Synthetic biology to access and expand nature's chemical diversity

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

Abstract | Bacterial genomes encode the biosynthetic potential to produce hundreds of thousands of complex molecules with diverse applications, from medicine to agriculture and materials. Accessing these natural products promises to reinvigorate drug discovery pipelines and provide novel routes to synthesize complex chemicals. The pathways leading to the production of these molecules often comprise dozens of genes spanning large areas of the genome and are controlled by complex regulatory networks with some of the most interesting molecules being produced by non-model organisms. In this Review, we discuss how advances in synthetic biology — including novel DNA construction technologies, the use of genetic parts for the precise control of expression and for synthetic regulatory circuits — and multiplexed genome engineering can be used to optimize the design and synthesis of pathways that produce natural products.

[1]*Department of Biochemistry, Molecular Biology, and Biophysics and the BioTechnology Institute, University of Minnesota Twin Cities, Saint Paul, Minnesota 55108, USA.*
[2]*Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.*
[3]*Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, California 94158, USA.*
[4]*Departments of Chemistry and Molecular Therapeutics, The Scripps Research Institute, Jupiter, Florida 33458, USA.*
Correspondence to C.A.V.
cavoigt@gmail.com

doi:10.1038/nrmicro.2015.24
Published online 15 Feb 2016

Natural products are specialized metabolites, produced by plants, animals and microorganisms, with diverse chemical structures and biological activities. These molecules are invaluable in the clinical setting, with half of the small-molecule drugs that were approved during the past three decades being derived from natural products[1]. Although these molecules are often used as antibiotics, cancer therapeutics and immunosuppressive agents, they are also used commercially as antivirals, anthelmintics, enzyme inhibitors, nutraceuticals (or 'health foods'), polymers, surfactants, bioherbicides and vaccines[2].

Claims that natural products are an inexhaustible resource[3] are based on the disparity between the staggering biological and chemical diversity in nature and the relatively low-throughput methods that are currently available to characterize these compounds. In the coming decades, advances in technology will close this gap and enable a more systematic characterization of natural products. Improving bioinformatic methods, combined with the dramatic increase in sequenced genomes, is shedding light on the potential number of undiscovered natural products[4–10]. For example, in *Streptomyces* spp. alone, conservative estimates put the number of natural products produced by these species at 150,000, of which <5% have been characterized[11]. Bioinformatic investigations of hundreds of genomes across many genera have estimated that there are hundreds of thousands of natural products[12] and the inclusion of less-studied classes, such as carbohydrates and lipids, substantially adds to the number of estimated molecules[6]. Some natural products remain uncharacterized as they are not produced by the respective microorganism during laboratory cultivation, are produced at sufficiently low levels to evade detection, or are encoded in strains that have not yet been investigated, or any combination of these three factors.

In many organisms that produce natural products, all of the genes required for the regulation, biosynthesis, export and protection against any harmful effects of the product are colocalized in the genome in compact biosynthetic gene clusters (BGCs) (FIG. 1a). These BGCs can range in size from several kilobases to more than 100 kilobases. BGCs are identified from genome sequence data based primarily on the presence of biosynthetic genes for one of the major classes of secondary metabolite, including polyketide synthases (PKSs), non-ribosomal peptide synthetases (NRPSs), terpene synthases and ribosomally synthesized and post-translationally modified peptide (RiPP) biosynthetic genes. The colocalization of genes that are involved in the production of a particular metabolite into BGCs greatly facilitates the characterization of biosynthetic pathways, as the effects of genetic manipulation can be readily determined using analytical chemistry. Because of this, the BGCs of natural products have been a rich resource for the discovery of new enzymes.

Currently, our ability to mine bacterial genomes, that is to assess the capability of microorganisms to produce natural products, is unable to keep pace with the identification of new BGCs by DNA sequencing and bioinformatics (FIG. 1b). Therefore, the most substantial issues

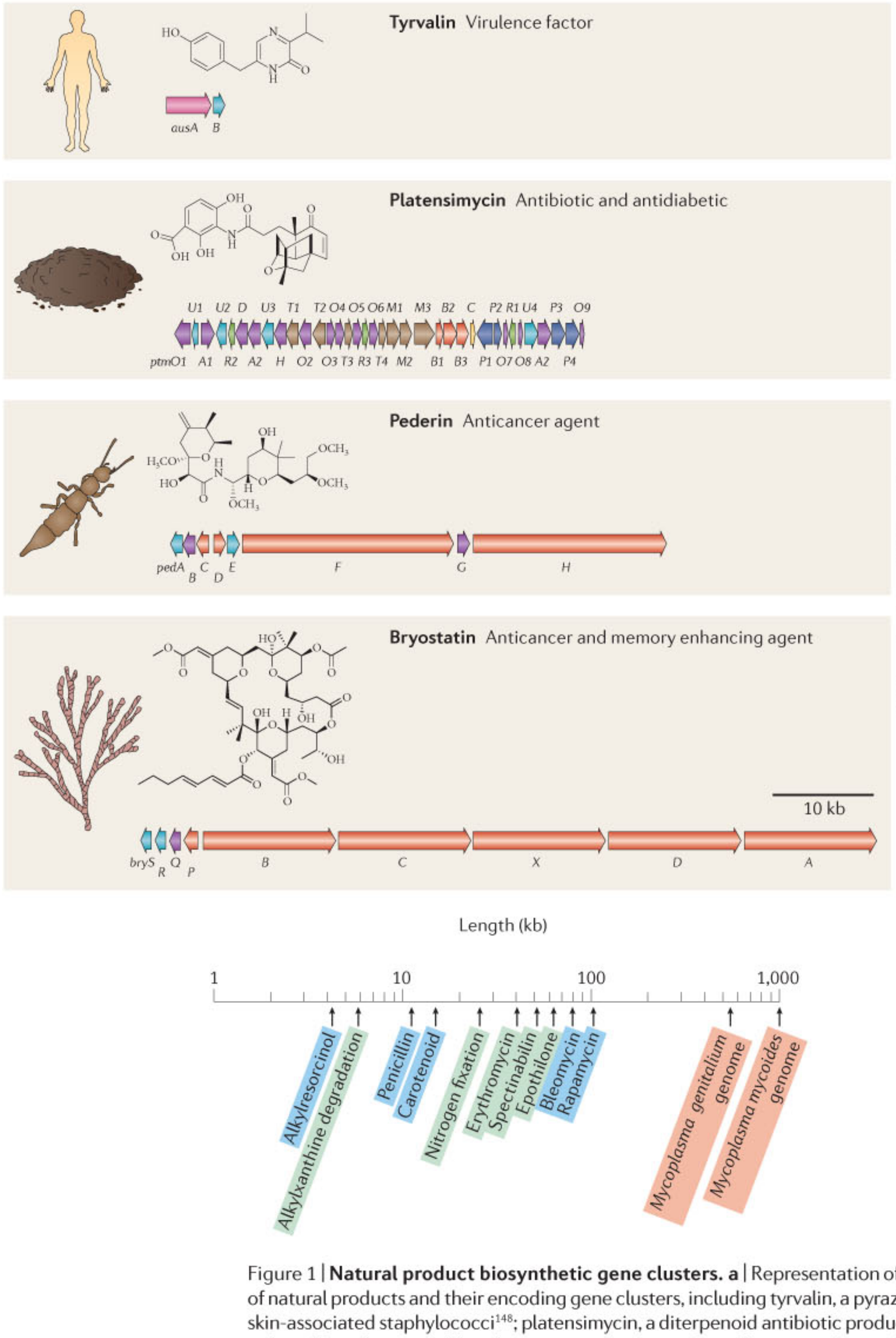


Figure 1 | **Natural product biosynthetic gene clusters. a** | Representation of the diversity of size and the complexity of natural products and their encoding gene clusters, including tyrvalin, a pyrazinone virulence factor produced by skin-associated staphylococci[148]; platensimycin, a diterpenoid antibiotic produced by soil-dwelling *Streptomyces* isolates[149]; pederin, a polyketide anticancer agent produced by an uncultivated symbiont of the *Paederus* spp. beetles[14]; and bryostatin, a macrocyclic lactone anticancer agent produced by a symbiont of a marine bryozoan[150]. The approximate sizes of the biosynthetic gene clusters (BGCs) for select natural products (blue), along with noteworthy examples of large systems that have been built with synthetic DNA technology in wild-type (red) or redesigned (green) genetic architecture, are shown at the bottom of the panel. **b** | The widening gap of uncharacterized polyketide synthases (PKSs) compared with characterized PKSs that have been experimentally linked to the production of a natural product (data until 2010 from REF. 151; data point from 2014 from REF. 152). The dashed line represents the best fit for the available data points. **c** | Recent history of the cost of DNA synthesis[153] (left axis) and the corresponding number of 25 kb gene clusters that could be synthesized with US$100,000 (right axis). The dotted lines predict future costs (shaded area) using the same trajectory that has been observed for the past 15 years.

that remain do not involve the discovery of new BGCs, but rather the design and the construction of pathways that lead to the optimal production of natural products. Notably, the cost of DNA sequencing and DNA synthesis continues to decrease (FIG. 1c) and future advances are projected to quickly make it possible to construct the DNA that encodes the components of many pathways that are responsible for the production of natural products. However, it is still challenging to rationally design a DNA sequence for a large pathway that will be functional in a model production host. Engineering the biosynthesis of natural products is still difficult for several reasons. Factors, such as transcription, translation, protein–protein interactions, availability of cofactors and precursors, export and resistance, all need to be accounted for. In addition, many of the organisms that harbour these BGCs are difficult to manipulate or cultivate[13,14], and the transfer of a BGC to a new host, for which there are successful examples[15], is by no means trivial.

Engineering the biosynthesis of natural products uses tools from various subfields in genetic engineering and chemistry. Protein engineering involves modifying the properties of individual proteins, including the activity, specificity and stability of enzymes. For example, the domains of different proteins can be recombined to diversify the products of large PKS and NRPS 'assembly lines' (REF. 16). Metabolic engineering is focused on understanding how several enzymes assemble into a pathway and how this affects fluxes through the metabolic network of the host[17]. For example, the flux of a precursor to a natural product could be boosted as part of optimizing the titre. Applied microbiology and strain engineering have been crucial for identifying organisms that produce high titres of natural products and for optimizing their productivity through processes such as random chemical mutagenesis[18]. Synthetic biology has focused on tools to accelerate and increase the scale of genetic engineering[19,20].

