Communication

Comparative Genomics of Closely-Related *Gordonia* Cluster DR Bacteriophages

Cyril J. Versoza [1,†], **Abigail A. Howell** [2,†], **Tanya Aftab** [3], **Madison Blanco** [3], **Akarshi Brar** [3], **Elaine Chaffee** [3], **Nicholas Howell** [3,4], **Willow Leach** [4], **Jackelyn Lobatos** [3], **Michael Luca** [3,5], **Meghna Maddineni** [5], **Ruchira Mirji** [3], **Corinne Mitra** [3], **Maria Strasser** [3], **Saige Munig** [3], **Zeel Patel** [3], **Minerva So** [3], **Makena Sy** [3], **Sarah Weiss** [3] and **Susanne P. Pfeifer** [1,*]

[1] Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ 85281, USA; cversoza@asu.edu
[2] Biodesign Institute, School of Life Sciences, Arizona State University, Tempe, AZ 85281, USA; aahowel3@asu.edu
[3] School of Life Sciences, Arizona State University, Tempe, AZ 85281, USA; taftab@asu.edu (T.A.); mmblanco@asu.edu (M.B.); akbrar@asu.edu (A.B.); emchaffe@asu.edu (E.C.); nwhowell@asu.edu (N.H.); jlobatos@asu.edu (J.L.); mluca@asu.edu (M.L.); rmirji@asu.edu (R.M.); cmitra@asu.edu (C.M.); mbstrass@asu.edu (M.S.); smunig@asu.edu (S.M.); zspatel1@asu.edu (Z.P.); minervas@asu.edu (M.S.); mrsy2@asu.edu (M.S.); sweiss13@asu.edu (S.W.)
[4] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85281, USA; waleach@asu.edu
[5] School of Molecular Sciences, Arizona State University, Tempe, AZ 85281, USA; mmaddine@asu.edu
[*] Correspondence: susanne.pfeifer@asu.edu
[†] These authors contributed equally to this work.

Abstract: Bacteriophages infecting bacteria of the genus *Gordonia* have increasingly gained interest in the scientific community for their diverse applications in agriculture, biotechnology, and medicine, ranging from biocontrol agents in wastewater management to the treatment of opportunistic pathogens in pulmonary disease patients. However, due to the time and costs associated with experimental isolation and cultivation, host ranges for many bacteriophages remain poorly characterized, hindering a more efficient usage of bacteriophages in these areas. Here, we perform a series of computational genomic inferences to predict the putative host ranges of all *Gordonia* cluster DR bacteriophages known to date. Our analyses suggest that BiggityBass (as well as several of its close relatives) is likely able to infect host bacteria from a wide range of genera—from *Gordonia* to *Nocardia* to *Rhodococcus,* making it a suitable candidate for future phage therapy and wastewater treatment strategies.

Keywords: bacteriophage; cluster DR; *Gordonia;* comparative genomics; host range

1. Introduction

Bacteriophages are one of the most abundant organisms on Earth, infecting a wide range of host bacteria present in almost any environment from common garden soil to volcanic substrates and from freshwater streams to oceans [1]. Among these hosts, members of the order *Corynebacteriales*—including *Gordonia, Mycobacterium, Nocardia,* and *Rhodococcus*—are of particular importance to agriculture, biotechnology, and medicine as the outer membrane of their bacterial cells, which consists of long-chain hydroxylated mycolic acids, frequently leads to complications during the prevention, treatment, and cure of opportunistic pathogens [2]. Moreover, due to the hydrophobic nature of this "mycomembrane", *Corynebacteriales* often cause severe problems during wastewater treatment as they can stabilize foams on the surface of aeration tanks during the activated

sludge phase [3], which not only complicates sludge management and increases maintenance costs but also poses a health hazard to wastewater treatment plant workers in their aerosolized form [4].

Owing to the growing scarcity of clean water across the globe, treated wastewater serves as an important alternative to freshwater for many nations with more than 35% of agricultural irrigation, 17% of landscape irrigation, and 12% of groundwater recharge in the United States stemming from treated wastewater [5]. However, microbial hazards, such as multi-drug resistant bacterial pathogens, are frequently discharged into sewage systems due to the common usage of antibiotics in animal farms and on crop fields. Consequently, effective wastewater treatment strategies are indispensable to combat environmental and health concerns for farmers and consumers alike [6].

