



## OPEN ACCESS

## EDITED BY

Sangryeol Ryu,  
Seoul National University,  
South Korea

## REVIEWED BY

Steven Batinovic,  
Yokohama National University,  
Japan  
Marcin Łoś,  
University of Gdansk,  
Poland

## \*CORRESPONDENCE

George P. C. Salmond  
gpcs2@cam.ac.uk

## †PRESENT ADDRESS

Samuel James Magaziner,  
Center for Human Genetics and Genomics,  
New York University Grossman School of  
Medicine, New York, NY, United States

## SPECIALTY SECTION

This article was submitted to  
Phage Biology,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 03 August 2022

ACCEPTED 10 October 2022

PUBLISHED 31 October 2022

## CITATION

Magaziner SJ and Salmond GPC (2022) A  
novel T4- and  $\lambda$ -based receptor binding  
protein family for bacteriophage therapy  
host range engineering.  
*Front. Microbiol.* 13:1010330.  
doi: 10.3389/fmicb.2022.1010330

## COPYRIGHT

© 2022 Magaziner and Salmond. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# A novel T4- and $\lambda$ -based receptor binding protein family for bacteriophage therapy host range engineering

Samuel J. Magaziner<sup>†</sup> and George P. C. Salmond<sup>\*</sup>

Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Widespread multidrug antimicrobial resistance in emerging pathogens has led to a renewed interest in phage therapy as an alternative or supplement to traditional small molecule drugs. The primary limiting factors of phage therapy deployment rest in the narrow host range specificity of phage as well as a poor understanding of many phages' unintended downstream effects on host physiology and microbiota as well as on adverse pathogen evolution. Consequently, this has made assembling well-defined and safe "phage-cocktails" of solely naturally occurring phages labor- and time-intensive. To increase the speed, efficacy, and safety of therapeutic deployment, there is exceptional interest in modulating the host ranges of well-characterized lytic phages (e.g., T4 and T7) by using synthetic strategies to the swap phage tail components, the receptor binding proteins (RBPs) key for host specificity. Here we identify the RBP of the *Citrobacter rodentium* temperate phage  $\Phi$ NP as ORF6. Through bioinformatic and phylogenetic assays, we demonstrate this RBP to be closely related to the known RBPs of T4 and  $\lambda$ . Further investigation reveals a novel, greater than 200 members RBP family with phages targeting several notable human pathogens, including *Klebsiella pneumoniae*, *Escherichia coli* O157:H7, *Salmonella* spp., and *Shigella* spp. With well characterized lytic members, this RBP family represents an ideal candidate for use in synthetic strategies for expanding therapeutic phage host ranges.

## KEYWORDS

bacteriophage, receptor binding proteins, antimicrobial resistance, phage therapy and biotechnology, phage T4, phage  $\lambda$ , *Citrobacter rodentium*

## Introduction

Bacteriophages, or phages, are obligate intracellular parasites of bacteria. With an estimated global  $10^{31}$  phage particles they are thought to be the most abundant biological entities on Earth and important drivers of bacterial host evolution and population dynamics (Salmond and Fineran, 2015). With the arrival of widespread antibiotic resistance in emerging pathogens (Fischbach and Walsh, 2009; Crofts et al., 2017), specific bacterial

targeting and killing with phage therapy, for a means of alternate or joint treatment with wide-spectrum antibiotics, has seen renewed interest with some notable triumphs (Gordillo Altamirano and Barr, 2019; Hatfull et al., 2022; Nick et al., 2022). However, some of the primary challenges facing the effective employment of traditional “cocktail” bacteriophage therapy within a clinical setting are the often highly specific host-range exhibited by a given phage and unintended downstream effects on host physiology and microbiota as well as unintended and adverse pathogen evolution (Sulakvelidze et al., 2001; Lu and Koeris, 2011; Nilsson, 2014). Host-range, defined at the strain level; downstream effects, such as endotoxin release, local inflammatory effect, or off-target microbiota lysis; and temperate or lytic nature can differ from phage to phage, even those targeting the same pathogen. Thus, there exist impracticalities behind the assessment of every newly isolated phage with enough detailed examination for safe, efficient, and speedy deployment in a human system.

Recently, advances in genome engineering and synthetic biology methodologies have provided an alternative to traditional bacteriophage therapy. Instead of employing a “cocktail” of naturally occurring phages of different genomic compositions and with unknown downstream effects, researchers have pursued phage genome engineering techniques (Kiro et al., 2014; Martel and Moineau, 2014) to generate single, well-characterized phages capable of targeting multiple, specific bacterial hosts through exploitation of receptor binding protein (RBP) modularity. Bacteriophage host recognition and binding (adsorption) to a cognate bacterial receptor is the first stage of phage infection and a primary limiting factor for phage propagation and host range determination. Receptors exploited by phages can be any component of the bacterial surface including capsular polysaccharides, residues of the lipopolysaccharide (LPS) in Gram-negatives, flagella, pili, or surface membrane proteins (Silva et al., 2016; Tzipilevich et al., 2017). Some phages employ a dual-receptor model of adsorption, such as phage T4 which requires both specific LPS residues as well as the outer membrane protein, OmpC, to infect (Henning and Jann, 1979; Washizaki et al., 2016). Phages recognize and bind to their host receptors through RBPs. Depending on phage morphology these are most often named tail-spikes, spike proteins, or tail fibers (Silva et al., 2016). The role of RBPs is to facilitate the adsorption and proper orientation of phages on the extracellular surface to initiate infection, with RBP-receptor interactions often resulting in a cascade of structural rearrangements, from baseplate alteration to tail sheath contraction (Hu et al., 2015). Phage RBPs have been shown to be very stable proteins and are highly resistant to proteases and detergents, presumably required for survival and functionality in harsh environments, such as the mammalian intestine. Moreover, RBPs can exhibit tight receptor specificity and affinities comparable to those of antibodies or lectins, making them particularly interesting molecules (Simpson et al., 2016).

Earlier studies focusing on the T3 and T7 phage families, successfully conferred expanded phage host ranges to lytic phage virions by the swapping of RBP tail fiber regions for phages

targeting pathogenic *Klebsiella* spp. and *Yersinia* spp. (Ando et al., 2015; Yosef et al., 2017). Other strategies, such as those explored by Ram, Ross, Novick and colleagues, focused on developing “antibacterial drones,” engineered and non-replicative phage particles lacking the ability to package endogenous phage DNA, to deliver antibacterial staphylococcal pathogenicity island cargo genes to a cognate bacterial target; this approach both cleared *Staphylococcus aureus* in mice infection while simultaneously mitigating unintended downstream effects of more traditional biologically-active phage therapies (Ram et al., 2018). While successful, expansion of these approaches to other pathogens or even strains of the same species would likely require a “ground-up” rebuild due to limited tools for enabling phage host-range versatility. Indeed, for engineered phage-based therapies to respond effectively to the constantly evolving pathogen landscape, these methodologies will have to be expanded to cover more pathogenic strains and, consequently, will require finer tuned phage RBP families for host range modulation. Here we identify such an RBP family based in the well characterized T4, Tula, Tulb,  $\lambda$ , and, here identified,  $\Phi$ NP lineage.

