



OPEN ACCESS

EDITED BY

Mercedes Gonzalez Moreno,
Leibniz Institute for Natural Product Research
and Infection Biology, Hans Knoll Institute,
Germany

REVIEWED BY

Robert Ramirez-Garcia,
Imperial College London, United Kingdom

*CORRESPONDENCE

Steffanie A. Strathdee
✉ sstrathdee@health.ucsd.edu

RECEIVED 14 April 2025

ACCEPTED 30 April 2025

PUBLISHED 27 May 2025

CITATION

Doud MB, Robertson JM and Strathdee SA
(2025) Optimizing phage therapy with
artificial intelligence: a perspective.
Front. Cell. Infect. Microbiol. 15:1611857.
doi: 10.3389/fcimb.2025.1611857

COPYRIGHT

© 2025 Doud, Robertson and Strathdee. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Optimizing phage therapy with artificial intelligence: a perspective

Michael B. Doud¹, Jacob M. Robertson²
and Steffanie A. Strathdee^{1*}

¹Division of Infectious Diseases and Global Public Health, Department of Medicine, University of California, San Diego, San Diego, CA, United States, ²Department of Ecology, Behavior & Evolution, School of Biological Sciences, University of California, San Diego, La Jolla, CA, United States

Phage therapy is emerging as a promising strategy against the growing threat of antimicrobial resistance, yet phage and bacteria are incredibly diverse and idiosyncratic in their interactions with one another. Clinical applications of phage therapy often rely on a process of manually screening collections of naturally occurring phages for activity against a specific clinical isolate of bacteria, a labor-intensive task that is not guaranteed to yield a phage with optimal activity against a particular isolate. Herein, we review recent advances in artificial intelligence (AI) approaches that are advancing the study of phage-host interactions in ways that might enable the design of more effective phage therapeutics. In light of concurrent advances in synthetic biology enabling rapid genetic manipulation of phages, we envision how these AI-derived insights could inform the genetic optimization of the next generation of synthetic phages.

KEYWORDS

phage therapy, artificial intelligence, phage specificity, gene discovery, phage engineering, machine learning, synthetic biology

Introduction

The nascent fields of phage therapy, synthetic biology and artificial intelligence (AI) are coalescing at a time in history when antimicrobial resistance (AMR) is increasingly an urgent global health threat (Strathdee et al., 2023). Recent estimates indicate that over the next twenty-five years, 39 million people will die from a multi-drug resistant bacterial infection (Naghavi et al., 2024). Bacteriophage (phage) therapy is emerging as a promising tactic to confront this crisis. Despite the fact that phage were discovered over one hundred years ago and played prominent roles in launching the fields of molecular biology and genetic engineering, clinical applications of phage to treat acute bacterial infections were largely limited to parts of the former Soviet Union and Poland until the past decade, when a series of high-profile case reports ushered in a new era of phage therapy in the West (Strathdee et al., 2023).

In his 2020 commentary published in *Frontiers in Microbiology*, Belgian phage researcher Jean-Paul Pirnay re-imagined phage therapy in the year 2035 (Pirnay, 2020). His vision leverages advances in synthetic biology whereby phage could be generated *de novo* based on genetic sequences of bacterial host isolates without the need to ship bacterial cultures to laboratories for phage matching. In cases where bacterial isolates could not be obtained from patients (e.g., due to antibiotic suppressive therapy), he posited that AI algorithms could be applied to metagenomic data to predict the most likely bacterial sequence to facilitate phage matching. Further, he envisioned a global phage governance platform that would create an efficient, standardized, sustainable and ethical phage supply chain.

How close are we to achieving these realities? In this perspective, we review available literature on emerging advances in AI and synthetic biology that could be used to understand and engineer host specificity and other phage functions, considering the entire phage life cycle (i.e., binding and entry, replication, and lysis). We also consider how AI could be used to mine large datasets for accessory gene discovery and phage genome annotation. Finally, given the need to take phage therapy to scale, we discuss future avenues for research that could further advance the field.