Due to their host specificity, lytic bacteriophages have been proposed as promising and environmentally-friendly bacterial treatment and control agents to remove harmful (or otherwise problematic) bacteria—such as gram-positive *Gordonia* which are associated with both systemic infections in immunocompromised and local infections in immunocompetent individuals [7,8] as well as sludge foaming [9,10]—while maintaining desirable microorganisms in the wastewater. To effectively guide these biological control strategies, bacteriophages and their host ranges (i.e., the bacterial genera and species a bacteriophage is able to infect) must be well-characterized—yet, the diversity of *Gordonia* bacteriophages remains largely unexplored.

As part of a course-based undergraduate research experience at Arizona State University, we computationally inferred putative host ranges of all *Gordonia* cluster DR bacteriophages known to date to aid the design and improvement of future wastewater treatment strategies.

2. Materials and Methods

Genomic data for *Gordonia* cluster DR bacteriophages (Supplementary Table S1) were explored using Phamerator [11] and phylogenetic relationships characterized together with representative *Microbacterium*, *Mycobacterium*, and *Streptomyces* bacteriophages as outgroups (Supplementary Table S2). Specifically, MAFFT v.7 [12] embedded within the EMBL-EBI Bioinformatics Toolkit [13,14] was used to generate a multiple-sequence alignment between the bacteriophages. The resulting alignment was then used to generate a neighbor-joining tree in MEGA X [15] using a phylogeny test with 10,000 bootstrap replicates. Nucleotide sequence relatedness was assessed using Gepard v.2.1.0 [16]. Pairwise average nucleotide identities (ANIs) were calculated using the "Genome Comparison" tool embedded within DNA Master v.5.23.6 and plotted using the ggplot2 package [17] in R v.4.1.0.

Following suggested best practices by Versoza and Pfeifer [18], a combination of exploratory and confirmatory methods was utilized to computationally predict host ranges of the closely-related *Gordonia* cluster DR bacteriophages. Specifically, putative host ranges were predicted using two machine-learning based prediction tools—CHERRY [19] and PHERI v.0.2 [20]—as well as the alignment-free prediction tool WIsH v.1.1 [21] together with genomic data from ten putative bacterial host species spanning three genera—*Gordonia*, *Nocardia*, *Rhodococcus*, and, as a negative control, *Escherichia* (Supplementary Table S3). All software was executed using default settings.

3. Results

To confirm cluster membership, the genomes of *Gordonia* cluster DR bacteriophages were investigated. They show a high level of sequence similarity with the left arm of the genomes mostly encoding well-conserved structural and assembly proteins (including a terminase, portal protein, capsid maturation protein as well as major capsid hexamer and pentamer proteins, a head-to-tail adaptor, tail assembly protein, tape measure protein, minor tail protein subunits, lysin A, lysin B, and several genes responsible for integration into the host). Thereby, the RuvC-like resolvase (Supplementary Figure S1), a Holliday

junction resolving enzyme that is a distant relative of the RuvC proteins present in gram-negative bacteria such as *Escherichia coli* [22], is of particular interest. It closely resembles the RuvC-like endonucleases found in select *Siphoviridae* and *Myoviridae* bacteriophages infecting *Streptococcus* and *Lactococcus* hosts [23,24], which may hint at a shared evolutionary history. The right arm of the genomes contains non-structural genes (including an exonuclease, DNA helicase, DNA polymerase, and HNH endonuclease). Notably, several cluster DR bacteriophages exhibit a partial toxin/antitoxin (TA) system (Supplementary Figure S2). Prevalent in many archaea and bacteria, TA systems encode a toxin protein and a corresponding antitoxin in the form of a protein or non-coding RNA that serves as a defense mechanism against invading bacteriophages [25,26]. As bacteriophages co-evolve with their bacterial hosts [27], adaptations to such defense mechanisms are common [28] to allow bacteriophages to inactivate bacteria-encoded toxins [29,30]. Indeed, the TA system of the cluster DR bacteriophages is homologous to the *hicA* TA system frequently present in *Burkholderia pseudomallei*, *E. coli*, and *Pseudomonas aeruginosa* [31–33].

To elucidate phylogenetic relationships, comparative analyses were performed between all *Gordonia* cluster DR bacteriophages known to date (Supplementary Table S1). Following Pope and colleagues [34], clustering was based on nucleotide similarity and shared gene content, with bacteriophages sharing at least 35% of genes being grouped into clusters. A neighbor-joining tree confirmed membership in the DR cluster (Supplementary Figure S3a)—an assignment that was further supported by both the dot plot analyses (Supplementary Figure S4) as well as the pairwise average nucleotide identities (Supplementary Figure S5). Interestingly, gene trees of the RuvC-like resolvase (Supplementary Figure S3b) and the *hicA*-like toxin (Supplementary Figure S3c) do not recapitulate the whole genome phylogeny—however, it is unclear whether this is due to inconsistent resampling during bootstrapping caused by the short sequence length [35] or the mosaic architecture of the genome caused by horizontal gene transfer by illegitimate recombination [36–38]. Compared to temperate bacteriophages, both gene acquisition and gene loss, in lytic bacteriophages is less well understood [39]. However, there have been previous reports of gene transfers in T4-like and T7-like bacteriophages [40,41], and lytic bacteriophages with large genomes have been suggested to have acquired genes from donor genomes [42].

