of the rounds of dual selection.

The improvements in the response functions are shown in Figure 2a (blue curves). The fit parameters to Equation 1 are provide in Table 1 (raw data are provided in Supplementary Figure 5). There is marked improvement in many of the response functions, sometimes showing orders of magnitude changes in the leakiness, dynamic range, and sensitivity. The Ery- and Acu- sensors showed slight improvements in their response functions, but were not chosen to be part of the final set because of their low dynamic range.

CymR, NahR, VanR, PhlF, and PcaU each responds to a substituted benzene. Therefore, we examined the activity of this set of sensors against all five inducers. The optimized sensors showed significant reduction in the crosstalk while maintaining high activity with their cognate inducer (Figure 2b,

Supplementary Figure 8). Improvements in the antagonism between CymR and Sal were also tested (Figure 2c, Supplementary Figure 7). In the presence of 100 µM Sal, the ability for Cuma to induce its sensor drops by 1200-fold. The sensor obtained by rounds of positive selection in the presence of Cuma reduces this by two orders of magnitude. There is also slight antagonism of AraC by IPTG, which also improved as a result of the selection (Supplementary Figure 7). Collectively, these improvements allow for all of these sensors to be used simultaneously in a single cell.

The best 12 sensors were then combined to form a "sensor array" that was inserted into the genome of *E. coli* MG1655 to create "Marionette-Wild" (Figure 3 and Supplementary Table 6). Genomic insertion has several benefits: it stabilizes the cluster and simplifies the use of multiple systems without building large plasmids containing the regulators. The array consists of the 12 regulatory genes and *araE* transporter organized into several operons (Figure 3a and Supplementary Figure 9). The genes were organized into three operons controlled by three medium-strength constitutive promoters. Each gene is encoded by its own ribosome binding site (RBS), which was rationally designed using the RBS Calculator[55] in order to achieve equivalent expression as when encoded on the plasmid (Methods and Supplementary Figure 10). Strong terminators were included before and after the sensor array in order to insulate the array from context effects. Phage transduction was used to move the sensor array to create two additional cell lines: the recA-deficient *E. coli* DH10B strain for cloning "Marionette-Clo" and the protease-deficient *E. coli* BL21 for protein expression "Marionette-Pro" (Methods).

The responses of the 12 genomically-encoded sensors in Marionette-Wild are shown in Figure 3b, the parameters derived from the fits to Equation 1 are shown in Table 1, and the raw data points are provided in Supplementary Figure 11 and Supplementary Appendix 1. Each response was measured by transforming the strain with a p15a (Supplementary Figure 2) plasmid containing the output promoter fused to YFP. Each response function shows at least 100-fold induction with similar levels of on- and off-states. The technical information for the use of each sensor is organized as a series of datasheets in the Supplementary Information. The response of each sensor was also measured in Marionette-Clo and Marionette-Pro (Supplementary Appendix 1). The performances of the sensors closely match that of Marionette-Wild with several exceptions; notably, the responses to IPTG ($P_{Tac}$) and choline ($P_{BetI}$) are leakier.

All of the sensors follow similar induction dynamics, with induction after 15 minutes and full induction by 2 hours. Interestingly, those sensors based on activators were slower to turn on, as compared to those based on repressors (Supplementary Figure 12). There is little cross reactivity from the non-cognate inducers (Figure 3c and Supplementary Figure 13). The response functions for three

different promoters were also measured in the presence of the maximum levels of all 11 other inducers, and there was little change in dynamic range (Supplementary Figure 14). The sensor responses were measured during exponential growth. To evaluate performance in stationary phase, cells were grown overnight (~20 hours) and response functions were measured (Methods). The responses closely match those measured during exponential growth, with several exceptions (Supplementary Figure 12). Induction on plates where the inducers are added to LB-agar resulted in similar responses as those observed in early stationary phase (Supplementary Figure 15).

We tested whether carrying the sensor array impacts the growth rate of the three Marionette strains (Supplementary Figure 16). The Marionette-Wild strain grows with a doubling time of 29.0 $\pm$ 1.9 min as compared 27.1 $\pm$ 1.2 for wild-type *E. coli* MG1655 (Methods). The other two Marionette strains grow slightly faster, albeit within error (Supplementary Figure 16). Growth of Marionette-Wild was also evaluated in the presence of all 12 inducers at maximum levels and only a modest effect was observed (Supplementary Figure 16).

Carrying the sensor array requires the continuous expression of 12 regulatory proteins and a transporter. This could lead to a draw on cellular resources that confers a selective advantage to eliminating the array. While genomic insertion improves evolutionary stability[21, 56-57], it could still be disrupted over time, particularly for the *recA*-positive[58] Marionette-Wild. To address this, we performed three independent experiments to assess the evolutionary stability of Marionette-Wild. First, we determined whether Marionette could reliably control a plasmid based promoter, even after extended passage. The 12 reporter strains were passaged for 14 days in liquid culture without inducer, diluting the cells $10^6$-fold each day ($10^{84}$-fold dilution total). On each day, a subset of cells from each line were grown and assayed with and without inducer (Figure 3d and Supplementary Figure 17). Second, the Marionette-Wild strain was passaged for nine days, streaking cultures on agar plates and inoculating single colonies into liquid culture each day. On the tenth day, the culture was transformed with each of the 12 reporter plasmids and assayed with and without the cognate inducer for each reporter. Third, we determined if serial transfer would lead to the emergences of subpopulations[59]. We passaged the Marionette-Wild strain for nine days in liquid culture, diluting the cells $10^6$-fold each day ($10^{54}$-fold dilution total). On the tenth day, the culture was transformed with each of the 12 reporter plasmids and assayed with and without the cognate inducer for each reporter. For all three evolution experiments, the sensors perform indistinguishably after growth and passaging (Supplementary Figure 17). There is no decline in the fold-induction over time and there was no emergence of "broken" subpopulations by flow cytometry (Supplementary Figure 18).

This work represents a dramatic expansion in our ability to study and control genes in cells. The Marionette strains enable the modular control of up to 12 genes, simply by placing each one under the control of a small (50 to 289 base pair) inducible promoter. This means that a single construct can be built and then the expression levels perturbed in many ways through the combination of different small molecules. This could be to determine the role of proteins in a natural system, picking apart the stoichiometric requirements for a molecular machine[60-61]. This can also be part of rapid optimization of metabolic pathways, where the ideal stoichiometry can be identified without the need to build megabase-scale libraries[62-63]. Through the use of CRISPRi[64], sRNA[65], or other tools, endogenous genes can be inducibly down-regulated as well as up-regulated, thus enabling exquisite control of natural processes and metabolic flux. Further, dynamic systems can be probed by controlling the timing of the induction of each component in order to determine the role of ordered gene expression[15, 66]. The ability to control gene expression has been a major limitation in genetic engineering; now, pulling the strings on Marionette enables unprecedented genetic control.

**Acknowledgements**

Supported by U.S. Office of Naval Research Multidisciplinary University Research Initiative grant # N00014-16-1-2388 (AJM, THSS, and CAV). We would like to thank Aditya M. Kunjapur and Kristala L. J. Prather for providing DNA templates for the amplification of $P_{Van}$, $P_{3B5}$, *vanR*, and *pcaU*.

