Exploring the synthetic biology potential of bacteriophages for engineering non-model bacteria

Eveline-Marie Lammens 1, Pablo Ivan Nikel 2 & Rob Lavigne 1

Non-model bacteria like Pseudomonas putida, Lactococcus lactis and other species have unique and versatile metabolisms, offering unique opportunities for Synthetic Biology (SynBio). However, key genome editing and recombineering tools require optimization and large-scale multiplexing to unlock the full SynBio potential of these bacteria. In addition, the limited availability of a set of characterized, species-specific biological parts hampers the construction of reliable genetic circuitry. Mining of currently available, diverse bacteriophages could complete the SynBio toolbox, as they constitute an unexplored treasure trove for fully adapted metabolic modulators and orthogonally-functioning parts, driven by the longstanding co-evolution between phage and host.

For decades, Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae and, to a lesser extent, Mycoplasma, have been the key reference models for synthetic biology (SynBio) and the go-to biological chassis for the implementation of genetic circuits. However, advances in metabolic and genome engineering enabled the exploration of non-traditional bacterial species as chassis, including Pseudomonas putida and Lactococcus lactis, for applications that extend far beyond the traditional biotechnology and food industry.

In spite of the progress, these industrially relevant organisms remain underutilized, due to the relatively limited toolboxes available for the construction of synthetic genetic circuits. The general practice of utilizing genome-editing enzymes and expression systems from the extensive E. coli toolbox in other hosts often leads to unreliable results, pointing to the general problem of portability in SynBio. To fully unlock the potential of non-model bacteria for SynBio, a set of standardized and reliable biological parts is urgently required, tailored to the specific requirements needed to establish genetic circuits in these organisms.

These parts could be mined from bacterial viruses (bacteriophages), as phage-based parts and enzymes are, by definition, fully adapted to the non-model host due to the longstanding co-evolution between phage and bacteria. Indeed, the study of phage-host interactions and phage biology has been a vast source of inspiration for SynBio tools, including CRISPR-Cas9-based technologies, recombineering and numerous DNA-editing enzymes, but has mainly focused on mining coliphages. In this Perspective, we discuss the potential of bacteriophages to address the lack of reliable, portable tools to bridge critical knowledge and application gaps, towards fully unlocking the potential of non-traditional bacteria for SynBio applications.

Rising stars in the field of synthetic biology
Several bacterial species have caught the attention of the SynBio community as potential biological chassis. This Perspective will focus on industrially relevant members of the Pseudomonas,
Klebsiella, Bacillus, Lactococcus, and Mycobacterium genera, for which a significant number of phage genomes are available (Table 1). Moreover, these species cover different bacterial lifestyles and a wide range of SynBio applications. For instance, the Gram-negative root-colonizer *P. putida* is endowed with a highly versatile metabolism, a significant tolerance for endogenous and exogenous stresses and the capability to produce toxic, value-added compounds, including cis,cis-muconic acid, cinnamate, and p-coumarate. *B. subtilis* and *L. lactis* are excellent producers and in vivo delivery vehicles for therapeutic products, including oral vaccines. The non-pathogenic *K. pneumoniae* is currently available for several non-model species, which is currently available for several non-model species, including *P. putida* and *B. subtilis*, allowing engineered processing of abundant waste streams, such as lignocellulosic biomass and molasses.

Apart from *P. putida*, other Gram-negative species are on their way to become valuable biological chassis. For example, the human pathogen *Klebsiella pneumoniae* has been recognized as an outstanding glycerol fermenter for value-added, reduced compounds (e.g. diols) and research towards a safe, non-pathogenic *K. pneumoniae* chassis is currently ongoing. However, not only Gram-negative bacteria are popular hosts for SynBio. Gram-positive organisms, such as *B. subtilis* and *L. lactis*, are excellent producers and in vivo delivery vehicles for therapeutic products, including oral vaccines. This is due to their lack of immunogenic lipopolysaccharides and their ‘Generally Regarded As Safe’ (GRAS) status conferred by the FDA. In addition, *B. subtilis* shows good growth and robustness in industrial fermentations and possesses a superior capacity of protein secretion in comparison to other hosts, whereas *L. lactis* has become indispensable in the dairy industry as a fermenter and producer of flavouring and texturizing compounds. Furthermore, other lesser-known Gram-positive bacteria show interesting features for SynBio applications. For example, *Mycobacterium smegmatis*, which is the non-pathogenic model organism for tuberculosis research, meets all the requirements to become a valuable production host for sterols and mycobacterial proteins, which often form inclusion bodies when expressed in *E. coli*.

**Table 1** Overview of non-model bacteria with high potential for SynBio applications.

| Bacterial species               | Gram-negative | Gram-positive |          |
|---------------------------------|---------------|---------------|----------|
| **Main feature**                | *Pseudomonas putida* | *Klebsiella pneumoniae* | *Bacillus subtilis* | *Lactococcus lactis* | *Mycobacterium smegmatis* |
| **Lifestyle**                   | High tolerance towards toxic compounds and stress Aerobic | Glycerol fermentation to value-added compounds Facultative anaerobic | Efficient protein secretion Facultative anaerobic | Production of flavour- and texturizing compounds Facultative anaerobic | Aerobic |
| **Biosafety**                   | L1-classified | L2-classified | Yes | Yes | Yes | L1-classified |
| **Number of identified Caudovirales phages for this genus** | 259 | 71 | 124 | 133 | 178 |

**Missing devices in the engineering toolbox of non-model bacteria**

Tools for large-scale genome engineering are still in their infancy. In the past two decades, the genome engineering scene has been revolutionized by ground-breaking advances in DNA-sequencing and RNA-sequencing, CRISPR-Cas9-based genome editing and innovative recombineering techniques. Metabolic and genome engineering of non-model organisms has come a long way and reliable tools for creating genetic deletions, insertions, and replacements are available. However, large-scale reprogramming of cells is still a tedious and time-consuming undertaking, due to cycles of plasmid construction, multiple cloning steps, and inefficient plasmid curing. These problems were largely overcome in *E. coli* and related Enterobacterial species like *K. pneumoniae*, with the development of DNA recombineering and their derived, large-scale technologies including MAGE, DivERGE, and TRMR (Box 1). While these technologies have been perfected for large-scale genome engineering of reference hosts, the upsampling of genome engineering for non-traditional microbes is trailing behind and often accompanied with unsatisfactory efficiencies. This can be explained by inadequate efficiencies of transformation methods and (counter)selection markers for non-model hosts. Moreover, the streamlined genomes of reference organisms and some key strains of *P. putida*, *B. subtilis*, and *L. lactis* allow for easy and efficient manipulation, whereas the manipulation of alternative hosts and wild strains is hampered by native endonucleases and other interfering genomic loci. Furthermore, the specific physiology of non-enteric hosts prevents the direct transfer of *E. coli*-optimized technologies, thus requiring (often tedious and time-consuming) species-specific optimization. As an example, different strategies to boost recombineering efficiencies of MAGE in non-model hosts are depicted in Fig. 1. Even with large efforts to implement and improve recombineering-based strategies in non-model hosts, these technologies are still in their infancy and further optimization is vital to allow these valuable bacterial strains to reach their full potential in the SynBio field.