Due to their bactericidal nature, bacteriophages are frequently used for a variety of agricultural, biotechnological, and medical applications [43]. To effectively guide the usage of bacteriophages in these areas, their host ranges have to first be determined (see discussion in [18]). To investigate the host ranges of the closely related cluster DR bacteriophages, a combination of exploratory and confirmatory prediction tools was utilized together with a dataset of ten putative bacterial host species and *E. coli* as a negative control (Supplementary Table S3). Specifically, the tested host dataset spans the three genera of the *Corynebacteriales* order—*Gordonia*, *Nocardia*, and *Rhodococcus*—that have been implicated in activated sludge foaming in wastewater treatment plants [44].

Using the exploratory method PHERI [20], seven out of nine cluster DR bacteriophages were predicted to infect hosts under the *Gordonia* genus (Table 1), with the exception of bacteriophages AnClar and Yago84. To make host range predictions for newly encountered bacteriophages, PHERI utilizes a decision tree classifier of annotated protein clusters of bacteriophages with known hosts. Consequently, bacteriophages will only be predicted to infect a particular host if their protein profile closely matches that of another bacteriophage known to infect that host. As minor tail proteins play an essential role in bacteriophage infection [45], the lack of similarity in the minor tail protein profiles of AnClar and Yago84 compared to those bacteriophages known to infect *Gordonia* hosts might explain why neither were predicted to infect the *Gordonia* genus, despite having been isolated in *G. terrae* (Supplementary Table S1). In fact, the clades observed within the gene tree of the minor tail protein shared across all cluster DR bacteriophages (Supplementary Figure S3d) reflects the clustering of the bacteriophages with respect to host range, reiterating the importance of tail proteins for host infection. Using the exploratory method

CHERRY [19]—a graph convolutional encoder and decoder that relies on a broader range of features including protein organization, sequence similarity, and *k*-mer frequency to predict host ranges—highlights *M. smegmatis*, *G. terrae*, and *R. hoagie* as the three most likely host candidates for all cluster DR bacteriophages (though the latter two scoring predictions fell below the recommended confidence threshold of 0.9). Conversely, the confirmatory method WIsH [21]—based on a Markov model that determines the *k*-mer similarity between bacteriophage and host genomes—predicted *G. hydrophobica*, *G. malaquae*, *G. rubripertincta*, and *G. terrae* as potential hosts for all nine cluster DR bacteriophages relative to the negative control, *E. coli* (Figure 1). Moreover, log likelihood values for putative *Nocardia* and *Rhodococcus* hosts were comparable to those of *Gordonia*, suggesting the potential for a much broader host range. Interestingly, BiggityBass exhibits the broadest predicted host range among all cluster DR bacteriophages, spread across five different phyla (Table 1), making it an appealing agent to explore for future wastewater treatment strategies [46].

In conclusion, computational methods can offer a first glimpse into the putative host ranges of newly discovered bacteriophages—yet, it is important to remember that these methods are predictive by their very nature. Thereby, each computational method exhibits their own advantages and limitations. For example, tools that rely solely on *k*-mer-based models can lead to an overprediction of host ranges if convergent evolution resulted in similar nucleotide frequency patterns [47], whereas tools that rely on machine-learning are inherently limited in their predictions by the bacteriophage-host datasets available for training [18]. Experimental validation through bacteriophage isolation and cultivation still remains the "gold standard" in determining bacteriophage host ranges—however, it certainly is not without its own limitations as not all microbial hosts are amendable to cultivation in the laboratory and, even if they are, results may depend on the conditions under which the experiments were performed [18]. Given the ever growing knowledge of bacteriophage diversity across the globe, it is our hope that future computational and experimental research will go hand in hand to further explore polyvalent bacteriophages as an interesting study system to gain a better understanding of the molecular and genetic determinants underlying host range.

Table 1. Putative host ranges as predicted by PHERI. Putative hosts of the nine *Gordonia* cluster DR bacteriophages included in this study (Supplementary Table S1) predicted by PHERI [20].

