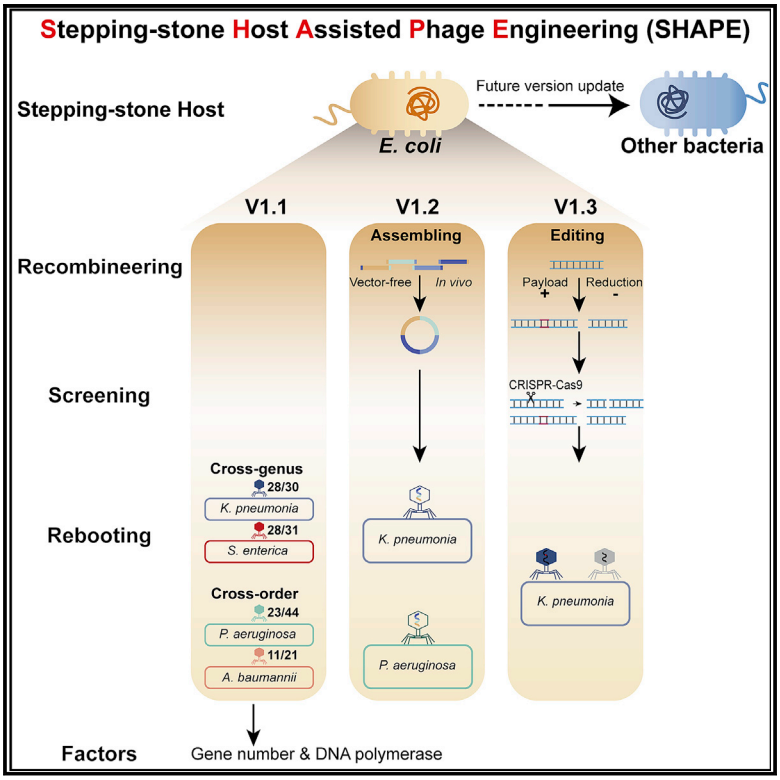
Cell Reports
Methods

Report

Harnessing stepping-stone hosts to engineer, select, and reboot synthetic bacteriophages in one pot

Graphical abstract



Highlights

- A stepping-stone host-assisted phage engineering framework is established
- Genome assembly, editing, and rebooting are achieved in a single cell type
- The stepping-stone supports cross-genus and cross-order phage rebooting
- Rebooting outcome is associated with phage gene number and polymerase availability

Authors

Li Cheng, Ziqing Deng, Haoran Tao, ..., Yue Shen, Junhua Li, Minfeng Xiao

Correspondence

xiaominfeng@genomics.cn

In brief

Cheng et al. develop a widely applicable framework, SHAPE, that utilizes user-friendly bacteria as stepping-stone hosts to complete multiplex manipulations of phage engineering in one pot. This framework makes tailored phages more approachable and has potential in phage therapy.

Cheng et al., 2022, Cell Reports Methods 2, 100217
May 23, 2022
https://doi.org/10.1016/j.crmeth.2022.100217

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Report

Harnessing stepping-stone hosts to engineer, select, and reboot synthetic bacteriophages in one pot

Li Cheng,[1,5,7,12] Ziqing Deng,[1,7,8,12] Haoran Tao,[1,6,7,13] Wenchen Song,[1,7,13] Bo Xing,[1,6,7] Wenfeng Liu,[1,7] Lingxin Kong,[1,6,7] Shengjian Yuan,[2] Yingfei Ma,[2] Yayun Wu,[3] Xun Huang,[3] Yun Peng,[4] Nai-Kei Wong,[4] Yingxia Liu,[4] Yun Wang,[1,9,10] Yue Shen,[1,9,11] Junhua Li,[1,5,7] and Minfeng Xiao[1,7,14,*]

[1]BGI-Shenzhen, Shenzhen 518083, China
[2]Shenzhen Key Laboratory of Synthetic Genomics, Guangdong Provincial Key Laboratory of Synthetic Genomics, CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[3]Infection Control Center, Xiangya Hospital, Central South University, Changsha 410008, Hunan, China
[4]Shenzhen Key Laboratory of Pathogen and Immunity, Guangdong Key Laboratory for Diagnosis and Treatment of Emerging Infectious Diseases, State Key Discipline of Infectious Disease, Second Hospital Affiliated to Southern University of Science and Technology, Shenzhen Third People's Hospital, 518112 Shenzhen, China
[5]School of Biology and Biological Engineering, South China University of Technology, Guangzhou 510006, China
[6]University of the Chinese Academy of Sciences, Beijing 101408, China
[7]Shenzhen Key Laboratory of Unknown Pathogen Identification, BGI-Shenzhen, Shenzhen 518083, China
[8]BGI-Beijing, Beijing 100101, China
[9]Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen 518120, China
[10]Guangdong Provincial Academician Workstation of BGI Synthetic Genomics, BGI-Shenzhen, Shenzhen 518120, China
[11]Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics, BGI-Shenzhen, Shenzhen 518120, China
[12]These authors contributed equally
[13]These authors contributed equally
[14]Lead contact
*Correspondence: xiaominfeng@genomics.cn
https://doi.org/10.1016/j.crmeth.2022.100217

MOTIVATION Synthetic genomics, a powerful approach to understand life and develop therapeutic agents, includes two crucial stages: genome synthesis and rebooting genome into life, both of which can be troublesome. For synthetic viruses, for example, engineering and rebooting are divided into separate steps and often involve multiple hosts or platforms, which is rather laborious and costly. Simplifying and integrating the two stages is a compelling approach to synthesize genomes easier and faster. We aimed to develop an alternative stepping-stone strategy to achieve genome refactoring of viruses in one pot in an efficient and economic manner, the stepping-stone host-assisted phage engineering (SHAPE) framework.

SUMMARY

Advances in synthetic genomics have led to a great demand for genetic manipulation. Trimming any process to simplify and accelerate streamlining of genetic code into life holds great promise for synthesizing and studying organisms. Here, we develop a simple but powerful stepping-stone strategy to promote genome refactoring of viruses in one pot, validated by successful cross-genus and cross-order rebooting of 90 phages infecting 4 orders of popular pathogens. Genomic sequencing suggests that rebooting outcome is associated with gene number and DNA polymerase availability within phage genomes. We integrate recombineering, screening, and rebooting processes in one pot and demonstrate genome assembly and genome editing of phages by stepping-stone hosts in an efficient and economic manner. Under this framework, *in vitro* assembly, yeast-based assembly, or genetic manipulation of native hosts are not required. As additional stepping-stone hosts are being developed, this framework will open doors for synthetic phages targeting more pathogens and commensals.

INTRODUCTION

Advances in synthetic biology are a boon for a deeper understanding of life and exploring novel therapeutic agents for diseases (Coradini et al., 2020; Elowitz and Lim, 2010). Efforts to synthesize genomes have spawned a wide range of synthetic species from viruses to bacteria and, more recently, complex eukaryotic chromosomes (Cello et al., 2002; Chan et al., 2005; Fredens et al., 2019; Gibson et al., 2010; Oldfield et al., 2017; Shen et al., 2017; Smith et al., 2003; van der Sloot and Tyers, 2017). Synthetic genomics usually require two crucial stages: genome synthesis and rebooting DNA into life (Coradini et al., 2020). *De novo* synthesis of genomes usually required a bottom-up approach involving cumbersome, expensive, and hierarchical assembly from smaller to larger fragments *in vitro* or *in vivo* (Baker, 2011; Chan et al., 2005; Fredens et al., 2019; Gibson et al., 2010; Oldfield et al., 2017; Shen et al., 2017; Smith et al., 2003; van der Sloot and Tyers, 2017). Previous studies have relied heavily on the yeast *Saccharomyces cerevisiae* to assemble chunks up to 10 kb or larger (Ando et al., 2015; Gibson et al., 2010; Oldfield et al., 2017; Shen et al., 2017; Thi Nhu Thao et al., 2020; van der Sloot and Tyers, 2017; Vashee et al., 2017). Bringing synthetic genomes into life is challenging. For synthetic viruses, for example, to reboot viral DNA into life, previous studies introduced the synthetic genome into natural hosts (Chan et al., 2005; Oldfield et al., 2017; Smith et al., 2003; Thi Nhu Thao et al., 2020; Vashee et al., 2017). However, this approach is hampered because many natural hosts are hard to manipulate. Simplifying the two stages is a compelling approach to synthesize genomes easier and faster.

Bacteriophages are fascinating organisms that play a key role in genetics and molecular biology and were crucial in establishing the central dogma of molecular biology because of their highly compact genomes and less complicated biological processes. More recently, engineered phages have emerged as versatile biological agents that efficiently detect and control multidrug-resistant (MDR) bacteria because of their advantages in tunable host range, killing efficiency, toxin expression, and so on (Ando et al., 2015; Citorik et al., 2014; Dedrick et al., 2019; Kilcher et al., 2018; Lemire et al., 2018; Lu and Collins, 2007, 2009; Yehl et al., 2019; Yosef et al., 2015). A variety of strategies have been proposed for phage engineering (Table S1; Kilcher and Loessner, 2019). These strategies fall into four categories: (1) genome editing and rebooting in native hosts, (2) genome assembly in yeast and rebooting in native or non-native hosts , (3) genome assembly *in vitro* and rebooting in native or cross-genus hosts , and (4) genome assembly *in vitro* and rebooting in a cell-free system. Representatives of strategy (1) are Bacteriophage Recombineering of Electroporated DNA (BRED)and BRED combined with CRISPR-Cas9 (CRISPY-BRED), which transform target DNA into native hosts by electroporation, accelerate DNA recombination and/or promote counterselection of recombinant phage genomes via plasmid-carried exogenous systems, and reboot engineered genomes in native hosts (Marinelli et al., 2008; Wetzel et al., 2021). Representative of strategy (2) is the yeast platform, which achieved genome rebooting of T7-family phages in *Escherichia coli* 10G (Ando et al., 2015; Latka et al., 2021). Representative of strategy (3) is the L-form bacteria platform, which successfully rebooted 9 *Listeria* phages, 2 *Bacillus* phages, and 2 *Staphylococcus aureus* phages in *Listeria* L-form cells (Kilcher et al., 2018; Meile et al., 2020). Representative of strategy (4) is the cell-free transcription-translation (TX-TL) system, which has been applied successfully to genome rebooting of 4 coliphages (MS2, phiX174, T7, and T4) (Garamella et al., 2016; Noireaux and Liu, 2020; Rustad et al., 2018; Shin et al., 2012). Strategy 1 has been used widely in phage engineering and applied successfully to host strains with a well-developed genetic manipulation system, but it is challenging when the host strain is MDR or biofilm productive. The applicable scales of strategies (2)–(4) remain to be fully evaluated because of the limited number of phages tested.

