

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Biotechnology Tools Derived from the Bacteriophage/Bacteria Arms Race

*Vitor B. Pinheiro*

## Abstract

The long association and intense competition between bacteria and their viruses have created a fertile ground for evolution to develop numerous tools for DNA modification, assembly and degradation. Many of these tools underpin the past 50 years of molecular biology, and others show great potential in shaping the next 50 years of the field. Here, I present some of the tools that have come out of the bacteria-bacteriophage arms race and discuss some of the concepts that may shape their future use. Molecular biology remains a fast-growing area increasingly limited solely by researcher ingenuity.

**Keywords:** molecular biology tools, orthogonality, DNA modifications

## 1. Introduction

The relationship between bacteriophages and bacteria is often explained in terms of an arms race: each ‘developing’ measures and countermeasures for attacking and defending itself from the other [1, 2]. The imagery of an arms race is a powerful metaphor to summarise the relationship between possibility and availability that have constrained the emergence, evolution and diversification of life on Earth.

Life is limited by what is possible. For instance, life can only exist because chemical information storage is possible. On the other hand, life as we know it has evolved around DNA and RNA because they are informational molecules that could function in the environment of the early Earth and whose building blocks were readily available.

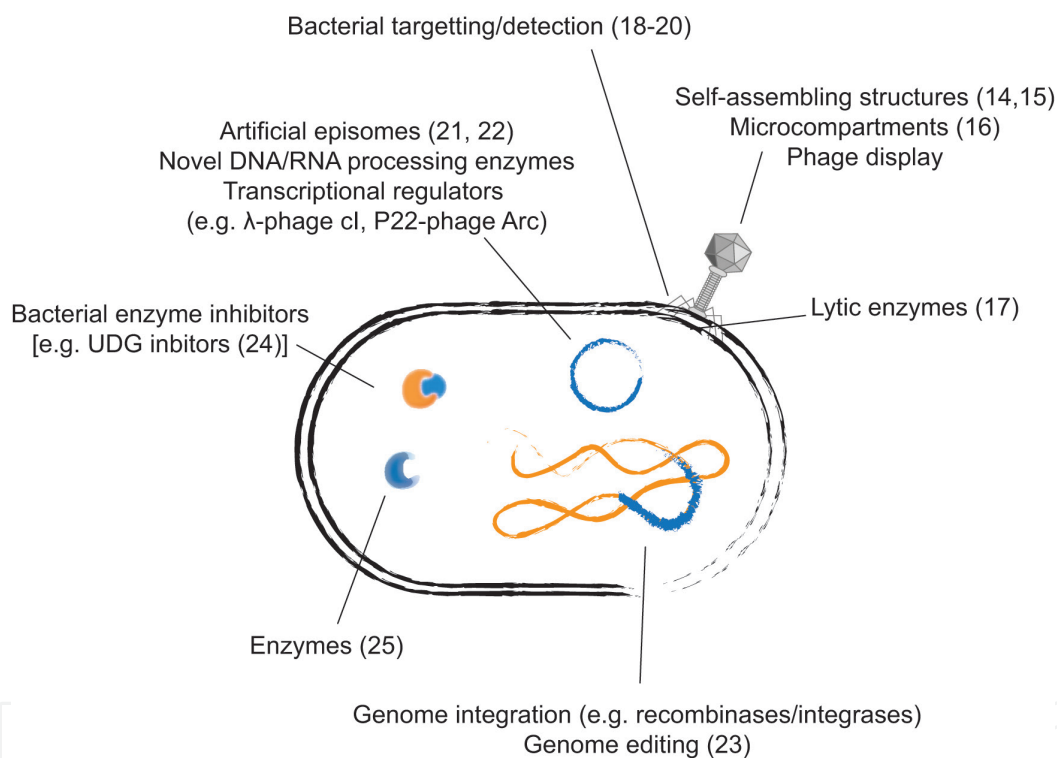
The relationship between bacteria and bacteriophages is similarly constrained. The emergence (or availability) of bacterial cells capable of establishing a rich internal environment (compared to the outside of the cell) creates the possibility for other organisms to evolve predatory or parasitic survival strategies, including bacteriophages. Once phages emerge, they alter the dynamics of the ecological niche and create an advantage to bacterial hosts that can reduce the success of phage infection—whether by hindering phage access to the cell cytoplasm, by interfering with phage survival or replication in the cell or by interfering with phage maturation and release [3].

Bacterial defences arise from any function already available in the host (e.g. uracil-DNA glycosylase involved in DNA repair) or that can be co-opted from available genetic resources in the cell or in the environment

(e.g. restriction-modification systems). Defences can also emerge from loss of function (e.g. mutations to the maltose porin LamB in *E. coli*, which make it resistant to bacteriophage  $\lambda$  infections [4]).

Given the prevalence of bacteria and phages in the environment, and given the evolutionary scale time of their arms race, the variety, complexity and efficiency of these attack and defence strategies are huge and can range from silent integration into the host genome (i.e. lysogeny) to enacting a hostile molecular takeover of the bacterial host cell. Despite our current efforts to map the genetic diversity available on Earth, it remains likely that new strategies are still to be identified and characterised.

