nature reviews bioengineering

https://doi.org/10.1038/s44222-025-00354-x

Review article

Check for updates

Non-model bacteria as platforms for endogenous gene expression in synthetic biology

Jorien Poppeliers[1,2], Maarten Boon[1], Marjan De Mey[3], Joleen Masschelein[2,4] & Rob Lavigne[1]

Abstract

Microbial synthetic biology seeks to engineer bacterial genomes for industrial and biomedical applications, typically by applying heterologous gene expression in well-characterized model organisms, such as *Escherichia coli*. However, heterologous gene expression might cause metabolic disruptions, thereby impacting production efficiency and yield. In this Review, we highlight non-model organisms, such as *Lacticaseibacillus* and pseudomonads, for endogenous compound production, taking advantage of their evolutionary optimization for the production of certain metabolites and proteins. We first outline key limitations of heterologous production and then examine endogenous production pathways in non-model organisms for biotechnological and therapeutic applications. In particular, multi-omics approaches enable the discovery and characterization of these organisms, and phage-based genome refactoring enhances genome engineering capabilities. Finally, we outline key bottlenecks in the application of non-model organisms in biotechnology, including scale-up, costs and safety.

Sections

[1]Laboratory of Gene Technology, Department of Biosystems, KU Leuven, Leuven, Belgium. [2]Laboratory for Biomolecular Discovery and Engineering, Department of Biology, KU Leuven, Leuven, Belgium. [3]Metabolic Engineering of Microorganisms group, Department of Biotechnology, UGent, Ghent, Belgium. [4]Center for Microbiology, VIB-KU Leuven, Leuven, Belgium. e-mail: rob.lavigne@kuleuven.be

Key points

- Chassis diversification is essential to unlock the full potential of microbial synthetic biology.
- Native producers have an evolutionary advantage over heterologous expression hosts, which can be leveraged for biotechnological applications.
- Multi-omics techniques enable the characterization of non-model organisms to become the next generation of synthetic biology chassis.
- Bacteriophage genomes are valuable sources for the discovery of synthetic biology tools to modify gene expression in non-model organisms.
- The switch towards endogenous gene expression can improve market adoption for the second generation of microbial cell factories.

Introduction

In microbial synthetic biology, genomes of microorganisms (that is, bacteria or fungi) are typically engineered to customize or create new biological processes[1]. In particular, applications in industry and biomedicine have emerged by recreating, mimicking or modifying microbial metabolic processes and exploiting molecular cloning and omics technologies[2]. Basic biological circuits can be designed to mimic logic behaviour, and entire biosynthetic pathways can be constructed to engineer synthetic cells[3–5]. Such engineered microorganisms can be harnessed as sustainable factories for the production of 'green' chemicals (for example, biofuels, surfactants and herbicides) or they can be transformed into fuel cells to generate renewable energy[6–10]. In addition, microorganisms can be designed to function as live-attenuated vaccines or probiotics[11,12].

In microbial synthetic biology, the term chassis is typically used as a metaphor for the framework or foundation upon which genetic engineering is built[13]. In particular, model organisms, such as *Escherichia coli*, have served as chassis in microbial synthetic biology[1,14]. Owing to their well-characterized metabolism and regulatory networks, the modification of their genome and the introduction of new genetic material is facilitated. Therefore, outcomes of genetic modifications can be anticipated, reducing the iterative process required to obtain an optimal host for the target application. Moreover, engineering protocols and DNA delivery tools are well established for model organisms[15,16] to serve a variety of applications[17–21]. Accordingly, many biotechnological applications are biased towards these model organisms, even if these demands often push them beyond their natural capacity.

An alternative approach is to select the host organism best suited for a given condition and application[22]. Application-oriented synthetic biology requires engineered chassis to function in industrially relevant environments[14], which are often characterized by alternative feedstock, such as waste streams, limited aeration and fluctuations in temperature or pH[23,24]. However, the performance of model organisms can be affected by environmental stresses, and therefore, their potential as fermentation hosts remains limited. Alternatively, bacteria, including *Lacticaseibacillus*, *Burkholderia*, *Streptomyces*, *Bacillus subtilis* and *Pseudomonas putida*, have been explored as chassis, as they are inherently better adapted to industrial conditions owing to the harsh environments in which they naturally occur[13,14].

Through evolution, many bacteria have naturally developed specialized genes or biosynthetic gene clusters that enable them to thrive in distinct environments. Their metabolic pathways are thus inherently optimized to produce certain compounds, as they already possess all the necessary resources, such as precursor molecules, antitoxins and protein-folding chaperones. Using these endogenous systems avoids the need to introduce large fragments of foreign DNA, thereby minimizing the metabolic burden associated with resource reallocation[25]. This makes endogenous gene expression, especially when combined with genome engineering, a powerful approach for leveraging non-model bacteria in biotechnological applications[22,26].

In this Review, we discuss non-model bacteria as native production hosts for applications in microbial synthetic biology. We first outline shortcomings of heterologous expression in model organisms and the evolutionary advantages that favour native producers for specific applications. We then provide a framework for the implementation of non-model bacteria in biotechnology. In particular, we highlight a multi-omics approach to guide engineering as well as to identify and address potential bottlenecks, and examine phage genomes as resources to extend the synthetic biology toolbox beyond model organisms. Finally, we give an overview of current and potential future industrial applications that use non-model organisms.

Limitations of model organisms for heterologous production

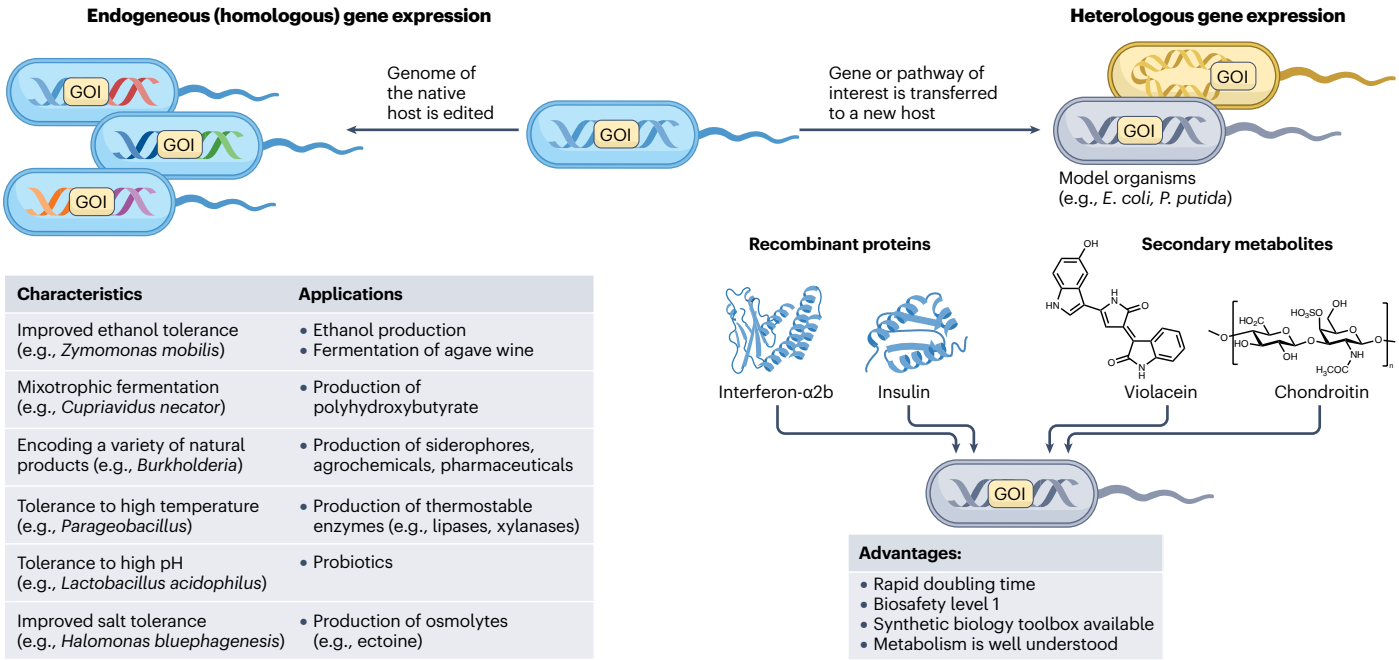
Heterologous gene expression requires the introduction, transcription and translation of foreign DNA to produce a functional output (for example, protein or metabolite). This approach is typically used if genetic manipulation or cultivation of the native host is challenging[19,27]. It has been applied for the production of human insulin, interferon-α2b, flavour compounds and biofuels[28–31] (Fig. 1). However, expressing genes outside their native biological context can result in reduced production yields, compared with the original producer[32].

Codon usage bias

A key challenge for heterologous gene expression is codon usage bias; here, codons of foreign genes may not match the preference of the new host. This can be mitigated by silent mutations, introduced through gene synthesis[33] or codon-randomized refactoring of biosynthetic gene clusters[34]. However, specific codon usage is considered a fine-tuned regulatory mechanism affecting both transcription and translation. Codon choice impacts posttranscriptional and posttranslational modifications, influencing mRNA levels as well as the structure and function of proteins[35]. In addition to reliance on gene synthesis, costly and labour-intensive screening is often needed to identify improved pathways. Alternatively, rare transfer RNA molecules can be coexpressed to reduce the demand of native transfer RNAs[36]. However, this approach can disrupt cellular resource allocation and impose fitness costs on engineered cells[36].

Integration of DNA

Synthesized DNA fragments must be cloned into a new host. In vitro assembly methods, including Gibson assembly and Golden Gate cloning, enable the integration of multiple DNA fragments into a plasmid backbone[37,38]. However, biosynthetic gene cluster fragmentation can disrupt the native transcriptional architecture and, consequently, decrease product yields. Direct pathway cloning, combining long



Characteristics	Applications
Improved ethanol tolerance (e.g., *Zymomonas mobilis*)	• Ethanol production • Fermentation of agave wine
Mixotrophic fermentation (e.g., *Cupriavidus necator*)	• Production of polyhydroxybutyrate
Encoding a variety of natural products (e.g., *Burkholderia*)	• Production of siderophores, agrochemicals, pharmaceuticals
Tolerance to high temperature (e.g., *Parageobacillus*)	• Production of thermostable enzymes (e.g., lipases, xylanases)
Tolerance to high pH (e.g., *Lactobacillus acidophilus*)	• Probiotics
Improved salt tolerance (e.g., *Halomonas bluephagenesis*)	• Production of osmolytes (e.g., ectoine)

Fig. 1 | Heterologous and endogenous gene expression in biotechnology. In endogenous, or homologous, gene expression, the genome of the native host is edited. Evolution has selected native producers to thrive under specific environmental conditions, making them valuable for diverse biotechnological applications. However, a synthetic toolbox to engineer their genomes is lacking. Consequently, heterologous expression in model organisms is frequently chosen as an alternative strategy. Here, the gene or biosynthetic gene cluster of interest (GOI) is transferred to a model organism to produce the desired protein or metabolite[218–222]. *E. coli*, *Escherichia coli*; *P. putida*, *Pseudomonas putida*.

amplification polymerase chain reaction with in vitro recombination, addresses this limitation by enabling the cloning of entire pathways into a plasmid backbone[39]. However, amplifying and cloning biosynthetic gene clusters of large size remains complex. Alternatively, artificial chromosomes can accommodate DNA fragments of up to 300 kb and can thus maintain and express large biosynthetic gene clusters[40,41]. Transformation-associated recombination, which is an in vivo DNA assembly method that relies on artificial chromosomes to express entire biosynthetic gene clusters[42,43], has been applied in *Streptomyces* to express the tautomycetin cluster (80 kb), which produces a potential anticancer molecule[44], and several aromatic polyketides (biosynthetic gene clusters of up to 213 kb) with antibacterial properties[45]. Although predominantly used in *Streptomyces*, artificial chromosomes have also been applied in other hosts, including *Pseudomonas*, *Salmonella* and *Agrobacterium*[46,47]. Nevertheless, constructing artificial chromosomes remains time-consuming and technically challenging, especially for GC-rich sequences, and does not guarantee heterologous expression[43]. In addition, these vectors typically occur in low copy numbers, which limits expression levels. Once assembled, the DNA needs to be transferred to the host of interest, typically by chemical transformation, electroporation or conjugation. Although well studied for model organisms, the efficiency of these methods is strain dependent, and implementation in newly discovered species can be difficult[48].

