Genetic characterization of bacteriophage evolution with multidrug-resistant *Escherichia coli*

Ricardo Enrique Abadie Saenz

Department of Medical Science

The Graduate School, Yonsei University

Genetic characterization of bacteriophage evolution with multidrug-resistant *Escherichia coli*

Directed by Professor Dongeun Yong

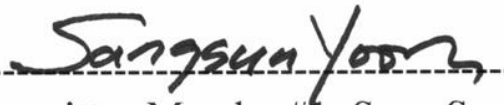
The Master's Thesis
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Master of Medical Science

Ricardo Enrique Abadie Saenz

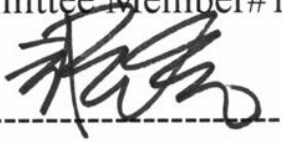
December 2023

This certifies that the Master's Thesis of Ricardo Enrique Abadie Saenz is approved.

Thesis Supervisor: Dongeun Yong



Thesis Committee Member#1: Sang Sun Yoon



Thesis Committee Member#2: Jun Yong Choi

The Graduate School

Yonsei University

December 2023

DEDICATORY

To my dear dad Guillermo, who supports me and watches over me from heaven. Your memory has been my guiding star, and I carry your love and wisdom with me in every step of this journey.

To my dear mother Balbina, your unwavering belief in me has been my strength. Your sacrifices and endless encouragement have made all the difference.

To my brothers Guimo and Lito, your companionship and shared laughter have lightened the load during the most challenging times.

This thesis is a tribute to our family's unity and the love that binds us even when distance separates us.

Ricardo Abadie

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my advisor, **Prof. Dongeun Yong**, for his unwavering guidance, support, and mentorship throughout the journey of completing this thesis. Your expertise, patience, and dedication have been instrumental in shaping my research and helping me navigate the challenges along the way. Your insightful feedback and constructive criticism have greatly enriched this work, and I am truly fortunate to have had you as my mentor.

I would like to thank my thesis committee members, **Prof. Sang Sun Yoon and Prof. Jun Yong Choi**, for their insightful comments and invaluable suggestions. Your feedback has been crucial in refining my research and ensuring its academic rigor.

I am also deeply thankful to my lab mates **Thao Nguyen Vu, Seongjun Yoo, and Hyunsook Lee**, whose camaraderie and collaboration made the research environment truly inspiring. Your diverse perspectives, willingness to share knowledge, and countless discussions during lab meetings were invaluable in shaping my ideas and improving the overall quality of this thesis.

I am grateful to the **researchers of Microbiotix ©., LTD** for their support and collaboration during the course of my research. Your expertise and contributions have played a significant role in shaping the direction and outcomes of this study.

I extend my appreciation to the **faculty members of the Medical Sciences Department** for their commitment to education and their role in fostering a stimulating academic atmosphere that encouraged intellectual growth and curiosity.

I am immensely grateful to the **National Institute for International Education (NIIED)** for their generous support through the **Korea Global Scholarship (GKS)**. This scholarship provided me with the opportunity to pursue my academic goals and played a pivotal role in making this research possible.

To my **friends and family**, thank you for your unwavering encouragement. Your support provided the emotional foundation that sustained me through the challenges of this journey.

This work would not have been possible without the collective support and collaboration of these individuals and organizations. I am deeply grateful for their contributions to my academic and personal growth.

Ricardo Enrique Abadie Saenz
Yonsei University College of Medicine
December 2023

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Genetic characterization of bacteriophage evolution with multidrug-resistant *Escherichia coli*

Ricardo Enrique Abadie Saenz

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Dongeun Yong)

Multidrug-resistant (MDR) bacteria are one of the major threats to global public health. MDR *Escherichia coli* (EC), especially carbapenem-resisters are one of the most urgent threats causing serious infections, necessitating novel treatment strategies. Bacteriophages (phages), viruses that can kill bacteria, are being explored as an alternative to combat MDR bacterial infections. However, phage resistance and limited host range are challenges. This study conducted a short-term lab experiment to evolve phages targeting multidrug-resistant *E. coli*. A cocktail of four phages was co-cultured with two sets of 11 *E. coli* strains each for 30 rounds, with the goal of broadening host range and understanding the underlying genetic mechanisms. Set-1 included phage-resistant mutant EC strains (ΦR-Mut), while set-2 included naturally phage-resistant EC strains (Nat-ΦR). After 30 rounds, two evolved phages (EC_7.1Φ and EC_9.1Φ) were isolated from set-1, showing recombination events and mutations affecting tail structures proteins, suggesting their potential contribution to the phages' re-adaptation to the ΦR-Mut strains. In set-2, three phages (EC_6.2Φ, EC_8.2Φ, and EC_11Φ) were obtained. EC_6.2Φ had insufficient titer for further analysis, while EC_8.2Φ and EC_11Φ were nearly genetically identical. EC_8.2Φ displayed genetic dissimilarity to the four original/parent phages but exhibited DNA homology with prophage regions of a Nat-ΦR strain (EC_5.2). These findings suggest that

phages evolve more rapidly to counteradapt against hosts that were previously sensitive to the phage, and prophages from bacterial genomes can "jump" to infect other bacterial strains. This research offers potential for developing phages capable of countering phage-resistant mutants and inducing prophages from bacterial genomes, potentially expanding their host range for more effective phage applications.

