Synthetic biology to access and expand nature's chemical diversity

Citation: Smanski, Michael J. et al. "Synthetic Biology to Access and Expand Nature's Chemical Diversity." Nature Reviews Microbiology 14.3 (2016): 135–149.

As Published: http://dx.doi.org/10.1038/nrmicro.2015.24

Publisher: Nature Publishing Group

Persistent URL: http://hdl.handle.net/1721.1/107490

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

HHS Public Access

Author manuscript

Nat Rev Microbiol. Author manuscript; available in PMC 2016 October 04.

Published in final edited form as:

Nat Rev Microbiol. 2016 March ; 14(3): 135–149. doi:10.1038/nrmicro.2015.24.

Synthetic biology to access and expand nature's chemical diversity

Michael J. Smanski[1], **Hui Zhou**[2], **Jan Claesen**[3], **Ben Shen**[4], **Michael Fischbach**[3], and **Christopher A. Voigt**[2]

[1]Department of Biochemistry, Molecular Biology, and Biophysics and BioTechnology Institute, University of Minnesota – Twin Cities, Saint Paul, MN, 55108, USA

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

[3]Department of Bioengineering and Therapeutic Sciences, University of California-San Francisco, San Francisco, CA 94158

[4]Departments of Chemistry and Molecular Therapeutics, The Scripps Research Institute, Jupiter, FL 33458

Abstract

Bacterial genomes encode the biosynthetic potential to produce hundreds of thousands of complex molecules with diverse applications, from medicine to agriculture and materials. Economically accessing the potential encoded within sequenced genomes promises to reinvigorate waning drug discovery pipelines and provide novel routes to intricate chemicals. This is a tremendous undertaking, as the pathways often comprise dozens of genes spanning as much as 100+ kiliobases of DNA, are controlled by complex regulatory networks, and the most interesting molecules are made by non-model organisms. Advances in synthetic biology address these issues, including DNA construction technologies, genetic parts for precision expression control, synthetic regulatory circuits, computer aided design, and multiplexed genome engineering. Collectively, these technologies are moving towards an era when chemicals can be accessed *en mass* based on sequence information alone. This will enable the harnessing of metagenomic data and massive strain banks for high-throughput molecular discovery and, ultimately, the ability to forward design pathways to complex chemicals not found in nature.

Introduction

Natural products (NPs) are specialized metabolites produced by plants, animals, and microorganisms with diverse chemical structures and biological activities. These molecules are valuable in the clinical setting, with half of small molecule drugs approved during the past three decades being derived from NPs[1]. While NPs are prevalent in the treatment of infection[2], cancer[3], and as immunosuppressive agents[4], they have also made it into commercial products as antivirals, anthelmintics, enzyme inhibitors, nutraceuticals,

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

polymers, surfactants, bioherbicides, and vaccines[5]. In many NP-producing organisms, all of the genes required for regulation, biosynthesis, export, and self-resistance are co-localized in the genome in compact 'biosynthetic gene clusters' (BGCs) (Figure 1a).

Claims that natural products are an inexhaustible resource[6] are based on the disparity between the staggering biological and chemical diversity on the planet and the relatively low-throughput methods currently available to characterize these compounds. In the coming decades, advances in technology will close this gap and allow for a more systematic characterization of global NP production. Improving bioinformatic methods, combined with the dramatic rise in sequenced genomes, is shedding light on the potential number of undiscovered natural products (Figure 1b)[7–17]. In *Streptomyces* alone, conservative estimates put the number of natural products at 150,000, of which <5% have been discovered[18]. Bioinformatic investigations of hundreds of genomes across genera estimate that there are 100,000s of NPs[19], and the inclusion of less-studied classes, such as saccharides and lipids, substantially add to the number of molecules[12].

Currently, our ability to mine bacterial genomes to produce NPs is unable to keep pace with the identification of new BGCs by DNA sequencing and bioinformatics. However, the cost of DNA sequencing and synthesis continues to drop (Figure 1c) and future advances are projected to quickly make it possible to build the DNA for many pathways. This leaves design as the biggest remaining issue, where it is still challenging to rationally compose a DNA sequence for a large pathway that will be functional in a model production host. Engineering NP biosynthesis is still difficult for several reasons. Factors like transcription, translation, protein-protein interactions, cofactor and precursor availability, export, and self-resistance all need to be accounted for in a final production strain. In addition, many of the organisms that harbor these BGCs are difficult to manipulate or cultivate[20,21] and the transfer of a BCG to a new host, for which there are successful examples[22], is by no means trivial.

The goal of this review is to highlight emerging technologies relevant to engineering multi-gene systems with a special focus on the application of methods from synthetic biology to the engineering of biosynthetic pathways and NP-producing organisms. Many of these technologies were developed in model lab organisms like *E. coli* or *S. cerevisiae*, so when relevant, technical difficulties associated with moving these into major NP production strains are discussed. Other areas relevant to NP discovery, including methods of NP identification, manipulation of global and pathway-specific regulators, and prioritization of BGCs identified via genome sequencing efforts will not be discussed in detail, as these were reviewed recently[23].

Synthetic biology to accelerate natural product discovery and production

Engineering NP biosynthesis draws tools from a variety of subfields in genetic engineering and chemistry. For example, *protein engineering* seeks to modify the properties of individual proteins, including the activity/specificity/stability of enzymes and in the recombination of domains to diversify the products of large PKS/NRPS "assembly lines"[24]. *Metabolic engineering* is focused on understanding how multiple enzymes assemble into a pathway and

how this impacts metabolic fluxes in the larger natural network[25]. For example, the flux of a precursor to a natural product could be boosted as part of optimizing the titer. *Applied microbiology* and *strain engineering* have been critical for identifying high-producing organisms and optimizing the titer and performance in a fermentation through processes such as the random chemical mutagenesis of the genome[26]. *Synthetic biology* has focused on tools to accelerate and increase the scale of genetic engineering[27,28]. The scope of this review is to cover recent advances in synthetic biology in the light of how they will impact the field of NPs, including the following technologies:

- *Abstraction of genetic functions into "parts."* There has been an emphasis on creating genetic parts, such as promoters, that generate precise levels of gene expression[29–32]. There has been focus on generating large libraries of well-characterized parts and the development of biophysical/bioinformatics models to predict part behavior[33–36]. Part libraries for different organisms will aid the transfer of BGCs between hosts[37–41].
- *Large-scale construction technologies*. DNA synthesis capacity has exploded over the last decade and it is routine to synthesize the 20kb-100kb needed for a large gene cluster[42–44]. In addition, new DNA assembly methods enable the rapid construction of different part permutations or to substitute many parts in a single step[45–47].
- *Design automation*. New computer aided design methods and work environments accelerate the process of designing a genetic system, scanning the system for errors, and to analyze screening and –omics datasets[48–50].
- *Synthetic regulation*. Genetic circuits have been constructed that function as logic, timers, switches, and oscillators[51–53]. Sensors have also been developed that respond to many inducible inputs as well as metabolite levels. These could be incorporated into natural product pathways to control the timing of expression of different genes or to implement feedback in response to a toxic intermediate[54].
- *Genome editing for host design*. It is often desirable to make many simultaneous genomic changes. Methods, such as CRISPR-Cas9, can target essentially any region of the genome and have been shown to function in many species, including several host species well-suited for industrial production of small molecules[55–58].

Reducing genetics to genetic parts

Natural product BGCs are large and unwieldy[59]. They can comprise several dozen genes, arranged in one or many operons facing either direction. Their expression often relies on regulatory elements that are overlapping or imbedded in neighboring genes. They are under multiple layers of complex regulation including transcriptional and translational control. These factors makes engineering gene expression in BGCs technically difficult. This organization is in contrast to the concept of 'genetic parts,' which are units of DNA with

defined and modular function that replace native regulation to provide finer control over biological processes[60].

Applying a parts-level approach to multi-gene systems is facilitated by ‘refactoring’, in which the natural genetics are re-written to make the systems more amenable to engineering efforts (Figure 2a)[43,61]. A refactored gene cluster has several advantages that lend themselves to high-throughput applications. First, the process of refactoring severs the native regulation, which is usually incompletely understood, and allows for synthetic control of gene expression. Second, it leads to a modular genetic architecture, which facilitates part-swapping and combinatorial optimization. An example of an application would be in the creation of diverse compounds by substituting variations of enzymes from homologous clusters (Figure 2b). Further, the expression levels of the enzymes are important and it is more difficult to control if the regulatory parts (*e.g.,* RBSs) cannot be exchanged without leading to a web of other effects. Refactored BGCs can undergo wholesale swapping of genetic parts to optimize expression levels and increase titers. The concept of refactoring has started to be applied to BCGs[62,63] (Figure 2c) around the substitution of some synthetic regulatory elements and as the parts and tools improve, this will expand to include the complete elimination of native regulation.

Precision control of gene expression

Many metabolic pathways and BGCs are highly sensitive to gene expression, where small changes can cause a loss of activity[64,65]. Recent work to create large libraries of regulatory parts enable the graded control of gene expression over many orders of magnitude[30,33,37,66–68] (Figure 3a). In addition, computational methods have been developed that will design a new part based on biophysical models of transcription or translation[34,35,69,70]. While much work has been done in *E. coli*, there have been efforts to expand part libraries into other relevant organisms, including *Streptomyces*[37–39], B*acillus*[40], and fungi[41,66,71](Figure 3b). This can often be stymied by issues that are taken for granted in model organisms. For example, single cell techniques, such as flow cytometry and fluorescent reporters, are difficult to use in *Streptomyces* because they have multicellular branched growth patterns[72,73], differences in the DNA copy number, and the stochastic nature of gene expression seen during mycelial growth[72]. Different paradigms for measuring part function are needed for these hosts. Part design is also complicated by other issues; for example, the high GC content of some genomes make it difficult to codon optimize genes and design RNA parts, including RBSs[36,74].