Disruption of metabolic flux

The introduction of new DNA fragments can also disrupt the metabolic flux of the host organism. Therefore, reallocation of cellular resources is required to express additional genes or biosynthetic gene clusters[18]. For example, the metabolic flux of *E. coli* can be reprogrammed to optimize production of chondroitin, a glycosaminoglycan that is used as a supplement to treat osteoarthritis[49], by introducing three heterologous enzymes; however, this requires changes in medium composition (that is, supplementation of the precursor) and in several metabolic pathways to attain relevant production levels[50]. Therefore, additional engineering is required to alter the metabolic flux in favour of the heterologous compound. Furthermore, the absence of certain resources, such as RNA-binding proteins to stabilize mRNA, chaperones to guide protein folding or export signals to avoid accumulation in the cell and facilitate downstream processing, can impede recombinant protein production, increasing both labour and cost[19,21,51]. These challenges in heterologous gene expression remain to be addressed.

Endogenous gene expression in non-model organisms

A second approach to alter gene expression in bacteria is endogenous (or homologous) expression (Fig. 1). Endogenous expression can be applied only if the coding sequence or biosynthetic gene cluster of interest is natively present in the chassis[52]. With over 10^5 bacterial species described so far, there is potential to diversify the range of chassis strains available for synthetic biology[53].

Evolutionary adaptation and fitness landscapes

Through evolution, bacterial populations have adapted to survive and function in specific ecological niches. The concept of fitness landscapes can be applied to describe how evolutionary pressure has shaped such adaptations[54] by mapping the genotype–phenotype

relationship as a 3D landscape with bacterial populations navigating peaks and valleys in response to environmental changes. Fitness landscapes change according to the dynamics of the environment, causing new peaks to emerge that drive the evolutionary trajectory of bacterial populations. This fine-tuned, dynamic interplay between genetic traits and environmental factors determines both bacterial fitness and phenotypic traits endogenously expressed by bacteria[54,55]. For example, *Pseudomonas* species have evolved to colonize different niches: *Pseudomonas aeruginosa* colonizes the human lungs, whereas *Pseudomonas syringae* and *Pseudomonas fluorescens* inhabit plants and soil, respectively. Consequently, distinct characteristics have emerged among these species, ranging from human and plant pathogenicity (for example, *P. aeruginosa* and *P. syringae*) to plant growth promotion (for example, *P. fluorescens*)[56–58].

Rapidly altering genetic (for example, through introduction of a new gene or pathway) or environmental (for example, laboratory-scale to industrial-scale) conditions can result in the maladaptation of engineered populations to the applied changes[59,60]. For example, the expression of green fluorescent protein (GFP) can serve as an indication for the stability of synthetic circuit expression in *E. coli* over multiple generations. Upon GFP expression from two circuits, induced by either *N*-acylhomoserine lacton or isopropyl β-D-1-thiogalactopyranoside), *E. coli* loses its ability to produce GFP after 20–50 generations, probably owing to the lack of selective pressure[61]. Therefore, gene–environment interactions should be considered as a coupled system, rather than as isolated variables. Moreover, it remains unclear whether new gene–environment combinations fail because they have not been favoured by evolutionary pathways shaped by past genetic changes, or because such combinations are inherently disadvantageous and, thus, selected against[62].

Applications of extremophiles

Instead of altering multiple variables in model bacteria, chassis can be selected that require limited genetic alterations (for example, promoter exchange) or that can thrive in environmental conditions that resemble the original habitat. For example, extremophiles can be applied in processes that require operation under harsh environmental conditions, such as high temperature, high pH, high salinity or the presence of toxic compounds (Fig. 1). Extremophiles encode specialized pathways or enzymes that function under these conditions[63]. For example, lipases and xylanases derived from thermophiles are commonly used in laundry detergents or for the degradation of paper waste, respectively[64,65]. Extremophiles also produce a range of metabolites that can be used for biomanufacturing. For example, halophiles produce the natural product ectoin, which is used in skin care products[66]. Furthermore, they are capable of producing high amounts of polyhydroxyalkanoates, which are biodegradable plastics with numerous applications in medicine, agriculture and packaging[67]. Beyond biomanufacturing, acidophiles, such as *Lactobacillus acidophilus*, are used as probiotics. Their tolerance for acidic environments allows them to produce metabolites that lower the pH, preventing gastrointestinal infections and reducing serum cholesterol levels[68]. In addition, the extremophilic bacterium *Bhargavaea beijingensis* has been applied for bioremediation and the production of bio-cement[69].

Biotechnological and therapeutic uses of mesophilic bacteria

Native bacteria can also be applied in vaccines to protect against bacterial pathogens[70]. For example, the EvoVax vaccine, an oral vaccine consisting of four peracetic acid-inactivated *Salmonella* strains, is being tested in piglets to treat salmonellosis, and the *Bacillus* Calmette–Guérin vaccine, a live-attenuated strain of *Mycobacterium bovis*, is approved to prevent tuberculosis[70,71]. In addition, mesophilic bacteria can be used for biomanufacturing. For example, *Cupravidus necator* is well suited for the production of biopolymers, including poly-3-hydroxybutyrate, owing to its ability to switch between a heterotrophic and autotrophic lifestyle. *Burkholderia* spp. have large genomes, in which they encode a high number of natural products, such as pharmaceuticals, lipopeptides and siderophores. Moreover, *Zymomonas mobilis* can produce and tolerate high levels of ethanol[72–74] (Fig. 1).

Screening by multi-omics approaches

Microbial chassis strains can be selected on the basis of their natural specializations to fully leverage endogenous gene expression in synthetic biology[26,75]. However, although non-model bacteria have an evolutionary advantage, the mechanisms underlying their adaptations need to be understood to select the most suitable chassis for a given biotechnological application. Many bacteria have been isolated for specific purposes, resulting in limited characterization beyond their initial application[76]. However, these underexplored strains may hold untapped potential for biotechnological applications.

Multi-omics approaches provide powerful tools for the identification, screening and selection of suitable microbial hosts for endogenous gene expression by offering insights into the working mechanisms of microorganisms (Fig. 2). Following isolation of a strain, genomics enables the identification of the bacterial strain and the compounds it encodes. However, genome annotations can be incomplete, necessitating additional experimental procedures, such as transcriptomics, proteomics and metabolomics, to refine annotations of regulatory elements, biosynthetic gene clusters and hypothetical proteins. Regulatory and metabolic bottlenecks can then be identified through a combination of transcriptomics, ribosomal profiling and fluxomics. Such multi-omics information can be integrated in a genome-scale metabolic model to simulate the genotype-to-phenotype relationship. This system-level understanding supports rational engineering of non-model bacteria into optimized microbial cell factories.

Genomics

DNA sequencing and genomics are key pillars in the identification and characterization of microorganisms. Currently, 715,230 bacterial genomes are publicly available in databases[53], and many more remain to be discovered. Therefore, identifying suitable chassis can be challenging. To capture the potential of bacteria to perform a certain function, accurate genome annotations are crucial, matching DNA sequences to their potential function. For example, bacterial genome annotation tools, such as prokka and bakta[77,78], perform annotation based on sequence similarity to proteins uploaded in databases, such as RefSeq and Uniprot[79,80]. Therefore, genome annotation is limited to proteins that have been (experimentally) characterized[81]. As the 3D structure of a protein is intrinsically related to its function, the development of open-source software to computationally predict protein structures (for example, ColabFold and RoseTTAFold) has greatly increased the number of available protein structures and their functions in databases[82,83]. To further optimize annotations, a multitude of tools can be applied for specific queries, including for biosynthetic gene clusters (for example, antiSMASH and MiBIG), genes related to the bacterial immune system (for example, Deepdefense) and mobile genetic elements (for example, geNomad and PHASTEST)[84–88]. Nevertheless, the accurate prediction and annotation of small open reading

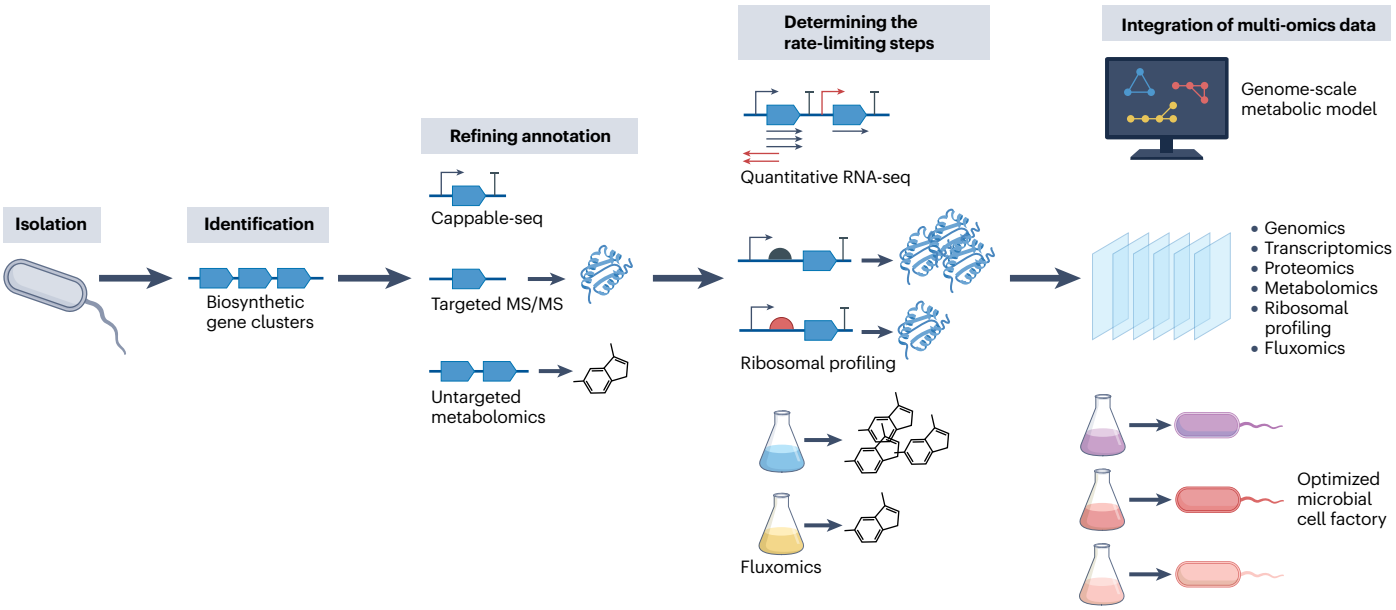


Fig. 2 | Omics technologies guiding the transition from isolation to microbial cell factories. Non-model bacteria can be developed into microbial cell factories using a multi-omics approach. Following isolation, genomics can aid in the identification of bacteria and biosynthetic gene clusters. However, genome annotations are not complete, and thus, additional experimental procedures (that is, transcriptomics, proteomics and metabolomics) can be applied to refine annotations with regards to regulatory elements, biosynthetic gene clusters and confirmation of hypothetical proteins. Once the compound of interest is identified, regulatory and metabolic bottlenecks can be identified through transcriptomics, ribosomal profiling and fluxomics. Finally, omics information can be integrated in a genome-scale metabolic model to simulate the genotype-to-phenotype relationship. MS/MS, tandem mass spectrometry; RNA-seq, RNA sequencing.

frames (for example, peptides), non-canonical proteins and regulatory sequences (for example, promoters, terminators, ribosomal binding sites and small RNAs) remains difficult, owing to the lack of experimental validation, limiting the use of genomics alone to select a suitable production host.

Transcriptomics

Understanding how production is regulated is key to overcome bottlenecks in production and identify strategies that target such bottlenecks. For example, transcriptomics and fluxomics data revealed that heterologous expression of the cellulase CelA in *Streptomyces lividans* resulted in upregulation of genes involved in secretion, making the secretion system a potential bottleneck that can be targeted for strain improvement[89]. Computational tools that can help in the identification of regulatory elements are typically based on experimental data from a limited number of model organisms. Therefore, extrapolating these results to non-model organisms might result in false positives[90]. Consequently, experimental approaches are required to elucidate the regulatory mechanisms of non-model bacteria.

At the transcriptome level, RNA sequencing (RNA-seq) provides information about gene expression under a set of prespecified conditions[91]. For example, RNA-seq was used to determine the genes involved in biofilm formation of *Streptococcus suis* to reveal its pathogenesis and identify potential drug targets[92]. Furthermore, the transcriptome can provide information on the location of regulatory elements (that is, promoters and terminators). However, RNA-seq can typically not differentiate between processing sites and transcriptional boundaries (that is, transcriptional start and termination sites). Therefore, experimental approaches, such as cappable-seq and terminator 5′-phosphate-dependent exonuclease treatment, are used[93,94] to target the characteristic 5′ triphosphate group at the transcriptional start site and accurately identify its position. Nevertheless, the identification of transcriptional termination sites and transcriptional units remains difficult owing to the short reads, which cannot capture full-length transcripts. To overcome this limitation, cappable-sequencing methodology can be combined with long-read sequencing technologies (that is, single molecule, real-time (SMRT)-cappable-seq and Oxford Nanopore Technologies (ONT)-cappable-seq)[95,96]. However, although long reads might enable the sequencing of full-length transcripts, degradation at the 3′ end limits the accurate identification of transcriptional termination sites[97].