Understanding phage-host specificity determinants using AI

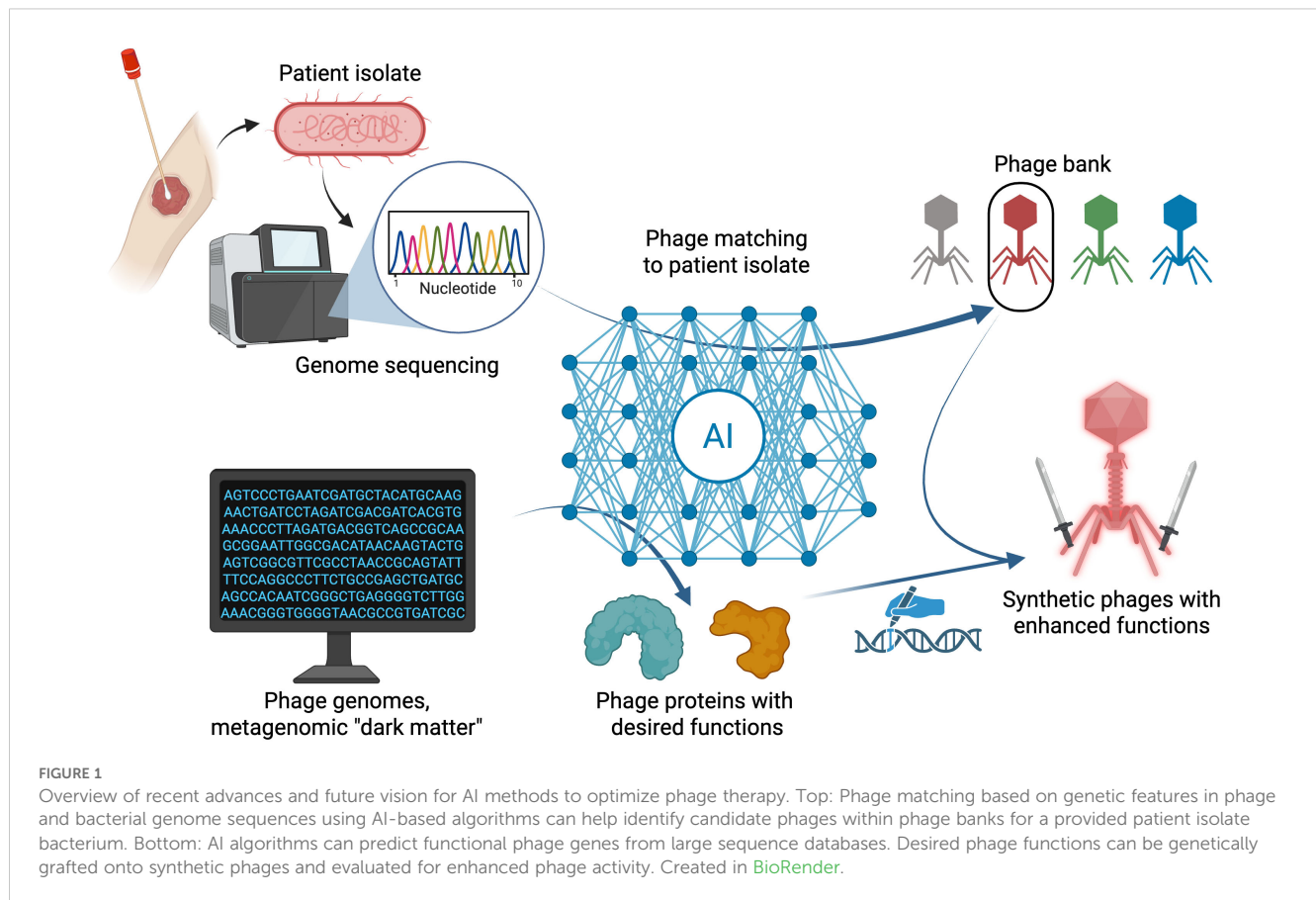
Identifying infectious phage strains for a given host is essential for phage therapy. However, it is logistically challenging to screen a panel of phage on each clinical isolate, especially when time is of the essence in the treatment of patients with multi-drug resistant bacterial infections. With rapid and inexpensive sequencing increasingly available, matching a potential phage to a target host based on bacterial whole genome or metagenomic data has the potential to accelerate these earliest stages of preparing a therapeutic phage. Recently, several groups have made promising use of AI to achieve strain-level prediction of phage infection from host genome sequences (Boeckaerts et al., 2024; Gaborieau et al., 2024).

Strain-level prediction of infectious phages for a given host genome has been reported for *Klebsiella* spp. (Boeckaerts et al., 2024) and *Escherichia* spp (Gaborieau et al., 2024), whereby shared aspects of these studies reflect the current state of the art. In both cases, machine learning algorithms were constructed using genotypic information as features and large phenotypic datasets (i.e., phage-bacteria infection networks, or PBINs) as outcomes in training data. Both studies made use of pre-existing tools to construct relevant features from genotypic information [e.g., Kaptive (Lam et al., 2022), ECtyper (Bessonov et al., 2021)]. The features predictive of infection were the attachment factors that phages of these genera tend to utilize: surface polysaccharide traits such as capsular K-serotype, lipopolysaccharide (LPS) outer core variations, or O-antigen serotypes. Impressively, both studies encompass genus-level diversity, yet can predict strain-level phage-host specificity.

Interestingly, outer membrane proteins are also frequent attachment factors for phages of *Escherichia* and other spp (Nobrega et al., 2018), yet they were not significantly associated with infection in the dataset analyzed by Gaborieau et al. (2024). Strain-specific amino-acid variation in phage receptor proteins has been shown to modulate phage infectivity (Suga et al., 2021), suggesting that future work may be necessary to develop AI-guided phage matching algorithms for outer membrane protein-targeting phages. One potential way to approach this problem is to use protein structural modeling to predict infection based on interactions between phage receptor variants and phage receptor-binding proteins (RBPs). However, this approach has not yet been evaluated to our knowledge and may be susceptible to false-positive results that accurately predict receptor-RBP binding for phage-host pairs that, for other reasons, do not result in productive phage infection. For instance, in an *in vitro* receptor binding experiment, the phage T4 RBP bound to 85% of the 72 strains in an *E. coli* reference collection, yet T4 phage only formed plaques on 11% of the collection (Farquharson et al., 2021). This indicates that, although receptor-RBD binding may be necessary for infection, it is not sufficient, suggesting that deeper understanding of phage-host interactions downstream from phage attachment may need to be incorporated to improve the accuracy of predictive algorithms.