In synthetic biology, the concept of the “expression cassette” has been expanded to include insulators that increase the reliability of part function in different genetic contexts[29–32](Figure 3a). This arose out of observations that different combinations of parts, for example promoters and RBSs, can lead to unexpected behaviors[30,75]. Some examples include ribozymes that decouple the promoter from the 5′-UTR[76] and bi-cistrons that decouple the 5′-UTR from the RBS[32]. These allow the promoters and RBS to be swapped to vary expression levels without impacting the behavior of neighboring parts. Similarly, long promoters and strong terminators have been developed to transcriptionally insulate the genes[33,77]. Collectively, this has led to genetic architectures that are more focused on the

control of individual genes as individual cistrons, as opposed to their organization into operons. Adopting these design principles for BGCs will be important for combinatorial optimization or in the exchange of parts to create chemical diversity.

High-throughput genetic optimization of multi-gene systems

Accessing new NPs from genome sequence information requires that the BGCs contained in sequence databases be converted into physical DNA constructs. High-throughput fabrication allows many designs to be tested in parallel, thereby increasing the probability of identifying a functional construct. This is beneficial both for chemical diversification via combinatorial biosynthesis[78–81], as well as for genetic optimization of pathway performance6[62,64].

There are two DNA construction technologies relevant to BGCs (Figure 4a). The first is *de novo* synthesis where genes or entire clusters are chemically constructed, typically by synthesis companies[82]. The cost has dropped dramatically in the last decade and it is possible to order hundreds of individual genes or full clusters[43,63,83]. While the cost has declined significantly, it is still expensive to build large clusters and building comprehensive sets of clusters out of the sequence databases is prohibitive. However, for NP biosynthetic classes such as RiPPs, this low cost synthetic DNA can be leveraged for combinatorial generation of new derivatives[84–86].

The second is DNA assembly, which constitutes the combination of parts to build a larger construct[46,87–94]. This enables many variants of gene clusters to be built based on a set of re-used underlying parts[45]. This is significantly cheaper that constructing *de novo* clusters for each variant that one wants to test. Many assembly methods are now available, including Isothermal assembly[91], Golden Gate assembly[95], ligase cycling reactions[94], scarless stitching[45], and recombination-based methods[92,96]. Automating these techniques using liquid handling robots enables hundreds or thousands of permuted combinations to be built[45,48,79,80,85,97].

Combinatorial optimization will be important in optimizing a BCG and transferring it between hosts. Epistasis in the expression levels of biosynthetic genes points to the importance of combinatorial optimization methods that look at more than one variable at a time[64,98]. Once pathways grow beyond a small handful of genes, complete exploration of the combinatorial gene expression space is impractical or impossible. This combinatorial space can be reduced via Multivariate Modular Metabolic Engineering by using *a priori* knowledge of a biochemical pathway (*e.g.,* enzyme kinetics, order of reactions, or pathway branching architecture) to constrain groups of genes into a small number of operons[99]. Combining metabolic modeling and RBS design, production levels of the isoprenoid neurosporene could be tuned over a continuous range from 0-300 μg per gram of dry cell weight[36]. Another approach is to use Design of Experiments (DOE) to reduce the number of experiments required to search a combinatorial space. These require that variables effecting construct performance can be readily manipulated, but do not require an understanding of the underlying biology. Recent examples in the neurosporene and violacein biosynthetic pathways shows that following a fractional factorial experiment, new positions in the sequence space can be accurately predicted[36,71].

Combinatorial assembly has been applied to the optimization of several NP pathways, as well as in the creation of chemical diversity. Several examples of optimizing BGCs, including for heterologous production of artimisinin[65], taxadiene[64], and opiate alkaloids[100] point to the importance of combinatorial engineering, as optimal production levels do not result from a blanket gene over-expression. Using a Gibson assembly protocol optimized for high GC content, the pathway specific regulators of pristinamycin biosynthesis were mutated in a combinatorial fashion to increase production levels to over 1 g/L[101]. An improved three-gene pathway for catechin production was created via combinatorial assembly by drawing from eight homologous biosynthetic genes from different plant species[102]. Certain classes of NPs, for example indolocarbazoles, have been greatly expanded using combinatorial DNA assembly with more than fifty derivatives created to date[78,103]. Finally, combinatorial assembly has been applied to probe the desire rules underlying large multi-modular enzymes[79–81,85,104,105]. For example, promiscuous polyketide donor or accepter modules were identified[79].

Host Transfer

Transferring a BGC between hosts is important for NP discovery, diversity screening, and optimization. This is particularly true if the BCG only appears in a sequence database and its native organism is unknown or inaccessible. Transferring a BCG would enable the new host to make the encoded compound.

However, the direct transfer of a BCG between even similar species can result in dramatic changes in the timing and relative expression levels for pathway genes (Figures 4c)[106]. This is unsurprising given the performance of genetic parts depends on components of the host cell's machinery[75]. These host-context effects can result from unintended crosstalk or interactions with native regulatory proteins[107], from limitations in host resources available for expressing heterologous constructs[108], and for NP production, from crosstalk with endogenous biosynthetic pathways. For example, transfer of the platencin gene cluster from *S. platensis* to the model host *S. lividans* resulted in the excess accumulation of shunt metabolites (Figure 4c,d)[106]. This was correlated with substantial changes in gene expression patterns. A likely explanation is that improperly balanced expression levels in the heterologous host led to build-up of pathway intermediates, which were then subject to modification by endogenous biosynthetic enzymes and thus diverted away from the desired product (Figure 4d). Even moving multi-gene systems just between different strains of the same species can likewise negatively impact performance[107,109,110].

One way to transfer the BCG more effectively would be to exchange parts for those that are known to function in the new species (Figure 4e). Effort could be taken to ensure that the desired expression levels are reached in the new host chassis. For example, differences in the anti-Shine-Dalgarno sequence in the small ribosomal subunit mean that RBS strengths are effected upon transfer to disparate hosts and this requires that all the RBSs be simultaneously redesigned[36]. This type of wholesale reassignment of regulatory parts is much easier in refactored genetic systems, whose modular architectures allow them to be built up from parts quickly. Also important is increasing the size of characterized part libraries for NP-relevant species[37–41,66,67,71,90,111,112]. For complex multigene systems, the

dramatic design changes that accompany host transfer will likely require further genetic optimization via combinatorial DNA assembly to re-tune expression levels and improve performance (previous section)[45].

Genetically-encoded biosensors for biosynthetic pathway engineering

The design and construction of DNA libraries has accelerated to the point where screening and making sense of this diversity is the bottleneck. While bioassay-guided prescreens or selections can reduce the number of strains that need to be investigated in detail[26], these do not replace direct measurements of titer. Current analytical chemistry methods have reduced analysis time to <3 min per sample[113], but this still limits throughput to ~10^3 per day. As an alternative to analytical chemistry, *in vivo* biosensors translate information about a chemical signal, *i.e.* the concentration of a natural product, into light or fluorescence-based output that can be measured by flow-cytometry to screen thousands genotypes per second[114]. As an example of the power of this approach, the inherent fluorescence of the carotenoid astaxanthine was exploited to track titers in single cells by flow cytometry and cell-sorting, enabling a 10,000-fold enrichment of over-producing strains compared to plate-based techniques[115].

Intracellular biosensors can be broadly grouped into three categories, RNA-based, protein-based, or enzyme-based, according to their biomolecular make-up and mechanism. RNA aptamers that bind to small molecule signals to actuate a response have been linked to readouts including fluorescence, enzyme activity, cell mobility, or viability[114,116]. There are diverse strategies for designing RNA biosensors[117–125]. RNA aptamers have been used to build biosensors for natural products and intermediates, including theophylline[121], tetracycline[124], neomycin[126], tobramycin[127], dopamine[128], and ochratoxin A[129].

Protein biosensors function by transmitting molecular binding information into a measurable output, usually in the form of allosterically-regulated transcriptional activators or repressors controlling fluorescent protein expression. Naturally-evolved biosensors (*e.g.* Figure 5a,b) can detect a wide range of molecular scaffolds, including tetracyclines[130,131], cationic lipids and plant alkaloids[132], and anthraquinones[133]. There are >4000 TetR proteins identified in sequence databases, and only a small fraction of the ligands are known[134]. While protein biosensors have been found or engineered for a number of target molecules, including aromatics[135,136] and branched-chain amino acids [137], these will need to be repurposed to sense new molecules for generalized use in NP discovery and optimization pipelines. Cirino and co-workers engineered the transcriptional activator AraC to recognize either mevalonate (a key precursor for isoprenoids)[138] or triacetic acid lactone (a simple polyketide product)[139]. In an alternative to modulating gene expression, the Liu group has engineered protein biosensors that sense small molecule ligands to actuate an intein-splicing event[140,141], thereby allowing a diverse set of proteins or enzymes to be activated only in the presence of the ligand.

Enzymatic biosensors recognize the desired metabolite and convert it into a pigmented or fluorescent molecule that can be easily detected by spectrophotometry or flow cytometry. To optimize the production of L-DOPA in yeast, Dueber and co-workers expressed a DOPA

dioxygenase intracellularly[142]. The DOPA dioxygenase results in an extradiol ring cleavage of L-DOPA to produce the aldehyde-containing betalamic acid, which spontaneously reacts with cellular amino acids to form fluorescent imines known collectively as betaxanthins[143].

Whole cell biosensors can be created without the need for complicated protein or aptamer engineering[144]. For example, recombinant *E. coli* made auxotrophic for mevalonate shows concentration-dependent growth rate changes in the presence of extracellular mevalonate[145]. By expressing GFP within this strain, mevalonate levels can be detected in high-throughput from the culture broths of a production strain[145]. In principle, this strategy can be used to quickly engineer a whole cell biosensor to any molecule that can be made essential for strain viability[146].