**Accession**

The following plasmids and strains can be acquired from Addgene: pAJM.711 ($P_{PhlF}$-YFP) 108512; pAJM.712 ($P_{CymRC}$-YFP) 108513; pAJM.713 ($P_{LuxB}$-YFP) 108514; pAJM.714 ($P_{VanCC}$-YFP) 108515; pAJM.715 ($P_{Tac}$-YFP) 108516; pAJM.717 ($P_{Tet*}$-YFP) 108517; pAJM.716 ($P_{BAD}$-YFP) 108518; pAJM.718 ($P_{BetI}$-YFP) 108519; pAJM.719 ($P_{Ttg}$-YFP) 108520; pAJM.1459 ($P_{3B5C}$-YFP) 108521; pAJM.721 ($P_{SalTTC}$-YFP) 108522; pAJM.944 ($P_{Cin}$-YFP) 108523; pAJM.847 ($PhlF^{AM}$ + $P_{PhlF}$-YFP) 108524; pAJM.657 ($CymR^{AM}$ + $P_{CymRC}$-YFP) 108525; pAJM.474 (LuxR + $P_{LuxB}$-YFP) 108526; pAJM.773 ($VanR^{AM}$ + $P_{VanCC}$-YFP) 108527; pAJM.336 ($LacI^{AM}$ + $P_{Tac}$-YFP) 108528; pAJM.011 (TetR + $P_{Tet*}$-YFP) 108529; pAJM.677 ($AraC^{AM}$ + AraE + $P_{BAD}$-YFP) 108530; pAJM.683 ($BetI^{AM}$ + $P_{BetI}$-YFP) 108531; pAJM.661 ($TtgR^{AM}$ + $P_{Ttg}$-YFP) 108532; pAJM.690 ($PcaU^{AM}$ + $P_{3B5B}$-YFP) 108533; pAJM.771 ($NahR^{AM}$ + $P_{SalTTC}$-YFP) 108534; pAJM.1642 ($CinR^{AM}$ + $P_{Cin}$-YFP) 108535; pAJM.884 ($AcuR^{AM}$ + $P_{Acu}$-YFP) 108536; pAJM.969 ($MphR^{AM}$ + EryR + $P_{Mph}$-YFP) 108537; sAJM.1504 (Marionette-Clo) 108251; sAJM.1505 (Marionette-Pro) 108253; sAJM.1506 (Marionette-Wild) 108254.

## Author contributions

A.J.M and C.A.V. conceived the study and designed the experiments; A.J.M performed the experiments; A.J.M and T.H.S.S. analyzed the data; A.J.M and C.A.V. wrote the manuscript with input from all the authors.

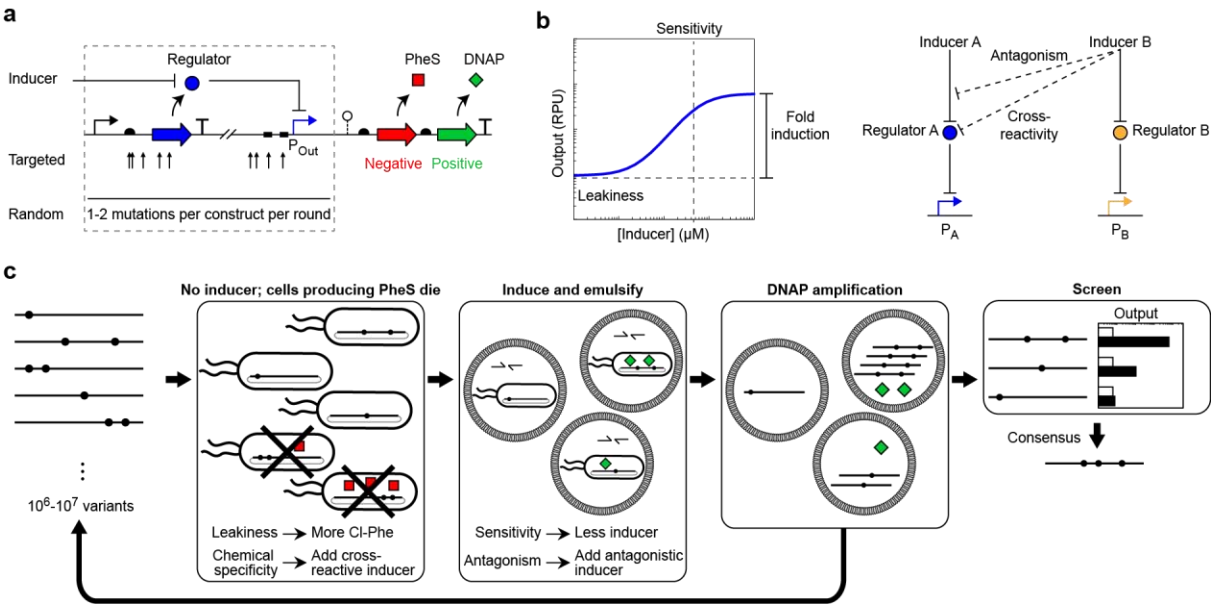


**Figure 1: A dual selection for sensor optimization. a)** A regulator (blue circle) is expressed from a weak constitutive promoter and controls expression from the output promoter ($P_{Out}$) while itself being controlled by an externally applied inducer molecule. Transcription from $P_{Out}$ determines the expression level of an aminoacyl-tRNA synthetase (PheS; red square) and a DNA polymerase (KOD or PK6; green diamond). Initial libraries may contain targeted degeneracy in the regulator RBS, regulator CDS, or $P_{Out}$ (arrows), and random mutations are added throughout the entire library during each round of selection. **b)** A response function captures the activity of $P_{Out}$ at various levels of inducer. Regulator A affects promoter A, and is itself affected by inducer A. Inducer B may affect Regulator A (chemical cross-reactivty) or interfere with inducer A's action (antagonism). **c)** The dual selection schema is shown (see text for details). Dots denote mutations; half arrows represent PCR primers.