RNA-seq can also be applied for identifying non-coding RNAs, which can occur as free-floating species that bind to genes or regulatory elements to fine-tune expression. Furthermore, they can act as riboswitches in the untranslated region of transcripts, where they form secondary structures that can interfere with transcription or translation[98].

Proteomics and metabolomics

Proteomics and metabolomics rely on mass spectrometry to identify compounds and biosynthesis pathways. In addition, reverse genetics approaches, such as gene knockout screens, can help link the genome to the proteome and metabolome[99]. In contrast to the DNA sequence, which typically remains stable over time, the protein and metabolite content of a cell greatly depend on the conditions under which the bacteria are analysed. For example, studying the proteome of *E. coli*

under 22 different conditions allowed the validation of only 55% of the predicted open reading frames[100]. Moreover, mass spectrometry sample preparation is time-consuming, limiting the number of experimental conditions that can be compared[101]. Alternatively, in acoustic ejection mass spectrometry, sound waves are applied to eject nanolitre droplets for analysis in a mass spectrometer, substantially increasing the sample throughput by facilitating sample preparation (that is, analysis can be performed in complex matrices without the need of extraction) for targeted proteomics and metabolomics[102,103]. Nevertheless, targeted approaches can be used only if the compound of interest is known. By contrast, untargeted approaches can provide an overview of the total protein or metabolite content in a cell. Similar to genome annotations, untargeted metabolomics relies on spectral data from previously identified compounds to identify metabolites in a sample. However, spectral databases have only recently become available to share metabolomics data, limiting the number of entries and thus the identification of metabolites[104].

Proteomics data can also be combined with in vivo translation data in an approach called ribosomal profiling (Ribo-seq) to identify regulatory mechanisms at the translational level. By fixing the ribosome to the mRNA, the positions to which the ribosomes bind can be identified, indicating active translation[105]. To fixate ribosomes at the translation initiation sites, retapamulin-assisted ribosomal profiling can be applied, which allowed the identification of more than a hundred non-conventional translation events in *E. coli*[106]. The translation initiation site affects the translation initiation rate and thus regulates protein synthesis. However, translation can also be regulated at other stages (that is, elongation, termination or ribosome recycling), which remains poorly understood and requires further research to reveal bottlenecks that could limit the production capacity of the host.

At the metabolome level, metabolic flux analysis, or fluxomics, can aid in the identification of metabolic bottlenecks in production. Mathematical models underlying this approach provide an overview of the energy distribution throughout the cell. For example, in ^{13}C-metabolic flux analysis, an isotope-labelled feedstock is used to grow bacteria and obtain an overview of the metabolites that are produced under the selected environmental conditions[107]. Using a combination of ^{13}C- and ^{2}H-metabolic flux analyses, the periplasmic gluconate shunt in *P. putida* could be identified as an alternative pathway to glycolysis to generates pyruvate, revealing new insights into the dynamics of glucose metabolism in *P. putida*[108]. Therefore, elucidating flux distribution can aid in rewiring the metabolism and optimizing flux towards the compound of interest.

Multi-omics integration

Although individual omics approaches can provide valuable insights into different aspects of cellular processes, their integration is crucial for obtaining a holistic understanding of bacterial metabolism. For example, the combination of transcriptomic, proteomic and metabolomic data of *P. putida* KT2440, grown on glucose in an oxygen-free anodic bio-electrochemical system, revealed that reduced acetate synthesis improves the production of 2-ketogluconate[109]. Similarly, transcriptomics and metabolomics data of *Lactiplantibacillus plantarum* can be integrated to study its stress response during the production of linoleic acid, which has beneficial effects on human health, including the maintenance of body weight[110].

Omics technologies can be integrated by generating a digital footprint of the bacterial metabolism to predict its performance for a certain application, for example, using genome-scale metabolic models. Such computational models can simulate bacterial metabolism under specified environmental conditions, taking into account all biochemical reactions known for the organism of interest[111]. As such, genome-scale metabolic models can predict phenotypical changes that emerge from varying environmental conditions[112]. In addition to biochemical reactions, these models can integrate data on gene regulation (for example, transcriptomics, Ribo-seq and fluxomics) and can thus be applied not only to investigate bacterial metabolism but also to optimize production.

If the exact wiring of metabolic networks is not yet known, gap-filling algorithms can be applied to add biochemical reactions and complete the model. However, these biochemical reactions remain hypothesized and might fail to predict the organism's capacity to produce or metabolize a certain compound[113]. For example, the prediction by a genome-scale metabolic model that *Rhodoccoccus jostii* RHA1 uses ethanol as a favourable carbon source over acetate for the production of microbial lipids could not be experimentally confirmed[114]. Experimental data can aid in refining and updating genome-scale metabolic models[115]. For example, genome-scale metabolic models have been optimized for the Gram-negative non-model organism *Zymomonas mobilis* to create microbial cell factories for the optimized production of several chemicals, including succinic acid, ethylene glycol, glycolic acid, 1,4-butanediol and 1,3-propanediol[116].

In addition, artificial intelligence, machine learning and knowledge engineering enable the integration of multi-omics data[117]. In particular, information obtained from model organisms can be combined with insights from non-model bacteria to train such algorithms. Importantly, training data need to be of high quality to build reliable models, which will require the establishment of reporting guidelines. Once optimized, these models might have a crucial role in the learning phase of the design–build–test–learn cycle[118] by reducing turnover rates to identify patterns and engineering targets within the complex metabolic networks of non-model bacteria[119]. For example, machine learning combined with multi-omics data allowed the elucidation of the genotype–phenotype relationship of *Synechococcus* sp. PCC 7002 under different growth conditions to identify metabolic bottlenecks that can be targeted and facilitate efficient phototrophic cultivation of *Synechococcus* sp. PCC 7002[120].

Phage-based genome refactoring

The development of non-model organisms for synthetic biology requires vectors that can replicate in the non-model host and effective transformation methods to deliver the genetic material. For example, in the case of *Lacticaseilactobacilli*, electroporation remains the only effective transformation method; however, its suboptimal performance limits DNA transfer into the cell, thereby complicating the implementation of genome engineering strategies in this organism[121]. Moreover, reliable genetic engineering tools must be established to enable precise modifications of the host genome. Finally, well-characterized genetic building blocks (for example, promoters, terminators and transcription factors) are essential for building synthetic circuits[14,26,122] (Table 1).

The majority of synthetic biology tools and genetic parts have been developed and optimized for model organisms, and their transfer to non-model hosts remains challenging. For example, the activity of promoters, such as *Pem7*, and the performance of expression systems, such as *pET* or *XylS/Pm*, can markedly vary across bacterial species[123–125]. In addition, integrases, such as ϕC31 and λ, rely on specific genomic attachment sites that may be absent in non-model organisms, limiting

Table 1 | Synthetic biology tools and parts for genome engineering

Function	Subtype	Applications in synthetic biology	Limitations	Strengths	Refs.
Promotors Responses vary depending on organism; orthogonality with respect to RNA polymerases	Constitutive	Synthetic (e.g., P_{em7}, P_{14a-g}, $P_{porin/polib}$, $P_{n25-n26}$ and $P_{s12(c)}$); native (e.g., P_{J23119}, P_{phaC1}, P_{gap} and P_{pdc}, P_{eno})	Typically screened for high strength; may result in toxicity	High expression levels; does not require inducers	176–181
	Inducible	Bacterially derived (e.g., LacI/PlacUV5, XylS/Pm, AraC/PBAD, rhaRS/Prham and pSynSens)	Inducers can be expensive	Additional level of regulation; timed expression (e.g., useful for growth decoupling)	182–186
	Orthogonal	Phage-derived (e.g., pET, pPUT, pL and pSP6)	Host-specific response; additional engineering required to incorporate all elements	Function orthogonal to the host's transcriptional machinery	124,187–189
Terminators Response varies depending on organism; degradation at 3′ end of transcripts renders identification of terminators difficult	Factor dependent	Rho-dependent	Cofactor is required to perform expression	Additional layer of regulation that can aid timed expression	190
	Factor independent	Phage-derived (e.g., λ: T_0, LUZ100: T_{19}, LUZ7: T_{50}, T7: TE); synthetic (T7 variants); bacterially derived (*E. coli* rrnB (T1))	Stability of the RNA secondary structure is affected by temperature, pH and ionic strength	Cofactor not required; based on secondary structure	135,191
Ribosome binding site	-	Synthetic (e.g., BCD1-22); native (e.g., *Streptomyces*)	Typically screened for strength; high expression may result in toxicity; experimental methods to identify ribosome binding sites remain limited	High expression levels; synthetic ribosome binding sites are designed such that they do not interfere with the secondary structure of the RNA	192,193
Aptamers	-	tBroccoli, malachite green aptamer, SP6 inhibitory aptamer, T7 inhibitory aptamer	Complex regulation; difficult to identify ligand; in vivo assays to evaluate functionality are lacking	Additional layer of regulation that can aid the development of complex circuits (e.g., biosensors)	194–197
Proteins	Integrases	ϕC31, ϕBT1, TG1, TP901-1, Bxb1, A118, Cre, FLP, λ, XerCD	Issues with orthogonality; integration site needs to be known and present	Development of complex synthetic circuits (e.g., memory and counting devices)	126,127,142, 198,199
	Transcription factors	Activators (e.g., ORF2 of phi31, AraC and SoxR); repressors (e.g., cI, cro, TetR and FadR), LysR	Regulation can be complex and intertwined with multiple processes, which can hinder the construction of synthetic circuits	Additional level of regulation; timed expression (e.g., useful for growth decoupling)	200–203
	Reporters	Fluorescent proteins (e.g., GFP, YFP, mCherry, RFP and BFP); luciferases (DTZ, Firefly and Renilla); chromoproteins (e.g., gfasPurple, amilCP, spisPink and eforRed)	Stability issue at high temperature or pH (e.g., thermophiles and halophiles)	Easy read-out to verify expression levels of regulatory elements enables large screens	204–208
Plasmid backbones	SEVA plasmids	pSEVA131, pSEVA237c, pSEVA2a2b8Rα	Available cargos are biased towards organisms often used for synthetic biology applications	Backbones can be used in a variety of organisms owing to the variety of available cargos	209
Genome editing tools (e.g., insertions, deletions and substitutions)	CRISPR–Cas engineering	CRISPR–Cas9, CRISPR–Cas3, CRISPR–Cas12a, CRISPR–nCas9, CRISPR–dCas9	Off-target effects; requirement of PAM sequence	Multiplex; high efficiency	128,130, 210,211
	Homologous recombination	Cre/*loxP*, phiC31/*att*, fluorescence-assisted genome engineering	Low efficiency; time-consuming; repetitive regions difficult to edit	No off-target effects	199,212
	Recombineering	λ-Red recombineering, RecT-mediated recombineering	Low efficiency; time-consuming	Short homology arms are required for recombination	213
	Transposons	Site-specific (e.g., Tn7/*glmS*); random mutagenesis (e.g., Tn5 and Mu), CAST; OMEGA	Edits are not clean; also contain transposon sequences; specificity and editing efficiency need to be improved	High-throughput	214,215

Table 1 (continued) | Synthetic biology tools and parts for genome engineering

Function	Subtype	Applications in synthetic biology	Limitations	Strengths	Refs.
Editing gene expression levels (e.g., knockdowns)	Gene silencing	RNA interference (e.g., siRNA and miRNA); xenogeneic silencing (e.g., H-NS, Rok and CpgS)	Challenges with stability and delivery; temporary effect	Essential genes can be targeted	216,217

BFP, blue fluorescent protein; CAST, CRISPR-associated transposase; Cas, CRISPR-associated protein; dCas9, catalytically dead Cas9; DTZ, diphenylterazine; *E. coli*, *Escherichia coli*; GFP, green fluorescent protein; mCherry, monomeric red fluorescent protein; miRNA, microRNA; nCas9, Cas9 nickase; OMEGA, Obligate Mobile Element Guided Activity; ORF2, open reading frame 2; RFP, red fluorescent protein; SEVA plasmids, Standard European Vector Architecture plasmids; siRNA, small interfering RNA; YFP, yellow fluorescent protein.

their utility[126,127]. Alternatively, analogous tools and parts can be directly isolated from the genome of non-model species or their phages, serving as blueprints for the rational design of customized synthetic biology toolkits. For example, owing to its high tolerance for chemicals and solvents, *P. putida* has been explored as a chassis for biotechnological applications[75], including the development of genome engineering tools, such as fluorescence-assisted genome engineering, CRISPR–Cas3 engineering as well as a cytidine base editor[128–130]. Furthermore, regulatory elements derived from bacteriophage phi15 have been used to develop a *Pseudomonas*-specific counterpart of the pET system, enabling tightly controlled recombinant protein expression within this genus[124].

