A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Molecular study of *Shigella flexneri* Bacteriophage-Host Receptor Interaction for Biocontrol

박테리오파지-숙주 수용체 상호작용 연구 및 이를 이용한 파지의 시겔라 플렉스네리 제어능 향상

February, 2023

Hojeong Choi

Department of Agricultural Biotechnology
College of Agriculture and Life Sciences
Seoul National University

Molecular study of Shigella flexneri Bacteriophage-Host Receptor Interaction for Biocontrol

지도교수 이 주 훈

이 논문을 농학석사 학위논문으로 제출함

2023년 2월

서울대학교 대학원

농생명공학부

최 호 정

최호정의 석사 학위논문을 인준함

2023년 2월

위 원 장 유 상 열 (인)

부 위 원 장 이 주 훈 (인)

위 원 하 남 출 (인)

ABSTRACT

Choi, Hojeong

Department of Agricultural Biotechnology

The Graduate School

Seoul National University

Shigella is an important causative agent of shigellosis causing diarrhea. Considering the antibiotic resistance, bacteriophage can be used as an antibiotics alternative for controlling *Shigella*. However, bacterium can become resistant to phage by modifying their cell surface receptor. Therefore, phage treatment that targets various host receptors is more effective than single phage treatment. In this study, new phages targeting *Shigella flexneri* were isolated and their specific host receptors were identified to understand the phage-host interaction. To identify the host receptors of the five phages (SFPB, SFP17, SFP20, SFP21A, SFP21B), various mutant libraries were constructed using lambda-red recombination and the Tn*5* transposon in *S. flexneri* 2a strain 2457T. The host receptor identification results showed that there were three types of receptors; an outer core of lipopolysaccharide (LPS), an O-antigen of LPS and a FhuA (ferrichrome porin). Phage SFPB uses the outer core as the receptor, while phage SFP20, SFP21A, and SFP21B use the O-antigen of LPS as the receptor. Phage which utilize FhuA for adsorption was SFP17. Interestingly, in the mutation of the *galU* gene that directly affects the synthesis of the *S. flexneri* LPS core, SFP21A showed different lytic activity than SFP20 and

SFP21B. For further understanding of the host-phage interaction, protein structure prediction was performed and showed that the Gp09 of SFP21A is a receptor binding protein having a homotrimeric structure. The binding activity using EGFP-Gp09 fusion proteins revealed that the C-terminal domain is involved in receptor binding. Furthermore, multiple sequence alignment revealed that one amino acid (V669) at the C-terminus of Gp09 was predicted to be a key residue affecting binding in the *galU* mutant. Point mutation analysis revealed that the single point mutation in the C-terminal domain (V669A) could no longer bind to the *galU* mutant, indicating that valine present in the outer loop of the binding domain may play an important role when binding to the host cell surface. Based on these characteristics, a spotting assay was performed to confirm the lytic activity of SFP21A in the phage SFPB-resistant mutant strains. As a result, despite being an O-antigen receptor, SFP21A effectively inhibited the SFPB-resistant bacteria. Therefore, phage cocktail consisting of SFP21A, SFPB, and SFP17 effectively inhibited the bacterial growth and the emergence of resistant strains than the single phages. These results suggest that phage cocktail using phage-host interactions can be applied as alternative biocontrol agents against *S. flexneri*.

***Keywords*: *Shigella flexneri*, Bacteriophage, Receptor, Tail fiber protein, Receptor binding protein**

Student Number: 2021-24995

CONTENTS

List of Figures

List of Tables

1. Introduction

The genus *Shigella* generally comprises four major serological groups: Group A (*S. dysenteriae*, 15 serotypes), Group B (*S. flexneri*, 19 serotypes), Group C (*S. boydii*, 19 serotypes), and Group D (*S. sonnei*, 1 serotype) (Sun, Lan et al. 2013). Among them, *S. flexneri* is a common predominant species and 2a is a major serotype in this species (Kotloff, Winickoff et al. 1999). The *Shigella* infection, a shigellosis, is an intestinal infection by *Shigella*, causing bloody diarrhea, fever, sever stomach cramping, nausea, and even vomiting (CDC 2018). In the United States between 1998 and 2008, major shigellosis outbreaks (58%) were food-borne, and these outbreaks (54%) were restaurant-associated (Nygren, Schilling et al. 2013). And raw food consumption and infected food handlers were major causes of these outbreaks and most of their bacterial origin (79%) was *S. sonnei*. During this period, 120 shigellosis outbreaks were reported and analyzed, resulting in 6,208 food-borne illnesses, 197 hospitalizations, and one death (Nygren, Schilling et al. 2013). In 20 countries of EU and UK for 2017, a total of 8,465 confirmed cases of shigellosis by *S. sonnei* were reported, and the highest notification rate was observed in children under five years of age (ECDC 2020). This high shigellosis risk to children was also supported by shigellosis

outbreaks by *S. flexneri* in sub-Saharan Africa and South Asia, reporting that nearly 70% of all deaths were children under the age of five (Jennison and Verma 2004). Interestingly, *S. sonnei* is a major bacterial origin for shigellosis in the developed countries, but *S. flexneri* is in the developing countries (Kotloff, Winickoff et al. 1999).

Since various antibiotics have been treated to control and regulate the shigellosis outbreaks, multidrug-resistant *Shigella* has been emerged globally, especially ampicillin, chloramphenicol, co-trimoxazole, tetracycline, and nalidixic acid (Bhattacharya, Sugunan et al. 2012, Control and Prevention 2018). Due to this multidrug resistance, WHO recommended antibiotics such as ciprofloxacin (1st stage) and pivmecillinam, ceftriaxone, or azithromycin (2nd stage) for treatment of shigellosis in 2005 WHO guidelines (Williams and Berkley 2018). However, a new control and regulation method for shigellosis needs to be developed to reduce the shigellosis without the emergence of this multidrug-resistance in *Shigella*. In addition, usage and application of antibiotics is strictly prohibited in foods (European Commission 2003). Therefore, bacteriophage has been suggested as an alternative approach to control and regulate *Shigella* in foods (Jun, Kim et al. 2013, Llanos-Chea, Citorik et al. 2019, Shahin, Zhang et al. 2021).

Bacteriophage is a bacterial virus and lyse a specific host bacterium via its distinct lytic and lysogenic life cycles without emergence of antibiotic

resistance, suggesting human safety. Because of these characteristics, bacteriophage has been considered as an alternative strategy for handling of antimicrobial-resistant infections (Criscuolo, Spadini et al. 2017, De Sordi, Lourenço et al. 2019, Dedrick, Guerrero-Bustamante et al. 2019). Previously, food application of bacteriophage was approved as a natural food preservative by US Food and Drug Administration (FDA) in 2006 (Lang 2006). Furthermore, a new phage product, ShigaShield, was developed with a phage cocktail of five *Shigella* lytic phages and commercialized by Intralytix (USA) (Shahin, Zhang et al. 2021). However, this phage cocktail was prepared with a random combination of those *Shigella* lytic phages without consideration of phage-resistant mutant formation. To enhance the host lysis activity of *Shigella* phages and to prevent phage-resistant mutant formation of the *Shigella* host strains, their further characterizations, associated with host receptor identification, its recognition, and even host lysis mechanisms, are required.

It has been known that bacterial host often protects phage infection by modification of its host receptor and consequently becomes resistant to the phage (Kim and Ryu 2012, Shin, Lee et al. 2012, Bai, Jeon et al. 2019). To overcome this limitation, it is important to identify the host receptor and to characterize the interaction between the host receptor and its specific phage to reduce the host resistance to phage infection (Kim, Kim et al. 2014, Bai,

Jeon et al. 2019, Yehl, Lemire et al. 2019). To initiate phage infection to specific bacterial host, phage binds to a specific receptor on the surface of the bacterial cell membrane, such as outer-membrane protein (OmpC), ferrichrome outter-membrane transporter (FhuA), vitamin B12 trasporter (BtuB), colicin E import outer-membrane (TolC), O-antigen of lipopolysaccharides (LPS) and flagella (Shin, Lee et al. 2012, Bai, Jeon et al. 2019). Therefore, the modification of the specific host receptor can develop the resistance of phage infection to the host strain. To reduce this phage resistance of the host, phage cocktail containing a few or several phages targeting different types of host receptors has been suggested (Kim, Kim et al. 2014, Bai, Jeon et al. 2019). As an example, treatment of a *Salmonella* phage cocktail containing three different phages targeting different host receptors (BSPM4, flagella; BSP101, O-antigen of LPS; BSP22A, BtuB) delayed effectively the emergence of phage-resistant *Salmonella* Typhimurium in fresh lettuce and cucumber (Bai, Jeon et al. 2019). In addition, a *Salmonella* phage cocktail containing two phages targeting two different regions of LPS (SSU14, O-antigen; SSU5, outer core) reduced the phage resistance of the host (Kim, Kim et al. 2014).

While only two *S. flexneri* phage Sf6 and A1-1 targeting three host receptors (O-antigen of LPS, OmpA, and OmpC) and OmpA were identified and characterized (Parent, Erb et al. 2014), no *Shigella* phage cocktails were

developed with consideration of host receptors to date (Shahin, Zhang et al. 2021, Shahin, Zhang et al. 2021). Treatment of a single phage vB_SflS-ISF001 to *S. flexneri*-contaminated raw and cooked chicken breast samples showed 2.0 log CFU reductions in 48 h and 24 h, respectively, suggesting that *S. flexneri* phage treatment was more effective to cooked sample (Llanos-Chea, Citorik et al. 2019). To increase the host lysis activity and to reduce the phage-resistance in the host strain, the phage cocktail has been prepared and evaluated. A phage cocktail containing *S. flexneri* phage SF-A2, *S. dysenteriae* phage SD-11, and *S. sonnei* phage SS-92 was treated to ready-to-eat spiced chicken contaminated with those host *Shigella* species, showing 2.0 log CFU reduction in 48 h (Zhang, Wang et al. 2013). In addition, an extended phage cocktail containing *S. flexneri* phages vB_SflS-ISF001/ vB_SflM_004, *S. dysenteriae* phages vB_SdyS-ISF003/vB_SdyM_006, and *S. sonnei* phages vB_SsoS-ISF002/vB_SsoS_008 was treated to four different food samples such as cherry tomato, Chinese cabbage, baked breast chicken meat, and yogurt, showing 3.8, 4.0, 3.3, and 3.4 log CFU reductions in 24 h, respectively (Shahin, Zhang et al. 2021). Based on these results, the phage treatment with a single phage, a cocktail of three different phages, and a cocktail of six different phages (two phages per *Shigella* species) revealed that the phage cocktail with more phage mixture has higher growth inhibition and host lysis activities of *Shigella* in foods, but did not reduce the phage

resistance of the host strains due to no consideration of host receptors. Therefore, further study regarding host receptor identification and recognition of *Shigella* phages, and their interactions for synergic host lysis activity must be performed to minimize the emergence of phage-resistant mutants. These advanced host receptor studies would provide important information for optimized formulation of phage cocktails to lengthen and maintain the host growth inhibition activity of the phage cocktails without formation of phage-resistant mutants. In addition, characterization of the phage receptor binding proteins (RBPs) in the phage tails would be important to understand the interaction between the host receptors and phage tails, because modification of small region in RBP of *E. coli* T3 phage tail fiber protein affected the binding activity to the host receptors (Yehl, Lemire et al. 2019). Therefore, characterization of RBPs in *Shigella* phages is necessary for further understanding of host recognition and phage infections between host receptors and phage tails.

In this study, five *S. flexneri* phages were newly isolated from various sewage samples and their phenotypic and genotypic characteristics were analyzed. In addition, their host receptors were identified and compared to elucidate the host specificities. Based on this, further receptor study was performed to differentiate host receptors targeting O-antigen of LPS. Furthermore, their RBPs were identified and characterized. Subsequently, a

phage cocktail was prepared with consideration of the host receptors and RBPs and it was evaluated for host lysis, growth inhibition, and even prevention of phage resistance of the host strains. Consequently, our findings would be useful for development of a novel phage cocktail as a promising alternative antimicrobial agent against *Shigella.*

2. Materials and methods

2.1. Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study were listed in Table 1 and 6, respectively. *Shigella flexneri* 2a strain 2457T was selected as an indicator strain for isolation and propagation of bacteriophages. *Shigella* and other gram-negative bacteria were grown in Luria-Bertani (LB) broth (Difco, USA) with shaking for 12 h at 37°C. Gram-positive bacteria were cultivated in Brain Heart Infusion (BHI) broth (Difco) for 24 h at 37°C. Agar and soft top agar were prepared with the associated broth supplemented with 1.8% agar and 0.6% agar (Difco), respectively.

Table 1. Bacterial strains and plasmids used in this study

Strain and plasmid	Relevant characteristic(s)[a]	References
Escherichia coli		
MFD*pir*	MG1655 RP4-2-Tc::[ΔMu1::Δ*aac(3)IV*-Δ*aphA*-Δ*nic35*-ΔMu2::zeo] Δ*dapA*::(*erm-pir*) *ΔrecA*	(Ferrières, Hémery et al. 2010)
DH5α λ*pir*	*Φ* 80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 hsdR17 recA1 endA1 gyrA96 deoR supE44 thi-1 relAI/λpir*	(Platt, Drescher et al. 2000)
BL21 (DE3)	F- *ompT hsdS*$_B$ (r_B^-, m_B^-) *dcm gal* (DE3)	Novagen
***S. flexneri* 2a strain 2457T**		
ΔwaaL	*waaL* deletion mutant of *S. flexneri*	This study
ΔwaaG	*waaG* deletion mutant of *S. flexneri*	This study
ΔflgK	*flgK* deletion mutant of *S. flexneri*	This study
ΔompA	*ompA* deletion mutant of *S. flexneri*	This study
ΔompC	*ompC* deletion mutant of *S. flexneri*	This study
ΔbtuB	*btuB* deletion mutant of *S. flexneri*	This study
ΔfhuA	*fhuA* deletion mutant of *S. flexneri*	This study
galU::Tn*5*	*S. flexneri* with transposon insertion in putative *galU*	This study
ΔwaaL+pWaaL	*ΔwaaL* complemented with WT *waaL* gene	This study
ΔwaaG+pWaaG	*ΔwaaG* complemented with WT *waaG* gene	This study
ΔfhuA+pFhuA	*ΔfhuA* complemented with WT *fhuA* gene	This study
galU::Tn*5*+pGalU	*galU*::Tn*5* complemented with WT *galU* gene	This study
Plasmid		
pRL27	Tn5-RL27 (Km^R-*oriR6K*γ) delivery vector	(Larsen, Wilson et al. 2002)
pKD46	P_{BAD}(*gam-bet-exo*) *oriR101 repA101*(Ts); Amp^R	(Datsenko and Wanner 2000)
pKD3	*oriR6K*γ Amp^R-FRT, and Cm^R-FRT	(Datsenko and Wanner 2000)
pKD20	*cI*857 *oripSC101*(Ts) λP_R-*flp*;Amp^R, Cm^R	(Datsenko and Wanner 2000)
pUHE21-2 *lacI*^{q}	rep_{pMB1} *lacI*^{q}; inducible Lac promoter; Amp^R	(Soncini, Véscovi et al. 1995)
pET28a	His-tagged expression vector, T7 promoter, Kan^R	Novagen
pET28a-EGFP	pET28a with EGFP	(Son, Kong et al. 2018)

pWaaL	pUHE21::*waaL*; AmpR	This study
pWaaG	pUHE21::*waaG*; AmpR	This study
pFhuA	pUHE21::*fhuA*; AmpR	This study
pGalU	pUHE21::*galU*; AmpR	This study
EGFP-Gp09	pET28a::EGFP-Gp09; AmpR	This study
EGFP-Gp09NTD	pET28a::EGFP-Gp09NTD; AmpR	This study
EGFP-Gp09CTD	pET28a::EGFP-Gp09CTD; AmpR	This study
EGFP-Gp09V669A	pET28a::EGFP-Gp09V669A; AmpR	This study

[a]Amp^R, ampicillin resistant, Kan^R, kanamycin resistant, Cm^R, chloramphenicol resistant

2.2. Bacteriophage isolation and stock preparation

Diverse sewage samples were collected from Jungnang, Seonam, Kyongan, Opo, Seongnam and Gonjiam Sewage treatment plants in South Korea. The phage isolation was performed according to the previously described protocol (Park, Kim et al. 2017). And then, the phage spot assay was performed to select virulent phages forming clear plaques with host lysis activity. After virulent phage selection using the spot assay, only one phage with the largest diameter of the clear plaques was chosen as a representative virulent phage per one sample. Before phage propagation, the *S. flexneri* host strain was incubated in 3 ml of fresh LB medium with shaking at 37°C up to OD_{600} = 0.5. A single phage plaque from the agar plate of selected phage was picked and added to the incubated culture, and additionally incubated at 37°C for 3 h. After incubation, the culture was centrifuged at 8,000 × g for 20 min and the supernatant was filtered with 0.45 μm syringe filter (Pall Co., USA) to remove the host strain. The filtrate containing the selected virulent phage was used for the next propagation step with 10 ml culture of the indicator strain. After that, this propagations step was repeated with 20, 200, and 800 ml of the culture to increase the culture volume. After serial phage propagations, the selected virulent phage was concentrated by precipitation with polyethylene glycol (PEG) 8000 (Daejung, Korea) and purified using Cesium chloride (CsCl) density gradient ultracentrifugation (78,500 × g for

2h at 4°C) (HIMAC CP100β; Hitachi, Japan). After ultracentrifugation, the phage band was extracted using a sterilized syringe, and then dialyzed with sodium chloride-magnesium sulfate (SM) buffer (100 mM NaCl, 10 mM $MgSO_4 \cdot 7H_2O$ and 50 mM Tris-HCl at pH 7.5). The concentrated phages were stored at 4°C for further experiments.