**Perfecting basic tools: learning how to walk before running.**

The principal requirement for efficient genome engineering of any host is a high-performance DNA transformation method, which is currently available for several non-model species, including *P. putida*, *L. lactis*, and *B. subtilis*. In contrast, transformation of *K. pneumoniae* and *M. smegmatis* is much more difficult, as these two strains are endowed with an outer polysaccharide capsule and a thick waxy layer of mycolic acid in the cell wall, respectively. This results in significantly reduced
Box 1 | ssDNA recombineering and derived, multiplexed technologies

ssDNA recombineering is a rapid and efficient method to introduce substitutions and relatively small insertions and deletions in a specific genomic locus15, ssDNA molecules carrying the desired mutation are electroporated into the host, after which the Beta protein of phage λ binds and anneals the ssDNA to a complementary single-strand near the replication fork. Contrary to traditional recombination methods, this method does not require constructing plasmids carrying genomic homology tails, thus shortening the genome engineering time. Moreover, this technique allows parallel and/or sequential integration of mutations by using ssDNA cocktails or multiple electroporation steps, respectively. This so-called multiplexing has led to the development of several techniques including, but not limited to, the following:

- **MAGE** or multiplexed automated genome engineering allows large-scale programming and evolution of cells by submitting the cells to several cycles of ssDNA recombineering with ssDNA cocktails carrying randomized mutations161. By using selective media for cell recovery, accelerated continuous evolution of the cell population can be carried out in a few days.
- **CRISPR-optimized MAGE (CRMAGE)** is a combination of CRISPR-Cas9 and MAGE to increase recombineering efficiencies162. The implementation of CRISPR-Cas9 allows to counterselect for wild-type sequences using targeted sgRNAs.
- **Trackable multiplex recombineering (TRMR)** is a recombineering-based method to alter gene expression levels at a genome-wide scale by introducing a strong promoter and RBS (gene overexpression) or RBS-less cassette (gene knock-down) at multiple loci163. The unique barcode integrated in the cassette allows to trace back the genomic alteration causing a specific host phenotype.
- **DivERGE** or directed evolution with random genomic mutations is very similar to MAGE, but employs completely random ssDNA oligonucleotides to accelerate directed evolution164. This method is specifically designed to generate resistant mutants with no prior knowledge of possible resistance mechanisms.
- **High-efficiency multi-site genomic editing (HEMSE)** is a multiplexed recombineering method optimized for *P. putida*21. Due to lower recombineering efficiencies compared to MAGE for *E. coli*, the automation aspect is not (yet) possible, hence the use of the novel HEMSE terminology.
- **ORBIT** or oligonucleotide-mediated recombineering followed by Bxb1 integrase targeting is a method for efficient genome engineering of larger genomic regions of mycobacteria in a selectable manner22. A Beta protein homologue allows for genomic insertion of an *attP* region. This *attP* site then acts as a landing path for a selectable cassette using the BxbA integrase to create genomic mutations, tagged genes, promoter replacements, and large insertions.

transformation efficiencies, in the range of only 10^4–10^5 transformants/µg DNA19,20. Since poor transformation performances will sabotage the efficacy of almost all genome-editing technologies, further optimization of transformation methods is absolutely essential. Moreover, most transformation methods have been optimized for effective uptake of plasmid DNA, whereas novel ssDNA and dsDNA recombineering techniques require efficient introduction of linear DNA. For instance, it has been shown that poor ssDNA uptake hampers the efficiency and multiplexing of recombineering-based techniques like MAGE and DivERGE in *P. putida*21, *L. lactis*22, and *M. smegmatis*23 (Box 1). A valid alternative for bacterial electroporation is transduction, in which phage particles take up DNA from a donor cell and deliver it to a recipient via phage infection. Transduction has been shown to be a superior method to achieve DNA transfer in some species and wild strains, including *Mycobacterium*, for which it can mediate an almost 100% transduction efficiency14. However, despite this excellent efficiency, transduction is more time-consuming compared to electroporation as it requires the generation of transducing phages14.

Besides efficient transformation methods, markers for selection and counter-selection are indispensable in genome engineering. Antibiotic-based markers are the state-of-the-art for transformant selection and are readily available for almost all bacterial hosts. This contrasts the lack of high-performance counter-selection markers for some non-model hosts, including *P. putida* and *M. smegmatis*. For instance, the well-known levansucrase-encoding sacB gene from *B. subtilis* efficiently inhibits cell growth in *E. coli* and related species like *K. pneumoniae* in the presence of sucrose24. However, its use in other hosts sometimes requires very high sucrose concentrations (>10% w/v), thus exerting high osmotic stress and inducing spontaneous mutations in *sacB*25,26.

To overcome this problem, novel counter-selection markers have been developed for several hosts inspired by the *URA3* and *upp* selection systems from yeast and *B. subtilis*, respectively. Although these markers show improved performance over *sacB* in *P. putida*, their utilization is strictly dependent on deletion mutants, thus limiting their application range25,27. To circumvent this issue, a plasmid curing system for *P. putida* based on inducible I-SceI-based plasmid restriction and conditional plasmid replication has been developed and has demonstrated to be widely applicable, shortening downtime in genome engineering28. This system, however, is more complex than traditional counter-selection markers.

**Urgent need for standards, norms, and regulations in synthetic biology.** Apart from advances in genome and metabolic engineering techniques, the positive impact of standardization on the SynBio field should not be underestimated, as it boosts interoperability, reproducibility of results, and reduces redundant work29. Although the level of standardization within SynBio is still far below those of the hard sciences, large efforts led by the National Institute of Standards and Technology, the iGEM foundation and several other organizations are putting us on the right track for the future by the creation of biological part databases, software for circuit design and standards for data representation and communication29. However, standardized biological parts and constructs, transparent data sharing and standard operating procedures (SOPs) are largely missing for non-model species. This makes the construction of reliable genetic circuits in non-traditional microbes a process of mere trial-and-error.

The design of genetic circuitry generally starts with the selection of a suitable backbone for the construct. The development of the standard European vector architecture (SEVA) for *Pseudomonas* and other Gram-negative species constituted a pivotal point for the uniformity of vectors and standardization of biotechnological research30. The SEVA 3.0 database offers a wide array of replicative and integrative ‘SEVA plasmids’ following a uniform design and nomenclature, some of them endowed with expression systems that are commonly used by the *Pseudomonas* community31. Besides the SEVA repository, several databases are available with biological parts for both Gram-negative and Gram-positive organisms29, but unfortunately, dedicated parts and
repositories for Gram-positive organisms remain scarce. In recent years, progress in this field has been made by developments including the iGEM Registry Probiotic Collection and the Bacillus BioBrick Box 2.0. Since their launching, hundreds of SEVA- and Bacillus BioBrick vectors have been requested and the BioBrick vectors have been received thousands of views, indicating the success and the need for dedicated databases for non-model hosts.

The vast majority of parts have been developed for E. coli and are barely validated nor characterized in other hosts. Moreover, even for this reference strain species, standardization of parts is still an illusion, as no clear consensus exists on characterization protocols for parts, let alone on standard units for the quantification of transcriptional and translational activity. Therefore, biological parts cannot be reliably compared and as a consequence, researchers still tend to construct genetic circuits with parts they are familiar with rather than those possessing the optimal features to reach their goal. The absence of standard units and SOPs has not gone unnoticed and several attempts have been made to implement these in the world of SynBio for processes like fluorescent-activated cell sorting, GFP quantification, and measurements of gene expression levels. Unfortunately, they have not (yet) been generally adopted by the SynBio community.