In this Review, we cover the recent advances in synthetic biology, taking into consideration how they will affect the field of natural products. We focus on different aspects of the design and construction of biosynthetic pathways, including libraries of genetic parts that can be used to construct BGCs, strategies used for the precise control of gene expression, and the use of large-scale construction technologies and genome editing techniques that enable the assembly of these synthetic pathways and their expression in engineered hosts. Other areas, including methods for the identification of natural products, the manipulation of global and pathway-specific regulators, the prioritization of BGCs identified by genome sequencing efforts and increasing flux to precursors through metabolic engineering, will not be discussed in detail as these have been reviewed recently[21–25].

Pathway design

Reducing genetics to genetic parts. The BGCs of natural products are large and unwieldy[26]. They can consist of several dozen genes, arranged in one or many operons facing either direction; their expression often relies on regulatory elements that are overlapping with or imbedded in neighbouring genes; and they are under several layers of complex regulation, including transcriptional and translational control. These factors make it difficult to change the expression level of individual genes within a natural BGC. This organization contrasts with 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[27].

Applying a parts-level approach to multigene systems is facilitated by 'refactoring', in which the natural genetics are rewritten to make the systems more amenable to engineering efforts[28,29] (FIG. 2a). During the refactoring process, known non-essential genes and natural regulatory elements, including promoters, ribosome binding sites (RBSs) and transcription factors, are removed. The remaining essential genes are codon-randomized *in silico* to remove any uncharacterized regulatory elements that are embedded within the coding DNA sequences (CDSs). Next, the codon-randomized CDSs are constructed using DNA synthesis. Synthetic regulation, in the form of well-characterized genetic parts, is added to control transcription and translation. This process has been demonstrated by the refactoring of a phage genome[29] and a bacterial gene cluster for nitrogen fixation[28].

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, enabling the synthetic control of gene expression. Second, it leads to a modular genetic architecture, which facilitates part-swapping and combinatorial optimization (see below). For example, refactoring could facilitate the creation of diverse compounds by substituting variations of enzymes from homologous clusters. This is especially relevant for families of natural products that consist of the differential modification of a common core scaffold, such as enediyne antitumour agents (FIG. 2b). Exploring core-modifications in a combinatorial fashion could inform structure–activity-relationship studies and lead to the efficient production of new analogues for preclinical evaluation. Furthermore, refactoring permits the exchange of regulatory parts, enabling the expression level of the enzymes in these pathways to be controlled (see below). Therefore, refactored BGCs can undergo the extensive swapping of genetic parts to optimize expression levels and increase titres.

The concept of refactoring has begun to be applied to BGCs[30–32] (FIG. 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. For example, the replacement of each of the native promoters in the spectinabilin gene cluster from *Streptomyces orinoci* with characterized strong promoters improved the production of spectinabilin in a heterologous host from undetectable amounts to 105 µg per litre[30]. The addition of strong promoters, known to function in a heterologous host, upstream of each gene in an uncharacterized BGC led to the discovery of new polycyclic tetramate macrolactam natural products from *Streptomyces griseus*[32]. The large 56 kb epothilone gene cluster was reconstructed using a modular-assembly

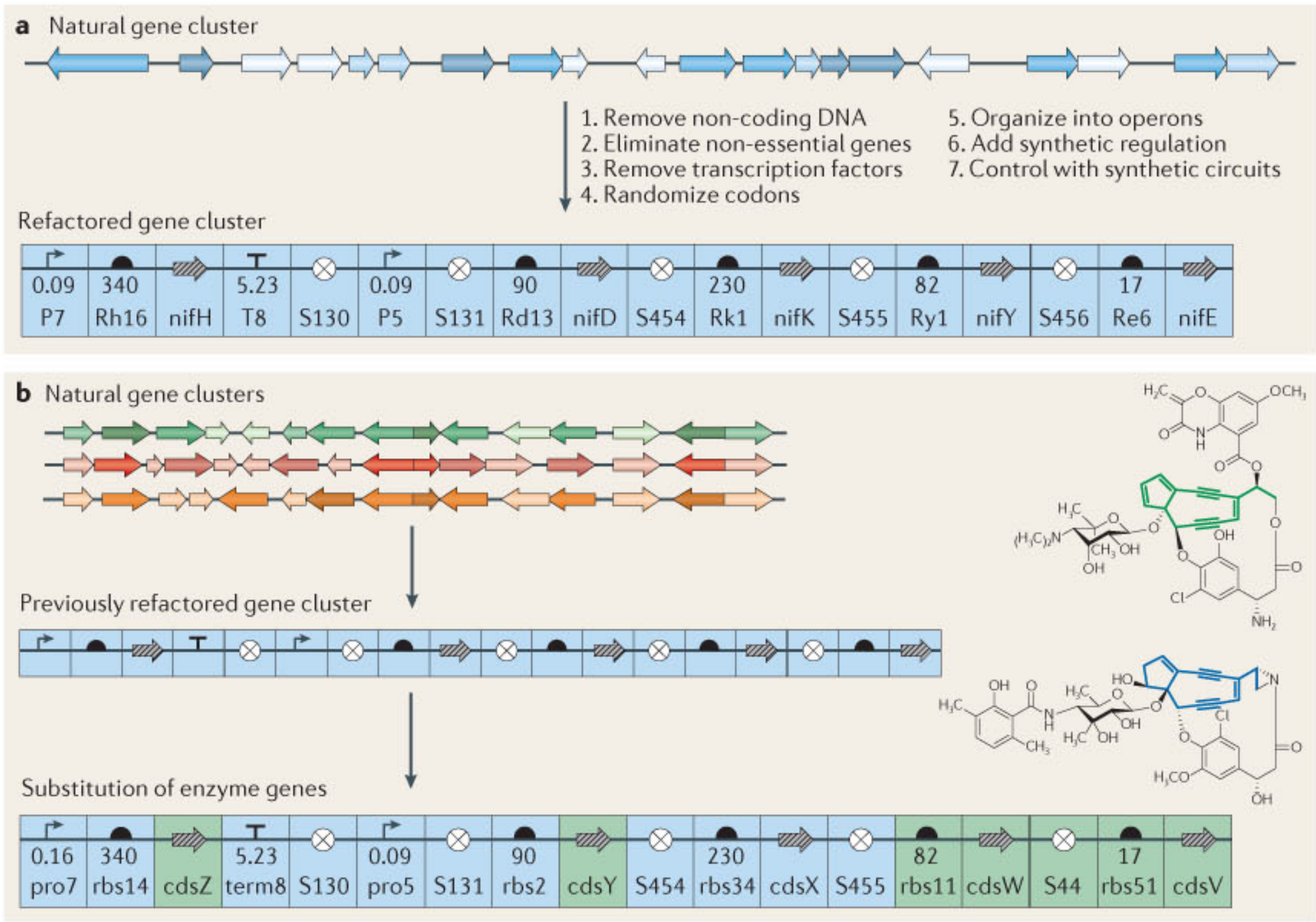
Polyketide
One of a family of natural products that share a common biosynthetic pathway through the decarboxylative condensation of substituted malonyl-CoA-derived extender units and acyl-CoA starter units on polyketide synthase enzymes.

Non-ribosomal peptide
One of a family of natural products that share a common biosynthetic pathway through the condensation of proteogenic or non-proteogenic amino acids on modular non-ribosomal peptide synthetase enzymes.

Terpene
One of a family of natural products that share a common biosynthetic pathway through the polymerization of branched five-carbon isoprene units and cyclization by terpene synthases.

Ribosomally synthesized and post-translationally modified peptide
(RiPP). One of a family of natural products, including the lanthipeptides, bacteriocins, and thiazole-modified or oxazole-modified microcins, that share a common biosynthetic pathway through the translation of an mRNA-encoded core peptide and subsequent modification.

Random chemical mutagenesis
A process by which cells or organisms are exposed to chemical mutagens to introduce mutations at random locations in the genome.



5′ untranslated region
(5′ UTR). The untranslated region of an mRNA transcript that is upstream of the start codon. The sequence of the 5′ UTR can influence translation initiation and mRNA stability.

Bicistrons
Transcripts that contain two coding DNA sequences (CDSs). For translational control, the first CDS encodes a short, non-functional peptide and is located immediately upstream of the ribosome binding site for the second CDS.

Actinorhodin
A benzoisochromanequinone polyketide pigment produced by *Streptomyces coelicolor*.

para-aminostyrene
An industrially relevant vinyl aromatic monomer with applications in materials and biomedicine.

approach that enabled the examination of the effect of the codon adaption index on heterologous expression[31]. The redesigned cluster achieved production titres of approximately 100 μg per litre in a heterologous host. In both cases, the modular DNA assembly scheme will facilitate future efforts to manipulate pathway regulation in a combinatorial fashion.

Precision control of gene expression. Many metabolic pathways and BGCs are highly sensitive to the level of gene expression, in which small changes can cause a loss of activity[33,34]. Recent work to create large libraries of regulatory parts has enabled the graded control of gene expression by many orders of magnitude[35–38] (FIG. 2d). For example, the RBS sequence affects the rate of translation initiation and thus the overall level of expression of a gene by many orders of magnitude. Important factors that contribute to the strength of a RBS include its affinity for the anti-Shine–Dalgarno loop of the small subunit rRNA, the space separating binding motifs and the initiation codon and secondary structural elements in the mRNA that prevent the ribosome from binding. A biophysical model of translation initiation that takes these factors into account enables new RBSs to be designed *de novo* to control gene expression over a 100,000-fold range[39]. This model extends to diverse bacterial phyla, including some with high-GC-content genomes[40] (FIG. 2e).