Potential applications for synthetic genetic circuits

The temporal control of expression is often important in building complex chemicals and materials[147,148]. In natural systems, this is implemented via regulatory networks consisting of interacting proteins, RNA, and DNA that collectively work to perform a computational operation. Synthetic genetic circuits have been built where a target behavior is achieved by artificially connecting regulatory proteins (Figure 5). This has been used to build cascades, bistable switches, pulse generators, oscillators, feedback/feedforward motifs, and logic gates[29,51]. NP pathways reveal complex and intricate control mechanisms including many of these same behaviors[149–151]. However, the naturally evolved regulation is not required for high-level production in defined culture conditions. In fact, its disruption often leads to improved production[152–154]. Replacing native regulation with synthetic circuits may implement the necessary feedback and dynamics without also having the environmental control that can inhibit production.

Feedback and feedforward regulation has been used to tie the accumulation of early-stage pathway intermediates with the expression of downstream processing genes. Feedforward regulation is seen in many NP biosynthetic pathways. In actinorhodin biosynthesis, the accumulation of pathway intermediates triggers expression of the efflux pump to export the final product[133]. In this case, the feed-forward motif helps to prevent the cell from deleterious effects of the accumulating high concentrations of the antibiotic within the cell. Modeling of a synthetic pathway for *para*-aminostyrene biosynthesis suggests that higher titers can be attained with dynamic regulation incorporating feedback/feedforward regulation compared to static regulation (Figure 5c)[155]. Using a fatty-acid biosensor to add feedforward/feedback regulation into a synthetic fatty acid ester pathway allowed Zhang *et al.* to boost bio-diesel production to 28% theoretical yield (Figure 5d)[54]. Additionally, positive and negative feedback loops can be exploited to control the allocation of cellular resources to secondary metabolism (Figure 5e)[156].

Synthetic circuits that act as metabolic “control valves” have been used to redirect carbon flux from primary to secondary pathways[157,158]. This dynamic control over central carbon metabolism is important, because if this diversion is constitutive, it slows growth to the point of decreasing productivity. Similar dynamic switching is seen in *Streptomyces* prior to antibiotic production[159]. The expression of housekeeping genes, particularly those involved

in translation, is strongly diminished prior to entering stationary phase, at which time actinorhodin and undecylprodigiosin production begins.

Control over the timing of gene expression can also be achieved using synthetic bistable switches, which reset slowly over time (Figure 5f)[160]. Bistable switches appear in many NP producing organisms, including *Streptomyces coelicolor* where it controls cryptic polyketide (*cpk*) production[151]. Engineered timing circuits can be used to separate cell growth and NP production phases, particularly in strains for which the regulatory connections to the natural metabolic switch[159] have been severed.

Genetic logic gates allow multiple input signals to be integrated before a pathway turns on. There is some evidence of logic in natural NP pathways, for example, the actinorhodin gene cluster is controlled by metabolite concentrations, stress response, and development program[149,150]. This can be used to turn on different sets of genes under different environmental conditions, the specificity of which improves as more signals are integrated. In synthetic biology, many logic gates have been built[29,52,107,110,161–163]. Connecting synthetic multi-input circuits to NP pathways could allow cells to sense cofactor levels, precursor abundance, dissolved oxygen content, or ATP charge before deciding whether or not to commit to NP biosynthesis. Such intracellular checkpoints could prevent the accumulation of unwanted intermediates and byproducts by cells that are not capable of making the final product.

CRISPRi has emerged as a powerful method to regulate gene expression, including those in the genome, for both prokaryotes and eukaryotes. It is based on the expression of a catalytically-inactive dCas9, which can be directed to a target when a sgRNA is transcribed[164]. dCas9 can serve as either a repressor or activator (CRISPRa) depending on domains to which it is fused[165]. CRISPRi has been shown to work in *E. coli*[164], fungi[165], actinobacteria[55], and plants[166]. Other organisms relevant to NP biosynthesis, such as cyanobacteria[167], burkholderia[168], pseudomonads[169], and myxobacteria[170], have endogenous CRISPR systems but have not as of yet been exploited as hosts for CRISPRi regulation. CRISPRi has already been exploited to control metabolic fluxes via multiplexed gene repression of endogenous pathways in *E. coli*[171] and in heterologous pathways in yeast (Figure 5g)[172].

New tools for combinatorial genome-scale engineering

Mutations in the genome outside of the BGC are required to optimize the titers of a NP. Some industrial strains can achieve gram per liter quantities, and this is usually achieved via random mutagenesis of the genome and screening[26]. Originally, the genetic diversity was generated using techniques such as chemical mutagenesis, but this has gotten more sophisticated with improved molecular biology methods. For example, the whole-genome shuffling of a tylosin producer yielded the same 6-fold improvement in 24,000 assays that had taken Eli Lilly 20 years and 1 million assays to achieve by classical methods[173]. Synthetic biology offers new techniques for generating genome diversity, from methods to replace parts or make defined mutations in a multiplexed manner to genome construction via *de novo* synthesis[44,174–176].

Multiplexed genome engineering strategies offer the ability to precisely target hundreds of loci in the genome for over- or under-expression in parallel. The first demonstration of massively-multiplexed recombineering in *E. coli* exploited oligo-mediated allelic replacement[177]. By mimicking Okazaki fragments at the replication fork, exogenous ssDNAs are able to anneal with the lagging strand of the genome serving as primers for DNA elongation and then get incorporated, which leads to simultaneous directed mutagenesis at multiple sites in the genome[178]. MAGE (Multiplexed Automated Genome Engineering) automates this process, which allows combinatorial exploration of mutations in a continuously evolving population. For example, twenty endogenous *E. coli* genes were targeted to optimize lycopene production and billions of variants were screened[177]. MAGE has also been used to insert regulatory parts into the genome, such as N- and C-terminal tags[179] and T7 RNAP promoters[176,180]. Recombineering has been shown to work in diverse organisms, including lactic acid bacteria[181], mycobacteria[182], corynebacteria[183], and fungi[184]. Getting MAGE to work in *Streptomyces* will be challenging due to the high GC-content, the lack of characterization of mismatch repair in this species, and the fact that the expression of some genes required by the technique (e.g., *bet*) is not likely to produce functional proteins[183], and transformation is significantly more challenging[185]. MAGE has already been shown to work in yeast and can be applied to pathways transferred to this host[184] and it may work in other NP-relevant fungi.

TRMR (trackable multiplex recombineering) is a related method, which was developed to rapidly map the effects of more than 95% of *E. coli* genes onto specific traits[174]. TRMR exploits array-based DNA synthesis to create barcoded oligonucleotides that target over 4,000 genes for either over-expression or repression. Following phenotypic enrichment, deep-sequencing allows the targeted mutations to be quickly mapped to identify the causal mutations, generating massive amounts of sequence-to-phenotype relationships[174,186]. Combining TRMR with multiplexed recombineering allowed the identification of 27 genome modification targets that accelerated growth in a target medium[186].

CRISPR techniques have revolutionized multiplexed genome engineering[55–58,187–197]. The Cas9 nuclease can be targeted to specific sequences by transcribing a sgRNA (Figure 6a). This system has been shown to work in nearly every organism that has been tried, including prokaryotes [55,57,171,187,189–191,195] such as *Streptomyces* (Figure 6b), eukaryotes[58,188,198], and higher organisms of relevance to NP production[56,199]. The lack of a canonical non-homologous end joining (NHEJ) DNA repair in some prokaryotes, including *Streptomyces*, lowers the efficiency of gene inactivation when CRISPR-Cas9 is used alone[55,200]. However, the efficiency of gene inactivation can be increased to over 75% by (i) including a double-stranded ‘repair fragment’ that can close the double-stranded DNA break via homologous recombination[57], or (ii) reconstituting a NHEJ pathway through heterologous expression of the ligase LigD during genome editing[55]. This system can be used to make five mutations in a single step (Figure 6c)[188], knockout 31 kilobase gene clusters[57], and insert DNA up to 9 kilobases in length[198].

Conclusions

New strategies from synthetic biology are enabling the engineering of large systems comprising many genetic parts, the control of gene expression with synthetic regulation, and efficient genome editing. New tools exist that provide precise control of gene expression from synthetic constructs, and the fabrication of large systems is made easier by abstracting designs in a parts-based approach to genetic engineering. While many of the approaches in this review were developed in a model organism, such as *E. coli*, over the last few years they have been increasingly ported to organisms of more direct relevance to NP discovery and production.

For NP chemists and biologists, the challenge will be determining how best to leverage the latest technologies in DNA fabrication and genetic control to probe NP pathways in new and insightful ways. Much in the same way that recombinant DNA technology revolutionized our ability to approach the molecular details of biology from a reductionist point of view, the ability to rapidly build large libraries of specifically designed gene clusters will provide greater opportunity to explore the effect of genetic design on the functional expression of BGCs. Continued research into the detailed regulatory mechanisms of employed within natural BGCs and the biochemistry of NP biosynthetic pathways will be paramount for forming hypotheses that can be tested using new bottom-up techniques.

Current high-throughput and multiplexed genetic engineering strategies can be harnessed to develop applications for NP-producers outside of the fermenter as well (Figure 7). This could have applications in environmental sensing, for example producing a small volatile metabolite in response to metal contamination in soils, or in the production of therapeutics by probiotic strains, for example making antibacterial compounds in response to a pathogen in the GI tract[201–204]. NPs are already used extensively in agriculture for crop protection[205], and gaining fine-tuned control over production dynamics either in soil microbial communities or in crop plants themselves could impact food production.

This is an exciting time for NP biosynthesis. The number of possible applications for NPs in medicine, industry, and agriculture is vast. The recent explosion in DNA sequencing technologies has revealed that BGCs for NP production are more widespread than previously imagined[12,19,206], and the ability to ‘write’ DNA into synthetic constructs is now catching up. New approaches for mining NPs from genomic sequences are needed more than ever to rejuvenate waning drug discovery pipelines, especially in light of the looming crisis in antibiotic resistance[207]. The development of state-of-the-art high-throughput screening platforms allow purified compounds or semi-pure extracts to be screened in hundreds of assays with less material required than in previous decades[208,209]. The suite of NPs present in nature is one of our most valuable natural resources, and we are now poised to more fully explore the extent of its depth and diversity.