Key words: bacteriophage, MDR *E. coli*, phage-resistant mutants, phage evolution, prophage induction.

The author of this thesis is a Global Korea Scholarship scholar sponsored by the Korean Government

Genetic characterization of bacteriophage evolution with multidrug-resistant *Escherichia coli*

Ricardo Enrique Abadie Saenz

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Dongeun Yong)

I. INTRODUCTION

Escherichia coli are a large and diverse group of bacteria of the order Enterobacterales found in the environment, foods, and intestines of people and animals. Although most strains of *E. coli* are harmless, others are pathogenic causing gastrointestinal infections, urinary tract infections, respiratory illness, pneumonia, and among others[1], with community-associated as well as nosocomial infections [2, 3].

The emergence of multidrug-resistant (MDR) bacteria is a global health threat and concert with an urgent need for attention and a solution. Multidrug-resistant *E. coli* are globally distributed in healthcare facilities and are increasingly being found in the community [4, 5]. Especially carbapenem-resistant Enterobacterales (CRE) where are included *E. coli*, have the major threat to the global health by increasing health care expenses, hospitalization time, morbidity, and death[6].

Bacteriophages, or phages, are the most abundant microorganisms in the planet where are present in all areas where bacteria grow and play a significant role in the population dynamics and evolution of their host[7]. Phages are viruses that can infect and kill specifically bacteria by hijacking the host bacteria's metabolic mechanisms to replicate and to produce multiple progeny phages that leads with the host lysis[8, 9].

Nowadays phages are being extensively studied as an alternative for treatment of multi-drug resistant (MDR) bacterial infection. And it was demonstrated that phages are safe and effective in reducing a pathogenic intestinal bacteria burden *in vivo* where in many cases the use of antibiotics to slow down the spread of the disease is not recommended because of their side effects on the resident microbiota and the selection of antibiotic-resistant bacteria[10]. Phage infectivity is very strain host-specific, infecting only the target bacteria while not affecting other bacteria or cell lines of another organism[11].

However, bacteria, alike with antibiotics, can also become resistant to bacteriophage infection[12-15], which could be unfavorable for phage therapy[12]. Bacteria can become resistant to phage through different mechanisms, including surface modification by spontaneous mutation, CRISPR-CAS bacterial adaptive immune system and restriction modification systems[12]. Additionally, another concert about phage application is the narrow host range of some phages, generating a necessity to isolate new phages with a wider host spectrum[16]. There are many efforts in research to improve the efficacy of phage application, including phage evolution or phage training, cocktail of phages targeting different receptors, genetically engineered phages, and combination of phages with antibiotics[13-15, 17-20].

The objectives of this study are 1) to generate novel and more effective evolved phages *in vitro* with an expanded host range, and adapting to phage-resistant *E. coli* mutants. This aims to address the limitations of phage narrow host ranges and phage resistance; and 2) to evaluate the genetic mechanism of this evolutionary process on the phages.

II. RESEARCH METHOD

1. Bacterial strains and original phages

A. *Escherichia coli* strains

Clinical *E. coli* strains were provided by our laboratory. The criteria for multidrug resistance were those *E. coli* strains that are resistant to one or more antibiotics in three or more antimicrobial classes[21]. The selection and initial typification of *E. coli* strains were according to their antibiotic and phage susceptibility profile. The antibiotic susceptibility test (AST) of the *E. coli* strains was determined by disk diffusion. All interpretations were according to the Clinical and Laboratory Standards Institute (CLSI) guideline[22]. In total 17 clinical *E. coli* strains with different antibiotic/phage susceptibility profiles were selected and distributed in 2 sets with 11 *E. coli* strains each to use them for the phage evolution experiment.

B. Original phages

The 4 parent bacteriophages targeting *E. coli* to form the cocktail were selected from our laboratory phage bank based on their different host range and 2 of them with high genetic homology. These 4 phages are the phage EC_1Φ (3697.1Φ), EC_2Φ (1005Φ), EC_3Φ (3415Φ), and the phage EC_4Φ (3697.2Φ). The phages' WGSs were previously performed.

C. Phage-resistant *E. coli* mutants

A group of *E. coli* strains that initially were sensitive to some of the four phages of the cocktail were selected. To make phage-resistant *E. coli* mutants, the phage with each selected *E. coli* strains were prolongly cultured (≥ 24 hours) in Luria-Bertani (LB) broth, after that, an inoculum from the co-culture were taken and streaked on MH agar plates and incubated overnight. Next day, up to 5

potential phage-resistant colonies per plate were selected and contrasted with the phage by spotting the phage onto lawns of each individual colony and incubated at 37°C overnight. Colonies that did not show lysis by the phage on the spot area were considered as phage-resistant mutants.

2. Bacteriophage training/evolution

A. *E. coli* strain sets conformation for the phage training

For the phage evolution/training experiment, we utilized 17 clinical *E. coli* strains selected and distributed across two sets, with 11 *E. coli* strains each. **(Table 1).**

Set-1 consisted of three strains that were susceptible to the phage cocktail, serving as phage propagation hosts. Additionally, it included six phage-resistant *E. coli* mutants against the phages in the cocktail, and two naturally phage-resistant *E. coli* strains to the cocktail's phages.