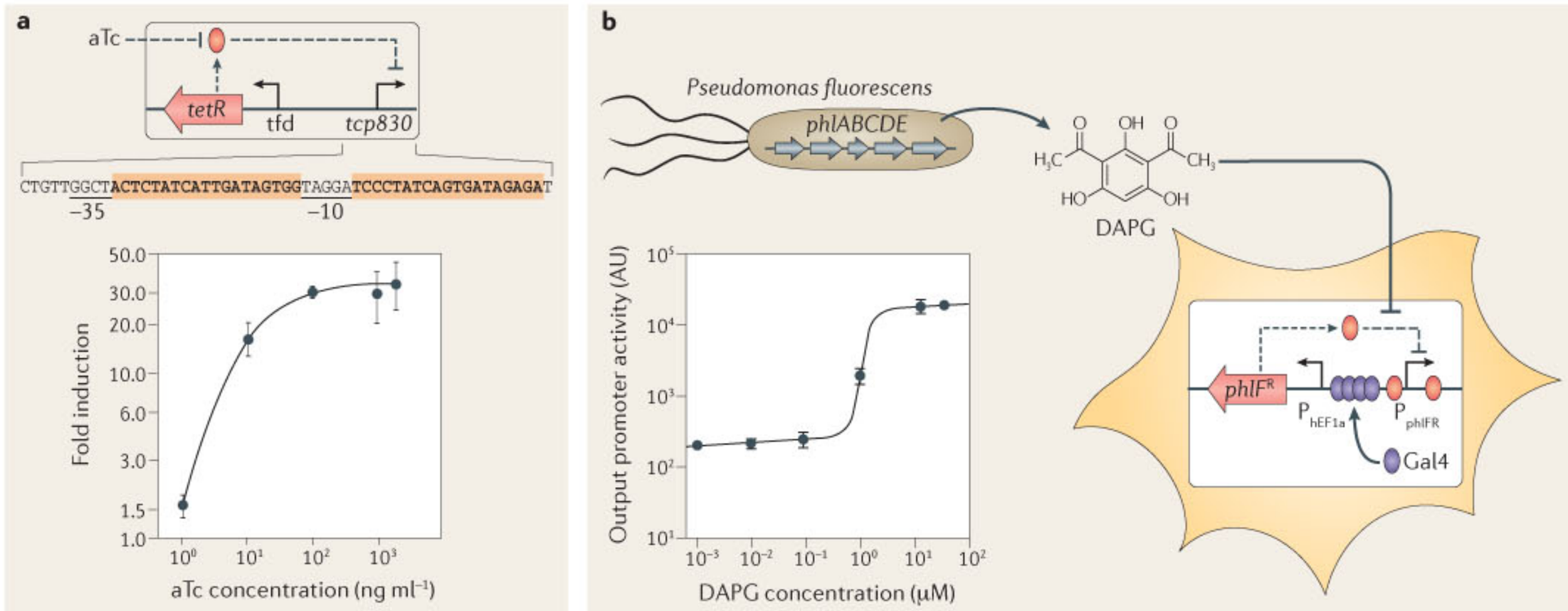
In synthetic biology, the concept of the 'expression cassette' has been expanded to include insulators that increase the reliability of part function in different genetic contexts[35,41–43] (FIG. 2d). This arose through observations that different combinations of parts — for example, promoters and RBSs — can lead to unpredictable levels of expression[35,44]. Some examples of these insulators include ribozymes that uncouple the promoter from the 5′ untranslated region (5′ UTR)[45,46], and bicistrons that uncouple the 5′ UTR from the RBS[43] (FIG. 2d). The ribozymes enzymatically cleave the variable 5′ terminus of mRNA transcripts that are generated from different promoters and thereby physically remove diverse sequences that could influence the levels of translation initiation. In bicistronic designs, a short leader peptide is translated immediately upstream of the designed RBS. It is thought that a ribosome that is actively translating on the 5′ UTR contributes to increased predictability by reducing variable secondary structure that would affect recruitment to the second RBS. These insulators enable the promoters and RBS to be swapped to vary levels of expression without affecting the behaviour of neighbouring parts. Similarly, long promoters and strong terminators have been developed to transcriptionally insulate the genes[36,47]. 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 is important for combinatorial optimization and in the exchange of parts to create chemical diversity.

Synthetic genetic circuits. The temporal control of gene expression is often important when constructing complex chemicals and materials[48,49]. In natural systems, this is implemented by regulatory networks, consisting of interacting proteins, RNA and DNA, that collectively work together to carry out information-processing operations. Synthetic genetic circuits have been created in which target behaviour is achieved by artificially linking the inputs and outputs of regulatory proteins to control cell behaviour. This has been used to create inducible systems (FIG. 3a,b), cascades, feedback and feedforward motifs (FIG. 3c–e), bistable switches (FIG. 3f), pulse generators, oscillators and logic gates[41,50]. Pathways that encode natural products reveal complex and intricate control mechanisms that include many of these same behaviours[51,52]. However, naturally evolved regulation is not required for high-level production under defined culture conditions. In fact, its disruption often leads to the improved production of natural products[53–55]. For example, deleting a component of a genetic feedback loop — *pspW* — from the planosporicin gene cluster results in a mutant that overexpresses biosynthetic genes and produces planosporicin earlier and to greater titres in both the native and heterologous hosts[53]. Therefore, replacing native regulation with synthetic circuits may implement the necessary feedback and dynamics without having the environmental control that can inhibit the production of natural products.

Feedback and feedforward regulation have been used to link the accumulation of early-stage pathway intermediates to the expression of downstream processing genes. Feedforward regulation is seen in many biosynthetic pathways of natural products. For example, in actinorhodin biosynthesis, the accumulation of pathway intermediates triggers the expression of an efflux pump to export the final product[56]. In this case, the feedforward motif helps to protect the cell from the deleterious effects of high concentrations of the accumulating antibiotic in the bacterial cell. Such regulation has begun to be included in the design of synthetic pathways. For example, modelling of a synthetic pathway for *para*-aminostyrene biosynthesis suggests that higher titres can be attained with dynamic regulation, which incorporates feedback and feedforward regulation, compared with static regulation[57] (FIG. 3c). The authors base the dynamic regulation models on a system of RNA aptamers that would respond to the key pathway intermediates

◀ Figure 2 | **Genetic refactoring and genetic parts for controlling levels of gene expression. a** | Schematic outline of the steps in the genetic refactoring process. **b** | The streamlined refactoring of homologous gene clusters by substituting coding sequences. New homologous clusters and their corresponding genetic parts are shown in green, orange and red, and the previously refactored gene cluster and genetic parts shown in blue. Bold lines on the chemical structures show a core scaffold that is conserved between two enediyne antitumour agents used as a hypothetical example. **c** | Refactored *epo* gene cluster, built into a two-plasmid system. An extracted ion chromatogram shows the production of epothilone A and epothilone B from the refactored gene cluster introduced into *Myxococcus xanthus* (see the inset chromatogram, top line), but not from the same *M. xanthus* strain lacking the gene cluster (see the inset chromatogram, bottom line)[31]. **d** | Characterization of the genetic parts in *Escherichia coli*, including (from left to right) promoter variants[35,42], ribosome insulators[45], bicistronic ribosome binding sites (RBSs)[43], computationally designed RBSs[39], and synthetic and natural terminators[36]. **e** | Genetic parts for engineering natural product-producing organisms, including promoter variants[46], ribosome insulators[46], computationally designed RBSs[40] and codon-optimized coding DNA sequence (CDS) parts[154]. PC, principal component.

para-aminophenylalanine or *para*-aminocinnamic acid to affect the stability of mRNA transcripts that encode biosynthetic or efflux genes[57]. In another example, using a fatty acid biosensor to add feedforward and feedback regulation into a synthetic fatty acid ester pathway enabled the production of biodiesel to be boosted to 28% of the theoretical maximal yield[58] (FIG. 3d). This system was designed using the fatty acid metabolism regulator protein (FadR), which is a transcriptional repressor that is allosterically inhibited by fatty acids and fatty



Undecylprodigiosin
A tripyrrole polyketide pigment produced by *Streptomyces coelicolor*.

acyl-coenzyme A. Using FadR-responsive promoters to control the transcription of late biosynthetic steps enabled a system to be constructed that dynamically responds to pathway intermediates to control the expression of the final biosynthetic enzymes, including the wax-ester synthase AtfA[58]. Additionally, positive feedback and negative feedback loops can be used to control the allocation of cellular resources to secondary metabolism[59] (FIG. 3e).

Synthetic circuits have also been used to act as metabolic 'control valves' that redirect carbon flux from primary pathways to secondary pathways[60,61]. For example, targeting the gene for glucokinase, a central enzyme in glucose utilization, with either antisense RNA or an inverting gene circuit, enabled glycolytic flux to be artificially controlled in an inducible manner[60]. With the constrained flux through glycolysis, the producing strain grew more slowly but was able to direct more glucose into an engineered biosynthetic pathway, thus increasing yields. Having 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* spp. prior to the production of antibiotics[62]. The expression of housekeeping genes, particularly those involved in translation, is markedly diminished before entering the stationary phase at which time the production of actinorhodin and undecylprodigiosin begins.

Control over the timing of gene expression can also be achieved with unstable genetic toggle switches, which reset slowly over time[63] (FIG. 3f). These reset-timing switches have the same mutual-repression motif that is used to construct bistable toggle switches[64]. To construct a reset timer, the relative expression level of each repressor is precisely tuned so that the system enters into a monostable regime, in which switch reset delay is dictated by the extent of imbalance between the expression levels of the two genes. Bistable switches are found in many organisms that produce natural products, including *Streptomyces coelicolor,* in which a switch controls the expression of the cryptic polyketide synthase gene cluster *cpk*[52]. Engineered timing circuits can be used to separate different phases of cell growth and the production of natural products, particularly in strains for which the regulatory connections to the natural metabolic switch have been severed[62].

Genetic logic gates enable several input signals to be integrated before a pathway switches on. There is some evidence of logic gates in natural product pathways — for example, the actinorhodin gene cluster is controlled by metabolite concentrations, stress responses and development programmes[51]. This can be used to turn on different sets of genes under varying environmental conditions with improved specificity when more signals are integrated. In synthetic biology, many logic gates have been constructed[41,65]. Connecting synthetic multiple-input circuits to natural product pathways could enable cells to sense cofactor levels, precursor abundance, dissolved oxygen content or ATP charge, before deciding whether to commit to the biosynthesis of the product. Such intracellular checkpoints could prevent the accumulation of unwanted intermediates and byproducts by cells that are not capable of generating the final product.

CRISPR interference (CRISPRi) has been developed as a powerful method to regulate gene expression, including in the genome for both bacteria and eukaryotes. It is based on the expression of a catalytically dead Cas9 (dCas9), which can be directed to a target when a single-guide RNA (sgRNA) is transcribed[66]. dCas9 can act as either a repressor or activator (CRISPRa) depending on the domains to which it is fused[67]. CRISPRi has been shown to function in *Escherichia coli*[66], fungi[67], actinobacteria[68] and plants[69]. Other organisms relevant to natural product biosynthesis, such as cyanobacteria, *Burkholderia* spp., pseudomonads and myxobacteria, have endogenous CRISPR–Cas systems but have not yet been used as hosts for CRISPRi regulation. CRISPRi has already been used to control metabolic fluxes through multiplexed gene repression of endogenous pathways in *E. coli*[70] and in heterologous pathways in yeast[71] (FIG. 3g). In yeast, guide RNAs were designed to upregulate *vioA* and *vioC* and to downregulate *vioD* in the violacein synthesis pathway. Combinatorial expression of the different sgRNAs in a host that expressed dCas9 provided a mechanism to control the output of the pathway with five unique chemical profiles for different combinations. In an example of how natural products research can influence synthetic biology, CRISPRa was made inducible by fusing the amino-terminal and carboxy-terminal halves of dCas9 and VP64 to the rapamycin binding domains, FKBP12-rapamycin binding (FRB) and FK506-binding protein (FKBP), respectively. This system requires the presence of a macrolide to bring the two fusion proteins together to activate transcription[72].