Traditionally, *de novo* identification of phage RBPs through known-RBP amino acid or DNA sequence homology analysis has proven difficult. This is in part due to the highly specific nature of each RBP-receptor pairing, phage modularity, the rapid evolution of phage genomes mediated by hyper-modification of phage DNA bases, and the sheer impracticality of examining every new phage at the structural level (Warren, 2003; Lima-Mendez et al., 2011). However, previously, we provided an in-depth characterization of the prophages of the non-motile, Gram-negative, enteric bacterium *Citrobacter rodentium* (Petty et al., 2011; Magaziner et al., 2019). This bacterium is a natural host-adapted intestinal mouse pathogen, causative agent of transmissible murine colonic hyperplasia, and important model organism for the study of enteric pathogens of the attaching and effacing (A/E) family (Barthold et al., 1978; Schauer and Falkow, 1993; Mundy et al., 2004; Wong et al., 2011; Collins et al., 2014). Previously, we showed that the ten prophages (complete and incomplete) of this bacterium represent a conserved family of horizontally-acquired mobile genetic elements associated with enteric evolution towards pathogenicity, playing a key role in horizontal gene transfer (HGT) within the enteric environment and carrying several virulence-associated cargo genes. We also showed that of these ten prophages, two, named  $\Phi$ NP and  $\Phi$ SM, actively excise spontaneously to form functional temperate phage particles. Furthermore, we showed that  $\Phi$ NP and  $\Phi$ SM utilize residue-specific lipopolysaccharide (LPS) components for adsorption and infection with  $\Phi$ NP utilizing the Glucose-II (Glc-II, catalyzed by WaaO) residue and  $\Phi$ SM utilizing the Glucose-III (Glc-III, catalyzed by WaaR) residue of the outermost LPS core sugars (Magaziner et al., 2019).

Here, we employ these prior insights and a suite of bioinformatics assays to identify a structural protein (ORF6) of  $\Phi$ NP showing extensive protein-level homology to the known RBPs of T4 and  $\lambda$ . Through adsorption kinetics assays and use of

TABLE 1 Bacterial strains used in this study.

Strain	Relevant Characteristics	Source/Reference
<i>Citrobacter rodentium</i>		
ICC168	Reference strain	Petty et al. (2011)
<i>Escherichia coli</i>		
MG1655	K-12 derivative; F <sup>-</sup> , λ <sup>-</sup> , <i>rph-1</i>	Blattner et al. (1997)
ER2507	K-12 derivative, F <sup>-</sup> , <i>ara-14 leuB6 flnA2 Δ(argF-lac)U169 lacY1 glnV44 galK2 rpsL20 xyl-5 mtl-5 Δ(malB) zjc::Tn5(KanR) Δ(mcrC-mrr)HB101</i>	New England Biotech
DH5α	K-12 derivative, F <sup>-</sup> , <i>Φ80lacZΔM15, Δ(lacZYA-argF)U169, endA1, recA1, hsdR17 (rk-mk+), deoR, thi-1, supE44, λ<sup>-</sup>, gyrA96, relA1</i>	Invitrogen
ER2507(ΦNP)	ER2507 ΦNP lysogen	Magaziner et al. (2019)
ER2507(ΦSM)	ER2507 ΦSM lysogen	Magaziner et al. (2019)
BW25113	K-12 derivative, <i>rrnB3, ΔlacZ4787, hsdR514, Δ(araBAD)567, Δ(rhaBAD)568, rph-1</i>	Baba et al. (2006)
JW3602-1	BW25113 <i>ΔwaaO::kan</i>	Baba et al. (2006)

TABLE 2 Bacteriophages used in this study.

Phage	Description	Origin	Reference
ΦNP	Temperate phage of <i>C. rodentium</i>	<i>C. rodentium</i> supernatant	Petty et al. (2011)

a sfGFP-tagged version of this structural protein and predicted tail fiber, we show that ORF6 is the putative RBP of ΦNP. We also show that the known ΦNP, T4, Tu1a, Tu1b, and Ur-λ RBPs all belong to a conserved family of tail fibers represented in the genomes of phages of several notable human pathogens including *Klebsiella pneumoniae*, *Escherichia coli* O157:H7, *Yersinia pestis*, *Salmonella* spp., and *Shigella* spp. While earlier literature focusing on the structure of the RBP of phage T4 has suggested the possible existence of such a family (Bartual et al., 2010), identification and verification of the RBP of ΦNP as one its members suggests that this greater, here characterized, family of RBPs represents an ideal candidate for use in the design and engineering of therapeutic bacteriophage technologies with expanded host ranges.

## Materials and methods

### Bacterial strains and culture conditions

The bacterial strains used in this study are listed in Table 1. Strains of *Citrobacter rodentium* and *Escherichia coli* were grown at 37°C. Overnight cultures were grown in 5 ml Luria broth (LB) in sterile 25 ml culture tubes placed on a rotary wheel. Bacterial growth was determined by measuring the optical density of the culture at a wavelength of 600 nm (OD<sub>600</sub>) using a Unicam Helios spectrophotometer and cuvettes with a 1 cm path length. Solid medium contained 1.5% weight by volume (w/v) agar with soft medium overlay (top agar) using 0.35% agar; both were made with LB unless otherwise noted. For long term storage, 800 μl of overnight cultures were mixed with 200 μl 80% (w/v) vacuum-sterilized glycerol. Samples were briefly vortexed, appropriately

labelled, and stored at -80°C. Phage buffer was composed of 10 mM Tris/HCl pH 7.4, 10 mM MgSO<sub>4</sub>, and 0.01% gelatin. Phage used in this study are listed in Table 2.

### Generation of high titer phage stock

A high titer phage sample was obtained by collecting top lawns displaying near confluent lysis into a glass universal container. The plate surface was then washed with 3 ml phage buffer and mixed with harvested top agar in the glass universal. Next, 500 μl of chloroform was added and the sample vortexed vigorously for 5 min. Following a sitting incubation period of 15–20 min at room temperature, the sample was spun down at 2220 × g and 4°C for 20 min. The resultant supernatant was then removed, transferred to a new glass bijoux container, mixed with a drop of chloroform to ensure sample sterility, and stored at 4°C until use. These steps were then repeated until a viral stock yielded a titer of > 10<sup>9</sup> pfu/ml and then purified.

### Phage purification

This protocol was undertaken as previously described (Boulanger, 2009). In brief, PEG-8000 was used to precipitate phage particles overnight from a high titer lysate. 5–6 confluent top lawns containing the phage of interest were harvested, extracted with chloroform and incubated with DNase I (1 μg/ml) and RNase A (1 μg/ml) for 30 min at room temperature. NaCl was dissolved into the lysate to the concentration of 0.5M and left to cool for 1 h at 4°C. Following centrifugation at 2220 × g for 10 min and 4°C, the crude lysate was filter sterilized and PEG-8000 was added to a final concentration of 10% and mixed until dissolved. This mixture was kept overnight at 4°C. The precipitated phage was then gently pelleted, supernatant removed, pellet resuspended in phage buffer (10% of original volume), and extracted with chloroform to yield PEG purified phage lysate. PEG precipitated phage lysate was further purified via a step gradient CsCl centrifugation followed by an CsCl isopycnic

TABLE 3 Oligonucleotides used in this study.

Name	Sequence (5'-3')	Comments	References
PF106	GACCACACGTCGACTAGTGCNNNNNNNNNAGAG	RP-PCR primer 1	Fineran et al. (2005)
PF107	GACCACACGTCGACTAGTGCNNNNNNNNNACGCC	RP-PCR primer 2	Fineran et al. (2005)
PF108	GACCACACGTCGACTAGTGCNNNNNNNNNGATAC	RP-PCR primer 3	Fineran et al. (2005)
PF109	GACCACACGTCGACTAGTGC	RP-PCR adapter primer	Fineran et al. (2005)
oREM7	CTAGAGTCGACCTGCAGGC	pDS1028 replicon clone sequencing primer	Monson et al. (2015)
SM.P45	TGCAATCTAAACTAGTAACATGCGTAAAGCGAAG	f; sfGFP addition of <i>SpeI</i> cutsite	
SM.P46	CAATTTTGGAAATTCACCAGAACC CGCGCAGAACCCG CAGAACCTTTGTACAGTTCATCCATACC	r; sfGFP addition of Linker and <i>EcoRI</i> cutsite	
SM.P47	GCGGGTCTGTTGAATTCCAAAAAATGGAGATCTCA	f; $\Phi$ NP ORF 6 addition of Linker and <i>EcoRI</i> cutsite	
SM.P49	TCGATATCAAGCTTTTATGCAAGCCTCACGAT	r; $\Phi$ NP ORF 6 addition of <i>HindIII</i> cutsite	
SM.P50	GCTTGCATAAAAGCTTGAAGGCATCAAATAAACGAAAG	f; pZA11 addition of <i>HindIII</i> cutsite	
SM.P51	TAAGTCTAGTTACTAGTTGATTTCTCCTCTTGTGC	r; pZA11 addition <i>SpeI</i> cutsite	

TABLE 4 Plasmids used in this study.