In contrast to parts standardization, methods for standard assembly of bacterial genetic constructs are much more established. The development of the restriction enzyme-based BioBrick and BglBrick standards has laid the foundation for more sophisticated assembly methods based on the Type IIs restriction enzymes, including Golden Gate assembly, GoldenBraid, Twin-Primer Assembly, but they lack the possibility to quickly generate full or semi-random combinatorial
libraries in multi-brick constructs, which is the principal strength of the aforementioned techniques. In summary, the first steps towards standardization have been taken in the field of SynBio with non-traditional hosts. To further unlock the SynBio potential of non-model microbes, there is an urgent need for a reliable, standardized, and well-characterized set of biological parts to build complex genetic circuitry.

**Constructing novel genetic circuits: the sum is more than its parts**

As our knowledge on metabolic engineering continues to increase, we have learned that balanced expression levels and feedback mechanisms often result in higher yields and healthier cell populations. To reach these desired expression levels, predictable, robust, and tuneable genetic constructs are a must and can only be built from well-characterized, reliable, and context-independent biological parts as demonstrated by Nielsen et al. Furthermore, more complex circuits require a set of analogous high-performance parts to avoid part reuse and unwanted recombination events between them. The available set of biological parts for non-model hosts is steadily expanding, but there is still a lack of much needed well-characterized, reliable, and orthogonal parts. A graphical summary of this section is provided in Fig. 2.

**Promoters as core devices to tune gene expression**

The choice for a specific promoter for both model and non-model hosts is very application-dependent, but in general, desired promoter features are predictability, tunability, independence, orthogonality, and composability. Constitutive promoters are the simplest subset of promoters and are the preferred choice for industrial-scale setups, due to their high reliability and a generally low context-dependency. A set of (synthetic) constitutive promoters is available for several non-model bacteria with extensive fold ranges, including *P. putida*, *B. subtilis*, *L. lactis*, *K. pneumoniae*, and *M. smegmatis*. Inducible promoters offer the user an extra level of control in terms of modulating and inducing gene expression at defined timepoints. Although many expression systems optimized for *E. coli* have been successfully transferred to other hosts, their characterization and optimization is generally limited (with a few exceptions). Due to the high context-dependent response of most expression systems, predicting protein production levels is almost impossible. Tedium optimization cycles of genetic circuit constructions are therefore required. Additionally, selecting an expression system with the optimal features for one’s goal is difficult, since each of them suffers from specific drawbacks and a centralized reporting system for expression system performance is lacking. As an example, Table 2 indicates the advantages and drawbacks of the most common expression systems used in *P. putida*. Other relevant inducible systems for *P. putida* are summarized in Supplementary Table 1 with their respective advantages and disadvantages. For other non-model hosts, the reader is referred to complementary reviews.

In contrast to the relatively large set of activators, the number of (characterized) repressors for *P. putida* is low and limits the construction of complex genetic feedback mechanisms for SynBio applications (Table 2 and Supplementary Table 1). In some applications, the lack of a suitable repressor can be substituted by CRISPR interference (CRISPRi) technology. This technology enables the simultaneous, tunable, and transient repression of multiple genes by expression of gene-specific sgRNAs and a single nuclease-deficient Cas9 (dCas9) and has been successfully applied in *P. putida*, *M. smegmatis*, *K. pneumoniae*, *B. subtilis*, and *L. lactis* with up to 98% repression. Although this technique offers great potential, one must note that its efficiency and efficacy remains dependent on (1) the presence of a PAM sequence within or near the promoter region for efficient transcriptional repression, (2) the strengths and weaknesses of the expression system used for dCas9 expression, (3) the burden imposed by accumulation of a large heterologous protein (dCas9), and (4) polar effects such as the repression of downstream operons.

Despite the many advantages of chemical inducers as stated above, they are often costly, toxic, and do not allow easily switching off induction. Alternative induction methods based on light, pH, temperature, and quorum sensing overcome these issues and are steadily gaining popularity in the SynBio field (Supplementary Table 1). Moreover, several of these systems are available and optimized for non-model hosts, though the number stays well below those for *E. coli*. Examples include the optogenetic CcaSR system, the thermosensitive cI857–P1 system, multiple quorum sensing-based expression systems, and pH-regulated systems.

**Ribosomal-binding sites: highly important, often overlooked.**

The past few years, the potential of ribosomal-binding sites (RBS) to tune expression levels has received increased attention, leading to the development of high-end RBS calculation tools and RBS libraries for some non-model hosts, including *P. putida* and *B. subtilis*. However, the performance of these RBS libraries is highly sequence-dependent, due to the formation of mRNA secondary structures at the 5’ end of the transcript, which potentially inhibit translation or cause instability of the mRNA. To overcome this problem, a library of RBS elements with a bicistronic design (BCD) was created for *E. coli*. The bicistronic layout consists of two Shine–Dalgarno (SD) sites: the first SD initiates translation of a leader peptide, which in its turn holds the second SD site in its coding sequence and allows translation of the downstream gene of interest. Translation of the standard leader peptide will thus disrupt secondary mRNA structures of the second SD and/or gene of interest and allows more reliable, sequence-independent translation levels than traditional RBS libraries. To our knowledge, this library has not been validated in one of the aforementioned hosts, but its BCD2 element has been deployed successfully in several genetic constructs for *P. putida* and the BCD concept has been tested with promising results in *B. subtilis*.

**Reliable transcriptional termination for qualitative genetic circuits.**

Proper insulation and termination of (synthetic) transcriptional units is key in genetic circuit design, as it avoids transcriptional read-through from the unit itself and interference from neighbouring promoters. A set of potent terminators, and more specifically bidirectional terminators (Fig. 2 (3)), is therefore an essential part of the SynBio toolbox, leading to the development of characterized (synthetic) terminators and terminator prediction tools for *B. subtilis* and *E. coli*. In contrast to initiation elements for which host-dependent performance has long been recognized, the set of optimized terminators for the Gram-negative and Gram-positive models *E. coli* and *B. subtilis* are often unduly considered to be generally functional in other hosts without validation. This results in poorly insulated circuits with all the undesired consequences thereof. For example, recent work demonstrated that the well-known rrrB-T1 terminator from *E. coli* does not fully repress gene expression in *P. putida*, despite its implementation in all SEVA vectors. Furthermore, several *E. coli*-based algorithms were used to mine putative rho-independent terminators from environmental DNA for *Pseudomonas* species. All predicted terminators functioned in *E. coli*, but
RNA-based and protein-based regulation are further to be explored. Besides the classical biological parts described above, RNA-based regulatory parts allow an additional level of regulation in genetic circuits and enable the construction of elegant logic gates. These RNA-based regulatory parts include riboswitches, RNA thermometers, translation riboregulators, ribozymes, antisense RNAs, and small transcription activating RNAs (STARs) with several functions. Their fast response times and minimal cell burden compared to protein-based systems have allowed RNA-based regulatory parts to become a valuable part of the E. coli SynBio toolbox, but to date, their use in non-model hosts is still rare.

Riboswitches and RNA thermometers are often used in SynBio to tune expression levels in response to their cognate ligand. For E. coli, genome-wide riboswitch identification, extensive characterization, optimization, and standardization of 5′ sequences have resulted in riboswitches with reliable functioning. In contrast, these parts are not yet exploited as regulatory tools. However, in contrast to RNA-based regulation, several examples are already available of controlled and/or inducible protein degradation via protein tags and orthogonal proteolysis machinery.

Riboswitches and RNA thermometers have only recently been characterized in vivo in non-traditional bacteria. Moreover, the utility as a glucosamine-6-phosphate responsive switch, RNA antisense RNAs and can be used for multiplexed genome-wide knock-down of gene expression, but their utilization as biological parts in non-traditional hosts is still limited. This might change in the near future as a synthetic sRNA toolbox for P. putida has recently been developed.