Nonetheless, many of these defence and attack strategies have also been harnessed for biotechnology applications, significantly beyond the simple use of bacteriophages (or bacteriophage proteins) as bacterial control agents [5–7]. Restriction-modification (RM) systems found in bacteria were among the very first tools isolated from the bacteria-phage arms race [8, 9]. They *de facto* represent the start of modern molecular biology, and they have remained key tools for over 50 years (**Figure 1**).



**Figure 1.**

*Potential biotechnology applications derived from bacteriophages. Almost every aspect of the bacteriophage life cycle can be exploited for the development of biotechnology tools. Aside from their natural bactericidal role, phages can be used for the delivery of engineered circuits [10, 11] and in directed evolution through phage display [12] and PACE [13]. Individual phage systems have been also successfully developed as tools.*

More recently, another tool derived from bacterial defence has been harnessed, with potentially equally transformative impact on how we interact with biology: clustered regularly interspersed palindromic repeats (CRISPR) [26–28]. CRISPR forms part of an adaptive immunity system in prokaryotes, but it is being harnessed to deliver a wide range of research and therapeutic tools.

Although RM systems and CRISPR are deservedly acknowledged as having a significant impact on molecular biology and biotechnology, many other tools have been or are being developed based on components isolated from the bacteria-phage arms race. This chapter focuses on some of those tools—their mechanisms, current and potential applications.

## 2. Common molecular biology tools and orthogonality

Because of the wide range of bacteriophage infection strategies available, it becomes difficult to introduce simple classification without recreating the complexity of approaches taken by phages. For instance, bacteriophage promoters can rely exclusively on host proteins (e.g. T4 early promoters), on a mixture of host and phage proteins (e.g. T4 middle promoters or P<sub>L</sub> promoter from  $\lambda$  phage) or exclusively on phage-derived factors (e.g. T7 RNA polymerase promoters). This provides a continuum that can be further dissected by analysing the mechanism of the hybrid promoters, with their specific host and phage dependencies.

That continuum maps how independent a phage system is from the host while still active within the host, i.e. it is a measure of the orthogonality of the system. Having evolved to survive in a changing environment, bacteria have complex layers of gene expression regulation with multiple feedback systems which are not necessarily easy to control independently, despite our advances in understanding bacterial metabolism [29, 30]. In that context, phage systems that have reduced dependencies on the host machinery (i.e. increased orthogonality) provide isolated systems that can be simpler to regulate and are, at least in part, shielded from variations in the cellular machinery—an approach that has dominated biotechnology until recently.

Many of the common phage-derived biotechnology tools have been developed from such systems, none more so than T7 RNA polymerase [31]. Isolated from T7 bacteriophage, this monomeric RNA polymerase can recognise a specific promoter sequence. The core T7 promoter sequence (TAATACGACTCACTATAG) is sufficient to trigger transcription in cells harbouring a T7 RNA polymerase gene. This is the strategy set up behind pET vectors, which contain a T7 promoter and rely on an *E. coli* host carrying a T7 RNA polymerase under an inducible promoter (e.g. *lacUV*)—usually the result of the introduction of a DE3 phage [32, 33].

Nevertheless, the context of the T7 promoter can have a significant impact on the expression level of the downstream genes, and at high polymerase concentrations, it is possible to drive transcription from suboptimal promoters—highlighting that the orthogonality of the system is limited.

Given its monomeric structure and orthogonal role in cellular transcription, T7 RNA polymerase became not only a useful tool in biotechnology but also an important model system for the study of transcription (reviewed in [34]). Because of its orthogonality, T7 RNA polymerase (and its promoter) can be harnessed for the regulation of transcription in a wide range of hosts beyond *E. coli*, including Gram-positive bacterial hosts [35], yeast [36, 37] and human cells [38, 39].

Its role in the regulation of transcription has also been expanded through the creation of more complex systems using split T7 RNA polymerase proteins. Surprisingly for a mesophilic highly dynamic enzyme, T7 RNA polymerase can be expressed in two [40, 41] or more [42] fragments that *in vivo* are able to reassemble and function as viable RNA polymerases.

While orthogonality can be a desirable feature for *in vivo* applications, it is wholly unnecessary for *in vitro* applications, where the key constraint lies on identifying reaction conditions in which expressed and purified proteins are sufficiently active to carry out the desired function. That is the case with T4 DNA ligase and T4 polynucleotide kinase, which were originally isolated from the *E. coli* T4 phage and remain important tools in molecular biology.

T4 DNA ligase has a central role in the replication and repair of the phage genome during its infection of *E. coli* [43]. This also entails coping with DNA modifications such as the full substitution of cytosine for 5-hydroxymethylcytosine and glucosyl-5-hydroxymethylcytosine that naturally occurs *in vivo* [44, 45]—this



is a phage defence mechanism discussed below. Although its structure has only recently been determined [46], the mechanism of action of this enzyme has long been characterised [47]. Even in the absence of other phage genes, it is active *in vivo* [48], but its main application in molecular biology remains its *in vitro* activity that, coupled with restriction endonucleases, has underpinned modern molecular biology—allowing a molecular cut and paste approach to DNA assembly.