◂ Figure 3 | **Advanced regulation relevant to natural product biosynthesis.** **a** | Inducible promoters for organisms that produce natural products[155]. **b** | A mammalian genetic circuit that is responsive to a bacterial metabolite[156]. **c** | Dynamic modelling results for a synthetic pathway for *para*-aminostyrene production[57]. **d** | A dynamic feedback and feedforward circuit for monitoring fatty acid ethyl ester (FAEE) production in *Escherichia coli* [58], in which '+' indicates the best titre with the dynamic regulation and '–' represents the best titre without dynamic regulation. **e** | A resource allocation system for controlling the transcription of a heterologous neurosporene operon in different hosts[59]. **f** | A genetic reset timer for controlled sedimentation in yeast[63]. **g** | Multiplexed transcriptional control of the violacein biosynthetic pathway using CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa)[71]. For the example of dynamic modelling (part **c**), the graphs shows the frequencies of expected yields for designs with static regulation (top), dynamic regulation (middle) or for the particular pattern of dynamic regulation pictured on the left (bottom). aTc, anhydrotetracycline; DAPG, diacetylphloroglucinol; *E. coli* (*Eco*), *Escherichia coli*; LAAO, L-amino acid oxidase; DCW, dry-cell weight; *P. putida* (*Ppu*), *Pseudomonas putida*; sgRNA, single-guide RNA.

Pathway construction

High-throughput genetic optimization of multigene systems. Accessing new natural products from genome sequence information requires the BGCs contained in sequence databases to be converted into physical DNA constructs. High-throughput fabrication of DNA enables many designs to be tested in parallel, thereby increasing the probability of identifying a functional construct. This is beneficial for both chemical diversification by combinatorial biosynthesis[73–76] and for the genetic optimization of pathway performance[30,33]. Notably, synthetic biology can also contribute to high-throughput

strain improvement efforts, from the perspective of analytical methods, through the development of next generation biosensors.

There are two DNA construction technologies relevant to BGCs (FIG. 4a). The first is *de novo* synthesis, in which genes or entire gene clusters are chemically constructed, typically by synthesis companies[77]. The cost has decreased substantially in the past decade and it is possible to order hundreds of individual genes or full gene clusters[28,31]. Although the cost has decreased, constructing large clusters is still expensive, and constructing comprehensive sets of gene clusters out of sequence databases is prohibitive. However, because the main chemical scaffolds for RiPP molecules are products of ribosomal translation, this low-cost synthetic DNA can be used for the combinatorial generation of new derivatives by introducing simple point mutations into the precursor peptide genes[78,79].

The second DNA construction technology relevant to BGCs is DNA assembly, which constitutes the combination of genetic parts to generate a larger construct. For example, once the CDSs that comprise a particular biosynthetic pathway have been printed using DNA synthesis, or cloned from the native host, they can be stitched together with different combinations of promoters to permute gene expression levels. This strategy enables the construction of many gene cluster variants based on a set of re-used underlying parts[80]. This is substantially cheaper than constructing *de novo* gene clusters for each variant that is to be tested. Many assembly methods, including isothermal assembly, Golden Gate assembly, ligase cycling reactions, scarless stitching and recombination-based methods, are available and have been reviewed recently[77]. Automating these techniques using liquid-handling robots enables the construction of hundreds or thousands of permuted combinations[74,75,78,80,81].This approach was used to improve the level of production of the antimalarial phosphonate natural product FR900098 (REF. 82) (FIG. 4b).

Optimization of the many variables that affect gene expression in a large BGC will be important for improving titres of natural products and transferring pathways between hosts. Epistasis in the expression levels of biosynthetic genes suggests the importance of combinatorial methods that enable the optimization of more than one variable at a time[33]. Once pathways grow beyond a small handful of genes, complete exploration of the combinatorial gene expression space is impractical or impossible. However, several rational strategies exist for simplifying this problem, which is common to all multivariate systems. In a process called multivariate modular metabolic engineering, subsets of genes are grouped together into operons based on prior knowledge about the biosynthetic pathway — such as enzyme kinetics, order of reactions or pathway branching architecture[83]. For example, a pathway might be divided into two operons, one for early biosynthesis genes and one for late biosynthesis genes. By constraining the genetic design in this way, the combinatorial expression space is markedly reduced. In this example, only the two promoters driving expression of the early and late operons need to be varied. In another example of this approach, metabolic modelling was used to constrain the range of RBSs that were explored for a neurosporene biosynthetic pathway, enabling production to be tuned over a continuous range from 0–300 μg per gram of dry cell weight[40].

Combinatorial assembly has been applied to the optimization of several natural product pathways, as well as the creation of chemical diversity. Several examples of optimizing BGCs, including for the heterologous production of artemisinin[34], taxadiene[33] and opiate alkaloids[84], suggest the importance of combinatorial engineering, as optimal levels of production are not achieved when each gene is expressed as strongly as possible. Using a Gibson assembly protocol that was optimized for high GC content, the pathway-specific regulators of pristinamycin II biosynthesis were mutated in a combinatorial fashion to increase the level of pristinamycin II production to more than 1 gram per litre[85]. Similarly, an improved three-gene pathway for the production of catechin was created through combinatorial assembly using eight homologous biosynthetic genes from different plant species[86]. Certain classes of natural product — for example, indolocarbazoles — have been greatly expanded using combinatorial DNA assembly with more than 50 derivatives created to date[73,87]. Finally, combinatorial assembly has been applied to probe the design rules underlying large multimodular enzymes[74–76,78,88]. For example, promiscuous polyketide donor or acceptor modules have been identified[74].

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

However, the direct transfer of a BGC even between similar species can result in substantial changes in the timing and expression level of genes involved in biosynthetic pathways[89] (FIG. 4c). These host–context effects can result from unintended crosstalk or interactions with native regulatory proteins[90], from limitations in host resources available for expressing heterologous constructs[91], and from crosstalk with endogenous biosynthetic pathways. For example, transfer of the platencin gene cluster from *Streptomyces platensis* to the model host organism *Streptomyces lividans* resulted in the excess accumulation of shunt metabolites that are structurally related to pathway intermediates[89] (FIG. 4c,d). This was correlated with substantial changes in the patterns of gene expression between the native and engineered host, which suggests a possible causal relationship between suboptimal levels of gene expression and the performance of the biochemical pathway (FIG. 4d). Similarly, even moving multigene systems between different strains of the same species can negatively affect performance[90,92]. One way to transfer the BGC between hosts more effectively would be to replace its genetic parts with those known to function in the new species (FIG. 4e). The extensive reassignment of regulatory parts that this would require is much easier in refactored genetic systems. For this to become routine, there has to be an increase in the size of the characterized part libraries for natural-product-relevant species[37,38,93,94].

Gibson assembly
A restriction-enzyme-independent method for the joining of several DNA fragments in a single isothermal reaction.

Pristinamycin II
One of two structurally unrelated chemical components of the clinical antibiotic pristinamycin. Pristinamycin II is a depsipeptide antibiotic produced by *Streptomyces pristinaespiralis*.

Catechin
A plant flavonoid with antioxidant properties.

Indolocarbazoles
A family of five-ring heterocyclic aromatic compounds that share a common biosynthetic pathway from two tryptophan molecules.

Platencin
An antibiotic produced by *Streptomyces platensis* containing an amino-dihydroxybenzoic acid moiety fused to a modified diterpene core.

Shunt metabolites
Chemically modified intermediates of a biosynthetic pathway that can no longer proceed through the biosynthetic pathway.