Bacteriophages as sources of regulatory elements

Bacteriophages provide a rich source of molecular biology tools and circuits not only for model organisms, such as *E. coli*, but also for non-model microbes (Table 1). These viruses specifically infect bacteria, hijacking their metabolism to propagate and form new virus particles[131]. Bacteriophages are equipped with a variety of genetic building blocks (that is, promoters, terminators, integrases and repressors) that function independently from their bacterial host and that are evolutionarily optimized to work well within that host. Therefore, their genomes can be mined to design a toolbox for non-model organisms that functions orthogonally to their native metabolism[132,133].

Phages are also valuable for the discovery of regulatory elements, for example, for the identification and functional characterization of phage-derived promoters and terminators (Fig. 3). Here, a transcriptomics approach (that is, ONT-cappable-seq) can be used to precisely locate regulatory elements in the phage genome[95,134]. A set of fluorescence-based in vivo assays can then be applied to validate the strength and functionality of these regulatory elements in the host of interest[135,136]. In the case of phage-derived promoters, a distinction needs to be made between host-specific phage promoters (that is, those recognized by the host RNA polymerase) and phage-specific phage promoters, whose expression requires the presence of a phage-encoded RNA polymerase. Cloning the promoter upstream of a fluorescent reporter (for example, GFP) can aid in assessing its strength[136] (Fig. 3). By contrast, for phage-derived terminators, the termination activity, which is a measure for the amount of read-through at the termination site, can be determined by cloning the phage-derived terminator in between two fluorescent proteins[135] (Fig. 3). As such, the functionality of genetic building blocks derived from phages can be screened in a rapid, high-throughput manner. Using this approach, a large set of phage promoters and terminators have been identified and their functionality tested in vivo in different *Pseudomonas* species[135,137,138]. Phages can also encode their own transcriptional machinery or sigma-like proteins that redirect the host RNA polymerase to favour phage-derived sequences, which can be used to further fine-tune transcriptional regulation[136,139,140]. Although efforts have mainly focused on transcriptional regulation thus far, phage genomes may also encode translational elements, such as ribosome binding sites and riboswitches, which could be exploited for synthetic biology in non-model hosts[132].

Tools for genome engineering and posttranslational modifications

Phages are also equipped with tools to facilitate genome engineering. For example, temperate phages naturally integrate into the bacterial chromosome using integrases. These enzymes, in particular, phage λ-derived red recombinase, can be exploited to provide access to the genome of non-model bacteria[141] (Table 1). Integrases can also be further engineered or additional ones can be mined from phages to provide accessibility to a larger number of genomes[142,143].

In addition to regulation at the DNA level, posttranslational modifications can fine-tune bacterial metabolism[144]. Posttranslational modifications, such as RNAylation, phosphorylation and acetylation, also occur in phages to regulate bacterial responses related to virus–bacteria interactions as well as bacterial transcription and translation[145]. Insights into the interplay between phages and their hosts could aid in improving phage resistance in bacterial fermentations, as phage contamination remains a risk in bacterial fermentations and solutions remain limited[146]. Finally, the ability of phages to kill their bacterial host can be used for engineering purposes. As genetically engineered organisms should not be released into the environment, microbial kill switches are necessary to contain modified organisms[147]. Phage proteins that are involved in lysing bacteria could be converted into synthetic circuits to improve biosafety[148].

Industrial applications

Endogenous gene expression in native hosts can be applied in biotechnology, including pharmaceutical, agricultural, fresh water, marine and desert biotechnology[149], in particular, in scenarios where heterologous gene expression underperforms owing to physiological constraints (Table 2).

Pharmaceutical industry

Non-model organisms can be used as microbial cell factories, for example, for the identification of new antimicrobials and anticancer agents[150–152]. Owing to their underexplored genomes, non-model organisms offer a valuable source of new bioactive compounds with pharmaceutical potential. However, the complexity of many biosynthetic pathways makes heterologous expression of the gene clusters encoding these compounds challenging. Native hosts are naturally optimized and have evolved innate resistance mechanisms to tolerate high quantities of many of these compounds and are thus suitable for antimicrobial and anticancer agent production (Table 2).

Chemical and agricultural industry

Endogenous expression in non-model organisms can also be applied in the chemical and agricultural industry (Table 2). For example, the methanotrophic bacteria *Methylococcus capsulatus* can be used to produce a protein-rich feed, which can be applied in agriculture, the pet industry and aquacultures[153]. In addition, lactic acid can be produced in *Lactobacilli* for use in bioplastics or as agrochemical in crop production. Moreover, entire bacteria, such as *Bacillus*, *Pediococcus* and *Lactobacillus*, can be applied in wastewater treatment[154]. Bioherbicidal compounds, such as coronatine, can also be produced in microorganisms. However, the production titres in the native host *Pseudomonas syringae* pv. *tomato* DC3000 are low. The coronatine biosynthetic gene cluster can also be heterologously expressed in *P. putida* KT2440, albeit with similarly low production titres. Moreover, heterologous production leads to an imbalance between the coronafacic and coronamic acid moieties, which together form coronatine. The engineered *P. putida* strain produces an excess of coronafacic acid relative to coronamic acid, which compromises the eventual titre of coronatine[155].

Furthermore, pigments for the textile and paint industry can be produced using colourants derived from *Streptomyces coelicolor*[156]. In addition, prodigiosin is a red pigment produced by *Serratia marcescens* at titres of up to 36 g l^{-1} (ref. 157). However, *S. marcescens* is pathogenic, and the regulatory mechanisms that drive prodigiosin production in this species are complex and intertwined with many metabolic processes, which complicates heterologous expression[158,159]. Alternatively, prodigiosin can be heterologously produced in *P. putida* KT2440, albeit at a low production titre of 0.094 g l^{-1}, probably owing to a stress response[160].

Oleaginous non-model bacteria can be valuable as palm oil substitutes and oil sources[117,161,162]. In addition, *Rhodococcus opacus* might serve as host for biodiesel production, because it can accumulate lipids of up to 70% of its dry weight. These lipids can be used as feedstocks for the production of biodiesel owing to their high calorific value[163]. Biosurfactants, which are commonly applied as detergents or emulsifiers, can also be synthetized by non-model organisms. For example, *Burkholderia* and *Pseudomonas* strains are natural producers of rhamnolipids, a class of biosurfactants. These native strains can achieve production titres up to 2.5 times higher than those observed in heterologous expression systems using model organisms[164,165]. Owing to the antimicrobial properties of rhamnolipids, it is hypothesized that native

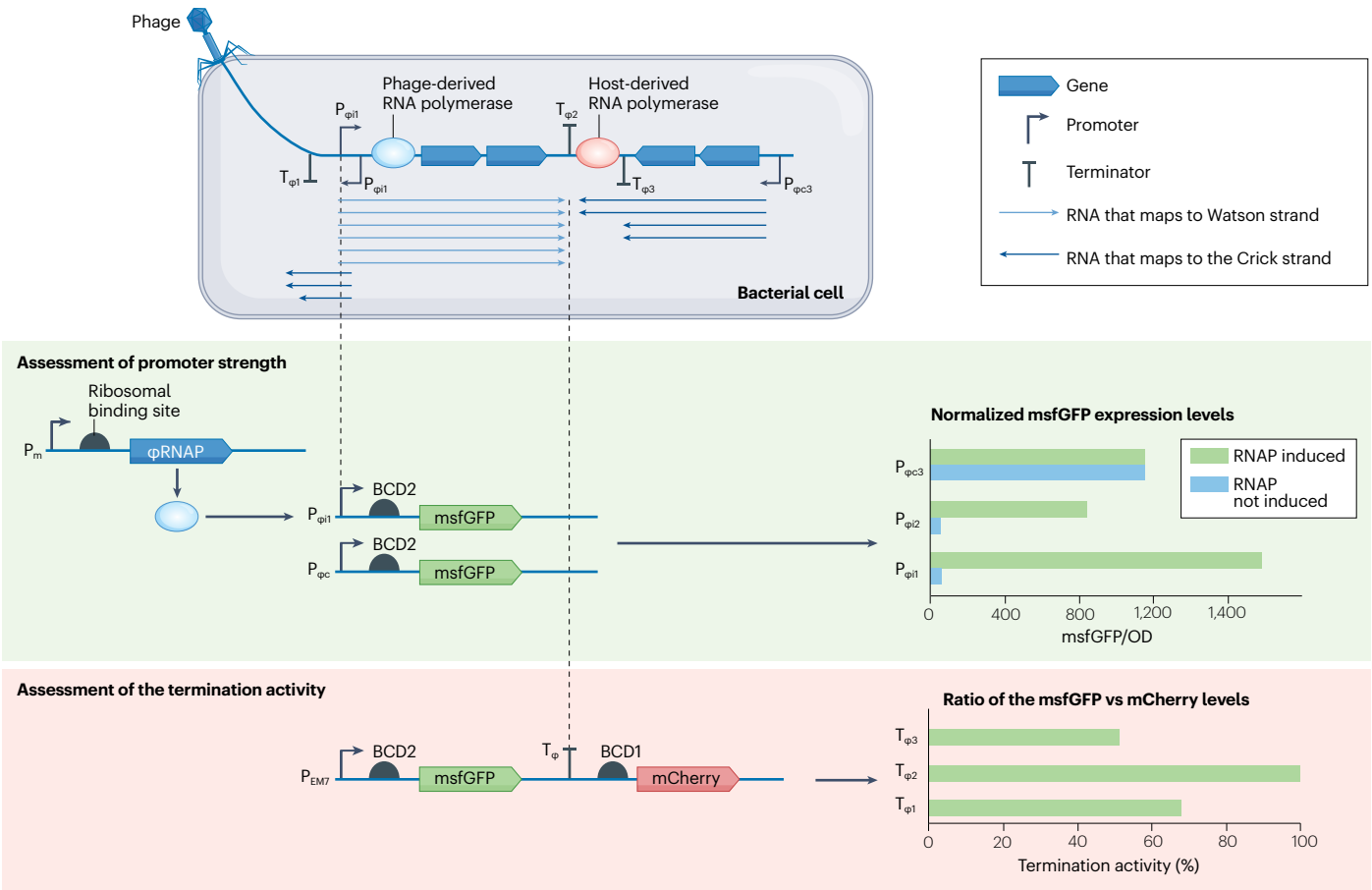


Fig. 3 | Pipeline to identify and test phage-derived regulatory elements. Transcriptomics is performed to identify promoter and terminator regions. Subsequently, their functionality can be tested in fluorescence-based assays to assess their functionality in the host of interest. In the case of promoters, the level of fluorescence intensity is assessed, whereas in the case of terminators, the termination activity is calculated. The latter is a measure of the fluorescence intensity of two subsequent fluorescent proteins, where a construct with a terminator is compared to a vector without terminator[135]. BCD, bicistronic design; GFP, green fluorescent protein; OD, optical density; P_{EM7}, constitutive promoter; P_m, inducible promoter; $P_{\phi c}$, host-specific phage promoter; $P_{\phi i}$, phage-specific phage promoter; RNAP, RNA polymerase; T_ϕ, phage terminator.