Ligase *in vitro* activity, particularly its ability to accept modified ligands, has been extensively explored for the assembly of heavily modified DNA sequences for aptamer selection [49, 50] and to explore a wider range of nucleic acid modifications, such as sugar-modified nucleic acids [51].

### 3. Second-generation tools and applications

The combination of different enzymatic functions has created novel applications, such as the large-scale DNA assembly in Gibson assembly through the combination of exonuclease, polymerase and ligase activities [52]. However, a whole range of novel applications are possible by harnessing additional bacteriophage proteins that have not yet been extensively explored.

Recombinases and integrases, enzymes that catalyse the sequence-specific insertion of a phage genome into the host chromosome [53, 54], were identified early in bacteriophage research (e.g. the  $\lambda$  integrase) but were not immediately harnessed for applications. These came substantially later once recombinases were shown to facilitate DNA assembly, whether by increasing the efficiency of subcloning such as in Gateway cloning [55], or enabling multipart assemblies needed for metabolic engineering [56].

In general, recombinases bind DNA specifically, as dimers, on recognition sites that are relatively short (usually between 30 and 50 bases) and partially palindromic, termed *attP* and *attB* (originally to distinguish phage and bacterial origins). These two sites have different sequences which contribute to making the recombination process unidirectional. The recombinases facilitate the break and religation of double-stranded DNA within the *att* sites resulting in chimeric sites that are then labelled *attR* and *attL* (from right and left sides, respectively). Insertion of a phage genome into the host is efficient and stable, but it can also be reversed with the contribution of a single host factor (reviewed in [57]).

Because of the high efficiency of integration, recombinases have been also developed as systems to facilitate homologous recombination in higher eukaryotes, such as mammalian cells [58], *Drosophila* organisms [59] and plants (reviewed in [60]). In the latter, recombinases were of particular interest because of their potential to remove transformation and selection markers from engineered crops, leaving behind only the genes responsible for the engineered trait. This idea of using recombinases to directly alter an organism's genome has been vastly expanded in the Yeast 2.0 project, through the design implementation of multiple *loxP* (the equivalent to *att* sites for Cre recombinase) in the synthetic yeast genome that can be activated, leading to large-scale genome rearrangement—termed synthetic chromosome rearrangement and modification by *loxP*-mediated evolution (SCRaMbLE) [61].

By enabling controllable chromosome rearrangements between the designed *loxP* sites, a synthetic yeast can delete, duplicate and reorder many of its genes, allowing *in vivo* selection for desirable traits such as increased alkali tolerance [62]. Alternatively, the system can be coupled to heterologously expressed genes allowing the rapid optimisation of pathways [63].

Like SCRaMbLE, protein-directed evolution relies on cycles of mutagenesis (to introduce diversity into a population) and selection (to reduce diversity towards

functional proteins) [64] which in some platforms can be achieved continuously *in vivo*—e.g. in PACE [13] or in some continuous culture approaches [65]. In both examples, mutation can be controlled by stressors or the induction of error-prone replication but are not necessarily limited to the area of interest (e.g. a single gene). Greater control of targeting is possible and has been reported through the use of an error-prone PolI [66]—which is necessary for the replication of some bacterial plasmids and can be used to drive diversification *in vivo* in the vicinity of plasmid replication initiation—and protein fusions that target an error-prone polymerase to a particular region of the genome [e.g. EvolvR and MutaT7 (reviewed in [67])].

However, one such system, termed diversity-generating retroelements (DGRs), has naturally emerged in bacteriophages and was first implicated in the tropism switching of *Bordetella* phages [68, 69] but that has since been identified in a wide range of bacterial and archaeal genomes [70, 71]. The system relies on an error-prone reverse transcriptase (RT) and on two DNA repeats—one operating as a template (template repeat) and the other as the target (variable repeat). RNA synthesised from the template is used to guide the error-prone synthesis of a DNA by the RT, which also coordinates its insertion at the target site. The error-profile of DGR RTs results in adenines being replaced with other bases, creating a directionality to the evolution that is always anchored by the template repeat [72]. Nevertheless, changes in DNA sequences involved in the targeting of the retroelement, termed initiation of mutagenic homing (IMH), can lead to both template and variable repeats being allowed to evolve at different rates—removing some of the directionality in evolution and freeing the system to explore the sequence space more thoroughly [68].