2.3. Transmission electron microscopy (TEM)

The diluted phage solution (1×10^9 PFU/ml) was negatively stained with 2% uranyl acetate (pH 4.0) on carbon-coated copper grids. Morphology of the stained phages was observed under transmission electron microscopy (TEM; JEM1010, Japan) at 80 kV in National Instrumentation Center for Environmental Management (NICEM) of Seoul National University (South Korea). The phages were identified and classified on the basis of International Committee on Taxonomy of Viruses (ICTV) classification (Lefkowitz, Dempsey et al. 2018).

2.4. Bacteriophage host range analysis

All selected strains for host range analysis were listed in Table 1. After bacterial incubation, 100 μl of the culture was added to 6 ml of 0.6% soft top agar and the mixture was poured onto the agar plate. The phage was 10-fold serially diluted with SM buffer and the serially diluted phage

suspension was spotted onto the prepared agar plates and they were incubated for 24 h at 37°C. After incubation, the number of formed single plaques was recorded to determine the phage sensitivity of the indicator strain. To clarify the ability of phage infection to specific indicator strain, the efficiency of plating (EOP) was determined by comparison of the phage titer of the selected indicator strain with that of the host strain, *S. flexneri* 2a strain 2457T.

2.5. Genome sequencing and bioinformatics analysis

Phage genomic DNAs were extracted using the Viral Gene-spin™ Viral DNA/RNA Extraction Kit (Intron, Korea), and sequenced using the Illumina HiSeq DNA sequencer (Illumina, USA) and MinION sequencer (Oxford Nanopore, USA). Open reading frames (ORFs) were predicted using Glimmers (Altermann and Klaenhammer 2003), FgenesV (Softeberry, Inc., USA), and GeneMarkS (Besemer, Lomsadze et al. 2001), and ribosomal binding sites were predicted using RBSfinder to confirm ORF predictions (J. Craig Venter Institute, USA). Functions of the predicted ORFs were confirmed and annotated using BLASTP (Altschul, Madden et al. 1997) and InterProScan programs (Quevillon, Silventoinen et al. 2005). BLAST Ring Image Generator (BRIG) was used to visualized genome of the phages (Alikhan, Petty et al. 2011). To find the receptor binding domain, the trimer structure of tail fiber protein (Gp09) was predicted using the AlphaFold v2.2

through a local installation (Jumper, Evans et al. 2021) with AlphaFold-multimer implementation (Evans, O'Neill et al. 2022) for the prediction of multiple chains. The resulting trimer structure was displayed by the PyMol program (DeLano, 2002). DNA and protein sequence alignment analyses were conducted using ClustalX2 (M.A. Larkin et al., 2007) and Genedoc (Karl B. Nicholas et al., 1997), respectively.

2.6. Mutant library constructions using Tn*5* transposon

A Tn*5* transposon mutant library was constructed as follows: A donor strain, *E. coli* MFD*pir* containing a suicide vector pRL27 and a recipient strain, *S. flexneri* 2a 2457T was harvested at OD_{600} = 0.6, and they were washed three times with 10 mM $MgSO_4$. A mixture of donor and recipient cells (3:1, v/v) was spotted onto the LB plate supplemented with 0.3 mM diaminopimelic acid (DAP). After 24 h incubation, the cell spot was resuspended with 1 ml of fresh LB broth medium, plated onto LB agar plate containing kanamycin sulfate (50 μg/ml), and incubated at 37°C for 24 h to obtain isolated colonies.

To isolate host receptor mutants for phage resistance from the mutant library, 5 μl of phage (10^8 PFU/ml) were spotted onto the mutant-spread agar plates and incubated at 37°C for 24 h. After obtaining the colonies, their phage resistance was verified by the spot assay with the host receptor mutants. To

determine the Tn*5* transposon insertion site, the genomic DNA of the selected transposon mutant was extracted and randomly digested with *Bam*HI (New England Biolabs, USA). After purification of *Bam*HI-digested DNA fragments, they were self-ligated with T4 ligase (Roche, Switzerland), transformed into *E. coli* DH5α λ*pir* by heat-shock method, and plated on the fresh LB plate containing kanamycin sulfate (50 μg/ml). Due to presence of R6Kγ origin in the Tn*5* transposon, only colonies harboring self-ligated DNA fragment with Tn*5* transposon were selected after incubation at 37°C for 24 h, and sequenced with a pair of sequencing primers, tpnRL17-1 (5'-AAC AAG CCA GGG ATG TAA CG-3') and tpnRL 13-2 (5'-CAG CAA CAC CTT CTT CAC GA-3') (Larsen, Wilson et al. 2002) to obtain DNA sequences neighboring Tn*5* transposon. And then Tn*5* transposon-neighboring DNA sequences were compared with complete genome sequence of *Shigella flexneri* 2a strain 2457T (GenBank accession number AE014073.1) to identify the Tn*5* transposon insertion site. Based on the Tn*5* transposon insertion site and the phage resistance of the mutant, the host receptor was identified.

2.7. Mutant library construction using lambda red mutation

To construct the mutant library by specific gene mutations, the lambda red recombination system was used (Datsenko and Wanner 2000,

Ranallo, Barnoy et al. 2006, Lee, Bai et al. 2016). The plasmid pKD46 was transformed into the host strain, *S. flexneri* 2a strain 2457T using electroporation (2.5 kV and 200 Ω), and then colonies of the transformant was selected on the LB agar plate containing ampicillin (50 μg/ml). For preparation of electrocompetent cells, the transformant strain was incubated in 50 ml of LB broth containing ampicillin (50 μg/ml) at 30°C up to OD_{600} = 0.5. After incubation, the culture was heat shocked at 42°C for 15 min, transferred to an ice-water bath for 10 min to enhance the electroporation efficiency. And then, bacterial cells are harvested by centrifugation at maximum speed for 15 min and resuspended with 1 mM MOPS and 20% glycerol solution. This step was repeated 3 times for electrocompetent cell washing. The chloramphenicol resistant (Cm^R) cassette from plasmid pKD3 was amplified using primers containing 50-mer homologous sequences of host receptor-associated genes and verifying sequence of pKD3 (Table 2). And then, electrocompetent cells were electroporated with the PCR fragment (0.1-0.4 μg) at 2.5 kV and 200 Ω. The transformed electrocompetent cells were recovered in 1 ml of fresh LB broth for at least 1 h at 30°C, and then plated on LB agar plates containing chloramphenicol (20 μg/ml). After colony selection, FLP recognition target site of the chloramphenicol resistant cassette was removed by additional electroporation of plasmid pCP20. Specific gene deletion was confirmed by PCR (Table 3). For complementation of the mutant,

each host receptor gene was amplified using PCR with the associated forward/reserve primer set (Table 4). The PCR product and pUHE21-2 *lacI*^{q} plasmid vector were double-digested with *Bam*HI/*Hind*III and ligated. After electroporation of the ligated vector into the associated *S. flexneri* mutant, the transformants were selected on fresh LB agar plates containing ampicillin (50 μg/ml). The cloned gene in the vector was verified by DNA sequencing with specific sequencing primer set (Table 4). The selected transformants were induced using 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) for complementation of the inactivated gene in the mutant.

Table 2. Primers used for construction of *S. flexneri* mutants

Gene target	Primer name	Oligonucleotide sequence (5'-3')[a]
waaL	waaL-RED-F	TTG TTT TTC ATC GCT AAT AAT AAG CCG GCG TAA ACG CCT AAT AAA TTT GGG TGT AGG CTG GAG CTG CTT C
	waaL-RED-R	CTC AAC ATT ATT TTT CTC TCT CGA GAA AAA AAA CTG GAT AGC GTA CTG GAA TGG GAA TTA GCC ATG GTC C
waaG	waaG-RED-F	GCT GCT GTC GAT AAA TTA CTG CCC TCC TCC ACG ACA GGT ACG TCG TTA TGC ATA TGA ATA TCC TCC TTA G
	waaG-RED-R	CGT GGC AAA CGG CTC TTT AAG TTC AAC CAT CCA GAC CAC CCG TTA TGA TAG TGT AGG CTG GAG CTG CTT CG
flgK	flgK-RED-F	AAC CTC CAT GTC CAG CTT GAT CAA TAA CGC CAT GAG CGG ACT GAA CGC GGG TGT AGG CTG GAG CTG CTT C
	flgK-RED-R	CCT TAG CGA ATG TTA ATC AGC GCA TCA AAA ATC GCG TTT GCC GTC TGC AGA TGG GAA TTA GCC ATG GTC C
ompA	ompA-RED-F	GCG GGG TTT TTC TAC CAG ACG ATA ACT TAA GCC TGC GGC TGA GTT ACA ACG TGT AGG CTG GAG CTG CTT C
	ompA-RED-R	GGA TGA TAA CGA GGC GCA AAA AAT GAA AAA GAC AGC TAT CGC GAT TGC AGA TGG GAA TTA GCC ATG GTC C
ompC	ompC-RED-F	GAC GCC GTT TAA AAT TCG TTT AGA AAA TAC GTT TAA CGA TAA TTA AAA GGG TGT AGG CTG GAG CTG CTT C
	ompC-RED-R	AGC ACT TAA ATC AAA AAC CAG CCC AGA GAA GGG CTG GTC AGG AGA GGA CGA TGG GAA TTA GCC ATG GTC C
btuB	btuB-RED-F	GTC TCA TAA TCT TTG TCG AAC AGG TTG GCT ATT TTA CCA CGA ACT GTC AGG TGT AGG CTG GAG CTG CTT C
	btuB-RED-R	TAT TGT GGA TGC TTT ACA ATG ATT AAA AAA GCT TCG CTG CTG ACG GCG TGT GGG AAT TAG CCA TGG TCC
fhuA	fhuA-RED-F	GGT GTC GAC TTT ATG CGT ATG CGT AAT GAC ATC AAC GCC TGG TTT GGT TAG TGT AGG CTG GAG CTG CTT C
	fhuA-RED-R	AAC AGC CAA CTT GTG AAA TGG GCA CGG AAA TCC GTG CCC CAA AAG AGA AAA TGG GAA TTA GCC ATG GTC C

[a], Sequences of the priming sites in pKD3 are underlined.

Table 3. Primers used for confirmation of *S. flexneri* mutants

Primer name	Oligonucleotide sequence (5'-3')	Reference
waaL-confirm-F	GGG ATG GCG TAA CTC AAA GA	This study
waaG-confirm-F	CGC CAT TTC AAA GCG CAG AG	This study
flgK-confirm-F	GGT CGC TGC CGA TAA TAC TC	This study
ompA-confirm-F	GGC ATT GCT GGG TAA GGA AT	This study
ompC-confirm-F	CGA CTT CGA TCA CAT ATC GA	This study
btuB-confirm-F	ACG TGG TTC AGA AGG TGT AG	This study
fhuA-confirm-F	GCA ATG TCG ATT TAT CAG CG	This study
C1	TTT TCA CCA TGG GCA AAT AT	(Datsenko and Wanner 2000)

Table 4. Primers used for complementary of *S. flexneri* mutants

Gene target	Primer name	Oligonucleotide sequence (5'-3')[a]
waaL	pUHE21-waaL-BamHI-F	GAA ATG GAT CCC AAT AAG TTG ACA TCG GAG
	pUHE21-waaL-HindIII-R	GGT TTG AAT AAA GCT TAA GGC CGC ATT ATG
waaG	pUHE21-waaG-BamHI-F	GGC AGA TGG ATC CGC TGC TGT CGA TAA ATT ACT
	pUHE21-waaG-HindIII-R	TTT TAA CTT CAA GCT TAG GAT CTT TGC CGC GCC
fhuA	pUHE21-fhuA-BamHI-F	GAG ATG GAT CCA TGG CGC GTT CCA AAA CTG C
	pUHE21-fhuA-HindIII-R	TGC CCC AAA AAA GCT TTT AGA AAC GAA AGG TTG
galU	pUHE21-galU-BamHI-F	CTT GGA TCC GCT GAT ATA CTG GGA TGC GA
	pUHE21-galU-HindIII-R	GTC AAG CTT TCA ACG CCG TTT CGT GGA TA

[a], Restriction sites are underlined.

2.8. Lipopolysaccharide (LPS) extraction and SDS-PAGE analysis

LPS samples were extracted from *S. flexneri* 2a strain 2457T and its LPS mutant strains using the phenol-water extraction method (Wang, Zhang et al. 2010) with modifications. One milliliter of bacterial cultures (1 × 10^8 CFU/ml) was harvested, washed once with 1 ml DPBS (Dulbecco's PBS containing 0.15 mM $CaCl_2$ and 0.5 mM $MgCl_2$), and resuspended in 300 µl molecular grade water. An equal volume of 68°C pre-heated Tris-saturated phenol solution (Sigma-Aldrich, USA) was added. The mixture was incubated at 68°C for 15 min with vigorous vortexing at every 5 min. After incubation, the mixture was chilled on ice for 5 min and the aqueous phases were separated by centrifugation at 10,000 × g, 4°C for 5 min. This LPS extraction was conducted again with the same procedure. 500 µl Sodium acetate (0.5 M, final concentration) and 10 ml absolute ethanol were added to the aqueous phase containing extracted LPS, and incubated at −80°C for 24 h. After incubation, the LPS was precipitated by centrifugation at 10,000 × g, 4°C for 5 min, resuspended in 100 µl molecular grade water, and precipitated again with absolute ethanol. Finally, the precipitated LPS was resolved in 50 µl molecular grade water and stored at −80°C until use. The purified LPS was separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) with a 12% acrylamide gel. The SDS-PAGE gel was fluorescently stained using a Pro-Q™ Emerald 300 Lipopolysaccharide Gel

Stain Kit (Molecular Probes, USA), according to the manufacturer's instructions and then visualized with Gel DOC^{TM} EZ Imager (Bio-Rad, USA).

2.9. Assays of inhibitory effects of LPS on phage infection

The extracted LPS (25 μg/ml) was added to phage suspension (7×10^3 PFU/ml) and incubated at 37°C for 2 h. A hundred microliter of each serially diluted mixture ($10^0 – 10^8$ dilutions) was added to 100 μl of an exponentially grown culture of host strain (1×10^7 CFU/ml). And then, the LPS-phage-host strain mixture was gently mixed with 6 ml of 0.6% soft top agar and poured onto fresh LB agar plate. After incubation at 37°C for 12 h, the titer number of each phage was determined. Phage without LPS was used as the negative control (phage titer, 100%). Lowering effect of LPS on phage infection were evaluated by EOP determination. Overall experiment was carried out in triplicates.

2.10. Bacteriophage adsorption assay

Bacterial cells were harvested at OD_{600} = 1.0 and diluted with fresh LB broth to OD_{600} = 0.1. Phage SFP21A was added to the diluted bacterial culture at a multiplicity of infection (MOI) of 0.01 and incubated at 37°C. One milliliter of the mixed culture was collected every 5 min and centrifuged at 16,000 × g for 1 min to remove bacteria cells. And then, the supernatant

was filtered using 0.22-um diameter syringe filter (Pall Co.). The number of PFU in the collected supernatant was determined by spotting and overlay assay using *S. flexneri* 2a strain 2457T as a bacterial host strain.