Set-2 was composed of the same three strains susceptible to the phage cocktail for phage propagation. It also included eight naturally resistant strains to the phages in the cocktail, with two strains overlapping with those in Set-1.

Table 1. *Escherichia coli* strain sets conformation for the phage evolution experiment

Set	*E. coli* strain sets conformation		
	phages' hosts	phage-refractory strains	
1	(3): EC_1 EC_2 EC_3	**Φ-R.Mut (6):** EC_4.1, EC_5.1, EC_6.1 EC_7.1, EC_8.1, EC_9.1	**Nat.Φ-R (2):** EC_10 EC_11
2		**Nat.Φ-R (6)** EC_4.2, EC_5.2, EC_6.2 EC_7.2, EC_8.2, EC_9.2	

Φ-R.Mut phage-resistant mutant strains
Nat.Φ-R naturally phage-resistant strains (no mutants)

B. Appelmans protocol for the phage evolution experiment

For the phage evolution/training experiment, a protocol called Appelmans protocol, was performed as described by Burrowes *et al.*, 2019 [17] with some adaptations. Phage titer of each phage were determined to make the initial input cocktail by combining the four phages (~10^{10} PFU/mL each). Then 100 µL of the phage cocktail and its serial dilutions (until 10^{-3} dilution) were added to the 96-well microplate for each strain of the bacteria sets and right after 100 µL of the bacterial strain suspension (10 µL of overnight culture in 1mL of LB broth) were added. **(Fig. 1).**

Later the 96-well microplate were incubated on a shaker incubator at 37°C and 150-200 rpm overnight (16 – 18 hr.). After incubation, the microplate was inspected for wells showing lysis. All wells showing complete lysis and their next dilution were pooled. If not lysis wells were seen, the undulated well was pooled. Pooled lysates were treated with chloroform 1% and centrifuged at 15000xg for 15 min and filtrated through a 0.22 µm filter. This filtration was the round 1 cocktail and the method was repeated until round 30. Host range was

performed for the cocktails of the 10th, 20th, and 30th rounds to determine the presence of evolved phages (phages that acquired new ability to infect new hosts).

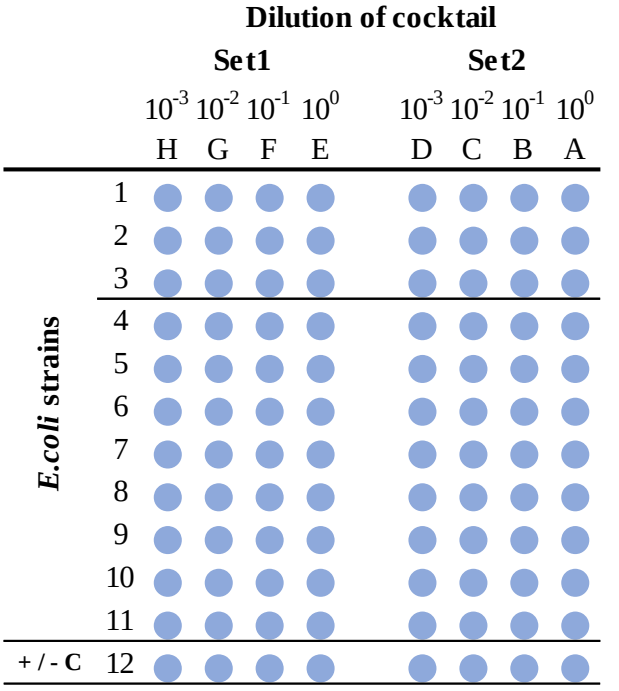


Figure 1. Schematic representation of the 96-well-plate for the phage evolution experiment. +/-C = positive and negative controls (only bacteria and only media, respectively).

C. Evolved phage isolation and propagation

From the round 30 pooled lysis cocktail, individual evolved phages that acquired new ability to infect new hosts (phage-resistant mutant strains or naturally phage-resistant strains) were 3 times purified on their respective new host by double-layer overlay technique[23]. Briefly 100 µl of the round 30 cocktail (diluted) was mixed with 100 µl of an overnight culture of the new host in a tube with 4 ml of molten soft agar at 55ºC, and then pouring on MH plate and incubated overnight at 37ºC. Next day, single phage plaques on a double-layer overlay were harvested by removing the plaque with an inoculation loop/needle and transferring it to a 1 ml of Sodium-Magnesium (SM) buffer. Then the double-layer overlay method was repeated 3 times and finally a single plaque was isolated

For the phage propagation and store we used the polyethylene glycol (PEG) - sodium chloride (NaCl) precipitation method [23, 24] with some adaptations. Briefly, first we made a stock solution of PEG/NaCl (3x) which consisted in PEG 6000 to 8000 (30% w/v) and NaCl (3 M) in ddH_20 and autoclaved. Then the purified phage was mixed with a fresh culture of its respective host in 30 ml of LB broth at multiplicity of infection (MOI) of ~0.1 and incubated overnight-shaking. Next day, phage-bacteria culture was centrifuged at 10,000 x g for 10 minutes at 4ºC, and the phage-rich supernatant was filtered through a 0.22 µm filter. 15 ml of the stock solution PEG/NaCl (3X) was mixed with the 30 ml filtered supernatant making a final concentration of PEG 10% and NaCl 1 M and incubated overnight at 4ºC. The following day, the PEG/NaCl-phage complex was obtained after centrifugation at 15,000 x g for 1 hour at 4ºC. Then, the supernatant was discarded and the pellet was resuspended in 1.5 ml of SM buffer and then filtered with a 0.22 µm filter and stored at 4ºC for further studies.

