Published in final edited form as:
Methods. 2013 March 15; 60(1): 81–90. doi:10.1016/j.ymeth.2012.03.009.

Directed Evolution as a Powerful Synthetic Biology Tool

Ryan E. Cobb[1,2,+], **Ning Sun**[3,+], and **Huimin Zhao**[1,2,3,4,*]

[1]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[2]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[3]Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801

[4]Departments of Chemistry and Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Abstract

At the heart of synthetic biology lies the goal of rationally engineering a complete biological system to achieve a specific objective, such as bioremediation and synthesis of a valuable drug, chemical, or biofuel molecule. However, the inherent complexity of natural biological systems has heretofore precluded generalized application of this approach. Directed evolution, a process which mimics Darwinian selection on a laboratory scale, has allowed significant strides to be made in the field of synthetic biology by allowing rapid identification of desired properties from large libraries of variants. Improvement in biocatalyst activity and stability, engineering of biosynthetic pathways, tuning of functional regulatory systems and logic circuits, and development of desired complex phenotypes in industrial host organisms have all been achieved by way of directed evolution. Here, we review recent contributions of directed evolution to synthetic biology at the protein, pathway, network, and whole cell levels.

1. Introduction

Synthetic biology aims to improve and design biological systems through construction of new biological components, such as enzymes, biosynthetic pathways, genetic circuits, and cells (recently reviewed by (1–8)). Synthetic biology has a wide variety of industrial and therapeutic applications, such as creating biosensors (9), generating biofuels (10–12), producing high-quality, inexpensive drugs (13–14), and remediating polluted sites (15). In addition to the practical applications, synthetic biology can also provide valuable scientific knowledge to increase our understanding of basic life sciences. Synthetic biology lies at the interface of many different research areas such as bioinformatics, functional genomics, protein engineering, metabolic engineering, and systems biology. As the field of synthetic biology rapidly expands, synthetic biologists have encountered new challenges. For example, heterologously expressed enzymes usually suffer from poor solubility and low thermal stability; gene expression levels of a biosynthetic pathway need to be tuned to achieve higher product yield; synthetic gene circuit components require optimization to

[*]To whom correspondence should be addressed. Phone: (217) 333-2631. Fax: (217) 333-5052. zhao5@illinois.edu.
[+]These two authors contributed equally.

function in the unnatural context; and the genomic context has to be modified in order to create cells with desired phenotypes. Due to limited knowledge of complex biological systems, limited success has been achieved in solving theses problems through rational design. In contrast, directed evolution, which does not require detailed functional, structural, or mechanistic information of a biological system *a priori*, shows great advantages in addressing these challenges in synthetic biology.

Directed evolution mimics the Darwinian evolution process in the test tube. In a typical directed evolution experiment, the gene encoding a macromolecule (i.e. an RNA aptamer or a protein) of interest is randomized and expressed in a suitable host. Appropriate screening or selection methods are then used to identify mutants that have particular properties, such as binding to a specific small molecule or catalyzing a desired chemical reaction. Through iterative cycles of mutagenesis and amplification of selected mutants, beneficial mutations accumulate as in genuine Darwinian evolution but on a much shorter time scale. In this way, populations of macromolecules may be deliberately evolved toward useful synthetic and therapeutic properties (16). In this review, we will discuss some of the most recent advances and applications of directed evolution in synthetic biology to improve or redesign biological components at different levels, including the protein level, pathway level, network level, and whole cell level.

2. Protein level

2.1 Improvement of biocatalysts

Enzymes are increasingly being used as biocatalysts in the generation of products such as pharmaceuticals, fine and bulk chemicals, and biofuels. A useful biocatalyst, especially one that is suitable for application in an industrial process, needs to have high catalytic turnover, high selectivity toward specific reactions, and high stability under chemical transformation conditions (17–21). In this respect, directed evolution is an efficient and powerful tool to improve and optimize natural enzymes in order to generate robust biocatalysts for practical applications. Numerous directed evolution endeavors have been carried out (recently reviewed by (20,22)). Here we only highlight a few examples of using directed evolution to change substrate specificity, increase catalytic turnover, and improve thermal stability.

Xylitol is a pentitol which can be used not only as a sweetener but also as a platform chemical for the production of industrially important chemicals. Biosynthetic routes of producing xylitol by enzymatic reduction of D-xylose from biomass using xylose reductases (XRs) serve as safer and more environment-friendly alternatives to traditional chemical reduction methods (23–24). However, natural XRs act as promiscuous aldose reductases that can also reduce L-arabinose which is abundant in plant hemicellulose, leading to production of the unwanted byproduct L-arabinitol. Nair and Zhao successfully engineered an XR preferring D-xylose over L-arabinose through a directed evolution strategy (25). Using the fungal XR from *Neurospora crassa* as a template, a mutant library was created by error-prone PCR-based random mutagenesis and a single mutant (L109Q) was isolated with 8.9-fold preference for D-xylose over L-arabinose. The subsequent iterative rounds of targeted site-saturation mutagenesis identified additional mutations and finally yielded a mutant with 16.5-fold preference for D-xylose. The xylose specific XR provided a viable method for circumventing purification issues during biosynthesis of xylitol.

Type III polyketide synthases (PKSs) from bacteria, fungi, and plants produce an amazing array of biologically and medically important polyketides. For example, PhlD is a PKS that synthesizes phloroglucinol, which is used in industry and medicine. Similar to other enzymes involved in secondary metabolism, type III PKSs generally exhibit low productivity, such as low activity and poor stability (26), thereby hampering their

application in practical synthesis of important polyketides. Zha and coworkers successfully improved the productivity of PhlD via directed evolution (27). They identified 52 homologous PhlD genes with 79%–99% sequence identity, and created a large library of 1.2×10^{11} variants through synthetic shuffling (28). By using a high throughput screening system, two PhlD mutants were identified that increased phloroglucinol production as a result of improved activity and/or thermostability. Mutational analysis revealed that single mutations far away from the active site can significantly improve the catalysis or stability of PhlD (Figure 1).

In addition to PKSs, tailoring enzymes found in natural product biosynthetic pathways are attractive candidates for the synthesis of semisynthetic derivatives and drug libraries (29). The diversified reactions catalyzed by tailoring enzymes can convert biologically inactive precursors into pharmaceutically active molecules via regioselective and stereoselective transformations. For example, LovD is an acyltransferase found in *Aspergillus terreus* that converts the inactive monacolin J acid (MJA) into lovastatin. Using MJA and a synthetic α-dimethylbutyryl thioester as substrates, LovD can be used to synthesize the blockbuster cholesterol-lowering drug simvastatin (SV). Since it has been removed from its natural context, LovD is catalytically suboptimal as a biocatalyst and suffers from poor thermal stability. Gao and coworkers thereby employed directed evolution to improve the SV synthase activity of LovD (13). They developed an agar-based diffusion screening method to correlate the enzyme activity with a detectable phenotype. After seven rounds of screening, LovD mutants with improved catalytic efficiency, solubility, and thermal stability were obtained, with the best mutant displaying an ~11-fold increase in an *E. coli*-based biocatalytic platform. The isolated mutants can serve as potential biocatalysts for the semisynthesis of SV in the pharmaceutical industry.

2.2 Incorporation of unnatural amino acids

With few exceptions, the genetic codes of all known organisms use only the 20 canonical amino acids for protein synthesis. It is possible, however, that encoding additional amino acids and their corresponding chemical functionalities would be evolutionarily advantageous. The incorporation of unnatural amino acids (UAAs) expands the structural and chemical diversity in proteins and provides an alternative way of generating proteins with novel functions (recently reviewed by (30–33)). Adding UAAs to the genetic code of a living cell requires three modules: a codon, a tRNA that recognizes that codon, and an orthogonal aminoacyl-tRNA synthetase (aaRS) that charges that tRNA (34). The rarely used UAG amber stop codon is most commonly used to code specific UAAs. Orthogonal aaRS-tRNA pairs can be obtained from different domains of life. To alter the amino acid charging specificity of a heterologous aaRS, the active site of the aaRS is randomized by structure-based mutagenesis followed by iterative positive and negative selections (35–36). This directed evolution approach has allowed >70 UAAs with diverse structures to be incorporated into bacterial, yeast, and mammalian cells (31).