Table 2 | Industrial applications of non-model bacteria and their metabolites

Industrial branch	Application	Compound/ enzyme	Microorganism	Examples of products[a]
Pharmaceutical industry	Antibiotic	Kanamycin	*Streptomyces kanamyceticus*	Kantrex
		Tetracycline	*Streptomyces aureofaciens*	Tetracyn, Panmycin
		Chloramphenicol	*Streptomyces venezuela*	Chloromycetin Ophthalmic, Chloroptic
		Erythromycin	*Saccharopolyspora erythraea*	Erythrocin, Erythroped, Erymax
		Vancomycin	*Amycolatopsis orientalis*	Edicin
		Clindamycin	*Streptomyces lincolnensis*	Duac Acne, Treclinac
		Rifamycin	*Amycolatopsis rifamycinica*	Aemcolo
		Cycloserine	*Streptomyces orchidaceus*	Seromycin
		Pristinamycin	*Streptomyces pristinaespiralis*	Pyostacine
		Fosfomycin	*Streptomyces fradiae*	Monurol, Monuril, Ivozfo
		Meroprenem	*Streptomyces cattleya*	Merrem
		Daptomycin	*Streptomyces roseosporus*	Cubicin, Dapzura RT
		Fidaxomicin	*Dactilosporangium aurantiacum* ssp. *hamdensis*	Dificlir
		Gramicidin	*Bacillus brevis*	Neosporin
		Bacitracin	*Bacillus subtilis*	BACiiM
		Colistin	*Paenibacillus polymyxa*	Xylistin, Coly-Mycin M, Colobreathe
		Mupirocin	*Pseudomonas fluorescens*	Bactroban, Centany
		Aztreonam	*Chromobacterium violaceum*	Azaxtam
	Anticancer agent	Bleomycin	*Streptomyces verticillus*	Blenoxane
		Dactinomycin	*Streptomyces parvullus*	Cosmegen
		Mitomycin C	*Streptomyces caespitosus*	Jelmyto, Mitosol
		Doxorubicin	*Streptomyces peucetius* var. *caesius*	Adriblastina, Caelyx, Doxorubin
		Daunorubicin	*Streptomyces peucetius*	Cerubidine
		Epirubicin	*Streptomyces peucetius*	Ellence
	Antifungal	Clotrimazole	*Streptomyces antibioticus*	Canesten
		Nystatin	*Streptomyces noursei*	Mycostatin, Nyamyc
		Amphotericin	*Streptomyces nodosus*	Fungizone, Mysteclin-F, AmBisome
	Antihelminthic	Ivermectin	*Streptomyces avermitilis*	Stromectol
	Probiotics	–	*Limosilactobacillus reuteri*	Pylopass
		–	*Lactobacilli*	ACN, VGN, FNG
Chemical industry	Pigments	Anthranilic acid	*Streptomyces coelicolor*	Pili biobased anthranolic acid
	Cleaning agents	–	Combination of 5 bacillus strains	Microvia
	Antiaging compound	Phytoene	*Deinococcus geothermalis*	Phyt-N-Resist
Agricultural industry	Biologicals	–	Combination of *Penicillium bilaiae*, *Bacillus amyloliquefaciens* and *Trichoderma virens*	BioniQ
		–	*Bradyrhizobium arachis*	Cell-Tech
		–	*Bacillus amyloliquefaciens*	Taegro 2
		–	*Bacillus thuringiensis*	VectoBak, Dipel, FlorBac
		–	*Bacillus amyloliquefaciens*	Toreda
		–	Consortium of aerobic bacteria in combination with *Rhodobacter* sp.	AMF

Table 2 (continued) | Industrial applications of non-model bacteria and their metabolites

Industrial branch	Application	Compound/ enzyme	Microorganism	Examples of products[a]
Agricultural industry (continued)	Probiotics	–	*Bacillus subtilis*	Alterion
	Protein-rich feed	–	*Methylococcus capsulatus*	Uniprotein
	Agrochemicals	Abamectin	*Streptomyces avermitilis*	Apache, Vivid II
		Spinosad	*Saccharopolyspora spinosa*	Tracer, Conserve
		Bilanafos	*Streptomyces hygroscopicus*	Bialaphos
	Antifouling agent	Entire secretome	*Pseudomonas* strain PF-11	BMX-11
Pet industry	Antihelminthic	Milbemycin oxime	*Streptomyces hygroscopicus aureolacrimosus*	Interceptor
	Ammonia removal in ponds	–	Nitrifying bacteria	Bactocool
Food industry	Low-lactose dairy products	Lactase	*Kluyveromyces lactis*	Lactozym, HA-Lactase, Maxilact
	Fish substitutes	Cellulose	*Komagataeibacter*, *Acetobacter*, and *Gluconacetobacter*	Tuna, Scallops
	Pulque (agave wine)	–	*Zymomonas mobilis*	The Queen La Reina Pulque
	Probiotics	–	*Lacticaseibacillus casei*, *Lactobacillus acidohpilus*, *Lactiplantibacillus plantarum*, *Bifidobacterium bifidum*	Activia probiotic yoghurt
		–	*Lactobacillus acidophilus*, *Bifidobacterium animalis* subsp. *lactis* and *Lacticaseibacillus rhamnosus*	Optimel Probiotica

[a]Product names are limited to three examples.

producers have evolved enhanced tolerance to these compounds, offering a potential advantage over heterologous hosts.

Industrial waste streams

Industrial waste streams might be reduced by adopting circular processes to recycle and reuse waste. In particular, industrial waste streams that contain nutrients can serve as feedstocks for bacteria to generate valuable new products. For example, lignin, with an annual global production of 34 billion tons[166], can be degraded by a variety of bacterial species, including *P. putida*, *E. coli* and *Corynebacterium glutamicum*. These microorganisms are capable of converting lignin into high-value compounds, such as vanillin or coumaric acid[166]. Crude glycerol and fruit peel waste streams can also be upcycled through microbial fermentation. These materials are rich in carbohydrates and, thus, require minimal supplementation to support bacterial metabolism[167], thereby reducing costs. However, the exploration of waste streams as feedstocks remains limited owing to the metabolic constraints of model organisms, which often lack the metabolic capacities to grow on these substrates[167].

Outlook

Endogenous gene expression in non-model organisms, which have been evolutionary optimized to produce certain proteins or metabolites in response to environmental stimuli, might offer an alternative route in synthetic biology[26,75,168]. In particular, multi-omics techniques enable the identification of bacteria for specific synthetic biology and biotechnological applications[168] (Box 1). The use of non-model organisms, however, requires a shift to fundamental rather than

Box 1 | High-throughput versus rational-based approaches

The microbial synthetic biology market is expanding. With a compound annual growth rate of 25%, the global market is projected to be worth US$148.93 billion by 2033[223]. For example, Ginkgo Bioworks has developed a service platform to engineer microbial strains for the production of compounds of interest. With the implementation of high-throughput automated strategies, a multitude of strains can be screened to optimize production. This platform runs on a limited set of microorganisms, which are trained to produce proteins or metabolites by adaptive laboratory evolution[224]. During this process, bacteria are exposed to specific environmental conditions to improve their survival and production rates under the selected conditions. However, adaptive laboratory evolution remains unpredictable and might cause mutations that do not favour the application of interest. Therefore, non-selective, high-throughput methods have to be applied to screen a large number of strains[225], requiring a large amount of resources. Furthermore, the number of generations required to reach an improved production strain are unpredictable, which makes this approach time-consuming and cost-prohibitve[226].

Alternatively, a deliberate selection of chassis can be tested for a given application to reduce the costs and time required to optimize microbial fermentations. This rational-based approach can aid in making targeted alterations in the bacterial genome to optimize production. Instead of bypassing evolution, fine-tuned, inherent processes can be exploited. For example, Geno(matica) applies this approach to improve production titres and for sustainable microbial fermentation[226]. Production costs remain a key hurdle in the marketing of microbial cell factories. Therefore, rational-based engineering approaches based on endogenous expression should be considered in microbial synthetic biology.

application-oriented research to improve our understanding of non-model bacteria and develop synthetic biology tools to access their genomes. In part, phages can offer a solution, as their genomes are equipped with genetic building blocks that are optimized to function in their bacterial hosts[132].

Synthetic biology has largely concentrated on replacing traditional petrochemical processes. As a result, bacteria have been engineered as microbial cell factories to produce a range of chemicals for use in agriculture, medicine and industrial biotechnology. However, even with optimized fermentation strategies, microbial production titres often fall short of those achieved through conventional chemical synthesis, limiting their competitiveness in certain applications. Accordingly, scale-up remains a key technical bottleneck to be addressed. In addition to high operational costs and expensive growth media, culture conditions substantially differ upon scale up, which should be taken into account in the development of a production strain[117]. To facilitate scale-up, bioreactors can be equipped with sensors (for example, optical and electrical sensors) to monitor the production process[169]. Furthermore, environmental bioreactor conditions can be mimicked by laboratory-scale equipment to optimize microbial cell factories as industrial-scale fermentation hosts[170]. Downstream processing operations can also vary in complexity and add to the production costs[171]. Moreover, the genetic stability of the strain needs to be considered, as bacteria can have spontaneous mutation rates of 10^{-9}–10^{-10} per nucleotide per generation[172]. As growth in bioreactors induces stress, these mutation rates can further increase. Therefore, the genetic stability of the strain should be traced over time to ensure steady production levels.

The application of non-model organisms in the field requires thorough knowledge of their biology. Under the current regulatory framework in Europe, only products derived from genetically modified microorganisms (GMMs) in contained use, such as in industrial fermentation, are permitted for commercialization. By contrast, the release and marketing of GMMs, particularly in healthcare and agriculture, are subject to much stricter regulations and are often restricted or prohibited[173]. A key concern is safety with regard to human health and the environment. Therefore, risk assessments are in place to assess their safety concerning toxicity, horizontal gene transfer and traceability[174]. This intensive risk assessment process, in combination with the lack of centralized regulation both in the EU and globally, might hamper the investment in technologies that explore GMMs[173,175]. Nevertheless, GMMs are being explored for a variety of applications, including live-attenuated vaccines, probiotics and bacterial batteries[10–12].

Published online: 24 September 2025

References

1. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
2. Lorenzo, V. de & Danchin, A. Synthetic biology: discovering new worlds and new words. The new and not so new aspects of this emerging research field. *EMBO Rep.* **9**, 822 (2008).
3. Xie, M. & Fussenegger, M. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nat. Rev. Mol. Cell Biol.* **19**, 507–525 (2018).
4. Adamala, K. P. et al. Present and future of synthetic cell development. *Nat. Rev. Mol. Cell Biol.* **25**, 162–167 (2024).
5. Volk, M. J. et al. Metabolic engineering: methodologies and applications. *Chem. Rev.* **123**, 5521–5570 (2023).
6. David, F. et al. A perspective on synthetic biology in drug discovery and development — current impact and future opportunities. *SLAS Discov.* **26**, 581–603 (2021).
7. Okoro, V., Azimov, U. & Munoz, J. Recent advances in production of bioenergy carrying molecules, microbial fuels, and fuel design — a review. *Fuel* **316**, 123330 (2022).
8. Kumari, R., Singha, L. P. & Shukla, P. Biotechnological potential of microbial bio-surfactants, their significance, and diverse applications. *FEMS Microbes* **4**, xtad015 (2023).
9. Wang, L., Zang, X. & Zhou, J. Synthetic biology: a powerful booster for future agriculture. *Adv. Agrochem.* **1**, 7–11 (2022).
10. Jalili, P., Ala, A., Nazari, P., Jalili, B. & Ganji, D. D. A comprehensive review of microbial fuel cells considering materials, methods, structures, and microorganisms. *Heliyon* **10**, e25439 (2024).
11. Mugwanda, K. et al. Recent advances in genetic tools for engineering probiotic lactic acid bacteria. *Biosci. Rep.* **43**, BSR20211299 (2023).
12. Cubillos-Ruiz, A. et al. Engineering living therapeutics with synthetic biology. *Nat. Rev. Drug Discov.* **20**, 941–960 (2021).
13. de Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the bioeconomy: define what a synthetic biology chassis is! *N. Biotechnol.* **60**, 44–51 (2021).
14. Adams, B. L. The next generation of synthetic biology chassis: moving synthetic biology from the laboratory to the field. *ACS Synth. Biol.* **5**, 1328–1330 (2016).
15. Pontrelli, S. et al. *Escherichia coli* as a host for metabolic engineering. *Metab. Eng.* **50**, 16–46 (2018).
16. Rosano, G. L., Morales, E. S. & Ceccarelli, E. A. New tools for recombinant protein production in *Escherichia coli*: a 5-year update. *Protein Sci.* **28**, 1412–1422 (2019).
17. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
18. Yang, D., Park, S. Y., Park, Y. S., Eun, H. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol.* **38**, 745–765 (2020).
19. Pouresmaeil, M. & Azizi-Dargahlou, S. Factors involved in heterologous expression of proteins in *E. coli* host. *Arch. Microbiol.* **205**, 212 (2023).
20. İncir, İ & Kaplan, Ö *Escherichia coli* as a versatile cell factory: advances and challenges in recombinant protein production. *Protein Expr. Purif.* **219**, 106463 (2024).
21. Mahalik, S., Sharma, A. K. & Mukherjee, K. J. Genome engineering for improved recombinant protein expression in *Escherichia coli*. *Microb. Cell Factories* **13**, 177 (2014).
22. Olicón-Hernández, D. R. et al. Fundaments and concepts on screening of microorganisms for biotechnological applications. Mini review. *Curr. Microbiol.* **79**, 373 (2022).
23. Lara, A. R., Galindo, E., Ramírez, O. T. & Palomares, L. A. Living with heterogeneities in bioreactors. *Mol. Biotechnol.* **34**, 355–381 (2006).
24. Mu, X. & Zhang, F. Diverse mechanisms of bioproduction heterogeneity in fermentation and their control strategies. *J. Ind. Microbiol. Biotechnol.* **50**, kuad033 (2023).
25. Snoeck, S., Guidi, C. & De Mey, M. "Metabolic burden" explained: stress symptoms and its related responses induced by (over)expression of (heterologous) proteins in *Escherichia coli*. *Microb. Cell Factories* **23**, 96 (2024).
26. Hwang, S., Joung, C., Kim, W., Palsson, B. & Cho, B.-K. Recent advances in non-model bacterial chassis construction. *Curr. Opin. Syst. Biol.* **36**, 100471 (2023).
27. Liu, J., Wang, X., Dai, G., Zhang, Y. & Bian, X. Microbial chassis engineering drives heterologous production of complex secondary metabolites. *Biotechnol. Adv.* **59**, 107966 (2022).
28. Nagata, S. et al. Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. *Nature* **284**, 316–320 (1980).
29. Carroll, A. L., Desai, S. H. & Atsumi, S. Microbial production of scent and flavor compounds. *Curr. Opin. Biotechnol.* **37**, 8–15 (2016).
30. Chance, R. E. & Frank, B. H. Research, development, production, and safety of biosynthetic human insulin. *Diabetes Care* **16**, 133–142 (1993).
31. Wang, C., Pfleger, B. F. & Kim, S.-W. Reassessing *Escherichia coli* as a cell factory for biofuel production. *Curr. Opin. Biotechnol.* **45**, 92–103 (2017).
32. Jiang, R. et al. Strategies to overcome the challenges of low or no expression of heterologous proteins in *Escherichia coli*. *Biotechnol. Adv.* **75**, 108417 (2024).
33. Sharp, P. M. & Li, W. H. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281 (1987).
34. Al-Hawash, A. B., Zhang, X. & Ma, F. Strategies of codon optimization for high-level heterologous protein expression in microbial expression systems. *Gene Rep.* **9**, 46–53 (2017).
35. Liu, Y. A code within the genetic code: codon usage regulates co-translational protein folding. *Cell Commun. Signal.* **18**, 145 (2020).
36. Love, A. M. & Nair, N. U. Specific codons control cellular resources and fitness. *Sci. Adv.* **10**, eadk3485 (2024).
37. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
38. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
39. Greunke, C. et al. Direct pathway cloning (DiPaC) to unlock natural product biosynthetic potential. *Metab. Eng.* **47**, 334–345 (2018).
40. Larionov, V. et al. Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl Acad. Sci. USA* **93**, 491–496 (1996).
41. Shizuya, H. et al. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA* **89**, 8794 (1992).
42. Noskov, V. N. et al. A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* **4**, 16 (2003).
43. Kouprina, N. & Larionov, V. Transformation-associated recombination (TAR) cloning and its applications for gene function; genome architecture and evolution; biotechnology and biomedicine. *Oncotarget* **14**, 1009 (2023).