Even with these recent advances in matching potential phages to target hosts, major gaps remain for strain-level matching in the context of phage therapy. The current AI models for phage matching from host genome sequences are highly specific to host genus, and no classifiers of this type are yet available for highly prevalent ESKAPE pathogens like *Staphylococcus aureus* or *Pseudomonas aeruginosa*. *S. aureus* exhibits surface polysaccharide diversity, with phage predominantly targeting wall teichoic acid (Krusche et al., 2025), so a machine learning approach using teichoic acid variations as features, analogous to using capsular types for *Klebsiella* (Boeckaerts et al., 2024), might be a promising approach. However, for *P. aeruginosa*, phage interactions may involve a larger diversity of receptor types, with resistance mutations in flagella, type IV pili, and LPS evolving against a single phage strain (Kortright et al., 2022). Moreover, some strains of *P. aeruginosa* harbor extensive defense systems (Costa et al., 2024), suggesting these may play a greater role in predicting phage effectiveness from host genomes than that observed in *E. coli* or *Klebsiella*. The degree to which the presence of defense systems is predictive of infection by *Pseudomonas* phages is an area of ongoing work (Müller et al., 2024). With these differences in mind, a holistic approach evaluating the importance of different *P. aeruginosa* genomic features across many phage host pairs (akin to what was undertaken for *E. coli* (Gaborieau et al., 2024) may be required to build a reliable model for this species. Nonetheless, there appears to be great potential for researchers to extend recent examples by leveraging AI to achieve effective strain-level phage matching models in other pathogenic bacteria (Figure 1).

Since the above models for phage selection are largely based on host-phage genotype matching, these models would likely struggle to predict the effects of novel resistance mutations outside the training data. Training data are often comprised of a variety of



strains of bacteria and phages representing a coarse sampling of genetic variation, but fine genetic variation (such as point mutations) arising during phage-bacteria coevolution can modulate infection outcome. Bacterial evolution of resistance to phage is frequently observed in clinical phage therapy settings (Schooley et al., 2017; Pirnay et al., 2024) and in the laboratory (Luria and Delbrück, 1943; Meyer et al., 2012). Experimental coevolution of bacteria and phages has revealed not only phages' ability to evolve counter-resistance, but the potential for long-term evolutionary and ecological conflict between phage and bacteria (Borin et al., 2021, 2023; Shaer Tamar and Kishony, 2022; Chen et al., 2024). By experimentally identifying how to position phage with an early evolutionary advantage in the microbial arms race, coevolutionary phage training has the potential to improve the effectiveness of phage therapy (Borin et al., 2021). Recent advances have used machine learning (e.g., L1-penalized logistic regression) to predict the outcome of fine genetic variation in mutation profiles within coevolutionary PBINs assembled through experimental coevolution (Shaer Tamar and Kishony, 2022; Lucia-Sanz et al., 2024). A future challenge will be to combine insights gained through these analyses of fine-scaled genetic variation in laboratory coevolution with predictive models trained on more coarse genetic variation at the strain level observed across natural isolates. Furthermore, while most of the effort in this space has focused on understanding phage-bacteria interactions at the

attachment step, which appears to provide the strongest predictive power for infectivity (Boeckaerts et al., 2024; Gaborieau et al., 2024), future work should identify how phage-bacteria interactions downstream of attachment modulate phage activity in ways that could improve efficacy of phage therapy.

Recent AI advances in the discovery of phage genes with specific functions

There has been an explosion in the number and variety of phage sequences available in public databases in recent years. This has been fueled by both the increase in the number of sequencing studies (often metagenome sequencing) as well as AI-driven improvements in identifying phage sequences in the "dark matter" of metagenome data (Hatfull, 2015). The applications of AI in discriminating phage from non-phage sequences in metagenomic datasets have been extensively reviewed (Nami et al., 2021; Flamholz et al., 2024). There is a wide collection of AI-driven tools for identifying phage genetic sequences (McNair et al., 2012; Auslander et al., 2020; Wu et al., 2021; Johansen et al., 2022; Shang et al., 2023) and classifying phage virion proteins (Thung et al., 2021; Ahmad et al., 2022; Fang et al., 2022). These tools are accelerating the annotation of predicted phage sequences,

allowing for a greater diversity of phage sequences to be used for comparative studies.

Ultimately, phage therapy requires cultivated phages, but the discovery of novel functional phage genes in sequence databases can be leveraged to program naturally occurring phages with specific biological functions. A key impediment to harnessing the huge and growing amount of phage sequencing data is that the number of potential genes of interest is vast, making it experimentally intractable to functionally screen sequences for biological function. Several recent studies (Concha-Eloko et al., 2024; Zhang et al., 2024; Yirmiyya et al., 2025) exemplify recent advances in leveraging AI to sift through large sequencing datasets to predict putative phage genes with specific functions, triaging labor-intensive experimental validation for predicted candidate genes. Once validated, these novel gene sequences can be tested for their ability enhance the efficacy of engineered therapeutic phages (Figure 1).

Anti-phage defense systems are diverse and heterogeneously distributed throughout bacteria (Doron et al., 2018; Bernheim and Sorek, 2020; Georjon and Bernheim, 2023), and there is a growing collection of phage genes that have been identified as counter-defenses against these bacterial immune systems (Vassallo et al., 2022; Mayo-Muñoz et al., 2024; Murtazaliev et al., 2024; Yirmiyya et al., 2024). It is likely that many phage counter-defenses are yet to be discovered. Yirmiyya et al. (Yirmiyya et al., 2025) used protein structure and interaction prediction using AlphaFold2-Multimer (Evans et al., 2022) to screen approximately two million phage genomes containing over 30 million phage genes, and identify phage proteins predicted to fold and interact with pre-chosen bacterial phage defense proteins. By carefully and iteratively designing a computational workflow, they were able to attain a ~50% success rate in identifying experimentally validated novel phage inhibitors of several well-characterized bacterial defense systems. Although there are limitations to this approach – including the need to select appropriate protein-binding partners such as a bacterial defense system of interest *a priori*, and a substantial false-positive rate – this approach more generally establishes a paradigm that leverages AI to predict novel phage genes that interact with specific proteins of interest. As additional bacterial anti-phage defense systems are functionally and structurally characterized, this approach will likely uncover additional novel phage counter-defense genes that may find utility in synthetic phages armed to match the capabilities of the target bacteria they are deployed against.