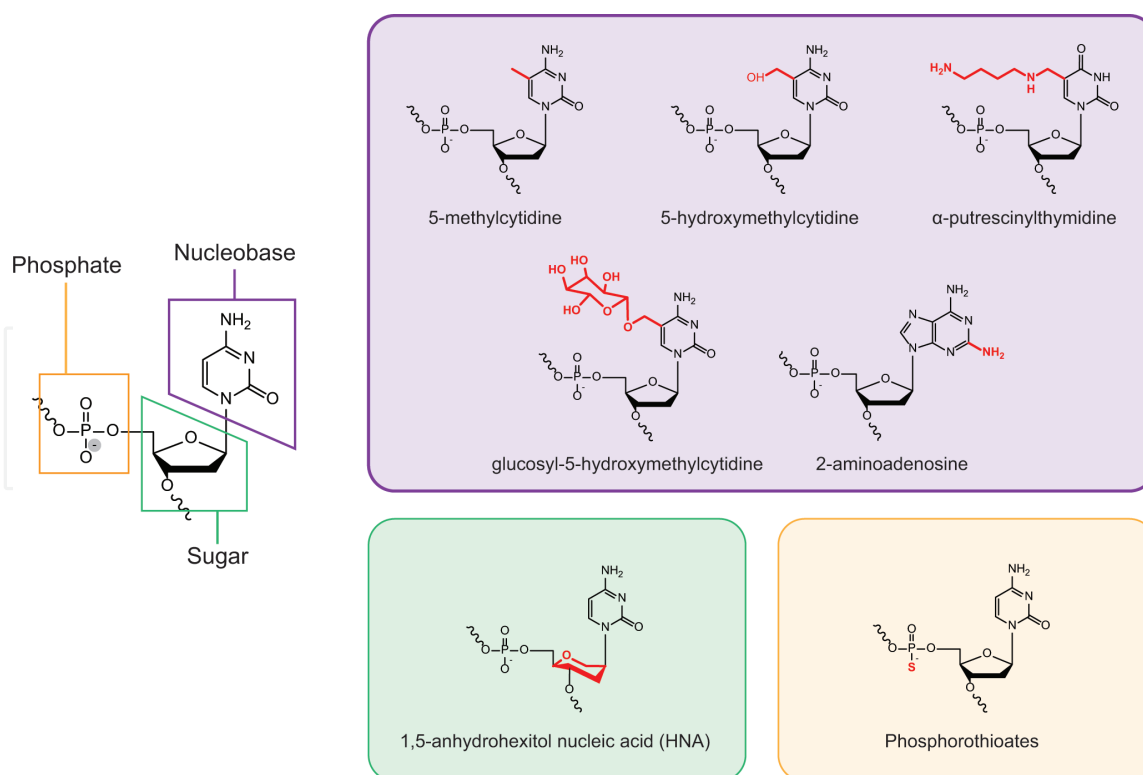
Despite its potential, it remains to be seen whether such a system can be harnessed for protein engineering. If the DGR systems can themselves be engineered, their targeting and error rate may be amenable to modulation opening possibilities to compete with the most recent Cas9-derived gene editors [73].

#### 4. Xenobiotic nucleic acids

Chemical modification of the phage's own genome is a widespread strategy that emerged multiple times in evolution to circumvent (or at least slow down) bacterial defence mechanisms that target the invading DNA for degradation: restriction endonucleases, exonucleases and CRISPR-Cas systems [44, 74]. Those modifications have been reported not only on the nucleobases, akin to eukaryotic epigenetic markers, but also on the nucleic acid backbone.

Reported nucleobase modifications suggest that the overwhelming majority of any such modifications is targeted to the C5 position in pyrimidines. They range from small modifications, such as methylation, to bulky modifications, such as putrescine and even carbohydrate moieties. While such modifications have long been known, new sequencing platforms capable of reading the DNA sequence without an amplification step, such as nanopore sequencing, hold great promise in enabling a more systematic mapping and characterisation of DNA modifications in phage genomes (**Figure 2**) [75].

DNA modifications, particularly modifications that bring chemical functionality not available in natural bases, such as glycosylation in *Bacillus subtilis* SP-15 phage [76], can be harnessed for function as has been achieved through the chemical modification of DNA bases and systematic evolution of ligands by exponential enrichment (SELEX) [50, 77]. Despite characterisation of the biosynthetic pathway for multiple-phage DNA modification systems, none have been implemented *in vitro* for applications.

**Figure 2.**

Examples of XNAs in phages. Nucleobases are composed of three different chemical moieties, and change to any one of them has the potential to alter duplex conformation and how easily they are recognised by natural nucleic acid processing enzymes [14–25]. The base, sugar and backbone discussed in the chapter are shown here.

One potential bottleneck lies on how the phage and the host handle those chemically modified DNAs. Upon infection, the mature phage DNA needs to be modified if that is an evolutionary strategy being exploited to slow down or avoid *in vivo* degradation. On the other hand, those chemical modifications can affect DNA structure and biophysical properties which may also be detrimental to bacteriophage replication—since this would require a DNA polymerase capable of processing such heavily modified genomes.

It is known that at least in some cases, this chemical cloaking is removed upon infection, such as in SP-15 [78], before unmodified DNA is replicated *in vivo*. But, given viral polymerases more permissive substrate specificity, it is possible that some systems can be replicated directly by highly adapted phage polymerases—either to DNA followed by reinstalling the chemical modifications or directly from modified DNA to modified DNA. In the case of SP-15, the bulkier modifications are rapidly removed prior to replication of DNA [78]. T4 seems to follow a similar pattern where glycosylation is removed, and DNA is replicated containing only the simpler 5-hydroxymethylation modification. This is further supported by the biochemical evidence that glycosylation is ‘installed’ on the replicated T4 DNA [44, 74]. Nevertheless, early T4 transcription occurs rapidly, and it is carried out by the host RNA polymerase, suggesting that natural RNA polymerases can still use the hypermodified bases in that template.