44. Nah, H.-J., Woo, M.-W., Choi, S.-S. & Kim, E.-S. Precise cloning and tandem integration of large polyketide biosynthetic gene cluster using Streptomyces artificial chromosome system. *Microb. Cell Factories* **14**, 140 (2015).
45. Kudo, K. et al. Capability of a large bacterial artificial chromosome clone harboring multiple biosynthetic gene clusters for the production of diverse compounds. *J. Antibiot.* **77**, 288–298 (2024).
46. Kakirde, K. S. et al. Gram negative shuttle BAC vector for heterologous expression of metagenomic libraries. *Gene* **475**, 57–62 (2011).
47. Zhang, J. J., Tang, X., Zhang, M., Nguyen, D. & Moore, B. S. Broad-host-range expression reveals native and host regulatory elements that influence heterologous antibiotic production in gram-negative bacteria. *mBio* **8**, e01291-17 (2017).
48. Ren, J., Karna, S., Lee, H.-M., Yoo, S. M. & Na, D. Artificial transformation methodologies for improving the efficiency of plasmid DNA transformation and simplifying its use. *Appl. Microbiol. Biotechnol.* **103**, 9205–9215 (2019).
49. Singh, J. A., Noorbaloochi, S., MacDonald, R. & Maxwell, L. J. Chondroitin for osteoarthritis. *Cochrane Database Syst. Rev.* **2015**, CD005614 (2015).
50. Zhao, C. et al. Reprogramming metabolic flux in *Escherichia coli* to enhance chondroitin production. *Adv. Sci.* **11**, 2307351 (2024).
51. Boo, S. H. & Kim, Y. K. The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.* **52**, 400–408 (2020).
52. Zhang, M. M., Wang, Y., Ang, E. L. & Zhao, H. Engineering microbial hosts for production of bacterial natural products. *Nat. Prod. Rep.* **33**, 963–987 (2016).
53. Parks, D. H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).
54. Fragata, I., Blanckaert, A., Dias Louro, M. A., Liberles, D. A. & Bank, C. Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82 (2019).
55. Bajić, D., Vila, J. C. C., Blount, Z. D. & Sánchez, A. On the deformability of an empirical fitness landscape by microbial evolution. *Proc. Natl Acad. Sci. USA* **115**, 11286–11291 (2018).
56. Sah, S., Krishnani, S. & Singh, R. *Pseudomonas* mediated nutritional and growth promotional activities for sustainable food security. *Curr. Res. Microb. Sci.* **2**, 100084 (2021).
57. Qin, S. et al. *Pseudomonas aeruginosa*: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. *Signal. Transduct. Target. Ther.* **7**, 1–27 (2022).
58. Xin, X.-F., Kvitko, B. & He, S. Y. *Pseudomonas syringae*: what it takes to be a pathogen. *Nat. Rev. Microbiol.* **16**, 316–328 (2018).
59. Aida, H. & Ying, B.-W. Data-driven discovery of the interplay between genetic and environmental factors in bacterial growth. *Commun. Biol.* **7**, 1–12 (2024).
60. Ying, B.-W., Seno, S., Matsuda, H. & Yomo, T. A simple comparison of the extrinsic noise in gene expression between native and foreign regulations in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **486**, 852–857 (2017).
61. Sleight, S. C., Bartley, B. A., Lieviant, J. A. & Sauro, H. M. Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng.* **4**, 12 (2010).
62. Radde, N. et al. Measuring the burden of hundreds of BioBricks defines an evolutionary limit on constructability in synthetic biology. *Nat. Commun.* **15**, 6242 (2024).
63. Kochhar, N. et al. Perspectives on the microorganism of extreme environments and their applications. *Curr. Res. Microb. Sci.* **3**, 100134 (2022).
64. Febriani, Aura, N., Kemala, P., Saidi, N. & Iqbalsyah, T. M. Novel thermostable lipase produced by a thermo-halophilic bacterium that catalyses hydrolytic and transesterification reactions. *Heliyon* **6**, e04520 (2020).
65. Phuyal, M. et al. Xylanase-producing microbes and their real-world application. *Int. J. Chem. Eng.* **2023**, 3593035 (2023).
66. Wang, K., Cui, B., Wang, Y. & Luo, W. Microbial production of ectoine: a review. *ACS Synth. Biol.* **14**, 332–342 (2025).
67. Mitra, R., Xu, T., Xiang, H. & Han, J. Current developments on polyhydroxyalkanoates synthesis by using halophiles as a promising cell factory. *Microb. Cell Factories* **19**, 86 (2020).
68. Gao, H. et al. The functional roles of Lactobacillus acidophilus in different physiological and pathological processes. *J. Microbiol. Biotechnol.* **32**, 1226–1233 (2022).
69. Gadhvi, M. S., Javia, B. M., Vyas, S. J., Patel, R. & Dudhagara, D. R. Bhargavaea beijingensis a promising tool for bio-cementation, soil improvement, and mercury removal. *Sci. Rep.* **14**, 23976 (2024).
70. Frost, I. et al. The role of bacterial vaccines in the fight against antimicrobial resistance: an analysis of the preclinical and clinical development pipeline. *Lancet Microbe* **4**, e113–e125 (2023).
71. Lentsch, V. et al. "EvoVax" — a rationally designed inactivated *Salmonella* Typhimurium vaccine induces strong and long-lasting immune responses in pigs. *Vaccine* **41**, 5545–5552 (2023).
72. Yang, S. et al. *Zymomonas mobilis* as a model system for production of biofuels and biochemicals. *Microb. Biotechnol.* **9**, 699–717 (2016).
73. Kunakom, S. & Eustáquio, A. S. *Burkholderia* as a source of natural products. *J. Nat. Prod.* **82**, 2018–2037 (2019).
74. Weldon, M. & Euler, C. Physiology-informed use of cupriavidus necator in biomanufacturing: a review of advances and challenges. *Microb. Cell Factories* **24**, 30 (2025).
75. Calero, P. & Nikel, P. I. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* **12**, 98–124 (2019).
76. Trüper, H. G. & Krämer, J. in *The Prokaryotes: a Handbook on Habitats, Isolation, and Identification of Bacteria* (eds Starr, M. P. et al.) 176–193 (Springer, 1981).
77. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
78. Schwengers, O. et al. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb. Genomics* **7**, 000685 (2021).
79. Ahmad, S. et al. The UniProt website API: facilitating programmatic access to protein knowledge. *Nucleic Acids Res.* **53**, W547–W553 (2025).
80. Goldfarb, T. et al. NCBI RefSeq: reference sequence standards through 25 years of curation and annotation. *Nucleic Acids Res.* **53**, D243–D257 (2024).
81. Lobb, B., Tremblay, B. J.-M., Moreno-Hagelsieb, G. & Doxey, A. C. An assessment of genome annotation coverage across the bacterial tree of life. *Microb. Genomics* **6**, e000341 (2020).
82. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
83. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
84. Blin, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **51**, W46–W50 (2023).
85. Terlouw, B. R. et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res.* **51**, D603–D610 (2023).
86. Hauns, S., Alkhnbashi, O. S. & Backofen, R. Deepdefense: annotation of immune systems in prokaryotes using deep learning. *GigaScience* **13**, giae062 (2024).
87. Camargo, A. P. et al. Identification of mobile genetic elements with geNomad. *Nat. Biotechnol.* **42**, 1303–1312 (2024).
88. Wishart, D. S. et al. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.* **51**, W443–W450 (2023).
89. Daniels, W. et al. Transcriptomic and fluxomic changes in *Streptomyces lividans* producing heterologous protein. *Microb. Cell Factories* **17**, 198 (2018).
90. Cassiano, M. H. A. & Silva-Rocha, R. Benchmarking bacterial promoter prediction tools: potentialities and limitations. *mSystems* **5**, e00439–20 (2020).
91. Costa-Silva, J., Domingues, D. & Lopes, F. M. RNA-Seq differential expression analysis: an extended review and a software tool. *PLoS ONE* **12**, e0190152 (2017).
92. Wang, H., Fan, Q., Wang, Y., Yi, L. & Wang, Y. Multi-omics analysis reveals genes and metabolites involved in Streptococcus suis biofilm formation. *BMC Microbiol.* **24**, 297 (2024).
93. Sharma, C. M. et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**, 250–255 (2010).
94. Ettwiller, L., Buswell, J., Yigit, E. & Schildkraut, I. A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome. *BMC Genomics* **17**, 199 (2016).
95. Putzeys, L. et al. Development of ONT-cappable-seq to unravel the transcriptional landscape of *Pseudomonas* phages. *Comput. Struct. Biotechnol. J.* **20**, 2624–2638 (2022).
96. Yan, B., Boitano, M., Clark, T. A. & Ettwiller, L. SMRT-cappable-seq reveals complex operon variants in bacteria. *Nat. Commun.* **9**, 3676 (2018).
97. Wang, X. et al. Processing generates 3′ ends of RNA masking transcription termination events in prokaryotes. *Proc. Natl Acad. Sci. USA* **116**, 4440–4445 (2019).
98. Papenfort, K. & Storz, G. Insights into bacterial metabolism from small RNAs. *Cell Chem. Biol.* **31**, 1571–1577 (2024).
99. Vitale, G. A. et al. Connecting metabolome and phenotype: recent advances in functional metabolomics tools for the identification of bioactive natural products. *Nat. Prod. Rep.* **41**, 885–904 (2024).
100. Schmidt, A. et al. The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
101. Fu, Q., Murray, C. I., Karpov, O. A. & Van Eyk, J. E. Automated proteomic sample preparation: the key component for high throughput and quantitative mass spectrometry analysis. *Mass. Spectrom. Rev.* **42**, 873–886 (2023).
102. Sinclair, I. et al. Novel acoustic loading of a mass spectrometer: toward next-generation high-throughput MS screening. *SLAS Technol.* **21**, 19–26 (2016).
103. Van Puyvelde, B. et al. Acoustic ejection mass spectrometry empowers ultra-fast protein biomarker quantification. *Nat. Commun.* **15**, 5114 (2024).
104. Aron, A. T. et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **15**, 1954–1991 (2020).
105. Brar, G. A. & Weissman, J. S. Ribosome profiling reveals the what, when, where, and how of protein synthesis. *Nat. Rev. Mol. Cell Biol.* **16**, 651–664 (2015).
106. Meydan, S. et al. Retapamulin-assisted ribosome profiling reveals the alternative bacterial proteome. *Mol. Cell* **74**, 481–493.e6 (2019).
107. Crown, S. B. & Antoniewicz, M. R. Publishing ^{13}C metabolic flux analysis studies: a review and future perspectives. *Metab. Eng.* **20**, 42–48 (2013).
108. Volke, D. C., Gurdo, N., Milanesi, R. & Nikel, P. I. Time-resolved, deuterium-based fluxomics uncovers the hierarchy and dynamics of sugar processing by *Pseudomonas putida*. *Metab. Eng.* **79**, 159–172 (2023).
109. Weimer, A. et al. Systems biology of electrogenic *Pseudomonas putida* — multi-omics insights and metabolic engineering for enhanced 2-ketogluconate production. *Microb. Cell Factories* **23**, 246 (2024).
110. Dou, H., Liu, X.-X., Wang, G.-Q., Xiong, Z.-Q. & Ai, L.-Z. Integration of transcriptomics and metabolomics provides comprehensive insights into metabolic and functional alterations in *Lactobacillus plantarum* AR195 exposed to excess linoleic acid. *Food Biosci.* **64**, 105957 (2025).