Plasmid	Relevant Characteristics	References
pZA11	High copy Ap <sup>r</sup> , sfGFP expressing plasmid	The Wang Lab, Columbia University
pBAD33	Arabinose-inducible Cm <sup>r</sup> plasmid	Guzman et al. (1995)
pZSM1 (sfGFP-L)	pZA11 derived plasmid containing a sfGFP and flexible linker sequence (N'-GSAGSAAGSGEF-C') under a tetracycline repressible promoter pL-TetO; Ap <sup>r</sup>	This study (Chen et al., 2013)
pZSM6b (sfGFP-L-ORF6)	pZSM1 derivative containing $\Phi$ NP ORF 6 sfGFP fusion construct	This study

equilibrium centrifugation step (gradient layers of 1.20, 1.30, 1.40, 1.50, and 1.70 g/ml; isopycnic gradient of 1.50 g/ml). Opalescent bands of phage were aspirated *via* syringe. CsCl was removed from the purified phage solution *via* dialysis with a 10 kDa MW cutoff membrane overnight at 4°C in 10,000x volume phage buffer. All ultracentrifugation steps were performed at 35,000 rpm (~150,000 × g) carried out in a Beckman Coulter Optima L-100 XP ultracentrifuge with a SW-40ti swinging bucket rotor.

## Phage genomic DNA extraction

Phage DNA was extracted using a phenol-chloroform protocol. In a phase-lock gel (PLG) tube, 450 µl of high titer phage lysate was incubated with 4.5 µl of 1 mg/ml DNase I and 2.5 µl of 10 mg/ml RNase A and incubated at 37°C for 30 min. The mixture was then added to 11.5 µl of 20% SDS and 4.5 µl of 10 mg/ml Proteinase K and incubated for another 30 min. DNA was extracted by adding 500 µl of a phenol:chloroform:isoamyl alcohol 25:24:1 mix and centrifuged at 1500 × g for 5 min. The supernatant was transferred to a new PLG tube and the previous step repeated. In a new PLG tube, the supernatant was supplemented with 500 µl of chloroform:isoamyl alcohol 24:1 and centrifuged at 1,500 × g for 5 min. The aqueous phase at the top was then incubated with 45 µl sodium acetate (3 mol/l, pH 5.2) and 500 µl of 100% isopropanol at room temperature for 45 min. The mixture was then subjected to centrifugation at 12,000 × g for 20 min, after which the pellet was washed at least twice with 70% ethanol and then re-suspended in dH<sub>2</sub>O.

## Phage chemical mutagenesis

Chemical mutagenesis was conducted utilizing hydroxylamine containing phosphate-EDTA buffer as previously described (Villafane, 2009).

## DNA manipulations, oligonucleotides, and sequencing

Unless otherwise stated, standard molecular biological methods were used for all DNA manipulations. Genomic and plasmid DNA were purified using the GeneJET Genomic DNA Purification Kit (Thermo Scientific) and GeneJET Plasmid Miniprep kit (Thermo Scientific) according to manufacturers' instructions. All restriction enzymes used were obtained from New England Biolabs and used according to manufacturer's protocols. DNA fragments were ligated using T4 DNA ligase (NEB). Oligonucleotides were obtained from Sigma Aldrich and are listed in Table 3. DNA sequencing of PCR and plasmid products was performed by GATC Biotech utilizing their Lightrun Tube-Barcode Sanger Sequencing.

## Generation of plasmids used in this study

The plasmids used in this study are listed in Table 4. pZSM1 was generated as a derivative of the pZA11 background, adding, in front of the native pL-TetO Tc repressible promoter and RBS a *SpeI* cutsite

and a terminal *Hind*III cutsite using primer SM.P51 and SM.P50, respectively. Next, the sfGFP element was extracted from pZA11 with primers SM.P45 and SM.P46 adding a 5' *Spe*I cutsite and a 3' flexible amino acid linker sequence (N'-GSAGSAAGSGEF-C') with a 3' *Eco*RI and *Hind*III cutsite. These two elements were then digested, ligated, and cloned to form vector pZSM1. Next,  $\Phi$ NP ORF6 was extracted from phage gDNA with primers SM.P47 and SM.P49 adding a 5' *Eco*RI cutsite and a 3' *Hind*III cutsite, respectively. Lastly, ORF6 and pZSM1 were digested using *Eco*RI and *Hind*III, ligated, and cloned to yield vector pZSM6b. Both vectors were sequence verified and assessed under light microscopy to verify the intact nature of the sfGFP marker.

## Fluorescence binding assay of sfGFP-L-ORF6

Overnight cultures (5 ml) containing sfGFP-L-ORF6 and sfGFP-L were either chloroform treated or sonicated for a 1 min at 15 micron amplitude with a Soniprep 150 (MSE). Chloroform treated samples were observed immediately by fluorescent light microscopy. Sonicated lysates were spun down at 2220 $\times$ g and 4°C for 10 min, followed by syringe-driven 0.22  $\mu$ m filter sterilization prior to use. 100  $\mu$ l of test and control crude protein extract was added to 900  $\mu$ l of WT ER2507 cells normalized to an OD<sub>600</sub> of 1.0 and samples taken every 5 min to be analyzed by fluorescence and light microscopy. Sonicated lysates were also plated on LB-agar plates and allowed to grow up overnight at 37°C to ensure observed fluorescence was a product of tail fiber binding and not exogenous, sfGFP+ bacteria.

## Adsorptions assays

Triplicate overnight cultures of strains to be tested were grown up in 5 ml of LB on a tube roller at 37°C. Pre-made Eppendorf tubes containing 900  $\mu$ l phage buffer and 30  $\mu$ l chloroform were labelled and set out for each sample and time point to be collected. To each sample of 5 ml overnight bacterial culture, a combination of either (1)  $\Phi$ NP + LB, (2)  $\Phi$ NP + ER2507, (3)  $\Phi$ NP + ER2507 + sfGFP-L lysate, (4)  $\Phi$ NP + sfGFP-L lysate, (5)  $\Phi$ NP + ER2507 + sfGFP-L-ORF6 lysate, or (6)  $\Phi$ NP + sfGFP-L-ORF6 lysate (MOI of 0.01) and mixed immediately. From these mixtures 100  $\mu$ l was removed for time point 0 min and added to the pre-made sample tubes and quickly vortexed for 5 s. Sample collection occurred in a similar fashion for the next 55 min every 5–10 min. Vortexed samples were then spun down and supernatant removed. These supernatants were then titrated in serial dilution on bacterial top lawns. The final adsorption curve was plotted by calculating the percentage of free phages in the culture against time.

## Bioinformatic analysis

Coding sequences and ORFs were determined by a combination of prior annotations and the Geneious R7 predictive

ORF function on known *C. rodentium* ICC168 genome and prophage sequences (accession number: NC\_013716; Kearsse et al., 2012). Protein functionality and homology were predicted using BLASTP.<sup>1</sup> Phage protein products of interest were initially identified using blastp, constraining search results to known viral sequences (taxid: 10239). Genome comparisons were generated by the Artemis Comparison Tool (ACT) (Carver et al., 2005) and EasyFigure (Sullivan et al., 2011).

## Phylogenetic analysis

Protein alignments were generated using ClustalOmega (Sievers et al., 2011). Phylogenetic trees were constructed by the maximum-likelihood method under the LG model (Le and Gascuel, 2008) with PhyML (Guindon and Gascuel, 2003) and verified using Bayesian likelihood using Blosom matrix support (Henikoff et al., 1992) plus Gamma model with MrBayes (Huelsenbeck and Ronquist, 2001). PhyML branch support was tested using 1,000 bootstrap replicates with Bayesian modelling further corroborating PhyML branches.