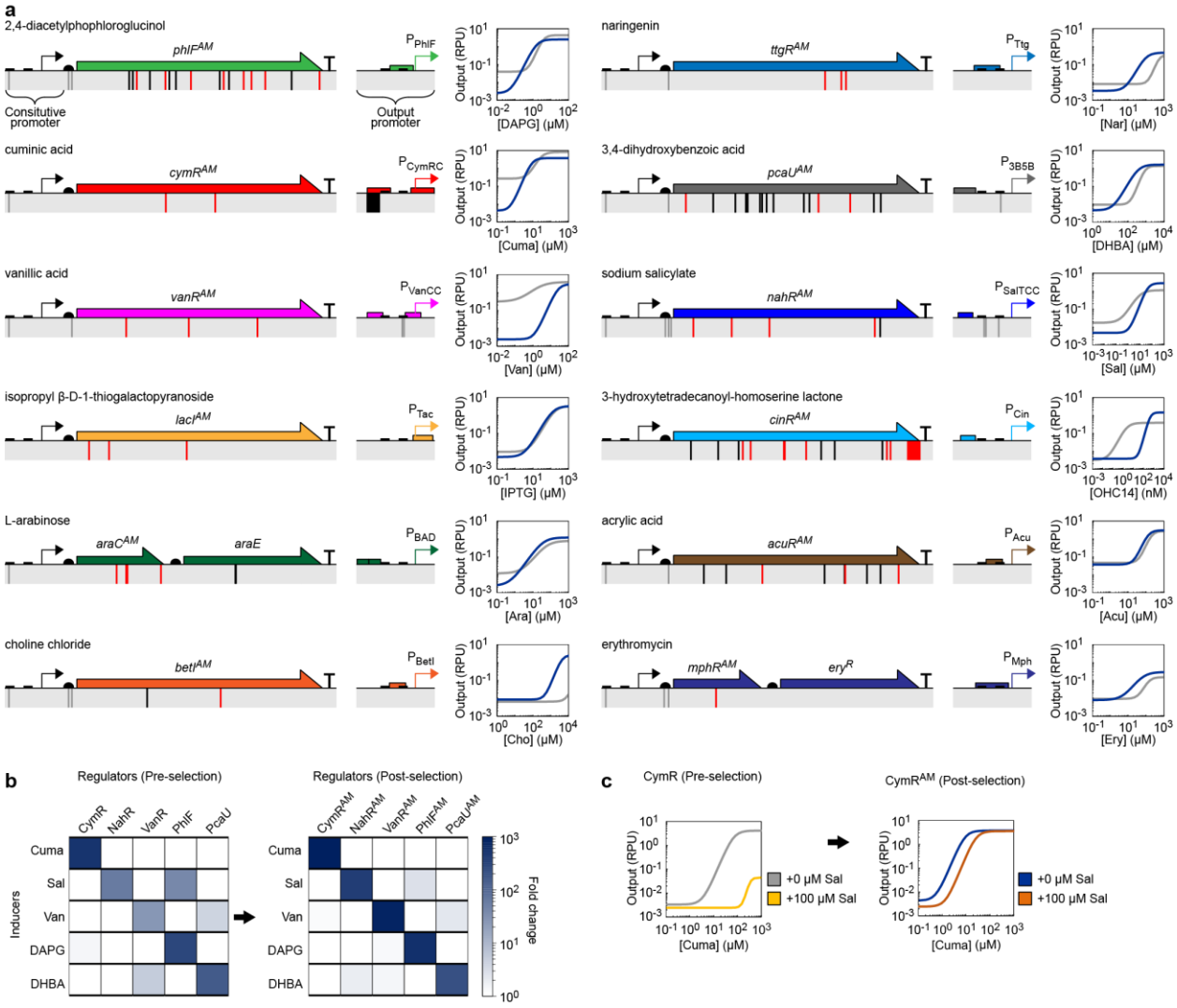


**Figure 2: Improved sensor performance. a)** Each genetically encoded sensor is shown, with coding (red) and non-coding (grey/black) mutations noted. Mutations in grey were also applied to the parental sensor. The corresponding response functions comparing the evolved (blue) and parental (grey) sensors. The fit of Equation 1 to the mean of three replicates performed on different days is shown (see Supplementary Figure 5 for data). The response function parameters for evolved sensors are provided in Table 1. **b)** Chemical cross-reactivity heat-map of reference (left) and evolved (right) sensors. Inducer concentrations were: 100 μM Cuma, 100 μM Sal, 100 μM Van, 10 μM DAPG, and 1 mM DHBA. The mean of three replicates performed on different days is shown (see Supplementary Figure 8 for data). **c)** Response function with Sal (light/dark orange) and without Sal (grey/blue) for parental CymR (left) and evolved CymR$^{AM}$ (right). The fit of Equation 1 to the mean of three replicates from different days is shown (see Supplementary Figure 7 for data). Sequences of promoters and regulators are provided in Supplementary Table 1 and 4.

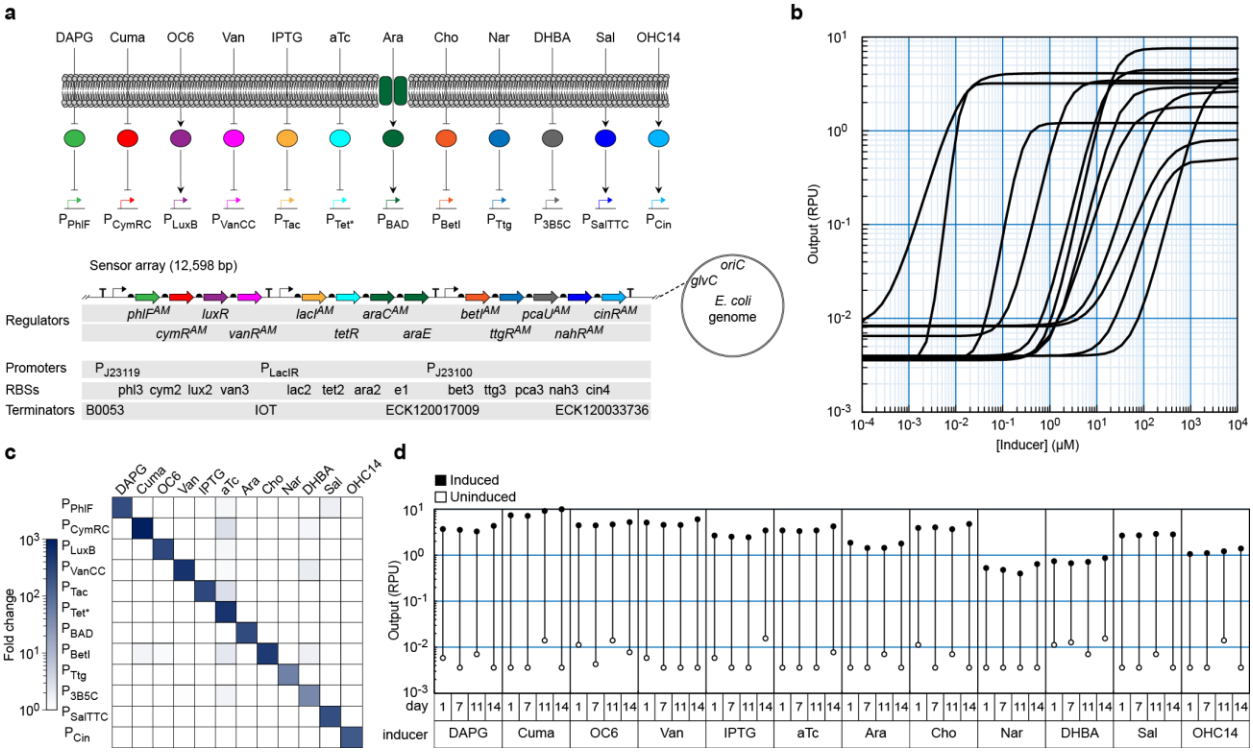


**Figure 3: Marionette-Wild performance. a)** The molecular and genetic schematic of the Marionette cluster and its location in the *E. coli* genome. The cluster was inserted in the direction of leading strand replication, between 3,860,010 and 3,861,627 in *E. coli* MG1655 (NCBI accession number NC_000913), between 3,957,594 and 3,959,211 in *E. coli* DH10B (NC_010473), and between 3,720,027 and 3,721,644 in *E. coli* BL21 (CP010816). **b)** Overlaid response functions for each of the 12 output promoters with their cognate inducers. The fit of Equation 1 to the mean of three replicates performed on different days is shown (see Supplementary Figure 11 for data). The response function parameters are provided in Table 1. **c)** Chemical cross-reactivity heat-map of the 12 output promoters with each inducer. Inducer concentrations were: 25 μM DAPG, 500 μM Cuma, 10 μM OC6, 100 μM Van, 1 mM IPTG, 200 nM aTc, 4 mM Ara, 10 mM Cho, 1 mM Nar, 2.5 mM DHBA, 250 μM Sal, and 10 μM OHC14. The mean of three replicates performed on different days is shown (see Supplementary Figure 13 for data). **d)** Output of uninduced (open circles) and induced (closed circles) cultures on day 1, 7, 11, and 14 days of passaging. Inducer concentrations were: 25 μM DAPG, 500 μM Cuma, 10 μM OC6, 100 μM Van, 1 mM IPTG, 200 nM aTc, 4 mM Ara, 10 mM Cho, 1 mM Nar, 2.5 mM DHBA, 250 μM Sal, and 10 μM OHC14. A single evolutionary trajectory is shown. Passaging details and data from other days are provided in Supplementary Figure 17.

**Table 1. Sensor response function parameters[a]**

| Inducer[d] | | Max inducer (μM)[b] | Genome-based regulator (Marionette-Wild[c]) $y_{max}$ (RPU) | $y_{min}$ (RPUx10$^3$) | $K$ (μM) | $n$ | Plasmid-based regulator (*E. coli* DH10B) $y_{max}$ (RPU) | $y_{min}$ (RPUx10$^3$) | $K$ (μM) | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| DAPG | 2,4-Diacetylphophloroglucinol | 25 | 3.4 | 6.5 | 2.1 | 2.3 | 2.5 | 2.5 | 1.7 | 2.1 |
| Cuma | Cuminic acid | 100 | 7.6 | 3.6 | 22 | 2.3 | 3.7 | 4.3 | 8.9 | 2.4 |
| OC6 | 3OC6-AHL | 10 | 4.1 | 8.6 | 0.012 | 1.7 | 1.3 | 2.4 | 0.12 | 1.8 |
| Van | Vanillic acid | 100 | 4.5 | 4.0 | 14 | 2.1 | 3.0 | 2.4 | 26 | 2.3 |
| IPTG | Isopropyl-beta-D-thiogalactoside | 1000 | 2.6 | 8.3 | 190 | 1.7 | 3.3 | 4.8 | 140 | 1.8 |
| aTc | Anhydrotetracycline HCl | 0.2 | 3.2 | 3.6 | 0.012 | 4.4 | 2.4 | 4.9 | 0.013 | 3.8 |
| Ara | L-Arabinose | 4000 | 1.8 | 3.6 | 43 | 1.7 | 1.2 | 2.4 | 37 | 1.5 |
| Cho | Choline chloride | 10000 | 3.7 | 4.0 | 1900 | 2.0 | 2.6 | 8.5 | 4100 | 2.7 |
| Nar | Naringenin | 1000 | 0.5 | 4.0 | 280 | 2.3 | 0.5 | 3.4 | 95 | 1.9 |
| DHBA | 3,4-Dihydroxybenzoic acid | 1000 | 0.8 | 8.0 | 240 | 1.5 | 1.6 | 4.5 | 370 | 1.8 |
| Sal | Sodium salicylate | 100 | 2.9 | 3.8 | 29 | 2.1 | 2.8 | 4.7 | 43 | 1.8 |
| OHC14 | 3OHC14:1-AHL | 10 | 1.2 | 3.6 | 0.25 | 3.0 | 1.5 | 3.0 | 0.43 | 2.3 |
| Acu | Acrylic acid | 1000 | | | | | 3.1 | 37.0 | 130 | 2.5 |
| Ery | Erythromycin | 125 | | | | | 0.3 | 8.0 | 65 | 1.5 |