3. Original and evolved phage characterization

Host range[23], adsorption, one-step growth, and lysis assays were conducted according previously described methods[23, 25] with some modification.

A. Host range

First host range screening of the original and evolved phages was performed against their respective set of *E. coli* strain with a high phage titer (≥ 10^9 PFU/ml) by spot test. For this the phage solution was then spotted (10 µl) onto lawns of each individual bacteria (double-layer overlay agar) and incubated at 37°C overnight and next day the plates were examined for the presence of lysis or plaques on the phage spot. Second host range screening of the original and evolved phages was performed against a panel of 40 *E. coli* strains, including the 17 strains from the two sets, and in here the high phage titer (~ 10^9 PFU/ml) and its serial dilutions were spotted on the bacteria lawn to observe the presence of plaques that could confirm real infection.

B. Adsorption rate assay

The adsorption rate assay determined the ratio of phage attachment in the bacterial cell to observe the impact of the phage on bacteria[23]. The procedure consisted in adding the phage to the host bacterial suspension (~10^6 CFU/ml) at a MOI of 0.001. Subsequently, samples were subjected to filtration through a 0.22 µm syringe filter at different time points (0, 1, 2, 3, 4, 5, 10, 20 and 30 minutes) to separate unabsorbed phage particles. The titer of the phage in the filtered samples was then determined using the double-layered agar method.

C. One-step growth assay

One-step growth assay allowed us to determine the bacteriophage yield during an infection cycle, and this experiment let us know, as well, the latent

period and burst size with the growth of the phage[23, 26]. Fir this test, 30 ml of host bacterial suspension (~10^7 CFU/ml) was centrifuged at 10,000 × g for 15 minutes at 4℃. After removing the supernatant, the bacteria-rich pellet was reconstituted with 10 ml of LB broth and the phage at MOI 0.001, and then the sample was incubated at room temperature for 10 to 30 minutes (depending of the phage adsorption rate time previously determined) to allow for phage adsorption. Following this, the sample was subjected to centrifugation at 12,000 × g for 10 min at 4°C, and the supernatant was discarded to isolate the phage that had attached to the bacteria located in the pellet. This pellet was then resuspended in 10 ml of fresh LB broth and subjected to incubation at 37°C shaking. Phage samples were collected at 5-minute intervals for 100 minutes, and their titers were determined using the double-layered agar method.

D. Lysis test

Lysis assay determines how effective the bacteriophage is at killing the bacteria at different multiplicity of infection (MOI) over a period of time. For this, the bacterial solution (~10^7 CFU/ml) was mixed with the phage at MOIs of 10, 1, and 0.1 in a 96-well plate. The plate was then incubated afor24 hours at 37°C and optical density (OD600) was measured at hourly intervals using the VersaMax™ ELISA Microplate Reader and data collection using the Softmax® Pro version 5.4.1 (Molecular Devices, San Jose, CA, USA).

4. Phage and bacteria genome analysis

A. DNA extraction and whole genome sequencing (WGS)

Phage DNA was isolated for WGS using the phage DNA Isolation Kit (Norgen Biotek Corp., Canada) according to manufacturer's instructions. Bacterial genomic DNA of all *E. coli* strains from both sets were isolated using FastDNA Spin Kit for Soil (MP Biomedicals, Irvine, CA, USA) following the manufacturer's instructions. The WGS of the phages was conducted by Macrogen, Inc (Seoul, Korea) utilizing the Illumina platform, and the DeNovo assembly was performed by various k-mer using SPAdes (Platanus-alle for EC_7.1Φ). The WGS of the bacteria strains was executed by the Yonsei Laboratory Medicine Department, and the DeNovo assembly was performed using Geneious (Geneious Prime 2023.2.1;).

B. Multi-locus sequence typing (MLST) and the phylogenetic tree construction

Multi-locus sequence typing (MLST) and the phylogenetic tree construction were performed to determine the genetic diversity of both sets of *E. coli* strains used in this study. For this we used the whole genome sequencing (WGS) information of the *E. coli* strains. The MLST and the phylogenetic tree (by CSIPhylogeny tool) were determined and constructed, respectively, using the free online bioinformatics services provided by Center for Genomic Epidemiology, DTU, Denmark[27, 28] at https://www.genomicepidemiology.org. And the visualization and design of the phylogenetic tree was made using Geneious Prime® 2023.2.1 program.

C. Resistance genes and point mutations in resistance-determining regions detection

The resistance genes and point mutations in resistance-determining regions from the WGS data of the 17 *E. coli* strains were detected using the tool ResFinder 4.1 on the free online bioinformatics services provided by Center for Genomic Epidemiology, DTU, Denmark[29-31]. (https://www.genomicepidemiology.org).

D. Prophage regions identification

Prophage regions in the bacterial genomes were identified using PHAge Search Tool Enhanced Release (PHASTER)[32] (https://phaster.ca/).