Here we take this a step further to expand the application of non-native hosts as inspired by strategies (2) and (3). Motivated by the term ''stepping-stone'' in evolutionary biology and virology, which is sometimes used to describe the notion of a virus taking advantage of an intermediate host to reach the final host, we adapted the term and upgraded the stepping-stone host-assisted strategy to a simple but powerful framework to promote virus synthesis in one pot. As a proof of concept, we build the first versions of stepping-stone hosts and tested cross-genus and cross-order rebooting of 126 T7/non-T7-family phages that originally infect common clinical MDR strains of *Klebsiella pneumoniae*, *Salmonella enterica*, *Pseudomonas aeruginosa*, and *Acinetobacter baumannii*. We also find underlying factors correlating with whether these phages are successfully rebooted by genomic sequencing of all 126 phages and analyzing ''genotype-phenotype'' association. Application of the stepping-stone host-assisted phage engineering (SHAPE) framework is highlighted by achieving phage engineering, including genome assembly and genome editing, and rebooting in one pot. SHAPE is a simple, efficient, and broadly applicable framework to build synthetic phages. This work provides insights into discovering more stepping-stone hosts to expand the application of SHAPE to the general field of synthetic genomics, where process improvement must be done to promote simple and fast streamlining of genetic code into life.

RESULTS

Construction and optimization of the stepping-stone host

First we tried an easy-to-manipulate strain, *E. coli* DH10B, as the first stepping-stone host. The laboratory *E. coli* K-12 strain DH10B is a MC1061 derivative, specifically designed for higher-efficiency cloning, and carries mutations that embrace large DNA uptake, enhance DNA stability, and protect foreign DNA from restriction systems (Durfee et al., 2008). To test the feasibility of the stepping-stone host DH10B, we first applied DH10B to reboot phage CPB0329, a *K. pneumoniae* phage isolated against clinical MDR strains. CPB0329 was chosen because it shares moderate homology with coliphage T7 (coverage, 65%; identity, 73.2%), and the T7 phage life

cycle is host independent (Qimron et al., 2010). As expected, the stepping-stone host DH10B successfully rebooted T7-family *K. pneumoniae* phage CPB0329 with high efficiency (Figure S1).

It has been reported that some viruses carry their own tRNAs to compensate for the deviation between host tRNA composition and virus codon use and to attain higher fitness (Bailly-Bechet et al., 2007; Delesalle et al., 2016). This led us to ask whether supplementary tRNAs could help phage rebooting in a non-native host. To test this, we optimized the stepping-stone host by introducing pRARE, a plasmid encoding six rare tRNAs and four common tRNAs in *E. coli* (Umlauf et al., 2015). The rebooting efficiency of *K. pneumoniae* phage CPB0329 was around one order of magnitude higher in DH10B/pRARE than in DH10B (Figure 1A). We observed that the rebooting efficiency decreased in DH10B and DH10B/pRARE after 24 h of recovery time immediately after transformation. It is likely that the optimal recovery time is affected by the balance between phage protein expression and degradation.

Cross-genus and cross-order rebooting of phages against MDR bacteria

To examine the scope of applying DH10B as a stepping-stone host, we rebooted 126 bacteriophages from the T7 family and non-T7 family, which infect clinical MDR *K. pneumoniae* (Kp), *S. enterica* (Se), *P. aeruginosa* (Pa), and *A. baumannii* (Ab) (Figure S2). Kp and Se phages represent cross-genus rebooting (*Escherichia* » *Klebsiella*, *Salmonella*), whereas Pa and Ab phages represent cross-order rebooting (Enterobacterales » Pseudomonadales), and all phages were unable to infect DH10B. Our results showed that the stepping-stone host DH10B successfully rebooted 93.3% (28 of 30) Kp phages, 90% (28 of 31) Se phages, 52.3% (23 of 44) Pa phages, and 52.4% (11 of 21) Ab phages (Figure 1B). Overall, the success rate of cross-genus rebooting is higher than that of cross-order rebooting. Among the successfully rebooted phages, 77.8% do not belong to the T7 family, which is an improvement over prior research (Ando et al., 2015; Latka et al., 2021). All 44 Pa phages were isolated against the same type strain, PAO1, but their rebooting outcome varied distinctly.

Only a handful of synthetic phages have been reported previously, and the generalization ability of the synthetic methods and the underlying factors were not clear. We explored the underlying factors affecting phage rebooting efficiency by genomic sequencing of all 126 tested phages and analyzing the association between genetic features and rebooting results (Figure 1C). To briefly summarize, the 126 tested phages are representative, covering Myoviridae, Siphoviridae, Podoviridae, a novel Ackermannviridae family, as well as an unassigned family (Figure 1C). Their genomes range from 20,741–299,545 bp (Figure 1C.), with a range of GC contents from 37.7%–64.6% (Figure S3B). Among all tested phages, 90 were successfully rebooted, including Myoviridae, Siphoviridae, Podoviridae, and an unassigned family (Figure S3A). Their genomes range from 20.7–156.8 kbp with 37.7%–64.6% GC content (Figures S3B and S3C). We analyzed the correlation between rebooting outcome and genes crucial for the phage life cycle, such as integrase, DNA polymerase, RNA polymerase, and tRNAs, via Fisher's exact test (Table S2). We found that tRNA and integrase (indicating a lytic or temperate lifestyle) show insignificant correlation with rebooting outcome (tRNA, $p = 0.09491$; integrase, $p = 0.1827$). The availability of DNA polymerase instead of RNA polymerase showed significant correlation with rebooting outcome (DNA polymerase, $p = 0.04362$; RNA polymerase, $p = 0.5257$). We examined the correlation between rebooting outcome and genomic features, including genome size, gene number, average protein length, transcriptional orientation, and transcriptional strand switch (Table S2). We found that the gene number of a phage shows negative correlation with its rebooting outcome ($p = 0.006151$, $r = -0.243$), suggesting that phages with fewer genes rather than smaller genomes are more likely to be rebooted successfully within the stepping-stone host. These results demonstrated that the stepping-stone host represents a highly versatile strategy accommodating cross-genus and even cross-order rebooting of phages. We showed that larger-scale experimentation could help to elucidate the underlying phage biology and drive the generalizability of the framework.

One-pot genome assembly and rebooting of synthetic phages

With the success of phage rebooting in the stepping-stone host, we explored the potential of the stepping-stone host in phage genome engineering (Figure 2). In general, synthetic biology includes two sub-fields: *de novo* synthesis of an organism and re-engineering an existing organism (Wang et al., 2018). We first tested *de novo* genome synthesis in the stepping-stone host. Putting together multiple chunks using existing methods can be time consuming, troublesome, and expensive (Baker, 2011). However, the capability of assembling multiple large fragments *in vivo*, especially into the full-length genome of viruses, by bacteria has not been tested.

To perform *in vivo* assembly, we utilized the plasmid pKD46 carrying the λ-Red recombination system, which has been validated extensively in engineering the *E. coli* genome (Datsenko and Wanner, 2000). We first evaluated its function in DH10B by assembling a 3.2-kb plasmid from two DNA fragments with or without arabinose induction of the λ-Red recombination system. We found that the assembly efficiency of DH10B without induction was 1,000-fold lower than that with induction (Figure S4). Then, to demonstrate phage genome assembly, we dissected Kp phage CPB0260 into 4 fragments of ~9.5 kb, sharing 40- to 60-bp overlaps with each other to facilitate homologous recombination. The results showed that the stepping-stone host carrying a functional recombination system supports assembly and rebooting of phage CPB0260 from four synthetic segments (Figure 3A). Considering that the assembly efficiency is dramatically affected by the number of DNA fragments and the size of DNA (Huang et al., 2017), we tested synthesizing Kp phage CPB0174, 45,798 bp in length, from 5, 10, 11, and 12 fragments. Assembling and rebooting CPB0174 from 5 and 10 fragments succeeded, but synthesizing the phage from 11 fragments and more failed, suggesting that the current settings of SHAPE could support *in vivo* assembly of phages with up to 10 DNA fragments (Figures 3B and 3D). To demonstrate the ability of synthesizing larger genomes, we also