2.11. Adsorption assay with periodate and proteinase K treatments

S. flexneri 2a strain 2457T was treated with periodate or proteinase K to examine the effect of the treatments on phage adsorption, according to methods that were described by Kiljunen et al. (Kiljunen, Datta et al. 2011), with some modifications. Briefly, When the OD_{600} of the bacterial culture reached 1.0, 1 mL of the culture was collected by centrifugation at 13,000 g for 1 min and washed with 1 mL of fresh LB broth. Proteinase K (0.2 mg/mL, final concentration) was added to the prepared sample and then incubated at 37°C for 2h. For periodate treatment, 2 mL of the culture was harvested by centrifugation and washed with 1 mL of LB broth. The pellet was then treated with 1 mL of sodium acetate (50 mM, adjusted to pH 5.2) or sodium acetate containing either 10- or 100-mM periodate for 2 h in the dark. After the treatments, cells were washed at least three times with 1 mL of LB broth, which was adjusted to an OD_{600} of 0.1, and then the phage adsorption assay was performed as described above.

2.12. Bacterial growth inhibition assay

Bacteria cells were incubated to 50 ml of fresh LB broth at 37°C. When the OD_{600} = 1.0, phage was added at a MOI of 0.1. During the incubation, the culture sample was collected every hour and the bacterial growth was monitored by the OD_{600} measurement. Instead of phage, SM buffer was added to the culture as a negative control.

2.13. Characterization of native tail fiber protein

2.13.1. PCR, Gene expression, and protein purification

The primer set for PCR amplification of a whole gene encoding Gp09 (tail fiber protein) consists of a forward primer containing *Eco*RI site (Gp09-EcoRI-F) and a reverse primer containing *Hind*III site (Gp09-HindIII-R) (Table 5). From the structure analysis, two primer sets were designed for PCR amplification of N-terminal domain and C-terminal domain with the same restriction enzyme sites (Gp09-EcoRI-F/$Gp09_{NTD}$-HindIII-R; $Gp09_{CTD}$-EcoRI-F/Gp09-HindIII-R), respectively (Table 5). For PCR amplification, PCR mixture (final volume, 50 µl) contained 1 µl of template DNA (40 ng/µl), 1 µl of forward and reverse primers (20 µM each), 25 µl of KAPA HiFi HotStart ReadyMix (2X), and 23 µl of molecular water. The PCR conditions were as follows: 1 cycle of 95℃ for 3 min; 35 cycles of 98℃ for 20 s, 60℃ for 15 s, and 72℃ for 30 s; and 1 cycle of 72°C for 1 min. Following PCR, 16S rRNA amplicons were purified using a NICSROprep™ PCR Clean-up

S & V Kit (Bionics, South Korea). After PCR amplification and purification, each PCR product and pET28a-EGFP vector were double-digested with *Eco*RI/*Hind*III (Son, Kong et al. 2018) and they were ligated using T4 DNA ligase (New England Biolabs). After heat shock of the ligated vector into *E. coli* BL21(DE3), the transformants were selected on fresh LB agar plate containing ampicillin (50 μg/ml). When the transformant cells reached to OD_{600} = 0.5, expression of the cloned gene was induced at 18°C for 16 h with 1 mM IPTG (final concentration). The cells were harvested by centrifugation at 7,000 × g for 15 min, and resuspended with 5 ml of lysis buffer (50 mM sodium dihydrogen phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.4). The resuspended cells were disrupted using a sonicator (ActiveMotif, USA) with the condition of 2 s pulse on and 4 s pulse off for total 5 min. After sonication, the lysate was centrifuged at 10,000 × g for 30 min and then the supernatant was transferred to a new 15-ml falcon tube (Corning, USA). The EGFP-tagged protein was purified by immobilized metal ion affinity chromatography using a Ni-NTA agarose column (Qiagen, Germany), according to the manufacturer's protocol. After purification, the EGFP-tagged protein was analyzed using 4-20% SDS-PAGE precast gradient gel with a Precision Plus Protein Dual Xtra Standard protein marker (Bio-Rad). In addition, as a negative control, the EGFP protein was obtained by heat-shock transformation with pET28a-EGFP without insert DNA, IPTG

induction, cell disruption with sonication, protein purification with Ni-NTA column with the same protocols. And the obtained EGFP protein was confirmed by SDS-PAGE analysis.

Table 5. Primers used for construction of EGFP fusion proteins

Primer name	Oligonucleotide sequences (5'-3')[a]
Gp09-EcoRI-F	GTA AAA GAA TTC GAT GGT CGT ACC GTT GAG CA
Gp09-HindIII-R	GGT ACG AAG CTT TGA GGC TTT TAC CAT AGA GG
$Gp09_{CTD}$-EcoRI-F	GTA AAA GAA TTC GAT GGT CGT ACC GTT GAG CA
$Gp09_{NTD}$-HindIII-R	GGT ACG AAG CTT TGA GGC TTT TAC CAT AGA GG
F1-SDM	CGG GAG TGC TAC CAT AGG CTC C
R1-SDM	GGA AGC TAT TAA GTG CGG TAG CCT GAG GC
F2-SDM	GCC TCA GGC TAC CGC ACT TAA TAG CTT CC
R2-SDM	GTA CTC CCT GTT CAG CGC CTC C

[a], Restriction sites are underlined.

2.13.2. Host binding assay

Five hundred microliter of host bacterial culture was harvested at OD_{600} = 0.8 by centrifugation at 13,000 × g for 1 min and the cell pellet was resuspended with 150 µl PBS buffer. And then 100 µl of purified EGFP-tagged protein (1 mg/ml) was added to the resuspended host cell solution and the mixture was incubated at room temperature for 15 min. The mixture was washed twice with 1 ml PBS buffer and resuspended again with 100 µl PBS buffer. Finally, 5 µl of the cell suspension were analyzed using a fluorescence microscope (DM2500; Leica, Germany). The purified EGFP protein was prepared and used as a negative control with the same procedures.

2.14. Characterization of the point-mutated tail fiber protein

2.14.1. Point mutation of native tail fiber protein

The site-directed mutagenesis (SDM) of a 2,511-bp gene encoding tail fiber protein (Gp09) was performed by PCR-mediated overlap extension with two different PCR reactions (Huss, Meger et al. 2021). The first PCR was performed to obtain two PCR amplicons (front part, 2,075-bp; rear part, 970-bp) with F1-SDM/R1-SDM primer set and F2-SDM/R2-SDM primer set (Table 5), and these PCR amplicons were overlapped in the middle of Gp09. Because R1-SDM and F2-SDM primers were chemically synthesized with two switched nucleotides by replacement of “TG” to “CA” in the overlapped

region, these two PCR amplicons also have point mutated “CA” in the sequences of overlapped region. And then, the second PCR was conducted with these two PCR amplicons as DNA templates and Gp09-EcoRI-F/Gp09-HindIII-R primers (Table 5) to obtain a whole gene encoding tail fiber protein with replacement of “TG” to “CA”. This final PCR amplicon was purified using a NICSROprep™ PCR Clean-up S & V Kit (Bionics, South Korea). This point mutated gene of tail fiber protein was used for further cloning and gene expression.

2.14.2. Gene expression, protein purification, and host binding assay of the point-mutated tail fiber protein

The point-mutated Gp09 and pET28a-EGFP vector were double-digested with *Eco*RI/*Hind*III and they were ligated. Point-mutated tail fiber protein was obtained by heat-shock transformation with pET28a-EGFP, IPTG induction, cell disruption with sonication, protein purification with Ni-NTA column with the same protocols. And the obtained EGFP-tagged point-mutated protein was confirmed by SDS-PAGE analysis. The binding assay of the point-mutated tail fiber protein was also performed with the same protocol.

2.15. Isolation of spontaneous phage-resistant mutants

To isolate phage resistant mutants, *S. flexneri* 2a strain 2457T host

strain was incubated up to OD_{600} = 0.5, and then infected with a phage SFPB at a MOI of 1. After incubation at 37℃ for 12 h, 100 µl of infected bacterial culture was plated on fresh LB agar and phage SFPB-resistant mutant colonies were obtained. To purify the phage-resistant mutants, picking of a single colony and its streaking on fresh agar plates were repeated at least 3 times. Finally, the purified phage-resistant mutants were verified using a spot assay with a host strain, *S. flexneri* 2a strain 2457T.

2.16. Data analysis

Statistical analysis was carried out using GraphPad Prism version 9.0.0 Software (GraphPad, USA). Experimental data are presented as the mean ± standard deviation (SD) from three independent experiments. Differences between means were tested by Student's *t*-test. Differences were defined as significant at $p < 0.05$.

2.17. Nucleotide sequence accession number

The complete genome sequences of *S. flexneri*-infecting phage SFPB, SFP17, SFP20, SFP21A, and SFP21B, with its annotation information, were deposited in the National Center for Biotechnology Information database under the GenBank accession no. OQ031072, MN432485, OQ031076, OQ031074, and OQ031073, respectively.

3. Results

3.1. Characterization of bacteriophages

3.1.1. Isolation and morphological determination of bacteriophages

S. flexneri phages were screened by a plaque assay from sewage samples collected from various regions in South Korea. Five phages were isolated, and transmission electron microscopy revealed that a phage SFPB was resembled *Myoviridae* family, and phage SFP17 was resembled *Siphoviridae* family. Phage SFP20, SFP21A and SFP21B were resembled *Podoviridae* family (Fig. 1). Three *Podoviridae* phages were characterized with an icosahedral head (53 ± 1 to 55 ± 6 nm) and a short non-contractile tail (13 ± 3 to 16 ± 2 nm). One *Myoviridae* phage was characterized with isomeric head (84 ± 5 nm) and contracted tail (126 ± 4 nm). One *Siphoviridae* phage was characterized with an icosahedral head (63 ± 1 nm) and noncontractile flexible tail (142 ± 7 nm).

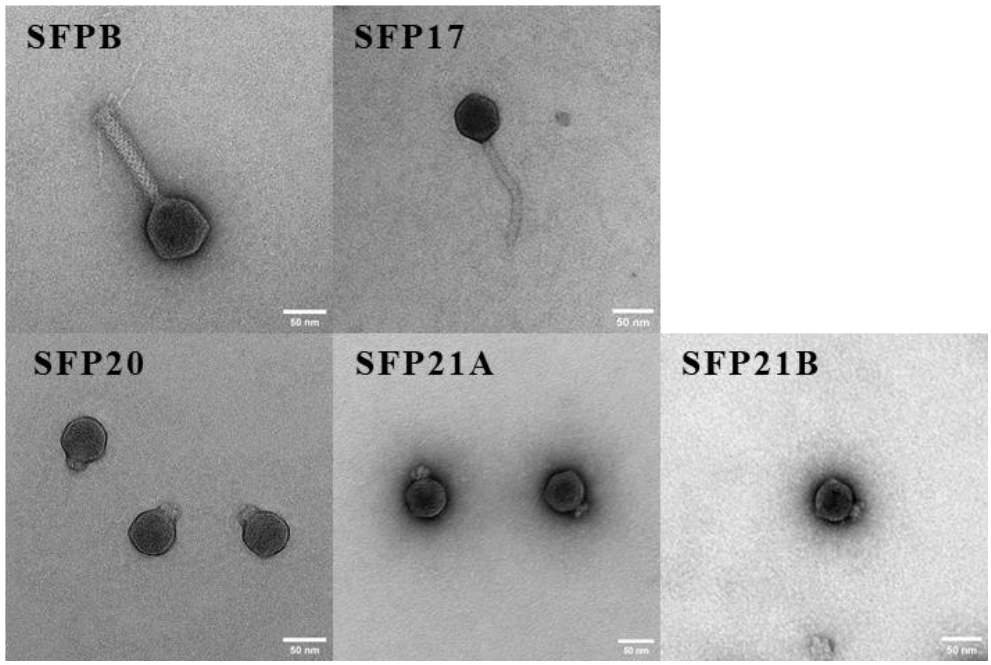

Fig. 1. Transmission electron microscopy image of *Shigella* phages. Each phage name is indicated in the upper left corner of the image. Scale bar, 50 nm.

3.1.2. Host range analysis

The host range analysis of five phages was determined using *S. flexneri* strains and other gram-negative and gram-positive bacteria (Table 6). The result showed that *Podoviridae* phages have narrow host range compared with the other two phages. Interestingly, phage SFP21B infects the K-12 strain of *E. coli* compared with the other two phages. The phage SFPB and SFP17 form clear plaques on *S. flexneri* ATCC 12022. In addition, phage SFPB also infect several strains of *E. coli* O157 and phage SFP17 can infect *E. coli* MG1655, indicating that these bacteriophages have a host range that extends to other *S. flexneri* strain or *E. coli* strains.

Table 6. Host range of *Shigella* phages

Bacterial strains	SFPB	SFP17	SFP20	SFP21A	SFP21B
Shigella flexneri					
2a strain 2457T	+++	+++	+++	+++	+++
ATCC 12022	+++	+++	-	-	-
Shigella sonnei					
KCTC 22530	-	-	-	-	-
Shigella boydii					
KCTC 22528	-	-	-	-	-
Gram-negative bacteria					
Escherichia coli					
O157:H7 ATCC 43895	+	-	-	-	-
O157:H7 ATCC 35150	+	-	-	-	-
O157:H7 ATCC 43888	+	-	-	-	-
O157:H7 ATCC 43890	+	-	-	-	-
MG 1655	-	++	-	-	+
Cronobacter sacazakii					
ATCC 29544	-	-	-	-	-
Pseudomonas aeruginosa					
KACC 10186	-	-	-	-	-
Yersinia enterocolitica					
ATCC 55075	-	-	-	-	-
Salmonella Typhimurium					
ATCC 14028	-	-	-	-	-
DT 104	-	-	-	-	-
Gram-positive bacteria					
Listeria monocytogenes					
ATCC 15313	-	-	-	-	-
Bacillus cereus					
ATCC 14579	-	-	-	-	-
ATCC 13061	-	-	-	-	-
Staphylococcus aureus					
ATCC 29213	-	-	-	-	-
Clostridium perfringens					
ATCC 3624	-	-	-	-	-

+++, EOP of 1 to 0.5; ++, EOP of 0.5 to 0.2; +, EOP less than 0.2; -, no susceptibility to phage

3.1.3. Genome sequence analysis

The complete genome sequence of the phages SFPB, SFP17, SFP20, SFP21A, and SFP21B were analyzed and their general genome features were summarized in Table 7. To understand their evolutionary relationships, comparative phylogenetic analyses with major capsid proteins and tail fiber proteins were performed (Fig. 2). Interestingly, these five phages were categorized into three different groups: group A (SFPB), group B (SFP17), and group C (SFP20, SFP21B, and SFP21A) in both selected proteins. The genome analysis results revealed that these three groups have different genome sizes (A > B > C) and G+C contents (A < B < C), but three phages in group C showed very similar genome sizes and G+C contents, suggesting that these three different phage groups might be originated from different ancestors, while three phages in group C might share the common ancestor. Previous host range analysis of three phage groups showed different host range patterns (Table 6), supporting these different evolutionary relationships from comparative genome analysis among phage groups. In addition, different ORF arrangement patterns among five phage genomes also support this (Fig. 3). However, these five phage genomes have all required functional core proteins, such as structural and package proteins, tail structure proteins, host lysis proteins, and DNA replication/modification proteins, for phage reconstruction and host infection (Table 7). It is noteworthy that all five phage

genomes have no virulence factor gene or antibiotic resistance gene (Fig. 3), suggesting their safety for human and food applications.