Although introduction of UAAs provides new chemical diversity and functionality, proteins containing UAAs usually show reduced activity and lower stability (37–40). Directed evolution has proved to be effective in improving proteins containing UAAs. Tirrell and colleagues have constructed a variant chloramphenicol acetyl transferase (CAT) and a green fluorescent protein (GFP) by global replacement of the leucine residues by 5,5,5-trifluoroleucine (TFL). However, the native functions of variant proteins were sacrificed after replacement. Three rounds of iterative random mutagenesis and screening generated a CAT TFL variant with similar stability to the wild type enzyme (39). Eleven rounds of error-prone PCR and fluorescence activated cell sorting (FACS) restored the fluorescent activity of TFL-replaced GFP (40). To discover and evolve novel binding proteins containing UAAs, Schultz and colleagues used a phage display strategy in combination with

X-E.coli (expressing orthogonal aaRS-tRNA pairs). They demonstrated that a 21 amino acid code (containing a genetically encoded sulfotyrosine) can confer a selective advantage in a phage-based system for the evolution of antibodies binding to gp120, the HIV coat protein (38). Later, they reported the generation of antibodies in the *Boro-X-E.coli* that encodes *p*-boronophenylalanine (BF), which has an intrinsic affinity for diols in glycans. From a mutant library generated by saturation mutagenesis, specific clones containing BF were identified from selection for binding to an acyclic glycan (37). These findings demonstrate that directed evolution is an effective tool for improving proteins containing UAAs with novel functions that cannot be obtained from Darwinian evolution in nature.

2.3 De novo protein design

While most useful biocatalysts were obtained through engineering of natural enzymes as mentioned above, the ability to *de novo* construct a functional enzyme capable of catalyzing a desired chemical reaction has great potential in a variety of applications. However, *de novo* protein design represents an overwhelming challenge. To address this challenge, the Baker group applied computational approaches to the *de novo* design and construction of proteins with particular properties (41–43). For example, Rothlisberger and coworkers used a computational enzyme design methodology to create a new enzyme catalyzing the Kemp elimination reaction, for which no naturally occurring enzyme exists (42). With the help of the RosettaMath algorithm (44), they created idealized active sites, introduced additional functional groups, and searched for protein backbone configurations. The designed enzymes had k_{cat}/K_m values in the range of 6 to 160$\mu M^{-1}\mu s^{-1}$, which was not optimized because the design process did not clearly model configurational entropy changes, longer-range second-shell interactions, and other dynamics parameters that can be important for efficient turnover. Therefore, a directed evolution strategy, without the need to model so many unknown parameters, was used to increase the catalytic activity of the designed enzyme. A total of seven rounds of random mutagenesis and DNA shuffling were carried out followed by screening in microtiter plates. Finally, several variants were successfully identified that showed an improvement of >200-fold in k_{cat}/K_m and contained 4-8 mutations. In this respect, directed evolution can serve as a valuable tool in improving the computationally designed biocatalysts. On the other hand, the improved catalysts from the directed evolution endeavor can provide valuable insight on the missing parts from the computational design and further improve the design methodology.

3. Pathway level

The synthesis of natural or unnatural products in microorganisms often involves the introduction of a metabolic pathway, which usually involves more than one heterologous gene. By coupling multiple enzymes to create a metabolic pathway in a single cell, one eliminates the need for purification of the chemical intermediates prior to the next reaction step. To achieve maximal product yield, the enzymatic activities of the pathway components need to be optimized and the expression levels of several genes of a metabolic pathway need to be regulated. In this respect, the directed evolution strategy has proven to be effective in addressing those challenges.

Biofuels synthesized from renewable resources are of increasing interest because of global energy and environmental problems. Previously, it has been demonstrated by the Liao group that the production of higher alcohols from *E. coli* can be achieved by using a 2-keto acid-based pathway (10). For example, 1-propanol and 1-butanol can be synthesized through 2-ketobutyrate, which is a degradation product of threonine. In other species, 2-ketobutyrate can be obtained through a shorter and more direct route, the citramalate pathway (45). In this pathway, (*R*)-citramalate synthesized from pyruvate and acetyl-CoA by citramalate synthase (CimA) is then converted to 2-ketobutyrate. Since the CimA from *Methanococcus*

jannaschii has low activity in *E. coli*, directed evolution was utilized to evolve the heterologous pathway (46). The selection was performed in L-isoleucine auxotrophic *E. coli*, whose growth in minimal medium can be rescued by an active citramalate pathway synthesizing 2-ketobutyrate. Thus, the growth rate of the cell in minimal medium reflected the activity of the citramalate pathway. After six rounds of mutagenesis and selection under increasing selection pressure, a CimA mutant was identified that was more active and was also insensitive to feedback inhibition. This variant enabled 9- and 22-fold higher production levels of 1-propanol and 1-butanol, respectively, compared to the strain expressing the wild-type CimA.

Similarly, based on the 2-keto acid pathway, Zhang and coworkers have expanded the natural metabolic capability of *E. coli* to synthesize a nonnatural amino acid L-homoalanine, which is an important chiral drug intermediate (47). In *E. coli*, native glutamate dehydrogenase (GDH) converts 2-ketoglutarate into L-glutamate in the presence of cofactor NADPH. To make GDH a biocatalyst producing L-homoalanine, they engineered GDH with new substrate specificity towards 2-ketobutyrate by directed evolution. Four conserved residues in the substrate binding pocket were randomized by site-saturation mutagenesis to make a library with 2 million independent clones. The selection was performed in valine auxotrophic *E. coli*, whose growth in minimal medium can be rescued by mutant GDH-catalyzed amination of the valine precursor 2-ketoisovalerate. As a result, the specificity constant k_{cat}/K_m of the best mutant towards 2-ketobutyrate is 50-fold higher than that towards the natural substrate 2-ketoglutarate. In a fermentation study, over-expression of the mutant GDH in the metabolically engineered *E. coli* strain yielded 5.4 g/L L-homoalanine from 30 g/L glucose, which is 26% of the theoretical maximum. These two examples demonstrate that directed evolution can be a powerful tool for improving the activity of certain pathway enzymes so as to maximize the production of non-growth-related products such as biofuels and pharmaceutical intermediates.

Instead of engineering proteins, Alper and coworkers showed that promoters can be engineered via directed evolution to achieve precise strengths and regulation to tune genetic control (48). In this study, a derivative of the constitutive bacteriophage P_L-γ promoter was mutated via error-prone PCR and cloned into a reporter plasmid upstream of a GFP gene. After screening in *E. coli* based on the fluorescence signal, 22 mutants spanning a wide range of GFP fluorescence were selected to form a functional promoter library. Three different orthogonal metrics were used to characterize the promoter library and ensure the consistent behavior of all its members for various genes and culture conditions. Next, they used the characterized library of promoters to evaluate the impact of phosphoenolpyruvate carboxylase (PPC) levels on cell growth yield and deoxyxylulose-5-phosphate synthase (DXS) levels on lycopene production. The multi-level characterization of promoter strengths enabled identification of optimal expression levels for PPC and DXS, which maximized the desired phenotype. In addition, they demonstrated that the promoter engineering concept could also be applied to *Saccharomyces cerevisiae*. Similarly, by screening a library of *TEF1* promoter mutants using yellow fluorescent protein (YFP) as a reporter, a promoter collection with various strengths was obtained. The applicability of this approach in both prokaryotic and eukaryotic organisms indicates that the promoter library method can be used to construct an integral platform to tune gene expression levels for synthetic biology, especially for functional genomics and metabolic engineering endeavors.

Grouping multiple, related genes into operons is a convenient means for regulating several genes simultaneously without the need for multiple promoters. To tune the expression level of several genes within operons, Pfleger and coworkers generated and screened large libraries of tunable intergenic regions (TIGRs) containing control elements that include mRNA secondary structures, RNase cleavage sites, and ribosome binding site (RBS)

sequestering sequences (49). Through an operon reporter system containing the genes encoding red fluorescent protein (RFP) and GFP, they showed that TIGRs can vary the relative expression of two reporter genes over a 100-fold range. Using the TIGR approach, they balanced the expression of three genes in an operon that encodes a heterologous mevalonate biosynthetic pathway (50), resulting in a 7-fold increase of mevalonate production in *E. coli*. This technology has great potential to optimize the expression of multiple genes in synthetic operons, in both prokaryotes and eukaryotes (Figure 2).

The inherent gene shuffling activity of the integron, which is a natural bacterial site-specific recombination system, was also used for construction and optimization of metabolic pathways (51). Integrons represent a significant proportion of bacterial genomes, which are composed of a tyrosine recombinase (the integrase, IntI), a primary recombination site (attI), and an array of gene cassettes (52). Integron recombinations represent extremely powerful evolutionary devices, whose success is exemplified by their ubiquitous spread associated with multiple antibiotic resistances. Bikard and coworkers used the unique recombination properties of the integron machinery to enable rapid and large-scale generation of combinations *in vivo* (51). As proof of principle, they constructed and optimized a functional tryptophan biosynthetic operon in *E. coli*. The *trpA-E* genes along with 'regulatory' elements were delivered as individual recombination cassettes in a synthetic integron platform. They generated thousands of genetic combinations via integrase-mediated recombination and isolated a large number of arrangements displaying various fitness and tryptophan production capacities. Several assemblages required as many as six recombination events and produced as much as 11-fold more tryptophan than the natural gene order in the same context. These results demonstrate that a synthetic integron can be used for efficient generation of functional gene combinations from a library of independent candidate gene cassettes.