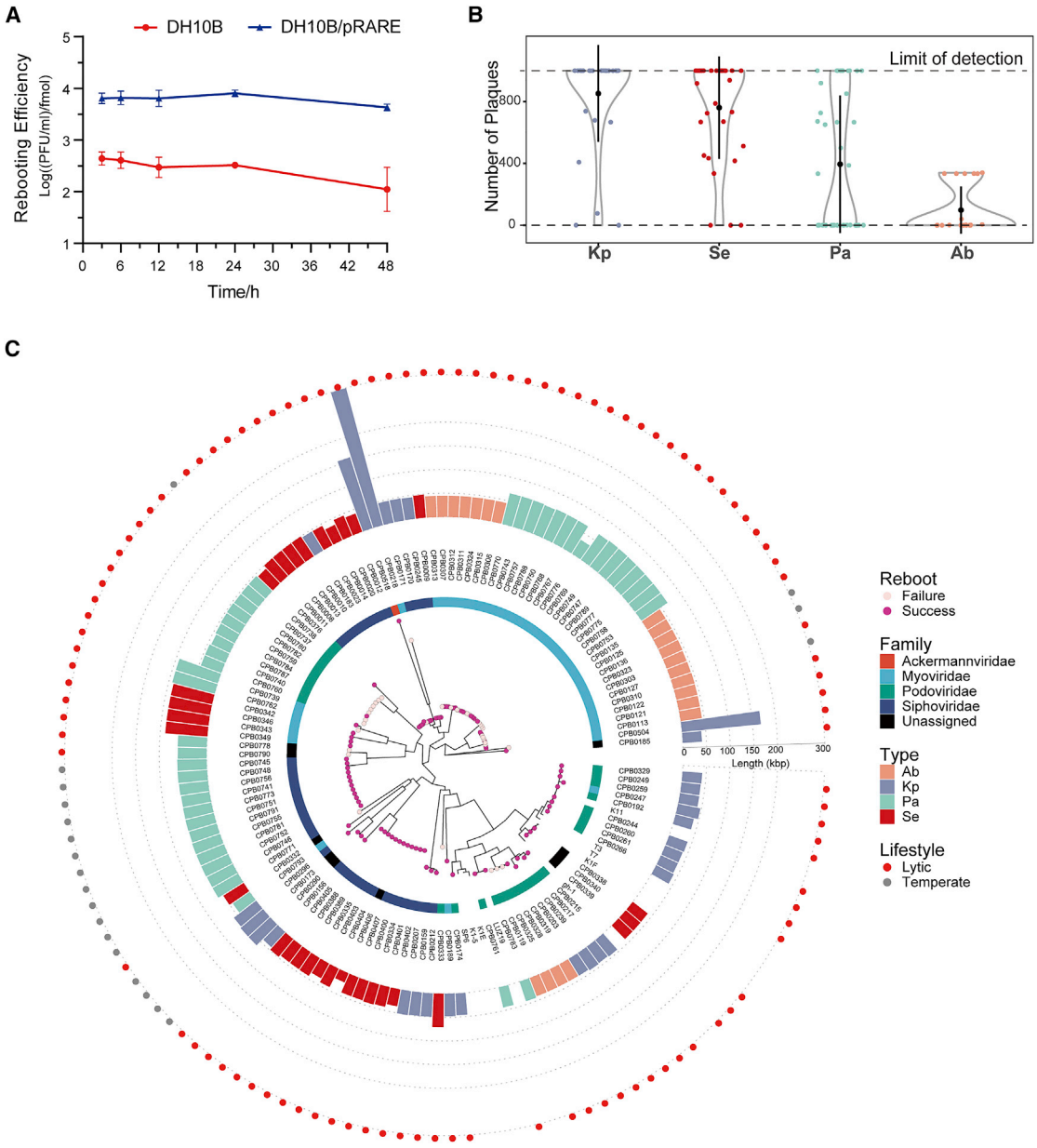


Figure 1. Cross-genus and cross-order phage rebooting using SHAPE

(A) Comparison of rebooting efficiency in stepping-stone host DH10B and DH10B/pRARE. The rebooted phage was assayed as plaque-forming units (PFUs) at the indicated time points. The data represent the mean of three independent experiments. Error bars represent the standard deviation.

(B) The rebooting efficiency of Kp, Se, Pa, and Ab phages in stepping-stone host DH10B/pRARE. Black dots represent the mean on each side of the violin plot. Colored dots represent the mean of three independent experiments. A negative control was performed by plating native host cells only. The detection limit of the reboot assay is less than 1,000 plaques.

(C) Phylogenetic tree based on the large terminase subunit of the 126 phages. The 9 T7-family phages tested previously were also included. The rebooting outcome is indicated by the color of tree tips (light salmon, failure; magenta, success). Phage taxonomy is indicated by the color of the inner circle. Genome size (height) and host species (color) are indicated by the bar chart in the middle circle. The lifestyle of phages is indicated by the color of the outer circle (red, lytic; grey, temperate). The tree was plotted using ggTree and ggTreeExtra in R.

See also Figures S1–S3 and Table S2.

tried to assemble and reboot a Pa phage of over 92,000 bp, CPB0739, and succeeded (Figure 3C). Compared with previously reported *in vitro* genome assembly or vector-based genome assembly in yeast, followed by a separate step of rebooting, the SHAPE framework simplifies the phage engineering procedures and reduces labor and experiments costs. It also indicates the

CellPress
OPEN ACCESS

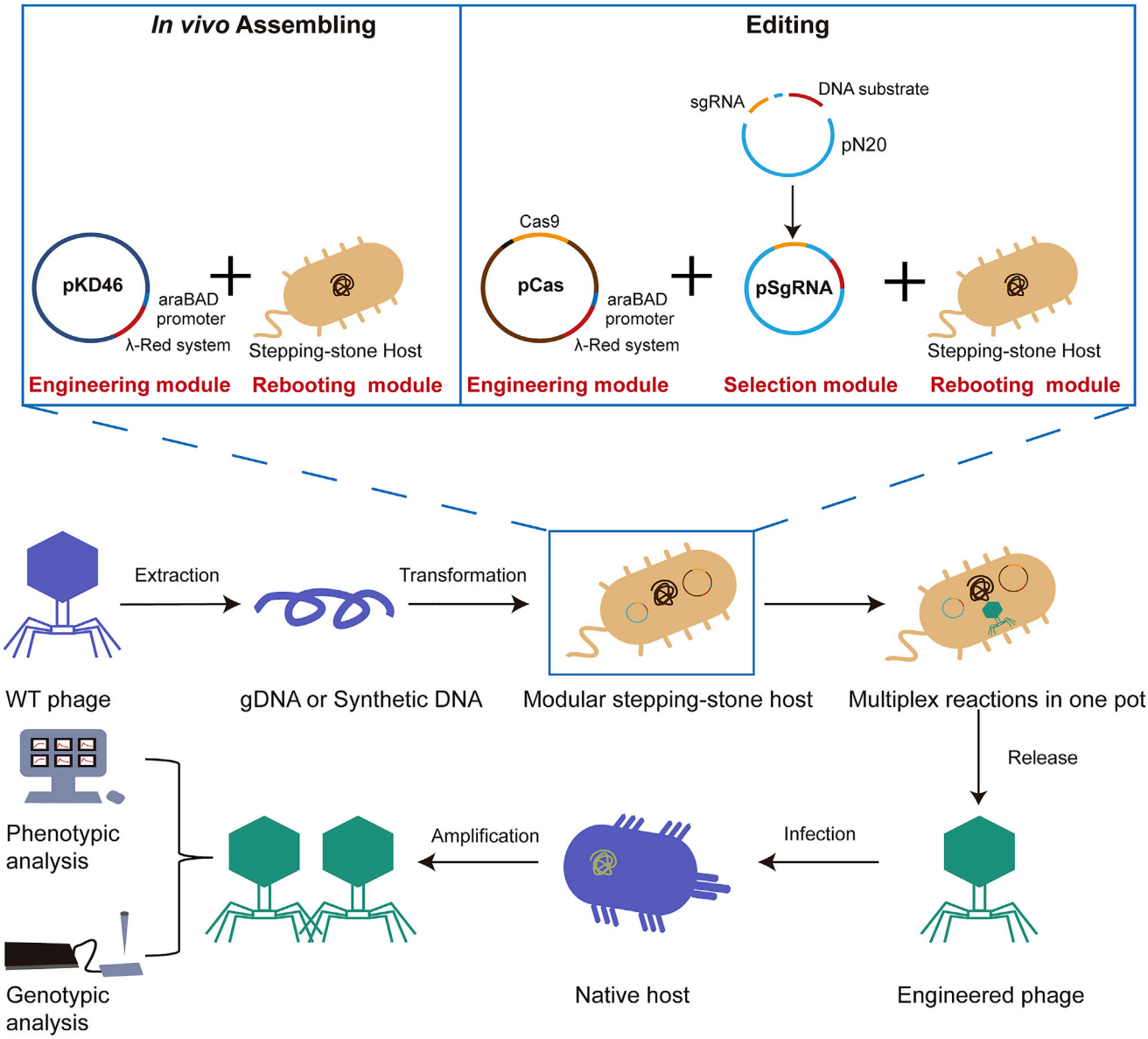


Figure 2. Schematic the workflow of SHAPE

SHAPE has two functions: *in vivo* assembly and DNA editing. For *in vivo* assembly, plasmid pKD46 is transformed into the stepping-stone host. The synthetic DNA fragments are transformed into the stepping-stone host harboring plasmid pKD46, and *de novo* synthetic phages are produced by the stepping-stone host and amplified on a lawn of a natural host. For DNA editing, sgRNA and DNA substrates are cloned in the pN20 vector. The resulting pSgRNA is co-transformed with pCas into the stepping-stone host. The phage genome is transformed into the stepping-stone host harboring the two plasmids, and engineered phages are produced by the stepping-stone host and amplified on a lawn of a natural host.

See also Figure S4 and Table S1.

great potential of using bacteria to assemble large genome fragments for other synthetic organisms.

Efficient phage engineering with genome payload and reduction

To perform genome editing of a T7-family Kp phage, CPB0329, we first tried to insert an exogenous DspB module (Lu and Collins, 2007; Figure 4A) into the phage genome mediated by the λ-Red recombination system alone. However, only wild-type phages were detected by PCR verification of more than 20 single clones (Table S3). Inspired by previous studies, we then utilized the CRISPR-Cas9 system for counterselection (Schilling et al., 2018; Shen et al., 2018; Wetzel et al., 2021). pCas carrying the CRISPR-Cas9 system under control of a constitutive promoter was co-transformed into the stepping-stone host with pSgRNA carrying the single guide RNA (sgRNA) and DNA substrate for recombineering. pSgRNA supports flexible cloning of new DNA substrates and sgRNA sequences.