E. Phage genomes alignment and annotation

The phage genome sequences were annotated using Rapid Annotations Subsystems Technology[33] (RAST; http://rast.nmpdr.org/) and BLASTP (NCBI; https://www.ncbi.nlm.nih.gov/) and genome alignment was performed using the NCBI Blastn suite, and the visualization/display was done using Progressive MAUVE and Easyfig version 2.2.5.

F. Detection of the genetic mechanisms of the evolved phages

For determining the genetic mechanisms of the evolved phages: origin, recombination events and mutations were visualized through the alignment between the annotated genomes of the original phages (or EC_5.2 prophage regions for EC.8.2 Φ) vs the evolved phages using MAFFT[34] through Geneious Prime® 2023.2.1 program, and BLASTP (NCBI; https://www.ncbi.nlm.nih.gov/) was performed to detect the substitution of amino acids in mutated ORFs

E. Prophage induction experiment

We executed experiment to induce prophages in the genome of the 17 *E. coli* strains using mitomycin C (mitC), which is well known for prophage induction**.** This experiment was conducted as described before with some modifications[35]. Briefly, 40µl of overnight culture of each bacterium were added to 2 tubes with 4 ml of LB broth each and incubated at 37°C shaking for 1 h (OD_{600} = 0.15 ~ 0.2). At that moment, one tube was left as a control were no inducing agent was added (to observe spontaneous induction of prophages), and in the second tube was added mitomycin C to a final concentration of 1µg/ml. Then both tubes were incubated for 4 h more at 37°C shaking to be after centrifuged at 7000 x g for 10 min at 4°C and the supernatant filtrated through a 0.22 µm syringe filter. Finally, the filtered supernatant was analyzed to observer the presence of induced prophages by spot test.

III. RESULTS

1. *Escherichia coli* strains and original phage

A. *Escherichia coli* strains

(1) Antibiotic susceptibility profile

Among the 17 isolates of *E. coli* strains, 88.2% (15/17) were resistant to ampicillin, 76.5% (13/17) to ciprofloxacin, 70.6% (12/17) to cefotaxime, 64.7% to amoxicillin-clavulanate, piperacillin-tazobactam, cefepime, ceftazidime and aztreonam each, 47.1 (8/17) to trimethoprim-sulfamethoxazole and ertapenem each, 41.2% (7/17) to cefoxitin and meropenem each, 35.3% (6/17) to tetracycline, 29.4% (5/17), and 11.8% (2/17) to amikacin. Among the 17 EC strains 82.4% (14/17) were considered as multi-drug resistant (MDR) strains (non-susceptible to 1 drug in 3 antimicrobial classes) and 47.1% (8/17) were carbapenem-resistant *E. coli* strains (CREC). It should be noted that EC strains EC_7.1 (1005_φR) and EC_8.1 (3415_ φR) are the phage-resistant mutants of the strains EC_2 (1005) and EC_3 (3415), respectively. Table 2.

(2) Multi-locus sequence type (MLST) and phylogenetic tree construction

Among all the 17 *E. coli* strains distributed in the 2 sets, the strain EC_5.1 belongs to the sequence type (ST) 95, the strain EC_4.1 to ST998 and the EC_9.2 to ST73, the strains EC_5.2, EC_7.2, EC_8.2 and EC_11 all belong to the ST131, the strains EC_1 and EC_6.1 belong to the ST1193, the strain EC_10 to ST48, the strains EC_2, EC_7.1 and EC_9.1 all belong to the ST224, the strain EC_4.2 to ST405, and the strains EC_3, EC_8.1 and EC_6.2 belong to the ST38. (Fig. 2).

Table 2. Antibiotic susceptibility profiles of the sets of *Escherichia coli* strains