Table 7. General genome characteristics of five isolated phages

Characteristics	**SFPB**	**SFP17**	**SFP20**	**SFP21A**	**SFP21B**
Genome size (bp)	88,550	52,151	39,499	39,004	39,356
G + C content (%)	38.6	44.59	53.86	53.25	53.62
Predicted ORFs	131	88	45	47	47
Structure/package proteins	5	6	11	9	9
Tail structure proteins	6	9	3	3	3
Host lysis related proteins	4	4	3	3	3
DNA replication/modification	15	3	9	8	8

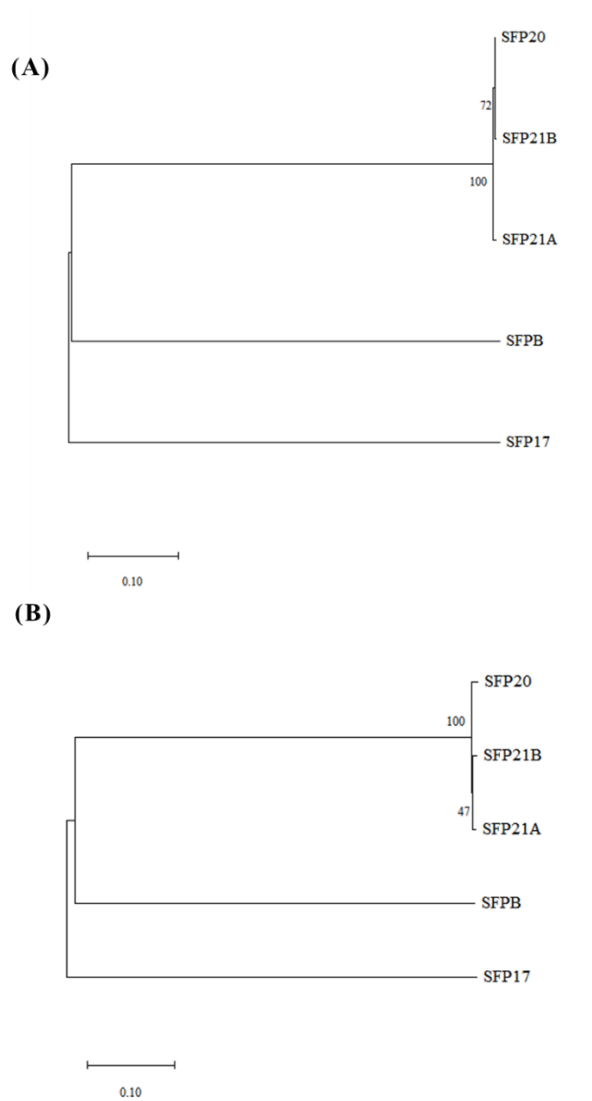

Fig. 2. Comparative phylogenetic analysis of conserved genes from five isolated phages. Phylogenetic tree constructed using MEGA software based on the amino acid sequence of the (A) major capsid protein and (B) tail fiber protein. Bootstrap values are in the phylogenetic tree.

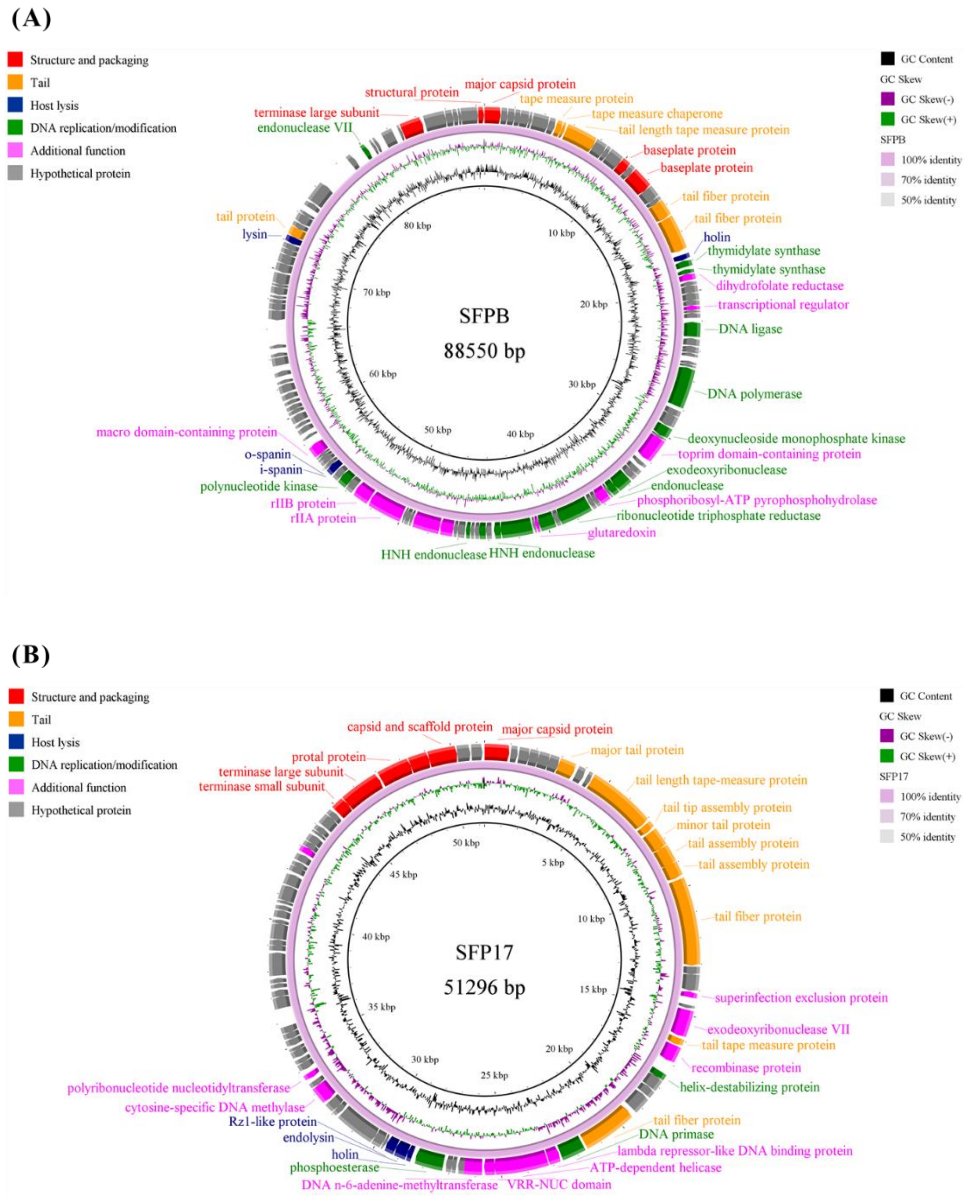

Fig. 3. Genome map of *Shigella* phages. (A) Circular genome map of phage SFPB. (B) Circular genome map of phage SFP17.

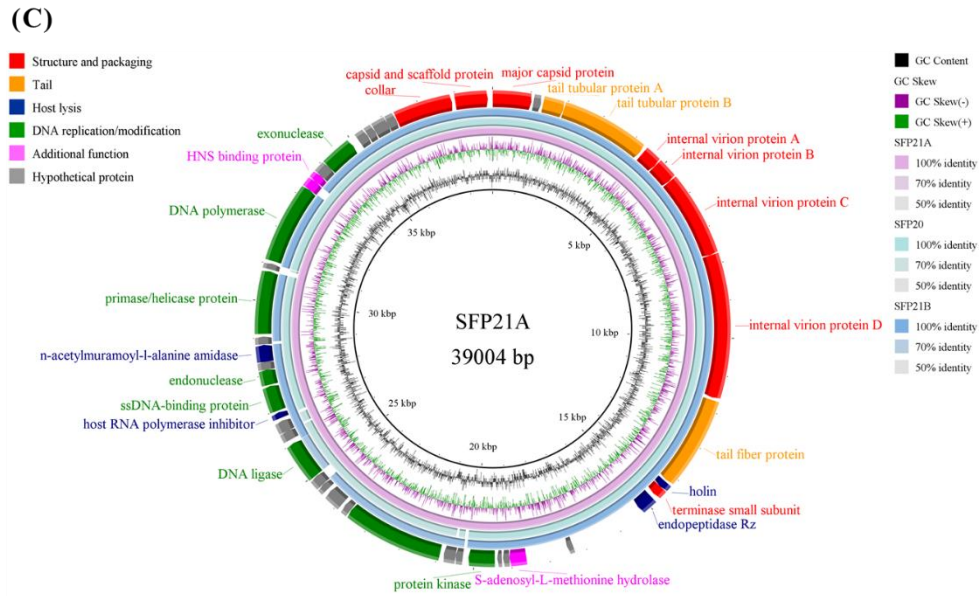

Fig. 3. Cont. Genome map of *Shigella* phages. (C) Comparative circular genome map of the three isolated *Podoviridae* phages. Phage SFP21A was used as the central reference sequence. The rings show BLASTn comparisons of three phage genomes against the SFP21A genome (SFP20, SFP21B).

3.1.4. Identification of host receptor

To identify the host receptors for the isolated phages, six well-known host receptor genes (Parent, Erb et al. 2014, Bertozzi Silva, Storms et al. 2016, Maffei, Shaidullina et al. 2021), which encode the O-antigen of lipopolysaccharides (LPS), FlgK of flagella (flagellar hook-filament junction protein), OmpA and OmpC (Outer membrane proteins), BtuB (Vitamin B12 transporter), and FhuA (ferric ion transporter), were determined and each host receptor gene was deleted to obtain *S. flexneri* mutant without specific host receptor using lambda-red recombination method. Among them, two mutants (*ΔwaaL* and *ΔfhuA*) with infection resistance to phages were selected (Table 8). The *ΔwaaL* mutant showed phage infection resistance to three *Podoviridae* phages (SFP20, SFP21A and SFP21B) and the *ΔfhuA* mutant showed resistance to a *Siphoviridae* phage (SFP17), substantiating that the specific host receptor of those *Podoviridae* phages are the O-antigen of LPS and one *Siphoviridae* phage is the ferrichrome porin. Interestingly, SFP21B showed resistance to *tonB* as well as *waaL* mutant, which suggests that the infection process of SFP21B requires the function of TonB (data not shown). Subsequent complementation experiments using pUHE21-*lac*^{q} expression vector system with cloning of wild-type *waaL* or *fhuA* confirmed that they are host receptors indeed.

To determine the host receptor for SFPB, a random mutant library of

S. flexneri 2a strain 2457T was constructed using Tn*5* transposon (Larsen, Wilson et al. 2002, Gordillo Altamirano, Forsyth et al. 2021). Of ~2,000 insertion mutants in the library, one phage-resistant mutant (*ΔgalU*) was selected. DNA sequencing of the Tn*5* insertion region in the mutant revealed one Tn*5* transposon insertion in the middle of *galU* gene encoding UTP-glucose-1-phosphate uridylyltransferase, regarding LPS outer core biosynthesis. In addition, the complementation experiment using pUHE21-lac^q expression vector system with cloning of wild-type *galU* gene showed the complete recovery of phage infection sensitivity, confirming that it is real host receptor. Therefore, each phage group recognizes and infects specific host receptor (outer core of LPS for group A, FhuA for group B, and O-antigen of LPS for group C), indicating that they do not share specific host receptor (Table 1). Therefore, these different host recognition and infection patterns among phage groups also support their different evolutionary relationships.

Table 8. Susceptibility of wild-type and mutant strains to *Shigella* phages

Phage	**Receptor**	***S. flexneri* 2a strain 2457T**				
		Wild-type	*ΔwaaL*	*galU*::Tn*5*	*ΔwaaG*	*ΔfhuA*
SFP17	Outer membrane	++	++	++	++	-
SFPB	LPS	++	++	-	-	++
SFP20	LPS	++	-	-	-	++
SFP21A	LPS	++	-	++	-	++
SFP21B	LPS	++	+	++	+	++

++, EOP of 1 to 0.5; +, EOP less than 0.5; -, no susceptibility to phage

3.2. Characterization of phage SFP21A using bacterial LPS mutants

3.2.1. LPS structure and spot formation of gene deletion mutants

The LPS is a gram-negative outer membrane component and is a major compound linked by three domains: lipid A, core oligosaccharide and O-antigen. Interestingly, although the phages SFP21A and SFP21B cannot infect the *ΔwaaL* mutant, due to O-antigen of LPS as a host receptor, they can infect the *galU*::Tn*5* mutant, suggesting that they may recognize and infect O-antigen as well as inner core of LPS (Table 8 and Fig. 4A). However, phage SFP21B had a receptor on the outer membrane, so it could be attached to parts other than the LPS. Based on these results, phage SFP21A was selected and characterized throughout the study.

The mutant of the *waaG* gene encoding glucose transferase completely loses the same outer core as *galU* mutant. However, previous studies have shown that the *galU*::Tn*5* mutant additionally lacks of the third heptose and N-acetylglucosamine substituent attached to the inner core of LPS (Molinaro, Silipo et al. 2008). Therefore, to determine whether the change in binding pattern is due to these additional deficiencies, the deletion in *waaG* gene was constructed and structure was confirmed by SDS-PAGE (Fig. 4B). Then, as a result of the spotting assay, surprisingly, the mutant of the *waaG* completely lost the same outer core as *galU*::Tn*5* mutant but did not form plaques when treated with SFP21A phage (Fig. 5). These

observations suggest that the GlcNAc substituent may interfere with host cell binding of SFP21A phage.

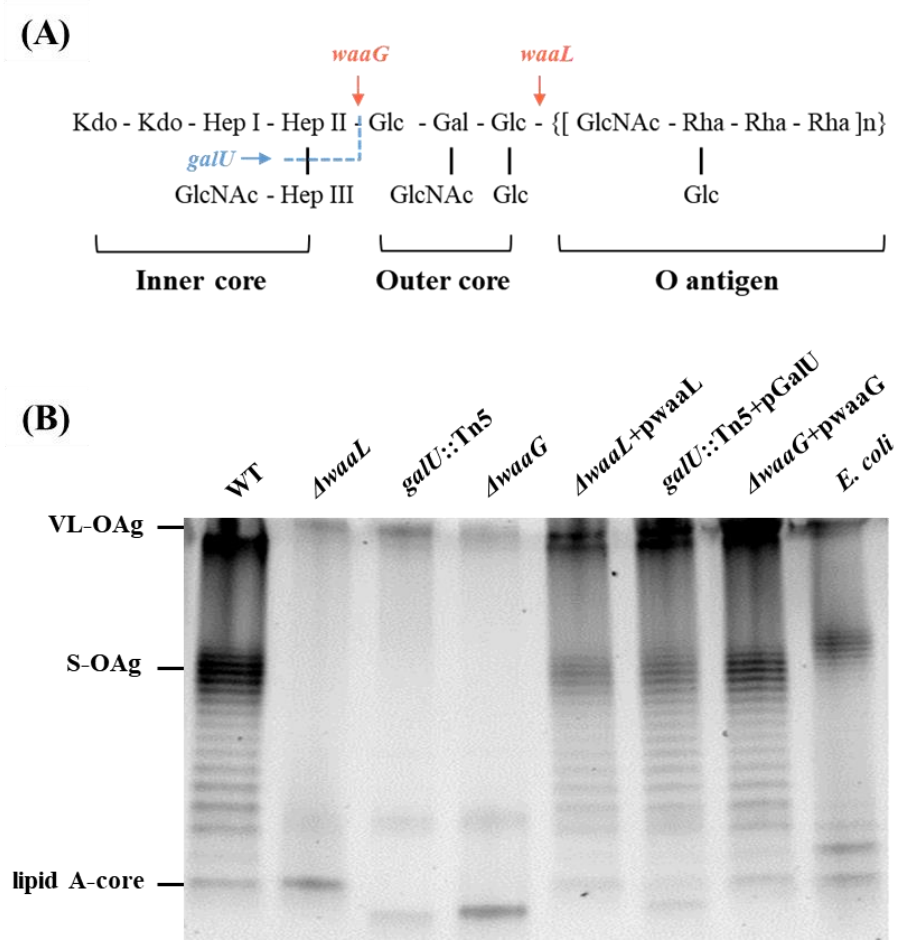

Fig. 4. SDS-PAGE of LPS from *S. flexneri* 2a strain 2457T and its mutants. (A) Structure of the LPS from *S. flexneri* 2a strain 2457T. Kdo 3-deoxy-D-manno-oct-2-ulosonic acid, Hep heptose, Glc glucose, Gal galactose, GlcNAc N-acetylglucosamine (B) SDS-PAGE of LPS from *S. flexneri* 2a strain 2457T and its mutants.