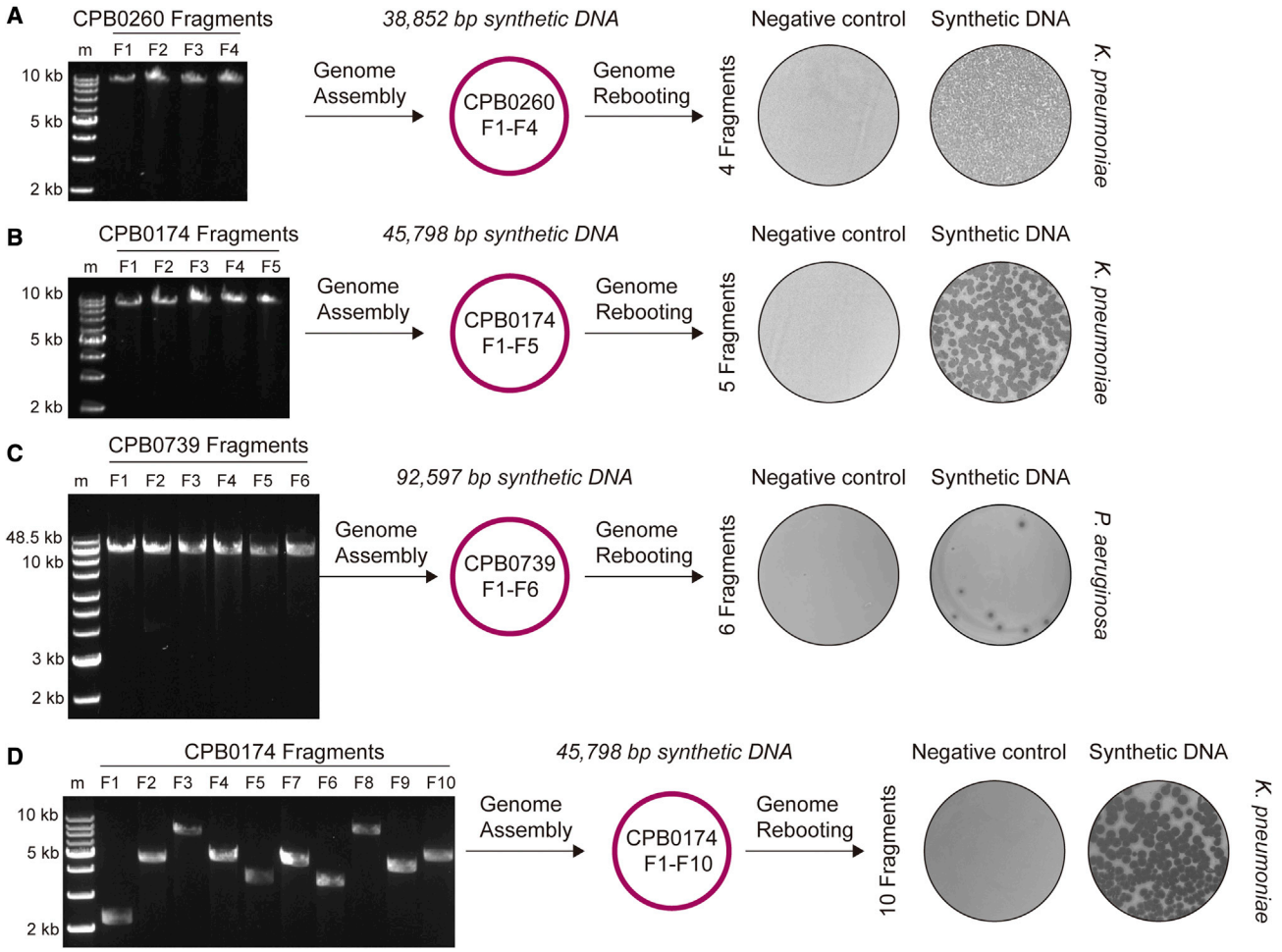


Figure 3. One-pot assembly and rebooting of phage genomes

(A–D) Phages with different genome sizes, synthesized from a series of number of chunks. Genome chunks were prepared by PCR and transformed into the stepping-stone host to simultaneously achieve *in vivo* assembly and phage genome rebooting. For visualization, the supernatants from rebooting reactions were mixed with native host Kp or Pa to perform double-layer plaque assays. Incomplete assemblies (n-1) were used as negative controls. Three technical replicates were included in each experiment.

See also Figure S4.

The SHAPE platform could accomplish three molecular reactions in one pot: recombination of designed DNA substrates, negative selection of native phage genomes, and rebooting of viral particles from engineered phage genomes (Figure 2).

With λ-Red recombination and CRISPR-Cas9 systems, we tried to engineer the genome of CPB0329 with a payload (Figure 4A) and reduction by knocking out a non-essential ligase gene (Masamune et al., 1971) (Figure 4B). To ensure successful counterselection, we designed different sgRNAs for independent experiments. Distinct engineering efficiencies were observed among these sgRNAs (Table S3; Figure 4E), suggesting that the design of sgRNAs can be crucial for counterselection efficiency by SHAPE. Two non T7-family Kp phages, CPB0170 (46,784 bp, Siphoviridae) and CPB0171 (46,784 bp, Siphoviridae), were also engineered. In this case, a DspB module was inserted downstream of a major capsid gene within the CPB0170 genome or between two hypothetical genes with opposite transcriptional directions within the CPB0171 genome (Figures 4C and 4D). Again, we observed 90%–100% and 60%–100% engineering efficiency in the two cases, respectively (Figure 4E.). In some cases, the engineering efficiency can be enhanced with an additional 15 h of recovery time after transformation (Table S3). These data illustrate the high efficiency of SHAPE in genome editing of T7- and non-T7-family phages.

DISCUSSION

We designed and implemented an easy and efficient framework: a stepping-stone strategy for synthetic phages. SHAPE incorporates a series of reactions in a single cell type and only requires the most widely used cloning techniques and laboratory conditions, which overcomes the limitation of native-host-based strategies (BRED or CRISPY-BRED) and simplifies the processes of

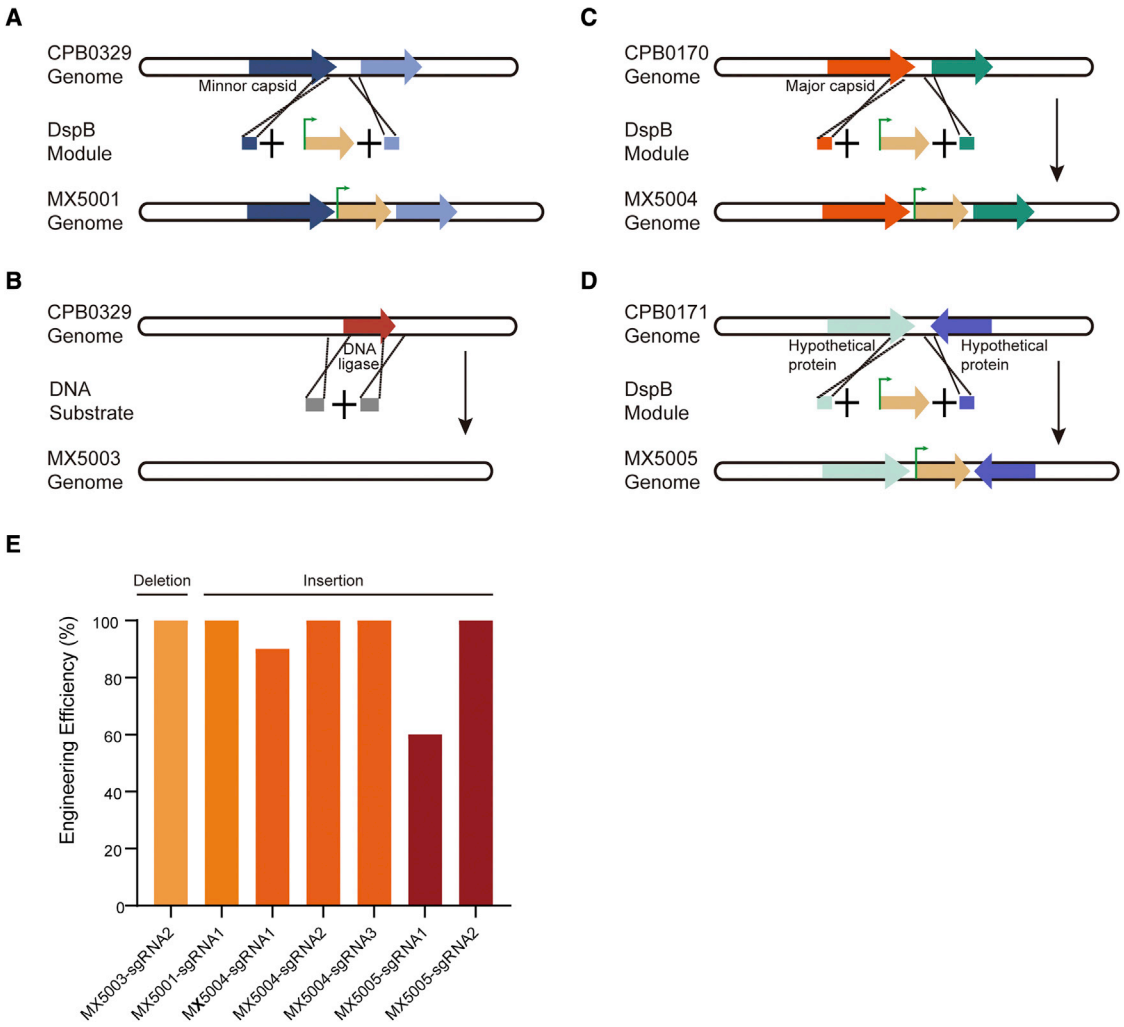


Figure 4. Efficient phage engineering with genome payload and reduction

(A–D) Schematics the designs of genome payload and reduction. The DspB module was inserted as a payload into CPB0329, CPB0170, and CPB0171 phage genomes to generate MX5001, MX5004, and MX5005, respectively; the DNA ligase gene was removed from the CPB0329 phage genome to generate MX5003.