Bacteria code	AMINOGLYCOSIDES		PENNICILINS		β-LACTAM COMBINATION AGENTS				CEPHALOSPORINS						MONOBACTAMS		FLUOROQUINOLONES		FOLATE PATHWAY		CARBAPENEMS						TETRACYCLINES		Drug-resistance definition and MDR %
	Amikacin		Ampicilin		Amoxicillin-clavulanate		Piperacillin-tazobactam		Cefepime		Cefotaxime		Cefoxitin		Ceftazidime		Aztreonam		Ciprofloxacin		Trimethoprim-sulfamethoxazole		Ertapenem		Imipenem		Meropenem		Tetracycline		
	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	mm	Int.	
3697 **EC_1**	24	S	6	R	6	R	10	R	18	R	6	R	6	R	6	R	6	R	6	R	6	R	24	S	20	I	28	S	23	S	MDR
1005 **EC_2**	20	S	6	R	7	R	12	R	17	R	6	R	17	I	8	R	6	R	6	R	26	S	14	R	15	R	16	R	21	S	MDR-CR
3415 **EC_3**	21	S	6	R	6	R	11	R	10	R	6	R	10	R	6	R	6	R	6	R	6	R	10	R	15	R	15	R	18	S	MDR-CR
6-5-1_qR **EC_4.1**	18	S	6	R	21	S	27	S	34	S	33	S	27	S	29	S	32	S	22	I	6	R	33	S	28	S	32	S	6	R	MDR
13-24-2_qR **EC_5.1**	20	S	21	S	22	S	29	S	35	S	33	S	27	S	29	S	32	S	38	S	26	S	33	S	29	S	32	S	23	S	non-MDR
21-16-1_qR **EC_6.1**	18	S	6	R	20	S	25	S	32	S	30	S	26	S	27	S	29	S	6	R	27	S	31	S	26	S	28	S	22	S	non-MDR
1005_qR **EC_7.1**	20	S	6	R	7	R	13	R	16	R	6	R	16	I	13	R	6	R	6	R	24	S	14	R	15	R	16	R	20	S	MDR-CR
3415_qR **EC_8.1**	23	S	6	R	6	R	10	R	7	R	6	R	7	R	6	R	6	R	6	R	6	R	10	R	13	R	10	R	26	S	MDR-CR
3959_qR **EC_9.1**	20	S	6	R	7	R	12	R	17	R	16	R	17	I	15	R	7	R	6	R	6	R	15	R	15	R	16	R	6	R	MDR-CR
28-1-1 **EC_10**	20	S	6	R	19	S	26	S	33	S	32	S	25	S	28	S	32	S	10	R	17	S	32	S	29	S	30	S	7	R	MDR
4178 **EC_11**	21	S	6	R	6	R	17	R	18	R	7	R	20	S	20	I	19	I	6	R	6	R	14	R	21	I	21	I	6	R	MDR-CR
2015-2 **EC_4.2**	19	S	6	R	6	R	14	R	20	SDD	6	R	6	R	6	R	7	R	6	R	6	R	23	S	23	S	29	S	6	R	MDR
2016-6 **EC_5.2**	10	R	6	R	6	R	11	R	6	R	6	R	12	R	7	R	6	R	6	R	25	S	9	R	24	S	17	R	20	S	MDR-CR
2016-9 **EC_6.2**	20	S	6	R	14	I	19	I	10	R	6	R	19	S	12	R	6	R	23	I	24	S	28	S	28	S	30	S	22	S	MDR
2017-6 **EC_7.2**	9	R	6	R	6	R	13	R	6	R	6	R	14	R	6	R	6	R	6	R	24	S	8	R	21	I	15	R	6	R	MDR-CR
2017-10 **EC_8.2**	18	S	6	R	9	R	15	R	7	R	6	R	7	R	6	R	6	R	6	R	6	R	19	I	30	S	25	S	23	S	MDR
20-7-1 **EC_9.2**	19	S	21	S	23	S	30	S	36	S	35	S	27	S	30	S	34	S	35	S	25	S	34	S	30	S	32	S	21	S	non-MDR
% of Resistance (R)	11.8		88.2		64.7		64.7		64.7		70.6		41.2		64.7		64.7		76.5		47.1		47.1		29.4		41.2		35.3		82.4

mm = diameter of inhibition in mm; int. = interpretation; S = susceptible; I = intermediate; R = resistant; MDR = multidrug-resistant; CR = carbapenem-resistant.

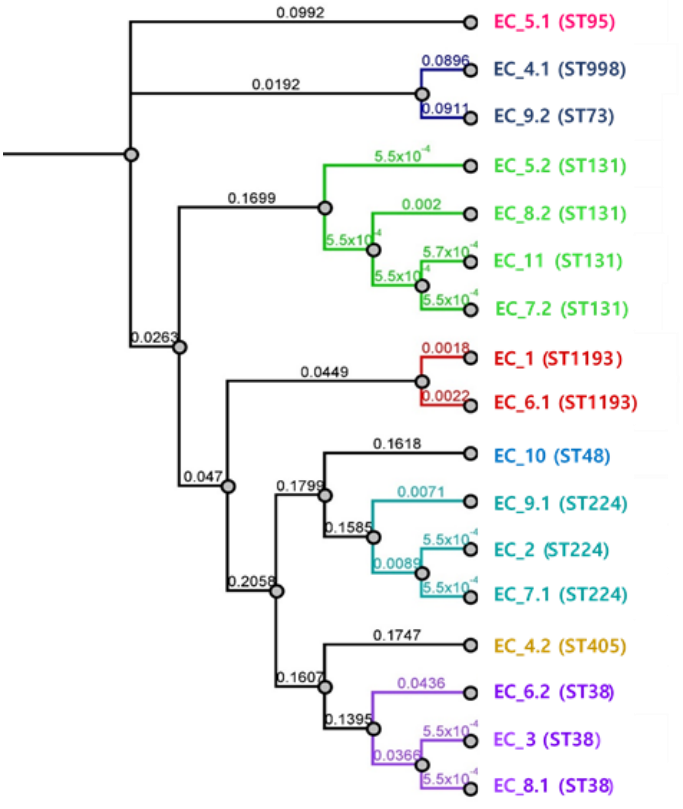


Figure 2. Phylogenetic tree and MLST of the 17 *E. coli* strains from the both sets. Same color means closer related strains.

(3) Antibiotic resistance genes and point mutations in resistance-determining regions

Among the WGS data for 17 *E. coli* strains, the following resistance genes and point mutations in resistance-determining regions were identified (either complete or partial), with the number of strains found indicated in square brackets. See details in Table 3.