Another set of undervalued RNA-based parts are ribozymes, whose SynBio applications include the creation of variant logic gates and detecting systems and who have demonstrated their worth as insulators of RNA parts. Only a few rare examples report the use of ribozymes as tools in non-models hosts, where they have shown their utility as a glucosamine-6-phosphate responsive switch, RNA antisense RNAs and can be used for multiplexed genome-wide knock-down of gene expression, but their utilization as biological parts in non-traditional hosts is still limited. This might change in the near future as a synthetic sRNA toolbox for P. putida has recently been developed.

To conclude, biological parts for RNA-based and protein-based regulation are highly underexploited for non-model hosts. Characterization and development of these parts (within a standardized framework) will expand the SynBio toolboxes of non-model bacteria with tools for gene expression manipulation at multiple levels.

**Phages are unexplored troves for synthetic biology tools and parts**

From the inception of molecular biology, bacteriophages have proven to be indispensable tools to conduct biomolecular...
### Table 2 Overview and comparison of relevant induction systems often used in *P. putida*.

| TF | Promoter(s) | Inducer(s) | Reported pros, cons and main applications in *P. putida* | Ref.(s) |
|----|-------------|-------------|--------------------------------------------------------|--------|
| Positively regulated induction systems | **XylS** | **Pin** | m-tolu ate and derivatives | + High expression levels  
- Catabolite repression  
- Dose-dependent response in absence of inducer  
- Expensive inducers  
- Metabolization pathways  
- Bimodal response  
- Leaky expression | 28, 49, 63, 130-139 |
|  |  |  |  |  | |
|  | **RhaRS** | **P_{raB}** | L-rhamnose | + No catabolite repression  
- No metabolization of rhamnose: dose-dependent response  
+ Tight regulation  
+ Non-toxic inducer  | Applications: Expression of I-sceI, dCas9, trfA, the λ red operon, and recombinases in genome engineering tools; expression of toxic genes for controlled autolysis; production of recombinant antibody fragments and p-coumaric acid | 132, 138, 140, 141 |
|  |  |  |  |  | |
|  | **AraC** | **P_{BAD}** | L-arabinose | + Characterization  
+ No catabolite repression  
+ No metabolization of arabinose: dose-dependent response  
+ Tight regulation  
+ Non-toxic inducer  | Applications: Expression of dCas9 for CRISPRi and Cre for genomic deletions; production of p-coumaric acid | 132, 142, 143 |
| Negatively regulated induction systems | **LacI** | **Plac, PlacUVS, Plac** | Lac tose, isopropyl β-D-1-thiogalactopyranoside (IPTG) | + Extensively characterized  
- Toxicity of IPTG  
- Leaky expression  
- Expensive inducers  | Applications: Expression of Bxb1 integrase for genomic integration; biodesulfurization; production of p-hydroxybenzoate, ethanol, and recombinant proteins | 132, 143-149 |
|  |  | **Ptrc** |  |  | |
|  | **TetR** | **P_{tet}** | Anhydrotetracycline (aTc) | + Tight regulation  
- Unreliable functioning (probably due to unbalanced TetR levels)  
- Toxicity of aTc  | Application: Tubulysin production | 143, 150-152 |

**Note:** TF transcription factor.
Box 2 | The diversity of RNA-based regulatory parts

RNA-based regulatory parts allow an additional level of gene expression regulation complementary to the classical biological parts such as promotors and RBSs. In general, RNA-based parts do not depend on additional proteins or translation. As such, they allow faster response times and reduced cell burden compared to protein-based systems. Moreover, since these parts mostly function in cis, the risk of off-target effects is limited and regulator concentrations do not need to be balanced. The RNA-based regulatory parts can be divided in distinct categories (Fig. 2 (4)):

- **Riboswitches** are cis-acting secondary RNA structures in untranslated mRNA, which control gene expression by obscuring the access to regulatory sequences like the RBS or by forming (anti)terminators in response to ligand binding.
- Similar to riboswitches are RNA **thermometers**. As the name implies, these riboswitch-like RNA structures alter their secondary structure as a function of temperature rather than ligand binding.
- **Translation riboregulators** and **antisense RNAs** are functionally equal to riboswitches, but are not part of the main mRNA transcript of a synthetic genetic construct. They regulate the construct’s expression in cis (encoded on the opposite strand of its target) or in trans (encoded in a different genomic location). Most commonly, these molecules regulate translation initiation by binding to 5' UTR regions to block the RBS or to induce rapid mRNA degradation.
- Apart from translational regulation, sRNA-based regulation of non-translation processes is also possible, as shown by the recently developed small **transcription activating RNAs (STARs)**. These small trans-acting RNAs induce transcription by disrupting an intrinsic transcription terminator upstream of the coding gene.
- A very unique set of catalytic RNA-based parts are called **ribozymes**, enzymatic RNA molecules able to catalyse biochemical reactions such as RNA cleavage and ligation. In some cases, the activity of these ribozymes is dependent on the binding of a specific ligand, which further broadens the application range of these RNA enzymes.

Efficient genome engineering using host-specific phage enzymes. Phage genomes encode many toxins to halt cellular growth, which have been successfully applied as treatments against bacterial infections including the engineered endolysin-based “Artilysin.” Besides their obvious applications as antimicrobials, phage toxins can also be used in SynBio as (inducible) counter-selection markers for non-traditional hosts (Fig. 3 (1)). Notably, several of these toxic proteins are host-specific, such as gp8 and gp18 from *Pseudomonas* phage LUZ7, a potentially interesting feature for manipulation of expression vectors. For instance, an expression vector which constitutively expresses a host-specific toxin, can be replicated easily in *E. coli*. After substituting the toxin for a gene of interest, the non-traditional expression host can be directly transformed with the recombinant vector, while the toxin counterselects against transformants carrying the unmodified vector. The toxin thus allows to skip intermediate cloning steps to *E. coli* saving valuable time and effort. Apart from toxins, some phages encode antitoxins or toxin-antitoxin modules to bypass host-encoded abortive infection mechanisms. These toxin–antitoxin pairs could be used to create plasmid-addiction systems or stabilize genomic inserts in situations where the use of antibiotics is undesirable or impossible, such as large bioreactors or human health applications (Fig. 3 (2)). The most well-known example of a phage-based addiction system is the *phd-doc* toxin–antitoxin module of temperate coliphage P1, but toxin–antitoxin modules have also been identified in prophages in other bacteria, including *txpA/ratA, bsrG/sr4*, and *yonT/as-yonT* in *B. subtilis* prophage elements.

The discovery of integrases and recombinases from temperate phages have revolutionized the SynBio field and resulted in a myriad of genome engineering tools. At first, these enzymes were primarily used for genomic integration of synthetic circuitry or to rearrange DNA segments. Soon after, their application range was expanded to include combinatorial and reversible DNA assembly methods, various logic gates, analogue-to-digital converters, memory devices, and multiplexed DNA editing via recombinering (Fig. 3 (4 and 5)). To date, the use of these versatile enzymes in non-model hosts has mostly remained limited to genomic integration tools and recombinering. The characterization of existing and novel integrases and recombinases from prophages will optimize the currently available recombinase-based tools for these bacteria. Additionally, they will enable the implementation of more advanced applications, including the construction of complex logic gates and memory devices, thus allowing non-model hosts to reach their full potential as valuable biological chassis for SynBio applications.