Notably, although phage polymerases replicate phage DNA *in vivo* harbouring simple chemical modifications, such as (in *Synechocystis* S-2 L phage) 2-amino-adenine [79] and uracil (in *Bacillus* phages AR9, PBS1 and PBS2, *Yersinia* phage PhiR1-37 and *Staphylococcus* phage S6), no viral polymerase has been described that is capable of selectively incorporating the modified bases. That is, although some bacteriophages make use of modified nucleobases and have evolved systems that lead to 100% incorporation of the modified bases in their genomes, their DNA polymerases have not specialised towards being able to only incorporate the modified nucleobases—they remain able to recognise unmodified triphosphates.

Still, the increased substrate flexibility of phage DNA polymerases may at least in part justify why a *Bacillus subtilis* Phi29 DNA polymerase required a single mutation for the synthesis of anhydrohexitol nucleic acids (HNA) [80] while an archaeal enzyme required in excess of seven mutations [81].

## 5. Conclusion

Bacteriophages remain a rich source of novel functionalities that can be harnessed to advance molecular biology (and synthetic biology). The examples here provided represent only a small fraction of the potential applications available, which also include medical applications from phage proteins [82, 83] and engineered phages [10, 11, 84].

In addition, bacteriophages have had a close relationship with directed evolution, either as a vehicle such as in phage display [85, 86] or by providing (in addition to the examples above) proteins to accelerate strain engineering, such as in multiplex automated genome engineering (MAGE) [87].

Finally, bacteriophages may also become an important tool in harnessing new non-model organisms in synthetic biology, as pre-optimised DNA delivery nanomachines for custom circuits.

### Author details


Vitor B. Pinheiro<sup>1,2</sup>

1 Rega Institute for Medical Research, KU Leuven, Leuven, Belgium

2 Department of Structural and Molecular Biology, University College London, London, UK

\*Address all correspondence to: [vitor.pinheiro@kuleuven.be](mailto:vitor.pinheiro@kuleuven.be)

### IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 



## References

- [1] Koonin EV, Makarova KS, Wolf YI. Evolutionary genomics of defense systems in Archaea and Bacteria. *Annual Review of Microbiology*. 2017;**71**(1):233-261. DOI: 10.1146/annurev-micro-090816-093830
- [2] Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. *Nature Reviews. Microbiology*. 2010;**8**:317. DOI: 10.1038/nrmicro2315
- [3] Seed KD. Battling phages: How bacteria defend against viral attack. *PLoS Pathogens*. 2015;**11**(6):e1004847. DOI: 10.1371/journal.ppat.1004847
- [4] Chatterjee S, Rothenberg E. Interaction of bacteriophage  $\lambda$  with its *E. coli* receptor, LamB. *Viruses*. 2012;**4**:3162-3178
- [5] Lemire S, Yehl KM, Lu TK. Phage-based applications in synthetic biology. *Annual Review of Virology*. 2018;**5**(1):453-476. DOI: 10.1146/annurev-virology-092917-043544
- [6] Salmond GPC, Fineran PC. A century of the phage: Past, present and future. *Nature Reviews. Microbiology*. 2015;**13**:777. DOI: 10.1038/nrmicro3564
- [7] Kortright KE, Chan BK, Koff JL, Turner PE. Phage therapy: A renewed approach to combat antibiotic-resistant bacteria. *Cell Host & Microbe*. 2019;**25**(2):219-232. Available from: <https://www.sciencedirect.com/science/article/pii/S1931312819300526?via%3Dihub>
- [8] Tock MR, Dryden DTF. The biology of restriction and anti-restriction. *Current Opinion in Microbiology*. 2005;**8**(4):466-472
- [9] Pingoud A, Wilson GG, Wende W. Type II restriction endonucleases—A historical perspective and more. *Nucleic Acids Research*. 2014;**42**(12):7489-7527. DOI: 10.1093/nar/gku447
- [10] Ando H, Lemire S, Pires DP, Lu TK. Engineering modular viral scaffolds for targeted bacterial population editing. *Cell Systems*. 2015;**1**(3):187-196
- [11] Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, et al. Exploiting CRISPR-cas nucleases to produce sequence-specific antimicrobials. *Nature Biotechnology*. 2014;**32**(11):1146-1150
- [12] Fernandez-Gacio A, Uguen M, Fastrez J. Phage display as a tool for the directed evolution of enzymes. *Trends in Biotechnology*. 2003;**21**(9):408-414
- [13] Esvelt KM, Carlson JC, Liu DR. A system for the continuous directed evolution of biomolecules. *Nature*. 2011;**472**(7344):499. DOI: 10.1038/nature09929
- [14] Moon J-S, Kim W-G, Kim C, Park G-T, Heo J, Yoo SY, et al. M13 bacteriophage-based self-assembly structures and their functional capabilities. *Mini-Reviews in Organic Chemistry*. 2015;**12**(3):271-281
- [15] Han J, Devaraj V, Kim C, Kim W-G, Han D-W, Hong SW, et al. Fabrication of self-assembled nanoporous structures from a self-templating M13 bacteriophage. *ACS Applied Nano Materials*. 2018;**1**(6):2851-2857. DOI: 10.1021/acsanm.8b00500
- [16] Krupovic M, Koonin EV. Multiple origins of viral capsid proteins from cellular ancestors. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;**114**(12):E2401-E2410
- [17] Wang I-N, Smith DL, Young R. Holins: The protein clocks of bacteriophage infections. *Annual Review of Microbiology*. 2000;**54**(1):799-825