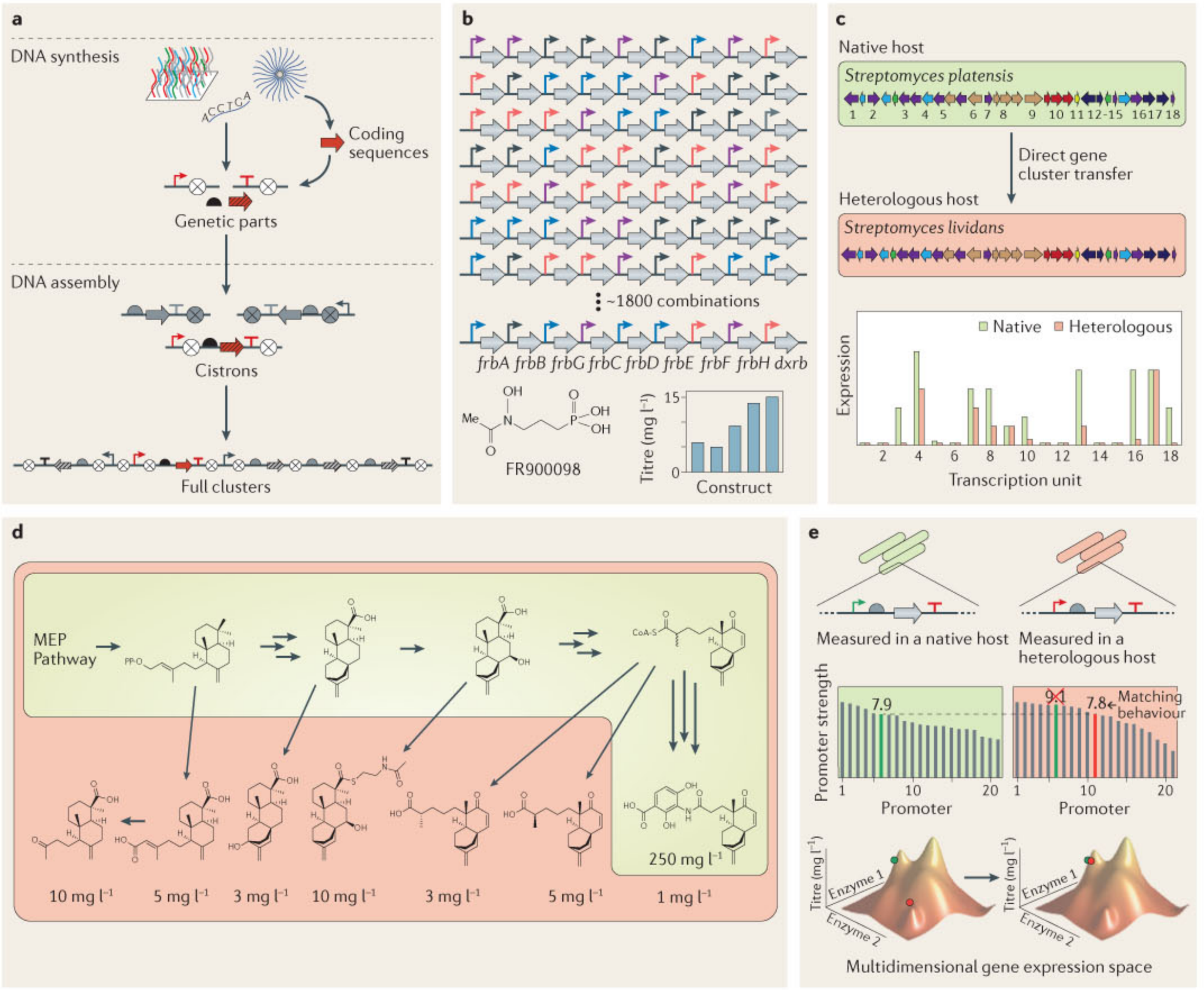


Figure 4 | **Using refactored systems for genetic optimization and host transfer. a** | Schematic representation of a DNA synthesis and assembly pipeline, wherein genetic parts are constructed from synthetic oligonucleotides and then assembled into unique combinations. **b** | High-throughput library design of permuted gene clusters for the antimalarial phosphonate FR900098. The bar graph shows characterized titres from constructs selected from iterative libraries, with successive libraries from left to right[82]. **c** | Experimental design for the heterologous expression of the *ptn* gene cluster and reverse transcription-PCR (RT-PCR) results for each operon in native and heterologous hosts[89]. **d** | The proposed platencin biosynthetic pathway along with several shunt metabolites isolated from a heterologous expression strain. Values shown in red are titres in a heterologous host, whereas the values shown in green are titres in the native producer. **e** | An illustration of behaviour-matching by part replacement during host transfer. The graphs represent the empirical characterization of genetic parts in a native host (green) and heterologous host (red). The landscape graphs show the effect on gene cluster performance, as measured by the titre of final metabolite in a multivariate system. MEP, methylerythritol phosphate.

Tylosin
A glycosylated macrolide antibiotic produced by *Streptomyces venezuelae*.

New tools for combinatorial genome-scale engineering. Mutations in the genome outside of the BGC are required to optimize the titres of a natural product. Some strains used in industrial production can achieve gram per litre quantities, and this is usually achieved through random mutagenesis of the genome and screening[18]. Originally, the genetic diversity was generated using techniques such as chemical mutagenesis, but this has become more sophisticated with improved methods in molecular biology. For example, the whole-genome shuffling of a tylosin-producing organism yielded the same sixfold improvement in 24,000 assays that had previously taken 20 years and one million assays to achieve using conventional methods[95]. Synthetic biology provides new techniques for generating genome diversity, from methods to replace parts or make defined mutations in a multiplexed manner to genome construction by *de novo* synthesis[96–99].

Multiplexed genome-engineering strategies provide the ability to precisely target hundreds of loci in a genome for overexpression or underexpression in

parallel. The first demonstration of massively multiplexed recombination-mediated genetic engineering (also known as recombineering) in *E. coli* used oligo-mediated allelic replacement[100]. By mimicking Okazaki fragments at the replication fork, exogenous single-stranded DNAs (ssDNAs) are able to anneal with the lagging strand of the genome acting as primers for DNA elongation. The ssDNAs then get incorporated, which leads to simultaneous directed mutagenesis at several sites in the genome[101]. Multiplexed automated genome engineering (MAGE) automates this process, which enables the combinatorial exploration of mutations in a continuously evolving population. For example, 20 endogenous *E. coli* genes were targeted to optimize the production of lycopene and billions of variants were screened[100]. MAGE has also been used to insert regulatory parts into the genome, such as N-terminal and C-terminal tags[102] and T7 RNA polymerase promoters[99,103]. Recombineering has been shown to be successful in diverse organisms, including lactic acid bacteria, mycobacteria, corynebacteria and fungi. However, success using MAGE in *Streptomyces* spp. will be challenging owing to the high GC content and the lack of characterization of mismatch repair in these species. Furthermore, the expression of some genes that are required for the technique (such as *bet*) is unlikely to produce functional proteins[104], and transformation is substantially less efficient in *Streptomyces* spp. than in *E. coli*. MAGE has already been shown to be successful in yeast and can be applied to pathways transferred to this host[105], and it may work in other fungi that are relevant to the research of natural products.

Trackable multiplex recombineering (TRMR) is a related method, which was developed to rapidly map the effects of more than 95% of *E. coli* genes onto specific traits[97]. TRMR uses array-based DNA synthesis to create barcoded oligonucleotides that target more than 4,000 genes for either overexpression or repression. Following phenotypic enrichment, deep-sequencing enables the targeted mutations to be quickly mapped to identify the causal mutations, which generates massive amounts of sequence-to-phenotype relationships[97,106]. In an example of combining TRMR with multiplexed recombineering, a total of 27 genome modifications that affected the growth of *E. coli* under industrially relevant conditions were targeted in a combinatorial fashion to generate mutants with improved growth rates in a target medium[106]. Importantly, the prevalence of positive and negative epistatic interactions between pairs of mutations suggests the need for high-throughput and multiplexed strategies that enable alternative hypotheses to be tested in parallel.

CRISPR techniques have revolutionized multiplexed genome engineering[68,107–115]. The Cas9 nuclease can be targeted to specific sequences by transcribing a sgRNA (FIG. 5a). This system has been shown to be successful in almost every organism that it has been tested in, including in bacteria (such as *Streptomyces* spp.)[68,108,112] (FIG. 5b), eukaryotes[109,111,116] and higher organisms (such as plants)[107]. The lack of a canonical non-homologous end joining (NHEJ) DNA repair system in some bacteria, including *Streptomyces* spp., lowers the efficiency of gene inactivation when using CRISPR–Cas9 alone[68]. However, the efficiency of gene inactivation can be increased to more than 75% by including a double-stranded 'repair fragment' that can close the double-stranded DNA break by homologous recombination[108], or by reconstituting a NHEJ pathway through heterologous expression of the ligase LigD during genome editing[68]. This system can be used to generate five mutations in a single step[111] (FIG. 5c), to knockout gene clusters of 31 kb[108] and insert DNA up to 9 kb in size[116].

Pathway screening

The design and construction of DNA libraries has accelerated to the point at which screening for natural products and making sense of the diversity is the bottleneck. Although bioassay-guided pre-screens or selections can reduce the number of strains that need to be investigated in detail[18], these methods do not replace direct measurements of product titre. Current analytical chemistry methods have reduced analysis time to less than 3 minutes per sample, but this still limits throughput to ~10^3 samples per day. As an alternative to analytical chemistry, *in vivo* biosensors translate information about a chemical signal — that is, the concentration of a natural product — into light or fluorescence-based output that can be measured by flow cytometry to screen thousands of genotypes per second[117]. For example, the inherent fluorescence of the carotenoid astaxanthin was used to track titres in single cells by flow cytometry and cell-sorting, enabling a 10,000-fold enrichment of over-producing strains compared with plate-based techniques[118].

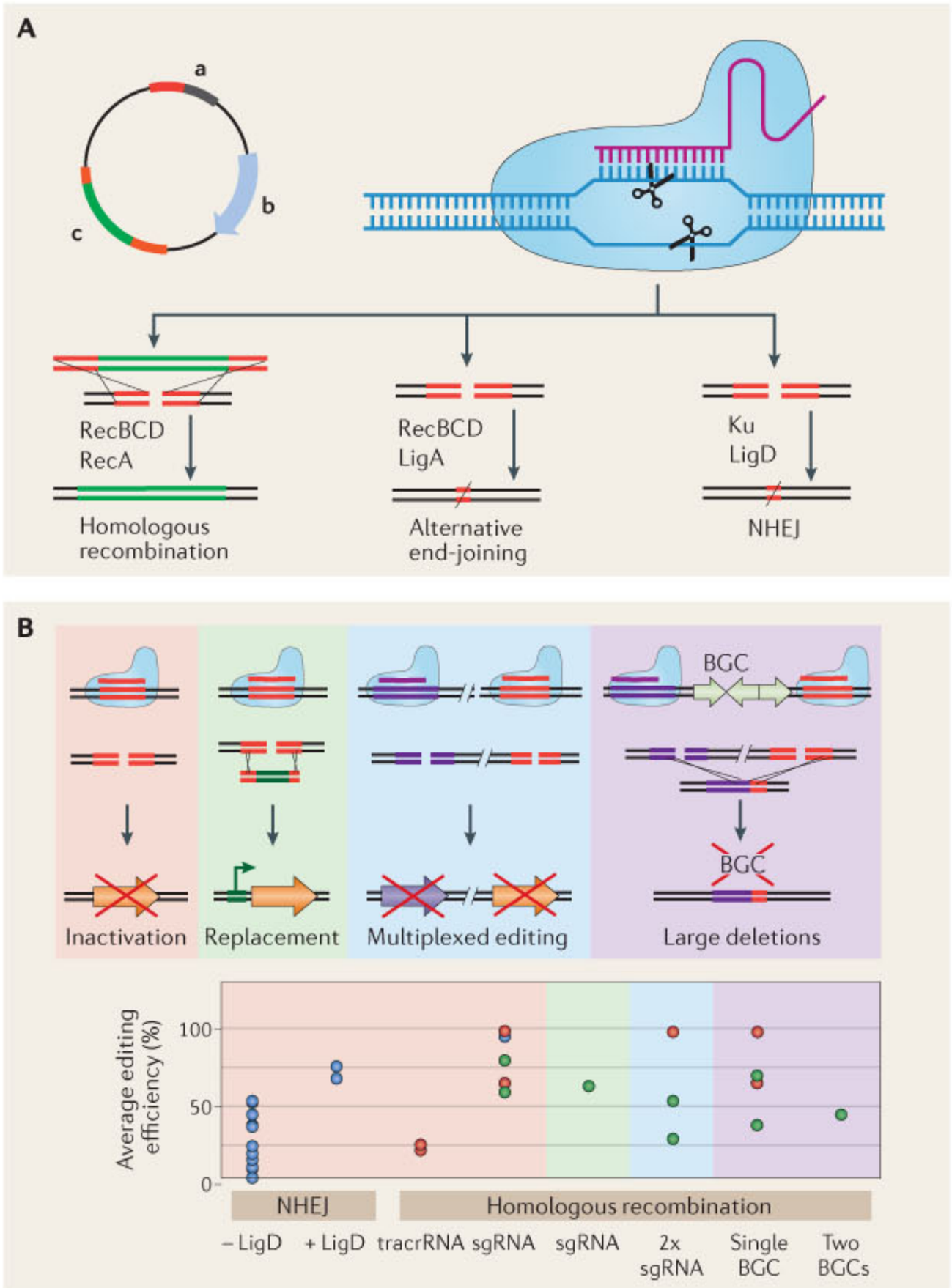
Intracellular biosensors can be broadly grouped into three categories, RNA-based, protein-based or enzyme-based, on the basis of to their biomolecular make-up and mechanism. RNA aptamers bind to small signalling molecules and have been linked to readouts including fluorescence, enzyme activity, cell mobility or viability[117,119]. There are diverse strategies for designing RNA biosensors[120–122]. RNA aptamers have been used to construct biosensors for natural products and intermediates, including theophylline[120], tetracycline[121], neomycin[123], tobramycin[124], dopamine[125] and ochratoxin A[126].