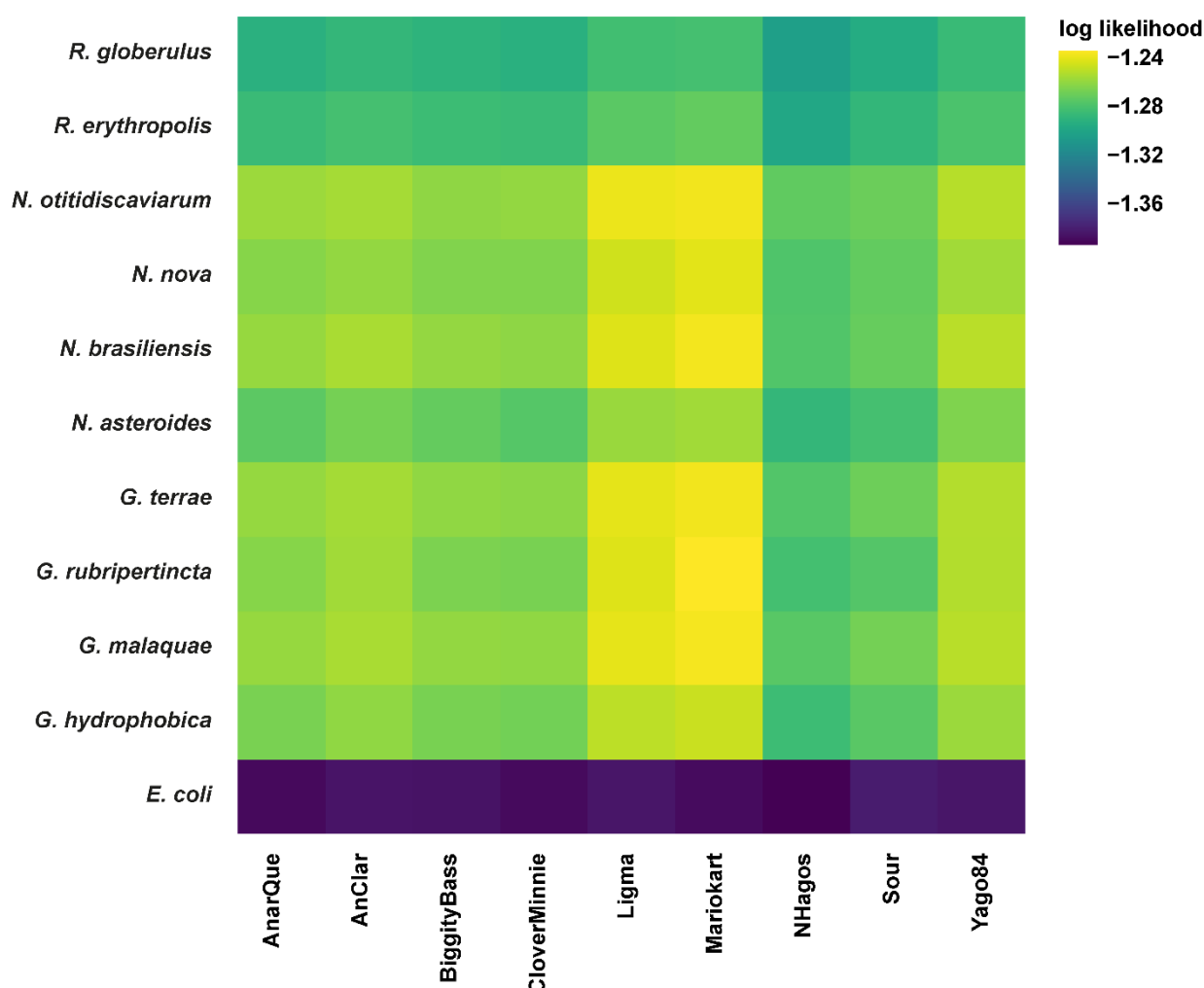|	*Gordonia*	*Arthrobacter*	*Aeromonas*	*Staphylococcus*	*Shigella*	*Corynebacterium*	*Stenotrophomonas*
AnarQue	√	√	√				
AnClar		√	√				
BiggityBass	√	√	√	√	√		
CloverMinnie	√	√	√				
Ligma	√	√	√			√	
Mariokart	√	√	√				
NHagos	√	√	√				
Sour	√	√	√				√
Yago84		√	√				