4. Network level

Directed evolution of individual biosynthetic enzymes or entire pathways to perform novel, useful activities is in itself a significant achievement. However, the goal of the synthetic biologist is not only to develop these novel functionalities, but to optimize their performance in the context of a biological framework. Living systems contain very dense, complex intrinsic regulatory networks to tightly control gene expression in response to a host of external and internal stimuli (53–54). So too should synthetic pathways be tightly regulated to optimize productivity without significant detriment to the host. Ideally, one would like to have at one's command a toolbox of tunable, interchangeable regulatory elements to incorporate into a novel pathway. Though this dream has yet to reach fruition, significant advances have been made in recent years through the use of directed evolution to develop regulatory elements suitable for application in the field of synthetic biology.

Yokobayashi and coworkers acknowledged the difficulty of rationally designing a genetic circuit and demonstrated the versatility of directed evolution to overcome this challenge (55). Their circuit consisted of two inhibitor/promoter pairs: LacI/P_{lac} and CI/λP_{RO12}. In their system, LacI was constitutively expressed, and the *cI* gene was placed under the control of the P_{lac} promoter. In the presence of IPTG, LacI can no longer bind to P_{lac}, and CI is transcribed and translated. This constitutes an IMPLIES gate with LacI and IPTG as its inputs and CI as the output. On a compatible plasmid, an enhanced YFP (EYFP) indicator was placed under the control of the λP_{RO12} promoter. This constitutes a logical inverter, where high CI levels result in low EYFP, and low CI levels result in high EYFP. When first constructed, however, no EYFP was observed at any IPTG concentration. This indicated that the two circuit elements (IMPLIES gate and inverter) were not properly tuned, likely due to leaky CI expression constitutively blocking EYFP expression. To rectify the two

components, directed evolution was carried out on the *cI* gene and its RBS. After two generations, several functional circuits were obtained with varying characteristics. DNA sequencing revealed mutations affecting either binding of CI to DNA or CI-CI dimerization, as well as mutations upstream of *cI*. This demonstrates the versatility of directed evolution for fine-tuning of genetic circuits, and introduces a platform for studying a variety of different circuit elements.

One of the more well-studied regulatory systems utilized in synthetic biology research is the quorum-sensing *lux* operon found in the marine bacterium *Vibrio fischeri*. In nature, *V. fischeri* can exist either as a free-floating bacterioplankton or as a bioluminescent endosymbiont of certain marine animals. Through the action of the LuxI protein, the diffusible small molecule 3-oxo-hexanoyl homoserine lactone (3OC6HSL) is produced and secreted. In the open sea environment, 3OC6HSL concentrations remain low, and transcription of the bioluminescence operon *luxICDABE* remains low. In the light-emitting organ of a host organism, however, 3OC6HSL concentrations rise significantly. This promotes binding to the LuxR transcriptional regulator, activating bioluminescence.

Through the application of directed evolution, the Arnold and Leadbetter groups have been able generate LuxR variants with altered small molecule specificities. They initially sought to broaden the substrate specificity of LuxR to respond to alternate acyl-homoserine lactones using a two plasmid screening method in *E. coli* (56). The first plasmid contained a LuxR variant generated by error-prone PCR, while the second contained a GFP variant under the control of the P_{luxI} promoter (Figure 3a). Approximately 20,000 colonies were screened in the presence of 200 nM octanoyl-homoserine lactone (C8HSL), leading to the identification of 6 LuxR variants. An additional round of DNA shuffling yielded 8 LuxR mutants that not only exhibited a 100-fold increase in sensitivity to C8HSL when compared to wild type LuxR, but also showed increased sensitivity to pentanoyl-HSL and tetradecanoyl-HSL. To further expand the substrate specificity of LuxR to smaller ligands, an additional cycle of error-prone PCR followed by DNA shuffling was employed (57). In this study, butanoyl-HSL was incorporated into the screening process, resulting in a mutant responsive to concentrations as low as 10 nM. For ideal application in a synthetic biology system, however, a transcription regulator should not only respond to an alternate small molecule input, but also be specific for that small molecule. To achieve this goal, a dual selection system was employed (58). “ON” selection in the presence of the desired ligand (decanoyl-HSL) utilized a chloramphenicol gene under the control of the P_{luxI} promoter, while “OFF” selection placed a β-lactamase inhibitory gene under P_{luxI} control, fostering carbenicillin resistance only with LuxR activation (Figure 3a). Three rounds of ON/OFF selection resulted in LuxR mutants that retained responsiveness to acyl-HSL ligands, but not to the wild type 3OC6HSL ligand.

To engineer the other component of this regulatory system, Kambam and coworkers have adapted the two-plasmid selection system to identify LuxI mutants with increased 3OC6HSL production (59). Both a mutant LuxI and wild type LuxR are placed on the first plasmid, while the second plasmid contains an ampicillin resistance gene under the control of the P_{luxI} promoter. Selection proceeds at high ampicillin concentrations, such that the quantity of 3OC6HSL produced by the wild type LuxI is not sufficient to produce enough β-lactamase to rescue growth. Using error-prone PCR, they identified LuxI mutants with 80-fold higher production of 3OC6HSL. As the authors note, this selection method could also be applied to the identification of LuxI mutants with altered substrate specificities by replacing the wild type LuxR with a LuxR mutant responsive only to an alternate ligand.

The application of directed evolution to network-level synthetic biology in recent years expands beyond just the LuxR/LuxI system. For example, Chockalingam and coworkers

engineered the human estrogen receptor α to respond to an orthogonal ligand, 4,4'-dihydroxybenzil (60). They performed site saturation mutagenesis at several amino acid positions in close contact with the bound wild type ligand, as well as error-prone PCR. Using a yeast two-hybrid assay coupling ligand-receptor affinity to growth on media lacking histidine, they were able to identify a mutant with ligand specificity altered by $>10^7$-fold. In a follow-up study, McLachlan and coworkers demonstrated the versatility of this approach by not only further improving the specificity of their previous mutant toward DHB, but also engineering a new mutant from the same human estrogen receptor α scaffold specific to a different orthogonal ligand, 2,4-di(4-hydroxyphenyl)-5-ethylthiazole (L9) (61). After applying a combination of site saturation mutagenesis and error-prone PCR, the final mutants both displayed an increase in specificity toward their orthogonal ligand of $>2 \times 10^9$-fold. To illustrate the utility of these engineered ligand-receptor pairs for synthetic biology applications, functional AND and NOR gates were constructed in HeLa cells.

Beyond the protein level of gene regulation, directed evolution has also been applied at the RNA level to generate ligand-dependent riboswitches. The Gallivan group has developed multiple methods to rapidly select such riboswitches. In one method, an aptamer sequence is placed 4 to 8 bp upstream of the RBS for a β-galactosidase gene (62). In the absence of ligand, the β-galactosidase gene should be inactive, and so white colonies are selected and re-screened in the presence of ligand for blue coloration. Using this method, a riboswitch with a 36-fold increase in protein expression in the presence of ligand was identified with a very low uninduced expression level. An alternate, more economical screening method was developed based on cell motility (63). Here, the RNA aptamer is placed upstream of the *cheZ* gene in *E. coli*, which is necessary for cell migration. In the absence of ligand, expression of CheZ in cells with functional riboswitches is low; thus, these cells tumble incessantly and do not migrate. As a result, cells that do not migrate on semisolid media are selected for subsequent rounds. Alternatively, in the presence of ligand, cells that have increased CheZ expression, and thus migrate through semisolid media, are selected.