Bacterial capsules and biofilms are virulence factors that pose challenges in the treatment of bacterial infections (Chang et al., 2022). Some phages rely upon recognition and digestion of polysaccharide components in bacterial capsules and biofilms as their first step in the infection process (Knecht et al., 2020). The use of phage as a strategy for overcoming biofilms in difficult-to-treat infections has long been proposed and is recognized to require very specific interactions between phage and host (Hughes et al., 1998; Mayorga-Ramos et al., 2024). Some phage genes necessary for biofilm and capsule degradation have been identified as

depolymerases that can attach to the distal tips of phage tails where they simultaneously act as enzymes that degrade polysaccharide components and as specific receptor-binding proteins for which presence of the cognate capsule is required for infection (Dunstan et al., 2021). Mirroring the genetic and antigenic diversity of capsular polysaccharide serotypes in pathogenic bacteria (Shu et al., 2009), phage depolymerase sequences are also quite diverse in amino-acid sequences (Knecht et al., 2020) and this high degree of sequence variability complicates the process of identifying novel depolymerases in sequence databases. To overcome these difficulties, Concha-Eloko et al. (2024) demonstrate how a protein language model fine-tuned for depolymerases and trained on carefully curated data has advanced the annotation of depolymerase genes and their respective enzymatic domains, beyond currently available computational tools. The improved AI-guided annotation of depolymerase genes enables further study of the use of diverse depolymerase genes in recombinant phages to reprogram specificity and enzymatic capabilities for targeted therapeutic applications against specific capsule types.

Phage lysins are enzymes that degrade peptidoglycans, playing an essential role in the phage life cycle by promoting host cell lysis and cell death. There has long been interest in developing recombinant lysins as treatments for bacterial infections since they can also lyse the cell from the outside (Fischetti, 2018). A recent Phase 3 clinical trial of a lysin targeting *Staphylococcus aureus* added to standard of care antibiotics was ended for futility after an interim efficacy analysis (ClinicalTrials.gov ID NCT04160468). However, there is potential to advance their use, using engineered lysins selected from combinatorial libraries recombining portions of known lysin sequences (Gerstmans et al., 2020), with the potential to develop novel synthetic lysins with fine-tuned specificity and activity. Further advances in lysin engineering – whether for use as therapeutic protein products or as genetic cargo in engineered therapeutic phages – has been limited by the lack of computational methods to comprehensively screen metagenomic or uncharacterized phage genome sequence data to identify new lysin genes. Recent work by Zhang et al. provides a machine learning based software package that identifies putative lysin genes from assembled contigs (Zhang et al., 2024). Among 17 predicted novel lysin sequences selected for experimental validation, seven exhibited appropriate activity. Similar to the approach used by Yirmiyya et al. (2025), there is a substantial false positive rate requiring rigorous experimental validation of AI-produced screening candidates, however, these are substantial advances in that they allow researchers to triage valuable time and resources validating candidate genes selected from otherwise intractably large datasets.

A key theme emerging from each of these studies is that an enormous amount of careful human planning, intuition of biological plausibility, and iterative human-driven improvement to AI algorithms is necessary to realize the potential of these approaches. By facilitating a computational screening process for specific biological functions, these emerging AI models are enabling

researchers to exploit vast troves of data to discover a diversity of new phage genes that can antagonize bacterial defense systems, degrade biofilm and lyse infected cells, each of which may be useful in the future design of therapeutic phages with desired functions.

Future outlook: AI-guided development of synthetic phages as enhanced therapeutics

The AI-driven advances described above are beginning to generate tools that can predict which naturally occurring phages are most likely to infect a target bacterium. A deeper level of understanding of which genetic determinants drive these predictions, and the underlying mechanisms behind productive infection, are beginning to emerge to enable phage specificity programming (Dunne et al., 2021). Synthetic biology methods are already available to modify many phage genomes (Jaschke et al., 2012; Ando et al., 2015; Kilcher et al., 2018; Pires et al., 2021; Adler et al., 2022; Assad-Garcia et al., 2022; Kamata et al., 2024) and have begun to be applied to study granular determinants of phage specificity through high-throughput mutational studies of phage receptor binding proteins (Dunne et al., 2019; Yehl et al., 2019; Andrews and Fields, 2021; Huss et al., 2021, 2023). Huss et al. have recently developed a method of analyzing deep mutational scanning data of a phage receptor binding protein to develop a motif-searching algorithm that identifies novel phage receptor-binding sequences from metagenomic data (Huss et al., 2024). Collections of new receptor-binding protein sequences from these and other studies can be used as the substrate for future AI-guided protein design, leveraging generative models of protein sequences (Hsu et al., 2024). In a manner analogous to using protein language models to accelerate directed evolution of antibody sequences targeting specific antigens (Hie et al., 2023), phage receptor-binding protein engineering may also be amenable to machine-learning-guided directed evolution to modulate receptor specificity. While such fine-tuning of phage receptor binding through protein design has the potential to generate phages with defined bacterial receptor targets, it is important to note that binding affinity alone is not always sufficient to confer productive infection (Farquharson et al., 2021), and more work is needed to understand the mechanisms of infection immediately downstream of receptor binding, which are incompletely understood for even some of the most heavily studied model phages (Hu et al., 2015; Ge and Wang, 2024). More broadly, generative models of entire genomes, including phage genomes, have recently been described (Nguyen et al., 2024; Shao and Yan, 2024), and although these models are in their infancy and do not yet produce biologically functional whole genomes, they are already able to recapitulate coarse genomic architectures similar to natural phage genomes and can even produce gene sequences for functionally active multicomponent systems (Nguyen et al., 2024).