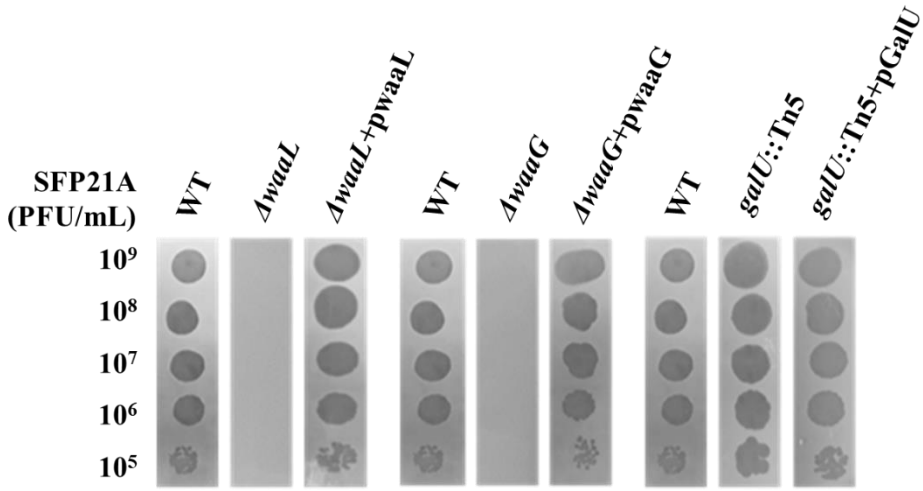

Fig. 5. Effect of various *S. flexneri* 2a strain 2457T mutants on SFP21A susceptibility. SFP21A spotting assay with three different LPS mutants. Serially diluted SFP21A lysates were spotted on the lawns of indicated *Shigella* mutants.

3.2.2. Inactivation assay of SFP21A by LPS

To confirm O-antigen of LPS is the receptor of phage SFP21A, we tested the effect of the extracted LPS on the infection by SFP21A. As shown in Fig. 6, preincubation of SFP21A with the LPS extracted from WT greatly reduced the EOP, whereas the LPS extracted from *waaL* mutant strain showed no such effect. These results showed that blocking SFP21A with the extracted LPS from WT significantly inhibited the efficiency of infection. Therefore, these results revealed that O-antigen of LPS is the host receptor for phage SFP21A adsorption.

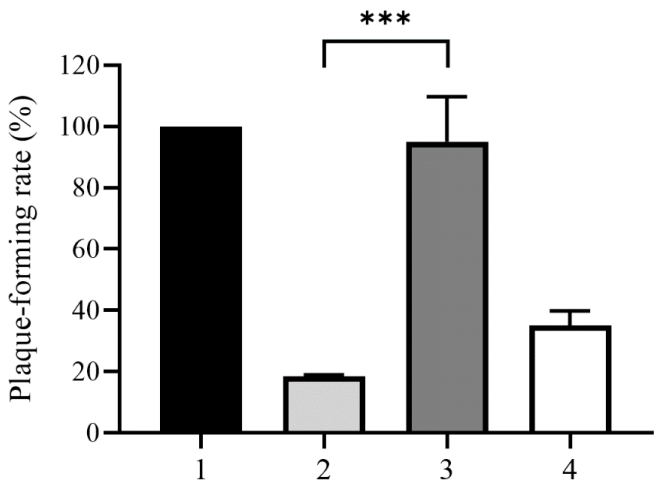

Fig. 6. Inactivation of SFP21A by LPS. Competitive binding assay is shown by plaque-forming rates of phage SFP21A preincubated with LPS extracted from different strains. Bars are as follows: 1: LPS-untreated; 2: WT-LPS treated; 3: *ΔwaaL*-LPS treated; 4: *ΔwaaL*+pWaaL-LPS treated. PFU of the group without LPS (black bar) was set as 100%. The experiments were independently replicated three times. Error bars show standard by Student's t test for comparison between the mutant and the WT group. ***, $p < 0.001$.

3.2.3. Bacteriophage adsorption assays

The initial step of phage infection is the attachment and adsorption to a susceptible host cell. Therefore, to verify whether SFP21A is affected by the mutation in the first step of infection, phage adsorption analysis was performed using *S. flenxeri* 2a strain 2457T and their LPS mutants such as *waaL*, *waaG* and *galU*. In the case of *waaL* and *waaG* mutants, it was confirmed that the result of no plaque activity was also reflected in the kinetics of adsorption rate. Interestingly, *galU*::Tn*5* mutant formed plaques similar to the wild-type, but the rate or extent of adsorption was significantly slower and lower than that of WT (Fig. 5 and Fig. 7). These results suggest that the third heptose residue and GlcNAc present in the inner core play an important role for adsorption of SFP21A.

To clarify whether the host receptor for phage SFP21A is a carbohydrate or protein, the host bacterial cells were treated with periodate or proteinase K prior to the phage adsorption. Proteinase K can degrade peptide bonds in cell surface protein structures, such as outer membrane proteins, while periodate cleaves carbohydrate structures containing a diol motif, such as oligosaccharides. When 10- or 100-mM periodate was treated to the host *S. flexneri* 2a strain 2457T cells, the adsorption of SFP21A was significantly inhibited compared to the non-targeted group (Fig. 8). In contrast, the group treated with protease K, the phage particles were adsorbed to the cells without

significant difference from the untreated group, and these results show that the protein substances did not serve as a host receptor.

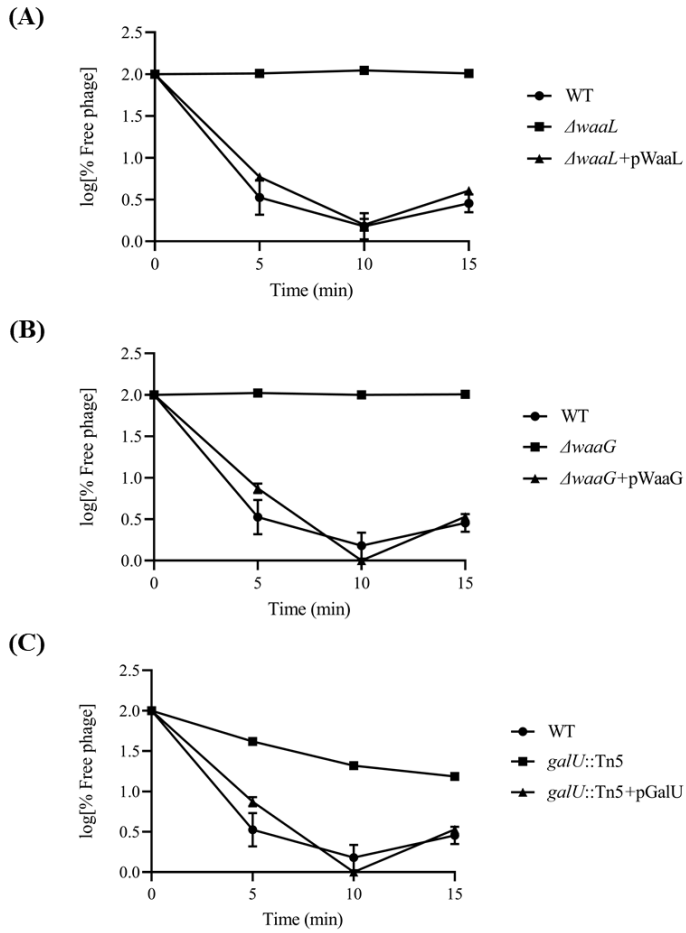

Fig. 7. Effect of various *S. flexneri* 2a strain 2457T mutants on SFP21A adsorption. (A) *ΔwaaL* and *ΔwaaL*+pWaaL strains. (B) *ΔwaaG* and *ΔwaaG*+pWaaG strains. (C) *galU*::Tn*5* and *galU*::Tn*5*+pGalU strains. Percent free phage was determined as a ratio of free phage at the time point divided by the total phage added at the beginning of the assay. Error bars present the standard deviations of three replicates.

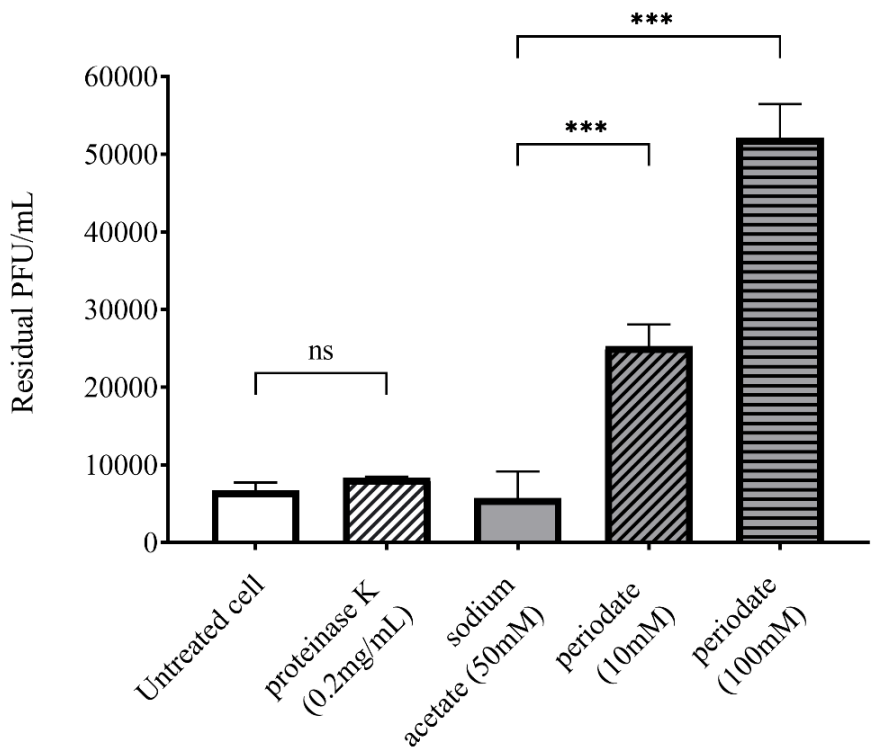

Fig. 8. Effect of periodate and proteinase K treatments on SFP21A adsorption. Cells suspended in LB (Untreated) and cells in acetate buffer were used as control. Error bars present the standard deviations of three replicates. Asterisks indicate significant differences (***P < 0.001).

3.2.4. Bacterial challenge assay

To evaluate bacterial growth inhibition activity of phage SFP21A, growth of *S. flexneri*, *galU*::Tn*5* mutant, and their complementation strain was periodically monitored after treatment of SFP21A at an MOI of 0.01. When the exponentially growing *S. flexneri* was infected by SFP21A, bacterial growth inhibition was shown at 1 hr post-infection and was continued for 8 hrs (Fig. 9). When the growing *galU*::Tn*5* mutant was infected by SFP21A, mutant strain showed sustained inhibition of bacterial cell growth, although it was not inhibited as effectively as the WT. The complementation of this gene restored the inhibitory ability of the phage SFP21A.

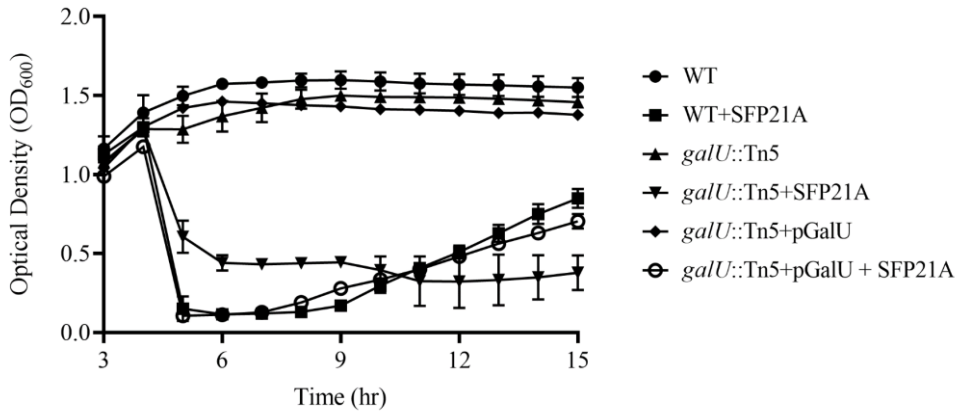

Fig. 9. Bacterial challenge assay of phage SFP21A against *S. flexneri* 2a strain 2457T and *galU*::Tn*5* mutant. The phage was added at a MOI of 0.01 to the bacterial culture after 3 hr incubation. The data are expressed as means and standard deviations of individual triplicate assays.

3.2.5. Binding activity of EGFP fusion protein

Phage tail fiber proteins recognize the host cell receptors and thus play an important role in the receptor binding process and in determining host specificity (Garcia-Doval and Raaij 2013). To understand the interaction of host binding and adsorption by the tail proteins of SFP21A phage, binding activity to the host bacteria and their mutant was performed. The fusion protein of SFP21A tail fiber protein and EGFP, designated as EGFP-Gp09, was expressed and purified (Fig. 10A). The results showed that EGFP-Gp09 bound to *Shigella flexneri* 2a strain 2457T and *galU*::Tn*5* mutant (Fig. 11). Interestingly, the recognition ability of mutant cells by EGFP-Gp09 was different from that of wild-type, which is the same as the previous decrease in adsorption rate (Fig. 7C).

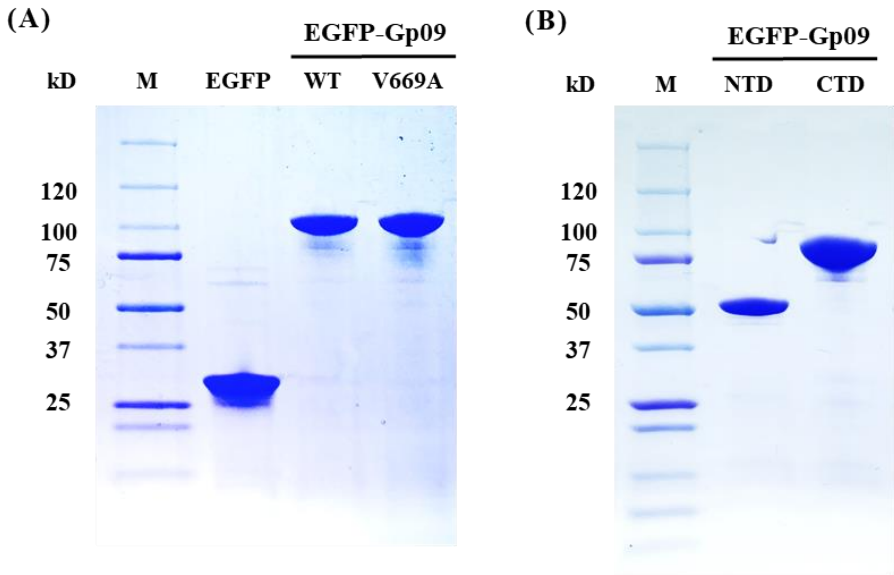

Fig. 10. SDS-PAGE analysis of purified EGFP-fused proteins. (A) SDS-PAGE of purified EGFP (26.9 kDa), SFP21A EGFP-Gp09 (118.13 kDa), and of EGFP-Gp09$_{V669A}$. Lane 1, protein standard (Bio-Rad); lane 2, purified EGFP; lane 3, purified EGFP-Gp09; lane 4, purified EGFP-Gp09$_{V669A}$. (B) SDS-PAGE of purified SFP21A EGFP-Gp09$_{NTD}$ (51.1 kDa) and of the EGFP-Gp09$_{CTD}$ (93.94 kDa). Lane 1, protein standard (Bio-Rad); lane 2, purified EGFP-Gp09$_{NTD}$; lane 3, purified EGFP-Gp09$_{CTD}$.