Dual selection systems are particularly effective in engineering gene switches, as they allow for selection of both the desired ON and OFF states. As a result, significant effort has been directed toward the development of such methods. One such method utilizes the single-gene selection marker *tetA*, a tetracycline/H^+ antiporter. When expressed, TetA confers resistance to tetracycline; however, it also renders cells more susceptible to toxic salts, such as $NiCl_2$. Muranaka and coworkers have adapted this system for high-efficiency enrichment of synthetic riboswitches (64). In the ON state, cells are selected with tetracycline resistance; in the OFF state, cells are selected with $NiCl_2$ resistance. To incorporate a quantitative handle into this assay, a GFP variant was fused to the TetA protein. This not only allowed for the rapid determination of the best riboswitches (based on GFP expression level with and without induction), but also eliminated the need for time-consuming and laborious subcloning of reporter genes into the many selected constructs. Alternatively, Tashiro and coworkers developed a different single-gene dual selection system based on Herpes simplex virus thymidine kinase (hsvTK) (65). In this system, ON selection is carried out in the presence of 2'-deoxy-5-fluorouridine (5FdU). 5FdU, when phosphorylated, inhibits *de novo* synthesis of dTMP; as a result, cells must synthesize dTMP from exogenous dT. Consequently, only cells expressing the selection gene hsvTK can grow in 5FdU-containing media. OFF selection occurs in the presence of 6-(β-D-2-deoxyribofuranosyl)-3,4-dihydro-8*H*-pyrimido[4,5-c][1,2]oxazin-7-one (dP), a synthetic nucleoside. When phosphorylated by hsvTK, dP can be incorporated into the genome of the cell, resulting in an increased mutation rate and decreased viability (Figure 3b). Thus, cells that do not effectively turn OFF the riboswitch will not be selected. Applying this method, they were able to achieve up to 32,000-fold enrichment of a functional gene circuit (consisting of an

AND gate, with LuxR and acyl-homoserine lactone as inputs and CI as output, and an inverter, with CI as input and GFP concentration as output) from non-functional variants.

As designed genetic circuits become more and more complex, directed evolution has proven to be a powerful, critical tool for tuning desired functionality. Coupling a NOR gate to a memory module via directed evolution, Lou and coworkers were able to construct a functional push-on push-off switch (Figure 4) (66). In their design, the memory module consists of mutually repressed repressors CI and CI434, each co-expressed with a fluorescent reporter gene. Each repressor not only inhibits the other repressor, but also activates its own expression, creating a bistable switch. The NOR gate consists of two protein inputs, LacI (co-cistronic with CI) and LexA (endogenously produced), both of which inhibit the output promoter P_{NOR} . A third repressor protein, CI_{ind-}, is placed under the control of the P_{NOR} promoter. CI_{ind-} behaves identically to CI, except that it is insensitive to the RecA* protease. Finally, UV radiation serves as an external input to the system by activating RecA*, which cleaves LexA, CI, and CI434. This clears the memory of the switch, allowing it accept a new state. To tune this gene circuit, the authors used directed evolution on the RBSs of both LacI and CI_{ind-}, the proteins that link the two circuit elements. After saturation mutagenesis at multiple key positions in both RBSs and two rounds of selection, three mutants were identified that could switch from green to red with high efficiency after UV stimulation, and could switch back from red to green (although at significantly lower efficiency) with a second UV dose.

5. Whole cell level

Although individual synthetic biology efforts may focus only on specific aspects of a biological system (i.e., a particular enzyme, pathway, or regulatory network), ultimately, all such elements function in the context of a complete, living cell. As a result, desired phenotypes can be influenced significantly by factors not directly related to the element of interest. To embrace this added level of complexity, numerous efforts have begun to focus on engineering at the level of the biological system itself. Rational design of a complete synthetic organism remains a distant goal; nevertheless, directed evolution has allowed for significant advances in this direction.

One of the powerful techniques for whole cell level synthetic biology is global transcription machinery engineering, or gTME, pioneered by the Stephanopoulos group (67–68). The gTME method seeks to elicit complex phenotypes by introducing diversity in proteins that regulate the transcriptome, rather than focusing just on the protein or pathway of interest (Figure 5). This allows for global alterations in behavior that may be otherwise inaccessible. One example of the gTME approach applied to *S. cerevisiae* led to the creation of a strain with improvements in both ethanol tolerance and production (69). In this study, the TATA-binding protein SPT15 and the TATA-binding protein-associated factor TAF25 were separately targeted for mutation by error-prone PCR. In both cases, the mutant protein was maintained on a plasmid while the wild type allele remained on the chromosome, allowing for identification of dominant mutations. After selection in “stressed” conditions (100 g/liter glucose, 5–6% ethanol), the mutant with the best improvement in growth yield was found in the SPT15 library and contain three amino acid substitutions. This mutant, called spt15–300, was the focus of subsequent analysis. Statistically significant improvements in cell viability were observed in this strain when compared to the wild type parent strain at ethanol concentrations of up to 20%. Transcriptional profiling under “unstressed” conditions (20 g/liter glucose, 0% ethanol), chosen to allow for similar growth rates between the mutant and parent strains, revealed differences in the expression levels of hundreds of genes from a variety of pathways and networks. Twelve of the genes with the greatest over-expression in the mutant strain were chosen for a loss-of-phenotype assay. In all but one of these cases,

deletion of the over-expressed gene resulted in a significant reduction or complete loss in the improved tolerance of the mutant strain. Further, individual over-expression in the parent strain of some of these genes could not induce the improved tolerance phenotype. These results suggest the existence of a complex, interconnected network requiring each of these components, which would be otherwise inaccessible at a more focused level of analysis.

Since this study, gTME has been applied to elicit a variety of complex phenotypes. For example, error-prone PCR-based mutagenesis of *SPT15* has also been used to generate an *S. cerevisiae* strain that can more efficiently ferment xylose and xylose-glucose mixtures (70). In *E. coli*, gTME has been applied to improve ethanol tolerance, increase lycopene production, and to simultaneously improve both ethanol and SDS tolerance (71). Here, the target of mutagenesis was *rpoD*, the gene encoding the main sigma factor σ^{70}. Selection for ethanol tolerance alone was carried out first in 50 g/L ethanol, followed by two additional rounds of mutation and selection in 60 g/L and 70 g/L ethanol. The resulting strain exhibited a higher growth rate and greater minimum inhibitory concentration than ethanol-tolerant strains identified through traditional strain improvement techniques. Improved lycopene production was identified via colorimetric screening of *rpoD* libraries in four different genetic backgrounds (a parent strain containing the lycopene biosynthetic pathway and three strains obtained through other improvement methods). In all four cases, application of gTME led to further strain improvement. Finally, to elicit a combined ethanol/SDS tolerance phenotype, four strategies were employed: isolation of an ethanol-tolerant mutant, followed by further selection for SDS tolerance; isolation of an SDS-tolerant mutant, followed by further selection for ethanol tolerance; simultaneous selection for SDS tolerance and ethanol tolerance; and independent selection of ethanol tolerance and SDS tolerance, followed by co-expression of the mutant sigma factors. Interestingly, the last strategy (independent selection followed by co-expression) proved to be the best, enabling significant tolerance to both ethanol and SDS without the tradeoff observed in either the sequential or simultaneous screening strategies. It is noteworthy that one of the two sigma factors identified was full-length, while the other was truncated, further illustrating the versatility of this strategy for eliciting complex phenotypes.

Another powerful technique that has seen significant application in whole cell level synthetic biology is genome shuffling. Genome shuffling takes advantage of whole-genome recombination between members of a diverse population to create complex, multi-parent progeny with significantly improved fitness in every successive generation (72). A variety of industrially relevant phenotypes have been pursued using a genome shuffling approach. Hida and coworkers, for example, sought to increase production of (2*S*,3*R*)-hydroxycitric acid (HCA) in the native producer *Streptomyces* sp. U121, the biosynthetic pathway of which is not well characterized (73). Typically, genome shuffling experiments proceed in two major steps: creation/identification of the parent strains to be shuffled, and subsequent recombination (via protoplast fusion) and selection to identify the desired progeny (Figure 5). In this study, the parent strains were created through random mutagenesis using nitrosoguanidine (NTG) and selected on media containing *trans*-epoxyaconitic acid (EAA), an antibiotic analog of HCA. Those mutants that showed increased resistance to EAA were thought to have either increased export of EAA (and thus of HCA as well) or increased ability to hydrate EAA, yielding the desired HCA product. From this initial round of mutagenesis, four strains were identified with increased HCA production (confirmed in liquid culture) and used for genome shuffling. To facilitate recombination, protoplasts of each of the parental strains were formed and mixed in equal proportion. One round of shuffling led to the identification of mutants with a small increase in HCA production, while a second round showed minimal improvement. Addition of EAA to the protoplast regeneration medium, however, led to a significant increase in HCA production, yielding four mutants with five- to six-fold improvement in product titer.