- [18] Denyes JM, Dunne M, Steiner S, Mittelviefhaus M, Weiss A, Schmidt H, et al. Modified bacteriophage S16 long tail fiber proteins for rapid and specific immobilization and detection of *Salmonella* cells. *Applied and Environmental Microbiology*. 2017;**83**(12):e00277-17
- [19] Schmelcher M, Loessner MJ. Bacteriophage: Powerful tools for the detection of bacterial pathogens. In: *Principles of Bacterial Detection: Biosensors, Recognition Receptors and Microsystems*. New York: Springer; 2008. p. 731-754. ISBN: 978-0-387-75112-2
- [20] Dunne M, Loessner MJ. Modified bacteriophage tail fiber proteins for labeling, immobilization, capture, and detection of bacteria. *Methods in Molecular Biology*. 2019;**1918**:67-86
- [21] Ravikumar A, Arrieta A, Liu CC. An orthogonal DNA replication system in yeast. *Nature Chemical Biology*. 2014;**10**(3):175-177
- [22] Mencía M, Gella P, Camacho A, de Vega M, Salas M. Terminal protein-primed amplification of heterologous DNA with a minimal replication system based on phage Phi29. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;**108**(46):18655-18660. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3219123&tool=pmcentrez&rendertype=abstract>
- [23] Carr PA, Wang HH, Sterling B, Isaacs FJ, Lajoie MJ, Xu G, et al. Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Research*. 2012;**40**(17):e132
- [24] Serrano-Heras G, Ruiz-Masó JA, Del Solar G, Espinosa M, Bravo A, Salas M. Protein p56 from the *Bacillus subtilis* phage Φ29 inhibits DNA-binding ability of uracil-DNA glycosylase. *Nucleic Acids Research*. 2007;**34**(16):5393-5401
- [25] Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe JA, et al. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;**108**(39):E757-E764
- [26] Lander ES. The heroes of CRISPR. *Cell*. 2016;**164**(1-2):18-28. Available from: <https://www.sciencedirect.com/science/article/pii/S0092867415017055?via%3Dihub#bib34>
- [27] Mojica FJM, Juez G, Rodriguez-Valera F. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Molecular Microbiology*. 1993;**9**(3):613-621. DOI: 10.1111/j.1365-2958.1993.tb01721.x
- [28] Jansen R, Embden JDA van, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*. 2002;**43**(6):1565-1575. DOI: 10.1046/j.1365-2958.2002.02839.x
- [29] Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, et al. Complete chemical synthesis, assembly, and cloning of a mycoplasma genitalium genome. *Science*. 2008;**319**(5867):1215-1220
- [30] Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, et al. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011. *Molecular Systems Biology*. 2011;**7**(1):535
- [31] Tabor S, Richardson CC. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes.

Proceedings of the National Academy of Sciences of the United States of America. 1985;**82**(4):1074-1078

[32] Rosenberg AH, Lade BN, Dao-shan C, Lin SW, Dunn JJ, Studier FW. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*. 1987;**56**(1):125-135

[33] Giordano TJ, Deuschle U, Bujard H, McAllister WT. Regulation of coliphage T3 and T7 RNA polymerases by the lac repressor-operator system. *Gene*. 1989;**84**(2):209-219

[34] Sousa R. T7 RNA polymerase. In: *Encyclopedia of Biological Chemistry*. 2nd ed. Elsevier; 2013. p. 355-359. ISBN: 9780123786319

[35] Kortmann M, Kuhl V, Klaffl S, Bott M. A chromosomally encoded T7 RNA polymerase-dependent gene expression system for *Corynebacterium glutamicum*: Construction and comparative evaluation at the single-cell level. *Microbial Biotechnology*. 2015;**8**(2):253-265

[36] Hobl B, Hock B, Schneck S, Fischer R, Mack M. Bacteriophage T7 RNA polymerase-based expression in *Pichia pastoris*. *Protein Expression and Purification*. 2013;**92**(1):100-104

[37] Benton BM, Eng WK, Dunn JJ, Studier FW, Sternglanz R, Fisher PA. Signal-mediated import of bacteriophage T7 RNA polymerase into the *Saccharomyces cerevisiae* nucleus and specific transcription of target genes. *Molecular and Cellular Biology*. 1990;**10**(1):353-360

[38] Aoki Y, Aizaki H, Shimoike T, Tani H, Ishii K, Saito I, et al. A human liver cell line exhibits efficient translation of HCV RNAs produced by a recombinant adenovirus expressing T7 RNA polymerase. *Virology*. 1998;**250**(1):140-150