Figure 1. Putative host ranges as predicted by WIsH. Heatmap of log-likelihoods of bacteriophage-host pairs—including nine *Gordonia* cluster DR bacteriophages (Supplementary Table S1) as well as ten potential bacterial hosts and *E. coli* as a negative control (Supplementary Table S3)—generated by the host prediction tool WIsH [21]. Higher values correspond to more likely interactions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rohwer, F. Global Phage Diversity. *Cell* **2003**, *113*, 141. https://doi.org/10.1016/S0092-8674(03)00276-9.
2. Dyson, Z.A.; Tucci, J.; Seviour, R.J.; Petrovski, S. Lysis to kill: Evaluation of the lytic abilities, and genomics of nine bacterio-phages infective for *Gordonia* spp. and their potential use in activated sludge foam biocontrol. *PLoS ONE* **2015**, *10*, e0134512. https://doi.org/10.1371/journal.pone.0134512.
3. Petrovski, S.; Seviour, R.J.; Tillett, D. Characterization of the genome of the polyvalent lytic bacteriophage GTE2, which has potential for biocontrol of *Gordonia-*, *Rhodococcus-*, and *Nocardia*-stabilized foams in activated sludge plants. *Appl. Environ. Microbiol.* **2011**, *77*, 3923–3929. https://doi.org/10.1128/AEM.00025-11.
4. Pal, P.; Kumar, R. Treatment of coke wastewater: A critical review for developing sustainable management strategies. *Sep. Purif. Rev.* **2014**, *43*, 89–123. https://doi.org/10.1080/15422119.2012.717161.
5. Kesari, K.K.; Soni, R.; Jamal, Q.M.S.; Tripathi, P.; Lal, J.A.; Jha, N.K.; Siddiqui, M.H.; Kumar, P.; Tripathi, V.; Ruokolainen, J. Wastewater treatment and reuse: A review of its applications and health implications. *Water Air Soil Pollut.* **2021**, *232*, 208. https://doi.org/10.1007/s11270-021-05154-8.
6. Dang, Q.; Tan, W.; Zhao, X.; Li, D.; Li, Y.; Yang, T.; Li, R.; Zu, G.; Xi, B. Linking the response of soil microbial community structure in soils to long-term wastewater irrigation and soil depth. *Sci. Total Environ.* **2019**, *688*, 26–36. https://doi.org/10.1016/j.scitotenv.2019.06.138.
7. Arenskötter, M.; Bröker, D.; Steinbüchel, A. Biology of the metabolically diverse genus *Gordonia*. *Appl. Environ. Microbiol.* **2004**, *70*, 3195–3204. https://doi.org/10.1128/AEM.70.6.3195-3204.2004.
8. Grisold, A.J.; Roll, P.; Hoenigl, M.; Feierl, G.; Vicenzi-Moser, R.; Marth, E. Isolation of *Gordonia terrae* from a patient with catheter-related bacteraemia. *J. Med. Microbiol.* **2007**, *56*, 1687–1688. https://doi.org/10.1099/jmm.0.47388-0.
9. De los Reyes, M.F.; de los Reyes, F.L.; Hernandez, M.; Raskin, L. Quantification of *Gordona amarae* strains in foaming activated sludge and anaerobic digester systems with oligonucleotide hybridization probes. *Appl. Environ. Microbiol.* **1998**, *64*, 2503–2512. https://doi.org/10.1128/AEM.64.7.2503-2512.1998.
10. Kragelund, C.; Remesova, Z.; Nielsen, J.L.; Thomsen, T.R.; Eales, K.; Seviour, R.; Wanner, J.; Nielsen, P.H. Ecophysiology of mycolic acid-containing Actinobacteria (Mycolata) in activated sludge foams. *FEMS Microbiol. Ecol.* **2007**, *61*, 174–184. https://doi.org/10.1111/j.1574-6941.2007.00324.x.
11. Cresawn, S.G.; Bogel, M.; Day, N.; Jacobs-Sera, D.; Hendrix, R.W.; Hatfull, G.F. Phamerator: A bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinform.* **2011**, *12*, 395. https://doi.org/10.1186/1471-2105-12-395.
12. Katoh, K.; Standley, D.M. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. https://doi.org/10.1093/molbev/mst010.
13. Zimmermann, L.; Stephens, A.; Nam, S.-Z.; Rau, D.; Kübler, J.; Lozajic, M.; Gabler, F.; Söding, J.; Lupas, A.N.; Alva, V. A completely reimplemented MPI Bioinformatics Toolkit with a new HHpred server at its core. *J. Mol. Biol.* **2018**, *430*, 2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.
14. Gabler, F.; Nam, S.; Till, S.; Mirdita, M.; Steinegger, M.; Söding, J.; Lupas, A.N.; Alva, V. Protein sequence analysis using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinform.* **2020**, *72*, e108. https://doi.org/10.1002/cpbi.108.
15. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. https://doi.org/10.1093/molbev/msy096.
16. Krumsiek, J.; Arnold, R.; Rattei, T. Gepard: A rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **2007**, *23*, 1026–1028. https://doi.org/10.1093/bioinformatics/btm039.
17. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2009; ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org (accessed on 15 May 2022).
18. Versoza, C.J.; Pfeifer, S.P. Computational prediction of bacteriophage host ranges. *Microorganisms* **2022**, *10*, 149. https://doi.org/10.3390/microorganisms10010149.
19. Shang, J.; Sun, Y. CHERRY: A Computational MetHod for AccuratE PRediction of Virus-PRokarYotic interactions using a graph encoder-decoder model. *Brief. Bioinform.* **2022**, bbac182. https://doi.org/10.1093/bib/bbac182.
20. Baláž, A.; Kajsík, M.; Budiš, J.; Szemeš, T.; Turňa, J. PHERI—Phage Host Exploration Pipeline. *bioRxiv* **2020**. https://www.bio-rxiv.org/content/10.1101/2020.05.13.093773v3 (accessed on 15 May 2022).
21. Galiez, C.; Siebert, M.; Enault, F.; Vincent, J.; Söding, J. WIsH: Who Is the Host? Predicting prokaryotic hosts from metagenomic phage contigs. *Bioinformatics* **2017**, *33*, 3113–3114. https://doi.org/10.1093/bioinformatics/btx383.
22. Lilley, D.M.J.; White, M.F. The junction-resolving enzymes. *Nat. Rev. Mol. Cell. Biol.* **2001**, *2*, 433–443. https://doi.org/10.1038/35073057x.
23. Bidnenko, E.; Ehrlich, S.D.; Chopin, M.-C. *Lactococcus lactis* phage operon coding for an endonuclease homologous to RuvC. *Mol. Microbiol.* **2002**, *28*, 823–834. https://doi.org/10.1046/j.1365-2958.1998.00845.x.
24. Curtis, F.A.; Reed, P.; Sharples, G.J. Evolution of a phage RuvC endonuclease for resolution of both Holliday and branched DNA junctions. *Mol. Microbiol.* **2004**, *55*, 1332–1345. https://doi.org/10.1111/j.1365-2958.2004.04476.x.
25. Unterholzner, S.J.; Poppenberger, B.; Rozhon, W. Toxin–antitoxin systems: Biology, identification, and application. *Mob. Genet. Elem.* **2013**, *3*, e26219. https://doi.org/10.4161/mge.26219.