The recent AI-guided advances outlined above identifying novel phage lysins, depolymerases, and counter-defenses to bacterial immune systems similarly lay the groundwork for incorporating these functions into designed, synthetic phages (Lenneman et al., 2021) or other therapeutics, expanding the armamentarium of engineered phages that can deliver heterologous effector proteins (Du et al., 2023; Gencay et al., 2023) or augment natural phage function in other ways favorable for therapy. Phages engineered to avoid lysogeny have already been used clinically (Dedrick et al., 2019). Future work will be necessary to identify whether the rational design of synthetic phages with other various functions can increase treatment efficacy. Additionally, the bioethical and environmental implications of treating patients with genetically engineered phages requires continued careful contemplation from a One Health perspective, since engineered phages have the potential to impact human-, animal-, and environment-associated microbial communities (Hernando-Amado et al., 2019; Nair and Khairnar, 2019; Banerjee and van der Heijden, 2023). Although the dream of instant AI-designed therapeutic phage synthesis for a provided target bacterium is unlikely to be achieved in the next 10 years (Pirnay, 2020), advances in AI are both accelerating the identification of naturally occurring phages for therapy, as well as enabling the distillation of useful knowledge from otherwise untenably large sequencing databases abundant with uncharacterized phage genes that could find utility in synthetic phages. In the meantime, coordinated efforts are needed to make the growing number of phage libraries across the world compatible with one another, and accessible for compassionate use cases, clinical trials, and translational research experiments.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

MD: Conceptualization, Writing – original draft, Writing – review & editing. JR: Writing – original draft, Writing – review & editing. SS: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. MD was supported by the National Institute of Allergy and Infectious Diseases (NIAID) T32AI007036. JR was supported by Howard Hughes Medical

Institute Emerging Pathogens Initiative grant 7012574. SAS acknowledges support from the Mallory Smith Legacy Fund and the Herbert W. Hoover Foundation.