## Results

### $\Phi$ NP ORF6 shares high similarity with the known T4 RBP (gp37)

To identify the potential RBPs of  $\Phi$ NP and  $\Phi$ SM, we utilized the recently reannotated  $\Phi$ NP and  $\Phi$ SM genomes (Magaziner et al., 2019) and extracted all coding regions predicted to encode structural proteins. We next took the translated amino acid gene products of these coding regions and searched for notable sequence homology, both within non-specified search parameters as well as within the viral NCBI database (taxid: 10239). Translated coding regions demonstrating protein-level homology to proteins with known or predicted non-RBP functionalities (such as capsid, sheath, or portal proteins) were removed from the analysis. As  $\Phi$ NP and  $\Phi$ SM are members of the *Myoviridae* and *Siphoviridae* families, respectively, their RBPs would be expected to be tail fibers or spikes. While no significant tail spike homology was found for the  $\Phi$ SM structural comparison, one coding region of  $\Phi$ NP, ORF6 displayed striking homology to a 215 AA stretch of the tail fiber protein product encoded by the phage T4/Tula/Tulb coding regions of gp37 (Figure 1A). ORF6 is encoded antisense in the  $\Phi$ NP prophage genome from base pair 3,252 to 2,404; this corresponds to base pairs 2,733,682 to 2,732,833 in the *Citrobacter rodentium* ICC168 genome (Genbank accession: FN543502.1).

<sup>1</sup> <https://blast.ncbi.nlm.nih.gov/Blast.cgi>



## ORF6 encodes the RBP of $\Phi$ NP

To determine whether ORF6 was indeed the RBP of  $\Phi$ NP, ORF6 was recombinantly expressed with a superfolder GFP (sfGFP) fusion tag, linked with a flexible linker sequence (N'-GSAGSAAGSGEF-C'; sfGFP-L-ORF6; Figure 2A; Chen et al., 2013). As a control, an identical vector and linker sequence, but lacking the ORF6 insertion, was utilized (sfGFP-L).

$\Phi$ NP host range has been shown to include *E. coli* K12 strain ER2507 (Magaziner et al., 2019). To examine ORF6 as the potential RBP of  $\Phi$ NP, the ability of sfGFP-L-ORF6 to localize to the extracellular surface of wild type (WT) ER2507 in culture was examined (Figures 2A–E). A crude protein extract was mixed with an ER2507 WT liquid culture and fluorescence localization was visualized over a time course of 55 min (Figure 2B). While no detectable fluorescence localization was seen between 0 and 20 min (Figure 2C), after 25 min (Figure 2E), most cells within the test culture displayed an apparent fluorescent localization. Notably, fluorescent localization was absent in the sfGFP-L control (Figures 2D,F). Previously we had identified the  $\Phi$ NP bacterial surface receptor as the lipopolysaccharide (LPS) residue GlcII, a residue dependent on the catalytic activity of the *waaO* gene in the core LPS biosynthesis cascade (Magaziner et al., 2019). As expected, no fluorescent localization was observed in tests conducted utilizing sfGFP-L-ORF6 with the  $\Phi$ NP receptor knockout strain ( $\Delta waaO$ ; data not shown).

Lastly, the effect of sfGFP-L-ORF6 on the adsorption of  $\Phi$ NP was examined. Were ORF6 the RBP of  $\Phi$ NP, it would likely compete with active phage virions for extracellular binding sites, lowering overall adsorption kinetics and total phage binding. An adsorption test was performed under standard conditions with or without the addition of a sfGFP-L-ORF6 crude protein extract. In the presence of GFP-L lysate, phage adsorption was consistent with WT (ER2507 only) kinetics and observed to plateau near 25 min (Figure 2G), consistent with observation of fluorescence localization of the sfGFP-L-ORF6 lysate. However, the presence of the sfGFP-L-ORF6 in the adsorption assay was found to both decrease the rate of adsorption as well as the total amount of adsorbed phage in the supernatant (Figure 2G). In addition, attempts to isolate ORF6 KO  $\Phi$ NP mutants through chemical mutagenesis proved not possible, suggesting it as an essential protein within the  $\Phi$ NP infection pathway. Taken together, these results are consistent with the view that ORF6 encodes the putative RBP of  $\Phi$ NP.

## Identification of a novel enteric RBP and RBP assembly family

Further bioinformatic interrogation revealed high conservation between a ~215 AA region of the tail fiber encoding ORF314 of phage Ur- $\lambda$  and ORF6. In addition, protein alignments suggested that  $\Phi$ NP, as well as the putative tail proteins of over 200 other phage including the *Klebsiella pneumoniae* phage ST405-OXA-48phi1.1, the *E. coli* O157:H7 typing phage 3, the *Yersinia pestis* phage  $\Phi$ D1, the *Salmonella* spp. phage SEN5, and the

*Shigella* spp. phage SHBML-50-1 are members of this same lineage (Figures 3A,B). A full table of hits can be found in Supplementary Table 1 (note: there are 505 homology hits for over 200 unique bacteriophage due to repeat hits). Notably, protein level homology was only seen within the last ~215 AAs of each phage tail fiber. Moreover, sequence divergence was seen in the same region of the T4 LTF shown to be responsible for its OmpC/LPS receptor specificity (Montag et al., 1990).

Furthermore, when we extended protein-level comparison to the immediate upstream and downstream structural modules of  $\Phi$ NP, ORF5 of  $\Phi$ NP, predicted to encode a structural assembly protein, showed striking homology to the T4 protein gp38. Indeed, this conservation was present within each of the other phages of the RBP lineage which we examined prior (Figures 4A–D). Earlier studies have shown gp38 to be an essential tail fiber assembly chaperone responsible for mediating the correct trimeric folding of gp37 (Beckendorf et al., 1973; Hashemolhosseini et al., 1996; Leiman et al., 2010). As such, it seems likely that versatility and modularity of this family of RBP tips is, at least in part, due to a conserved assembly protein found directly upstream of the RBP coding region (Figure 4D). When phylogenetically examined, two clear families of assembly proteins emerged: (1) one containing  $\Phi$ NP, SEN5, Ur- $\lambda$ , and ST405-OXA-48phi1.1 (Figures 4A,C) and (2) a more recently evolved family containing T4, Tula, Tulb, the O157:H7 typing phage 3,  $\Phi$ D1, and SHBML-50-1 (Figures 4B,C). Excepting Ur- $\lambda$ , with whole genome alignment suggesting large modules of  $\Phi$ NP morphogenesis (ORFs 5–7, 21 and, 23–26) are of Ur- $\lambda$  descent (Magaziner et al., 2019), low translational similarity exists between the remainder of the whole phage genomes of  $\Phi$ NP to other RBP family members (Supplementary Figure 1). We examined the other structural ORFs of several dozen of the phages within this RBP family and found no further homology at any level, including auxiliary tail fiber domains. This would suggest that this RBP and tail fiber assembly protein lineage is not constrained functionally by phage morphology or unique structural proteins motifs but represents an advantageous and conserved adsorption module in phage evolution.