26. Song, S.; Wood, T.K. A primary physiological role of toxin/antitoxin systems is phage inhibition. *Front. Microbiol.* **2020**, *11*, 1895. https://doi.org/10.3389/fmicb.2020.01895.

27. Stern, A.; Sorek, R. The phage-host arms race: Shaping the evolution of microbes. *Bioessays* **2011**, *33*, 43–51. https://doi.org/10.1002/bies.201000071.

28. Rauch, B.J.; Silvis, M.R.; Hultquist, J.F.; Waters, C.S.; McGregor, M.J.; Krogan, N.J.; Bondy-Denomy, J. Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell* **2017**, *168*, 150–158.e10. https://doi.org/10.1016/j.cell.2016.12.009.

29. Otsuka, Y.; Yonesaki, T. Dmd of Bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. *Mol. Microbiol.* **2012**, *83*, 669–681. https://doi.org/10.1111/j.1365-2958.2012.07975.x.

30. Wei, Y.; Gao, Z.; Zhang, H.; Dong, Y. Structural characterizations of phage antitoxin Dmd and its interactions with bacterial toxin RnlA. *Biochem. Biophys. Res. Commun.* **2016**, *472*, 592–597. https://doi.org/10.1016/j.bbrc.2016.03.025.

31. Yamaguchi, Y.; Inouye, M. Regulation of growth and death in *Escherichia coli* by toxin–antitoxin systems. *Nat. Rev. Microbiol.* **2011**, *9*, 779–790. https://doi.org/10.1038/nrmicro2651.

32. Butt, A.; Higman, V.A.; Williams, C.; Crump, M.P.; Hemsley, C.M.; Harmer, N.; Titball, R.W. The *hicA* toxin from *Burkholderia pseudomallei* has a role in persister cell formation. *Biochem. J.* **2014**, *459*, 333–344. https://doi.org/10.1042/BJ20140073.