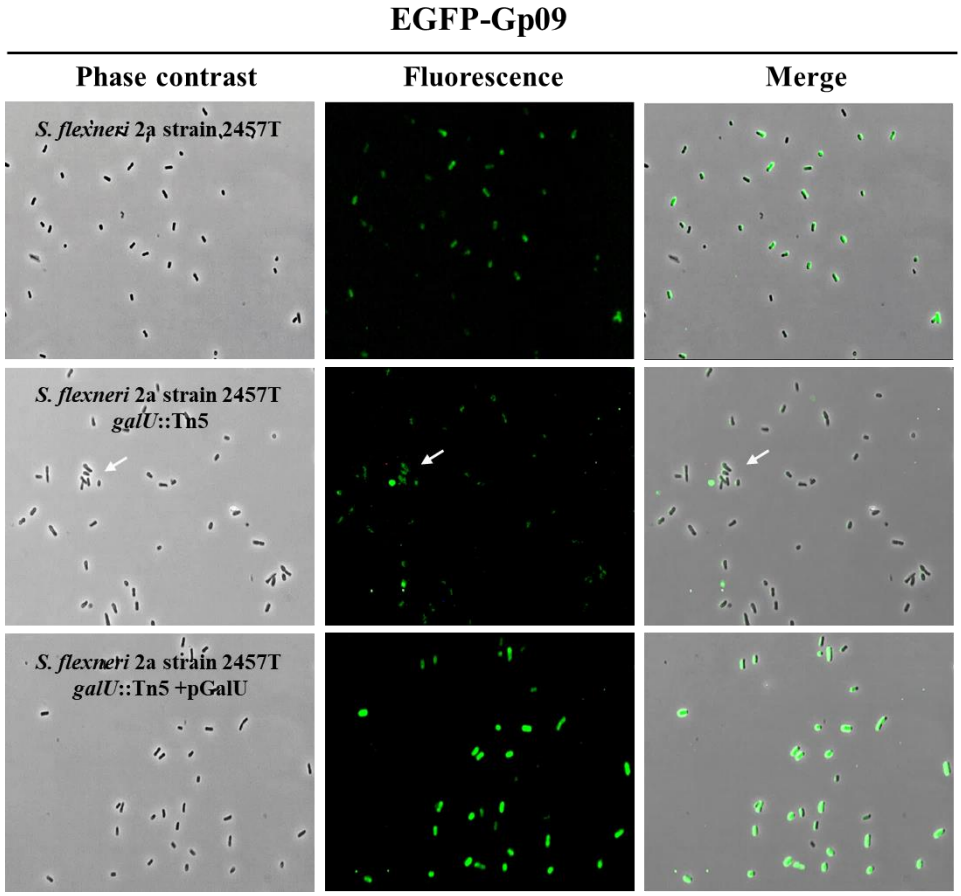

Fig. 11. Binding activity of EGFP-Gp09. The ability of this protein to bind *Shigella* cells with *galU*::Tn*5* and *galU*::Tn*5*+pGalU was visualized by confocal microscopy.

3.3. Identification of phage SFP21A receptor binding protein

3.3.1. Bioinformatic analysis of Gp09

To find out which part of the phage tail protein affects this binding pattern, protein structure analysis was performed using the structural prediction program AlphaFold 2 (Jumper, Evans et al. 2021, Varadi, Anyango et al. 2022). The SFP21A tail fiber protein was predicted to have a complete trimeric structure with three distinct domains including an N-terminal domain, a neck, and a C-terminal domain (Fig. 12A and 12B). In addition, the predicted AlphaFold structure showed structural similarities between the N-terminal of the phage SFP21A tail fiber protein and the N-terminal of the phage T7 tail fiber protein. In addition, a portion of the C-terminal domain of SFP21A (amino acid 262-461) was found to be structurally similar to a lipolytic enzyme. However, the tip of the C-terminal domain had no homology with other phage tail fiber proteins or spike proteins (Fig. 12C). These results suggest that the C-terminal domain of SFP21A Gp09 may be related to the receptor binding region.

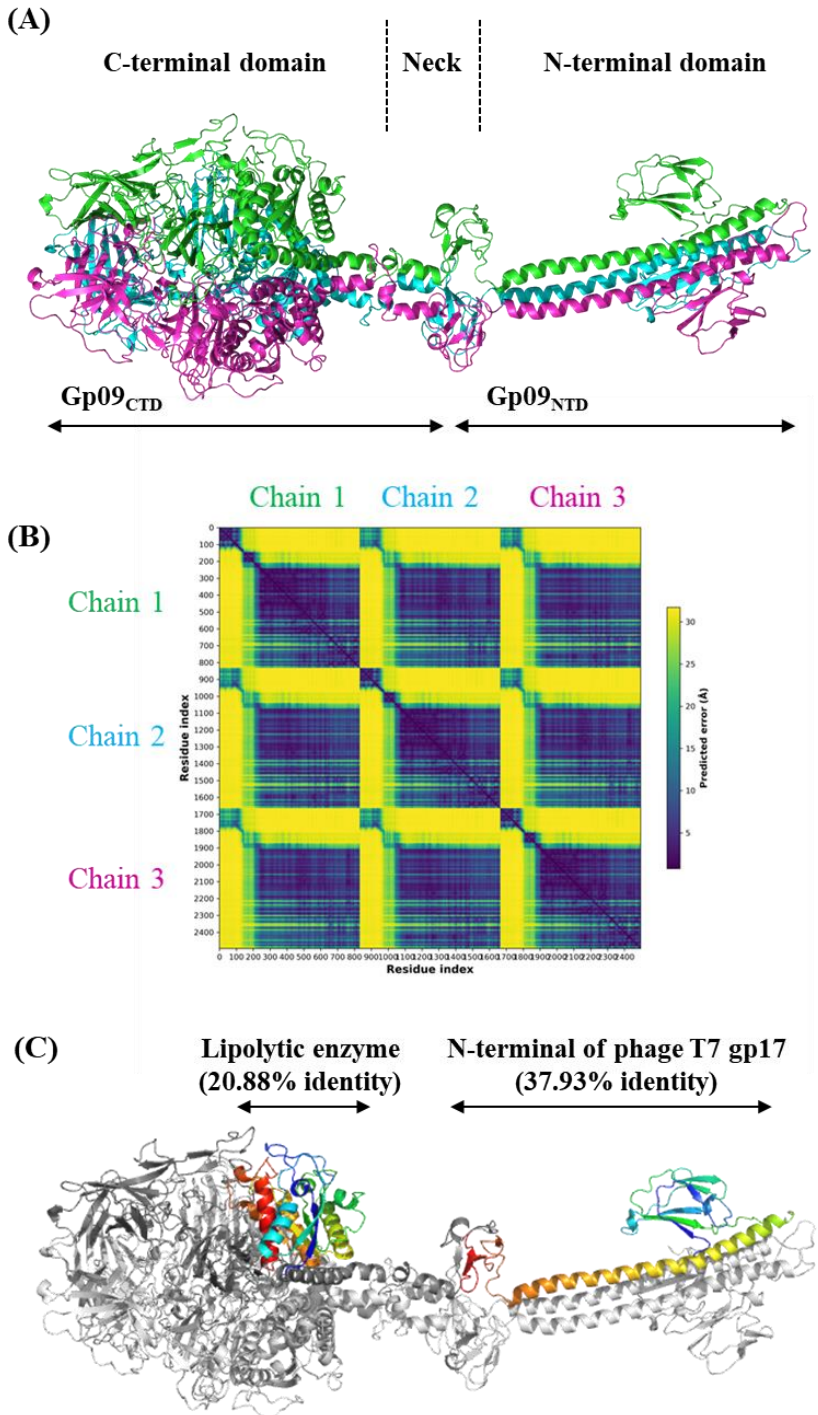
(A)
C-terminal domain
Neck
N-terminal domain
Gp09CTD
Gp09NTD
(B)
Chain 1
Chain 2
Chain 3
Residue index
Predicted error (Å)
(C)
Lipolytic enzyme
(20.88% identity)
N-terminal of phage T7 gp17
(37.93% identity)

Fig. 12. Predicted 3D structure of tail fiber protein of phage SFP21A.

(A) Predicted AlphaFold structure of the Gp09 homotrimeric complex with individual chains colored blue, pink and green. (B) Predicted aligned error (PAE) plot of AlphaFold structure of the Gp09 homotrimeric complex model. Note the high confidence prediction (low PAE) for the interface between each chain. (C) Structural similarity to the N-terminal and portion of the C-terminal regions of SFP21A.

3.3.2. Binding activity of EGFP-Gp09$_{CTD}$ and EGFP-Gp09$_{NTD}$

To verify the carboxyl terminal domain of Gp09 has a direct effect on cell binding, the EGFP-Gp09 fusion proteins using the C-terminal domain (amino acid 228-832) and N-terminal domain (amino acid 1-227) of SFP21A tail fiber protein, designated as EGFP-Gp09$_{CTD}$ and EGFP-Gp09$_{NTD}$, were expressed and purified (Fig. 10B). The observed bacterial binding indicated that unlike the N-terminal domain (NTD), positive binding by EGFP-Gp09$_{CTD}$ revealed that the C-terminal domain confers receptor binding (Fig. 13).

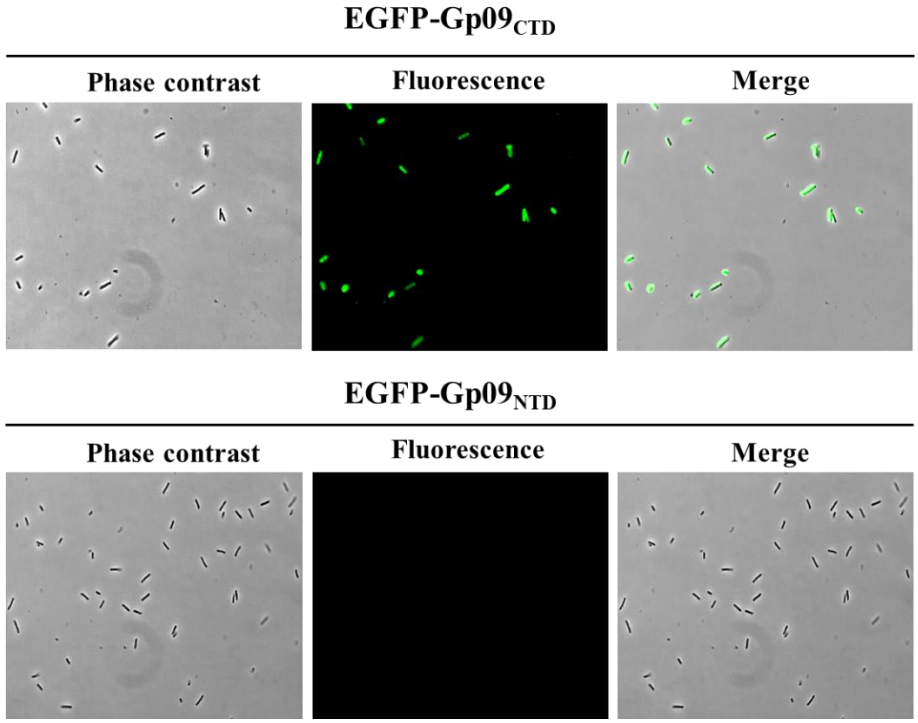

Fig. 13. Binding activity of EGFP-Gp09$_{CTD}$ and EGFP-Gp09$_{NTD}$.

The ability of these proteins to bind *Shigella flexneri* 2a strain 2457T was visualized by confocal microscopy.

3.3.3. Binding activity of EGFP-Gp09$_{V669A}$

To determine which amino acids of the tail fiber protein of SFP21A interact with the *Shigella galU*::Tn*5* mutant, the isolated three *Podoviridae* phage tail fiber proteins were compared by multiple sequence alignment (Fig. 14). The result shows that the tail fiber proteins share 98.68 to 98.92% identity at the protein level. Among them, there were seven different amino acid regions of the tail fiber protein, which is the same as the phage infection pattern in the *galU*::Tn*5* mutant: SFP21A and SFP21B with clear plaques and SFP20 without plaques. In addition, host-receptor interaction changes most likely occur in the distal loop of the C-terminal domain of the tail fiber protein (Washizaki, Yonesaki et al. 2016, Dunne, Rupf et al. 2019, Islam, Fokine et al. 2019). This suggested that V669 could be responsible for the *galU*::Tn*5* mutant infection pattern. To verify whether changes in amino acid alter binding activity to the host bacteria and their mutant, the fusion protein using EGFP and the single point mutation of SFP21A tail fiber protein, designated as EGFP-Gp09$_{V669A}$ was expressed and purified (Fig. 10A). The results showed that EGFP-Gp09$_{V669A}$ bound to the wild-type but failed to bind to the *galU*::Tn*5* mutant (Fig. 15). These suggest that V669 located in the C-terminal domain plays an important role in binding *galU*::Tn*5* mutant.

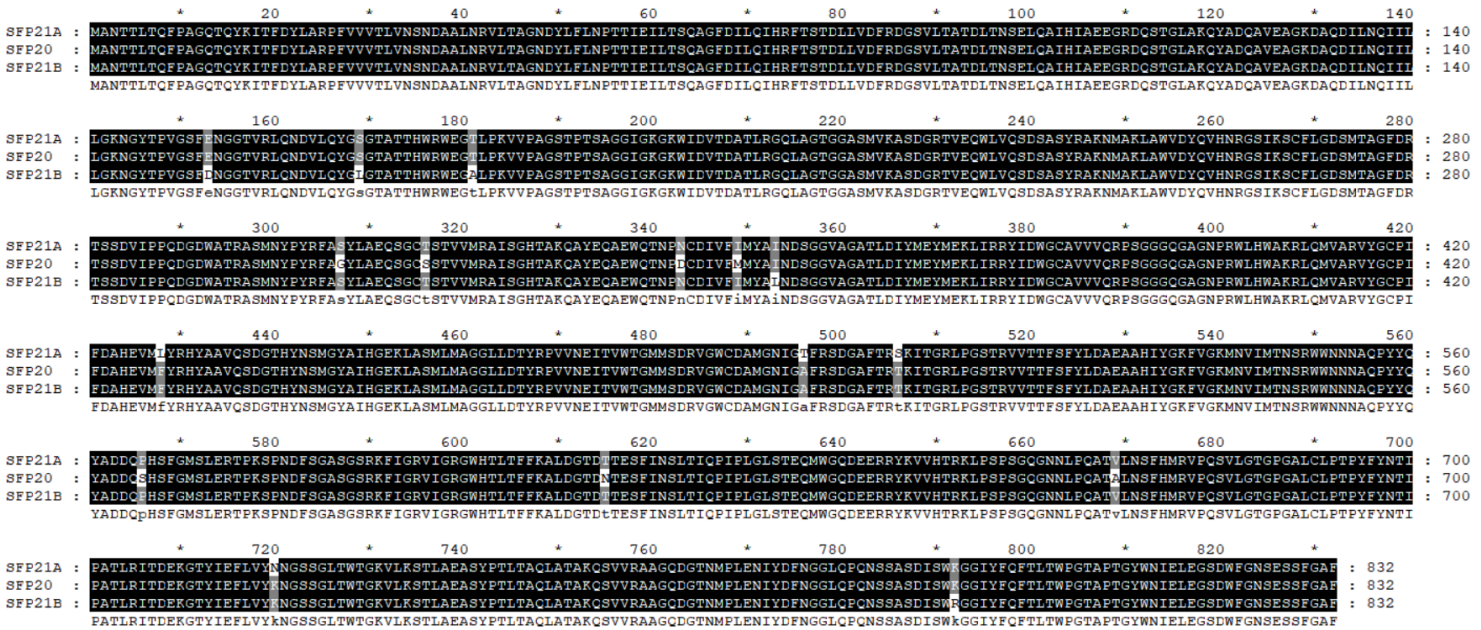

Fig. 14. Amino acid sequence alignment among Gp09 related tail fiber protein.

Amino acid alignment of SFP21A with *S. flexneri* phage SFP20 tail fiber protein and *S. flexneri* phage SFP21B tail fiber protein.