## Discussion

The advent of widespread multidrug antibacterial resistance in emerging pathogens has led to a renewed interest in therapeutic alternatives to small molecule drugs. One such revitalized area of focus has been in bacteriophage, or phage, therapy. The primary limitations of phage therapy stem from the incredible host range specificity displayed by phages; in turn this leads to large “cocktails” of therapeutic phages and lengthy screening periods required to treat a small range of pathogenic strains. With impracticalities behind rigorous characterization of each naturally occurring phage used in each “cocktail,” there can often be unforeseen downstream consequences on host physiology and microbiota as well as risks of adverse pathogen evolution in the event of temperate bacteriophage inclusion within the cocktail. To combat this, recent approaches have focused on generating single,

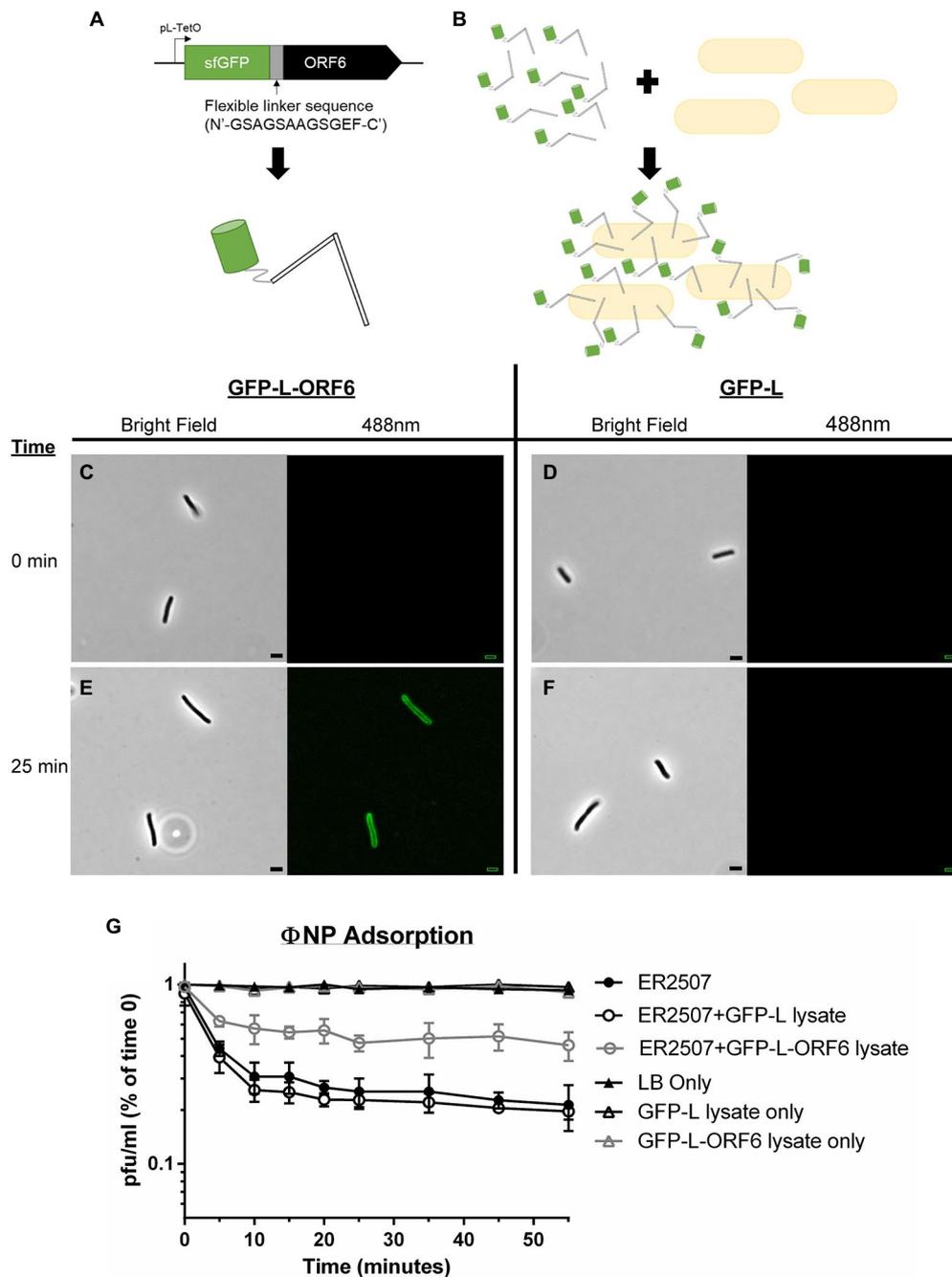


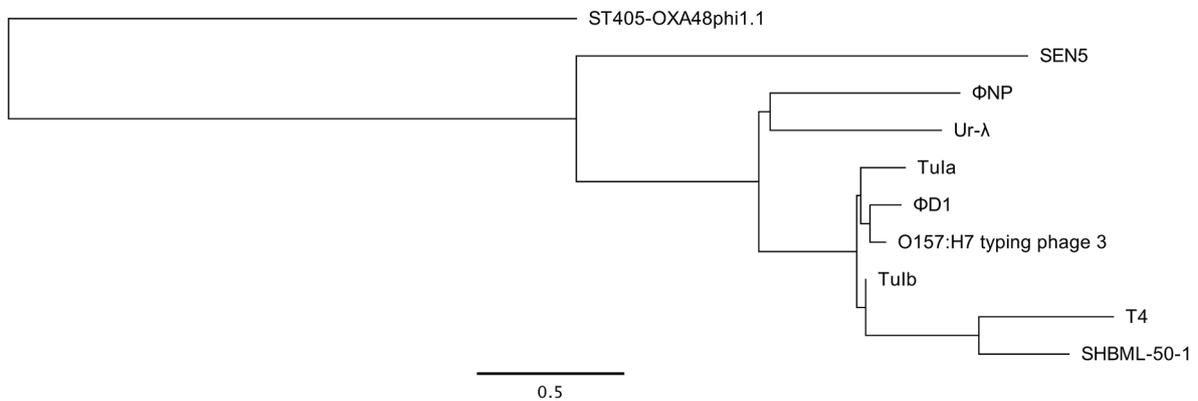
FIGURE 2

Adsorption kinetics and fluorescence assays show that ORF6 encodes the putative RBP of  $\Phi$ NP. (A) A cartoon depicting the construct utilized to assess ORF6 binding. ORF6 was recombinantly expressed with a superfolder GFP (sfGFP) fusion tag, linked with a flexible linker sequence (N'-GSAGSAAGSGEF-C'), and under control of a pL-TetO Tc repressible promoter (named sfGFP-L-ORF6). As a control an identical construct lacking ORF6 was built (named GFP-L). (B) Cartoon depicting fluorescence localization to the surface of host bacteria (not to scale). (C-F) Fluorescence localization assay of crude sfGFP-L-ORF6 lysate (A,E) and control crude GFP-L lysate (D,F) using host bacteria *E. coli* K12 strain ER2507 under both brightfield and 488nm emission. Notably, at 0min (C,D) no background fluorescence is seen in either sample. Following 25min (E,F) fluorescence can be seen localized to the periphery of bacteria in the sample containing crude sfGFP-L-ORF6 lysate (E) while none is seen with samples containing GFP-L crude lysate (F). (G) Representative graph showing the mean results of triplicate adsorption assays of  $\Phi$ NP over 55min onto *E. coli* K-12 strain ER2507 in the presence of no lysate, crude sfGFP-L-ORF6 lysate, or crude sfGFP-L lysate with error bars denoting SD. Also shown are the results of *E. coli* K-12 strain ER2507 in the presence of no lysate, crude sfGFP-L-ORF6 lysate, or crude sfGFP-L lysate with no addition of  $\Phi$ NP. As can be seen, addition of GFP-L-ORF6 lysate to of *E. coli* K-12 strain ER2507 in the presence of  $\Phi$ NP greatly reduces  $\Phi$ NP adsorption kinetics. The left axis shows phage left unadsorbed as a percentage of phage at time 0min in pfu/ml.