Genome shuffling is a very versatile technique that can be applied in a variety of different hosts. For example, application in the strict anaerobe *Clostridium diolis* DSM 15410 led to improved 1,3-propanediol (1,3-PD) production (74). In this study, NTG-based chemical mutagenesis was employed to generate two sets of parental strains: those improved in glycerol (substrate) tolerance, and those improved in 1,3-PD (product) tolerance. Four rounds of genome shuffling in the presence of 1,3-PD and the side products acetic and butyric acid were carried out, yielding a strain with 80% greater volumetric production of 1,3-PD. In *S. cerevisiae*, genome shuffling was applied to simultaneously improve thermotolerance and ethanol tolerance (75). UV irradiation was used to generate the initial mutant library, from which two strains were identified that could grow at 43 °C on plates containing 15% v/v ethanol. These two mutants, which also exhibited improved ethanol production, served as the parental strains for the first round of protoplast fusion. Following two subsequent rounds, a final strain was identified that could grow at up to 55 °C and tolerate 25% v/v ethanol. At 45 °C, this strain could completely consume 20% w/v glucose, producing 9.95% w/v ethanol. Further application of genome shuffling in *Propionibacterium shermanii* and *Penicillium decumbens* JU-A10 have been recently explored as well, resulting in mutants with increased production of vitamin B12 and cellulase, respectively (76-77).

In light of the persistently decreasing cost of oligonucleotide synthesis, the Church group has recently developed a powerful tool for large-scale, parallel modification of a genome at many sites (78). Termed multiplex automated genome engineering, or MAGE, this technique makes use of the bacteriophage λ-Red ssDNA-binding protein β to achieve allelic replacement of desired gene loci with synthetic oligos. One cycle of MAGE includes transformation of cells with the synthetic oligos, recovery and growth to mid-log phase, and heat shock to induce β protein expression and recombination. With successive cycles, increasing variation can be introduced into the host genome (Figure 5). Depending on the oligo design, this technique can be used to introduce mutations, insertions, or deletions, all in a simultaneous fashion. As proof of concept, 24 different genes were simultaneously targeted to induce a lycopene overproduction phenotype. Twenty of these genes had been reported to increase lycopene production, and so oligos with degenerate RBS sequences were designed to enhance translation efficiency. The remaining four genes came from alternate pathways and were targeted for inactivation via oligos containing nonsense mutations. After 5–35 cycles of MAGE, up to five-fold greater lycopene production was observed when compared to the ancestral strain, yielding titers of ~9000 ppm. As this method requires prior knowledge of the specific gene targets for modification, it serves as a unique complement to the techniques previously described.

6. Conclusion and perspectives

Synthetic biology has progressed significantly since the term was first introduced by Waclaw Szybalski in the 1970s, and reemphasized at the turn of this century (79–80). In recent years, significant advances have been made in the field, including the identification of the approximate minimal genome required for life and the complete synthesis of an entire functional genome (81–82). In light of such achievements, it seems that the goal of a complete designer organism is not far from fruition. Nevertheless, the fact remains that in most cases, our current level of knowledge is still insufficient to design an organism displaying a desired phenotype through strictly rational means. As a result, directed evolution has been continually called upon as a powerful tool to bridge the gap between our current knowledge and our desired destination. Whether it is through optimization of a particular enzyme or biosynthetic pathway, tuning of a complex regulatory circuit for practical application, or even genome-scale development of an ideal host organism, directed

evolution has found numerous applications at all levels of biological complexity and will likely continue to do so for many years to come.

Acknowledgments

We thank the National Institutes of Health (GM077596), the National Academies Keck Futures Initiative on Synthetic Biology, the Biotechnology Research and Development Consortium (BRDC) (Project 2-4-121), the British Petroleum Energy Biosciences Institute, and the National Science Foundation as part of the Center for Enabling New Technologies through Catalysis (CENTC), CHE-0650456, for financial support in our synthetic biology projects. R. Cobb also acknowledges support from the US National Institutes of Health under Ruth L. Kirschstein National Research Award 5 T32 GM070421 from the National Institute of General Medical Sciences.

References

1. Yeh BJ, Lim WA. Synthetic biology: lessons from the history of synthetic organic chemistry. Nat Chem Biol. 2007; 3:521–525. [PubMed: 17710092]
2. Keasling JD. Synthetic biology for synthetic chemistry. ACS Chem Biol. 2008; 3:64–76. [PubMed: 18205292]
3. Mukherji S, van Oudenaarden A. Synthetic biology: understanding biological design from synthetic circuits. Nat Rev Genet. 2009; 10:859–871. [PubMed: 19898500]
4. Purnick PE, Weiss R. The second wave of synthetic biology: from modules to systems. Nat Rev Mol Cell Biol. 2009; 10:410–422. [PubMed: 19461664]
5. Khalil AS, Collins JJ. Synthetic biology: applications come of age. Nat Rev Genet. 2010; 11:367–379. [PubMed: 20395970]
6. Liang J, Luo Y, Zhao H. Synthetic biology: putting synthesis into biology. Wiley Interdiscip Rev Syst Biol Med. 2010; 3:7–20. [PubMed: 21064036]
7. Neumann H, Neumann-Staubitz P. Synthetic biology approaches in drug discovery and pharmaceutical biotechnology. Appl Microbiol Biotechnol. 2010; 87:75–86. [PubMed: 20396881]
8. Young E, Alper H. Synthetic biology: tools to design, build, and optimize cellular processes. J Biomed Biotechnol. 2010; 2010:130781. [PubMed: 20150964]
9. Rajendran M, Ellington AD. Selection of fluorescent aptamer beacons that light up in the presence of zinc. Anal Bioanal Chem. 2008; 390:1067–1075. [PubMed: 18049815]
10. Atsumi S, Hanai T, Liao JC. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature. 2008; 451:86–89. [PubMed: 18172501]
11. Steen EJ, Chan R, Prasad N, Myers S, Petzold CJ, Redding A, Ouellet M, Keasling JD. Metabolic engineering of Saccharomyces cerevisiae for the production of n-butanol. Microb Cell Fact. 2008; 7:36. [PubMed: 19055772]
12. Waks Z, Silver PA. Engineering a synthetic dual-organism system for hydrogen production. Appl Environ Microbiol. 2009; 75:1867–1875. [PubMed: 19201964]
13. Gao X, Xie X, Pashkov I, Sawaya MR, Laidman J, Zhang W, Cacho R, Yeates TO, Tang Y. Directed evolution and structural characterization of a simvastatin synthase. Chem Biol. 2009; 16:1064–1074. [PubMed: 19875080]
14. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 2006; 440:940–943. [PubMed: 16612385]
15. Gilbert ES, Walker AW, Keasling JD. A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion. Appl Microbiol Biotechnol. 2003; 61:77–81. [PubMed: 12658518]
16. Jackel C, Kast P, Hilvert D. Protein design by directed evolution. Annu Rev Biophys. 2008; 37:153–173. [PubMed: 18573077]
17. Fox RJ, Clay MD. Catalytic effectiveness, a measure of enzyme proficiency for industrial applications. Trends Biotechnol. 2009; 27:137–140. [PubMed: 19193465]
18. Panke S, Held M, Wubbolts M. Trends and innovations in industrial biocatalysis for the production of fine chemicals. Curr Opin Biotechnol. 2004; 15:272–279. [PubMed: 15357999]