**Autographiviridae** are a goldmine for orthogonal transcription. A specific family of phages, the *Autographiviridae*, encode their own transcriptional machinery for viral replication. The archetype of this class is the well-known T7 phage, which has a small, single-subunit RNA polymerase (RNAP) and strong T7 promotors, RBSs and terminators, which have made their mark on the SynBio field. Apart from the broad-use, fully orthogonal PET expression system for *E. coli*, several other T7-based applications emerged, such as AND-gates based on fragmented T7 RNAPs and a resource allocator. So far, the use of the pET system in other hosts has been restricted by severe cytotoxicity of the T7 RNAP. We argue that T7-like phages, like gh-1 and KPO1K2 infecting *P. putida* and *K. pneumoniae*, respectively, should be explored as a source of similar orthogonal RNAPs and biological parts specifically tailored to their Gram-negative bacterial host, with reduced toxicity (Fig. 3 (6 and 7)). Beyond this family, the N4-like bacteriophages encode multiple phage RNAPs...
and corresponding RNAP-specific promoters, allowing a gradual shift from early to middle and late gene expression\(^9\). The timed aspect of this transcriptional scheme shows high potential for sequential gene expression in synthetic constructs.

In contrast to the *Autographiviridae*, other phage families are completely dependent on the host transcription machinery and use different strategies to regulate transcriptional expression. For one, *Pseudomonas* phage 14-1 and *Bacillus* phage SP01 encode sigma factor-like proteins gp12 and gp28, respectively, to alter the RNAP's specificity towards phage-specific promoters\(^96,97\). These proteins, in combination with their native phage promoters, could be used to favour transcription of heterologous genes and at the same time avoid cellular stress due to overexpression by balancing the total transcriptional load, which is a common concern with the PET system (Fig. 3 (8))\(^91\). Secondly, temperate phages rely on transcriptional repressors for their lysogenic switch, which can be exploited for SynBio purposes like the cl and Cro repressors from coliphage λ (Fig. 3 (9))\(^98\). Besides repressors, also transcriptional activators have been identified in phage genomes, including ORF2 of *Lactococcus* phage phiKZ\(^99\). Thirdly, apart from transcriptional initiation, other steps in the protein production process are also prone to viral interference, which could inspire novel SynBio tools or interesting approaches to increase bioprocess yields. For example, phage mRNA translation is influenced by secondary mRNA structures, self-splicing introns, frameshifting and the remarkable translational bypassing phenomenon\(^100\), which could lead to novel translation regulatory tools. Programmed translational frameshifting is the process where a repeat of nucleotides allows the ribosome to slip and continue translation in a different reading frame, as observed in many dsDNA phages including *Lactococcus* phage Q54\(^101\). These viral frameshifting events could inspire synthetic programmed frameshift to generate signal splitters in genetic circuitry: e.g. the expression of one coding sequence results in two enzymes with different functioning\(^102\). The translational bypassing phenomenon has so far only been observed in coliphage T4, where the mRNA of gp60 contains a stretch of 50 non-coding nucleotides, but is still translated into a single polypeptide. Instead of dissociating from the gp60 mRNA when reaching the non-coding gap, the ribosome “glides” over this non-coding region and continues translation at the second coding region\(^103\). Further research on this phenomenon may indicate if engineered translational bypassing can be used in controlled expression of designer proteins\(^103\). Furthermore, RNA turnover is altered during phage infection by stabilizing viral mRNA via asRNA-masking of RNase E degradation sites or inhibition of the RNA degradosome, e.g. by Dip of *Pseudomonas* phage phiKZ\(^104,105\). These could potentially become interesting approaches to elevate mRNA abundance of a desired transcript and increase protein yield (Fig. 3 (10))\(^106\). Lastly, *Pseudomonas* phages EL and OBP and *Bacillus* phage AR9 promote protein folding with efficient phage chaperonins which do not require any additional co-factors\(^107–109\). These chaperonins can be co-expressed in industrial bioreactors to elevate soluble and functional levels of recombinant protein, a proven method for *E. coli*-based recombinant protein production using more complex bacterial chaperonins (Fig. 3 (11))\(^110\).

![Fig. 3 Overview of SynBio applications of phage proteins.](image-url)

Phage proteins show high potential for use in genome engineering (left), manipulation of transcription and translation processes (middle) and cellular engineering (right). Besides phage proteins, phage genomes are a large source of orthogonally functioning biological parts for synthetic genetic circuitry (middle). The RNA, DNA, and bacteriophage icons were adapted from ‘waves Icon #459777’, ‘DNA Icon #3383522’, and ‘bacteriophage Icon #467715′ by Dmitry Kovalev, Muhammad Tajudin, and Dong Ik Seo, respectively, from the nounproject.com.
Beyond traditional engineering: phage-based cellular regulation. During phage infection, phages do not only alter transcription and translation-related processes as described above. They remodel the entire cell metabolism, core processes, and cell physiology to create an optimal environment for phage replication. To do so, they have acquired high-end cell modulators, optimized to their host’s specific cell physiology. These modulators constitute a goldmine of revolutionary, strain-optimized tools to expand the SynBio toolbox, as we illustrate here using selected examples. First, when entering the cell, the phage protects itself from several host defence mechanisms by using anti-CRISPR proteins (Acr) or by integrating non-canonical nucleotides into their genome to avoid host restriction enzymes. Since the discovery of Acr proteins in Pseudomonas phages, many other Acr proteins have been identified and used in a range of SynBio applications, including regulation of CRISPR-Cas-based tools at multiple stages, reduction of CRISPR-Cas toxicity and the development of Acr-based CRISPRi and selection markers.

To reach this goal, perfecting current genome engineering technologies, creating biological standards and optimizing and expanding the set of biological parts, tailored to each host, are absolutely essential. Bacteriophages are not only a vast source of potential host-specific biological parts and genome engineering tools, their genomes also encode many metabolic modulators and other interesting enzymes for directed cell manipulation, which can extend and revolutionize the SynBio toolbox. These phage genomes are becoming increasingly available and can be readily exploited for Pseudomonas, Klebsiella, Bacillus, Lactococcus and Mycobacterium genera (Table 1). In contrast, for other industrially relevant species (including the ultrafast-growing and upcoming host Vibrio natriegens), the number of identified phages is still extremely low, thus limiting the SynBio potential of phages for these species for the time being.

To date, protein expression levels are mostly manipulated on the transcriptional level. However, recent research has revealed the importance of the amino acid composition of N-terminal and C-terminal protein regions, post-translation modifications (PTMs) and cellular spatial organization on protein expression, stability, and enzymatic activity. These novel approaches to tune expression levels could soon constitute a new generation of established SynBio tools for reference bacteria. Interestingly, phages employ these strategies naturally to regulate expression levels and manipulate the cell, and could be exploited in phage-based SynBio tools for multilevel cellular engineering in the future. In contrast to regulation of transcription and translation, PTMs allow a very quick adaptation of protein function, localization, or stability to a changing environment. In spite of this major advantage, the manipulation of PTMs for SynBio purposes remains limited due to the difficulty to alter specific PTMs in vivo. The use of phage kinases and acetylases could facilitate PTM engineering in the future, as such enzymes have been discovered to specifically manipulate host proteins of metabolic pathways and host defence mechanisms.