**A**



**B**



**FIGURE 3**  
 Expansion of bioinformatic search reveals the RBP of ΦNP to be a member of a larger novel enteric RBP family based in a T4- and λ-lineage. **(A)** Amino acid (AA) alignment by ClustalOmega depicting high similarity between ΦNP ORF6 and the distal 215 AAs of the known T4 RBP (gp37), ORF314 of phage λ, and the distal 215 AAs of uncharacterized structural proteins for the *Klebsiella pneumoniae* phage ST405-OXA-48phi1.1, the *E. coli* O157:H7 typing phage 3, the *Yersinia pestis* phage ΦD1, the *Salmonella* spp. phage SEN5, and the *Shigella* spp. phage SHBML-50-1. Note the sequence similarity with primary divergence centered around the putative receptor specificity domain of phage T4. **(B)** Maximum likelihood tree of the putative RBPs of the newly identified enteric RBP family. Trees were constructed using PhyML with default parameters and the General time reversible (GTR)+Gamma model. Branch support was tested using approximate likelihood ratio test (aLRT) based on the Shimodaira-Hasegawa-like (SH-like) procedures and re-tested using 1,000 bootstrap replicates. Tree structures and root positions were verified by Bayesian phylogenetic analysis using BEAST2 under a GTR substitution model, which yielded results consistent with PhyML. Trees are drawn to scale. Scale bars represent the number of substitutions per site.

well-characterized strictly lytic phages capable of targeting multiple, specific bacterial hosts through exploitation of RBP modularity. Here we identified a new RBP family containing over

200 unique sequenced bacteriophage members including the here elucidated RBP of *Citrobacter rodentium* temperate phage ΦNP as well as the RBPs of phages T4, Ur-λ, the *Klebsiella pneumoniae*

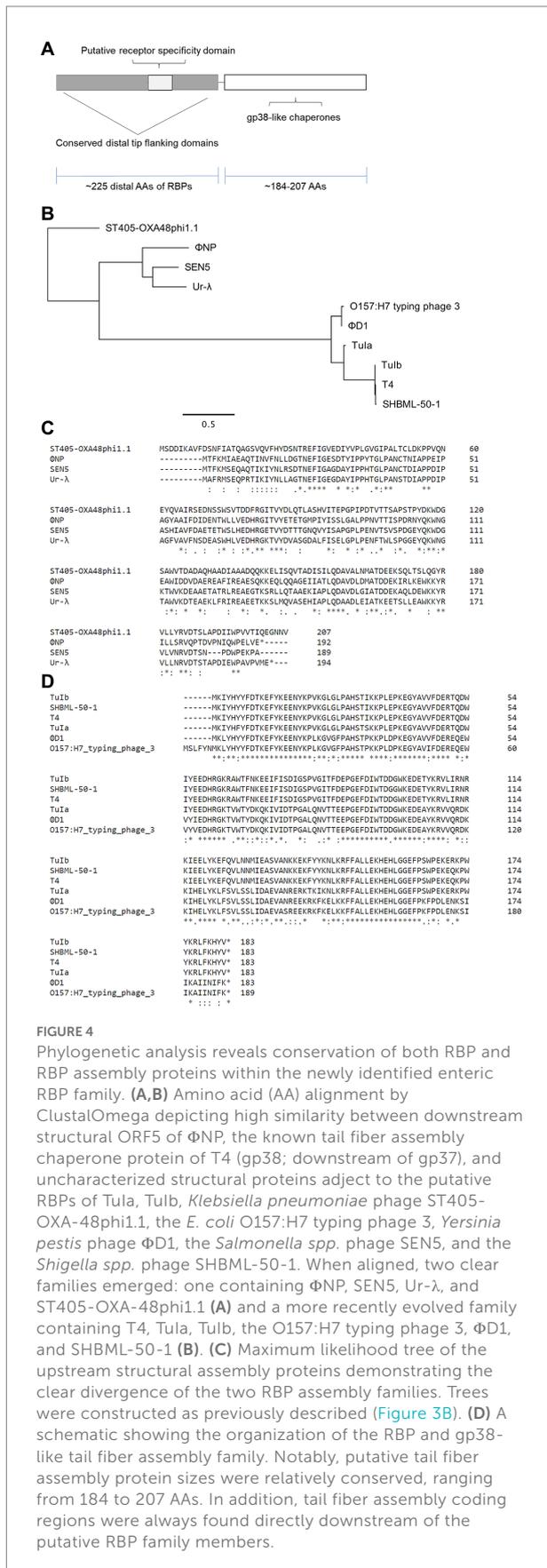


FIGURE 4

Phylogenetic analysis reveals conservation of both RBP and RBP assembly proteins within the newly identified enteric RBP family. (A,B) Amino acid (AA) alignment by ClustalOmega depicting high similarity between downstream structural ORF5 of ΦNP, the known tail fiber assembly chaperone protein of T4 (gp38; downstream of gp37), and uncharacterized structural proteins adjacent to the putative RBPs of Tula, TuIb, *Klebsiella pneumoniae* phage ST405-OXA-48phi1.1, the *E. coli* O157:H7 typing phage 3, *Yersinia pestis* phage ΦD1, the *Salmonella* spp. phage SEN5, and the *Shigella* spp. phage SHBML-50-1. When aligned, two clear families emerged: one containing ΦNP, SEN5, Ur-λ, and ST405-OXA-48phi1.1 (A) and a more recently evolved family containing T4, Tula, TuIb, the O157:H7 typing phage 3, ΦD1, and SHBML-50-1 (B). (C) Maximum likelihood tree of the upstream structural assembly proteins demonstrating the clear divergence of the two RBP assembly families. Trees were constructed as previously described (Figure 3B). (D) A schematic showing the organization of the RBP and gp38-like tail fiber assembly protein. Notably, putative tail fiber assembly protein sizes were relatively conserved, ranging from 184 to 207 AAs. In addition, tail fiber assembly coding regions were always found directly downstream of the putative RBP family members.

phage ST405-OXA-48phi1.1, the *E. coli* O157:H7 typing phage 3, the *Yersinia pestis* phage ΦD1, the *Salmonella* spp. phage SEN5, and the *Shigella* spp. phage SHBML-50-1.

Previously, we had extensively characterized the ten prophages of non-motile, Gram-negative, enteric bacterium *Citrobacter rodentium*, including the temperate phage ΦNP (Magaziner et al., 2019). Here we used these prior insights to identify the uncharacterized structural protein ORF6 of ΦNP as having extensive AA similarity to the known RBP of phage T4, gp37. This was initially quite exciting as T4 represents one of the most, if not the most, well characterized lytic bacteriophages of the last 100 years (Karam and Miller, 2010). Identifying a potential family of RBPs which might serve to expand T4's host range would be incredibly valuable in the effort to design a modulated therapeutic alternative to naturally occurring "phage-cocktails." Through adsorption kinetics assays and use of a sfGFP-tagged version of this structural protein and predicted tail fiber we identified ORF6 to be the putative RBP of ΦNP. Corroborating the essential role of ORF6 in ΦNP infection, attempts *via* chemical mutagenesis to obtain an ORF6 KO were not successful. Notably, this approach of fluorescence localization and adsorption occlusion of WT phage virions in the presence of the tagged tail fiber represents a facile means of examining supposed RBPs of other phages.

We also identified that the putative RBP of ΦNP was closely related to ORF314 of phage Ur-λ. However, it is well known that the RBP of the λ bacteriophage is the tail fiber encoded by gene J which allows binding to its putative receptor LamB (Werts et al., 1994). The early laboratory history of phage λ explains this finding. The "true" wild type prophage λ is designated as Ur-λ; yet, the most common laboratory strain referred to as WT-λ is in fact a variant named PaPa-λ (Pasadena/Paris; Casjens and Hendrix, 2015). PaPa-λ, while possessing the tail fiber encoded by gene J, lacks the short tail fibers (*stf* gene) of Ur-λ; this *stf* gene deletion was first identified by Sanger and colleagues to be the result of a frameshift, generating two truncated reading frames in PaPa-λ ORF401 and ORF314 (Karam and Miller, 2010). Rather interestingly, *stf* has been shown to encode tail fibers that specifically adsorb to LPS residues, not unlike gp37 of phage T4 and ORF6 ΦNP (Sanger et al., 1982; Hendrix and Duda, 1992). It is thus likely that a unifying feature of this RBP family may be LPS specificity (either as secondary to a putative receptor, as with T4 and Ur-λ, or as the sole LPS residue-specific binding site, as with ΦNP). This universality of LPS binding might help explain why this novel RBP family is conserved across such varying bacteriophage.