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. Journal of Natural Products. 2012; 75:311–335. [PubMed: 22316239]

2. Clardy J, Fischbach MA, Walsh CT. New antibiotics from bacterial natural products. Nat. Biotechnol. 2006; 24:1541–1550. [PubMed: 17160060]
3. Cragg GM, Grothaus PG, Newman DJ. Impact of natural products on developing new anti-cancer agents. Chem. Rev. 2009; 109:3012–3043. [PubMed: 19422222]
4. Mann J. Natural products as immunosuppressive agents. Nat. Prod. Rep. 2001; 18:417–430. [PubMed: 11548050]
5. Demain AL. Importance of microbial natural products and the need to revitalize their discovery. J. Ind. Microbiol. Biotechnol. 2014; 41:185–201. [PubMed: 23990168]
6. Davies J. How to discover new antibiotics: Harvesting the parvome. Current Opinion in Chemical Biology. 2011; 15:5–10. [PubMed: 21111668]
7. Starcevic A, et al. ClustScan: An integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. Nucleic Acids Res. 2008; 36:6882–6892. [PubMed: 18978015]
8. Weber T, et al. CLUSEAN: A computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. J. Biotechnol. 2009; 140:13–17. [PubMed: 19297688]
9. Bachmann BO, Ravel J. Methods for *in silico* prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. Methods Enzymol. 2009; 458:181–217. [PubMed: 19374984]
10. Röttig M, et al. NRPSpredictor2 - A web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res. 2011; 39
11. Kim J, Yi G-S. PKMiner: a database for exploring type II polyketide synthases. BMC Microbiology. 2012; 12:169. [PubMed: 22871112]
12. Cimermancic P, et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. Cell. 2014; 158:412–421. [PubMed: 25036635]
13. Van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP. BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acids Res. 2013; 41
14. Ansari MZ, Sharma J, Gokhale RS, Mohanty D. In silico analysis of methyltransferase domains involved in biosynthesis of secondary metabolites. BMC Bioinformatics. 2008; 9:454. [PubMed: 18950525]
15. Kamra P, Gokhale RS, Mohanty D. SEARCHGTr: A program for analysis of glycosyltransferases involved in glycosylation of secondary metabolites. Nucleic Acids Res. 2005; 33
16. Blin K, et al. antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. 2013; 41
17. Medema MH, et al. AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 2011; 39
18. Watve MG, Tickoo R, Jog MM, Bhole BD. How many antibiotics are produced by the genus Streptomyces? Archives of Microbiology. 2001; 176:386–390. [PubMed: 11702082]
19. Doroghazi JR, et al. A roadmap for natural product discovery based on large-scale genomics and metabolomics. Nat. Chem. Biol. 2014; 10
20. Galm U, et al. In vivo manipulation of the bleomycin biosynthetic gene cluster in Streptomyces verticillus ATCC15003 revealing new insights into its biosynthetic pathway. J. Biol. Chem. 2008; 283:28236–28245. [PubMed: 18697737]
21. Piel J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles. Proc. Natl. Acad. Sci. U. S. A. 2002; 99:14002–14007. [PubMed: 12381784]
22. Galm U, Shen B. Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis. Expert Opinion on Drug Discovery. 2006; 1:409–437. [PubMed: 23495943]
23. Rutledge PJ, Challis GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat. Rev. Microbiol. 2015; doi: 10.1038/nrmicro3496

24. Williams GJ. Engineering polyketide synthases and nonribosomal peptide synthetases. Current Opinion in Structural Biology. 2013; 23:603–612. [PubMed: 23838175]
25. Pickens LB, Tang Y, Chooi Y-H. Metabolic Engineering for the Production of Natural Products. Annual Review of Chemical and Biomolecular Engineering. 2011; 2:211–236.
26. Demain AL, Adrio JL. Strain improvement for production of pharmaceuticals and other microbial metabolites by fermentation. Progress in Drug Research. 2008; 65:252–289.
27. Voigt CA. Synthetic biology. ACS Synth. Biol. 2012; 1:1–2. [PubMed: 23651002]
28. Way JC, Collins JJ, Keasling JD, Silver P. a. Integrating biological redesign: Where synthetic biology came from and where it needs to go. Cell. 2014; 157:151–161. [PubMed: 24679533]
29. Nielsen AAK, Segall-Shapiro TH, Voigt CA. Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression. Current Opinion in Chemical Biology. 2013; 17:878–892. [PubMed: 24268307]
30. Mutalik VK, et al. Quantitative estimation of activity and quality for collections of functional genetic elements. Nat. Methods. 2013; 10:347–53. [PubMed: 23474467]
31. Kosuri S, et al. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:14024–9. [PubMed: 23924614]
32. Mutalik VK, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. Nat. Methods. 2013; 10:354–60. [PubMed: 23474465]
33. Chen Y-J, et al. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat. Methods. 2013; 10:659–64. [PubMed: 23727987]
34. Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. 2009; 27:946–950. [PubMed: 19801975]
35. Brewster RC, Jones DL, Phillips R. Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia coli. PLoS Comput. Biol. 2012; 8
36. Farasat I, et al. Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. Mol. Syst. Biol. 2014; 10:731. [PubMed: 24952589]
37. Seghezzi N, Amar P, Koebmann B, Jensen PR, Virolle MJ. The construction of a library of synthetic promoters revealed some specific features of strong Streptomyces promoters. Appl. Microbiol. Biotechnol. 2011; 90:615–623. [PubMed: 21243353]
38. Siegl T, Tokovenko B, Myronovskyi M, Luzhetskyy A. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. Metab. Eng. 2013; 19:98–106. [PubMed: 23876413]
39. Sohoni SV, Fazio A, Workman CT, Mijakovic I, Lantz AE. Synthetic promoter library for modulation of actinorhodin production in Streptomyces coelicolor A3(2). PLoS One. 2014; 9
40. Phan TTP, Nguyen HD, Schumann W. Development of a strong intracellular expression system for Bacillus subtilis by optimizing promoter elements. J. Biotechnol. 2012; 157:167–172. [PubMed: 22100269]
41. Hartner FS,FS, et al. Promoter library designed for fine-tuned gene expression in Pichia pastoris. Nucleic Acids Res. 2008; 36
42. Kodumal SJ, et al. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:15573–15578. [PubMed: 15496466]
43. Temme K, Zhao D, Voigt CA. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. Proceedings of the National Academy of Sciences. 2012; 109:7085–7090.
44. Gibson DG, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Science. 2010; 329:52–56. [PubMed: 20488990]
45. Smanski MJ, et al. Functional optimization of gene clusters by combinatorial design and assembly. Nat. Biotechnol. 2014; doi: 10.1038/nbt.3063
46. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One. 2011; 6
47. Gibson D, Smith H, C. H. Chemical synthesis of the mouse mitochondrial genome. Nat. Methods. 2010; 7

48. Hillson NJ, Rosengarten RD, Keasling JD. J5 DNA assembly design automation software. ACS Synth. Biol. 2012; 1:14–21. [PubMed: 23651006]
49. Chen J, Densmore D, Ham TS, Keasling JD, Hillson NJ. DeviceEditor visual biological CAD canvas. Journal of Biological Engineering. 2012; 6:1. [PubMed: 22373390]
50. Bilitchenko L, et al. Eugene - A domain specific language for specifying and constraining synthetic biological parts, devices, and systems. PLoS One. 2011; 6
51. Brophy JAN, Voigt CA. Principles of genetic circuit design. Nat. Methods. 2014; 11:508–20. [PubMed: 24781324]
52. Hasty J, McMillen D, Collins JJ. Engineered gene circuits. Nature. 2002; 420:224–230. [PubMed: 12432407]
53. Balázsi G, Van Oudenaarden A, Collins JJ. Cellular decision making and biological noise: From microbes to mammals. Cell. 2011; 144:910–925. [PubMed: 21414483]
54. Zhang F, Carothers JM, Keasling JD. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nature Biotechnology. 2012; 30:354–359.
55. Tong Y, Charusanti P, Zhang L, Weber T, Lee SY. CRISPR-Cas9 based engineering of actinomycetal genomes. ACS Synth. Biol. 2015; 150325161425009. doi: 10.1021/acssynbio.5b00038
56. Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods. 2013; 9:39. [PubMed: 24112467]
57. Cobb RE, Wang Y, Zhao H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. ACS Synth. Biol. 2014; 4:723–728. [PubMed: 25458909]
58. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339:819–23. [PubMed: 23287718]
59. Fischbach M, Voigt CA. Prokaryotic gene clusters: A rich toolbox for synthetic biology. Biotechnology Journal. 2010; 5:1277–1296. [PubMed: 21154668]
60. Endy D. Foundations for engineering biology. Nature. 2005; 438:449–53. [PubMed: 16306983]
61. Chan LY, Kosuri S, Endy D. Refactoring bacteriophage T7. Mol. Syst. Biol. 2005; 1:2005–0018. [PubMed: 16729053]
62. Shao Z, et al. Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. ACS Synth. Biol. 2013; 2:662–669. [PubMed: 23968564]
63. Oßwald C, et al. Modular construction of a functional artificial epothilone polyketide pathway. ACS Synth. Biol. 2012; 121025132953007. doi: 10.1021/sb300080t
64. Ajikumar PK, et al. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science. 2010; 330:70–74. [PubMed: 20929806]
65. Paddon CJ, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013; 496:528–32. [PubMed: 23575629]
66. 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]
67. Braatsch S, Helmark S, Kranz H, Koebmann B, Jensen PR. Escherichia coli strains with promoter libraries constructed by Red/ET recombination pave the way for transcriptional fine-tuning. Biotechniques. 2008; 45:335–337. [PubMed: 18778259]
68. Stanton BC, et al. Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat. Chem. Biol. 2014; 10:99–105. [PubMed: 24316737]
69. Rhodius VA, Mutalik VK. Predicting strength and function for promoters of the Escherichia coli alternative sigma factor, sigmaE. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:2854–2859. [PubMed: 20133665]
70. De Mey M, Maertens J, Lequeux GJ, Soetaert WK, Vandamme EJ. Construction and model-based analysis of a promoter library for E. coli: an indispensable tool for metabolic engineering. BMC Biotechnol. 2007; 7:34. [PubMed: 17572914]