Resistance genes: *For beta-lactam resistance:* blaTEM-1B [6], blaCTX-M-15 [5], blaOXA-1 [3], **blaKPC-2 [3]**, blaSHV-182 [3], **blaKPC-3 [2]**, blaOXA-9 [2], blaTEM-1A [2], blaCMY-2 [2], blaCTX-M-14 [2], **blaOXA-48 [1]**, blaCTX-M-27 [1] and blaCTX-M-55 [1]. Genes in **bold** encode the production of carbapenemases. *For aminoglycoside resistance*: aadA5 [5], aadA2 [3], aac(3)-IId [3], aph(3")-Ib [3], aph(6)-Id [3], aph(3')-Ia [2], aac(3)-Iia [2] and aadA1 [1]. *For fluoroquinolone and aminoglycoside resistance*: aac(6')-Ib-cr [3]. *For fluoroquinolone resistance:* qnrS1 [1]. *For macrolide resistance*: mph(A) [8] and erm(B) [2]. *For phenicol resistance*: catB3 [3], cmlA1 [1], floR [1]. *For folate pathway antagonist resistance:* sul1 [7] and sul2 [2] for sulfamethoxazole resistance, dfrA17 [5], dfrA12 [3] and dfrA15 [1] for trimethoprim resistance. *For tetracycline resistance:* tet(A) [5] and tet(M) [1]. *For disinfectant resistance:* sitABCD [13] for hydrogen peroxide resistance, qacE [7] and qacL [1] for quaternary ammonium compound resistance.

Point mutations in resistance-determining regions: *Point mutations in quinolone resistance-determining regions for nalidixic acid and ciprofloxacin resistance:* gyrA p.S83L [14], gyrA p.D87N [11], gyrA p.D87Y [2], parC p.S80I [13], parE p.S458A [6], parC p.E84V [4], parE p.I529L [4], parE p.L416F [2].

Table 3. Antibiotic resistance genes and point mutations in resistance-determining regions of the 17 *E. coli* strains from both sets

Bacteria Code	Resistance gene	Mutation	Identity %	Coverage %	Phenotype	Accession no.
EC_1 (3697)	aadA5		100	100	Aminoglycoside resistance	AF137361
	blaCMY-2		100	100	Beta-lactam resistance	X91840
	dfrA17		100	100	Trimethoprim resistance	FJ460238
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sitABCD		97.34	99.74	Disinfectant resistance	AY598030
	sul1		100	100	Sulphonamide resistance	U12338
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.L416F			Quinolone resistance	
EC_2 (1005)	blaCTX-M-15		100	100	Beta-lactam resistance Alternate name; UOE-1	AY044436
	blaKPC-2		100	100	Beta-lactam resistance	AY034847
	blaOXA-9		99.88	100	Beta-lactam resistance	KQ089875
	blaTEM-1A		99.88	100	Beta-lactam resistance Alternate name; RblaTEM-1	HM749966
	sitABCD		97.42	99.60	Disinfectant resistance	AY598030
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_3 (3415)	aac(3)-IIa		99.88	100	Aminoglycoside resistance	CP023555
	aadA2		100	100	Aminoglycoside resistance	JQ364967
	blaCTX-M-14		100	100	Beta-lactam resistance	AF252622
	blaKPC-3		100	100	Beta-lactam resistance	HM769262
	blaSHV-182		99.88	100	Beta-lactam resistance	KP050489
	dfrA12		100	100	Trimethoprim resistance	AM040708
	erm(B)		99.86	100	Macrolide resistance	JN899585
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sul1		100	100	Sulphonamide resistance	U12338
		gyrA p.D87Y			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_4.1 (6-5-1_φR)	aac(3)-IId		99.88	100	Aminoglycoside resistance	EU022314
	aadA5		100	100	Aminoglycoside resistance	AF137361
	aph(3")-Ib		100	100	Aminoglycoside resistance	AF321551
	aph(6)-Id		100	100	Aminoglycoside resistance	CP000971
	blaTEM-1B		100	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	dfrA17		100	100	Trimethoprim resistance	FJ460238
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sitABCD		97.34	99.74	Disinfectant resistance	AY598030
	sul1		100	100	Sulphonamide resistance	U12338
	sul2		100	100	Sulphonamide resistance	AY034138
	tet(A)		100	100	Tetracycline resistance	AJ517790
		gyrA p.S83L			Quinolone resistance	
EC_5.1 (13-14-2_φR)	mph(A)		100	100	Macrolide resistance	D16251
	sitABCD		99.1	99.74	Disinfectant resistance	AY598030
	sitABCD		97.75	99.74	Disinfectant resistance	AY598030
EC_6.1 (21-16-1_φR)	aac(3)-IId		99.88	100	Aminoglycoside resistance	EU022314
	blaTEM-1B		100	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	sitABCD		97.34	99.74	Disinfectant resistance	AY598030
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.L416F			Quinolone resistance	

Table 3. Antibiotic resistance genes and point mutations in resistance-determining regions of the 17 *E. coli* strains from both sets (cont.)