Further protein level alignments suggested a RBP lineage including structural protein conservation in the sequenced genomes of over 200 other bacteriophages, predominantly of those whose bacterial hosts are of the *Enterobacteriaceae*. Notably, protein level homology was only seen within the last ~215 AAs of each phage structural protein and likely tail fiber, with primary sequence divergence centered around the putative receptor specificity domain of T4 gp37 product LTF (shown to be responsible for T4's OmpC/LPS receptor specificity; Montag et al., 1990). We then extended protein-level comparison to the immediate upstream and downstream structural modules of these phages to the T4 protein gp38, a key chaperone for tail fiber folding. Similar to the tail fiber itself, ΦNP

ORF5- and T4 gp38-like chaperones encoded genomically adjacent to the tail fibers were identified in each RBP family member. This included the well characterized *stf* (tail fiber) and *lyfa* (chaperone) of Ur- $\lambda$ , suggesting that the RBP family included both the tail fiber and tail fiber assembly module. Fascinatingly, an early study showed that the N-terminal region of Ur- $\lambda$  *stf* can functionally substitute for gp37 in receptor binding; similarly the Ur- $\lambda$  tail fiber assembly protein (*ltfa*) can functionally substitute for gp38 in mediating the correct folding of gp37 (Montag et al., 1989). Notably, no other homology with any other tail fiber or structural proteins was observed between RBP family phage at either the nucleotide or amino acid level. In addition, phage genome sizes (and thus likely, corresponding structural sizes and variations) ranged widely from as small as 20kb base pairs to as large as several 100kb base pairs. While modularity is a core feature of bacteriophage evolution and expansion (Lima-Mendez et al., 2011), this was still surprising given the incredibly dissonant whole genome profiles of each member. With the majority family predominantly constrained to enterics, the wide prevalence of this RBP family speaks to its versatility and efficacy within the host and gut environment.

Testament to the overarching limitations present with traditional phage therapy, only 5 ( $\Phi$ NP, T4, TuIa, TuIb, and Ur- $\lambda$ ) of the over 200 members of the RBP family here reported are well characterized with known extracellular bacterial receptors. Interestingly, despite RBP and chaperone conservation,  $\Phi$ NP, T4, TuIa, TuIb, and Ur- $\lambda$  rely on different extracellular receptors for adsorption, LPS residue GlcII, OmpC, OmpF, OmpC/LamB, and LamB, respectively, with all excepting  $\Phi$ NP possessing a secondary, non-essential, irreversible binding affinity for LPS elements (Henning and Jann, 1979; Hendrix and Duda, 1992; Magaziner et al., 2019). These differing receptor specificities are likely explained by the ~30–70 AA unique, receptor-specific domain of each RBS sandwiched by the conserved RBP-tail fiber tip domain outlined in this study.

Taken together with previous studies, our findings highlight a conserved modular domain for bacteriophage host range specificity. It is worth noting, however, that while initial absorption to a bacterial surface often serves as the greatest filter to host range specificity there are a suite of other considerations in the engineering of successful recombinant phage. For example, proper DNA injection and transfer, native bacterial anti-viral defenses (e.g., CRISPR-Cas9 or Toxin-Antitoxin altruistic suicide cassettes) or proper development of phage can all be rate-limiting steps in bacteriophage infection and propagation post-absorption (Pires et al., 2016; Łobocka et al., 2021). Various phage engineering techniques have been pioneered to address these issues including a wide range of phage genome engineering strategies to circumvent native antibacterial resistances or non-compatible genomic islands or GC contents for more efficient phage co-opting of bacterial host replication machinery (Duong et al., 2020; Mahler et al., 2022). This having been said, recombinant expression of the modular RBP and tail fiber assembly lineage reported here might serve to help expand the host range of several very well characterized phages to several clinically important human pathogens for use in therapeutic applications and be of notable interest for phage-based technologies at large.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

SM and GS: conceptualization, methodology, project administration, visualization, writing: original draft preparation, and writing: review and editing. SM: data curation, formal analysis, investigation, and validation. GS: funding acquisition and supervision. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC, UKRI) through awards BB/W000105/1, BB/T006668/1, and PHJZ/744 to GS.

## Acknowledgments

We would like to acknowledge the other members of the Salmond Laboratory, especially Rita Monson and lab alumni Ray Chai and Bihe Chen for their kind and helpful discussion.

## 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.

## 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.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1010330/full#supplementary-material>