(E) Efficiency of phage genome engineering with various genome payload/reduction designs and sgRNA designs. Engineering efficiency is defined as the number of positive clones validated by PCR proportional to the total number of phage clones (plaques) picked for validation. The engineering efficiency with sgRNAs from N.D. groups is not shown. N.D. means not determined, indicating that no plaques were observed in spot assays.

See also Table S3.

previous non-native-host-based strategies (yeast- or *in-vitro*-based platforms). SHAPE powerfully complements current approaches. For example, employing the technique of *in vitro* genome assembly and rebooting in L-form bacteria seems to be a wise choice for phages targeting Gram-positive bacteria, particularly *Listeria*, *Bacillus*, and *Staphylococcus*, as demonstrated by Kilcher et al. (2018). As for phages targeting Gram-negative bacteria, three approaches can be chosen. If the native host bacteria are easy to manipulate and compatible recombineering systems are available, a native-host-based strategy would be a fair option; if the native host bacteria are easy to manipulate but no compatible recombineering systems are available, yeast/*in vitro* assembly and native-host-based rebooting are applicable. In other cases, the benefits of SHAPE could be overwhelming, especially for engineering phages targeting *Klebsiella*, *Salmonella*, *Pseudomonas*, and *Acinetobacter* phages, as verified in this study.

With development more stepping-stone hosts, SHAPE could be applicable for phage engineering targeting more pathogens and commensals. An example is gut commensal bacteria, many of which play a key role in human chronic diseases such as inflammatory bowel disease (IBD), type II diabetes, and liver disease (Clemente et al., 2012). Most gut commensal bacteria are strictly anaerobic and require a special growth medium and environment. Massive expansion of gut bacteriophages identified *in silico* has created new capabilities to further investigate interactions among phages, gut bacteria, immunity, and disease (Benler et al., 2021; Camarillo-Guerrero et al., 2021; Devoto

et al., 2019; Gregory et al., 2020; Nayfach et al., 2021; Yutin et al., 2018). But efforts to isolate gut bacteriophages have proven difficult, and only a limited number of gut phages have been isolated to date (Guerin et al., 2018, 2021; Hryckowian et al., 2020; Porter et al., 2020). In these cases, SHAPE might be an interesting option for building synthetic phages to treat chronic gut microbiome-related diseases (Dong et al., 2020; Duan et al., 2019; Kabwe et al., 2021; Yuan et al., 2019; Zheng et al., 2019).

De novo genome assembly is one of the bottlenecks of genome writing, although the progress of genome writing technologies have been witnessed (Chari and Church, 2017; Farzadfard et al., 2021). Compared with the dramatical rise in throughput and drop in cost of genome reading (sequencing), efforts to improve genome writing technologies are still required (Chari and Church, 2017; Wang et al., 2018). At present, it is still laborious and costly to perform large-scale genome synthesis. The SHAPE framework integrates multiple steps in one pot and provides insights into reducing the cost and labor of genome writing. With further standardization of the framework and advances in automation (Hillson et al., 2019; Holowko et al., 2021), fewer modules and higher throughput can be realized in biofoundries to enable massively parallel construction of synthetic organisms.

Limitations of the study

The study presented version 1.0 of SHAPE, and there is room for updates of a few aspects. For instance, SHAPE's efficiency is associated with transformation efficiency, which reflects the stepping-stone host's ability to take up large phage genomes; therefore, *in vitro* circularization or spermidine treatment might be useful to improve the transformation efficiency of large DNA (Gosule and Schellman, 1976). Using a stepping-stone host with high transformation efficiency (e.g. Stellar-competent cells), could also be helpful. To some degree, the engineering efficiency of SHAPE is limited by the design of sgRNA. The recently developed near-Protospacer Adjacent Motif (PAM)-less CRISPR-Cas9 variants might be used in the next version of SHAE to avoid the NGG-PAM limitation of Cas9 targeting (Walton et al., 2020), which would provide higher flexibility in programming phages. Multiple DNA substrates and sgRNAs could be introduced in pSgRNA, enabling multi-site modifications in one step. Most importantly, diverse stepping-stone hosts must be validated to generalize application of the SHAPE framework in genome assembly, editing, and rebooting of random phages.

We did not demonstrate application of tailored phages for clinically relevant purposes in this study, but hopefully distinctive designs and applications of synthetic phages can be purred among the readers of this paper. The study also did not investigate whether the genetic factors correlating with phage rebooting are actually the cause, and this can be done through vigorous experimental verification, including exogenous expression of genes in the stepping-stone host cells. To better understand the mechanisms of phage rebooting, more phage-host pairs can be tested, especially in non-native hosts. We believe that gathering "genotypic-phenotypic" mapping data via SHAPE application, by us and by the research community working with a wide variety of bacteria and phages, will provide a better understanding of phage biology. This, in turn, will drive the evolution of SHAPE and could eventually lead to universal application of this strategy by selecting stepping-stone hosts with rationale.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Isolation and sequencing of phages
  - Genome rebooting of phages
  - *In vivo* assembly of phages
  - Genome editing of phages
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2022.100217.

ACKNOWLEDGMENTS

We are grateful for inspiring discussion with Prof. Lingchong You (Duke University, USA) and Dr. Franklin L. Nobrega (Delft University of Technology, the Netherlands). Genome sequencing and DNA synthesizing is supported by China National GeneBank and the Global Phage Hub project initiated by BGI-Shenzhen. This work is supported by the National Key R&D Program of China (SQ2018YFA090004-3 and 2020YFA0908700), the Hunan Health Commission Research Program (20201422), the Guangdong Provincial Key Laboratory of Genome Read and Write (2017B030301011), the Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics (DRC-SZ [2016]884), and the Guangdong Provincial Academician Workstation of BGI Synthetic Genomics (2017B090904014).

AUTHOR CONTRIBUTIONS

M.X. conceived the study. L.C. and M.X. designed the experiments. W.S., YayunW., X.H., S.Y., Y.M., Y.P., N.-K.W., and Y.L. isolated the bacteria and phages. L.C., Z.D., H.T., B.X., W.L., L.K., YunW., and Y.S. carried out all other experiments. L.C., W.S., B.X., and J.L. performed bioinformatics and computational analyses. L.C., Z.D., and M.X. interpreted the results and wrote the manuscript.

DECLARATION OF INTERESTS

The authors have filed a patent application.

Received: December 3, 2021
Revised: March 7, 2022
Accepted: April 25, 2022
Published: May 23, 2022

REFERENCES

Ando, H., Lemire, S., Pires, D.P., and Lu, T.K. (2015). Engineering modular viral scaffolds for targeted bacterial population editing. Cell Syst. *1*, 187–196. https://doi.org/10.1016/j.cels.2015.08.013.

Bailly-Bechet, M., Vergassola, M., and Rocha, E. (2007). Causes for the intriguing presence of tRNAs in phages. Genome Res. *17*, 1486–1495. https://doi.org/10.1101/gr.6649807.

Baker, M. (2011). The next step for the synthetic genome. Nature *473*, 403–408. https://doi.org/10.1038/473403a.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Benler, S., Yutin, N., Antipov, D., Rayko, M., Shmakov, S., Gussow, A.B., Pevzner, P., and Koonin, E.V. (2021). Thousands of previously unknown phages discovered in whole-community human gut metagenomes. Microbiome *9*, 78. https://doi.org/10.1186/s40168-021-01017-w.

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J.R., Kropinski, A.M., Krupovic, M., Lavigne, R., et al. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat. Biotechnol. *37*, 632–639. https://doi.org/10.1038/s41587-019-0100-8.

Camarillo-Guerrero, L.F., Almeida, A., Rangel-Pineros, G., Finn, R.D., and Lawley, T.D. (2021). Massive expansion of human gut bacteriophage diversity. Cell *184*, 1098–1109.e9. https://doi.org/10.1016/j.cell.2021.01.029.

Cello, J., Paul, A.V., and Wimmer, E. (2002). Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science *297*, 1016–1018. https://doi.org/10.1126/science.1072266.

Chan, L.Y., Kosuri, S., and Endy, D. (2005). Refactoring bacteriophage T7. Mol. Syst. Biol. *1*, 2005 0018. https://doi.org/10.1038/msb4100025.

Chan, Patricia P., Lin, Brian Y., Mak, Allysia J., and Lowe, Todd M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. *49*, 9077–9096. https://doi.org/10.1093/nar/gkab688.

Chari, R., and Church, G.M. (2017). Beyond editing to writing large genomes. Nat. Rev. Genet. *18*, 749–760. https://doi.org/10.1038/nrg.2017.59.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018a). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., et al. (2018b). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience *7*, 1–6. https://doi.org/10.1093/gigascience/gix120.

Citorik, R.J., Mimee, M., and Lu, T.K. (2014). Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat. Biotechnol. *32*, 1141–1145. https://doi.org/10.1038/nbt.3011.

Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. Cell *148*, 1258–1270. https://doi.org/10.1016/j.cell.2012.01.035.

Coradini, A.L.V., Hull, C.B., and Ehrenreich, I.M. (2020). Building genomes to understand biology. Nat. Commun. *11*, 6177. https://doi.org/10.1038/s41467-020-19753-2.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. *97*, 6640–6645. https://doi.org/10.1073/pnas.120163297.

Dedrick, R.M., Guerrero-Bustamante, C.A., Garlena, R.A., Russell, D.A., Ford, K., Harris, K., Gilmour, K.C., Soothill, J., Jacobs-Sera, D., Schooley, R.T., et al. (2019). Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. Nat. Med. *25*, 730–733. https://doi.org/10.1038/s41591-019-0437-z.