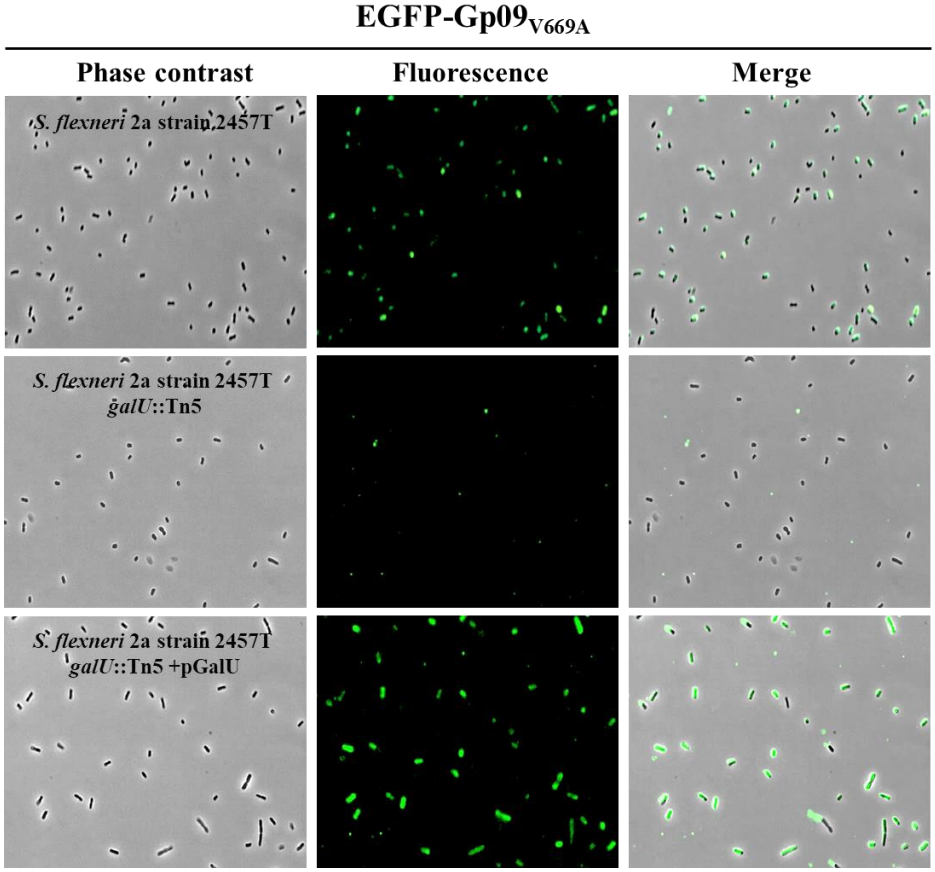

Fig. 15. Binding activity of EGFP-Gp09$_{V669A}$.

The ability of these proteins to bind *Shigella* cells with *galU*::Tn*5* and *galU*::Tn*5*+pGalU was visualized by confocal microscopy.

3.4. Enhanced inhibition of *Shigella* by three different phages

3.4.1. Susceptibility of BIMs to other receptor-targeting phages

Bacteria often modify their surface structures to avoid phage adsorption and become resistant to phage. In a previous study, a phage cocktail using three different host receptors including O-antigen of LPS, outer membrane protein BtuB and flagella was able to reduce the emergence of phage-resistant *Salmonella* Typhimurium (Bai, Jeon et al. 2019). The receptors of the five isolated *Shigella* phages are mostly LPS, but receptor assay was predicted that unlike the phages of other O-antigen receptor groups, phage SFP21A will be able to independently infect hosts using their own receptors regardless of the development of bacterial resistance to phage SFPB. To investigate whether a similar phenomenon could occur, spontaneous host-resistant SFPB mutants were isolated through the high-titer overlay assay as previous described (Hesse, Rajaure et al. 2020). Twelve mutants were obtained, and the resistance against SFPB infection was verified by the spot assay (data not shown). When the BIM strains were treated with SFP20, which had high activity among the same O-antigen receptor group, and SFP21A, respectively, most of the BIM strains appeared by SFPB treatment were more effectively infected with SFP21A than SFP20 (Fig. 16).

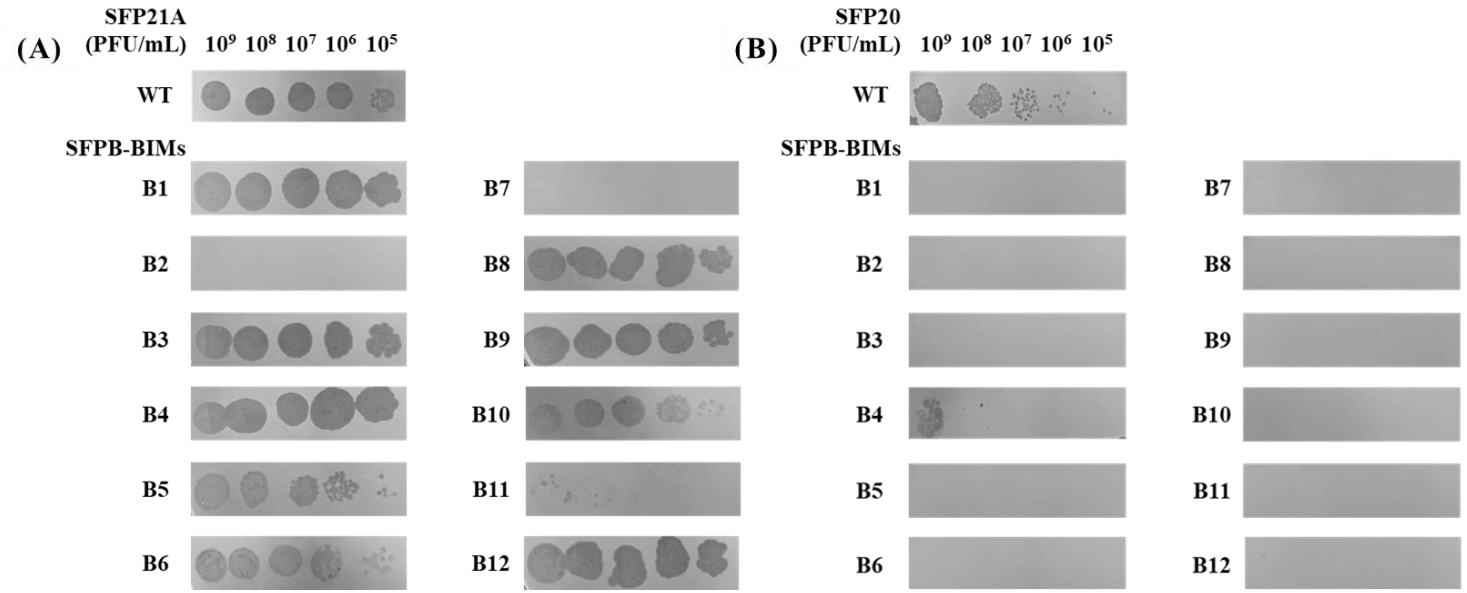

Fig. 16. Susceptibility of the mutants resistant to LPS core-specific phages.

The ability of the two O-antigen receptor group phages to infect BIMs obtained after other SFPB phage tre atment. (A) phage SFP21A and (B) phage SFP20.

3.4.2. Growth inhibition assay by a phage cocktail

The BIMs susceptibility analysis suggests that a phage-cocktail including SFP21A and SFPB may have a synergistic effect on host growth inhibition. Therefore, the combination of SFP17 phage, which had the outer membrane protein FhuA as a receptor, SFP21A and SFPB was suggested as a phage of the cocktail to reduce the development of phage-resistant mutants. Although SFP21A phage and SFPB phage were the same LPS receptor phages, treatment of three phages was enhanced lytic activity and diminished emergence of phage-resistant mutant than treatment with each phage (Fig. 17). These results suggest that even when the same LPS receptor is used as a cocktail, it can be used as a promising strategy for controlling *S. flexneri* because it can effectively inhibit the growth of bacteria according to the specific receptor pattern of SFP21A.

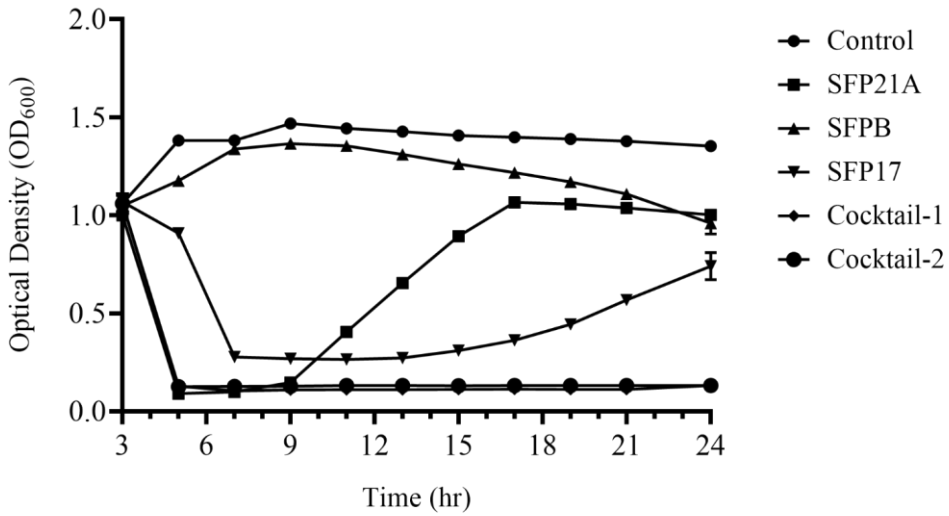

Fig. 17. Bacterial challenge assay with phage cocktail.

Growth curves of *S. flexneri* 2a strain 2457T in the absence and presence of phage predation. Cultures were grown in the absence of phage or with SFP21A, SFPB, SFP17, at a multiplicity of infection of 0.01 or the phage cocktail at a multiplicity of infection of 0.01 or 0.001. Error bars present the standard deviations of three replicates.

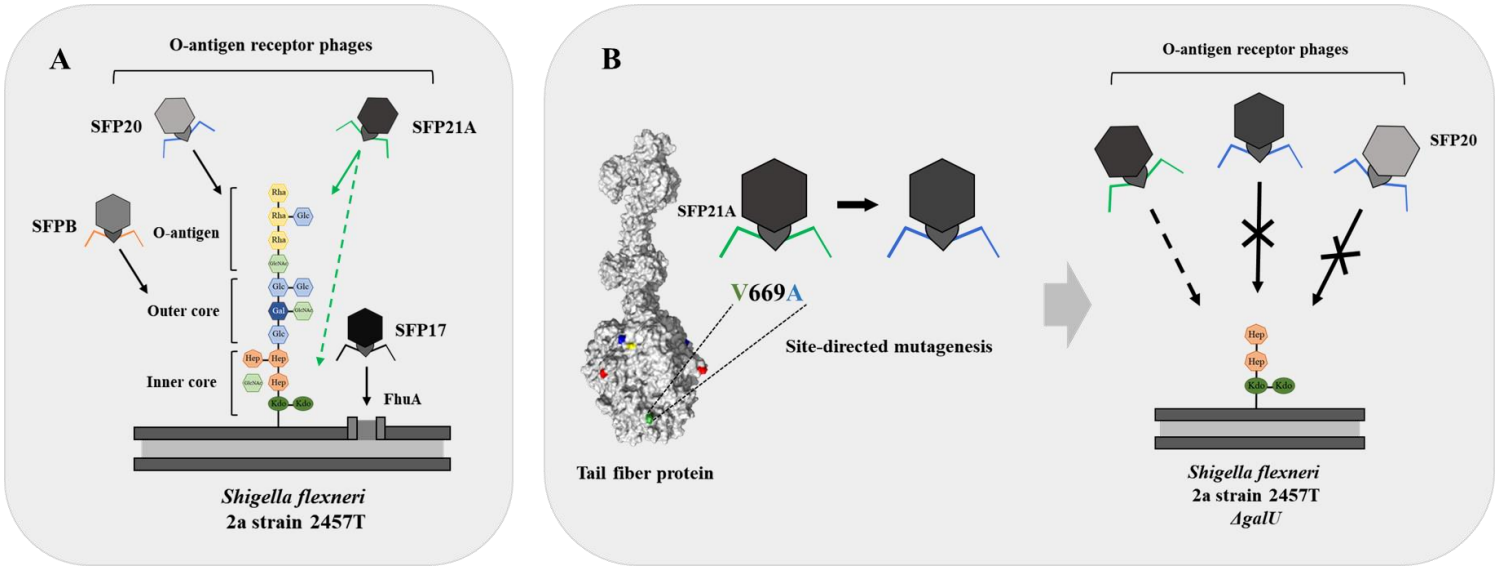

Fig. 18. Proposed model of phage recognition for the host receptor.

4. Discussion

Bacteria of the *Shigella* group are human pathogens that cause an intestinal infection called shigellosis, and infections usually through unsafe water or contaminated food. Phages have long been considered an alternative antimicrobial agent for the biocontrol of *Shigella flexneri* but the evolution of phage resistance bacteria should be overcome to develop more practical phage therapies (Llanos-Chea, Citorik et al. 2019). Biological control of bacteria with phage cocktails more effectively inhibits the emergence of phage resistance when the cocktails are composed of phages using different receptor types (Kim, Kim et al. 2014, Bai, Jeon et al. 2019). Also, since the emergence of phage resistance in most cases occurs by modifying their receptors of the cell surface, understanding the RBP-host receptor interaction is an important prerequisite for the development of effective phage therapies against pathogens. However, until the present, *Shigella* infecting phages have not been extensively studied, and their host receptors have also not yet been widely characterized. In this study, identification of various host receptors and structural analysis of the RBP of phage SFP21A showing a novel binding pattern of LPS were performed to suppress the development of phage

resistance.

Isolation of five phages infecting *Shigella flexneri* from various sewage treatment plants were characterized and determined that they have distinct host ranges, and genome sequences. In addition, five novel virulent phages targeting different receptors on *S. flexneri* (i.e., O-antigen, LPS outer core, and FhuA) were identified by screening phages against a library of bacterial mutants (Fig. 18A). A very interesting results is that the O-antigen receptor phage SFP21A was able to infect *galU* mutant that lacks all the outer core of LPS (Fig. 18A). The gene cluster of *waa* (*rfa*) are involved in the biosynthesis of the inner core and outer core region of the LPS. Among them, the *waaG* gene encoding glucose transferase is characterized by a complete loss of the same outer core as the *galU* mutant. However, phage SFP21A was able to infect the *galU*::Tn*5* mutation, but not the mutant of the *waaG* gene. Previous studies have shown that the *galU* mutation additionally lacks a third heptose and N-acetylglucosamine substituent attached to the inner core of LPS (Molinaro, Silipo et al. 2008). These results suggest that the GlcNAc substitution may interfere with host cell binding of SFP21A phage.

Bacteriophages have the characteristic of specifically recognizing target bacteria using receptor binding protein (RBP). Among them, the C-terminal domain of most tail fiber proteins that form trimers enables host recognition or receptor binding (Dunne, Rupf et al. 2019, Abdelkader,

Gutiérrez et al. 2022). Previous studies have shown that T3 phage initially binds to bacterial LPS through the C-terminus of the tail fiber for host recognition (Ando, Lemire et al. 2015). Similarly, the C-terminal domain of the Lambda phage tail fiber protein is responsible for the recognition of the receptor protein LamB (Spinelli, Veesler et al. 2014). In this study, the protein prediction for RBP of SFP21A showed that it is composed of a homotrimer of the gene *gp09* and the C-terminal domain (residues 228-832) is involved in receptor binding. In addition, as a result of binding analysis by modifying a specific amino acid of the C-terminal domain, it was confirmed that the binding pattern may be changed even by one amino acid present in the outer loop (Fig. 18B). Previous studies have shown that single point mutations in *Listeria* phage PSA RBP (S302R, I306K/R, A332V, S334R and S354T) transformed the host range from *Listeria* SV 4b to 4d, supporting this finding (Dunne, Rupf et al. 2019). These changes suggest that mutation V669A does not alter hydrophobicity or polarity, respectively, but the feature of having slightly larger side chains may have an effect.

In general, bacteria modify surface receptors to develop defenses against phages. As a result of comparing the sensitivity of two o-antigen receptor phages to SFPB BIM having an LPS outer core as a receptor, only SFP21A, which showed a re-binding pattern in the *galU*::Tn*5* mutant, effectively inhibited bacteria. This novel infection pattern of LPS suggests

that even the same LPS receptor group can kill bacteria more effectively when treated together. In addition, phage cocktail was designed that consists of three different receptor phages including SFP21A in *S. flexneri* 2a strain 2457T to delay the emergence of bacterial resistance. The results showed that treatment with a combination of phages targeting three different receptors inhibited bacterial growth much more than treatment with each single phage. Therefore, the synergistic effect of these three phages, SFP21A, SFPB, and SFP17, in killing *S. flexneri* was demonstrated. These results are similar to those of previous studies showing that combinations of phages targeting different host receptors can reduce the likelihood and incidence of phage-resistant mutants ((Kim, Kim et al. 2014, Bai, Jeon et al. 2019).

These results suggest that the preparation of phage cocktails that target multiple host receptors would be helpful for the development of effective alternative strategies to control *S. flexneri* with phages instead of using conventional antibiotics. It can also provide insight into a more detailed understanding of the functional and structural interactions between host receptor and tail fiber protein.