[39] Dunn JJ, Kripl B, Bernstein KE, Westphal H, William SF. Targeting bacteriophage T7 RNA polymerase to the mammalian cell nucleus. *Gene*. 1988;**68**(2):259-266

[40] Schaerli Y, Gili M, Isalan M. A split intein T7 RNA polymerase for transcriptional AND-logic. *Nucleic Acids Research*. 2014;**42**(19):12322-12328

[41] Shis DL, Bennett MR. Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;**110**(13):5028-5033

[42] Segall-Shapiro TH, Meyer AJ, Ellington AD, Sontag ED, Voigt CA. A 'resource allocator' for transcription based on a highly fragmented T7 RNA polymerase. *Molecular Systems Biology*. 2014;**10**(7):742

[43] Mueser TC, Hinerman JM, Devos JM, Boyer RA, Williams KJ. Structural analysis of bacteriophage T4 DNA replication: A review in the virology journal series on bacteriophage T4 and its relatives. *Virology Journal*. 2010;**7**(1):359

[44] Weigle P, Raleigh EA. Biosynthesis and function of modified bases in bacteria and their viruses. *Chemical Reviews*. 2016;**116**(20):12655-12687

[45] Vlot M, Houkes J, SJA L, Swarts DC, Zheng P, Kunne T, et al. Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR-Cas effector complexes. *Nucleic Acids Research*. 2018;**46**(2):873-885

[46] Shi K, Bohl TE, Park J, Zasada A, Malik S, Banerjee S, et al. T4 DNA ligase structure reveals a prototypical ATP-dependent ligase with a unique mode of sliding clamp



interaction. *Nucleic Acids Research*. 2018;**46**(19):10474-10488

*Journal of Molecular Biology*. 2004;**335**(3):667-678

[47] Rossi R, Montecucco A, Ciarrocchi G, Biamonti G. Functional characterization of the T4 DNA ligase: A new insight into the mechanism of action. *Nucleic Acids Research*. 1997;**25**(11):2106-2113

[55] Hartley JL, Temple GF, Brasch MA. DNA cloning using *in vitro* site-specific recombination. *Genome Research*. 2000;**10**(11):1788-1795

[48] Su T, Liu F, Chang Y, Guo Q, Wang J, Wang Q, et al. The phage T4 DNA ligase mediates bacterial chromosome DSBs repair as single component non-homologous end joining. *Synthetic and Systems Biotechnology*. 2019;**4**(2):107-112

[56] Colloms SD, Merrick CA, Olorunniji FJ, Stark WM, Smith MCM, Osbourn A, et al. Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. *Nucleic Acids Research*. 2014;**42**(4):e23

[49] Hili R, Niu J, Liu DR. DNA ligase-mediated translation of DNA into densely functionalized nucleic acid polymers. *Journal of the American Chemical Society*. 2013;**135**(1):98-101. DOI: 10.1021/ja311331m

[57] Rutherford K, Van Duyne GD. The ins and outs of serine integrase site-specific recombination. *Current Opinion in Structural Biology*. 2014;**24**:125-131

[50] Kong D, Yeung W, Hili R. In vitro selection of diversely functionalized aptamers. *Journal of the American Chemical Society*. 2017;**139**(40):13977-13980

[58] Thyagarajan B, Olivares EC, Hollis RP, Ginsburg DS, Calos MP. Site-specific genomic integration in mammalian cells mediated by phage  $\phi$ C31 integrase. *Molecular and Cellular Biology*. 2001;**21**(12):3926-3934

[51] Kestemont D, Renders M, Leonczak P, Abramov M, Schepers G, Pinheiro VB, et al. XNA ligation using T4 DNA ligase in crowding conditions. *Chemical Communications*. 2018;**54**(49):6408-6411. DOI: 10.1039/C8CC02414F

[59] Groth AC, Fish M, Nusse R, Calos MP. Construction of transgenic *Drosophila* by using the site-specific integrase from phage  $\phi$ C31. *Genetics*. 2004;**166**(4):1775-1782

[52] Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*. 2009;**6**(5):343-345

[60] Ow DW. The long road to recombinase-mediated plant transformation. *Plant Biotechnology Journal*. 2016;**14**(2):441-447

[53] Landy A. The  $\lambda$  integrase site-specific recombination pathway. In: *Mobile DNA III*. American Society of Microbiology; 2015. p. 91-118. ISSN: 2165-0497

[61] Shen Y, Stracquadanio G, Wang Y, Yang K, Mitchell LA, Xue Y, et al. SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Research*. 2016;**26**(1):36-49

[54] Groth AC, Calos MP. Phage integrases: Biology and applications.