After successful cell entry, the phage hijacks the cell’s metabolism to favour viral replication. Examples of altered metabolic pathways include the phosphate and nitrogen metabolism, photosynthesis, the pentose phosphate pathway, and nucleic acid synthesis. For example, Pseudomonas phage phiKZ promotes the de novo synthesis of pyrimidine nucleotides by reprogramming the pyrimidine pathway with seven predicted modulators. These metabolic modulators could aid to overcome bottlenecks in metabolic engineering or could indicate key targets of metabolic pathways. In a similar fashion, these inhibitory proteins could be used to uncouple cellular growth from protein production to maximize substrate-to-product conversion and product yield in bioreactors, as demonstrated successfully in E. coli. In conclusion, phage genomes are an underexploited and used in a range of SynBio applications, including regulation of CRISPR-Cas-based tools at multiple stages, reduction of CRISPR-Cas toxicity and the development of Acr-based CRISPRi and selection markers.

In recent years, several alternative bacterial hosts for SynBio applications have emerged alongside E. coli, the reference Gram-negative bacterium for such purposes. With further optimization, these hosts could be developed into an established set of biological chassis with an ideal option for every possible SynBio application. To reach this goal, perfecting current genome engineering technologies, creating biological standards and optimizing and expanding the set of biological parts, tailored to each host, are absolutely essential. Bacteriophages are not only a vast source of potential host-specific biological parts and genome engineering tools, their genomes also encode many metabolic modulators and other interesting enzymes for directed cell manipulation, which can extend and revolutionize the SynBio toolbox. These phage genomes are becoming increasingly available and can be readily exploited for Pseudomonas, Klebsiella, Bacillus, Lactococcus and Mycobacterium genera (Table 1). In contrast, for other industrially relevant species (including the ultrafast-growing and upcoming host Vibrio natriegens), the number of identified phages is still extremely low, thus limiting the SynBio potential of phages for these species for the time being.

Concluding remarks and outlook

In recent years, several alternative bacterial hosts for SynBio applications have emerged alongside E. coli, the reference Gram-negative bacterium for such purposes. With further optimization, these hosts could be developed into an established set of biological chassis with an ideal option for every possible SynBio application. To reach this goal, perfecting current genome engineering technologies, creating biological standards and optimizing and expanding the set of biological parts, tailored to each host, are absolutely essential. Bacteriophages are not only a vast source of potential host-specific biological parts and genome engineering tools, their genomes also encode many metabolic modulators and other interesting enzymes for directed cell manipulation, which can extend and revolutionize the SynBio toolbox. These phage genomes are becoming increasingly available and can be readily exploited for Pseudomonas, Klebsiella, Bacillus, Lactococcus and Mycobacterium genera (Table 1). In contrast, for other industrially relevant species (including the ultrafast-growing and upcoming host Vibrio natriegens), the number of identified phages is still extremely low, thus limiting the SynBio potential of phages for these species for the time being.

To date, protein expression levels are mostly manipulated on the transcriptional level. However, recent research has revealed the importance of the amino acid composition of N-terminal and C-terminal protein regions, post-translation modifications (PTMs) and cellular spatial organization on protein expression, stability, and enzymatic activity. These novel approaches to tune expression levels could soon constitute a new generation of established SynBio tools for reference bacteria. Interestingly, phages employ these strategies naturally to regulate expression levels and manipulate the cell, and could be exploited in phage-based SynBio tools for multilevel cellular engineering in the future. In contrast to regulation of transcription and translation, PTMs allow a very quick adaptation of protein function, localization, or stability to a changing environment. In spite of this major advantage, the manipulation of PTMs for SynBio purposes remains limited due to the difficulty to alter specific PTMs in vivo. The use of phage kinases and acetylases could facilitate PTM engineering in the future, as such enzymes have been discovered to specifically manipulate host proteins of metabolic pathways and host defence mechanisms.

Some phages build their own cytoskeleton to centrally position their genome in the host to optimize phage-particle production. To our knowledge, physical positioning of RNA as an optimization strategy for recombinant protein production is unknown territory so far, but could significantly increase protein yields in the future. Furthermore, it has been proposed that jumbo phages encapsulate their genome in a nucleus-like proteinoceous sheath upon cell entry to evade CRISPR-Cas. Similar strategies including synthetic physical encapsulation or co-localization of enzymes via scaffold molecules have already proven effective to protect molecules from degradation by host enzymes and, moreover, to increase metabolic flux and reduce cross-talk. By further studying phage encapsulation and shell proteins, this could lead to improved synthetic encapsulation and co-localization of enzymes.

On a critical note, one potential limitation of phage-based parts and tools is that they are mostly species-specific, thus limiting portability to other species. However, recent examples indicate that some phage recombinases may display broad host specificity. Rather than solving the portability issue within SynBio, we call for caution when using E. coli-optimized tools blindly in other hosts and promote the development of host-specific toolboxes instead. Furthermore, while several examples of potential novel tools have been proposed and discussed here, it should be obvious that we have only just scratched the surface of the potential bacteriophages hold for SynBio. Indeed, the majority of world’s bacteriophages remains unidentified and up to 40–90% of viral sequences cannot be aligned. The functional potential to be found in this ‘viral dark matter’ is boundless and one can only imagine the great innovations that remain to be discovered.
References

1. Calero, P. & Nikel, P. L. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. Microb. Biotechnol. 12, 98–124 (2019).

2. Lemire, S., Yehl, K. M. & Lu, T. K. Phage-based applications in synthetic biology. Annu. Rev. Virol. 5, 453–476 (2018).

3. Nadeem, S. M. et al. Potential, limitations and future prospects of Pseudomonas spp. for sustainable agriculture and environment: a review. Soil Environ. 35, 106–145 (2016).

4. Nikel, P. I. & de Lorenzo, V. Pseudomonas putida as a functional chassis for industrial biocatalysis: from native biochemistry to trans-metabolism. Metab. Eng. 50, 142–155 (2018).

5. Kumar, V. & Park, S. Potential and limitations of Mycobacterium smegmatis as a microbial cell factory utilizing glycerol as the carbon source. Biotechnol. Adv. 36, 150–167 (2018).

6. van Tilburg, A. Y., Cao, H., van der Meulen, S. B., Solopova, A. & Kuipers, O. P. Metabolic engineering and synthetic biology employing Lactococcus lactis and Bacillus subtilis cell factories. Curr. Opin. Biotechnol. 59, 1–7 (2019).

7. Liu, Y., Liu, L., Li, D., Gu & Chen, J. Synthetic biology toolbox and chassis development in Bacillus subtilis. Trends Biotechnol. 37, 548–562 (2019).

8. Galán, B. et al. Mycobacterium smegmatis is a suitable cell factory for the production of steroidic synths. Microb. Biotechnol. 10, 138–150 (2017).

9. Bashiri, G. & Baker, E. N. Production of recombinant proteins in Lactococcus lactis and Bacillus subtilis. Microb. Cell Fact. 6, 1–10 (2017).

10. Martínez-García, E., Nikel, P. I., Aparicio, T. & de Lorenzo, V. Perspective review from classical to non-traditional microorganisms.

11. Gay, P. et al. Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. J. Bacteriol. 164, 918–921 (1985).

12. Galvão, T. C. & de Lorenzo, V. Adaptation of the yeast URA3 selection system to Gram-negative bacteria and generation of a JhblcDE Pseudomonas putida strain. Appl. Environ. Microbiol. 71, 883–892 (2005).

13. Pelicic, V., Reyrat, J. M. & Gicquel, B. Expression of the Bacillus subtilis sacB gene confers sucrose sensitivity on Mycobacteria. J. Bacteriol. 178, 1197–1199 (1996).

14. Graf, N. & Altenbuchner, J. Development of a method for markerless gene deletion in Pseudomonas putida. Appl. Environ. Microbiol. 77, 5549–5552 (2011).

15. Voilke, D. C., Friis, L., Wirth, N. T., Turlin, J. & Nikel, P. L. 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).