19. Pollard DJ, Woodley JM. Biocatalysis for pharmaceutical intermediates: the future is now. Trends Biotechnol. 2007; 25:66–73. [PubMed: 17184862]
20. Turner NJ. Directed evolution drives the next generation of biocatalysts. Nat Chem Biol. 2009; 5:567–573. [PubMed: 19620998]
21. Wen F, Nair NU, Zhao H. Protein engineering in designing tailored enzymes and microorganisms for biofuels production. Curr Opin Biotechnol. 2009; 20:412–419. [PubMed: 19660930]
22. Dougherty MJ, Arnold FH. Directed evolution: new parts and optimized function. Curr Opin Biotechnol. 2009; 20:486–491. [PubMed: 19720520]
23. Granstrom TB, Izumori K, Leisola M. A rare sugar xylitol. Part II: biotechnological production and future applications of xylitol. Appl Microbiol Biotechnol. 2007; 74 :273–276. [PubMed: 17216458]
24. Granstrom TB, Izumori K, Leisola M. A rare sugar xylitol. Part I: the biochemistry and biosynthesis of xylitol. Appl Microbiol Biotechnol. 2007; 74:277–281. [PubMed: 17216457]
25. Nair NU, Zhao H. Evolution in reverse: engineering a D-xylose-specific xylose reductase. Chembiochem. 2008; 9:1213–1215. [PubMed: 18383056]
26. Watanabe K, Praseuth AP, Wang CC. A comprehensive and engaging overview of the type III family of polyketide synthases. Curr Opin Chem Biol. 2007; 11:279–286. [PubMed: 17482864]
27. Zha W, Rubin-Pitel SB, Zhao H. Exploiting genetic diversity by directed evolution: molecular breeding of type III polyketide synthases improves productivity. Mol Biosyst. 2008; 4:246–248. [PubMed: 18437267]
28. Ness JE, Kim S, Gottman A, Pak R, Krebber A, Borchert TV, Govindarajan S, Mundorff EC, Minshull J. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. Nat Biotechnol. 2002; 20:1251–1255. [PubMed: 12426575]
29. Zhou H, Xie X, Tang Y. Engineering natural products using combinatorial biosynthesis and biocatalysis. Curr Opin Biotechnol. 2008; 19:590–596. [PubMed: 18992814]
30. Brustad EM, Arnold FH. Optimizing non-natural protein function with directed evolution. Curr Opin Chem Biol. 2010
31. Liu CC, Schultz PG. Adding new chemistries to the genetic code. Annu Rev Biochem. 2010; 79:413–444. [PubMed: 20307192]
32. Voloshchuk N, Montclare JK. Incorporation of unnatural amino acids for synthetic biology. Mol Biosyst. 2010; 6:65–80. [PubMed: 20024068]
33. Wang Q, Parrish AR, Wang L. Expanding the genetic code for biological studies. Chem Biol. 2009; 16:323–336. [PubMed: 19318213]
34. Filipovska A, Rackham O. Building a Parallel Metabolism within the Cell. ACS Chem Biol. 2008; 3:51–63. [PubMed: 18173237]
35. Santoro SW, Wang L, Herberich B, King DS, Schultz PG. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat Biotechnol. 2002; 20:1044–1048. [PubMed: 12244330]
36. Melancon CE 3rd, Schultz PG. One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases. Bioorg Med Chem Lett. 2009; 19:3845–3847. [PubMed: 19398201]
37. Liu CC, Mack AV, Brustad EM, Mills JH, Groff D, Smider VV, Schultz PG. Evolution of proteins with genetically encoded "chemical warheads". J Am Chem Soc. 2009; 131:9616–9617. [PubMed: 19555063]
38. Liu CC, Mack AV, Tsao ML, Mills JH, Lee HS, Choe H, Farzan M, Schultz PG, Smider VV. Protein evolution with an expanded genetic code. Proc Natl Acad Sci U S A. 2008; 105:17688–17693. [PubMed: 19004806]
39. Montclare JK, Tirrell DA. Evolving proteins of novel composition. Angew Chem Int Ed Engl. 2006; 45:4518–4521. [PubMed: 16763955]
40. Yoo TH, Link AJ, Tirrell DA. Evolution of a fluorinated green fluorescent protein. Proc Natl Acad Sci U S A. 2007; 104:13887–13890. [PubMed: 17717085]
41. Jiang L, Althoff EA, Clemente FR, Doyle L, Rothlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF 3rd, et al. De novo computational design of retro-aldol enzymes. Science. 2008; 319:1387–1391. [PubMed: 18323453]

42. Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, et al. Kemp elimination catalysts by computational enzyme design. Nature. 2008; 453:190–195. [PubMed: 18354394]
43. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, St Clair JL, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL, et al. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. Science. 2010; 329:309–313. [PubMed: 20647463]
44. Zanghellini A, Jiang L, Wollacott AM, Cheng G, Meiler J, Althoff EA, Rothlisberger D, Baker D. New algorithms and an in silico benchmark for computational enzyme design. Protein Sci. 2006; 15:2785–2794. [PubMed: 17132862]
45. Howell DM, Xu H, White RH. (R)-citramalate synthase in methanogenic archaea. J Bacteriol. 1999; 181:331–333. [PubMed: 9864346]
46. Atsumi S, Liao JC. Directed evolution of Methanococcus jannaschii citramalate synthase for biosynthesis of 1-propanol and 1-butanol by Escherichia coli. Appl Environ Microbiol. 2008; 74:7802–7808. [PubMed: 18952866]
47. Zhang K, Li H, Cho KM, Liao JC. Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. Proc Natl Acad Sci U S A. 2010; 107:6234–6239. [PubMed: 20332210]
48. Alper H, Fischer C, Nevoigt E, Stephanopoulos G. Tuning genetic control through promoter engineering. Proc Natl Acad Sci U S A. 2005; 102:12678–12683. [PubMed: 16123130]
49. Pfleger BF, Pitera DJ, Smolke CD, Keasling JD. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat Biotechnol. 2006; 24:1027–1032. [PubMed: 16845378]
50. Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat Biotechnol. 2003; 21:796–802. [PubMed: 12778056]
51. Bikard D, Julie-Galau S, Cambray G, Mazel D. The synthetic integron: an in vivo genetic shuffling device. Nucleic Acids Res. 2010; 38:e153. [PubMed: 20534632]
52. Mazel D. Integrons: agents of bacterial evolution. Nat Rev Microbiol. 2006; 4:608–620. [PubMed: 16845431]
53. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, et al. Transcriptional regulatory networks in Saccharomyces cerevisiae. Science (New York, NY). 2002; 298:799–804.
54. Gama-Castro S, Salgado H, Peralta-Gil M, Santos-Zavaleta A, Muniz-Rascado L, Solano-Lira H, Jimenez-Jacinto V, Weiss V, Garcia-Sotelo JS, Lopez-Fuentes A, et al. RegulonDB version 7.0: transcriptional regulation of Escherichia coli K-12 integrated within genetic sensory response units (Gensor Units). Nucleic acids research. 2011; 39 :D98–D105. [PubMed: 21051347]
55. Yokobayashi Y, Weiss R, Arnold FH. Directed evolution of a genetic circuit. Proc Natl Acad Sci U S A. 2002; 99:16587–16591. [PubMed: 12451174]
56. Collins CH, Arnold FH, Leadbetter JR. Directed evolution of Vibrio fischeri LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. Mol Microbiol. 2005; 55:712–723. [PubMed: 15660998]
57. Hawkins AC, Arnold FH, Stuermer R, Hauer B, Leadbetter JR. Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. Appl Environ Microbiol. 2007; 73:5775–5781. [PubMed: 17675429]
58. Collins CH, Leadbetter JR, Arnold FH. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. Nat Biotechnol. 2006; 24:708–712. [PubMed: 16715074]
59. Kambam PK, Sayut DJ, Niu Y, Eriksen DT, Sun L. Directed evolution of LuxI for enhanced OHHL production. Biotechnol Bioeng. 2008; 101:263–272. [PubMed: 18428113]
60. Chockalingam K, Zhao H. Creating new specific ligand-receptor pairs for transgene regulation. Trends Biotechnol. 2005; 23:333–335. [PubMed: 15978316]
61. McLachlan MJ, Chockalingam K, Lai KC, Zhao H. Directed evolution of orthogonal ligand specificity in a single scaffold. Angewandte Chemie (International ed. 2009; 48:7783–7786.

62. Lynch SA, Desai SK, Sajja HK, Gallivan JP. A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function. Chemistry & biology. 2007; 14:173–184. [PubMed: 17317571]
63. Topp S, Gallivan JP. Random walks to synthetic riboswitches--a high-throughput selection based on cell motility. Chembiochem. 2008; 9:210–213. [PubMed: 18098254]
64. Muranaka N, Sharma V, Nomura Y, Yokobayashi Y. An efficient platform for genetic selection and screening of gene switches in Escherichia coli. Nucleic Acids Res. 2009; 37:e39. [PubMed: 19190095]
65. Tashiro Y, Fukutomi H, Terakubo K, Saito K, Umeno D. A nucleoside kinase as a dual selector for genetic switches and circuits. Nucleic acids research. 2011; 39:e12. [PubMed: 21062820]
66. Lou C, Liu X, Ni M, Huang Y, Huang Q, Huang L, Jiang L, Lu D, Wang M, Liu C, et al. Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. Mol Syst Biol. 2010; 6:350. [PubMed: 20212522]
67. Lam FH, Hartner FS, Fink GR, Stephanopoulos G. Enhancing stress resistance and production phenotypes through transcriptome engineering. Methods Enzymol. 2010; 470:509–532. [PubMed: 20946823]
68. Tyo KE, Alper HS, Stephanopoulos GN. Expanding the metabolic engineering toolbox: more options to engineer cells. Trends Biotechnol. 2007; 25:132–137. [PubMed: 17254656]
69. Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G. Engineering yeast transcription machinery for improved ethanol tolerance and production. Science. 2006; 314:1565–1568. [PubMed: 17158319]
70. Liu H, Yan M, Lai C, Xu L, Ouyang P. gTME for improved xylose fermentation of Saccharomyces cerevisiae. Appl Biochem Biotechnol. 2010; 160:574–582. [PubMed: 19067246]
71. Alper H, Stephanopoulos G. Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab Eng. 2007; 9:258–267. [PubMed: 17292651]
72. Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WP, del Cardayre SB. Genome shuffling leads to rapid phenotypic improvement in bacteria. Nature. 2002; 415:644–646. [PubMed: 11832946]
73. Hida H, Yamada T, Yamada Y. Genome shuffling of Streptomyces sp. U121 for improved production of hydroxycitric acid. Applied microbiology and biotechnology. 2007; 73:1387–1393. [PubMed: 17043823]
74. Otte B, Grunwaldt E, Mahmoud O, Jennewein S. Genome shuffling in Clostridium diolis DSM 15410 for improved 1,3-propanediol production. Applied and environmental microbiology. 2009; 75:7610–7616. [PubMed: 19854917]
75. Shi DJ, Wang CL, Wang KM. Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of Saccharomyces cerevisiae. J Ind Microbiol Biotechnol. 2009; 36:139–147. [PubMed: 18846398]
76. Zhang Y, Liu JZ, Huang JS, Mao ZW. Genome shuffling of Propionibacterium shermanii for improving vitamin B12 production and comparative proteome analysis. Journal of biotechnology. 2010; 148:139–143. [PubMed: 20553774]
77. Cheng Y, Song X, Qin Y, Qu Y. Genome shuffling improves production of cellulase by Penicillium decumbens JU-A10. Journal of applied microbiology. 2009; 107:1837–1846. [PubMed: 19486209]
78. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009; 460:894–898. [PubMed: 19633652]
79. Szybalski W, Skalka A. Nobel prizes and restriction enzymes. Gene. 1978; 4:181–182. [PubMed: 744485]
80. Rawls RL. 'Synthetic biology' makes its debut. Chem Eng News. 2000; 78:49–53.
81. Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC. Essential genes of a minimal bacterium. Proc Natl Acad Sci U S A. 2006; 103:425–430. [PubMed: 16407165]
82. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Science (New York, NY). 2010; 329:52–56.