Reference

Abdelkader, K., D. Gutiérrez, A. Latka, D. Boeckaerts, Z. Drulis-Kawa, B. Criel, H. Gerstmans, A. Safaan, A. S. Khairalla and Y. Gaber (2022). "The Specific Capsule Depolymerase of Phage PMK34 Sensitizes Acinetobacter baumannii to Serum Killing." Antibiotics **11**(5): 677.

Alikhan, N.-F., N. K. Petty, N. L. Ben Zakour and S. A. Beatson (2011). "BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons." BMC genomics **12**(1): 1-10.

Altermann, E. and T. R. Klaenhammer (2003). "GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes." Omics A Journal of Integrative Biology **7**(2): 161-169.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic acids research **25**(17): 3389-3402.

Ando, H., S. Lemire, D. P. Pires and T. K. Lu (2015). "Engineering modular viral scaffolds for targeted bacterial population editing." Cell systems **1**(3): 187-196.

Bai, J., B. Jeon and S. Ryu (2019). "Effective inhibition of Salmonella Typhimurium in fresh produce by a phage cocktail targeting multiple host receptors." Food microbiology **77**: 52-60.

Bertozzi Silva, J., Z. Storms and D. Sauvageau (2016). "Host receptors for bacteriophage adsorption." FEMS microbiology letters **363**(4): fnw002.

Besemer, J., A. Lomsadze and M. Borodovsky (2001). "GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions." Nucleic acids research **29**(12): 2607-2618.

Bhattacharya, D., A. Sugunan, H. Bhattacharjee, R. Thamizhmani, D. Sayi, K. Thanasekaran, S. P. Manimunda, A. Ghosh, A. Bharadwaj and M. Singhania (2012). "Antimicrobial resistance in Shigella-rapid increase & widening of spectrum in Andaman Islands, India." The Indian journal of medical research **135**(3): 365.

European Commission (2003). IP/03/1058 Council and Parliament Prohibit Antibiotics as Growth Promoters: Commissioner Byrne Welcomes Adoption of Regulation on Feed Additives. European Commission: Brussels, Belgium, 2003.

Centers for Disease Control and Prevention (CDC). (2018). Shigella-Shigellosis [Internet]. The Center for Disease Control and Prevention. https://www. cdc. gov/shigella/general-information. html.

Criscuolo, E., S. Spadini, J. Lamanna, M. Ferro and R. Burioni (2017). "Bacteriophages and their immunological applications against infectious threats." Journal of immunology research **2017**.

Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." Proceedings of the National Academy of Sciences **97**(12): 6640-6645.

De Sordi, L., M. Lourenço and L. Debarbieux (2019). "The battle within: interactions of bacteriophages and bacteria in the gastrointestinal tract." Cell host & microbe **25**(2): 210-218.

Dedrick, R. M., C. A. Guerrero-Bustamante, R. A. Garlena, D. A. Russell, K. Ford, K. Harris, K. C. Gilmour, J. Soothill, D. Jacobs-Sera and R. T. Schooley (2019). "Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus." Nature medicine **25**(5): 730-733.

Dunne, M., B. Rupf, M. Tala, X. Qabrati, P. Ernst, Y. Shen, E. Sumrall, L. Heeb, A. Plückthun and M. J. Loessner (2019). "Reprogramming bacteriophage host range

through structure-guided design of chimeric receptor binding proteins." Cell reports **29**(5): 1336-1350. e1334.

ECDC, S. (2020). Shigellosis, Annual Epidemiological report for 2017. Surveill Rep, 9

Evans, R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell and J. Yim (2022). "Protein complex prediction with AlphaFold-Multimer." BioRxiv: 2021.2010. 2004.463034.

Ferrières, L., G. Hémery, T. Nham, A.-M. Guérout, D. Mazel, C. Beloin and J.-M. Ghigo (2010). "Silent mischief: bacteriophage Mu insertions contaminate products of Escherichia coli random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery." Journal of bacteriology 192(24): 6418-6427.

Garcia-Doval, C. and M. J. v. Raaij (2013). "Bacteriophage receptor recognition and nucleic acid transfer." Structure and Physics of Viruses: 489-518.

Gordillo Altamirano, F., J. H. Forsyth, R. Patwa, X. Kostoulias, M. Trim, D. Subedi, S. K. Archer, F. C. Morris, C. Oliveira and L. Kielty (2021). "Bacteriophage-resistant Acinetobacter baumannii are resensitized to antimicrobials." Nature microbiology **6**(2): 157-161.

Hesse, S., M. Rajaure, E. Wall, J. Johnson, V. Bliskovsky, S. Gottesman and S. Adhya (2020). "Phage resistance in multidrug-resistant Klebsiella pneumoniae ST258 evolves via diverse mutations that culminate in impaired adsorption." MBio **11**(1): e02530-02519.

Huss, P., A. Meger, M. Leander, K. Nishikawa and S. Raman (2021). "Mapping the functional landscape of the receptor binding domain of T7 bacteriophage by deep mutational scanning." Elife **10**.

Islam, M. Z., A. Fokine, M. Mahalingam, Z. Zhang, C. Garcia-Doval, M. J. van Raaij, M. G. Rossmann and V. B. Rao (2019). "Molecular anatomy of the receptor binding module of a bacteriophage long tail fiber." PLoS Pathogens **15**(12): e1008193.

Jennison, A. V. and N. K. Verma (2004). "Shigella flexneri infection: pathogenesis and vaccine development." FEMS microbiology reviews **28**(1): 43-58.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek and A. Potapenko (2021). "Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.

Jun, J. W., J. H. Kim, S. P. Shin, J. E. Han, J. Y. Chai and S. C. Park (2013). "Characterization and complete genome sequence of the Shigella bacteriophage pSf-1." Research in microbiology **164**(10): 979-986.

Kiljunen, S., N. Datta, S. V. Dentovskaya, A. P. Anisimov, Y. A. Knirel, J. A. Bengoechea, O. Holst and M. Skurnik (2011). "Identification of the lipopolysaccharide core of Yersinia pestis and Yersinia pseudotuberculosis as the receptor for bacteriophage φA1122." Journal of bacteriology **193**(18): 4963-4972.

Kim, M., S. Kim, B. Park and S. Ryu (2014). "Core lipopolysaccharide-specific phage SSU5 as an auxiliary component of a phage cocktail for Salmonella biocontrol." Applied and environmental microbiology **80**(3): 1026-1034.

Kim, M. and S. Ryu (2012). "Spontaneous and transient defence against bacteriophage by phase-variable glucosylation of O-antigen in S almonella enterica serovar T yphimurium." Molecular microbiology **86**(2): 411-425.

Kotloff, K. L., J. P. Winickoff, B. Ivanoff, J. D. Clemens, D. L. Swerdlow, P. J. Sansonetti, G. Adak and M. Levine (1999). "Global burden of Shigella infections: implications for vaccine development and implementation of control strategies." Bulletin of the World Health Organization **77**(8): 651.

Lang, L. H. (2006). "FDA approves use of bacteriophages to be added to meat and poultry products." Gastroenterology **131**(5): 1370.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm and R. Lopez (2007). "Clustal W and Clustal X version 2.0." bioinformatics **23**(21): 2947-2948.

Larsen, R. A., M. M. Wilson, A. M. Guss and W. W. Metcalf (2002). "Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria." Archives of microbiology **178**(3): 193-201.

Lee, J.-H., J. Bai, H. Shin, Y. Kim, B. Park, S. Heu and S. Ryu (2016). "A novel bacteriophage targeting Cronobacter sakazakii is a potential biocontrol agent in foods." Applied and environmental microbiology **82**(1): 192-201.

Lefkowitz, E. J., D. M. Dempsey, R. C. Hendrickson, R. J. Orton, S. G. Siddell and D. B. Smith (2018). "Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV)." Nucleic acids research **46**(D1): D708-D717.

Llanos-Chea, A., R. J. Citorik, K. P. Nickerson, L. Ingano, G. Serena, S. Senger, T. K. Lu, A. Fasano and C. S. Faherty (2019). "Bacteriophage therapy testing against Shigella flexneri in a novel human intestinal organoid-derived infection model." Journal of pediatric gastroenterology and nutrition **68**(4): 509.

Maffei, E., A. Shaidullina, M. Burkolter, Y. Heyer, F. Estermann, V. Druelle, P. Sauer, L. Willi, S. Michaelis and H. Hilbi (2021). "Systematic exploration of Escherichia coli phage–host interactions with the BASEL phage collection." PLoS biology **19**(11): e3001424.

Molinaro, A., A. Silipo, C. D. Castro, L. Sturiale, G. Nigro, D. Garozzo, M. L. Bernardini, R. Lanzetta and M. Parrilli (2008). "Full structural characterization of

Shigella flexneri M90T serotype 5 wild-type R-LPS and its Δ galU mutant: glycine residue location in the inner core of the lipopolysaccharide." Glycobiology **18**(3): 260-269.

Nicholas, K. B. (1997). "GeneDoc: analysis and visualization of genetic variation." Embnew. news 4: 14.

Nygren, B., K. Schilling, E. Blanton, B. Silk, D. Cole and E. Mintz (2013). "Foodborne outbreaks of shigellosis in the USA, 1998–2008." Epidemiology & Infection **141**(2): 233-241.

Parent, K. N., M. L. Erb, G. Cardone, K. Nguyen, E. B. Gilcrease, N. B. Porcek, J. Pogliano, T. S. Baker and S. R. Casjens (2014). "OmpA and OmpC are critical host factors for bacteriophage S f6 entry in S higella." Molecular microbiology **92**(1): 47-60.

Park, E.-A., Y.-T. Kim, J.-H. Cho, S. Ryu and J.-H. Lee (2017). "Characterization and genome analysis of novel bacteriophages infecting the opportunistic human pathogens Klebsiella oxytoca and K. pneumoniae." Archives of virology **162**(4): 1129-1139.

Platt, R., C. Drescher, S.-K. Park and G. J. Phillips (2000). "Genetic system for reversible integration of DNA constructs and lacZ gene fusions into the Escherichia coli chromosome." Plasmid 43(1): 12-23.

Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler and R. Lopez (2005). "InterProScan: protein domains identifier." Nucleic acids research **33**(suppl_2): W116-W120.

Ranallo, R. T., S. Barnoy, S. Thakkar, T. Urick and M. M. Venkatesan (2006). "Developing live Shigella vaccines using λ Red recombineering." FEMS Immunology & Medical Microbiology **47**(3): 462-469.

Shahin, K., L. Zhang, H. Bao, A. Hedayatkhah, A. Soleimani-Delfan, M. Komijani, T. He, M. Barazandeh, M. Mansoorianfar and M. Bouzari (2021). "An in-vitro study on a novel six-phage cocktail against multi-drug resistant-ESBL Shigella in aquatic environment." Letters in Applied Microbiology **72**(3): 231-237.

Shahin, K., L. Zhang, A. S. Delfan, M. Komijani, A. Hedayatkhah, H. Bao, M. Barazandeh, M. Mansoorianfar, M. Pang and T. He (2021). "Effective control of Shigella contamination in different foods using a novel six-phage cocktail." LWT **144**: 111137.

Shin, H., J.-H. Lee, H. Kim, Y. Choi, S. Heu and S. Ryu (2012). "Receptor diversity and host interaction of bacteriophages infecting Salmonella enterica serovar Typhimurium."

Son, B., M. Kong and S. Ryu (2018). "The auxiliary role of the amidase domain in cell wall binding and exolytic activity of staphylococcal phage endolysins." Viruses **10**(6): 284.

Soncini, F. C., E. G. Véscovi and E. A. Groisman (1995). "Transcriptional autoregulation of the Salmonella typhimurium phoPQ operon." Journal of bacteriology 177(15): 4364-4371.

Spinelli, S., D. Veesler, C. Bebeacua and C. Cambillau (2014). "Structures and host-adhesion mechanisms of lactococcal siphophages." Frontiers in microbiology **5**: 3.

Sun, Q., R. Lan, J. Wang, S. Xia, Y. Wang, Y. Wang, D. Jin, B. Yu, Y. A. Knirel and J. Xu (2013). "Identification and characterization of a novel Shigella flexneri serotype Yv in China." PloS one **8**(7): e70238.

Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood and A. Laydon (2022). "AlphaFold Protein Structure Database:

massively expanding the structural coverage of protein-sequence space with high-accuracy models." Nucleic acids research **50**(D1): D439-D444.

Wang, X., C. Zhang, F. Shi and X. Hu (2010). Purification and characterization of lipopolysaccharides. Endotoxins: Structure, Function and Recognition, Springer**:** 27-51.

Washizaki, A., T. Yonesaki and Y. Otsuka (2016). "Characterization of the interactions between Escherichia coli receptors, LPS and OmpC, and bacteriophage T4 long tail fibers." Microbiologyopen **5**(6): 1003-1015.

Williams, P. C. and J. A. Berkley (2018). "Guidelines for the treatment of dysentery (shigellosis): a systematic review of the evidence." Paediatrics and international child health **38**(sup1): S50-S65.

Yehl, K., S. Lemire, A. C. Yang, H. Ando, M. Mimee, M. D. T. Torres, C. de la Fuente-Nunez and T. K. Lu (2019). "Engineering phage host-range and suppressing bacterial resistance through phage tail fiber mutagenesis." Cell **179**(2): 459-469. e459.

Zhang, H., R. Wang and H. Bao (2013). "Phage inactivation of foodborne Shigella on ready-to-eat spiced chicken." Poultry Science **92**(1): 211-217.

국문 초록

시겔라 플렉스네리는 설사를 일으키는 세균성 이질의 중요한 원인균이다. 박테리오파지는 시겔라 제어를 위한 항생제 대안으로 사용될 수 있다. 그러나 박테리아는 자신의 세포 표면의 수용체를 변형하여 파지에 대한 내성을 나타낼 수 있으므로 파지와 숙주 세포의 수용체의 상호작용을 파악하는 것이 중요하다. 본 논문에서는 숙주 수용체와의 상호작용을 기반으로 플렉스네리균의 제어 효율을 증대를 시키고자 시겔라 플렉스네리를 감염시키는 파지들을 분리하였다. 다섯 개 파지의 숙주 수용체를 알아보기 위하여, 시겔라 플렉스네리 2a strain 2457T 야생형 균주에 람다-레드 재조합과 Tn*5* 트랜스포존을 이용한 다양한 돌연변이 라이브러리를 구축하였다. 그 결과, 세 가지 유형으로 분류할 수 있었다; 지질다당체(LPS)의 outer core, LPS의 O항원, 및 페리크롬-철 수용체인 FhuA. 그 중 파지 SFP21A는 LPS O항원이 수용체로 작용하지만 LPS core의 합성에 직접적으로 영향을 미치는 *galU* 유전자의 Tn*5*에 의한 돌연변이 균주를 감염하는 현상을 확인할 수 있었다. 생물정보학적 분석을 통해 SFP21A 파지의 유전정보 안에 감염 특이성을 결정하는 꼬리 섬유 단백질로 추정되는 단백질 Gp09를 찾아냈다. SFP21A의 Gp09는 세 개의 도메인으로 구성된 동종삼량체를 갖으며, 이 중 C-말단 도메인이 수용체 결합에 관여하는 것을 확인하였다. 또한 점 돌연변

이 기법을 통한 Gp09의 아미노산 돌연변이는 더 이상 Tn*5*에 의한 *galU* 돌연변이 균주에 결합을 하지 못함을 볼 수 있었다. 이는 결합 도메인의 외부 루프에 존재하는 발린이 초기 감염에 중요한 역할을 할 수 있음을 예상할 수 있었다. 결과적으로 이러한 감염 패턴을 통해 LPS outer core가 수용체로 작용하는 SFPB에 대한 저항성 균들을 사용한 닷팅 분석시험에서 SFP21A는 다른 O항원 수용체 유형 파지와 다르게 대부분 감염을 할 수 있는 것을 확인하였다. 따라서 SFP21A, SFPB, 및 FhuA를 수용체로 사용하는 파지 SFP17로 구성된 파지 칵테일은 각각의 파지를 단독으로 처리했을 때에 비해 효과적으로 24시간 동안 박테리아 성장을 저해하였다. 결과적으로 파지와 숙주 수용체와의 상호작용을 알아봄으로써 시겔라 균을 제어할 수 있는 더 효과적인 파지 칵테일을 제조 가능할 것이라고 생각된다.

주요어: 시겔라 플렉스네리, 박테리오파지, 수용체, 꼬리 섬유 단백질, 수용체 결합 단백질

학번: 2021-24995