16. Decoene, T. et al. Standardization in synthetic biology: an engineering discipline coming of age. Crit. Rev. Biotechnol. 38, 647–656 (2018).

17. Silva-Rocha, R. et al. The Standard European Vector Architecture (SEVA): a cohesive platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res. 41, 666–675 (2013). The development of the modular SEVA vectors constitutes a major step in the standardization of SynBio research in Gram-negative hosts.

18. Martínez-García, E. et al. SEVA 3.0: an update of the Standard European Vector Architecture for enabling portability of genetic constructs among diverse bacterial hosts. Nucleic Acids Res. 48, D1164–D1170 (2020).

19. Pedrolli, D. B. et al. Engineering microbial living therapeutics: the synthetic biology toolbox. Trends in biotechnology 37, 100–115 (2019).

20. Popp, P. F., Dotzlér, M., Radeck, J., Bartels, J. & Mascher, T. The Bacillus BioBrick Box 2.0: expanding the genetic toolbox for the standardized work with Bacillus subtilis. Sci. Rep. 7, 1–13 (2017).

21. Engel, K., Kandzra, R. & Mariënou, S. A one pot, one step, precise cloning method with high throughput capacity. PLoS ONE 3, e3648 (2008).

22. Vázquez-Vilar, M. et al. GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. Nucleic Acids Res. 45, 2196–2209 (2017).

23. Van Hove, B., Gaïdi, C., De Wannezaeker, L., Maertens, J. & De Mey, M. Recursive DNA assembly using protected oligonucleotide duplex assisted cloning (PODAC). ACS Synth. Biol. 6, 943–949 (2017).

24. Jin, P., Ding, W., Gu, C., Chen, J. & Kang, Z. DATESL: A Scarless and sequence-independent DNA assembly method using thermostable exonuclease and ligase. ACS Synth. Biol. 5, 1028–1032 (2016).

25. Liang, J., Liu, Z., Low, X. & Zhao, H. Twin primer non-enzymatic DNA assembly: an efficient and accurate multi-part DNA assembly method. Nucleic Acids Res. 45, e94 (2017).

26. Kent, R. & Dixon, N. Contemporary tools for regulating gene expression in bacteria. Trends Biotechnol. 38, 316–333 (2020).

27. Nielsen, A. A. K. et al. Genetic circuit design automation. Science 352, aac3741 (2016).

28. Lucks, J. B., Qi, L., Whitaker, W. R. & Arkin, A. P. Toward scalable parts families for predictable design of biological circuits. Curr. Opin. Microbiol. 11, 567–573 (2008).

29. Gilman, J. & Love, J. Synthetic promoter design for new microbial chassis. Biochem. Soc. Trans. 44, 731–737 (2016).

30. Zobel, S. et al. Tin7-based device for calibrated heterologous gene expression in Pseudomonas putida. ACS Synth. Biol. 4, 1341–1351 (2015).

31. Guizzoni, S. et al. A part toolbox to tune genetic expression in Bacillus subtilis. Nucleic Acids Res. 44, 7495–7508 (2016).

32. Zhu, D. et al. Isolation of strong constitutive promoters from Lactococcus lactis subsp. lactis N8. FEMS Microbiol. Lett. 362, 1–6 (2015).

33. Jiao, X. et al. Vector promoters used in Lactococcus lactis. Microbiol. Methods 80, 715–724 (2010).

34. Spratt, J. M., Britton, W. J. & Triccas, J. A. Identification of strong promoter elements of Mycobacterium smegmatis and their utility for foreign gene expression in mycobacteria. FEMS Microbiol. Lett. 224, 139–142 (2003).

35. Kok, J. et al. The evolution of gene regulation research in Lactococcus lactis. FEMS Microbiol. Rev. 41, S220–S243 (2017).

36. Batianc, C. et al. An expanded CRISPRi toolbox for tunable control of gene expression in Pseudomonas putida. Microb. Cell Fact. 13, 368–385 (2020).

37. Rock, J. M. et al. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. Nat. Microbiol. 2, 1–9 (2017).

38. Peters, J. M. et al. Enabling genetic analysis of diverse bacteria with Mobile-CRISPR. Nat. Microbiol. 4, 244–250 (2019).

39. Berlec, A., Škrlček, K., Kočjan, J., Olenic, M. & Štrukelj, B. Single plasmid systems for inducible dual protein expression and for CRISPR-Cas9/CRISPRi gene regulation in lactic acid bacterium Lactococcus lactis. Sci. Rep. 8, 1–11 (2018).

40. Hueso-Gil, A., Nyerges, Á., PáI, C., Calles, B. & de Lorenzo, V. Multiple-site diversification of regulatory sequences enables interspecies operability of genetic devices. ACS Synth. Biol. 9, 104–114 (2020).

41. Castillo-Hair, S. M., Baerman, E. A., Fujita, M., Igoшин, O. A. & Tabor, J. J. Optogenetic control of Bacillus subtilis gene expression. Nat. Commun. 10, 1–11 (2019).

42. Aparicio, T., de Lorenzo, V. & Martínez-García, E. Improved thermotolerance of genome-reduced Pseudomonas putida EM42 enables effective functioning of the PL/LBS57 system. Biotechnol. J. 14, 1–8 (2019).

43. Serrano-Heras, G., Salas, M. & Bravo, A. A new plasmid vector for regulated gene expression in Bacillus subtilis. Plasmid 54, 278–282 (2005).
57. Meyers, A., Furtmann, C., Gesing, K., Torzikidas, I. E. P. & Jose, J. Cell density-dependent auto-inducible promoters for expression of recombinant proteins in Pseudomonas. Microb. Biotechnol. 12, 1062–1073 (2019).

58. Guan, C. et al. Construction and development of an auto-regulatory gene expression system in Bacillus subtilis. Microb. Cell Fact. 14, 1–15 (2015).

59. Jørgensen, C. M., Vrang, A. & Madsen, S. M. Recombinant protein expression in Lactococcus lactis using the P170 expression system. FEMS Microbiol. Lett. 270, 178–183 (2008).

60. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosomal binding sites to control protein expression. Nat. Biotechnol. 27, 946–950 (2009). This research article highlights the importance of synthetic ribosomal binding site design to create genetic circuitry with reliable and predictable expression levels.

61. Engstrom, M. D. & Pfeger, B. F. Transcription control engineering and applications in synthetic biology. Synth. Syst. Syst. Biotechnol. 2, 176–191 (2017).

62. Noll, P. et al. Precise quantification of translation inhibition by mRNA structures that overlap with the ribosomal footprint in N-terminal coding sequences. Nucleic Acids Res. 45, 5437–5448 (2017).

63. Wirth, N. T., Kozoeva, E. & Nikel, P. I. Accelerated genome engineering of Pseudomonas putida by I-CceI—mediated recombination and CRISPR-Cas9 counterselection. Microb. Biotechnol. 13, 233–249 (2019).

64. Wagemans, J. et al. Functional elucidation of antibacterial phage ORFans targeting Pseudomonas aeruginosa. Cell. Microbiol. 16, 1822–1835 (2014).

65. Bernard, P., Gabant, E. & Couturier, M. Positive-selection vectors using the F plasmid colE8 killer gene. Gene. 148, 71–74 (1994).

66. Psaty, J. H. & Yonesaki, T. D. Mdm of bacteriophage T4 functions as an antitoxin against Escherichia coli Lsod and RlnA toxins. Mol. Microbiol. 83, 669–681 (2012).