a. Response data from at least days were averaged and fit to Equation 1 (Methods). Full response functions are provided in Supplementary Appendix 1.
b. Growth defects are observed above this concentration.
c. Based on *E. coli* MG1655. Data for Marionette-Clo (DH10B) and Marionette-Pro (BL21) are provided in Supplementary Appendix 1.
d. DAPG-Santa Cruz sc-206518; Cuma-Sigma 268402; OC6-Sigma K3007; Van-Sigma 94770; IPTG-Gold I2481C25; aTc-Sigma 37919; Ara-Sigma A3256; Cho-Sigma C7017; Nar-Sigma N5893; DHBA-Sigma 37580; Sal-Sigma S3007; OHC14-Sigma 51481; Acu-Sigma 147230; Ery-Sigma E5389.

## Methods

**Strains, plasmids, and media.** *Escherichia coli* DH10B (New England Biolabs, Ipswich, MA – USA) was used for all routine cloning and directed evolution. Plasmid-based regulator systems were characterized in *E. coli* DH10B. Marionette-Wild, -Clo, and -Pro were derived from *E. coli* MG1655[67], *E. coli* DH10B, and *E. coli* BL21 (New England Biolabs, Ipswich, MA – USA) cells respectively. *E. coli* JTK164H was used to clone RK6 suicide vectors[68]. All other plasmids contain p15A origins of replication and kanamycin resistance (Supplementary Figures 2 and 6). LB-Miller media (BD, Franklin Lakes, NJ - USA) was used for directed evolution, stability assays, and cytometry assays unless specifically noted. 2xYT liquid media (BD, Franklin Lakes, NJ - USA) and LB + 1.5% agar (BD, Franklin Lakes, NJ – USA) plates were used for routine cloning and strain maintenance. M9 media (1x M9 Salts [Millipore Sigma, St. Louis, MO - USA], 2 mM $MgSO_4$, 100 µM $CaCl_2$, and 0.2% Casamino acids) supplemented with either 0.4% glucose or 0.4% glycerol was used for cytometry assays where noted.

**Chemical transformation.** For routine transformations, strains were made competent for chemical transformation. Overnight cultures (250 µl for *E. coli* DH10B derived cells, 100 µl for *E. coli* MG1655 and BL21 derived cells) were subcultured into 100 ml SOB media (BD, Franklin Lakes, NJ – USA) and grown at 37 °C, 250 rpm for 3 hours. Cultures were centrifuged (4500*g*, 4 °C, 10 minutes) and pellets were resuspended in 15 ml TFBI buffer[51] (30 mM KOAc, 50 mM $MnCl_2$, 100 mM RbCl, 10 mM $CaCl_2$, and 15% v/v glycerol, pH 5.0). After 1 hour on ice, cells were centrifuged (4500*g*, 4 °C, 10 minutes) and pellets were resuspended in 2 ml TFBII buffer (10 mM NaMOPS pH 7.0, 75 mM $CaCl_2$, 10 mM RbCl, and 15% v/v glycerol). Competent cells were stored at -80 °C until use.

**Response function measurements (mid-log phase).** All measurements shown were taken by cytometry of cells in mid-log growth except when noted. Glycerol stocks of strains containing the plasmids of interest were streaked on LB + 1.5% Agar plates and grown overnight at 37 °C. Single colonies were inoculated into 1 ml LB + antibiotics in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB + antibiotics in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours the growths were diluted (*E. coli* DH10B/Marionette-Clo 1:500; *E. coli* BL21/Marionette-Pro 1:2,000; *E. coli* MG1655/Marionette-Wild 1:5,000) into prewarmed LB + antibiotics + inducer where necessary in

2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 5 hours. After growth, 20 µl of culture sample was diluted into 180 µl PBS + 200 µg/ml kanamycin to inhibit translation.

**Response function measurements (stationary phase).** Measurements were taken from cells in stationary phase to generate data shown in Supplementary Appendix 1 and Supplementary Figures 12 and 13. Glycerol stocks of strains containing the plasmids of interest were streaked on LB + 1.5% Agar plates and grown overnight at 37 °C. Single colonies were inoculated into 1 ml LB + antibiotics in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB + antibiotics in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours the growths were diluted (*E. coli* DH10B/Marionette-Clo 1:500; *E. coli* BL21/Marionette-Pro 1:2,000; *E. coli* MG1655/Marionette-Wild 1:5,000) into prewarmed LB + antibiotics + inducer where necessary in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 20 hours. After growth, 2 µl of culture sample was diluted into 198 µl PBS + 200 µg/ml kanamycin to inhibit translation.

**Time course (mid-log phase)**. For Supplemental Figure 12: mid-log induction time course, glycerol stocks of strains containing the plasmids of interest were streaked on LB + 1.5% Agar plates and grown overnight at 37 °C. Single colonies were inoculated into 1 ml LB + antibiotics in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB + antibiotics in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours the growths were diluted 1:500 into prewarmed LB + antibiotics. After 0, 1, 2, 3, 3.5, 4, 4.25, 4.5, or 4.75 hours, cultures were further diluted 1:10 into prewarmed LB + antibiotics + inducer where necessary in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 5, 4, 3, 2, 1.5, 1, 0.75, 0.5, or 0.25 hours (5 hours total after the initial growth). After growth, 20 µl of culture sample was diluted into 180 µl PBS + 200 µg/ml kanamycin to inhibit translation.

**Time course (mid-log phase to stationary phase)**. For Supplemental Figure 12: mid-log to stationary induction time course, glycerol stocks of strains containing the plasmids of interest were streaked on LB + 1.5% Agar plates and grown overnight at 37 °C. Single colonies were inoculated into 1 ml LB + antibiotics

in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB + antibiotics in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours the growths were diluted 1:5000 into prewarmed LB + antibiotics + inducer where necessary in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 5, 6, 7, 8, 9, 10, or 20 hours. After growth, 2 to 20 µl of culture sample was diluted into 180-198 µl PBS + 200 µg/ml kanamycin to inhibit translation.

**Cytometry analysis.** Fluorescence characterization with cytometry was performed on a BD LSR Fortessa flow cytometer with HTS attachment (BD, Franklin Lakes, NJ - USA). Cells diluted in PBS + kanamycin were run at a rate of 0.5 µl/s. The events were gated by forward scatter height (mid-log: 1,000-10,000; stationary: 500-5000) and side scatter area (mid-log: 1,000-10,000; stationary: 500-5,000) to reduce false events. After gating, thousands of events were used for analysis. For each sample, the median YFP fluorescence was calculated. All output values are reported in terms of relative promoter units (RPU). For a given promoter measurement, the strain (*E. coli* DH10B, Marionette-Wild, *etc*) is transformed with the plasmid. The strain is then assayed alongside a strain containing the RPU standard plasmid (Supplementary Figure 1, Supplementary Table 8) as well as an autofluorescence control. The median autofluorescence value is subtracted from the all other median fluorescence values, including that of the RPU standard. The experimental sample value is then divided by the RPU standard value.