Acknowledgments

We are grateful to members of the Meyer laboratory for fruitful discussions and to Dr. Jean-Paul Pirnay for inspiration.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Adler, B. A., Hessler, T., Cress, B. F., Lahiri, A., Mutalik, V. K., Barrangou, R., et al. (2022). Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing. *Nat. Microbiol.* 7, 1967–1979. doi: 10.1038/s41564-022-01258-x
- Ahmad, S., Charoenkwan, P., Quinn, J. M. W., Moni, M. A., Hasan, M. M., Lio, P., et al. (2022). SCORPION is a stacking-based ensemble learning framework for accurate prediction of phage virion proteins. *Sci. Rep.* 12, 4106. doi: 10.1038/s41598-022-08173-5
- 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. doi: 10.1016/j.cels.2015.08.013
- Andrews, B., and Fields, S. (2021). Balance between promiscuity and specificity in phage λ host range. *ISME J.* 15, 2195–2205. doi: 10.1038/s41396-021-00912-2
- Assad-Garcia, N., D'Souza, R., Buzzeo, R., Tripathi, A., Oldfield, L. M., Vashee, S., et al. (2022). Cross-genus “Boot-up” of synthetic bacteriophage in *staphylococcus aureus* by using a new and efficient DNA transformation method. *Appl. Environ. Microbiol.* 88, e0148621. doi: 10.1128/AEM.01486-21
- Auslander, N., Gussow, A. B., Benler, S., Wolf, Y. I., and Koonin, E. V. (2020). Seeker: alignment-free identification of bacteriophage genomes by deep learning. *Nucleic Acids Res.* 48, e121. doi: 10.1093/nar/gkaa856
- Banerjee, S., and van der Heijden, M. G. A. (2023). Soil microbiomes and one health. *Nat. Rev. Microbiol.* 21, 6–20. doi: 10.1038/s41579-022-00779-w
- Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* 18, 113–119. doi: 10.1038/s41579-019-0278-2
- Bessonov, K., Laing, C., Robertson, J., Yong, I., Ziebell, K., Gannon, V. P. J., et al. (2021). ECTyper: in silico *Escherichia coli* serotype and species prediction from raw and assembled whole-genome sequence data. *Microb. Genomics* 7, 728. doi: 10.1099/mgen.0.000728
- Boeckaerts, D., Stock, M., Ferriol-González, C., Oteo-Iglesias, J., Sanjuán, R., Domingo-Calap, P., et al. (2024). Prediction of *Klebsiella* phage-host specificity at the strain level. *Nat. Commun.* 15, 4355. doi: 10.1038/s41467-024-48675-6
- Borin, J. M., Avrani, S., Barrick, J. E., Petrie, K. L., and Meyer, J. R. (2021). Coevolutionary phage training leads to greater bacterial suppression and delays the evolution of phage resistance. *Proc. Natl. Acad. Sci.* 118, e2104592118. doi: 10.1073/pnas.2104592118
- Borin, J. M., Lee, J. J., Lucia-Sanz, A., Gerbino, K. R., Weitz, J. S., and Meyer, J. R. (2023). Rapid bacteria-phage coevolution drives the emergence of multiscale networks. *Science* 382, 674–678. doi: 10.1126/science.adf5536
- Chang, C., Yu, X., Guo, W., Guo, C., Guo, X., Li, Q., et al. (2022). Bacteriophage-mediated control of biofilm: A promising new dawn for the future. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.825828
- Chen, L., Zhao, X., Wongso, S., Lin, Z., and Wang, S. (2024). Trade-offs between receptor modification and fitness drive host-bacteriophage co-evolution leading to phage extinction or co-existence. *ISME J.* 18, wrae214. doi: 10.1093/ismej/wrae214
- Concha-Eloko, R., Stock, M., Baets, B. D., Briers, Y., Sanjuán, R., Domingo-Calap, P., et al. (2024). DepoScope: Accurate phage depolymerase annotation and domain delineation using large language models. *PLoS Comput. Biol.* 20, e1011831. doi: 10.1371/journal.pcbi.1011831
- Costa, A. R., van den Berg, D. F., Esser, J. Q., Muralidharan, A., van den Bossche, H., Bonilla, B. E., et al. (2024). Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*. *Sci. Adv.* 10, ead0341. doi: 10.1126/sciadv.ad0341
- Dedrick, R. M., Guerrero-Bustamante, C. A., Garlena, R. A., Russell, D. A., Ford, K., Harris, K., et al. (2019). Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.* 25, 730–733. doi: 10.1038/s41591-019-0437-z
- Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., et al. (2018). Systematic discovery of anti-phage defense systems in the microbial pan-genome. *Science* 359, eaar4120. doi: 10.1126/science.aar4120
- Du, J., Meile, S., Baggenstos, J., Jäggi, T., Piffaretti, P., Hunold, L., et al. (2023). Enhancing bacteriophage therapeutics through *in situ* production and release of heterologous antimicrobial effectors. *Nat. Commun.* 14, 4337. doi: 10.1038/s41467-023-39612-0
- Dunne, M., Prokhorov, N. S., Loessner, M. J., and Leiman, P. G. (2021). Reprogramming bacteriophage host range: design principles and strategies for engineering receptor binding proteins. *Curr. Opin. Biotechnol.* 68, 272–281. doi: 10.1016/j.copbio.2021.02.006
- Dunne, M., Rupf, B., Tala, M., Qabrati, X., Ernst, P., Shen, Y., et al. (2019). Reprogramming bacteriophage host range through structure-guided design of chimeric receptor binding proteins. *Cell Rep.* 29, 1336–1350.e4. doi: 10.1016/j.celrep.2019.09.062
- Dunstan, R. A., Bamert, R. S., Belousoff, M. J., Short, F. L., Barlow, C. K., Pickard, D. J., et al. (2021). Mechanistic Insights into the Capsule-Targeting Depolymerase from a *Klebsiella pneumoniae* Bacteriophage. *Microbiol. Spectr.* 9. doi: 10.1128/spectrum.01023-21
- Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., et al. (2022). Protein complex prediction with AlphaFold-Multimer 463034. doi: 10.1101/2021.10.04.463034
- Fang, Z., Feng, T., Zhou, H., and Chen, M. (2022). DeePVP: Identification and classification of phage virion proteins using deep learning. *GigaScience* 11, giac076. doi: 10.1093/gigascience/giac076
- Farquharson, E. L., Lightbown, A., Pulkkinen, E., Russell, T., Werner, B., and Nugen, S. R. (2021). Evaluating phage tail fiber receptor-binding proteins using a luminescent flow-through 96-well plate assay. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.741304
- Fischetti, V. A. (2018). Development of phage lysins as novel therapeutics: A historical perspective. *Viruses* 10, 310. doi: 10.3390/v10060310
- Flamholz, Z. N., Li, C., and Kelly, L. (2024). Improving viral annotation with artificial intelligence. *mBio* 15, e03206–e03223. doi: 10.1128/mbio.03206-23
- Gaborieau, B., Vaysset, H., Tesson, F., Charachon, I., Dib, N., Bernier, J., et al. (2024). Prediction of strain level phage-host interactions across the *Escherichia* genus using only genomic information. *Nat. Microbiol.* 9, 2847–2861. doi: 10.1038/s41564-024-01832-5
- Ge, X., and Wang, J. (2024). Structural mechanism of bacteriophage lambda tail's interaction with the bacterial receptor. *Nat. Commun.* 15, 4185. doi: 10.1038/s41467-024-48686-3
- Gencay, Y. E., Jasinskytė, D., Robert, C., Semsey, S., Martínez, V., Petersen, A. Ø, et al. (2023). Engineered phage with antibacterial CRISPR-Cas selectively reduce *E. coli* burden in mice. *Nat. Biotechnol.* 42, 265–274. doi: 10.1038/s41587-023-01759-y
- Georjon, H., and Bernheim, A. (2023). The highly diverse antiphage defence systems of bacteria. *Nat. Rev. Microbiol.* 21, 686–700. doi: 10.1038/s41579-023-00934-x
- Gerstmans, H., Grimon, D., Gutiérrez, D., Lood, C., Rodríguez, A., van Noort, V., et al. (2020). A VersaTile-driven platform for rapid hit-to-lead development of engineered lysins. *Sci. Adv.* 6, eaaz1136. doi: 10.1126/sciadv.aaz1136