Protein biosensors function by transmitting molecular binding information into a measurable output, which is usually in the form of allosterically regulated transcriptional activators or repressors that control the expression of fluorescent proteins. Naturally evolved biosensors (see FIG. 3a,b for examples) can detect a wide range of molecular scaffolds, including tetracyclines[127], cationic lipids and plant alkaloids[128], and anthraquinones[56]. More than 4,000 Tet repressor proteins were identified in sequence databases but only a small fraction of the ligands are known[129]. Although protein biosensors have been found or engineered for several target molecules, including aromatic molecules[130] and branched-chain amino acids[131], these will need to be repurposed to sense new molecules for generalized use in natural product discovery and optimization pipelines. Cirino and colleagues genetically engineered the transcriptional activator AraC to recognize either mevalonate (a key precursor for isoprenoids)[132] or triacetic acid lactone (a simple polyketide product)[133]. Alternatively to modulating gene expression,

Okazaki fragments
Short, newly synthesized single-stranded DNA oligomers that are formed on the lagging template strand during DNA replication.

Lycopene
A symmetrical tetraterpene pigment formed by the tail-to-tail condensation of two molecules of geranylgeranyl diphosphate.

T7 RNA polymerase promoters
Short DNA sequences that are recognized by T7 RNA polymerase to initiate transcription.



protein biosensors can sense small-molecule ligands to control intein-splicing events[134,135], thereby enabling 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. For example, a 3,4-dihydroxyphenylalanine (DOPA) dioxygenase was used to optimize the production of L-DOPA in yeast[136]. This enzyme cleaves the extradiol ring 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[137].

Whole-cell biosensors can be created without the need for complicated protein or aptamer engineering[138]. For example, recombinant *E. coli* that was made auxotrophic for mevalonate showed concentration-dependent growth rate changes in the presence of extracellular mevalonate[139]. Expressing GFP in this strain enabled the high-throughput detection of mevalonate levels in the culture broths of production strains[139]. 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[140].

Conclusions

New strategies from synthetic biology have enabled the engineering of large systems comprising many genetic parts, the synthetic control of gene expression and efficient genome editing. New tools have been developed 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. Although many of the approaches in this Review were developed in model organisms, such as *E. coli*, over the past few years these approaches have been increasingly transferred to organisms of more direct relevance to the production and delivery of natural products.

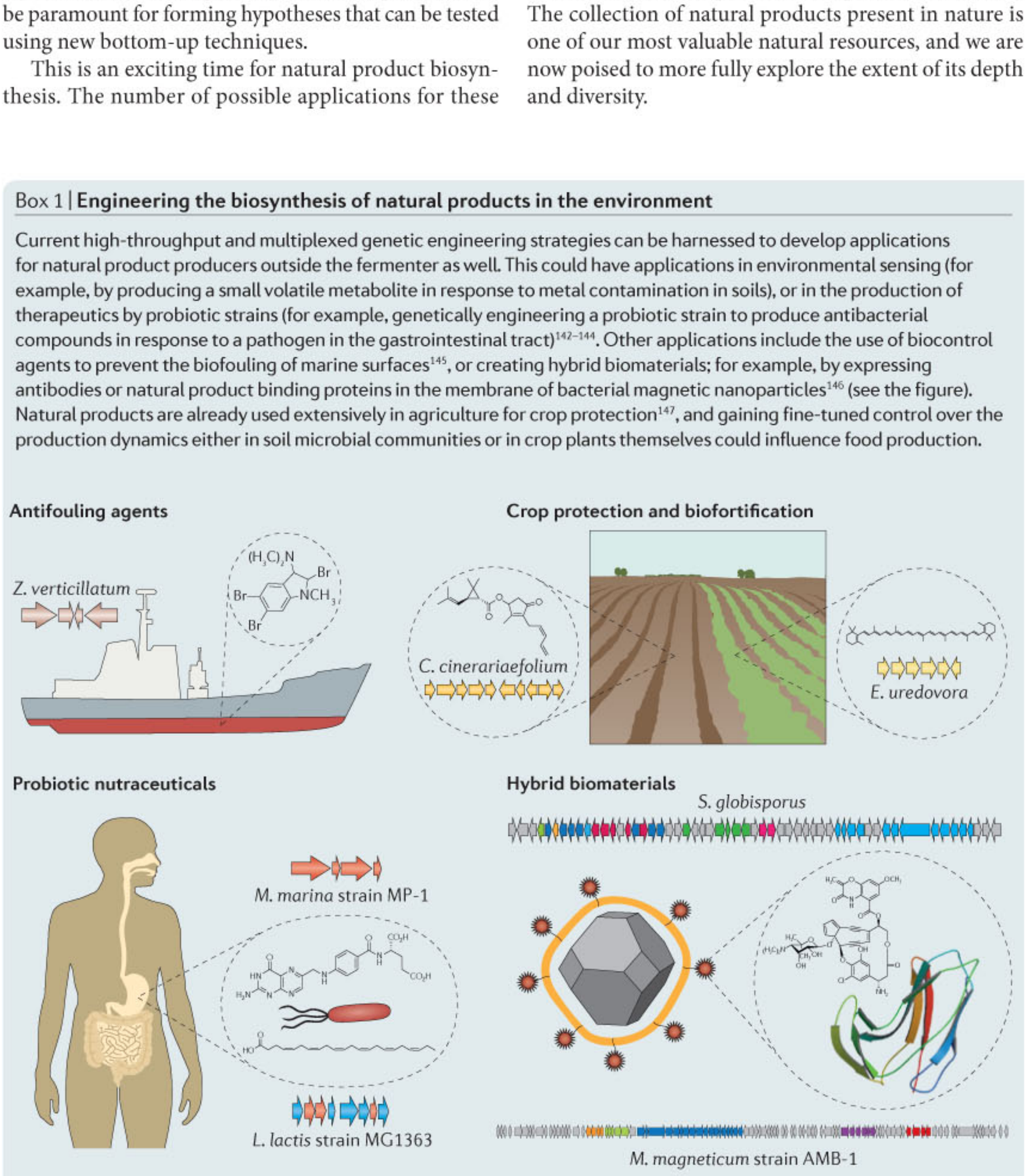
Figure 5 | **Multiplexed genome editing with CRISPR–Cas9. A** | Minimal genome editing construct design, including a single-guide RNA (sgRNA; part **a**), Cas9 from *Streptococcus pyogenes* (part **b**) and an optional 'repair fragment' (part **c**). Three routes of DNA repair are shown, including homologous recombination, alternative end-joining and non-homologous end-joining (NHEJ)[157]. **B** | Applications of CRISPR-mediated genome editing in *Streptomyces* spp. The graph at the bottom of the panel shows the reported efficiencies for experiments grouped by application with the background colours matching the diagrams above. Protocol differences are labelled below the graph, and data points are coloured according to the published study (blue[68], red[108] and green[112]). **C** | An example of multiplexed CRISPR editing for engineering the production of mevalonate. The bar graph at the bottom of the panel shows the editing efficiency (grey) and levels of mevalonate (green) averaged across several different combinations of gene deletions (the number of combinations is indicated in parentheses)[111]. BGC, biosynthetic gene cluster; tracrRNA, trans-activating CRISPR RNA.

For chemists and biologists whose research focuses on natural products, the challenge is to determine how best to leverage the latest technologies in DNA fabrication and genetic control to probe natural product pathways in new and insightful ways. Much in the same way that recombinant DNA technology has revolutionized our ability to approach the molecular details of biology from a reductionist point of view, the ability to rapidly construct 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 used in natural BGCs and the biochemistry of the biosynthetic pathways of natural products will be paramount for forming hypotheses that can be tested using new bottom-up techniques.

This is an exciting time for natural product biosynthesis. The number of possible applications for these molecules in medicine, industry and agriculture is vast (see BOX 1 for non-fermenter applications). The recent explosion in DNA sequencing technologies has identified that BGCs encoding natural products are more widespread than previously thought[6,12], and the ability to 'write' DNA into synthetic constructs is now catching up. New approaches for mining these molecules from genomic sequences are needed now more than ever to rejuvenate drug discovery pipelines, especially in light of the looming crisis of antibiotic resistance. The development of state-of-the-art high-throughput screening platforms enables purified compounds or semi-pure extracts to be screened in hundreds of assays with less material required than in previous decades[141]. The collection of natural products 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.

Box 1 | **Engineering the biosynthesis of natural products in the environment**

Current high-throughput and multiplexed genetic engineering strategies can be harnessed to develop applications for natural product producers outside the fermenter as well. This could have applications in environmental sensing (for example, by producing a small volatile metabolite in response to metal contamination in soils), or in the production of therapeutics by probiotic strains (for example, genetically engineering a probiotic strain to produce antibacterial compounds in response to a pathogen in the gastrointestinal tract)[142–144]. Other applications include the use of biocontrol agents to prevent the biofouling of marine surfaces[145], or creating hybrid biomaterials; for example, by expressing antibodies or natural product binding proteins in the membrane of bacterial magnetic nanoparticles[146] (see the figure). Natural products are already used extensively in agriculture for crop protection[147], and gaining fine-tuned control over the production dynamics either in soil microbial communities or in crop plants themselves could influence food production.



C. cinerariaefolium, *Chrysanthemum cinerariaefolium*; *E. uredovora*, *Erwinia uredovora*; *L. lactis*, *Lactococcus lactis*; *M. magneticum*, *Magnetospirillum magneticum*; *M. marina*, *Moritella marina*; *S. globisporus*, *Streptomyces globisporus*; *Z. verticillatum*, *Zoobotryon verticillatum*.

1. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335 (2012).
2. Demain, A. L. Importance of microbial natural products and the need to revitalize their discovery. *J. Ind. Microbiol. Biotechnol.* **41**, 185–201 (2014).
3. Davies, J. How to discover new antibiotics: harvesting the parvome. *Curr. Opin. Chem. Biol.* **15**, 5–10 (2011).
4. Röttig, M. *et al.* NRPSpredictor2 — a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39**, W362–W367 (2011).
5. Kim, J. & Yi, G.-S. PKMiner: a database for exploring type II polyketide synthases. *BMC Microbiol.* **12**, 169 (2012).
6. Cimermancic, P. *et al.* Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**, 412–421 (2014).
This study provides a comparative analysis of 33,000 putative BGCs present in more than 1,000 sequenced bacterial and archaeal genomes. A large family of aryl polyene gene clusters was characterized as a result.
7. van Heel, A. J., de Jong, A., Montalbán-López, M., Kok, J. & Kuipers, O. P. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res.* **41**, W448–W453 (2013).
8. Blin, K. *et al.* antiSMASH 2.0 — a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **41**, W204–W212 (2013).
9. Medema, M. H. *et al.* AntiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).
10. Weber, T. *et al.* antiSMASH 3.0 — a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43**, W237–W243 (2015).
11. Watve, M. G., Tickoo, R., Jog, M. M. & Bhole, B. D. How many antibiotics are produced by the genus *Streptomyces*? *Arch. Microbiol.* **176**, 386–390 (2001).
12. Doroghazi, J. R. *et al.* A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* **10**, 963–968 (2014).
This study describes a multi-parameter distance metric for comparing BGCs and applies this metric to organize more than 11,000 actinobacterial BGCs into families of gene clusters.
13. 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.* **283**, 28236–28245 (2008).
14. Piel, J. A polyketide synthase–peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc. Natl Acad. Sci. USA* **99**, 14002–14007 (2002).
15. Galm, U. & Shen, B. Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis. *Expert Opin. Drug Discov.* **1**, 409–437 (2006).
16. Williams, G. J. Engineering polyketide synthases and nonribosomal peptide synthetases. *Curr. Opin. Struct. Biol.* **23**, 603–612 (2013).
17. Pickens, L. B., Tang, Y. & Chooi, Y. H. Metabolic engineering for the production of natural products. *Annu. Rev. Chem. Biomolecular Engineer.* **2**, 211–236 (2011).
18. Demain, A. L. & Adrio, J. L. Strain improvement for production of pharmaceuticals and other microbial metabolites by fermentation. *Prog. Drug Res.* **65**, 252–289 (2008).
19. Voigt, C. A. Synthetic biology. *ACS Synth. Biol.* **1**, 1–2 (2012).
20. Way, J. C., Collins, J. J. & Keasling, J. D. & Silver, P. A. Integrating biological redesign: where synthetic biology came from and where it needs to go. *Cell* **157**, 151–161 (2014).
21. Medema, M. H., Breitling, R., Bovenberg, R. & Takano, E. Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. *Nat. Rev. Microbiol.* **9**, 131–137 (2011).
22. Luo, Y., Cobb, R. E. & Zhao, H. Recent advances in natural product discovery. *Curr. Opin. Biotechnol.* **30**, 230–237 (2014).
23. Luo, Y. *et al.* Engineered biosynthesis of natural products in heterologous hosts. *Chem. Soc. Rev.* **44**, 5265–5290 (2015).
24. Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523 (2015).
25. Keasling, J. D. Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* **14**, 189–195 (2012).
26. Fischbach, M. & Voigt, C. A. Prokaryotic gene clusters: a rich toolbox for synthetic biology. *Biotechnol. J.* **5**, 1277–1296 (2010).
27. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
28. Temme, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl Acad. Sci. USA* **109**, 7085–7090 (2012).
29. Chan, L. Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol. Syst. Biol.* **1**, 2005.0018 (2005).
30. Shao, Z. *et al.* Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth. Biol.* **2**, 662–669 (2013).
31. Oßwald, C. *et al.* Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* **3**, 759–772 (2012).
32. Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 2894 (2013).
33. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**, 70–74 (2010).
This study showcases the approach of multivariate modular metabolic engineering for optimizing the biosynthesis of a Taxol precursor to 1 gram per litre in *E. coli*.
34. Paddon, C. J. *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528–532 (2013).
35. Mutalik, V. K. *et al.* Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods* **10**, 347–353 (2013).
36. Chen, Y.-J. *et al.* Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659–664 (2013).
37. Seghezzi, N., Amar, P., Koebmann, B., Jensen, P. R. & Virolle, M. J. The construction of a library of synthetic promoters revealed some specific features of strong *Streptomyces* promoters. *Appl. Microbiol. Biotechnol.* **90**, 615–623 (2011).
38. Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc. Natl Acad. Sci. USA* **102**, 12678–12683 (2005).
39. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
40. Farasat, I. *et al.* Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Mol. Syst. Biol.* **10**, 731 (2014).
41. Nielsen, A. A. K., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr. Opin. Chem. Biol.* **17**, 878–892 (2013).
42. Kosuri, S. *et al.* Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **110**, 14024–14029 (2013).
43. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
44. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–866 (2012).
45. Lou, C., Stanton, B., Chen, Y.-J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
46. Bai, C. *et al.* Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc. Natl Acad. Sci. USA* **112**, 12181–12186 (2015).
47. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131–1141 (2011).
48. Kalir, S. *et al.* Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292**, 2080–2083 (2001).
49. 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.* **377**, 47–61 (2008).
50. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520 (2014).
51. Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol. Mol. Biol. Rev.* **77**, 112–143 (2013).
52. Chatterjee, A. *et al.* Convergent transcription in the butyrolactone regulon in *Streptomyces coelicolor* confers a bistable genetic switch for antibiotic biosynthesis. *PLoS ONE* **6**, e21974 (2011).
53. Sherwood, E. J. & Bibb, M. J. The antibiotic planosporicin coordinates its own production in the actinomycete *Planomonospora alba*. *Proc. Natl Acad. Sci. USA* **110**, E2500–E2509 (2013).
54. Chen, Y., Smanski, M. J. & Shen, B. Improvement of secondary metabolite production in *Streptomyces* by manipulating pathway regulation. *Appl. Microbiol. Biotechnol.* **86**, 19–25 (2010).
55. Smanski, M. J., Peterson, R. M., Rajski, S. R. & Shen, B. Engineered *Streptomyces platensis* strains that overproduce antibiotics platensimycin and platencin. *Antimicrob. Agents Chemother.* **53**, 1299–1304 (2009).
56. Tahlan, K. *et al.* Initiation of actinorhodin export in *Streptomyces coelicolor*. *Mol. Microbiol.* **63**, 951–961 (2007).
57. Stevens, J. T. & Carothers, J. M. Designing RNA-based genetic control systems for efficient production from engineered metabolic pathways. *ACS Synth. Biol.* **4**, 107–115 (2015).
58. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor–regulator system for production of chemicals and fuels derived from fatty acids. *Nat Biotechnol.* **30**, 354–359 (2012).
59. Kushwaha, M. & Salis, H. M. A portable expression resource for engineering cross-species genetic circuits and pathways. *Nat. Commun.* **6**, 7832 (2015).
60. Solomon, K. V., Sanders, T. M. & Prather, K. L. J. A dynamic metabolite valve for the control of central carbon metabolism. *Metab. Eng.* **14**, 661–671 (2012).
61. Brockman, I. M. & Prather, K. L. J. Dynamic knockdown of *E. coli* central metabolism for redirecting fluxes of primary metabolites. *Metab. Eng.* **28**, 104–113 (2015).
62. Nieselt, K. *et al.* The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* **11**, 10 (2010).
63. Ellis, T., Wang, X. & Collins, J. J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.* **27**, 465–471 (2009).
64. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
65. Hasty, J., McMillen, D. & Collins, J. J. Engineered gene circuits. *Nature* **420**, 224–230 (2002).
66. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
This study describes CRISPRi, which utilizes a catalytically inactive dCas9 protein to block gene expression by RNA polymerase.
67. Gilbert, L. A. *et al.* XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
This study demonstrates that the CRISPR–dCas9 system can be used to activate gene expression by fusing transcriptional activator domains to dCas9.
68. Tong, Y., Charusanti, P., Zhang, L., Weber, T. & Lee, S. Y. CRISPR–Cas9 based engineering of actinomycetal genomes. *ACS Synth. Biol.* **4**, 1020–1029 (2015).
69. Piatek, A. *et al.* RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol. J.* **13**, 578–589 (2015).
70. 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.* **29**, 160–168 (2015).
71. Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339–350 (2014).