**Library generation.** Portions of the initial libraries were created by PCR using degenerate oligonucleotides (Integrated DNA Technologies Coralville, IA - USA). These fragments were joined into a degenerate, full-length sensor module by overlap PCR. Sensor modules were assembled into selection vectors by Golden Gate assembly[70]. Linear insert and plasmid selection vector were mixed at 1:1 molar ratio totaling ~ 1 µg DNA along with 5 µl 10x T4 ligase Buffer, 1 µl T4 DNA ligase (2,000,000 U/ml), 2 µl BbsI (10,000 U/ml) (all from New England Biolabs, Ipswich, MA - USA) in 50 µl total. Reactions were cycled 45 times between 2 minutes at 37 °C and 5 minutes at 16 °C, and then incubated for 30 minutes at 50 °C, 30 minutes at 37 °C, and 10 minutes at 80 °C in a DNA Engine cycler (Bio-Rad, Hercules, CA – USA). An additional 1 µl BbsI was then added, and the assembly was incubated for 1 hour at 37 °C. Assemblies were then purified using Zymo Spin I columns (Zymo Research, Irvine, CA - USA). Host cells were transformed with library plasmid by electroporation. Supplementary Appendix 2 contains a depiction of all degeneracy found in the "Initial library" for each selection.

**Negative selection.** LB media + 8 mM Cl-Phe (4-Chloro-DL-phenylalanine [Millipore Sigma, St. Louis, MO - USA]) was mixed, autoclaved, and stored at room temperature. Cl-Phe has a tendency to adhere to glassware, and care was taken to avoid disturbing the water + LB powder + Cl-Phe powder mixture prior to autoclaving. LB media + 8 mM Cl-Phe was mixed with plain LB media to achieve the desired concentration of Cl-Phe (See Interventions:[Cl-Phe] for each selection in Supplementary Appendix 2). Following transformation and outgrowth, cultures were diluted into 7 ml LB + Cl-Phe supplemented with antibiotics as well as any cross-reactive inducers (See Interventions:[Negative ligand] for each selection in Supplementary Appendix 2). Cultures were grown at 37 °C for 12-16 hours.

**Positive selection.** Following negative selection, cultures were diluted 1:100 into 2 ml LB + antibiotics in culture tubes and grown at 37 °C, 250 rpm for 2 hours. The cultures were then diluted 1:100 into 2 ml prewarmed LB + antibiotics + inducer (See Interventions:[Positive ligand] for each selection in Supplementary Appendix 2) and grown at 37 °C, 250 rpm for 4 hours. Following induction, 1 ml of culture was centrifuged at (5,000*g*, 25 °C, 10 minutes). Supernatant was removed and the cell pellet was resuspended in 50 µl 1x CPR buffer (50 mM Tris-HCl pH 8.8, 10 mM KCl, 2 mM $MgSO_4$, 10 mM $(NH_4)_2SO_4$). 5 µl resuspension was added to 95 µl of 1 x CPR buffer plus 0.4 µM CPR primers and 200 µM dNTPs. This aqueous phase added to a 2 ml centrifuge tube containing 438 µl Tegosoft DEC (Evonik, Essen, Germany), 42 µl AbilWE09 (Evonik, Essen, Germany) and 120 µl mineral oil (Millipore Sigma, St. Louis, MO - USA) and the rubber stopped from a 1 ml syringe. The mixture was vortexed at maximum setting for 2 minutes. The emulsion was split evenly five 0.2 ml PCR tubes, and thermal cycled (95 °C for 5 minutes; 20 cycles of [95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 2 minutes/kb]; and 72 °C for 5 min). Emulsions were then centrifuged (10,000*g*, 25 °C, 10 minutes) and the upper (oil) phase was removed. Then, 100 µl $H_2O$ and 500 µl chloroform were added and the mixture was pipetted to disrupt the pellet. The resuspension was then transfered to a 1.5 ml heavy-gel phase-lock tube (5 Prime, San Francisco, CA - USA) and centrifuged (16,000*g*, 25 °C, 2 minutes). The upper (aqueous) phase was collected, and DNA was purified using Zymo Spin I columns. The library was amplified in a recovery PCR using Accuprime *Pfx* and nested recovery primers, gel purified, and assembled as described above. A depiction of the CPR primer and recovery primer amplification are provided in Supplementary Figure 6.

**Library shuffling.** Between some rounds of selection (See Interventions:Notes for each selection in Supplementary Appendix 2), libraries were shuffled upon themselves such that fragments of each mutant

mutually serve as both primer and template for a primer-less PCR amplification[54]. 1 µg linear, library DNA (from the recovery PCR) was added to a mild DNAse reaction (500 mM Tris pH 7.4, 100 mM MnCl2, 0.5 U DNAse [New England Biolabs, Ipswich, MA - USA]) and lightly digested for 3 minutes at 15 °C. Fragmented DNA was purified using Zymo Spin I columns and reassembled in a primer-less PCR in 1 x KAPA HiFi Master Mix (KAPA Biosystems, Wilmington, MA – USA) by thermal cycling (95 ˚C for 2 minutes; 35 cycles of [95 ˚C for 30 seconds, 65 ˚C for 90 seconds, 62 ˚C for 90 seconds, 59 ˚C for 90 seconds, 56 ˚C for 90 seconds, 53 ˚C for 90 seconds, 50 ˚C for 90 seconds, 47 ˚C for 90 seconds, 44 ˚C for 90 seconds, 41 ˚C for 90 seconds, 68 ˚C for 90 seconds]; and 72 ˚C for 4 minutes). The reassembly was purified using Zymo Spin I columns, reamplified using Accuprime *Pfx* and CPR primers, gel purified, and assembled as described above.

**On/off screen.** At the end of each selection, libraries were assembled into the YFP screening plasmid (Supplementary Figure 6). Host cells were transformed and plated on LB-agar. Between 35 and 92 individual clones were picked and assayed by cytometry as described. Cells were grown with no inducer, in the presence of cognate inducer, and when necessary in the presence of relevant non-cognate inducers. Measurements were made of cells in mid-log phase by cytometry as described. The most promising clones, as judged by dynamic range and orthogonality, were mini-prepped and the sensor region was sequenced.

**Genomic integration.** In preparation of recombineering, cells were transformed with a plasmid containing Ara-inducible λ Red recombination machinery with a temperature sensitive origin of replication[71]. 50 µl of overnight culture was subcultured in 50 ml LB medium and grown at 30 °C, 250 rpm for 2 hours. 2 mM Ara was added, and the culture continued to grow at 30 °C, 250 rpm for 3 hours. The culture was then centrifuged (4500*g*, 4 °C, 10 minutes) and washed with ice cold 10% glycerol four times, with the fourth resuspension in 200 µl 10% glycerol. Recombineering-ready cells were stored at -80 °C until use. For the first insertion, six genes (*phlF*$^{AM}$, *cymR*$^{AM}$, *luxR*, *vanR*$^{AM}$, *lacI*$^{AM}$, and *tetR*) were Golden Gate assembled using BsaI into an RK6 suicide vector (which needs Pir protein in order to propagate)[68]. Pir-expressing *E. coli* JTK164H cells were transformed, and plasmids were purified, verified, and linearized with BpiI leaving homology to the *glvC* pseudogene. Recombineering-ready *E. coli* MG1655 cells were electroporated and transformed with gel-purified, linearized inserts. After an outgrowth of one hour at 37 °C, transformations were plated on LB-agar plates + antibiotic (5 µg/ml chloramphenicol). Colonies were picked and grown at 37˚C in LB + antibiotic, and the presence of the insert was verified by colony PCR. For the second insertion, this strain was made recombineering-ready and the process was repeated

with the next set of genes (*araC*$^{AM}$, *araE*, *betI*$^{AM}$, and *ttgR*$^{AM}$) with the insert containing homology to *tetR* and *glvC* and conferring resistance to 20 µg/ml spectinomycin. For the third insertion, this strain was made recombineering-ready and the process was repeated with the final set of genes (*pcaU*$^{AM}$, *nahR*$^{AM}$, and *cinR*$^{AM}$) with the insert containing homology to *ttgR*$^{AM}$ and *glvC* and conferring resistance to 5 µg/ml chloramphenicol. See Supplementary Figure 9 for schematic details.