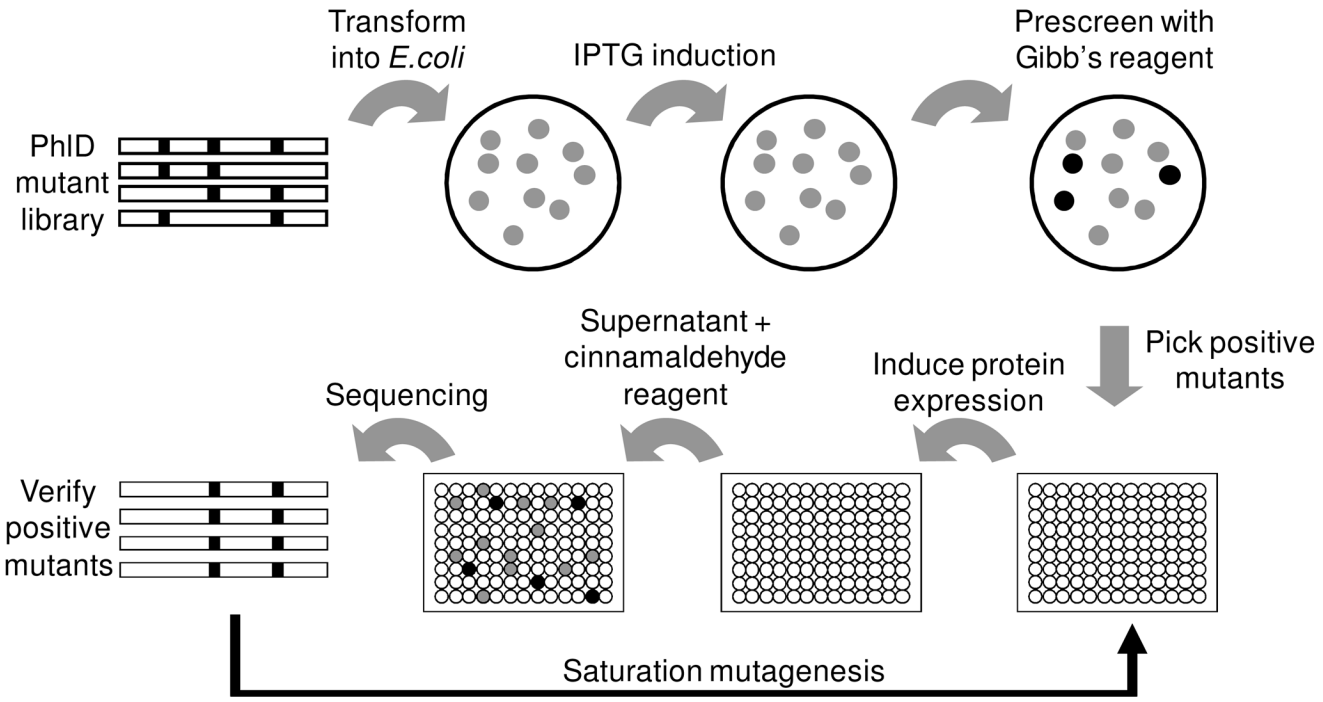

Figure 1.
Identification of PhlD mutants with improved phloroglucinol production. A PhlD mutant library was constructed through synthetic shuffling. After transformation of the library, membrane-based pre-screening was used to rapidly isolate active clones in a high throughput manner. All active mutants were then picked directly from the membrane into 96-well plates for more quantitative analysis. Phloroglucinol production in cell-free supernatant was quantified using the colorimetric reaction between cinnamaldehyde and phloroglucinol. After characterization of positive mutants, saturation mutagenesis was carried out to evaluate the effect of individual mutations.

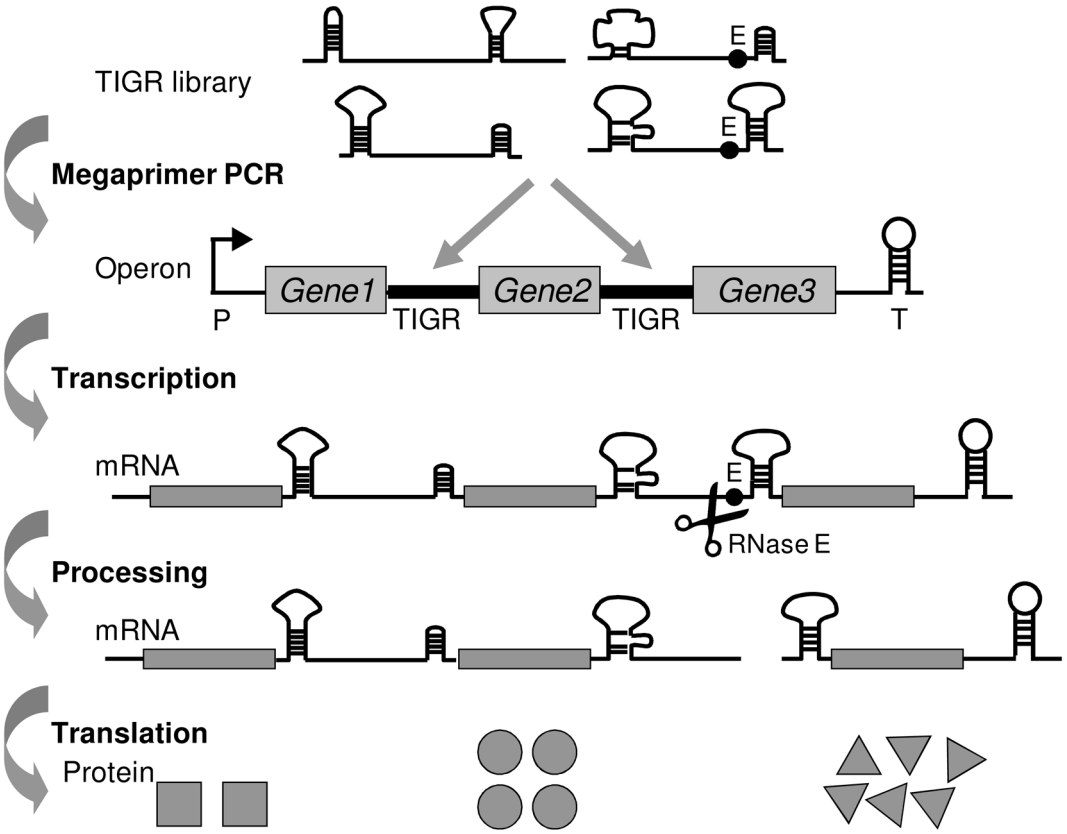

Figure 2.

Control of multiple gene expression levels within operons through tunable intergenic regions (TIGR). A library of TIGR sequences was assembled combinatorially from four sets of oligonucleotides using PCR. A megaprimer PCR approach was used to simultaneously place TIGR libraries within the intergenic regions of an operon with promoter (designated P) and terminator (designated T). The various TIGRs caused the mRNA coding regions to be differentially processed and translated, resulting in different levels of the various proteins encoded in the operon. In some TIGRs, RNase E sites (designated E) were incorporated to decouple the stability of the coding regions and thus the production of the corresponding proteins.