Delesalle, V.A., Tanke, N.T., Vill, A.C., and Krukonis, G.P. (2016). Testing hypotheses for the presence of tRNA genes in mycobacteriophage genomes. Bacteriophage *6*, e1219441. https://doi.org/10.1080/21597081.2016.1219441.

Devoto, A.E., Santini, J.M., Olm, M.R., Anantharaman, K., Munk, P., Tung, J., Archie, E.A., Turnbaugh, P.J., Seed, K.D., Blekhman, R., et al. (2019). Megaphages infect Prevotella and variants are widespread in gut microbiomes. Nat. Microbiol. *4*, 693–700. https://doi.org/10.1038/s41564-018-0338-9.

Dong, X., Pan, P., Zheng, D.W., Bao, P., Zeng, X., and Zhang, X.Z. (2020). Bioinorganic hybrid bacteriophage for modulation of intestinal microbiota to remodel tumor-immune microenvironment against colorectal cancer. Sci. Adv. *6*, eaba1590. https://doi.org/10.1126/sciadv.aba1590.

Duan, Y., Llorente, C., Lang, S., Brandl, K., Chu, H., Jiang, L., White, R.C., Clarke, T.H., Nguyen, K., Torralba, M., et al. (2019). Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease. Nature *575*, 505–511. https://doi.org/10.1038/s41586-019-1742-x.

Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., 3rd, Burland, V., Mau, B., Petrosino, J.F., Qin, X., Muzny, D.M., Ayele, M., et al. (2008). The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse. J. Bacteriol. *190*, 2597–2606. https://doi.org/10.1128/JB.01695-07.

Elowitz, M., and Lim, W.A. (2010). Build life to understand it. Nature *468*, 889–890. https://doi.org/10.1038/468889a.

Farzadfard, F., Gharaei, N., Citorik, R.J., and Lu, T.K. (2021). Efficient retroelement-mediated DNA writing in bacteria. Cell Syst. *12*, 860–872.e5. https://doi.org/10.1016/j.cels.2021.07.001.

Fredens, J., Wang, K., de la Torre, D., Funke, L.F.H., Robertson, W.E., Christova, Y., Chia, T., Schmied, W.H., Dunkelmann, D.L., Beranek, V., et al. (2019). Total synthesis of Escherichia coli with a recoded genome. Nature *569*, 514–518. https://doi.org/10.1038/s41586-019-1192-5.

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The all E. coli TX-TL toolbox 2.0: a platform for cell-free synthetic biology. ACS Synth. Biol. *5*, 344–355. https://doi.org/10.1021/acssynbio.5b00296.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., et al. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. Science *329*, 52–56. https://doi.org/10.1126/science.1190719.

Gosule, L.C., and Schellman, J.A. (1976). Compact form of DNA induced by spermidine. Nature *259*, 333–335. https://doi.org/10.1038/259333a0.

Gregory, A.C., Zablocki, O., Zayed, A.A., Howell, A., Bolduc, B., and Sullivan, M.B. (2020). The gut virome database reveals age-dependent patterns of virome diversity in the human gut. Cell Host Microbe. *28*, 724–740.e8. https://doi.org/10.1016/j.chom.2020.08.003.

Guerin, E., Shkoporov, A., Stockdale, S.R., Clooney, A.G., Ryan, F.J., Sutton, T.D.S., Draper, L.A., Gonzalez-Tortuero, E., Ross, R.P., and Hill, C. (2018). Biology and taxonomy of crAss-like bacteriophages, the most abundant virus in the human gut. Cell Host Microbe *24*, 653–664.e6. https://doi.org/10.1016/j.chom.2018.10.002.

Guerin, E., Shkoporov, A.N., Stockdale, S.R., Comas, J.C., Khokhlova, E.V., Clooney, A.G., Daly, K.M., Draper, L.A., Stephens, N., Scholz, D., et al. (2021). Isolation and characterisation of ΦcrAss002, a crAss-like phage from the human gut that infects Bacteroides xylanisolvens. Microbiome *9*, 89. https://doi.org/10.1186/s40168-021-01036-7.

Hillson, N., Caddick, M., Cai, Y., Carrasco, J.A., Chang, M.W., Curach, N.C., Bell, D.J., Feuvre, R.L., Friedman, D.C., Fu, X., et al. (2019). Author Correction: building a global alliance of biofoundries. Nat. Commun. *10*, 3132. https://doi.org/10.1038/s41467-019-10862-1.

Holowko, M.B., Frow, E.K., Reid, J.C., Rourke, M., and Vickers, C.E. (2021). Building a biofoundry. Synth. Biol. (Oxf) *6*, ysaa026. https://doi.org/10.1093/synbio/ysaa026.

Hryckowian, A.J., Merrill, B.D., Porter, N.T., Van Treuren, W., Nelson, E.J., Garlena, R.A., Russell, D.A., Martens, E.C., and Sonnenburg, J.L. (2020). Bacteroides thetaiotaomicron-infecting bacteriophage isolates inform sequence-based host range predictions. Cell Host Microbe *28*, 371–379.e5. https://doi.org/10.1016/j.chom.2020.06.011.

Huang, F., Spangler, J.R., and Huang, A.Y. (2017). *In vivo* cloning of up to 16 kb plasmids in E. coli is as simple as PCR. PLoS One *12*, e0183974. https://doi.org/10.1371/journal.pone.0183974.

Kabwe, M., Dashper, S., Bachrach, G., and Tucci, J. (2021). Bacteriophage manipulation of the microbiome associated with tumour microenvironments-can this improve cancer therapeutic response? FEMS Microbiol. Rev. *45*, fuab017. https://doi.org/10.1093/femsre/fuab017.

Kilcher, S., and Loessner, M.J. (2019). Engineering bacteriophages as versatile biologics. Trends Microbiol. *27*, 355–367. https://doi.org/10.1016/j.tim.2018.09.006.

Kilcher, S., Studer, P., Muessner, C., Klumpp, J., and Loessner, M.J. (2018). Cross-genus rebooting of custom-made, synthetic bacteriophage genomes in L-form bacteria. Proc. Natl. Acad. Sci. U S A. *115*, 567–572. https://doi.org/10.1073/pnas.1714658115.

Latka, A., Lemire, S., Grimon, D., Dams, D., Maciejewska, B., Lu, T., Drulis-Kawa, Z., and Briers, Y. (2021). Engineering the modular receptor-binding proteins of Klebsiella phages switches their capsule serotype specificity. mBio *12*, e00455-21. https://doi.org/10.1128/mBio.00455-21.

Lemire, S., Yehl, K.M., and Lu, T.K. (2018). Phage-based applications in synthetic biology. Annu. Rev. Virol. *5*, 453–476. https://doi.org/10.1146/annurev-virology-092917-043544.

Lu, T.K., and Collins, J.J. (2007). Dispersing biofilms with engineered enzymatic bacteriophage. Proc. Natl. Acad. Sci. U S A *104*, 11197–11202. https://doi.org/10.1073/pnas.0704624104.

Lu, T.K., and Collins, J.J. (2009). Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. Proc. Natl. Acad. Sci. U S A *106*, 4629–4634. https://doi.org/10.1073/pnas.0800442106.

Marinelli, L.J., Piuri, M., Swigonova, Z., Balachandran, A., Oldfield, L.M., van Kessel, J.C., and Hatfull, G.F. (2008). BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS One *3*, e3957. https://doi.org/10.1371/journal.pone.0003957.

Masamune, Y., Frenkel, G.D., and Richardson, C.C. (1971). A mutant of bacteriophage T7 deficient in polynucleotide ligase. J. Biol. Chem. *246*, 6874–6879. https://doi.org/10.1016/s0021-9258(19)45927-9.

Meile, S., Sarbach, A., Du, J., Schuppler, M., Saez, C., Loessner, M.J., and Kilcher, S. (2020). Engineered reporter phages for Rapid bioluminescence-based detection and differentiation of viable Listeria cells. Appl. Environ. Microbiol. *86*, e00442-20. https://doi.org/10.1128/AEM.00442-20.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Nayfach, S., Paez-Espino, D., Call, L., Low, S.J., Sberro, H., Ivanova, N.N., Proal, A.D., Fischbach, M.A., Bhatt, A.S., Hugenholtz, P., and Kyrpides, N.C. (2021). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nat. Microbiol. *6*, 960–970. https://doi.org/10.1038/s41564-021-00928-6.

Noireaux, V., and Liu, A.P. (2020). The new age of cell-free biology. Annu. Rev. Biomed. Eng. *22*, 51–77. https://doi.org/10.1146/annurev-bioeng-092019-111110.

Oldfield, L.M., Grzesik, P., Voorhies, A.A., Alperovich, N., MacMath, D., Najera, C.D., Chandra, D.S., Prasad, S., Noskov, V.N., Montague, M.G., et al. (2017). Genome-wide engineering of an infectious clone of herpes simplex virus type 1 using synthetic genomics assembly methods. Proc. Natl. Acad. Sci. U S A *114*, E8885–E8894. https://doi.org/10.1073/pnas.1700534114.

Porter, N.T., Hryckowian, A.J., Merrill, B.D., Fuentes, J.J., Gardner, J.O., Glowacki, R.W.P., Singh, S., Crawford, R.D., Snitkin, E.S., Sonnenburg, J.L., and Martens, E.C. (2020). Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. Nat. Microbiol. *5*, 1170–1181. https://doi.org/10.1038/s41564-020-0746-5.

Qimron, U., Tabor, S., and Richardson, C.C. (2010). New details about bacteriophage T7-host interactions. Microbe *5*, 117–122.

Rustad, M., Eastlund, A., Jardine, P., and Noireaux, V. (2018). Cell-free TXTL synthesis of infectious bacteriophage T4 in a single test tube reaction. Synth. Biol. (Oxf) *3*, ysy002. https://doi.org/10.1093/synbio/ysy002.

Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W., Goodale, A., Root, D.E., Piccioni, F., and Doench, J.G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat. Commun. *9*, 5416. https://doi.org/10.1038/s41467-018-07901-8.