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Hatfull, G. F. (2015). Dark matter of the biosphere: the amazing world of bacteriophage diversity. *J. Virol.* 89, 8107–8110. doi: 10.1128/JVI.01340-15
- Hernando-Amado, S., Coque, T. M., Baquero, F., and Martínez, J. L. (2019). Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat. Microbiol.* 4, 1432–1442. doi: 10.1038/s41564-019-0503-9
- Hie, B. L., Shanker, V. R., Xu, D., Bruun, T. U. J., Weidenbacher, P. A., Tang, S., et al. (2023). Efficient evolution of human antibodies from general protein language models. *Nat. Biotechnol.* 42, 275–283. doi: 10.1038/s41587-023-01763-2
- Hsu, C., Fannjiang, C., and Listgarten, J. (2024). Generative models for protein structures and sequences. *Nat. Biotechnol.* 42, 196–199. doi: 10.1038/s41587-023-02115-w
- Hu, B., Margolin, W., Molineux, I. J., and Liu, J. (2015). Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl. Acad. Sci. U. S. A.* 112, E4919–E4928. doi: 10.1073/pnas.1501064112
- Hughes, K. A., Sutherland, I. W., and Jones, M. V. (1998). Biofilm susceptibility to bacteriophage attack: the role of phage-borne polysaccharide depolymerase. *Microbiol. Read. Engl.* 144, 3039–3047. doi: 10.1099/0021287-144-11-3039
- Huss, P., Chen, J., and Raman, S. (2023). High-throughput approaches to understand and engineer bacteriophages. *Trends Biochem. Sci.* 48, 187–197. doi: 10.1016/j.tibs.2022.08.012
- Huss, P., Kieft, K., Meger, A., Nishikawa, K., Anantharaman, K., and Raman, S. (2024). Engineering bacteriophages through deep mining of metagenomic motifs 527309. doi: 10.1101/2023.02.07.527309
- Huss, P., Meger, A., Leander, M., Nishikawa, K., and Raman, S. (2021). Mapping the functional landscape of the receptor binding domain of T7 bacteriophage by deep mutational scanning. *eLife* 10, e63775. doi: 10.7554/eLife.63775
- Jaschke, P. R., Lieberman, E. K., Rodriguez, J., Sierra, A., and Endy, D. (2012). A fully decompressed synthetic bacteriophage øX174 genome assembled and archived in yeast. *Virology* 434, 278–284. doi: 10.1016/j.virol.2012.09.020
- Johansen, J., Plichta, D. R., Nissen, J. N., Jespersen, M. L., Shah, S. A., Deng, L., et al. (2022). Genome binning of viral entities from bulk metagenomics data. *Nat. Commun.* 13, 965. doi: 10.1038/s41467-022-28581-5
- Kamata, K., Birkholz, N., Ceelen, M., Fagerlund, R. D., Jackson, S. A., and Fineran, P. C. (2024). Repurposing an endogenous CRISPR-cas system to generate and study subtle mutations in bacteriophages. *CRISPR J.* 7 (6), 343–354. doi: 10.1089/crispr.2024.0047
- 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. doi: 10.1073/pnas.1714658115
- Knecht, L. E., Veljkovic, M., and Fieseler, L. (2020). Diversity and function of phage encoded depolymerases. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.02949
- Kortright, K. E., Chan, B. K., Evans, B. R., and Turner, P. E. (2022). Arms race and fluctuating selection dynamics in *Pseudomonas aeruginosa* bacteria coevolving with phage OMK01. *J. Evol. Biol.* 35, 1475–1487. doi: 10.1111/jeb.14095
- Krusche, J., Beck, C., Lehmann, E., Gerlach, D., Daiber, E., Mayer, C., et al. (2025). Characterization and host range prediction of *Staphylococcus aureus* phages through receptor-binding protein analysis. *Cell Rep.* 44, 115369. doi: 10.1016/j.celrep.2025.115369
- Lam, M. M. C., Wick, R. R., Judd, L. M., Holt, K. E., and Wyres, K. L. (2022). Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. *Microb. Genomics* 8, 800. doi: 10.1099/mgen.0.000800
- Lenneman, B. R., Fernbach, J., Loessner, M. J., Lu, T. K., and Kilcher, S. (2021). Enhancing phage therapy through synthetic biology and genome engineering. *Curr. Opin. Biotechnol.* 68, 151–159. doi: 10.1016/j.copbio.2020.11.003
- Lucia-Sanz, A., Peng, S., Leung, C. Y. J., Gupta, A., Meyer, J. R., and Weitz, J. S. (2024). Inferring strain-level mutational drivers of phage-bacteria interaction phenotypes arising during coevolutionary dynamics. *Virus Evol.* 10, veae104. doi: 10.1093/ve/veae104
- Luria, S. E., and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28, 491–511. doi: 10.1093/genetics/28.6.491
- Mayo-Muñoz, D., Pinilla-Redondo, R., Camara-Wilpert, S., Birkholz, N., and Fineran, P. C. (2024). Inhibitors of bacterial immune systems: discovery, mechanisms and applications. *Nat. Rev. Genet.* 25, 237–254. doi: 10.1038/s41576-023-00676-9
- Mayorga-Ramos, A., Carrera-Pacheco, S. E., Barba-Ostria, C., and Guzmán, L. P. (2024). Bacteriophage-mediated approaches for biofilm control. *Front. Cell. Infect. Microbiol.* 14. doi: 10.3389/fcimb.2024.1428637
- McNair, K., Bailey, B. A., and Edwards, R. A. (2012). PHACTS, a computational approach to classifying the lifestyle of phages. *Bioinform. Oxf. Engl.* 28, 614–618. doi: 10.1093/bioinformatics/bts014
- Meyer, J. R., Dobias, D. T., Weitz, J. S., Barrick, J. E., Quick, R. T., and Lenski, R. E. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* 335, 428–432. doi: 10.1126/science.1214449