33. Li, G.; Shen, M.; Lu, S.; Le, S.; Tan, Y.; Wang, J.; Zhao, X.; Shen, W.; Guo, K.; Yang, Y.; et al. Identification and characterization of the *hicAB* toxin-antitoxin system in the opportunistic pathogen *Pseudomonas aeruginosa*. *Toxins* **2016**, *8*, 113. https://doi.org/10.3390/toxins8040113.

34. Pope, W.H.; Mavrich, T.N.; Garlena, R.A.; Guerrero-Bustamante, C.A.; Jacobs-Sera, D.; Montgomery, M.T.; Russell, D.A.; Warner, M.H.; Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES); Hatfull, G.F. Bacteriophages of *Gordonia* spp. display a spectrum of diversity and genetic relationships. *mBio* **2017**, *8*, e01069-17. https://doi.org/10.1128/mBio.01069-10.

35. Lawrence, J.G.; Hatfull, G.F.; Hendrix, R.W. Imbroglios of viral taxonomy: Genetic exchange and failings of phenetic approaches. *J. Bacteriol.* **2002**, *184*, 4891–4905. https://doi.org/10.1128/JB.184.17.4891-4905.2002.

36. Ford, M.E.; Sarkis, G.J.; Belanger, A.E.; Hendrix, R.W.; Hatfull, G.F. Genome structure of Mycobacteriophage D29: Implications for phage evolution 1. Edited by J. Karn. *J. Mol. Biol.* **1998**, *279*, 143–164. https://doi.org/10.1006/jmbi.1997.1610.

37. Hatfull, G.F.; Pedulla, M.L.; Jacobs-Sera, D.; Cichon, P.M.; Foley, A.; Ford, M.E.; Gonda, R.M.; Houtz, J.M.; Hryckowian, A.J.; Kelchner, V.A.; et al. Exploring the mycobacteriophage metaproteome: Phage genomics as an educational platform. *PLoS Genet.* **2006**, *2*, e92. https://doi.org/10.1371/journal.pgen.0020092.

38. Pedulla, M.L.; Ford, M.E.; Houtz, J.M.; Karthikeyan, T.; Wadsworth, C.; Lewis, J.A.; Jacobs-Sera, D.; Falbo, J.; Gross, J.; Pannunzio, N.R.; et al. Origins of highly mosaic mycobacteriophage genomes. *Cell* **2003**, *113*, 171–182. https://doi.org/10.1016/S0092-8674(03)00233-2.

39. Moura de Sousa, J.A.; Pfeifer, E.; Touchon, M.; Rocha, E.P.C. Causes and consequences of bacteriophage diversification via genetic exchanges across lifestyles and bacterial taxa. *Mol. Biol. Evol.* **2021**, *38*, 2497–2512. https://doi.org/10.1093/molbev/msab044.

40. Filée, J.; Bapteste, E.; Susko, E.; Krisch, H.M. A selective barrier to horizontal gene transfer in the T4-type bacteriophages that has preserved a core genome with the viral replication and structural genes. *Mol. Biol. Evol.* **2006**, *23*, 1688–1696. https://doi.org/10.1093/molbev/msl036.

41. Dekel-Bird, N.P.; Avrani, S.; Sabehi, G.; Pekarsky, I.; Marston, M.F.; Kirzner, S.; Lindell, D. Diversity and evolutionary relationships of T7-like podoviruses infecting marine cyanobacteria. *Environ. Microbiol.* **2013**, *15*, 1476–1491. https://doi.org/10.1111/1462-2920.12103.

42. Mesyanzhinov, V.V.; Robben, J.; Grymonprez, B.; Kostyuchenko, V.A.; Bourkaltseva, M.V.; Sykilinda, N.N.; Krylov, V.N.; Volckaert, G. The genome of bacteriophage ΦKZ of *Pseudomonas aeruginosa*. *J. Mol. Biol.* **2002**, *317*, 1–19. https://doi.org/10.1006/jmbi.2001.5396.

43. Sharma, S.; Chatterjee, S.; Datta, S.; Prasad, R.; Dubey, D.; Prasad, R.K.; Vairale, M.G. Bacteriophages and its applications: An overview. *Folia Microbiol.* **2017**, *62*, 17–55. https://doi.org/10.1007/s12223-016-0471-x.

44. Goodfellow, M.; Alderson, G.; Chun, J. Rhodococcal systematics: Problems and developments. *Antonie Leeuwenhoek* **1998**, *74*, 3–20. https://doi.org/10.1023/A:1001730725003.