111. Gu, C., Kim, G. B., Kim, W. J., Kim, H. U. & Lee, S. Y. Current status and applications of genome-scale metabolic models. *Genome Biol.* **20**, 121 (2019).
112. Pandey, V., Hadadi, N. & Hatzimanikatis, V. Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models. *PLoS Comput. Biol.* **15**, e1007036 (2019).
113. Chen, C., Liao, C. & Liu, Y.-Y. Teasing out missing reactions in genome-scale metabolic networks through hypergraph learning. *Nat. Commun.* **14**, 2375 (2023).
114. Chen, K. et al. Electro-biodiesel empowered by co-design of microorganism and electrocatalysis. *Joule* **9**, 101769 (2025).
115. Gong, Z. et al. Genome-scale metabolic network models for industrial microorganisms metabolic engineering: current advances and future prospects. *Biotechnol. Adv.* **72**, 108319 (2024).
116. Yan, X. et al. Paradigm of engineering recalcitrant non-model microorganism with dominant metabolic pathway as a biorefinery chassis. *Nat. Commun.* **15**, 10441 (2024).
117. Yan, X., He, Q., Geng, B. & Yang, S. Microbial cell factories in the bioeconomy era: from discovery to creation. *BioDesign Res.* **6**, 0052 (2024).
118. Kitano, S., Lin, C., Foo, J. L. & Chang, M. W. Synthetic biology: learning the way toward high-precision biological design. *PLoS Biol.* **21**, e3002116 (2023).
119. Goshisht, M. K. Machine learning and deep learning in synthetic biology: key architectures, applications, and challenges. *ACS Omega* **9**, 9921–9945 (2024).
120. Vijayakumar, S., Rahman, P. K. S. M. & Angione, C. A hybrid flux balance analysis and machine learning pipeline elucidates metabolic adaptation in *Cyanobacteria*. *iScience* **23**, 101818 (2020).
121. Fristot, E. et al. An optimized electrotransformation protocol for *Lactobacillus jensenii*. *PLoS ONE* **18**, e0280935 (2023).
122. Bales, M. K., Vergara, M. M. & Eckert, C. A. Application of functional genomics for domestication of novel non-model microbes. *J. Ind. Microbiol. Biotechnol.* **51**, kuae022 (2024).
123. García-Gutiérrez, C. et al. Multifunctional SEVA shuttle vectors for actinomycetes and gram-negative bacteria. *MicrobiologyOpen* **9**, e1024 (2020).
124. Lammens, E.-M. et al. Engineering a phi15-based expression system for stringent gene expression in *Pseudomonas putida*. *Commun. Biol.* **8**, 1–15 (2025).
125. Gawin, A., Valla, S. & Brautaset, T. The XylS/Pm regulator/promoter system and its use in fundamental studies of bacterial gene expression, recombinant protein production and metabolic engineering. *Microb. Biotechnol.* **10**, 702–718 (2017).
126. Landy, A. The λ integrase site-specific recombination pathway. *Microbiol. Spectr.* **3**, MDNA3-0051–2014 (2015).
127. MacDonald, A. I. et al. Variable orthogonality of serine integrase interactions within the ϕC31 family. *Sci. Rep.* **14**, 26280 (2024).
128. Lammens, E.-M. et al. A SEVA-based, CRISPR–Cas3-assisted genome engineering approach for *Pseudomonas* with efficient vector curing. *Microbiol. Spectr.* **11**, e02707–e02723 (2023).
129. Volke, D. C., Wirth, N. T. & Nikel, P. Rapid genome engineering of *Pseudomonas* assisted by fluorescent markers and tractable curing of plasmids. *Bio Protoc.* **11**, e3917 (2021).
130. Volke, D. C., Martino, R. A., Kozaeva, E., Smania, A. M. & Nikel, P. I. Modular (de) construction of complex bacterial phenotypes by CRISPR/nCas9-assisted, multiplex cytidine base-editing. *Nat. Commun.* **13**, 3026 (2022).
131. Salmond, G. P. C. & Fineran, P. C. A century of the phage: past, present and future. *Nat. Rev. Microbiol.* **13**, 777–786 (2015).
132. Lammens, E.-M., Nikel, P. I. & Lavigne, R. Exploring the synthetic biology potential of bacteriophages for engineering non-model bacteria. *Nat. Commun.* **11**, 5294 (2020).
133. Lemire, S., Yehl, K. M. & Lu, T. K. Phage-based applications in synthetic biology. *Annu. Rev. Virol.* **5**, 453–476 (2018).
134. Putzeys, L. et al. Refining the transcriptional landscapes for distinct clades of virulent phages infecting *Pseudomonas aeruginosa*. *microLife* **5**, uqae002 (2024).
135. Lammens, E.-M., Putzeys, L., Boon, M. & Lavigne, R. Sourcing phage-encoded terminators using ONT-cappable-seq for SynBio applications in *Pseudomonas*. *ACS Synth. Biol.* **12**, 1415–1423 (2023).
136. Lammens, E.-M., Feyaerts, N., Kerremans, A., Boon, M. & Lavigne, R. Assessing the orthogonality of phage-encoded RNA polymerases for tailored synthetic biology applications in *Pseudomonas* species. *Int. J. Mol. Sci.* **24**, 7175 (2023).
137. Putzeys, L. et al. Transcriptomics-driven characterization of LUZ100, a T7-like *Pseudomonas* phage with temperate features. *mSystems* **8**, e01189-22 (2023).
138. Poppeliers, J. et al. Assessing the transcriptional landscape of *Pseudomonas* phage 201ϕ2-1: uncovering the small regulatory details of a giant phage. *Microb. Biotechnol.* **17**, e70037 (2024).
139. Schwartz, D. A., Lehmkuhl, B. K. & Lennon, J. T. Phage-encoded sigma factors alter bacterial dormancy. *mSphere* **7**, e00297-22 (2022).
140. Nechaev, S. & Severinov, K. Bacteriophage-induced modifications of host RNA polymerase. *Annu. Rev. Microbiol.* **57**, 301–322 (2003).
141. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009).
142. Zhang, Y., Ba, F., Huang, S., Liu, W.-Q. & Li, J. Orthogonal serine integrases enable scalable gene storage cascades in bacterial genome. *ACS Synth. Biol.* **13**, 3022–3031 (2024).
143. Durrant, M. G. et al. Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat. Biotechnol.* **41**, 488–499 (2023).
144. Pisithkul, T., Patel, N. M. & Amador-Noguez, D. Post-translational modifications as key regulators of bacterial metabolic fluxes. *Curr. Opin. Microbiol.* **24**, 29–37 (2015).
145. Longin, H., Broeckaert, N., van Noort, V., Lavigne, R. & Hendrix, H. Posttranslational modifications in bacteria during phage infection. *Curr. Opin. Microbiol.* **77**, 102425 (2024).
146. Zou, X., Mo, Z., Wang, L., Chen, S. & Lee, S. Y. Overcoming bacteriophage contamination in bioprocessing: strategies and applications. *Small Methods* **9**, 2400932 (2024).
147. Ma, Y., Manna, A. & Moon, T. S. Advances in engineering genetic circuits for microbial biocontainment. *Curr. Opin. Syst. Biol.* **36**, 100483 (2023).
148. Simon, A. J. & Ellington, A. D. Recent advances in synthetic biosafety. *F1000Research* **5**, F1000 Faculty Rev-2118 (2016).
149. Barcelos, M. C. S., Lupki, F. B., Campolina, G. A., Nelson, D. L. & Molina, G. The colors of biotechnology: general overview and developments of white, green and blue areas. *FEMS Microbiol. Lett.* **365**, fny239 (2018).
150. Hutchings, M. I., Truman, A. W. & Wilkinson, B. Antibiotics: past, present and future. *Curr. Opin. Microbiol.* **51**, 72–80 (2019).
151. Atanasov, A. G., Zotchev, S. B., Dirsch, V. M. & Supuran, C. T. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov.* **20**, 200–216 (2021).
152. Orhan, I. E., Suntar, I. & Emerce, E. Leading role of natural products in pharmaceutical, biotechnological, and cosmeceutical applications (GPSS-2021). *Phytochem. Rev.* **22**, 1–2 (2023).
153. Jørgensen, J. B. Exploiting microorganisms for animal feed production. In *7th Malaysia International Agro-Bio Business Conference* (2011).
154. Carpenter, R., Showell, M. S., Barnes, J. & Pal, N. Microbial-based waste water treatment compositions and methods of use thereof. European patent EP2999801B1 (2021).
155. Gemperlein, K. et al. Synthetic biology approaches to establish a heterologous production system for coronatines. *Metab. Eng.* **44**, 213–222 (2017).
156. Landrain, T., Adenis, M. S., Blache, J. & Boissonnat, G. Use of actinorhodine and its derivatives as a coloring agent. WIPO patent WO2018138089A1 (2018).
157. Giri, A. V., Anandkumar, N., Muthukumaran, G. & Pennathur, G. A novel medium for the enhanced cell growth and production of prodigiosin from *Serratia marcescens* isolated from soil. *BMC Microbiol.* **4**, 11 (2004).
158. Kwon, S.-K., Park, Y.-K. & Kim, J. F. Genome-wide screening and identification of factors affecting the biosynthesis of prodigiosin by Hahella chejuensis, using *Escherichia coli* as a surrogate host. *Appl. Environ. Microbiol.* **76**, 1661–1668 (2010).
159. Choi, S. Y., Lim, S., Yoon, K., Lee, J. I. & Mitchell, R. J. Biotechnological activities and applications of bacterial pigments violacein and prodigiosin. *J. Biol. Eng.* **15**, 10 (2021).
160. Domröse, A. et al. Efficient recombinant production of prodigiosin in *Pseudomonas putida*. *Front. Microbiol.* **6**, 972 (2015).
161. Whiffin, F., Santomauro, F. & Chuck, C. J. Toward a microbial palm oil substitute: oleaginous yeasts cultured on lignocellulose. *Biofuels Bioprod. Bioref.* **10**, 316–334 (2016).
162. Nunes, D. D., Pillay, V. L., Van Rensburg, E. & Pott, R. W. M. Oleaginous microorganisms as a sustainable oil source with a focus on downstream processing and cost-lowering production strategies: a review. *Bioresour. Technol. Rep.* **26**, 101871 (2024).
163. Goswami, L., Tejas Namboodiri, M. M., Vinoth Kumar, R., Pakshirajan, K. & Pugazhenthi, G. Biodiesel production potential of oleaginous *Rhodococcus opacus* grown on biomass gasification wastewater. *Renew. Energy* **105**, 400–406 (2017).
164. Beuker, J. et al. High titer heterologous rhamnolipid production. *AMB Express* **6**, 124 (2016).
165. Müller, M. M., Hörmann, B., Kugel, M., Syldatk, C. & Hausmann, R. Evaluation of rhamnolipid production capacity of *Pseudomonas aeruginosa* PAO1 in comparison to the rhamnolipid over-producer strains DSM 7108 and DSM 2874. *Appl. Microbiol. Biotechnol.* **89**, 585–592 (2011).
166. Zhang, Y. et al. Microbial upcycling of depolymerized lignin into value-added chemicals. *BioDesign Res.* **6**, 0027 (2024).
167. Lad, B. C., Coleman, S. M. & Alper, H. S. Microbial valorization of underutilized and nonconventional waste streams. *J. Ind. Microbiol. Biotechnol.* **49**, kuab056 (2021).
168. Bertile, F., Matallana-Surget, S., Tholey, A., Cristobal, S. & Armengaud, J. Diversifying the concept of model organisms in the age of -omics. *Commun. Biol.* **6**, 1–4 (2023).
169. Mitra, S. & Murthy, G. S. Bioreactor control systems in the biopharmaceutical industry: a critical perspective. *Syst. Microbiol. Biomanuf.* **2**, 91–112 (2022).
170. Fink, M., Cserjan-Puschmann, M., Reinisch, D. & Striedner, G. High-throughput microbioreactor provides a capable tool for early stage bioprocess development. *Sci. Rep.* **11**, 2056 (2021).
171. Wang, Y., Ling, C., Chen, Y., Jiang, X. & Chen, G.-Q. Microbial engineering for easy downstream processing. *Biotechnol. Adv.* **37**, 107365 (2019).
172. Schroeder, J. W., Yeesin, P., Simmons, L. A. & Wang, J. D. Sources of spontaneous mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* **53**, 29–48 (2018).
173. Wesseler, J., Kleter, G., Meulenbroek, M. & Purnhagen, K. P. EU regulation of genetically modified microorganisms in light of new policy developments: possible implications for EU bioeconomy investments. *Appl. Econ. Perspect. Policy* **45**, 839–859 (2022).
174. Lensch, A. et al. Safety aspects of microorganisms deliberately released into the environment. *EFB Bioecon. J.* **4**, 100061 (2024).
175. Shams, A., Fischer, A., Bodnar, A. & Kliegman, M. Perspectives on genetically engineered microorganisms and their regulation in the United States. *ACS Synth. Biol.* **13**, 1412–1423 (2024).
176. Yan, Q. & Fong, S. S. Study of in vitro transcriptional binding effects and noise using constitutive promoters combined with UP element sequences in *Escherichia coli*. *J. Biol. Eng.* **11**, 33 (2017).
177. Zobel, S. et al. Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. *ACS Synth. Biol.* **4**, 1341–1351 (2015).