Schilling, T., Dietrich, S., Hoppert, M., and Hertel, R. (2018). A CRISPR-cas9-based toolkit for fast and precise in vivo genetic engineering of Bacillus subtilis phages. Viruses *10*, 241. https://doi.org/10.3390/v10050241.

Shen, J., Zhou, J., Chen, G.Q., and Xiu, Z.L. (2018). Efficient genome engineering of a virulent Klebsiella bacteriophage using CRISPR-cas9. J. Virol. *92*, e00534–e00618. https://doi.org/10.1128/JVI.00534-18.

Shen, Y., Wang, Y., Chen, T., Gao, F., Gong, J., Abramczyk, D., Walker, R., Zhao, H., Chen, S., Liu, W., et al. (2017). Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. Science *355*, eaaf4791. https://doi.org/10.1126/science.aaf4791.

Shin, J., Jardine, P., and Noireaux, V. (2012). Genome replication, synthesis, and assembly of the bacteriophage T7 in a single cell-free reaction. ACS Synth. Biol. *1*, 408–413. https://doi.org/10.1021/sb300049p.

Smith, H.O., Hutchison, C.A., 3rd, Pfannkoch, C., and Venter, J.C. (2003). Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides. Proc. Natl. Acad. Sci. U S A *100*, 15440–15445. https://doi.org/10.1073/pnas.2237126100.

Thi Nhu Thao, T., Labroussaa, F., Ebert, N., V'Kovski, P., Stalder, H., Portmann, J., Kelly, J., Steiner, S., Holwerda, M., Kratzel, A., et al. (2020). Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature *582*, 561–565. https://doi.org/10.1038/s41586-020-2294-9.

Umlauf, B.J., McGuire, M.J., and Brown, K.C. (2015). Introduction of plasmid encoding for rare tRNAs reduces amplification bias in phage display biopanning. Biotechniques *58*, 81–84. https://doi.org/10.2144/000114256.

van der Sloot, A., and Tyers, M. (2017). Synthetic genomics: rewriting the genome chromosome by chromosome. Mol. Cell *66*, 441–443. https://doi.org/10.1016/j.molcel.2017.05.007.

Vashee, S., Stockwell, T.B., Alperovich, N., Denisova, E.A., Gibson, D.G., Cady, K.C., Miller, K., Kannan, K., Malouli, D., Crawford, L.B., et al. (2017). Cloning, assembly, and modification of the primary human cytomegalovirus isolate toledo by yeast-based transformation-associated recombination. mSphere *2*, e00331–e00417. https://doi.org/10.1128/mSphereDirect.00331-17.

Walton, R.T., Christie, K.A., Whittaker, M.N., and Kleinstiver, B.P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science *368*, 290–296. https://doi.org/10.1126/science.aba8853.

Wang, Y., Shen, Y., Gu, Y., Zhu, S., and Yin, Y. (2018). Genome writing: current progress and related applications. Genomics Proteomics Bioinf. *16*, 10–16. https://doi.org/10.1016/j.gpb.2018.02.001.

Wetzel, K.S., Guerrero-Bustamante, C.A., Dedrick, R.M., Ko, C.C., Freeman, K.G., Aull, H.G., Divens, A.M., Rock, J.M., Zack, K.M., and Hatfull, G.F. (2021). CRISPY-BRED and CRISPY-BRIP: efficient bacteriophage engineering. Sci. Rep. *11*, 6796. https://doi.org/10.1038/s41598-021-86112-6.

Xie, S., Shen, B., Zhang, C., Huang, X., and Zhang, Y. (2014). sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PLoS One *9*, e100448. https://doi.org/10.1371/journal.pone.0100448.

Xu, S., Dai, Z., Guo, P., Fu, X., Liu, S., Zhou, L., Tang, W., Feng, T., Chen, M., Zhan, L., et al. (2021). ggtreeExtra: compact visualization of richly annotated phylogenetic data. Mol. Biol. Evol. *38*, 4039–4042. https://doi.org/10.1093/molbev/msab166.

Yehl, K., Lemire, S., Yang, A.C., Ando, H., Mimee, M., Torres, M.D.T., de la Fuente-Nunez, C., and Lu, T.K. (2019). Engineering phage host-range and suppressing bacterial resistance through phage tail Fiber Mutagenesis. Cell *179*, 459–469.e9. https://doi.org/10.1016/j.cell.2019.09.015.

Yosef, I., Manor, M., Kiro, R., and Qimron, U. (2015). Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. Proc. Natl. Acad. Sci. U S A *112*, 7267–7272. https://doi.org/10.1073/pnas.1500107112.

Yuan, J., Chen, C., Cui, J., Lu, J., Yan, C., Wei, X., Zhao, X., Li, N., Li, S., Xue, G., et al. (2019). Fatty liver disease caused by high-alcohol-producing Klebsiella pneumoniae. Cell Metab. *30*, 1172–1688. https://doi.org/10.1016/j.cmet.2019.11.006.

Yutin, N., Makarova, K.S., Gussow, A.B., Krupovic, M., Segall, A., Edwards, R.A., and Koonin, E.V. (2018). Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. Nat. Microbiol. *3*, 38–46. https://doi.org/10.1038/s41564-017-0053-y.

Zheng, D.W., Dong, X., Pan, P., Chen, K.W., Fan, J.X., Cheng, S.X., and Zhang, X.Z. (2019). Phage-guided modulation of the gut microbiota of mouse models of colorectal cancer augments their responses to chemotherapy. Nat. Biomed. Eng. *3*, 717–728. https://doi.org/10.1038/s41551-019-0423-2.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli DH10B	Thermo Fisher Scientific	N/A
Escherichia coli DH10B/pSgRNA+pCas	This work	N/A
Escherichia coli DH10B/pRARE	This work	N/A
Escherichia coli DH10B/pRARE+pKD46	This work	N/A
See Table S2	This work	N/A
Chemicals, peptides, and recombinant proteins		
Q5 High-Fidelity 2X Master Mix	NEB	M0494S
DNase I	SIGMA	Cat # DN25
RNaseA	Invitrogen	Cat # 12091021
Proteinase K	NEB	P8107S
BsaI-HFv2	NEB	R3733S
BbsI-HF	NEB	R3539S
T4 DNA ligase	NEB	M0202S
L(+)-Arabinose	Sangon Biotech	Cat # A610071-0025
Agarose	BIO ROAD	Cat # 1613101
PEG8000	SIGMA	P5413-2KG
0.5M EDTA	INVITROGEN	Cat # AM9261
SDS Solution	INVITROGEN	Cat # AM9820
Critical commercial assays		
AxyPrep DNA Gel Extraction Kit	AXYGEN	AP-GX-250
Zymoclean Large Fragment DNA Recovery	ZYMO RESEARCH	Cat # 4045
TIANprep Rapid Mini Plasmid Kit	TIANGEN	Cat # 4992192
GENECLEAN Turbo Kit	MP Biomedicals	Cat # MP111102400
Qubit™ dsDNA HS Assay Kit	INVITROGEN	Q32854
Qubit™ dsDNA BR Assay Kit	INVITROGEN	Q32850
Oligonucleotides		
See Table S2	This work	N/A
DspB module	*De novo* synthesized	WP_005546617.1
Recombinant DNA		
pSgRNA	This work	N/A
MX5001	This work	N/A
MX5003	This work	N/A
MX5004	This work	N/A
MX5005	This work	N/A
Software and algorithms		
Snapgene v1.1.3	Snapgene Software	https://www.snapgene.com/
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/
CRISPick	Sanson et al., 2018	https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design
sgRNAcas9	Xie et al., 2014	N/A
Fastp	Chen et al., 2018a	N/A
SOAPnuke	Chen et al., 2018b	https://github.com/BGI-flexlab/SOAPnuke
SPAdes v3.13.0	Bankevich et al., 2012	N/A
tRNAscan-SE v2.0.9	Chan et al., 2021	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
vConTACT2	Bin Jang et al., 2019	N/A
IQ-TREE	Minh et al., 2020	N/A
ggtreeExtra	Xu et al., 2021	N/A
RStudio Version 4.1.1	RStudio, Inc., Boston, MA	N/A
GraphPad Prism v8.0.1	Graphpad Software	https://www.graphpad.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Minfeng Xiao (xiaominfeng@genomics.cn).

Materials availability

This study did not generate new unique reagents. Plasmids generated in this study are available from the lead contact with a completed material transfer agreement.

Data and code availability

- All data generated in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Kanamycin resistant and temperature sensitive plasmid pCas expresses Cas9 protein continuously and lambda Red recombineering proteins under the control of inducible arabinose promoter. Chloramphenicol resistant vector pSgRNA contains sgRNA and recombineering DNA substrates. Ampicillin resistant vector pKD46 only expressing lambda Red recombineering proteins under the control of inducible arabinose promoter has been described previously (Datsenko and Wanner, 2000). *E. coli* strain DH10B in this research was adopted as a stepping-stone host. DH10B/pRARE was constructed via transforming chloramphenicol resistant plasmid pRARE (Umlauf et al., 2015) that encodes six rare tRNAs into DH10B. DH10B harboring pCas was cultured in LB broth with 50 μg mL^{-1} kanamycin at 30°C, and DH10B harboring pN20 or pSgRNA was cultured in LB broth with 25 μg mL^{-1} chloramphenicol at 37°C. DH10B harboring pCas and pSgRNA was cultured in LB broth with 50 μg mL^{-1} kanamycin and 25 μg mL^{-1} chloramphenicol at 30°C. DH10B/pRARE harboring pKD46 was cultured in LB broth with 100 μg mL^{-1} ampicillin and 25 μg mL^{-1} chloramphenicol at 30°C. All other strains were cultured in LB broth at 37°C.