**Phage transduction to transfer the Marionette cluster.** 50 µl of overnight culture of Marionette-Wild was subcultured in 5 ml LB medium + 0.2% glucose and 5 mM $CaCl_2$ and grown at 37 °C, 250 rpm for 30 minutes. 100 µl p1 phage lysate ($10^9$ pfu/ml) was added and the culture was grown at 37 °C, 250 rpm until lysis (3 hours)[72]. 50 µl chloroform was added and the culture continued at 37 °C, 250 rpm for 5 minutes. Culture was centrifuged (9200*g*, 25 °C, 10 minutes). Supernatant was filtered (0.45 µM) and stored at 4 °C until use. 1.5 ml of overnight culture of *E. coli* MG1655, BL21, or DH10B (the DH10B cells contained a temperature-sensitive plasmid transiently expressing RecA) was centrifuged (10,000*g*, 25 °C, 5 minutes). The cell pellet was resuspended in 750 µl P1 buffer (5 mM $MgSO_4$ and 10 mM $CalCl_2$) and up to 100 µl Marionette-P1 phage lysate was added. After 30 minutes at 25 °C, 1 ml LB + 200 µl 1 M sodium citrate was added and the culture was grown at 37 °C, 250 rpm for 30 minutes. The culture was centrifuged (10,000*g*, 25 °C, 2 minutes) and the pellets were resuspended in 100 µl LB and plated on LB-agar plates + 5 µg/ml chloramphenicol and 5 mM sodium citrate. Colonies were picked and grown at 37 ˚C in LB + 5 µg/ml chloramphenicol and 5 mM sodium citrate, and the presence of the insert was verified by colony PCR. The entire Marionette cluster was verified by sequencing PCR amplicons of the cluster from the genome of Marionette-Wild.

**Sensor induction on plates**. Cells were transformed with reporter plasmids and, following outgrowth at 37˚C for one hour, plated on LB-agar (LB-Miller powder + 1.5 % agar [BD, Franklin Lakes, NJ – USA]) with and without the appropriate inducer. After an overnight incubation at 37˚C for 16 hours, plates were incubated at 4˚C for one hour and imaged using a ChemiDocMP Imaging system (Bio-Rad, Hercules, CA - USA) and Image Lab 4.0 software (Bio-Rad, Hercules, CA - USA) employing blue epi illumination, a 530/28 filter and a 0.1 second exposure. Raw images were rotated and cropped using XnView (XnSoft, Reims, France). Individual colonies were scraped from the plate, resuspended in 200 µl PBS + 200 µg/ml kanamycin, and assayed by cytometry as described in Cytometry analysis.

**Plate reader assays to measure growth rates**. Glycerol stocks of strains of interest were streaked on LB + 1.5% Agar plates and grown overnight at 37 °C. Single colonies were inoculated into 1 ml LB in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours the growths were diluted (*E. coli* DH10B/Marionette-Clo 1:500; *E. coli* BL21/Marionette-Pro 1:2,000; *E. coli* MG1655/Marionette-Wild 1:5,000) into prewarmed LB + inducer where necessary in 2-ml 96-deepwell plates. 100 µl of this culture was immediately transferred to a 300-µl 96-well black walled optical bottom plates (Thermo Scientific Nunc, Waltham, MA – USA) sealed with a BreathEasy film (Sigma-Aldrich, St. Louis, MO – USA), and grown in a Synergy H1 plate reader (BioTek, Winooski, VT – USA) at 37 °C, 1000 rpm. $OD_{600}$ was measured every 20 minutes over 12 hours of growth. $OD_{600}$ readings were also taken from wells containing media with no cells, and for each time point, readings from such wells were subtracted from the appropriate sample measurements to remove background. $OD_{600}$ values were converted to equivalent 1 cm path length measurements using a standard curve.

**Calculation of growth rates.** To calculate doubling times, the last measurement with $OD_{600} < 0.1$ and the first measurement with $OD_{600} > 0.4$ were identified, and the doubling time was calculated assuming exponential growth between those two points. Doubling time is calculated as elapsed time (in minutes) divided by the number of doublings that occurred in that time ($\log_2$[final $OD_{600}$/initial $OD_{600}$]).

**Evolutionary stability (passaging with reporters).** Marionette-Wild was transformed with each of the 12 reporter plasmids and plated on LB-agar. Single colonies were inoculated into 1 ml LB + antibiotics in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm overnight in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB + antibiotics in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours, the growths were diluted 1:5000 into prewarmed LB + antibiotics + inducer where necessary in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 5 hours. After growth, 20 µl of culture sample was diluted into 180 µl PBS + 200 µg/ml kanamycin to inhibit translation and assayed by cytometry as described in Cytometry analysis. Uninduced cultures were allowed to grow at 37 °C overnight (17 hours) and the process (beginning with the 1:200 dilution of the overnight culture) was repeated each day for 14 days.

**Evolutionary stability (passaging on plates).** A single colony of Marionette-Wild was inoculated into 1 ml LB in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm for 12 hours in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growth was streaked onto LB-agar plates and grown at 37 °C for 12 hours. This process was repeated 9 times. The final culture was made chemically competent as described in Chemical transformation, transformed with relevant reporter plasmids, and assayed with cytometry as described in Cytometry analysis.

**Evolutionary stability (passaging in liquid culture).** A single colony of Marionette-Wild was inoculated into 1 ml LB in 2-ml 96-deepwell plates (USA Scientific, Orlando, FL - USA) sealed with an AeraSeal film (Excel Scientific, Victorville, CA - USA) and grown at 37 °C, 900 rpm for 12 hours in a Multitron Pro shaker incubator (INFORS HT, Bottmingen, Switzerland). The overnight growths were diluted 1:200 into 1 ml LB in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm. After 2 hours, the growths were diluted 1:5000 into prewarmed LB in 2-ml 96-deepwell plates + AeraSeal film and grown at 37 °C, 900 rpm for 22 hours. This process was repeated 9 times. The final culture was made chemically competent as described in Chemical transformation, transformed with relevant reporter plasmids, and assayed with cytometry as described in Cytometry analysis.

**Computational methods (fitting to the response function).** To parameterize the response function, error minimization was performed using the Solver function in Excel software (Microsoft, Redmond, WA - USA). Equation 1 was entered with $y_{min}$ , $y_{max}$ , $K$, and $n$ as tunable parameters, $x$ (inducer concentration) as the independent variable and $y$ as the output. For each $x$, the error between the measured RPU value and the output of the function was determined. The total error was determined by summing the normalized square of the error ([$y$-measured RPU value]$^2$ /$y$) for each $x$. The Solver function minimized the total error by tuning $y_{min}$ , $y_{max}$ , $K$, and $n$.

**Computational methods (partial randomization of amino acids).** Initial libraries for the selection of LacI and AraC utilized partial randomization of amino acids. Previous literature informed decisions regarding the desirability of including each possible residue in the library. The CASTER 2.0 tool[73] was used to design degenerate oligodeoxynucleotides that can sample the desired amino acids while limiting undesired amino acids and stop codons.

**Computational methods (RBS design).** The initial RBS for each regulator in the Marionette cluster was designed using the RBS Calculator v1.1[55] using the *E. coli* DH10B 16S rRNA setting. The "Pre-Sequence" included the last 100 bp of the upstream CDS (where appropriate) as well as any scars used in Golden Gate assembly. The "Target Translation Initiation Rates" were chosen based on the RBS strength of the plasmid-based sensor (found using the "Reverse Engineer RBSs" tool) and compensating for the reduction in copy number and increase in promoter strength associated with the move to the genomic system. After initial assembly and testing, a decision was made based on the appearance of the response function as to whether the RBS was too low, too high, or good. Low RBSs were rationally mutated to more closely resemble the consensus Shine-Dalgarno sequence (TAAGGAGGT) while high RBSs were rationally mutated to less closely resemble the consensus Shine-Dalgarno sequence. All rationally designed RBS variants were checked using the "Reverse Engineer RBSs" to guard against making dramatic, unanticipated changes to the RBS strength. Rational mutations are noted in Supplementary Table 2. Translation Initiation Rate (in arbitrary units) for each RBS variant is provided in Supplementary Figure 10.