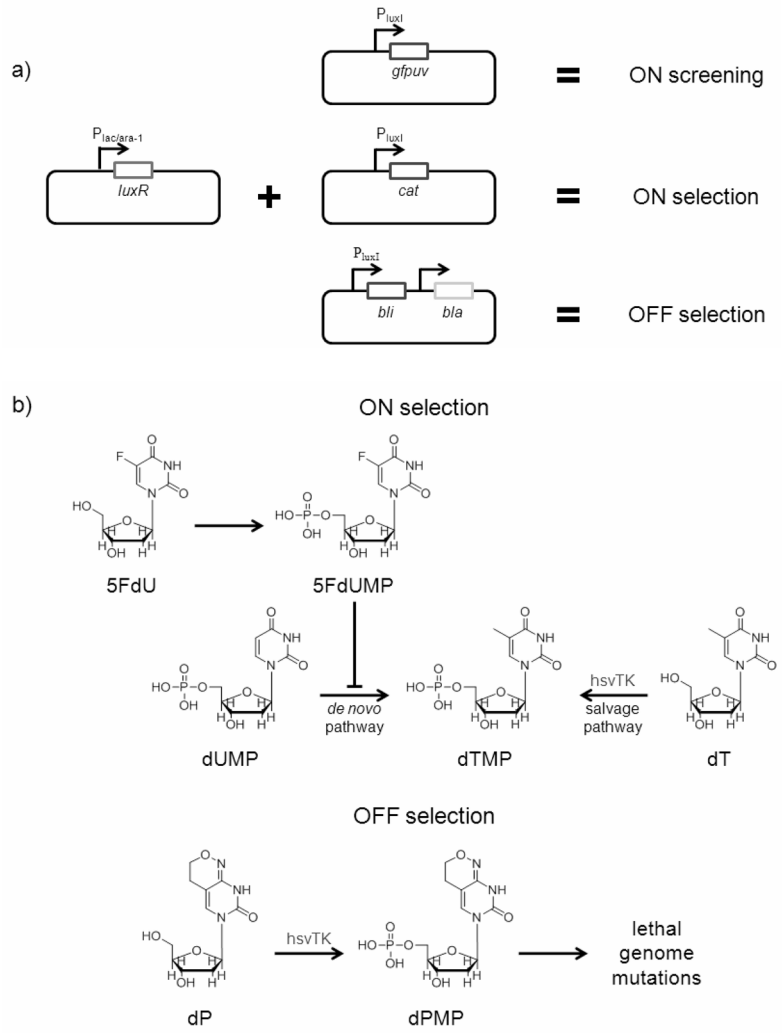

Figure 3.

Network level directed evolution strategies. (a) Directed evolution of the LuxR transcriptional regulator has been achieved through screening, ON selection, or OFF selection, where active LuxR stimulates transcription of a GFP marker, a chloramphenicol resistance gene, or a β-lactamase inhibitor, respectively. (b) Single gene selection markers such as *hsvTK* allow selection of both ON and OFF states. In ON selection, exogenously supplied 2'-deoxy-5-fluorouridine (5FdU) is converted by the host to 2'-deoxy-5-fluorouridine monophosphate (5FdUMP), which blocks *de novo* dTMP biosynthesis. Expression of *hsvTK* activates dTMP biosynthesis via the salvage pathway, rescuing the host. In OFF selection, leaky hsvTK expression produces the mutagenic nucleotide monophosphate dPMP from exogenously supplied dP, leading to host cell death.

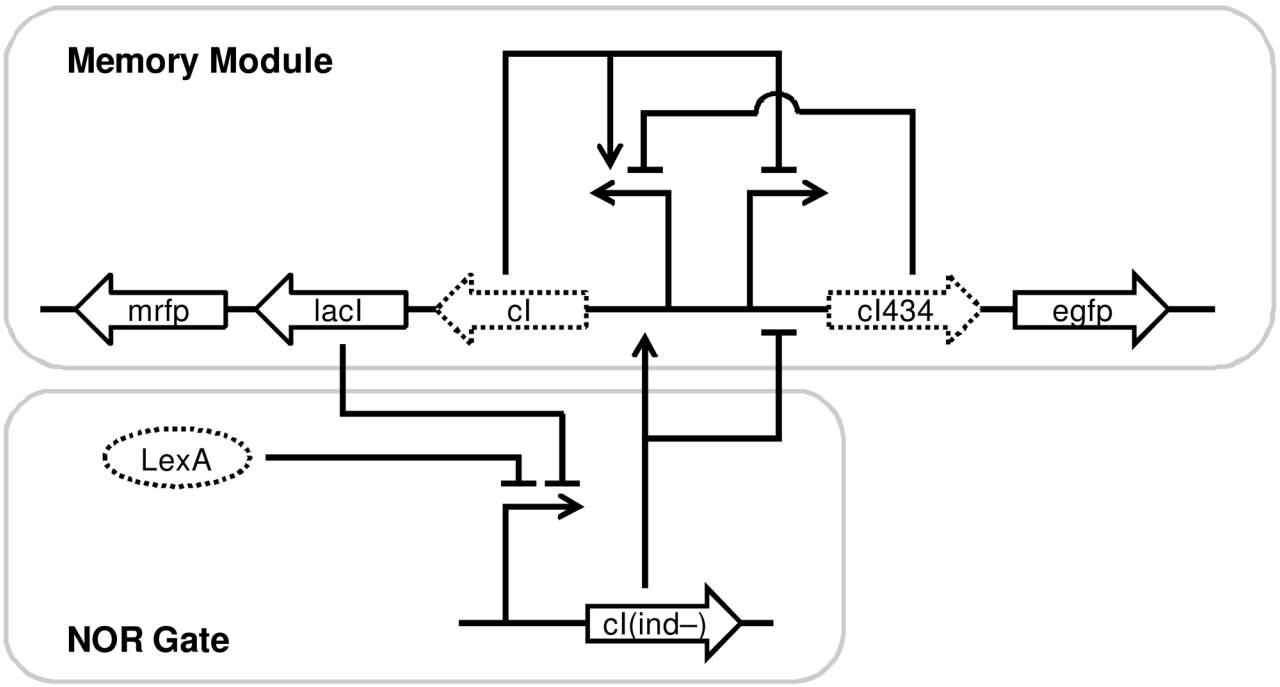

Figure 4.

An engineered bistable push-on push-off switch, adapted from reference 66. Dashed lines indicate gene products that can be cleaved by RecA*, which is activated by UV radiation. Arrows represent activation of the target promoter, while blunt ends represent repression. For a more detailed explanation, refer to the text.

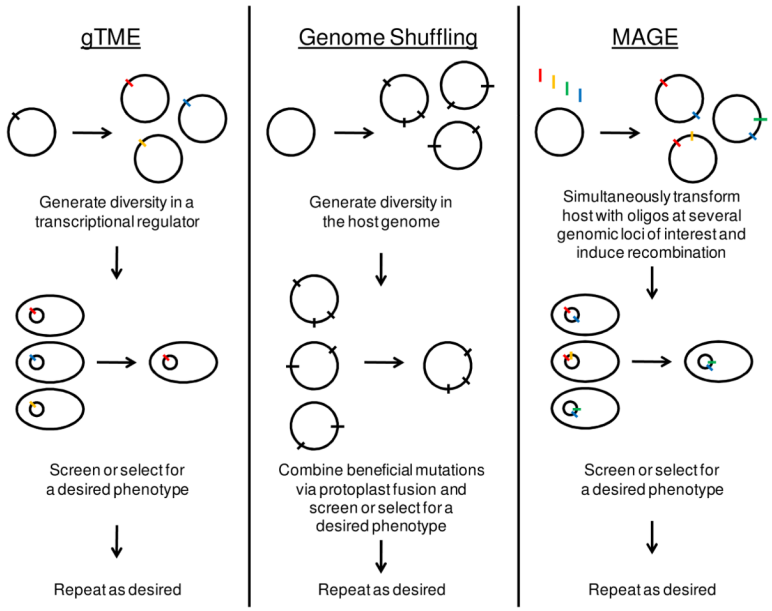

Figure 5.

Whole cell level directed evolution strategies. In gTME, diversity is generated (typically via random mutagenesis) in a particular transcriptional regulator of interest in the host genome. Subsequent screening or selection identifies mutants with the desired phenotype. In genome shuffling, diversity is generated throughout the genome, and beneficial mutations are subsequently combined through protoplast fusion. In MAGE, specific genomic loci are first identified and targeted (for mutation or deletion) with synthetic oligos, followed by several rounds of transformation and recombination. In all three strategies, several iterations yield the best results.