METHOD DETAILS

Isolation and sequencing of phages

Isolation of natural phages. Phages used in this study were isolated from sewage of 14 cities in China with a wide range of bacterial strains (Table S2). 50 mL of sewage water was centrifuged at 5,000 g for 20 min at 4°C to and the supernatant was passed through a 0.45 μm membrane filter (VWR). 40 mL of filtered sewage was co-incubated with 500 μL of overnight bacterial culture in a 250 mL flask containing 10 mL 5x LB broth, at 37°C, 220 rpm for overnight. 200 μL of the filter-sterilized culture was mixed with 100 μL of log-phase bacteria culture. Double-overlay agar assays were performed, and the plates were incubated at 37°C for overnight to obtain phage plaques. The phages were further purified, i.e. until all phage plaques are uniform, using double-overlay agar assays which usually takes 3-5 rounds of purification.

DNA extraction. Phage lysates were prepared by lysing 40 mL of logarithmically growing cells with the appropriate phage at a MOI of 0.1–0.01 and incubating the cultures until clearance. Lysates were centrifuged at 10,000 g for 10 min, sterilized with 0.45 μm membrane filters (VWR), incubated with 10 μg mL^{-1} DNase and RNase at 37°C for 1 h. Inactivate DNase and RNase at 65°C for 15 min. Lysates were treated with precipitate solution (10% PEG-8000,1 M NaCl final) at 4°C overnight. Phage particles were collected by spinning down the lysate at 10,000 g, 4°C for 20 min, suspended in 200 μL of TE buffer (0.5 M EDTA pH8, 0.1 M Tris·HCl pH7.4 final), and incubated with 0.5% (w/v) SDS and 10 μL of 20 mg mL^{-1} Proteinase K in 56°C for 2 h. Genomic DNA was extracted from the supernatants using GENECLEAN Turbo Kit according to the manufacturer's instructions (MP Biomedicals), and the concentration was determined by Qubit™ dsDNA BR Assay (Invitrogen).

Genome sequencing and assembly. Paired-end libraries with an insert size of 200–400 bp were constructed and sequenced on the MGISEQ-2000 (MGI, BGI-Shenzhen) platform to obtain about 1000x clean data (phage genome). Reads were filtered with SOAPnuke (Chen et al., 2018b) (https://github.com/BGI-flexlab/SOAPnuke) and fastp (Chen et al., 2018a), and clean reads were assembled with SPAdes v3.13.0(Bankevich et al., 2012). Annotation of phage genomes was conducted using prodigal, BLASTp searches against NCBI nr database (snapshot of 2019-07-17), and HMM searches against UniProt/Swiss-Prot database (snapshot of 2019-07-17). tRNAscan-SE (Chan et al., 2021) (version 2.0.5) was used to search for tRNA genes. Taxonomic classification of bacteriophages was performed by vConTACT2 (Bin Jang et al., 2019). Large terminase subunit proteins and major capsid proteins of phages were used to construct maximum-likelihood tree via IQ-TREE (Minh et al., 2020). The tree was visualized with ggtree and ggtreeExtra (Xu et al., 2021) in R.

Genome rebooting of phages

Preparation of competent cells. DH10B and DH10B/pRARE were grown in 3 mL of LB broth without or with 25 μg mL^{-1} chloramphenicol at 37°C overnight. 500 μL of overnight cultures were diluted 1:100 with 50 mL of fresh LB broth, and incubated at 37°C for ~3 h upon reaching absorbance OD_{600} = 0.5–0.6. Cells were harvested by centrifugation at 4,000 g at 4°C, washed with 12.5 mL of sterilized ice-cold 0.1 M $CaCl_2$ solution, and suspended in 500 μL of 0.1 M sterilized ice-cold $CaCl_2$ solution.

Rebooting of phages in stepping-stone hosts. Phage genomic DNA were gently mixed with 200 μL of competent cells, ice-bathed for 30 min, heat-shocked at 42°C for 2 min, and then ice-bathed for another 3 min. The transformants were recovered in 1 mL of pre-warmed LB broth at 37°C, 220 rpm for 4 h or over a period of 48 h to assess rebooting kinetics in DH10B and DH10B/pRARE. 5 mM Ca^{2+} was supplemented in the culture of *K. pneumoniae* phages, 5 mM Ca^{2+} and Mg^{2+} were supplemented in the culture of *S. enterica* phages, 5 mM Mg^{2+} was supplemented in the culture of *A. baumannii* phages or *P. aeruginosa* phages. 5% (v/v) chloroform was added to the culture and vortexed rigorously to lyse the cells and release phage particles. After centrifugation at 12,000 g for 5 min, 300 μL of supernatants were mixed with 200 μL of log-phase host bacteria and incubated at 37°C for 1–5 h. Specifically, *K. pneumoniae* or *S. enterica* phages were incubated for 1–3 h, while *A. baumannii* phages or *P. aeruginosa* phages were incubated for 4–5 h. The incubation step was skipped in order to measure the rebooting efficiency in Figure 1A. Double-overlay agar assays were performed, and the plates were incubated at 37°C or room temperature (~22°C) to obtain plaques. 5 mM Ca^{2+} was supplemented in the top-layer agar for *K. pneumoniae* phages, 5 mM Ca^{2+} and Mg^{2+} were supplemented in the top-layer agar for *S. enterica* phages, 5 mM Mg^{2+} was supplemented in the top-layer agar for *A. baumannii* phages or *P. aeruginosa* phages. Phage rebooting experiments were carried out at least three times. Negative control was included to avoid false reboot caused by contamination.

In vivo assembly of phages

In vivo assembly of synthetic genomes. Bacteriophage genomes were split into four to twelve DNA fragments with 40 bp to 300 bp overlaps. DH10B/pRARE plus pKD46 was used for *in vivo* assembling. DH10B/pRARE+pKD46 was inoculated in 3 mL of LB broth containing 100 μg mL^{-1} ampicillin and 25 μg mL^{-1} chloramphenicol at 30°C overnight. The overnight culture was diluted 1:100 with fresh LB broth containing 10 mM arabinose, 100 μg mL^{-1} ampicillin and 25 μg mL^{-1} and grown upon reaching absorbance OD_{600} = 0.5–0.6 at 30°C. Competent cells were then prepared. For *in vivo* assembling, DNA mixture of 0.06 pmol of each purified DNA fragment was transformed into 200 μL competent DH10B/pRARE+pKD46. The transformants were recovered in 1 mL of pre-warmed LB broth at 37°C,220 rpm for 6 h. After 5% chloroform treatment and centrifugation, 200 μL of supernatant were incubated with 200 μL of log-phase host bacteria at 37°C for 5 h. Double-overlay agar assays were performed, and the plates were incubated at 37°C to obtain plaques.

Genome editing of phages

Construction of pSgRNA plasmid. sgRNAs targeting the engineering site were designed using CRISPRko (Sanson et al., 2018) and sgRNAcas9 (Xie et al., 2014). sgRNA oligos were annealed to form double-strands with 4 nt sticky ends. DNA substrate for recombineering was constructed depending on specific engineering demand. 100 bp of homologous fragments upstream and downstream of the engineering site were amplified with primers carrying BbsI restriction site. For deletions, a sgRNA and two 100-bp homologous fragments were cloned into vector pN20 using golden gate assembly to generate the final homologous arms. To insert DspB module, a sgRNA, two 100-bp homologous fragments and synthetic DspB module were cloned into vector pN20 using golden gate assembly to generate DspB module flanking 100-bp homologous arms. The synthetic DspB module contained a copy of CPB0329 major capsid promoter and RBS, upstream of a gene encoding biofilm-dispersing enzyme - *dspB* from *Aggregatibacter actinomycetemcomitans* HK1651(ResSeq: NZ_CP007502.1). Plasmids were prepared with TIANprep Rapid Mini Plasmid Kit (TIANGEN).

Engineering of phages in stepping-stone hosts. pSgRNA and pCas were co-transformed into 200 μL of competent DH10B cells and recovered in 1 mL of LB broth at 30°C for 1.5 h. Transformants were selected on LB agar containing 50 μg mL^{-1} kanamycin and 25 μg mL^{-1} chloramphenicol. A positive transformant was inoculated in 3 mL of LB broth containing 50 μg mL^{-1} kanamycin and 25 μg mL^{-1} chloramphenicol at 30°C overnight. The overnight culture was diluted 1:100 with fresh LB broth containing 10 mM arabinose, 50 μg mL^{-1} kanamycin and 25 μg mL^{-1} chloramphenicol and grown upon reaching absorbance OD_{600} = 0.5–0.6, i.e., for 6 h at 30°C. Competent cells were then prepared, 1.5–3 μg of Kp phage genomic DNA was transformed into 200 μL of competent cells. The transformants were recovered in 1 mL of pre-warmed LB broth at 37°C, 220 rpm for 3 h or 37°C,

220 rpm for 3 h then place on the bench at room temperature (22°C) for 15 h. Spot test assays were performed, and the plates were incubated at 37°C for 4–16 h to obtain plaques. One to ten plaques were picked for screening using PCR, and positive plaques were then purified for three rounds for further PCR verification.

QUANTIFICATION AND STATISTICAL ANALYSIS

As shown in Figure 1A, rebooting efficiency was defined as phage plaque forming units per fmol of genomic DNA. All rebooting efficiencies are log10 transformed, presented as the mean, and the error bars represent the standard deviation from triplicate measures. As shown in Results "Cross-genus and cross-order rebooting of phages against MDR bacteria", correlation analysis of rebooting outcome (success or failure) and genome feature of 126 phages was performed by point-biserial in R. 5 parameters representing genome features were used in our analysis: (i) Transcription orientation, the ratio of the number of phage genes in the longest stretch of consecutive genes in the same direction to the total number of genes in phage genome, (ii) Average protein length, (iii) Gene numbers, (iv) Genome size, (v) Transcriptional strand switch, the ratio of the number of transcriptional strand switches to the total gene number in phage genomes. Fisher's exact test was performed to determine associations between rebooting outcome and DNA polymerase or RNA polymerase or Integrase or tRNA.