71. Lee ME, Aswani A, Han AS, Tomlin CJ, Dueber JE. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Res. 2013; 41:10668–10678. [PubMed: 24038353]
72. Jakimowicz D, Van Wezel GP. Cell division and DNA segregation in Streptomyces: How to build a septum in the middle of nowhere? Mol. Microbiol. 2012; 85:393–404. [PubMed: 22646484]
73. McCormick JR, Su EP, Driks A, Losick R. Growth and viability of Streptomyces coelicolor mutant for the cell division gene ftsZ. Mol. Microbiol. 1994; 14:243–254. [PubMed: 7830569]
74. Gustafsson C, Govindarajan S, Minshull J. Codon bias and heterologous protein expression. Trends in Biotechnology. 2004; 22:346–353. [PubMed: 15245907]
75. Cardinale S, Arkin AP. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. Biotechnology Journal. 2012; 7:856–866. [PubMed: 22649052]
76. Lou C, Stanton B, Chen Y-J, Munsky B, Voigt CA. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nature Biotechnology. 2012; doi: 10.1038/nbt.2401
77. Davis JH, Rubin AJ, Sauer RT. Design, construction and characterization of a set of insulated bacterial promoters. Nucleic Acids Res. 2011; 39:1131–1141. [PubMed: 20843779]
78. Sánchez C, et al. Combinatorial biosynthesis of antitumor indolocarbazole compounds. Proc. Natl. Acad. Sci. U. S. A. 2005; 102:461–466. [PubMed: 15625109]
79. Menzella HG, et al. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. Nat. Biotechnol. 2005; 23:1171–1176. [PubMed: 16116420]
80. Kakule, TB.; Lin, Z.; Schmidt, EW. Combinatorialization of Fungal Polyketide Synthase – Peptide Synthetase Hybrid Proteins. 2014.
81. Nguyen KT, et al. Combinatorial biosynthesis of novel antibiotics related to daptomycin. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:17462–17467. [PubMed: 17090667]
82. Kosuri S, Church GM. Large-scale de novo DNA synthesis: technologies and applications. Nat. Methods. 2014; 11:499–507. [PubMed: 24781323]
83. Bayer TS, et al. Synthesis of methyl halides from biomass using engineered microbes. J. Am. Chem. Soc. 2009; 131:6508–6515. [PubMed: 19378995]
84. Donia MS, et al. Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. Nat. Chem. Biol. 2006; 2:729–35. [PubMed: 17086177]
85. Ru, DE.; Schmidt, EW.; Heemstra, JR. Assessing the Combinatorial Potential of the RiPP Cyanobactin tru Pathway. 2015.
86. Mitchell DA, et al. Structural and functional dissection of the heterocyclic peptide cytotoxin streptolysin S. J. Biol. Chem. 2009; 284:13004–13012. [PubMed: 19286651]
87. Werner S, Engler C, Weber E, Gruetzner R, Marillonnet S. Fast track assembly of multigene constructs using golden gate cloning and the MoClo system. Bioeng. Bugs. 2012; 3:38–43. [PubMed: 22126803]
88. Chen WH, Qin ZJ, Wang J, Zhao GP. The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly. Nucleic Acids Res. 2013; 41
89. Anderson JC, et al. BglBricks: A flexible standard for biological part assembly. J. Biol. Eng. 2010; 4:1. [PubMed: 20205762]
90. Sarrion-Perdigones A, et al. GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. Plant Physiol. 2013; 162:1618–31. [PubMed: 23669743]
91. Gibson DG, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods. 2009; 6:343–345. [PubMed: 19363495]
92. Shao Z, Zhao H, Zhao H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic Acids Res. 2009; 37
93. Briggs AW, et al. Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res. 2012; 40
94. Kok, S. De, et al. Rapid and reliable DNA assembly via ligase cycling reaction. ACS Synth. Biol. 2014; 3:97–106. [PubMed: 24932563]
95. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One. 2008; 3

96. Zhang Y, Werling U, Edelmann W. SLiCE: A novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 2012; 40

97. Appleton E, Tao J, Haddock T, Densmore D. Interactive assembly algorithms for molecular cloning. Nat. Methods. 2014; 11:657–62. [PubMed: 24776633]

98. Xu P, et al. Modular optimization of multi-gene pathways for fatty acids production in E. coli. Nat. Commun. 2013; 4:1409. [PubMed: 23361000]

99. Biggs BW, De Paepe B, Santos CNS, De Mey M, Ajikumar PK. Multivariate modular metabolic engineering for pathway and strain optimization. Curr. Opin. Biotechnol. 2014; 29:156–162. [PubMed: 24927371]

100. Thodey K, Galanie S, Smolke CD. A microbial biomanufacturing platform for natural and semisynthetic opioids. Nat. Chem. Biol. 2014; 10:837–844. [PubMed: 25151135]

101. Li L, et al. A stepwise increase in pristinamycin II biosynthesis by Streptomyces pristinaespiralis through combinatorial metabolic engineering. Metab. Eng. 2015; 29:12–25. [PubMed: 25708513]

102. Zhao S, et al. Improvement of catechin production in Escherichia coli through combinatorial metabolic engineering. Metab. Eng. 2015; 28:43–53. [PubMed: 25527438]

103. Du Y-L, Ryan KS. Expansion of Bisindole Biosynthetic Pathways by Combinatorial Construction. ACS Synth. Biol. 2015; 150102122128000. doi: 10.1021/sb5003218

104. Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature. 2006; 440:1078–1082. [PubMed: 16495946]

105. Chemler, J. a., et al. Evolution of Efficient Modular Polyketide Synthases by Homeologous Recombination. J. Am. Chem. Soc. 2015; 150731185018005. doi: 10.1021/jacs.5b04842

106. Smanski MJ, et al. Expression of the platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. J. Nat. Prod. 2012; 75:2158–2167. [PubMed: 23157615]

107. Wang B, Kitney RI, Joly N, Buck M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nat. Commun. 2011; 2:508. [PubMed: 22009040]

108. Hajimorad M, Gray PR, Keasling JD. A framework and model system to investigate linear system behavior in Escherichia coli. J. Biol. Eng. 2011; 5:3. [PubMed: 21510907]

109. Egbert RG, Klavins E. Fine-tuning gene networks using simple sequence repeats. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:16817–22. [PubMed: 22927382]

110. Moser F, et al. Genetic circuit performance under conditions relevant for industrial bioreactors. ACS Synth. Biol. 2012; 1:555–564. [PubMed: 23656232]

111. Yamamoto YY, Obokata J. ppdb: A plant promoter database. Nucleic Acids Res. 2008; 36

112. Boyle PM, et al. A BioBrick compatible strategy for genetic modification of plants. J. Biol. Eng. 2012; 6:8. [PubMed: 22716313]

113. Xu RN, Fan L, Rieser MJ, El-Shourbagy TA. Recent advances in high-throughput quantitative bioanalysis by LC-MS/MS. Journal of Pharmaceutical and Biomedical Analysis. 2007; 44:342–355. [PubMed: 17360141]

114. Michener JK, Thodey K, Liang JC, Smolke CD. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. Metab. Eng. 2012; 14:212–222. [PubMed: 21946159]

115. An GH, Bielich J, Auerbach R, Johnson EA. Isolation and characterization of carotenoid hyperproducing mutants of yeast by flow cytometry and cell sorting. Biotechnology. 1991; 9:70–73. [PubMed: 1367215]

116. Cho EJ, Lee J-W, Ellington AD. Applications of aptamers as sensors. Annu. Rev. Anal. Chem. (Palo Alto. Calif). 2009; 2:241–264. [PubMed: 20636061]

117. Desai SK, Gallivan JP. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. J. Am. Chem. Soc. 2004; 126:13247–13254. [PubMed: 15479078]

118. Ogawa A, Maeda M. An artificial aptazyme-based riboswitch and its cascading system in E. coli. Chembiochem. 2008; 9:206–9. [PubMed: 18098257]

119. Suess B, et al. Conditional gene expression by controlling translation with tetracycline-binding aptamers. Nucleic Acids Research. 2003; 31:1853–1858. [PubMed: 12655001]

120. Wieland M, Hartig JS. Improved aptazyme design and in vivo screening enable riboswitching in bacteria. Angew. Chemie - Int. Ed. 2008; 47:2604–2607.

121. Wachsmuth M, Findeiss S, Weissheimer N, Stadler PF, Morl M. De novo design of a synthetic riboswitch that regulates transcription termination. Nucleic Acids Res. 2012; :1–11. DOI: 10.1093/nar/gks1330

122. Buskirk AR, Landrigan A, Liu DR. Engineering a ligand-dependent RNA transcriptional activator. Chem. Biol. 2004; 11:1157–1163. [PubMed: 15324817]

123. Thompson KM, Syrett HA, Knudsen SM, Ellington AD. Group I aptazymes as genetic regulatory switches. BMC Biotechnol. 2002; 2:21. [PubMed: 12466025]

124. Weigand JE, Suess B. Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. Nucleic Acids Res. 2007; 35:4179–4185. [PubMed: 17567606]

125. Michener JK, Smolke CD. High-throughput enzyme evolution in Saccharomyces cerevisiae using a synthetic RNA switch. Metab. Eng. 2012; 14:306–316. [PubMed: 22554528]

126. Weigand JE, et al. Screening for engineered neomycin riboswitches that control translation initiation. RNA. 2008; 14:89–97. [PubMed: 18000033]

127. Schoukroun-Barnes LR, Wagan S, White RJ. Enhancing the analytical performance of electrochemical RNA aptamer-based sensors for sensitive detection of aminoglycoside antibiotics. Anal. Chem. 2014; 86:1131–1137. [PubMed: 24377296]

128. Farjami E, et al. RNA aptamer-based electrochemical biosensor for selective and label-free analysis of dopamine. Anal. Chem. 2013; 85:121–128. [PubMed: 23210972]

129. Chen J, Fang Z, Liu J, Zeng L. A simple and rapid biosensor for ochratoxin A based on a structure-switching signaling aptamer. Food Control. 2012; 25:555–560.

130. Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. Nat. Struct. Biol. 2000; 7:215–219. [PubMed: 10700280]

131. Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. U. S. A. 1992; 89:5547–5551. [PubMed: 1319065]

132. Grkovic S, Hardie KM, Brown MH, Skurray RA. Interactions of the QacR Multidrug-Binding Protein with Structurally Diverse Ligands: Implications for the Evolution of the Binding Pocket. Biochemistry. 2003; 42:15226–15236. [PubMed: 14690433]

133. Tahlan K, et al. Initiation of actinorhodin export in Streptomyces coelicolor. Mol. Microbiol. 2007; 63:951–961. [PubMed: 17338074]

134. Ramos JL, et al. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. 2005; 69:326–356. [PubMed: 15944459]

135. Mohn WW, Garmendia J, Galvao TC, De Lorenzo V. Surveying biotransformations with à la carte genetic traps: Translating dehydrochlorination of lindane (gamma-hexachlorocyclohexane) into lacZ-based phenotypes. Environ. Microbiol. 2006; 8:546–555. [PubMed: 16478460]

136. Van Sint Fiet S, van Beilen JB, Witholt B. Selection of biocatalysts for chemical synthesis. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:1693–1698. [PubMed: 16446453]

137. Mustafi N, Grünberger A, Kohlheyer D, Bott M, Frunzke J. The development and application of a single-cell biosensor for the detection of l-methionine and branched-chain amino acids. Metab. Eng. 2012; 14:449–457. [PubMed: 22583745]

138. Tang SY, Cirino PC. Design and application of a mevalonate-responsive regulatory protein. Angew. Chemie - Int. Ed. 2011; 50:1084–1086.

139. Tang SY, et al. Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter. J. Am. Chem. Soc. 2013; 135:10099–10103. [PubMed: 23786422]

140. Buskirk AR, Ong Y-C, Gartner ZJ, Liu DR. Directed evolution of ligand dependence: small-molecule-activated protein splicing. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:10505–10510. [PubMed: 15247421]

141. Peck SH, Chen I, Liu DR. Directed evolution of a small-molecule-triggered intein with improved splicing properties in mammalian cells. Chem. Biol. 2011; 18:619–630. [PubMed: 21609843]

142. DeLoache WC, et al. An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. Nat. Chem. Biol. 2015; 11:465–471. [PubMed: 25984720]
143. Gandía-Herrero F. García-Carmona, F. Biosynthesis of betalains: Yellow and violet plant pigments. Trends Plant Sci. 2013; 18:334–343. [PubMed: 23395307]
144. Meyer A, et al. Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors. Nat. Chem. 2015; 7:673–678. [PubMed: 26201745]
145. Pfleger BF, Pitera DJ, Newman JD, Martin VJJ, Keasling JD. Microbial sensors for small molecules: Development of a mevalonate biosensor. Metab. Eng. 2007; 9:30–38. [PubMed: 17002894]
146. Bertels F, Merker H, Kost C. Design and characterization of auxotrophy-based amino acid biosensors. PLoS One. 2012; 7
147. Kalir S, et al. Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Science. 2001; 292:2080–2083. [PubMed: 11408658]
148. Temme K, et al. Induction and Relaxation Dynamics of the Regulatory Network Controlling the Type III Secretion System Encoded within Salmonella Pathogenicity Island 1. J. Mol. Biol. 2008; 377:47–61. [PubMed: 18242639]
149. Liu G, Chater KF, Chandra G, Niu G, Tan H. Molecular regulation of antibiotic biosynthesis in streptomyces. Microbiol. Mol. Biol. Rev. 2013; 77:112–43. [PubMed: 23471619]
150. Park SS, et al. Mass spectrometric screening of transcriptional regulators involved in antibiotic biosynthesis in streptomyces coelicolor A3(2). J. Ind. Microbiol. Biotechnol. 2009; 36:1073–1083. [PubMed: 19468766]
151. Chatterjee A, et al. Convergent transcription in the butyrolactone regulon in streptomyces coelicolor confers a bistable genetic switch for antibiotic biosynthesis. PLoS One. 2011; 6
152. Sherwood EJ, Bibb MJ. The antibiotic planosporicin coordinates its own production in the actinomycete Planomonospora alba. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:E2500–9. [PubMed: 23776227]
153. Chen Y, Smanski MJ, Shen B. Improvement of secondary metabolite production in Streptomyces by manipulating pathway regulation. Applied Microbiology and Biotechnology. 2010; 86:19–25. [PubMed: 20091304]
154. Smanski MJ, Peterson RM, Rajski SR, Shen B. Engineered streptomyces platensis strains that overproduce antibiotics platensimycin and platencin. Antimicrob. Agents Chemother. 2009; 53:1299–1304. [PubMed: 19164156]
155. Stevens, JT.; Carothers, JM. Designing RNA-Based Genetic Control Systems for E ffi cient Production from Engineered Metabolic Pathways. 2015.
156. Kushwaha M, Salis HM. A portable expression resource for engineering cross-species genetic circuits and pathways. Nat. Commun. 2015; 6:7832. [PubMed: 26184393]
157. Solomon KV, Sanders TM, Prather KLJ. A dynamic metabolite valve for the control of central carbon metabolism. Metab. Eng. 2012; 14:661–671. [PubMed: 23026120]
158. Brockman IM, Prather KLJ. Dynamic knockdown of E. coli central metabolism for redirecting fluxes of primary metabolites. Metab. Eng. 2015; 28:104–113. [PubMed: 25542851]
159. Nieselt K, et al. The dynamic architecture of the metabolic switch in Streptomyces coelicolor. BMC Genomics. 2010; 11:10. [PubMed: 20053288]
160. Ellis T, Wang X, Collins JJ. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. Nat. Biotechnol. 2009; 27:465–471. [PubMed: 19377462]
161. Anderson JC, Voigt CA, Arkin AP. Environmental signal integration by a modular AND gate. Mol. Syst. Biol. 2007; 3:133. [PubMed: 17700541]
162. Tamsir A, Tabor JJ, Voigt CA. Robust multicellular computing using genetically encoded NOR gates and chemical ‘wires’. Nature. 2011; 469:212–215. [PubMed: 21150903]
163. Moon TS, Lou C, Tamsir A, Stanton BC, Voigt CA. Genetic programs constructed from layered logic gates in single cells. Nature. 2012; doi: 10.1038/nature11516
164. Qi LS, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013; 152:1173–83. [PubMed: 23452860]

165. Gilbert LA, et al. XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013; 154
166. Piatek A, et al. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. Plant Biotechnol. J. 2014; n/a–n/a. doi: 10.1111/pbi.12284
167. Cai F, Axen SD, Kerfeld C. a. Evidence for the widespread distribution of CRISPR-Cas system in the Phylum Cyanobacteria. RNA Biol. 2013; 10:687–93. [PubMed: 23628889]
168. Seo Y-S, et al. Comparative genome analysis of rice-pathogenic Burkholderia provides insight into capacity to adapt to different environments and hosts. BMC Genomics. 2015; 16
169. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole G. a. The CRISPR/Cas adaptive immune system of Pseudomonas aeruginosa mediates resistance to naturally occurring and engineered phages. J. Bacteriol. 2012; 194:5728–5738. [PubMed: 22885297]
170. Wallace, R. a.; Black, WP.; Yang, X.; Yang, Z. A CRISPR with Roles in Myxococcus xanthus Development and Exopolysaccharide Production. J. Bacteriol. 2014; 196:4036–4043. [PubMed: 25201946]
171. Lv L, Ren Y, Chen J, Wu Q, Chen G. Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes , a case study : Controllable P (3HB- co -4HB) biosynthesis. Metab. Eng. 2015; 29:1–9. [PubMed: 25662836]
172. Zalatan JG, et al. Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. Cell. 2014; 160:339–350. [PubMed: 25533786]
173. Zhang Y, Perry K, Vinci V, Powell K. Genome shuffling leads to rapid phenotypic improvement in bacteria. Nature. 2002; 415:5–7.
174. Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LBA, Gill RT. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. Nat. Biotechnol. 2010; 28:856–862. [PubMed: 20639866]
175. Isaacs FJ, et al. Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. Science (80-.). 2011; 333:348–353.
176. Wang HH, et al. Genome-scale promoter engineering by coselection MAGE. Nat. Methods. 2012; 9:591–593. [PubMed: 22484848]
177. Wang HH, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009; 460:894–898. [PubMed: 19633652]
178. Wang HH, Church GM. Multiplexed genome engineering and genotyping methods: Applications for synthetic biology and metabolic engineering. Methods Enzymol. 2011; 498:409–426. [PubMed: 21601688]
179. Wang HH, et al. Multiplexed in vivo his-tagging of enzyme pathways for in vitro single-pot multienzyme catalysis. ACS Synth. Biol. 2012; 1:43–52. [PubMed: 22737598]
180. Bonde MT, et al. Direct mutagenesis of thousands of genomic targets using microarray-derived oligonucleotides. ACS Synth. Biol. 2015; 4:17–22. [PubMed: 24856730]
181. Van Pijkeren J-P, Britton R. a. High efficiency recombineering in lactic acid bacteria. Nucleic Acids Res. 2012; 40:e76. [PubMed: 22328729]
182. Van Kessel JC, Hatfull GF. Recombineering in Mycobacterium tuberculosis. Nat. Methods. 2007; 4:147–152. [PubMed: 17179933]
183. Binder S, Siedler S, Marienhagen J, Bott M, Eggeling L. Recombineering in Corynebacterium glutamicum combined with optical nanosensors: A general strategy for fast producer strain generation. Nucleic Acids Res. 2013; 41:6360–6369. [PubMed: 23630315]
184. Dicarlo JE, et al. Yeast Oligo-Mediated Genome Engineering (YOGE). ACS Synth. Biol. 2013; 2:741–749. [PubMed: 24160921]
185. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. Practical Streptomyces Genetics. John Innes Cent. Ltd. 2000; 529doi: 10.4016/28481.01
186. Sandoval NR, et al. Strategy for directing combinatorial genome engineering in Escherichia coli. Proceedings of the National Academy of Sciences U.S.A. 2012; 109:10540–10545.
187. Jiang W, Bikard D, Cox D, Zhang F, Marraffini L. a. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. 2013; 31:233–9. [PubMed: 23360965]