72. Zetsche, B., Volz, S. E. & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* **33**, 139–142 (2015).
73. Sánchez, C. *et al.* Combinatorial biosynthesis of antitumor indolocarbazole compounds. *Proc. Natl Acad. Sci. USA* **102**, 461–466 (2005).
74. Menzella, H. G. *et al.* Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005).
75. Kakule, T. B., Lin, Z. & Schmidt, E. W. Combinatorialization of fungal polyketide synthase–peptide synthetase hybrid proteins *J. Am. Chem. Soc.* **136**, 17882–17890 (2014).
76. Nguyen, K. T. *et al.* Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proc. Natl Acad. Sci. USA* **103**, 17462–17467 (2006).
77. Kosuri, S. & Church, G. M. Large-scale *de novo* DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499–507 (2014).
78. Ru, D. E., Schmidt, E. W. & Heemstra, J. R. Assessing the combinatorial potential of the RiPP cyanobactin tru pathway. *ACS Synth. Biol.* **4**, 482–492 (2015).
The authors use high-throughput mutagenesis and analytical chemistry to create more than 300 structural variants of the cyanobacterial RiPP, trunkamide. The data were analysed to establish rules for amino acid preference at various positions along the core scaffold.
79. Mitchell, D. A. *et al.* Structural and functional dissection of the heterocyclic peptide cytotoxin streptolysin S. *J. Biol. Chem.* **284**, 13004–13012 (2009).
80. Smanski, M. J. *et al.* Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* **32**, 1241–1249 (2014).
81. Appleton, E., Tao, J., Haddock, T. & Densmore, D. Interactive assembly algorithms for molecular cloning. *Nat. Methods* **11**, 657–662 (2014).
82. Freestone, T. S. & Zhao, H. Combinatorial pathway engineering for optimized production of the anti-malarial FR900098. *Biotechnol. Bioeng.* **113**, 384–392 (2015).
83. Biggs, B. W., De Paepe, B., Santos, C. N. S., De Mey, M. & Ajikumar, P. K. Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* **29**, 156–162 (2014).
84. Thodey, K., Galanie, S. & Smolke, C. D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* **10**, 837–844 (2014).
85. Li, L. *et al.* A stepwise increase in pristinamycin II biosynthesis by *Streptomyces pristinaespiralis* through combinatorial metabolic engineering. *Metab. Eng.* **29**, 12–25 (2015).
86. Zhao, S. *et al.* Improvement of catechin production in *Escherichia coli* through combinatorial metabolic engineering. *Metab. Eng.* **28**, 43–53 (2015).
87. Du, Y.-L. & Ryan, K. S. Expansion of bisindole biosynthetic pathways by combinatorial construction. *ACS Synth. Biol.* **4**, 682–688 (2015).
88. Chemler, J. a. *et al.* Evolution of efficient modular polyketide synthases by homeologous recombination. *J. Am. Chem. Soc.* **137**, 10603–10609 (2015).
89. Smanski, M. J. *et al.* Expression of the platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. *J. Nat. Prod.* **75**, 2158–2167 (2012).
90. Wang, B., Kitney, R. I., Joly, N. & Buck, M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat. Commun.* **2**, 508 (2011).
91. Hajimorad, M., Gray, P. R. & Keasling, J. D. A framework and model system to investigate linear system behavior in *Escherichia coli*. *J. Biol. Eng.* **5**, 3 (2011).
92. Moser, F. *et al.* Genetic circuit performance under conditions relevant for industrial bioreactors. *ACS Synth. Biol.* **1**, 555–564 (2012).
93. 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.* **19**, 98–106 (2013).
94. Sarrion-Perdigones, A. *et al.* GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol.* **162**, 1618–1631 (2013).
95. Zhang, Y., Perry, K., Vinci, V. & Powell, K. Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* **415**, 5–7 (2002).
96. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).
97. Warner, J. R., Reeder, P. J., Karimpour-Fard, A., Woodruff, L. B. A. & Gill, R. T. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.* **28**, 856–862 (2010).
98. Isaacs, F. J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
99. Wang, H. H. *et al.* Genome-scale promoter engineering by coselection MAGE. *Nat. Methods* **9**, 591–593 (2012).
100. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
101. Wang, H. H. & Church, G. M. Multiplexed genome engineering and genotyping methods: applications for synthetic biology and metabolic engineering. *Methods Enzymol.* **498**, 409–426 (2011).
102. Wang, H. H. *et al.* Multiplexed *in vivo* his-tagging of enzyme pathways for *in vitro* single-pot multienzyme catalysis. *ACS Synth. Biol.* **1**, 43–52 (2012).
103. Bonde, M. T. *et al.* Direct mutagenesis of thousands of genomic targets using microarray-derived oligonucleotides. *ACS Synth. Biol.* **4**, 17–22 (2015).
In this study, the authors develop a workflow for the integration of microarray-based DNA synthesis with MAGE and apply their approach to introduce T7 promoters upstream of 2,500 operons in the *E. coli* genome.
104. 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.* **41**, 6360–6369 (2013).
105. Dicarlo, J. E. *et al.* Yeast oligo-mediated genome engineering (YOGE). *ACS Synth. Biol.* **2**, 741–749 (2013).
106. Sandoval, N. R. *et al.* Strategy for directing combinatorial genome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **109**, 10540–10545 (2012).
107. 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* **9**, 39 (2013).
108. Cobb, R. E., Wang, Y. & Zhao, H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* **4**, 723–728 (2014).
This is the first study to demonstrate CRISPR–Cas9 genome editing in *Streptomyces* spp.; the authors achieve up to 100% editing efficiency for the deletion of single genes or whole gene clusters.
109. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
110. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR–Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
111. Jakočiūnas, T. *et al.* Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* **28**, 213–222 (2015).
112. 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).* **47**, 231–243 (2015).
113. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature* **517**, 583–588 (2015).
114. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR–Cas9. *Science* **346**, 1–9 (2014).
115. Bao, Z., Cobb, R. E. & Zhao, H. Accelerated genome engineering through multiplexing. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **8**, 5–21 (2015).
116. Chen, C., Fenk, L. a. & De Bono, M. Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. *Nucleic Acids Res.* **41**, e193 (2013).
117. Michener, J. K., Thodey, K., Liang, J. C. & Smolke, C. D. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* **14**, 212–222 (2012).
118. An, G. H., Bielich, J., Auerbach, R. & Johnson, E. A. Isolation and characterization of carotenoid hyperproducing mutants of yeast by flow cytometry and cell sorting. *Biotechnology* **9**, 70–73 (1991).
119. Cho, E. J., Lee, J.-W. & Ellington, A. D. Applications of aptamers as sensors. *Annu. Rev. Anal. Chem. (Palo Alto. Calif.).* **2**, 241–264 (2009).
120. Wachsmuth, M., Findeiss, S., Weissheimer, N., Stadler, P. F. & Morl, M. *De novo* design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.* **41**, 2541–2551 (2012).
121. Weigand, J. E. & Suess, B. Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.* **35**, 4179–4185 (2007).
122. Michener, J. K. & Smolke, C. D. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metab. Eng.* **14**, 306–316 (2012).
123. Weigand, J. E. *et al.* Screening for engineered neomycin riboswitches that control translation initiation. *RNA* **14**, 89–97 (2008).
124. Schoukroun-Barnes, L. R., Wagan, S. & White, R. J. Enhancing the analytical performance of electrochemical RNA aptamer-based sensors for sensitive detection of aminoglycoside antibiotics. *Anal. Chem.* **86**, 1131–1137 (2014).
125. Farjami, E. *et al.* RNA aptamer-based electrochemical biosensor for selective and label-free analysis of dopamine. *Anal. Chem.* **85**, 121–128 (2013).
126. 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* **25**, 555–560 (2012).
127. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl Acad. Sci. USA* **89**, 5547–5551 (1992).
128. Grkovic, S., Hardie, K. M., Brown, M. H. & Skurray, R. A. Interactions of the QacR multidrug-binding protein with structurally diverse ligands: implications for the evolution of the binding pocket. *Biochemistry* **42**, 15226–15236 (2003).
129. Ramos, J. L. *et al.* The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev.* **69**, 326–356 (2005).
130. Mohn, W. W., Garmendia, J., Galvao, T. C. & De Lorenzo, V. Surveying biotransformations with *à la carte* genetic traps: translating dehydrochlorination of lindane (γ-hexachlorocyclohexane) into *lacZ*-based phenotypes. *Environ. Microbiol.* **8**, 546–555 (2006).
131. 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.* **14**, 449–457 (2012).
132. Tang, S. Y. & Cirino, P. C. Design and application of a mevalonate-responsive regulatory protein. *Angew. Chem. Int. Ed. Engl.* **50**, 1084–1086 (2011).
133. Tang, S. Y. *et al.* Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. *J. Am. Chem. Soc.* **135**, 10099–10103 (2013).
134. Buskirk, A. R., Ong, Y.-C., Gartner, Z. J. & Liu, D. R. Directed evolution of ligand dependence: small-molecule-activated protein splicing. *Proc. Natl Acad. Sci. USA* **101**, 10505–10510 (2004).
135. Peck, S. H., Chen, I. & Liu, D. R. Directed evolution of a small-molecule-triggered intein with improved splicing properties in mammalian cells. *Chem. Biol.* **18**, 619–630 (2011).
136. DeLoache, W. C. *et al.* An enzyme-coupled biosensor enables (*S*)-reticuline production in yeast from glucose. *Nat. Chem. Biol.* **11**, 465–471 (2015).
137. Gandía-Herrero, F. & García-Carmona, F. Biosynthesis of betalains: yellow and violet plant pigments. *Trends Plant Sci.* **18**, 334–343 (2013).
138. Meyer, A. *et al.* Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors. *Nat. Chem.* **7**, 673–678 (2015).
139. Pfleger, B. F., Pitera, D. J., Newman, J. D., Martin, V. J. J. & Keasling, J. D. Microbial sensors for small molecules: development of a mevalonate biosensor. *Metab. Eng.* **9**, 30–38 (2007).
140. Bertels, F., Merker, H. & Kost, C. Design and characterization of auxotrophy-based amino acid biosensors. *PLoS ONE* **7**, e41349 (2012).
141. Bugni, T. S. *et al.* Marine natural product libraries for high-throughput screening and rapid drug discovery. *J. Nat. Prod.* **71**, 1095–1098 (2008).

142. Kotula, J. W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl Acad. Sci. USA* **111**, 4838–4843 (2014).
143. Claesen, J. & Fischbach, M. A. Synthetic microbes as drug delivery systems. *ACS Synth. Biol.* **4**, 358–364 (2015).
144. Fischbach, M. A., Bluestone, J. A. & Lim, W. A. Cell-based therapeutics: the next pillar of medicine. *Sci. Transl. Med.* **5**, 179ps7 (2013).
145. Qian, P.-Y., Xu, Y. & Fusetani, N. Natural products as antifouling compounds: recent progress and future perspectives. *Biofouling* **26**, 223–234 (2010).
146. Yoshino, T. & Matsunaga, T. Development of efficient expression system for protein display on bacterial magnetic particles. *Biochem. Biophys. Res. Commun.* **338**, 1678–1681 (2005).
147. Dayan, F. E., Cantrell, C. L. & Duke, S. O. Natural products in crop protection. *Bioorg. Med. Chem.* **17**, 4022–4034 (2009).
148. Zimmermann, M. & Fischbach, M. A. A family of pyrazinone natural products from a conserved nonribosomal peptide synthetase in *Staphylococcus aureus*. *Chem. Biol.* **17**, 925–930 (2010).
149. Smanski, M. J. *et al.* Dedicated *ent*-kaurene and *ent*-atiserene synthases for platensimycin and platencin biosynthesis. *Proc. Natl Acad. Sci. USA* **108**, 13498–13503 (2011).
150. 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.* **70**, 67–74 (2007).
151. Wong, F. T. & Khosla, C. Combinatorial biosynthesis of polyketides — a perspective. *Curr. Opin. Chem. Biol.* **16**, 117–123 (2012).
152. Medema, M. H. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).
153. Carlson, R. Time for new DNA synthesis and sequencing cost curves. *Synthesis* [online], http://www.synthesis.cc/2014/02/time-for-new-cost-curves-2014.html (2014).
154. Gustafsson, C., Govindarajan, S. & Minshull, J. Codon bias and heterologous protein expression. *Trends Biotechnol.* **22**, 346–353 (2004).
155. Rodrīguez-Garcīa, A., Combes, P., Pērez-Redondo, R., Smith, M. C. A. & Smith, M. C. M. Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Res.* **33**, e87 (2005).
156. Stanton, B. C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth. Biol.* **19**, 880–891 (2014).
157. Selle, K. & Barrangou, R. Harnessing CRISPR–Cas systems for bacterial genome editing. *Trends Microbiol.* **23**, 225–232 (2015).

Competing interests statement
The authors declare no competing interests.