[62] Ma L, Li Y, Chen X, Ding M, Wu Y, Yuan YJ. SCRaMbLE generates evolved yeasts with increased alkali tolerance. *Microbial Cell Factories*. 2019;**18**(1):52

[63] Liu W, Pham NT, Tuck L, Pérez-Pi I, Shen Y, French C, et al. Rapid pathway



prototyping and engineering using *in vivo* and *in vivo* synthetic genome SCRaMbLE-in methods. Nature Communications. 2018;**9**(1):1936

[64] Tizei PAG, Csibra E, Torres L, Pinheiro VB. Selection platforms for directed evolution in synthetic biology. Biochemical Society Transactions. 2016;**44**(4):1165-1175. DOI: 10.1042/BST20160076

[65] Marlière P, Patrouix J, Döring V, Herdewijn P, Tricot S, Cruveiller S, et al. Chemical evolution of a bacterium's genome. Angewandte Chemie, International Edition. 2011;**50**(31):7109-7114

[66] Camps M, Naukkarinen J, Johnson BP, Loeb LA. Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. Proceedings of the National Academy of Sciences of the United States of America. 2003;**100**(17):9727-9732

[67] Yang J, Kim B, Kim GY, Jung GY, Seo SW. Synthetic biology for evolutionary engineering: From perturbation of genotype to acquisition of desired phenotype. Biotechnology for Biofuels. 2019;**12**(1):113

[68] Doulatov S, Hodes A, Dal L, Mandhana N, Liu M, Deora R, et al. Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. Nature. 2004;**431**(7007):476-481

[69] Liu M, Deora R, Doulatov SR, Gingery M, Eiserling FA, Preston A, et al. Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. Science. 2002;**295**(5562):2091-2094

[70] Paul BG, Burstein D, Castelle CJ, Handa S, Arambula D, Czornyj E, et al. Retroelement-guided protein diversification abounds in vast lineages of bacteria and archaea. Nature Microbiology. 2017;**2**(6):17045

[71] Paul BG, Bagby SC, Czornyj E, Arambula D, Handa S, Sczyrba A, et al. Targeted diversity generation by intraterrestrial archaea and archaeal viruses. Nature Communications. 2015;**6**(1):6585

[72] Guo H, Tse LV, Barbalat R, Sivaamnuaiphorn S, Xu M, Doulatov S, et al. Diversity-generating retroelement homing regenerates target sequences for repeated rounds of codon rewriting and protein diversification. Molecular Cell. 2008;**31**(6):813-823

[73] Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019;**576**(7785):149-157. DOI: 10.1038/s41586-019-1711-4

[74] Warren RAJ. Modified bases in bacteriophage DNAs. Annual Review of Microbiology. 1980;**34**(1):137-158

[75] Liu Q, Fang L, Yu G, Wang D, Xiao C-L, Wang K. Detection of DNA base modifications by deep recurrent neural network on Oxford nanopore sequencing data. Nature Communications. 2019;**10**(1):2449. DOI: 10.1038/s41467-019-10168-2

[76] Ehrlich M, Ehrlich KC. A novel, highly modified, bacteriophage DNA in which thymine is partly replaced by a phosphoglucuronate moiety covalently bound to 5-(4',5'-dihydroxypentyl) uracil. The Journal of Biological Chemistry. 1981;**256**(19):9966-9972

[77] Davies DR, Gelinas AD, Zhang C, Rohloff JC, Carter JD, O'Connell D, et al. Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. Proceedings of the National Academy of Sciences. 2012;**109**(49):19971-19976. DOI: 10.1073/pnas.1213933109

[78] Neubort S, Marmur J. Synthesis of the unusual DNA of *Bacillus subtilis*

bacteriophage SP 15. *Journal of Virology*. 1973;**12**(5):1078-1084

[79] Kirnos MD, Khudyakov IY, Alexandrushkina NI, Vanyushin BF. 2-Aminoadenine is an adenine substituting for a base in S-2L cyanophage DNA. *Nature*. 1977;**270**(5635):369-370

[80] Torres LL, Pinheiro VB. Xenobiotic nucleic acid (XNA) synthesis by Phi29 DNA polymerase. *Current Protocols in Chemical Biology*. 2018;**10**(2):e41. DOI: 10.1002/cpch.41

[81] Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, et al. Synthetic genetic polymers capable of heredity and evolution. *Science*. 2012;**336**(6079):341-344. Available from: <http://www.sciencemag.org/content/336/6079/341%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/22517858>

[82] Saier MH, Reddy BL. Holins in bacteria, eukaryotes, and archaea: Multifunctional xenologues with potential biotechnological and biomedical applications. *Journal of Bacteriology*. 2015;**197**(1):7-17. Available from: <http://jlb.asm.org/content/197/1/7.abstract>

[83] São-José C. Engineering of phage-derived lytic enzymes: Improving their potential as antimicrobials. *Antibiotics*. 2018;**7**(2):29

[84] Pires DP, Cleto S, Sillankorva S, Azeredo J, Lu TK. Genetically engineered phages: A review of advances over the last decade. *Microbiology and Molecular Biology Reviews*. 2016;**80**(3):523-543

[85] McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: Filamentous phage displaying antibody variable domains. *Nature*. 1990;**348**:552. DOI: 10.1038/348552a0

[86] Barbas CF, Kang AS, Lerner RA, Benkovic SJ. Assembly of combinatorial

antibody libraries on phage surfaces: The gene III site. *Proceedings of the National Academy of Sciences*. 1991;**88**(18):7978-7982. DOI: 10.1073/pnas.88.18.7978

[87] Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009;**460**(7257):894-898. DOI: 10.1038/nature08187