45. Jacobs-Sera, D.; Marinelli, L.J.; Bowman, C.; Broussard, G.W.; Guerrero Bustamante, C.; Boyle, M.M.; Petrova, Z.O.; Dedrick, R.M.; Pope, W.H.; Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES); et al. On the nature of mycobacteriophage diversity and host preference. *Virology* **2012**, *434*, 187–201. https://doi.org/10.1016/j.virol.2012.09.026.

46. Ross, A.; Ward, S.; Hyman, P. More is better: Selecting for broad host range bacteriophages. *Front. Microbiol.* **2016**, *7*, 1352. https://doi.org/10.3389/fmicb.2016.01352.

47. Ahlgren, N.A.; Ren, J.; Young Lu, Y.; Fuhrman, J.A.; Sun, F. Alignment-free d_2^* oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. *Nucleic Acids Res.* **2016**, *45*, 39–53. https://doi.org/10.1093/nar/gkw1002.

48. Curran, E.; Callaway, S.E.; Dumanlang, R.R.; Harshaw, A.V.; Palacio, P.N.; Nakamura, Y.; Kimberley, K.W.; Theoret, J.R.; Yoon, E.J.; Windsor, E.J.; et al. Genome sequences of *Gordonia rubripertincta* bacteriophages AnarQue and Figliar. *Microbiol. Resourc. Announc.* **2022**, *11*, e01085-21.

49. Erill, I.; Caruso, S.M. Complete genome sequence of *Streptomyces* bacteriophage abt2graduateex2. *Genome Announc*. **2018**, *6*, e01480-17.
50. Han, S.-S.; Kang, H.K.; Jo, B.Y.; Ryu, B.-G.; Jin, H.M.; Chung, E.J.; Jung, J.Y. Complete genome sequence of *Gordonia rubripertincta* SD5, a soil bacterium isolated from a Di-(2-Ethylhexyl) Phthalate-degrading enrichment culture. *Microbiol. Resourc. Announc*. **2020**, *9*, e01087-20.
51. Harrington, D.A.L.; Stevens, J.L.; Johnson, M.J.; Pochiro, S.J.; Moriarty, M.M.; Robertson, M.E.; Sanchez, A.; Whitby, O.G.; Kimberley, K.W.; McKenna, C.C.; et al. Genome sequences of *Gordonia rubripertincta* bacteriophages Jellybones and NHagos. *Microbiol. Resourc. Announc*. **2020**, *9*, e00935-20.
52. Kang, Y.; Yuan, L.; Shi, X.; Chu, Y.; He, Z.; Jia, X.; Lin, Q.; Ma, Q.; Wang, J.; Xiao, J.; et al. A fine-scale map of genome-wide recombination in divergent *Escherichia coli* population. *Brief. Bioinform*. 2021, 22, bbaa335.
53. Lozano-Andrade, C.N.; Strube, M.L.; Kovács, T. Complete genome sequences of four soil-derived isolates for studying synthetic bacterial community assembly. *Microbiol. Resourc. Announc*. **2021**, *10*, e00848-21.
54. Luo, Q.; Hiessl, S.; Poehlein, A.; Daniel, R.; Steinbüchel, A. Steinbüchel. Insights into the microbial degradation of rubber and gutta-percha by analysis of the complete genome of *Nocardia nova* SH22a. *Appl. Environ. Microbiol*. **2014**, *80*, 3895–3907.
55. Pope, W.H.; Bowman, C.A.; Russell, D.A.; Jacobs-Sera, D.; Asai, D.J.; Cresawn, S.G.; Jacobs, W.R.; Jr Hendrix, R.W.; Lawrence, J.G.; Hatfull, G.F.; et al. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife*. **2015**, *4*, e06416.
56. Pope, W.H.; Butela, K.A.; Garlena, R.A.; Jacobs-Sera, D.; Russell, D.A.; Warner, M.H.; Hatfull, G.F.; University of Pittsburgh SEA-PHAGES. Genome sequences of 20 bacteriophages isolated on *Gordonia terrae*. *Microbiol. Resourc. Announc*. **2020**, *9*, e01489-19.
57. Sichtig, H.; Minogue, T.; Yan, Y.; Stefan, C.; Hall, A.; Tallon, L.; Sadzewicz, L.; Nadendla, S.; Klimke, W.; Hatcher, E.; et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. *Nat. Commun*. **2019**, *10*, 3313.
58. Versoza, C.J.; Howell, A.A.; Aftab, T.; Blanco, M.; Brar, A.; Chaffee, E.; Howell, N.; Leach, W.; Lobatos, J.; Luca, M.; et al. The complete genome sequence of the Gordonia bacteriophage BiggityBass. *Microbiol. Resour. Announc.* 2022, *accepted*.