## References

1. Zuo, J.; Chua, N. H., Chemical-inducible systems for regulated expression of plant genes. *Curr Opin Biotechnol* **2000,** *11* (2), 146-51.
2. Keyes, W. M.; Mills, A. A., Inducible systems see the light. *Trends Biotechnol* **2003,** *21* (2), 53-5.
3. Mijakovic, I.; Petranovic, D.; Jensen, P. R., Tunable promoters in systems biology. *Curr Opin Biotechnol* **2005,** *16* (3), 329-35.
4. de Boer, H. A.; Comstock, L. J.; Vasser, M., The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci U S A* **1983,** *80* (1), 21-5.
5. Lutz, R.; Bujard, H., 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* **1997,** *25* (6), 1203-10.
6. Urban, J. H.; Vogel, J., Translational control and target recognition by Escherichia coli small RNAs in vivo. *Nucleic Acids Res* **2007,** *35* (3), 1018-37.
7. Cookson, N. A.; Mather, W. H.; Danino, T.; Mondragon-Palomino, O.; Williams, R. J.; Tsimring, L. S.; Hasty, J., Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Mol Syst Biol* **2011,** *7*, 561.
8. Elowitz, M. B.; Leibler, S., A synthetic oscillatory network of transcriptional regulators. *Nature* **2000,** *403* (6767), 335-8.
9. Gardner, T. S.; Cantor, C. R.; Collins, J. J., Construction of a genetic toggle switch in Escherichia coli. *Nature* **2000,** *403* (6767), 339-42.
10. Becskei, A.; Serrano, L., Engineering stability in gene networks by autoregulation. *Nature* **2000,** *405* (6786), 590-3.
11. Elowitz, M. B.; Levine, A. J.; Siggia, E. D.; Swain, P. S., Stochastic gene expression in a single cell. *Science* **2002,** *297* (5584), 1183-6.
12. Ozbudak, E. M.; Thattai, M.; Kurtser, I.; Grossman, A. D.; van Oudenaarden, A., Regulation of noise in the expression of a single gene. *Nat Genet* **2002,** *31* (1), 69-73.
13. Isaacs, F. J.; Dwyer, D. J.; Ding, C.; Pervouchine, D. D.; Cantor, C. R.; Collins, J. J., Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* **2004,** *22* (7), 841-7.
14. Golding, I.; Cox, E. C., RNA dynamics in live Escherichia coli cells. *Proc Natl Acad Sci U S A* **2004,** *101* (31), 11310-5.
15. Kalir, S.; Alon, U., Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell* **2004,** *117* (6), 713-20.
16. Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C., Real-time kinetics of gene activity in individual bacteria. *Cell* **2005,** *123* (6), 1025-36.
17. Balagadde, F. K.; Song, H.; Ozaki, J.; Collins, C. H.; Barnet, M.; Arnold, F. H.; Quake, S. R.; You, L., A synthetic Escherichia coli predator-prey ecosystem. *Mol Syst Biol* **2008,** *4*, 187.
18. Friedland, A. E.; Lu, T. K.; Wang, X.; Shi, D.; Church, G.; Collins, J. J., Synthetic gene networks that count. *Science* **2009,** *324* (5931), 1199-202.
19. Juminaga, D.; Baidoo, E. E.; Redding-Johanson, A. M.; Batth, T. S.; Burd, H.; Mukhopadhyay, A.; Petzold, C. J.; Keasling, J. D., Modular engineering of L-tyrosine production in Escherichia coli. *Appl Environ Microbiol* **2012,** *78* (1), 89-98.
20. Daugherty, P. S.; Olsen, M. J.; Iverson, B. L.; Georgiou, G., Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. *Protein Eng* **1999,** *12* (7), 613-21.
21. Caliando, B. J.; Voigt, C. A., Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nat Commun* **2015,** *6*, 6989.

661 22. Schnider-Keel, U.; Seematter, A.; Maurhofer, M.; Blumer, C.; Duffy, B.; Gigot-Bonnefoy, C.;
662 Reimmann, C.; Notz, R.; Defago, G.; Haas, D.; Keel, C., Autoinduction of 2,4-diacetylphloroglucinol
663 biosynthesis in the biocontrol agent Pseudomonas fluorescens CHA0 and repression by the bacterial
664 metabolites salicylate and pyoluteorin. *J Bacteriol* **2000,** *182* (5), 1215-25.
665 23. Lee, S. K.; Chou, H. H.; Pfleger, B. F.; Newman, J. D.; Yoshikuni, Y.; Keasling, J. D., Directed
666 evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl*
667 *Environ Microbiol* **2007,** *73* (18), 5711-5.
668 24. Grant, P. K.; Dalchau, N.; Brown, J. R.; Federici, F.; Rudge, T. J.; Yordanov, B.; Patange, O.;
669 Phillips, A.; Haseloff, J., Orthogonal intercellular signaling for programmed spatial behavior. *Mol Syst Biol*
670 **2016,** *12* (1), 849.
671 25. Scott, S. R.; Hasty, J., Quorum Sensing Communication Modules for Microbial Consortia. *ACS*
672 *Synth Biol* **2016,** *5* (9), 969-77.
673 26. Tashiro, Y.; Kimura, Y.; Furubayashi, M.; Tanaka, A.; Terakubo, K.; Saito, K.; Kawai-Noma, S.;
674 Umeno, D., Directed evolution of the autoinducer selectivity of Vibrio fischeri LuxR. *J Gen Appl Microbiol*
675 **2016,** *62* (5), 240-247.
676 27. Halleran, A.; Murray, R. M., Cell-free and in vivo characterization of Lux, Las, and Rpa quorum
677 activation systems in E. coli. *ACS Synth Biol* **2017**.
678 28. Gyorgy, A.; Jimenez, J. I.; Yazbek, J.; Huang, H. H.; Chung, H.; Weiss, R.; Del Vecchio, D., Isocost
679 Lines Describe the Cellular Economy of Genetic Circuits. *Biophys J* **2015,** *109* (3), 639-46.
680 29. Callura, J. M.; Cantor, C. R.; Collins, J. J., Genetic switchboard for synthetic biology applications.
681 *Proc Natl Acad Sci U S A* **2012,** *109* (15), 5850-5.
682 30. Carrier, T. A.; Keasling, J. D., Investigating autocatalytic gene expression systems through
683 mechanistic modeling. *J Theor Biol* **1999,** *201* (1), 25-36.
684 31. Bintu, L.; Buchler, N. E.; Garcia, H. G.; Gerland, U.; Hwa, T.; Kondev, J.; Phillips, R., Transcriptional
685 regulation by the numbers: models. *Curr Opin Genet Dev* **2005,** *15* (2), 116-24.
686 32. Salis, H.; Tamsir, A.; Voigt, C., Engineering bacterial signals and sensors. *Contrib Microbiol* **2009,**
687 *16*, 194-225.
688 33. Daber, R.; Sochor, M. A.; Lewis, M., Thermodynamic analysis of mutant lac repressors. *J Mol Biol*
689 **2011,** *409* (1), 76-87.
690 34. Gatti-Lafranconi, P.; Dijkman, W. P.; Devenish, S. R.; Hollfelder, F., A single mutation in the core
691 domain of the lac repressor reduces leakiness. *Microb Cell Fact* **2013,** *12*, 67.
692 35. Ike, K.; Arasawa, Y.; Koizumi, S.; Mihashi, S.; Kawai-Noma, S.; Saito, K.; Umeno, D., Evolutionary
693 Design of Choline-Inducible and -Repressible T7-Based Induction Systems. *ACS Synth Biol* **2015,** *4* (12),
694 1352-60.
695 36. Mannan, A. A.; Liu, D.; Zhang, F.; Oyarzun, D. A., Fundamental Design Principles for
696 Transcription-Factor-Based Metabolite Biosensors. *ACS Synth Biol* **2017,** *6* (10), 1851-1859.
697 37. Yokobayashi, Y.; Weiss, R.; Arnold, F. H., Directed evolution of a genetic circuit. *Proc Natl Acad*
698 *Sci U S A* **2002,** *99* (26), 16587-91.
699 38. Tang, S. Y.; Fazelinia, H.; Cirino, P. C., AraC regulatory protein mutants with altered effector
700 specificity. *J Am Chem Soc* **2008,** *130* (15), 5267-71.
701 39. Tang, S. Y.; Cirino, P. C., Design and application of a mevalonate-responsive regulatory protein.
702 *Angew Chem Int Ed Engl* **2011,** *50* (5), 1084-6.
703 40. Tashiro, Y.; Fukutomi, H.; Terakubo, K.; Saito, K.; Umeno, D., A nucleoside kinase as a dual
704 selector for genetic switches and circuits. *Nucleic Acids Res* **2011,** *39* (3), e12.
705 41. Tang, S. Y.; Qian, S.; Akinterinwa, O.; Frei, C. S.; Gredell, J. A.; Cirino, P. C., Screening for
706 enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed
707 triacetic acid lactone reporter. *J Am Chem Soc* **2013,** *135* (27), 10099-103.