## 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. doi: 10.1016/j.cels.2015.08.013
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2:2006.0008. doi: 10.1038/msb4100050
- Barthold, S. W., Coleman, G. L., Jacoby, R. O., Livstone, E. M., and Jonas, A. M. (1978). Transmissible murine colonic hyperplasia. *Vet. Pathol.* 15, 223–236. doi: 10.1177/030098587801500209
- Bartual, S. G., Otero, J. M., Garcia-Doval, C., Llamas-Saiz, A. L., Kahn, R., Fox, G. C., et al. (2010). Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20287–20292. doi: 10.1073/pnas.1011218107
- Beckendorf, S. K., Kim, J. S., and Lielausis, I. (1973). Structure of bacteriophage T4 genes 37 and 38. *J. Mol. Biol.* 73, 17–35. doi: 10.1016/0022-2836(73)90156-3
- Blattner, F. R., Plunkett, G. III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462. doi: 10.1126/science.277.5331.1453
- Boulanger, P. (2009). Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. *Methods Mol. Biol.* 502, 227–238. doi: 10.1007/978-1-60327-565-1\_13
- Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics* 21, 3422–3423. doi: 10.1093/bioinformatics/bti553
- Casjens, S. R., and Hendrix, R. W. (2015). Bacteriophage lambda: early pioneer and still relevant. *Virology* 479–480, 310–330. doi: 10.1016/j.virol.2015.02.010
- Chen, X., Zaro, J. L., and Shen, W. C. (2013). Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* 65, 1357–1369. doi: 10.1016/j.addr.2012.09.039
- Collins, J. W., Keeney, K. M., Crepin, V. F., Rathinam, V. A. K., Fitzgerald, K. A., Finlay, B. B., et al. (2014). *Citrobacter rodentium*: infection, inflammation and the microbiota. *Nat. Rev. Microbiol.* 12, 612–623. doi: 10.1038/nrmicro3315
- Crofts, T. S., Gasparrini, A. J., and Dantas, G. (2017). Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* 15, 422–434. doi: 10.1038/nrmicro.2017.28
- Duong, M. M., Carmody, C. M., Ma, Q., Peters, J. E., and Nugen, S. R. (2020). Optimization of T4 phage engineering via CRISPR/Cas9. *Sci. Reports* 10, 1–9. doi: 10.1038/s41598-020-75426-6
- Fineran, P. C., Everson, L., Slater, H., and Salmond, G. P. C. (2005). A GntR family transcriptional regulator (PigT) controls gluconate-mediated repression and defines a new, independent pathway for regulation of the tripyrrole antibiotic, prodigiosin, Serratia. *Microbiology* 151, 3833–3845. doi: 10.1099/mic.0.28251-0
- Fischbach, M. A., and Walsh, C. T. (2009). Antibiotics for emerging pathogens. *Science* 325, 1089–1093. doi: 10.1126/science.1176667
- Gordillo Altamirano, F. L., and Barr, J. J. (2019). Phage therapy in the postantibiotic era. *Clin. Microbiol. Rev.* 32, e00066–18. doi: 10.1128/CMR.00066-18
- Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520
- Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. *J. Bacteriol.* 177, 4121–4130. doi: 10.1128/jb.177.14.4121-4130.1995
- Hashemolhosseini, S., Stierhof, Y. D., Hindennach, I., and Henning, U. (1996). Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and  $\lambda$ . *J. Bacteriol.* 178, 6258–6265. doi: 10.1128/jb.178.21.6258-6265.1996
- Hatfull, G. F., Dedrick, R. M., and Schooley, R. T. (2022). Phage therapy for antibiotic-resistant bacterial infections. *Annu. Rev. Med.* 73, 197–211. doi: 10.1146/annurev-med-080219-122208
- Hendrix, R. W., and Duda, R. L. (1992). Bacteriophage lambda PaPa: not the mother of all lambda phages. *Science* 258, 1145–1148. doi: 10.1126/science.1439823
- Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* 89, 10915–10919. doi: 10.1073/pnas.89.22.10915
- Henning, U., and Jann, K. (1979). Two-component nature of bacteriophage T4 receptor activity in *Escherichia coli* K-12. *J. Bacteriol.* 137, 664–666. doi: 10.1128/jb.137.1.664-666.1979
- 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
- Huelsensbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754
- Karam, J. D., and Miller, E. S. (2010). Bacteriophage T4 and its relatives. *Virology* 403, 293–303. doi: 10.1016/j.virol.2010.06.028
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kiro, R., Shitrit, D., and Qimron, U. (2014). Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol.* 11, 42–44. doi: 10.4161/rna.27766
- Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067
- Leiman, P. G., Arisaka, F., van Raaij, M. J., Kostyuchenko, V. A., Aksyuk, A. A., Kanamaru, S., et al. (2010). Morphogenesis of the T4 tail and tail fibers. *Virology* 403, 355–365. doi: 10.1016/j.virol.2010.06.006
- Lima-Mendez, G., Toussaint, A., and Leplae, R. (2011). A modular view of the bacteriophage genomic space: identification of host and lifestyle marker modules. *Res. Microbiol.* 162, 737–746. doi: 10.1016/j.resmic.2011.06.006
- Łobocka, M., Dąbrowska, K., and Górski, A. (2021). Engineered bacteriophage therapeutics: rationale, challenges and future. *BioDrugs* 35, 255–280. doi: 10.1007/s40259-021-00480-z
- Lu, T. K., and Koeris, M. S. (2011). The next generation of bacteriophage therapy. *Curr. Opin. Microbiol.* 14, 524–531. doi: 10.1016/j.mib.2011.07.028
- Magaziner, S. J., Zeng, Z., Chen, B., and Salmond, G. P. C. (2019). The prophages of *Citrobacter rodentium* represent a conserved family of horizontally acquired mobile genetic elements associated with enteric evolution towards pathogenicity. *J. Bacteriol.* 201, e00638–18. doi: 10.1128/JB.00638-18
- Mahler, M., Costa, A. R., van Beljouw, S. P. B., Fineran, P. C., and Brouns, S. J. J. (2022). Approaches for bacteriophage genome engineering. *Trends Biotechnol.* 40, 15–28. doi: 10.1016/j.tibtech.2022.08.008 [Epub ahead of print]
- Martel, B., and Moineau, S. (2014). CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages. *Nucleic Acids Res.* 42, 9504–9513. doi: 10.1093/nar/gku628
- Monson, R., Smith, D. S., Matilla, M. A., Roberts, K., Richardson, E., Drew, A., et al. (2015). A plasmid-transposon hybrid mutagenesis system effective in a broad range of enterobacteria. *Front. Microbiol.* 6:1442. doi: 10.3389/fmicb.2015.01442
- Montag, D., Hashemolhosseini, S., and Henning, U. (1990). Receptor-recognizing proteins of T-even type bacteriophages. The receptor-recognizing area of proteins 37 of phages T4 Tula and Tulb. *J. Mol. Biol.* 216, 327–334. doi: 10.1016/S0022-2836(05)80324-9
- Montag, D., Schwarz, H., and Henning, U. (1989). A component of the side tail fiber of *Escherichia coli* bacteriophage  $\lambda$  can functionally replace the re.Pdf. *J. Bacteriol.* 171, 4378–4384. doi: 10.1128/jb.171.8.4378-4384.1989
- Mundy, R., Petrovska, L., Smollett, K., Simpson, N., Wilson, R. K., Yu, J., et al. (2004). Identification of a novel *Citrobacter rodentium* type III secreted protein, Esp I, and roles of this and other secreted proteins in infection. *Infect. Immun.* 72, 2288–2302. doi: 10.1128/IAI.72.4.2288-2302.2004
- Nick, J. A., Dedrick, R. M., Gray, A. L., Vladar, E. K., Smith, B. E., Freeman, K. G., et al. (2022). Host and pathogen response to bacteriophage engineered against mycobacterium abscessus lung infection. *Cells* 185, 1860–1874.e12. doi: 10.1016/j.cell.2022.04.024
- Nilsson, A. S. (2014). Phage therapy—constraints and possibilities. *Ups. J. Med. Sci.* 119, 192–198. doi: 10.3109/03009734.2014.902878
- Petty, N. K., Feltwell, T., Pickard, D., Clare, S., Toribio, A. L., Fookes, M., et al. (2011). *Citrobacter rodentium* is an unstable pathogen showing evidence of significant genomic flux. *PLoS Pathog.* 7:e1002018. doi: 10.1371/journal.ppat.1002018
- Pires, D. P., Cleto, S., Sillankorva, S., Azeredo, J., and Lu, T. K. (2016). Genetically engineered phages: a review of advances over the last decade. *Microbiol. Mol. Biol. Rev.* 80, 523–543. doi: 10.1128/MMBR.00069-15
- Ram, G., Ross, H. F., Novick, R. P., Rodriguez-Pagan, I., and Jiang, D. (2018). Conversion of staphylococcal pathogenicity islands to crispr-carrying antibacterial agents that cure infections in mice. *Nat. Biotechnol.* 36, 971–976. doi: 10.1038/nbt.4203
- Salmond, G. P. C., and Fineran, P. C. (2015). A century of the phage: past, present and future. *Nat. Rev. Microbiol.* 13, 777–786. doi: 10.1038/nrmicro3564

- Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F., and Petersen, G. B. (1982). Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* 162, 729–773. doi: 10.1016/0022-2836(82)90546-0
- Schauer, D. B., and Falkow, S. (1993). Attaching and effacing locus of a *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infect. Immun.* 61, 2486–2492. doi: 10.1128/iai.61.6.2486-2492.1993
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Silva, J. B., Storms, Z., and Sauvageau, D. (2016). Host receptors for bacteriophage adsorption. *FEMS Microbiol. Lett.* 363:fnw002. doi: 10.1093/femsle/fnw002
- Simpson, D. J., Sacher, J. C., and Szymanski, C. M. (2016). Development of an assay for the identification of receptor binding proteins from bacteriophages. *Viruses* 8:17. doi: 10.3390/v8010017
- Sulakvelidze, A., Alavidze, Z., Glenn, J., and Jr, M. (2001). Bacteriophage therapy. *Antimicrob. Agents Chemother.* 45, 649–659. doi: 10.1128/AAC.45.3.649
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). EasyFigure: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Tzipilevich, E., Habusha, M., and Ben-Yehuda, S. (2017). Acquisition of Phage Sensitivity by bacteria through exchange of phage receptors. *Cells* 168, 186–199. doi: 10.1016/j.cell.2016.12.003
- Villafane, R. (2009). Construction of phage mutants. *Methods Mol. Biol.* 501, 223–237. doi: 10.1007/978-1-60327-164-6\_20
- Warren, R. A. J. (2003). Modified bases in bacteriophage DNAs. *Annu. Rev. Microbiol.* 34, 137–158. doi: 10.1146/annurev.mi.34.100180.001033
- Washizaki, A., Yonesaki, T., and Otsuka, Y. (2016). Characterization of the interactions between *Escherichia coli* receptors, LPS and Omp C, and bacteriophage T4 long tail fibers. *Microbiology* 5, 1003–1015. doi: 10.1002/mbo3.384
- Werts, C., Michel, V., Hofnung, M., and Charbit, A. (1994). Adsorption of bacteriophage lambda on the lam B protein of *Escherichia coli* K-12: point mutations in gene J of lambda responsible for extended host range. *J. Bacteriol.* 176, 941–947. doi: 10.1128/jb.176.4.941-947.1994
- Wong, A. R. C., Pearson, J. S., Bright, M. D., Munera, D., Robinson, K. S., Lee, S. F., et al. (2011). Enteropathogenic and enterohaemorrhagic *Escherichia coli*: even more subversive elements. *Mol. Microbiol.* 80, 1420–1438. doi: 10.1111/j.1365-2958.2011.07661.x
- Yosef, I., Goren, M. G., Globus, R., Molshanski-Mor, S., and Qimron, U. (2017). Extending the host range of bacteriophage particles for DNA transduction. *Mol. Cell* 66, 721–728. doi: 10.1016/j.molcel.2017.04.025