67. Lehnherr, H., Maguin, E., Jafari, S. & Yarmolinsky, M. B. Plasmid addiction genes of bacteriophage P1 doc, which causes cell death on curing of prophage, and phl, which prevents host death when prophage is retained. J. Mol. Biol. 233, 414–428 (1993).

68. Kushwaha, M., Rostain, W., Prakash, S., Duncan, J. N. & Jaramillo, A. Using transcription control engineering and its application to deplete the beta subunit of RNA polymerase. Proc. Natl Acad. Sci. USA 117, 20005858 (2020).

69. Adriaenssens, E. M. et al. Taxonomy of probakryotic viruses: 2018–2019 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. Archiv. Virusb. 165, 1253–1260 (2020).

70. Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D. & Voigt, C. A. A transcription based on a highly fragmented T7 RNA polymerase. Mol. Cell. Biol. 10, 742 (2014).

71. Liang, X. C., Wang, C. L., Li, Q. Integrating T7 RNA polymerase and its cognate transcriptional units for a host-independent and stable expression system in single plasmid. ACS Synth. Biol. 7, 1424–1435 (2018).

72. Kovalyova, I. V. & Kropinski, A. M. The complete genomic sequence of lytic bacteriophage gh-1. dna4 – the gene for the int core for the asymmetric homologous recombination events. Virology 344, 508–518 (2007).

73. Verma, V., Harjai, K. & Chhibber, S. Characterization of a T7-Like lytic bacteriophage of Klebsiella pneumoniae b5055: a potential therapeutic agent. Carr. Microbiol. 59, 274–281 (2009).

74. Lenneman, B. R. & Rothman-Denes, L. B. Structural and biochemical investigation of bacteriophage T4-encoded RNA polymerases. Biopolymers 5, 647–667 (2015).

75. Van Den Bossche, A. et al. Systematic identification of hypothetical bacteriophage proteins targeting key protein complexes of Pseudomonas aeruginosa. J. Proteome Res. 13, 4436–4456 (2014). Concept research paper that identifies phage-encoded modulators of key bacterial protein complexes.

76. Fox, T. D., Losick, R. & Pero, J. Regulatory gene 28 of bacteriophage SP01 codes for a phage-induced subunit of RNA polymerase. J. Mol. Biol. 101, 427–433 (1976).

77. Kotula, J. W. et al. Programmable bacteriophage detect and record an environmental signal in the mammalian gut. Proc. Natl Acad. Sci. USA 111, 4838–4843 (2014).

78. Walker, S. A. & Klaenhammer, T. R. Molecular characterization of a phage-inducible middle promoter and its transcriptional activator from the lacticococcal bacteriophage P31. J. Bacteriol. 180, 921–931 (1998).

79. Kutter, E., Raya, R. & Carlson, K. Phage-related RNA: structures and unexpected functions. In Bacteriophages: Biology and Applications (eds Kutter, E. & Sulakvelidze, A.) Ch. 7, 197–200 (CRC Press, 2004).

80. Fortier, L. C., Bransi, A. & Moineau, S. Genome sequence and global gene expression of T7-Like-like bacteriophage of Klebsiella pneumoniae. BioSyst. 311, 305–315 (2003).

81. Lenneman, B. R. & Rothman-Denes, L. B. Structural and biochemical investigation of bacteriophage T4-encoded RNA polymerases. Biopolymers 5, 647–667 (2015).

82. Van Den Bossche, A. et al. Systematic identification of hypothetical bacteriophage proteins targeting key protein complexes of Pseudomonas aeruginosa. J. Proteome Res. 13, 4436–4456 (2014). Concept research paper that identifies phage-encoded modulators of key bacterial protein complexes.

83. Fortier, L. C., Bransi, A. & Moineau, S. Genome sequence and global gene expression of T4, a new phage species linking the 936 and c2 phage species of Lactococcus lactis. J. Bacteriol. 188, 6110–6114 (2006).

84. Dinman, J. D. Translational recording signals: expanding the synthetic biology toolbox. J. Biol. Chem. 294, 7537–7545 (2019).

85. Samatova, E., Konevva, A. L., Wills, N. M., Atkins, J. F. & Rodnina, M. V. High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. Nat. Commun. 5, 4459 (2014).

86. Kutter, E., Raya, R. & Carlson, K. Phage-related RNA: structures and unexpected functions. In Bacteriophages: Biology and Applications (eds Kutter, E. & Sulakvelidze, A.) Ch. 7, 197–200 (CRC Press, 2004).

87. Fortier, L. C., Bransi, A. & Moineau, S. Genome sequence and global gene expression of Q5, a new phage species linking the 936 and c2 phage species of Lactococcus lactis. J. Bacteriol. 188, 6110–6114 (2006).

88. Dinman, J. D. Translational recording signals: expanding the synthetic biology toolbox. J. Biol. Chem. 294, 7537–7545 (2019).

89. Samatova, E., Konevva, A. L., Wills, N. M., Atkins, J. F. & Rodnina, M. V. High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. Nat. Commun. 5, 4459 (2014).

90. Stacic, D., Lindell, D. & Steglich, C. Antisense RNA protects mRNA from RNAS e degradation by RNA–RNA duplex formation during phage infection. Nucleic Acids Res. 39, 11689–11697 (2011).

91. Van Den Bossche, A. et al. Structural elucidation of a novel mechanism for the bacteriophage-based inhibition of the RNA degradation. Elife 5, 1–20 (2016).

92. Neves, D., Vos, S., Blank, L. M. & Ebert, B. E. Pseudomonas mRNA 2.0: boosting gene expression through enhanced mRNA stability and translational efficiency. Front. Bioeng. Biotechnol. 7, 1–11 (2020).

93. Kurochkina, L. P. et al. Potential and functional characterization of the first bacteriophage-encoded chaperonin. J. Virol. 86, 10103–10111 (2012).

94. Semenyuk, P. I., Orlov, V. N., Sokolova, O. S. & Kurochkina, L. P. New GroEL-like chaperonin of bacteriophage OB Pseudomonas fluorescens suppresses thermal protein aggregation in an ATP-dependent manner. Mol. Cell. Proteomics 17, 473–483 (2018).

95. Semenyuk, P. I., Moiseenko, A. V., Sokolova, O. S., Muronetz, V. I. & Kurochkina, L. P. Structural and functional diversity of novel and known...
162. Ronda, C., Pedersen, L. E., Sommer, M. O. A. & Nielsen, A. T. CRMAGE: CRISPR Optimized MAGE Recombineering. Scit. Rep. 6, 1–11 (2016).
163. Mansell, T. J., Warner, J. R. & Gill, R. T. Trackable multiplex recombineering for gene-trait mapping in E. coli. In Systems Metabolic Engineering. Methods in Molecular Biology (Methods and Protocols), vol. 985 (ed. Alper, H.) 223–247 (Humana Press, 2013).
164. Nyerges, Á. et al. Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. Proc. Natl Acad. Sci. USA 115, E5726–E5735 (2018).

Acknowledgements
This paper was supported by funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 819800) awarded to R.L. P.I.N. acknowledges financial support by The Novo Nordisk Foundation (grants NNF10CC1016517 and NNF18OC0034818), the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 814418, SinFonia), and the Danish Council for Independent Research (SWEET, DFF-Research Project 8021-00039B).

Author contributions
All three authors contributed to the overall structure and design of the manuscript. Principal writing was performed by E.-M.L., with corrections and modifications by P.I.N. and R.L.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-19124-x.

Correspondence and requests for materials should be addressed to R.L.

Peer review information Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020