188. Jakočiūnas T, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab. Eng. 2015; 28:213–222. [PubMed: 25638686]
189. Citorik RJ, Mimee M, Lu TK. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat. Biotechnol. 2014; 32:1141–1145. [PubMed: 25240928]
190. Bikard D, et al. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat. Biotechnol. 2014; 32:1146–1150. [PubMed: 25282355]
191. Huang H, Zheng G, Jiang W, Hu H, Lu Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. Acta Biochim. Biophys. Sin. (Shanghai). 2015; 47:231–243. [PubMed: 25739462]
192. Esvelt KM, Wang HH. Genome-scale engineering for systems and synthetic biology. Mol. Syst. Biol. 2013; 9:641. [PubMed: 23340847]
193. Hemphill, J.; Borchardt, EK.; Brown, K.; Asokan, A.; Deiters, A. Optical Control of CRISPR / Cas9 Gene Editing. p. 1-11.
194. Konermann S, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. 2015; 517:2–11.
195. Selle K, Barrangou R. Harnessing CRISPR-Cas systems for bacterial genome editing. Trends Microbiol. 2015; 23:1–8. [PubMed: 25435136]
196. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science. 2011; 333:1307–1311. [PubMed: 21885784]
197. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. Science (80-.). 2014; 346:1–9.
198. Chen C, Fenk L. a. De Bono M. Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic Acids Res. 2013; 41
199. Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. 2015; 201420294. doi: 10.1073/pnas.1420294112
200. Bowater R, Doherty AJ. Making ends meet: Repairing breaks in bacterial DNA by non-homologous end-joining. PLoS Genet. 2006; 2:93–99.
201. Kotula JW, et al. Programmable bacteria detect and record an environmental signal in the mammalian gut. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:4838–43. [PubMed: 24639514]
202. Claesen J, Fischbach MA. Synthetic microbes as drug delivery systems. ACS Synth. Biol. 2014; doi: 10.1021/sb500258b
203. Fischbach MA, Bluestone JA, Lim WA. Cell-based therapeutics: the next pillar of medicine. Sci. Transl. Med. 2013; 5:179ps7.
204. Arkin AP. A wise consistency: Engineering biology for conformity, reliability, predictability. Curr. Opin. Chem. Biol. 2013; 17:893–901. [PubMed: 24268562]
205. Dayan FE, Cantrell CL, Duke SO. Natural products in crop protection. Bioorg. Med. Chem. 2009; 17:4022–4034. [PubMed: 19216080]
206. Donia MS, et al. A Systematic Analysis of Biosynthetic Gene Clusters in the Human Microbiome Reveals a Common Family of Antibiotics. Cell. 2014; 158:1402–1414. [PubMed: 25215495]
207. Zotchev SB, Sekurova ON, Katz L. Genome-based bioprospecting of microbes for new therapeutics. Curr. Opin. Biotechnol. 2012; 23:941–947. [PubMed: 22560158]
208. Bugni TS, et al. Marine natural product libraries for high-throughput screening and rapid drug discovery. J. Nat. Prod. 2008; 71:1095–1098. [PubMed: 18505284]
209. Koehn FE, Carter GT. The evolving role of natural products in drug discovery. Nat. Rev. Drug Discov. 2005; 4:206–220. [PubMed: 15729362]
210. Zimmermann M, Fischbach MA. A family of pyrazinone natural products from a conserved nonribosomal peptide synthetase in staphylococcus aureus. Chem. Biol. 2010; 17:925–930. [PubMed: 20851341]
211. Smanski MJ, et al. Dedicated ent-kaurene and ent-atiserene synthases for platensimycin and platencin biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 2011; 108:13498–13503. [PubMed: 21825154]

212. Sudek S, et al. Identification of the putative bryostatin polyketide synthase gene cluster from ‘Candidatus Endobugula sertula’, the uncultivated microbial symbiont of the marine bryozoan Bugula neritina. J. Nat. Prod. 2007; 70:67–74. [PubMed: 17253852]

213. Wong FT, Khosla C. Combinatorial biosynthesis of polyketides--a perspective. Curr. Opin. Chem. Biol. 2012; 16:117–23. [PubMed: 22342766]

214. Carlson, R. Time for new cost curves. 2014. http://www.synthesis.cc/2014/02/time-for-new-cost-curves-2014.html

215. Freestone TS, Zhao H. Combinatorial pathway engineering for optimized production of the anti-malarial FR900098. Biotechnol. Bioeng. 2015; n/a–n/a. doi: 10.1002/bit.25719

216. Rodríguez-García A, Combes P, Pérez-Redondo R, Smith M. C. a, Smith MCM. Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria Streptomyces. Nucleic Acids Res. 2005; 33:1–8. [PubMed: 15640442]

217. Dangel V, Westrich L, Smith MCM, Heide L, Gust B. Use of an inducible promoter for antibiotic production in a heterologous host. Appl. Microbiol. Biotechnol. 2010; 87:261–269. [PubMed: 20127238]

218. Stanton BC, et al. Systematic transfer of prokaryotic sensors and circuits to Mammalian cells. ACS Synth. Biol. 2014; doi: 10.1021/sb5002856

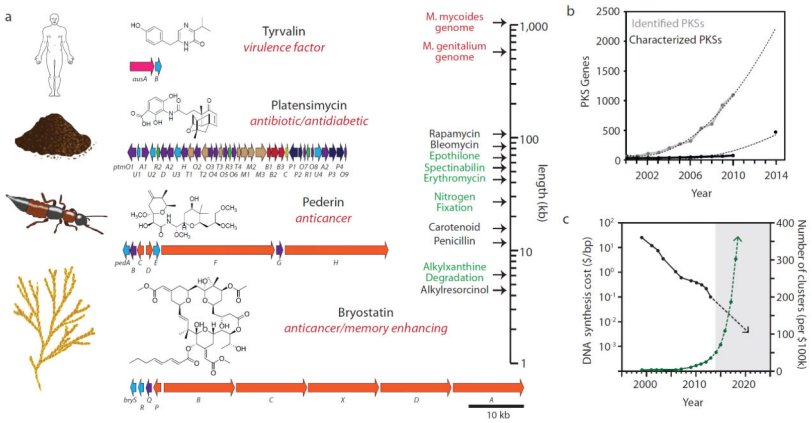

Figure 1.

Natural product biosynthetic gene clusters. (a) Representation of the diversity of size and complexity of NPs and their encoding gene clusters, including tyrvalin, a pyrazinone virulence factor from skin-associated staphylococci[210], platensimycin, a diterpenoid antibiotic from soil-dwelling *Streptomyces* isolates[211], pederin, a polyketide anticancer agent produced by an uncultivated symbiont of the Paederus spp. beetles[21], and bryostatin, a macrocyclic lactone anticancer agent produced by a symbiont of a marine bryozoan[212]. Approximate sizes of BGCs for select NPs (black), along with noteworthy examples of large systems that have been built with synthetic DNA technology in wild type (red) or re-designed (green) genetic architecture. (b) Widening gap of uncharacterized PKS enzymes (grey) compared to biochemically characterized PKSs (black) since 2000 (data to 2010 reproduced from Wong and Khosla[213]; 2014 data point from Marnix Medema, personal communication). Dashed line represents best fit to available data points. (c) Recent history of DNA synthesis costs[214] and the corresponding number of 50 kb gene clusters that could be synthesized with $100k. Dotted lines project to the future along the same trajectory of the past 15 years.

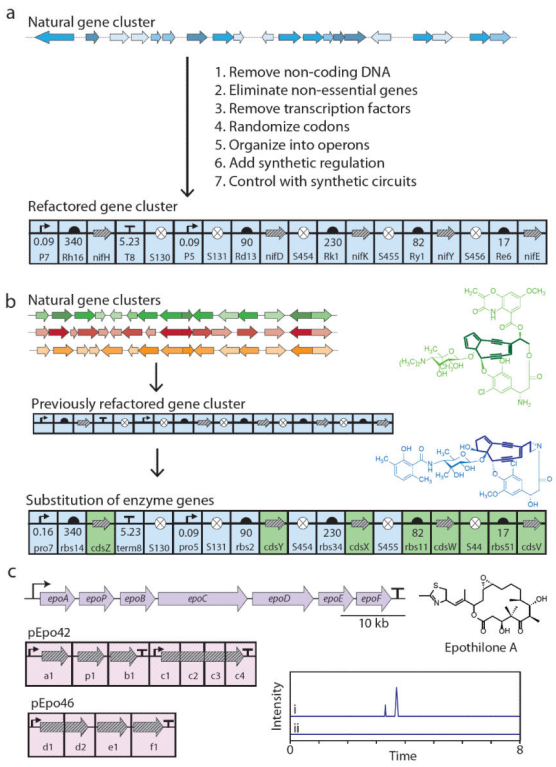

Figure 2.

Genetic refactoring. (a) Schematic outline of refactoring process, and (b) the streamlined refactoring of homologous gene clusters by substituting coding sequences. New homologous cluster and corresponding genetic parts are shown in green, and previously refactored cluster and parts are shown in blue. Bold lines on chemical structures show conserved core scaffold between two enediynes used as a hypothetical example. (c) Refactored *epo* gene cluster, built into a two plasmid system. Extracted ion chromatogram shows production of epothilones A and B from the refactored gene cluster introduced to *M. xanthus* (i), but not from the wild-type host (ii)[63].