- Müller, D. M., Pourtois, J. D., Kim, M. K., Targ, B., Burgener, E. B., Milla, C., et al. (2024). Bacterial Receptors but Not Anti-Phage Defence Mechanisms Determine Host Range for a Pair of *Pseudomonas aeruginosa* Lytic Phages 591980. doi: 10.1101/2024.04.30.591980
- Murtazaliev, K., Mu, A., Petrovskaya, A., and Finn, R. D. (2024). The growing repertoire of phage anti-defence systems. *Trends Microbiol.* 32 (12), 1212–1228. doi: 10.1016/j.tim.2024.05.005
- Naghavi, M., Vollset, S. E., Ikuta, K. S., Swetschinski, L. R., Gray, A. P., Wool, E. E., et al. (2024). Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet* 404, 1199–1226. doi: 10.1016/S0140-6736(24)01867-1
- Nair, A., and Khairnar, K. (2019). Genetically engineered phages for therapeutics: proceed with caution. *Nat. Med.* 25, 1028–1028. doi: 10.1038/s41591-019-0506-3
- Nami, Y., Imani, N., and Panahi, B. (2021). Application of machine learning in bacteriophage research. *BMC Microbiol.* 21, 193. doi: 10.1186/s12866-021-02256-5
- Nguyen, E., Poli, M., Durrant, M. G., Kang, B., Katrekara, D., Li, D. B., et al. (2024). Sequence modeling and design from molecular to genome scale with Evo. *Science* 386, ead9336. doi: 10.1126/science.ad9336
- Nobrega, F. L., Vlot, M., de Jonge, P. A., Dreesens, L. L., Beaumont, H. J. E., Lavigne, R., et al. (2018). Targeting mechanisms of tailed bacteriophages. *Nat. Rev. Microbiol.* 16, 760–773. doi: 10.1038/s41579-018-0070-8
- Pires, D. P., Monteiro, R., Mil-Homens, D., Fialho, A., Lu, T. K., and Azeredo, J. (2021). Designing *P. aeruginosa* synthetic phages with reduced genomes. *Sci. Rep.* 11, 2164. doi: 10.1038/s41598-021-81580-2
- Pirnay, J.-P. (2020). Phage therapy in the year 2035. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.01171
- Pirnay, J.-P., Djebara, S., Steurs, G., Griselain, J., Cochez, C., De Soir, S., et al. (2024). Personalized bacteriophage therapy outcomes for 100 consecutive cases: a multicentre, multinational, retrospective observational study. *Nat. Microbiol.* 9, 1434–1453. doi: 10.1038/s41564-024-01705-x
- Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J., Lessor, L., et al. (2017). Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *acinetobacter baumannii* infection. *Antimicrob. Agents Chemother.* 61, e00954–e00917. doi: 10.1128/AAC.00954-17
- Shaer Tamar, E., and Kishony, R. (2022). Multistep diversification in spatiotemporal bacterial-phage coevolution. *Nat. Commun.* 13, 7971. doi: 10.1038/s41467-022-35351-w
- Shang, J., Tang, X., and Sun, Y. (2023). PhaTYP: predicting the lifestyle for bacteriophages using BERT. *Brief. Bioinform.* 24, bbac487. doi: 10.1093/bib/bbac487
- Shao, B., and Yan, J. (2024). A long-context language model for deciphering and generating bacteriophage genomes. *Nat. Commun.* 15, 9392. doi: 10.1038/s41467-024-53759-4
- Shu, H.-Y., Fung, C.-P., Liu, Y.-M., Wu, K.-M., Chen, Y.-T., Li, L.-H., et al. (2009). Genetic diversity of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* clinical isolates. *Microbiol. Read. Engl.* 155, 4170–4183. doi: 10.1099/mic.0.029017-0
- Strathdee, S. A., Hatfull, G. F., Mutalik, V. K., and Schooley, R. T. (2023). Phage therapy: From biological mechanisms to future directions. *Cell* 186, 17–31. doi: 10.1016/j.cell.2022.11.017
- Suga, A., Kawaguchi, M., Yonesaki, T., and Otsuka, Y. (2021). Manipulating interactions between T4 phage long tail fibers and *escherichia coli* receptors. *Appl. Environ. Microbiol.* 87, e00423–e00421. doi: 10.1128/AEM.00423-21
- Thung, T. Y., White, M. E., Dai, W., Wilksch, J. J., Bamert, R. S., Rocker, A., et al. (2021). Component parts of bacteriophage virions accurately defined by a machine-learning approach built on evolutionary features. *mSystems* 6. doi: 10.1128/mSystems.00242-21
- Vassallo, C. N., Doering, C. R., Littlehale, M. L., Teodoro, G. I. C., and Laub, M. T. (2022). A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome. *Nat. Microbiol.* 7, 1568–1579. doi: 10.1038/s41564-022-01219-4
- Wu, S., Fang, Z., Tan, J., Li, M., Wang, C., Guo, Q., et al. (2021). DeePhage: distinguishing virulent and temperate phage-derived sequences in metavirome data with a deep learning approach. *GigaScience* 10, giab056. doi: 10.1093/gigascience/giab056
- Yehl, K., Lemire, S., Yang, A. C., Ando, H., Mimeo, M., Torres, M. D. T., et al. (2019). Engineering phage host-range and suppressing bacterial resistance through phage tail fiber mutagenesis. *Cell* 179, 459–469.e9. doi: 10.1016/j.cell.2019.09.015
- Yirmiya, E., Hobbs, S. J., Leavitt, A., Osterman, I., Avraham, C., Hochhauser, D., et al. (2025). Structure-guided discovery of viral proteins that inhibit host immunity. *Cell* 188 (6), 1681–1692.e17. doi: 10.1016/j.cell.2024.12.035
- Yirmiya, E., Leavitt, A., Lu, A., Ragucci, A. E., Avraham, C., Osterman, I., et al. (2024). Phages overcome bacterial immunity via diverse anti-defence proteins. *Nature* 625, 352–359. doi: 10.1038/s41586-023-06869-w
- Zhang, Y., Li, R., Zou, G., Guo, Y., Wu, R., Zhou, Y., et al. (2024). Discovery of antimicrobial lysins from the “Dark matter” of uncharacterized phages using artificial intelligence. *Adv. Sci.* 11, 2404049. doi: 10.1002/adv.202404049