42. Taylor, N. D.; Garruss, A. S.; Moretti, R.; Chan, S.; Arbing, M. A.; Cascio, D.; Rogers, J. K.; Isaacs, F. J.; Kosuri, S.; Baker, D.; Fields, S.; Church, G. M.; Raman, S., Engineering an allosteric transcription factor to respond to new ligands. *Nat Methods* **2016,** *13* (2), 177-83.
43. Maranhao, A. C.; Ellington, A. D., Evolving Orthogonal Suppressor tRNAs To Incorporate Modified Amino Acids. *ACS Synth Biol* **2017,** *6* (1), 108-119.
44. Xiu, Y.; Jang, S.; Jones, J. A.; Zill, N. A.; Linhardt, R. J.; Yuan, Q.; Jung, G. Y.; Koffas, M. A. G., Naringenin-responsive riboswitch-based fluorescent biosensor module for Escherichia coli co-cultures. *Biotechnol Bioeng* **2017,** *114* (10), 2235-2244.
45. Daber, R.; Lewis, M., Towards evolving a better repressor. *Protein Eng Des Sel* **2009,** *22* (11), 673-83.
46. Kast, P., pKSS--a second-generation general purpose cloning vector for efficient positive selection of recombinant clones. *Gene* **1994,** *138* (1-2), 109-14.
47. Thyer, R.; Filipovska, A.; Rackham, O., Engineered rRNA enhances the efficiency of selenocysteine incorporation during translation. *J Am Chem Soc* **2013,** *135* (1), 2-5.
48. Takagi, M.; Nishioka, M.; Kakihara, H.; Kitabayashi, M.; Inoue, H.; Kawakami, B.; Oka, M.; Imanaka, T., Characterization of DNA polymerase from Pyrococcus sp. strain KOD1 and its application to PCR. *Appl Environ Microbiol* **1997,** *63* (11), 4504-10.
49. Gollihar, J. D., Jr. Methods in protein engineering and screening: from rational design to directed evolution and beyond Ph.D. Dissertation, University of Texas, at Austin, 2015.
50. Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D., Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nature Biotechnology* **2014,** *32* (1), 97-101.
51. Abil, Z.; Ellefson, J. W.; Gollihar, J. D.; Watkins, E.; Ellington, A. D., Compartmentalized partnered replication for the directed evolution of genetic parts and circuits. *Nat Protoc* **2017,** *12* (12), 2493-2512.
52. Khlebnikov, A.; Datsenko, K. A.; Skaug, T.; Wanner, B. L.; Keasling, J. D., Homogeneous expression of the P(BAD) promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology* **2001,** *147* (Pt 12), 3241-7.
53. Rogers, J. K.; Guzman, C. D.; Taylor, N. D.; Raman, S.; Anderson, K.; Church, G. M., Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res* **2015,** *43* (15), 7648-60.
54. Meyer, A. J.; Ellefson, J. W.; Ellington, A. D., Library generation by gene shuffling. *Current protocols in molecular biology* **2014**, 15.12. 1-15.12. 7.
55. Salis, H. M.; Mirsky, E. A.; Voigt, C. A., Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **2009,** *27* (10), 946-50.
56. Lee, J. W.; Gyorgy, A.; Cameron, D. E.; Pyenson, N.; Choi, K. R.; Way, J. C.; Silver, P. A.; Del Vecchio, D.; Collins, J. J., Creating Single-Copy Genetic Circuits. *Mol Cell* **2016,** *63* (2), 329-336.
57. Bassalo, M. C.; Garst, A. D.; Halweg-Edwards, A. L.; Grau, W. C.; Domaille, D. W.; Mutalik, V. K.; Arkin, A. P.; Gill, R. T., Rapid and Efficient One-Step Metabolic Pathway Integration in E. coli. *ACS Synth Biol* **2016,** *5* (7), 561-8.
58. West, S. C., Enzymes and molecular mechanisms of genetic recombination. *Annu Rev Biochem* **1992,** *61*, 603-40.
59. Barrick, J. E.; Lenski, R. E., Genome dynamics during experimental evolution. *Nat Rev Genet* **2013,** *14* (12), 827-39.
60. Alon, U.; Surette, M. G.; Barkai, N.; Leibler, S., Robustness in bacterial chemotaxis. *Nature* **1999,** *397* (6715), 168-71.
61. Zhang, Y.; Lara-Tejero, M.; Bewersdorf, J.; Galan, J. E., Visualization and characterization of individual type III protein secretion machines in live bacteria. *Proc Natl Acad Sci U S A* **2017,** *114* (23), 6098-6103.

62. Zelcbuch, L.; Antonovsky, N.; Bar-Even, A.; Levin-Karp, A.; Barenholz, U.; Dayagi, M.; Liebermeister, W.; Flamholz, A.; Noor, E.; Amram, S.; Brandis, A.; Bareia, T.; Yofe, I.; Jubran, H.; Milo, R., Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res* **2013,** *41* (9), e98.

63. Smanski, M. J.; Bhatia, S.; Zhao, D.; Park, Y.; L, B. A. W.; Giannoukos, G.; Ciulla, D.; Busby, M.; Calderon, J.; Nicol, R.; Gordon, D. B.; Densmore, D.; Voigt, C. A., Functional optimization of gene clusters by combinatorial design and assembly. *Nat Biotechnol* **2014,** *32* (12), 1241-9.

64. Larson, M. H.; Gilbert, L. A.; Wang, X.; Lim, W. A.; Weissman, J. S.; Qi, L. S., CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **2013,** *8* (11), 2180-96.

65. Ghodasara, A.; Voigt, C. A., Balancing gene expression without library construction via a reusable sRNA pool. *Nucleic Acids Res* **2017,** *45* (13), 8116-8127.

66. Gupta, A.; Reizman, I. M.; Reisch, C. R.; Prather, K. L., Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol* **2017,** *35* (3), 273-279.

67. Blattner, F. R.; Plunkett, G., 3rd; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J. D.; Rode, C. K.; Mayhew, G. F.; Gregor, J.; Davis, N. W.; Kirkpatrick, H. A.; Goeden, M. A.; Rose, D. J.; Mau, B.; Shao, Y., The complete genome sequence of Escherichia coli K-12. *Science* **1997,** *277* (5331), 1453-62.

68. Kittleson, J. T.; Cheung, S.; Anderson, J. C., Rapid optimization of gene dosage in E. coli using DIAL strains. *J Biol Eng* **2011,** *5*, 10.

69. Nielsen, A. A.; Der, B. S.; Shin, J.; Vaidyanathan, P.; Paralanov, V.; Strychalski, E. A.; Ross, D.; Densmore, D.; Voigt, C. A., Genetic circuit design automation. *Science* **2016,** *352* (6281), aac7341.

70. Weber, E.; Engler, C.; Gruetzner, R.; Werner, S.; Marillonnet, S., A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **2011,** *6* (2), e16765.

71. Datsenko, K. A.; Wanner, B. L., One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **2000,** *97* (12), 6640-5.

72. Thomason, L. C.; Costantino, N.; Court, D. L., E. coli genome manipulation by P1 transduction. *Curr Protoc Mol Biol* **2007,** *Chapter 1*, Unit 1 17.

73. Reetz, M. T.; Carballeira, J. D., Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc* **2007,** *2* (4), 891-903.