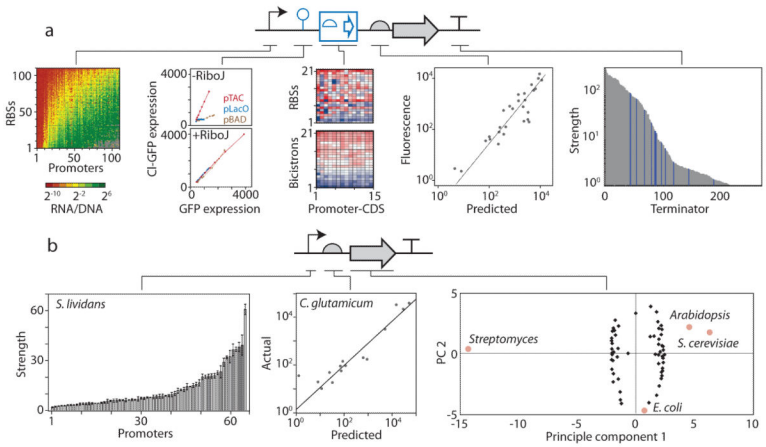

Figure 3.

Genetic parts for controlling gene expression levels. (a) Characterization of genetic parts in *E. coli*, including (from left to right), promoter variants[30,31], ribosome insulators[76], bicistronic RBSs[32], computationally designed RBSs[34], and synthetic and natural terminators[33]. (b) Genetic parts for engineering NP-producing organisms, including promoter variants[37–39], computationally designed RBSs[36], and codon-optimized CDS parts[74].

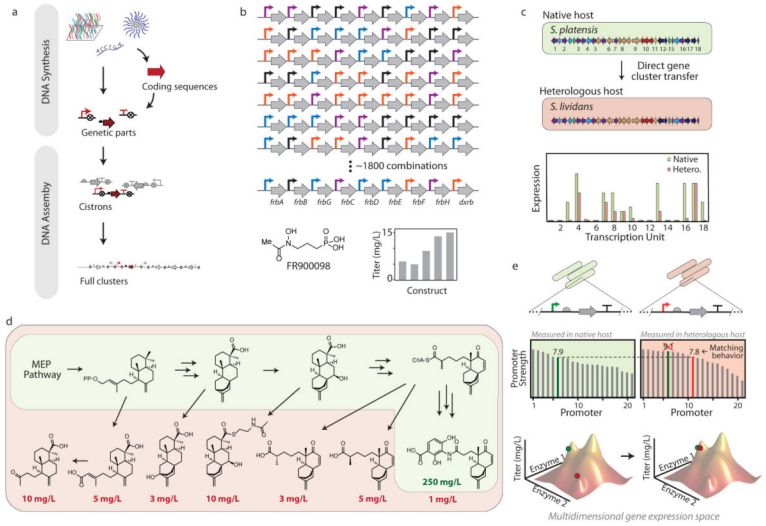

Figure 4.

Exploiting refactored genetics for host transfer of multi-gene devices. (a) Schematic representation of a DNA synthesis and assembly pipeline, wherein genetic parts are constructed from synthetic oligonucleotides and then assembled into unique combinations. (b) High-throughput library design of permuted gene clusters for antimalarial phosphonate FR900098. Bar graph shows characterized titers from constructs selected from iterative libraries, with successive libraries from left to right[215]. (c) Experimental design for heterologous expression of *ptn* gene cluster and RT-PCR results for each operon in native and wild-type hosts[106]. (d) The proposed platencin biosynthetic pathway, along with several shunt metabolites isolated from a heterologous expression strain. Values show in red are titers in heterologous host, while those shown in green are titers in the native producer. (e) Illustration of behavior-matching via part replacement during host transfer. Graphs represent empirical characterization of genetic parts in native host (green), and new host (red). Landscape graphs show effect on gene clusters performance, as measured by titer of final metabolite, in a multivariate system.

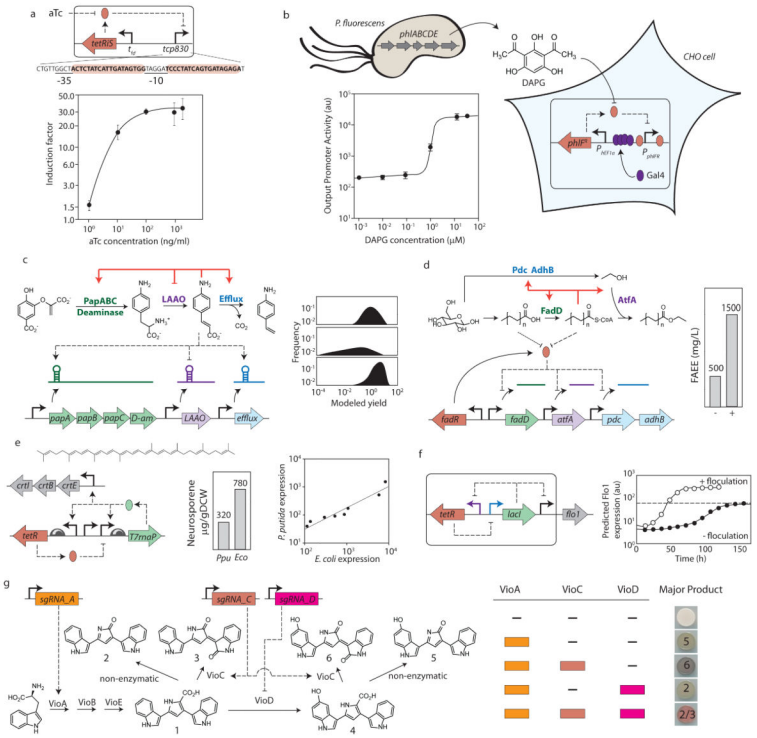

Figure 5.

Advanced regulation relevant to NP biosynthesis. Examples include (a) inducible promoters for NP producing organisms[216,217], (b) a mammalian genetic circuit responsive to a bacterial metabolite[218], (c) dynamic modeling results for a synthetic pathway for *para*-aminostyrene production[155], (d) a dynamic feedback/feedforward circuit for monitoring fatty acid ethyl ester production in *E. coli*[54], (e) a resource allocation system for controlling transcription of a heterologous neurosporene operon in different hosts[156], (f) a genetic reset timer for controlled sedimentation in yeast[160], and (g) multiplexed transcriptional control of the violacein biosynthetic pathway using CRISPRi/CRISPRa[172]. For dynamic modeling example (c), graphs show frequencies of expected yields for designs with static regulation (top), dynamic regulation (middle), or for the particular pattern of dynamic regulation pictured at left (bottom).

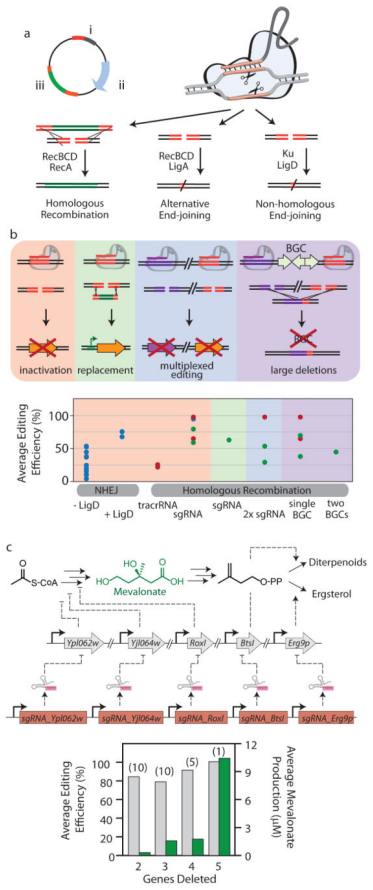

Figure 6.

Multiplexed genome editing with CRISPR/Cas9. (a) Minimal genome editing construct design, including (i) sgRNA, (ii) *S. pyogenes* Cas9, and (iii) optional 'repair fragment'. Three routes to DNA repair are shown, including homologous recombination (HR, left), alternative end-joining (AEJ, center), and non-homologous end-joining (NHEJ, center)[195]. (b) Applications of CRISPR-mediated genome editing in *Streptomyces*. Graph at bottom shows reported efficiencies for experiments grouped by application with background color matching illustrations above. Protocol differences are labeled below graph, and data points are colored according to published study (blue[55], red[57], green[191]). (c) Example of multiplexed CRISPR editing for engineering mevalonate levels. Bar graph at bottom shows editing efficiency (grey) and mevalonate levels (green), averaged across multiple different combinations of gene deletions (number of combinations indicated in parentheses)[188].

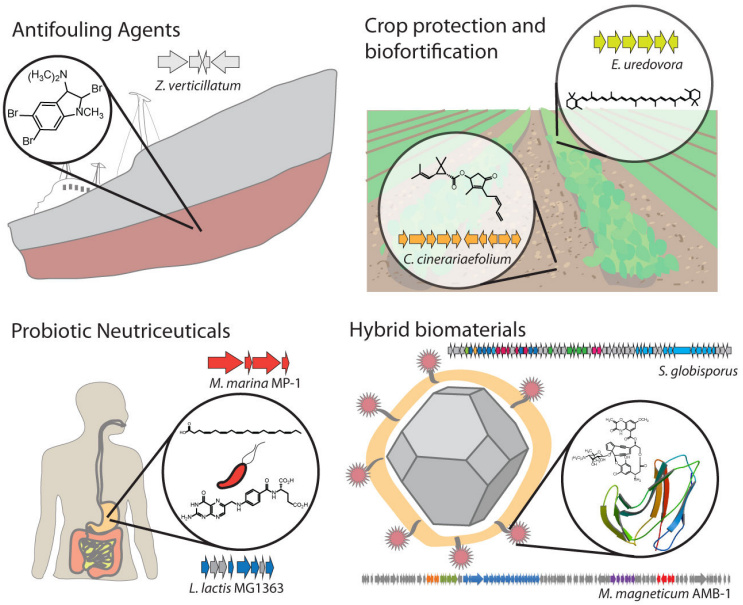

Figure 7.

Diverse applications of engineering NP biosynthesis. Structures of NPs are shown alongside a representation of their BGCs with native producing organisms noted.