178. Johnson, A. O., Gonzalez-Villanueva, M., Tee, K. L. & Wong, T. S. An engineered constitutive promoter set with broad activity range for *Cupriavidus necator* H16. *ACS Synth. Biol.* **7**, 1918–1928 (2018).
179. Li, T. et al. Engineering of core promoter regions enables the construction of constitutive and inducible promoters in *Halomonas* sp. *Biotechnol. J.* **11**, 219–227 (2016).
180. Yang, Y. et al. Prediction and characterization of promoters and ribosomal binding sites of *Zymomonas mobilis* in system biology era. *Biotechnol. Biofuels* **12**, 52 (2019).
181. Yu, M. & Tsang, J. S. H. Use of ribosomal promoters from *Burkholderia cenocepacia* and *Burkholderia cepacia* for improved expression of transporter protein in *Escherichia coli*. *Protein Expr. Purif.* **49**, 219–227 (2006).
182. De Paepe, B., Maertens, J., Vanholme, B. & De Mey, M. Modularization and response curve engineering of a naringenin-responsive transcriptional biosensor. *ACS Synth. Biol.* **7**, 1303–1314 (2018).
183. Worsey, M. J. & Williams, P. A. Metabolism of toluene and xylenes by *Pseudomonas* (putida (arvilla) mt-2: evidence for a new function of the TOL plasmid. *J. Bacteriol.* **124**, 7–13 (1975).
184. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
185. Wanner, B. L., Kodaira, R. & Neidhardt, F. C. Physiological regulation of a decontrolled lac operon. *J. Bacteriol.* **130**, 212–222 (1977).
186. Egan, S. M. & Schleif, R. F. A regulatory cascade in the induction of *rhaBAD*. *J. Mol. Biol.* **234**, 87–98 (1993).
187. Rosenberg, A. H. et al. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125–135 (1987).
188. Cheng, X. & Patterson, T. A. Construction and use of lambda PL promoter vectors for direct cloning and high level expression of PCR amplified DNA coding sequences. *Nucleic Acids Res.* **20**, 4591–4598 (1992).
189. Mead, D. A., Skorupa, E. S. & Kemper, B. Single stranded DNA SP6 promoter plasmids for engineering mutant RNAs and proteins: synthesis of a 'stretched' preproparathyroid hormone. *Nucleic Acids Res.* **13**, 1103–1118 (1985).
190. Delaleau, M. et al. Rho-dependent transcriptional switches regulate the bacterial response to cold shock. *Mol. Cell* **84**, 3482–3496.e7 (2024).
191. Silva-Rocha, R. et al. The standard European vector architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, D666–D675 (2013).
192. 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).
193. Mutalik, V. K. et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
194. Ohuchi, S., Mori, Y. & Nakamura, Y. Evolution of an inhibitory RNA aptamer against T7 RNA polymerase. *FEBS Open. Bio* **2**, 203–207 (2012).
195. Mori, Y., Nakamura, Y. & Ohuchi, S. Inhibitory RNA aptamer against SP6 RNA polymerase. *Biochem. Biophys. Res. Commun.* **420**, 440–443 (2012).
196. Kolpashchikov, D. M. Binary malachite green aptamer for fluorescent detection of nucleic acids. *J. Am. Chem. Soc.* **127**, 12442–12443 (2005).
197. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).
198. Sirois, S. & Szatmari, G. Detection of XerC and XerD recombinases in gram-negative bacteria of the family *Enterobacteriaceae*. *J. Bacteriol.* **177**, 4183–4186 (1995).
199. Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M. & Smith, M. C. M. New applications for phage integrases. *J. Mol. Biol.* **426**, 2703–2716 (2014).
200. Walker, S. A. & Klaenhammer, T. R. Molecular characterization of a phage-inducible middle promoter and its transcriptional activator from the *Lactococcal Bacteriophage* φ31. *J. Bacteriol.* **180**, 921–931 (1998).
201. Engstrom, M. D. & Pfleger, B. F. Transcription control engineering and applications in synthetic biology. *Synth. Syst. Biotechnol.* **2**, 176–191 (2017).
202. De Paepe, B. & De Mey, M. Biological switches: past and future milestones of transcription factor-based biosensors. *ACS Synth. Biol.* **14**, 72–86 (2025).
203. Pu, W. et al. Directed evolution of linker helix as an efficient strategy for engineering LysR-type transcriptional regulators as whole-cell biosensors. *Biosens. Bioelectron.* **222**, 115004 (2023).
204. De Wannemaeker, L. et al. Standardization of fluorescent reporter assays in synthetic biology across the visible light spectrum. *ACS Synth. Biol.* **12**, 3591–3607 (2023).
205. Lorenz, W. W., McCann, R. O., Longiaru, M. & Cormier, M. J. Isolation and expression of a cDNA encoding Renilla reniformis luciferase. *Proc. Natl Acad. Sci. USA.* **88**, 4438–4442 (1991).
206. Inouye, S. Firefly luciferase: an adenylate-forming enzyme for multicatalytic functions. *Cell. Mol. Life Sci.* **67**, 387–404 (2009).
207. Yeh, A. H.-W. et al. De novo design of luciferases using deep learning. *Nature* **614**, 774–780 (2023).
208. Ahmed, F. H. et al. Over the rainbow: structural characterization of the chromoproteins gfasPurple, amilCP, spisPink and eforRed. *Acta Crystallogr. Sect. Struct. Biol.* **78**, 599–612 (2022).
209. Martínez-García, E. et al. SEVA 4.0: an update of the standard European vector architecture database for advanced analysis and programming of bacterial phenotypes. *Nucleic Acids Res.* **51**, D1558–D1567 (2023).
210. Chen, W. et al. CRISPR/Cas9-based genome editing in *Pseudomonas aeruginosa* and cytidine deaminase-mediated base editing in *Pseudomonas* species. *iScience* **6**, 222–231 (2018).
211. Cong, L. & Zhang, F. in *Chromosomal Mutagenesis* (ed. Pruett-Miller, S. M.) 197–217 (Springer, 2015).
212. Volke, D. C., Friis, L., Wirth, N. T., Turlin, J. & Nikel, P. I. Synthetic control of plasmid replication enables target- and self-curing of vectors and expedites genome engineering of *Pseudomonas putida*. *Metab. Eng. Commun.* **10**, e00126 (2020).
213. Li, R., Li, A., Zhang, Y. & Fu, J. The emerging role of recombineering in microbiology. *Eng. Microbiol.* **3**, 100097 (2023).
214. Green, B., Bouchier, C., Fairhead, C., Craig, N. L. & Cormack, B. P. Insertion site preference of Mu, Tn5, and Tn7 transposons. *Mob. DNA* **3**, 3 (2012).
215. Chang, C.-W., Truong, V. A., Pham, N. N. & Hu, Y.-C. RNA-guided genome engineering: paradigm shift towards transposons. *Trends Biotechnol.* **42**, 970–985 (2024).
216. Wiechert, J., Gätgens, C., Wirtz, A. & Frunzke, J. Inducible expression systems based on xenogeneic silencing and counter-silencing and design of a metabolic toggle switch. *ACS Synth. Biol.* **9**, 2023–2038 (2020).
217. Saberi, F., Kamali, M., Najafi, A., Yazdanparast, A. & Moghaddam, M. M. Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. *Cell. Mol. Biol. Lett.* **21**, 6 (2016).
218. Carpine, R., Olivieri, G., Hellingwerf, K. J., Pollio, A. & Marzocchella, A. Industrial production of poly-?-hydroxybutyrate from CO_2: can *Cyanobacteria* meet this challenge? *Processes* **8**, 323 (2020).
219. Rajpurohit, H., Sharma, P., Sharma, S. & Bhandari, A. Polymers for colon targeted drug delivery. *Indian. J. Pharm. Sci.* **72**, 689–696 (2010).
220. Venegas, F. A. et al. The bacterial product violacein exerts an immunostimulatory effect via TLR8. *Sci. Rep.* **9**, 13661 (2019).
221. Zaykov, A. N., Mayer, J. P. & DiMarchi, R. D. Pursuit of a perfect insulin. *Nat. Rev. Drug. Discov.* **15**, 425–439 (2016).
222. Knox, C. et al. DrugBank 6.0: the DrugBank knowledgebase for 2024. *Nucleic Acids Res.* **52**, D1265–D1275 (2023).
223. Precedence Research. Synthetic biology market size to hit USD 148.93 billion by 2033 https://www.precedenceresearch.com/synthetic-biology-market (2025).
224. Switzer, A., Šusta, L. & Mines, P. Routes to industrial scalability to maximise investment in engineering biology in the UK — a bioplastics small and medium-sized enterprise perspective. *Eng. Biol.* **8**, 41–46 (2024).
225. Mavrommati, M., Daskalaki, A., Papanikolaou, S. & Aggelis, G. Adaptive laboratory evolution principles and applications in industrial biotechnology. *Biotechnol. Adv.* **54**, 107795 (2022).
226. Asin-Garcia, E., Fawcett, J. D., Batianis, C. & Martins Dos Santos, V. A. P. A snapshot of biomanufacturing and the need for enabling research infrastructure. *Trends Biotechnol.* **43**, 1000–1014 (2024).

Acknowledgements

J.P. holds a predoctoral SB scholarship from FWO (grant no. 1S18723N). M.B., M.D.M. and R.L. are supported by a grant from the Special Research Fund (grant no. iBOF/21/092). J.M. and R.L. are supported by a grant from FWO (grant no. G044624N).

Author contributions

J.P. researched data for the article and contributed to discussion of content and writing. M.B., M.D.M., J.M. and R.L. contributed to discussion and reviewing/editing the manuscript before submission.

Citation diversity statement

We acknowledge that research authored by scholars from historically excluded groups is systematically under-cited. Every attempt has been made to reference relevant research in a manner that is equitable in terms of racial, ethnic, gender and geographical representation.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Bioengineering* thanks Shihui Yang, Joshua Yuan and the other, anonymous, reviewers(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.