Bacteria Code	Resistance gene	Mutation	Identity %	Coverage %	Phenotype	Accession no.
EC_7.1 (1005_φR)	blaCTX-M-15		100	100	Beta-lactam resistance Alternate name; UOE-1	AY044436
	blaKPC-2		100	100	Beta-lactam resistance	AY034847
	blaOXA-9		99.88	100	Beta-lactam resistance	KQ089875
	blaTEM-1A		99.88	100	Beta-lactam resistance Alternate name; RblaTEM-1	HM749966
	sitABCD		97.42	99.60	Disinfectant resistance	AY598030
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_8.1 (3415_φR)	aac(3)-IIa		99.88	100	Aminoglycoside resistance	CP023555
	aadA2		100	100	Aminoglycoside resistance	JQ364967
	blaCTX-M-14		100	100	Beta-lactam resistance	AF252622
	blaKPC-3		100	100	Beta-lactam resistance	HM769262
	blaSHV-182		99.88	100	Beta-lactam resistance	KP050489
	dfrA12		100	100	Trimethoprim resistance	AM040708
	erm(B)		99.86	100	Macrolide resistance	JN899585
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sul1		100	100	Sulphonamide resistance	U12338
		gyrA p.D87Y			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_9.1 (3959_φR)	blaKPC-2		100	100	Beta-lactam resistance	AY034847
	blaSHV-182		99.88	100	Beta-lactam resistance	KP050489
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_10 (28-1-1)	aadA1		100	100	Aminoglycoside resistance	JQ414041
	aadA2		100	100	Aminoglycoside resistance	JQ364967
	blaTEM-1B		100	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	cmlA1		99.84	100	Phenicol resistance	M64556
	floR		98.11	99.92	Phenicol resistance	AF118107
	dfrA12		100	100	Trimethoprim resistance	AM040708
	qacL		92.22	94.00	Disinfectant resistance	NG_048048
	tet(A)		99.92	100	Tetracycline resistance	AF534183
	tet(M)		96.15	99.95	Tetracycline resistance	X04388
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
EC_11 (4178)	aadA5		100	100	Aminoglycoside resistance	AF137361
	aph(3')-Ia		100	100	Aminoglycoside resistance	V00359
	aph(3'')-Ib		100	100	Aminoglycoside resistance	AF321551
	aph(6)-Id		100	100	Aminoglycoside resistance	CP000971
	blaCTX-M-27		100	100	Beta-lactam resistance	AY156923
	blaOXA-48		100	100	Beta-lactam resistance	AY236073
	dfrA15		100	100	Trimethoprim resistance	AF221900
	dfrA17		100	100	Trimethoprim resistance	FJ460238
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sul1		100	100	Sulphonamide resistance	U12338
	sul2		100	100	Sulphonamide resistance	AY034138
	tet(A)		100	100	Tetracycline resistance	AJ517790
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.E84V			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.I529L			Quinolone resistance	

Table 3. Antibiotic resistance genes and point mutations in resistance-determining regions of the 17 *E. coli* strains from both sets (cont.)

Bacteria Code	Resistance gene	Mutation	Identity %	Coverage %	Phenotype	Accession no.
EC_4.2 (2015-2)	aac(3)-IId		99.88	100	Aminoglycoside resistance	EU022314
	aadA5		100	100	Aminoglycoside resistance	AF137361
	aph(3")-Ib		100	100	Aminoglycoside resistance	AF321551
	aph(6)-Id		100	100	Aminoglycoside resistance	CP000971
	blaCMY-2		100	100	Beta-lactam resistance	X91840
	blaTEM-1B		100	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	dfrA17		100	100	Trimethoprim resistance	FJ460238
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sitABCD		97.31	99.60	Disinfectant resistance	AY598030
	sul1		100	100	Sulphonamide resistance	U12338
	sul2		100	100	Sulphonamide resistance	AY034138
	tet(A)		100	100	Tetracycline resistance	AJ517790
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.S458A			Quinolone resistance	
EC_5.2 (2016-6)	aac(6')-Ib-cr		100	100	Fluoroquinolone and aminoglycoside resistance	DQ303918
	blaCTX-M-55		100	100	Beta-lactam resistance	DQ810789
	blaOXA-1		100	100	Beta-lactam resistance	HQ170510
	catB3		100	69.83	Phenicol resistance	AJ009818
	sitABCD		97.34	99.74	Disinfectant resistance	AY598030
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.E84V			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.I529L			Quinolone resistance	
EC_6.2 (2016-9)	aph(3')-Ia		100	100	Aminoglycoside resistance	EU722351
	blaCTX-M-15		100	100	Beta-lactam resistance Alternate name; UOE-1	AY044436
	qnrS1		100	100	Quinolone resistance	AB187515
EC_7.2 (2017-6)	aac(6')-Ib-cr		100	100	Fluoroquinolone and aminoglycoside resistance	DQ303918
	blaCTX-M-15		100	100	Beta-lactam resistance Alternate name; UOE-1	AY044436
	blaOXA-1		100	100	Beta-lactam resistance	HQ170510
	blaTEM-1B		100	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	catB3		100	69.83	Phenicol resistance	AJ009818
	sitABCD		97.34	99.74	Disinfectant resistance	AY598030
	tet(A)		100	100	Tetracycline resistance	AJ517790
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.E84V			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.I529L			Quinolone resistance	
EC_8.2 (2017-10)	aac(6')-Ib-cr		100	100	Fluoroquinolone and aminoglycoside resistance	DQ303918
	aadA5		100	100	Aminoglycoside resistance	AF137361
	blaCTX-M-15		100	100	Beta-lactam resistance Alternate name; UOE-1	AY044436
	blaOXA-1		100	100	Beta-lactam resistance	HQ170510
	blaTEM-1B		99.88	100	Beta-lactam resistance Alternate name; RblaTEM-1	AY458016
	catB3		100	69.83	Phenicol resistance	AJ009818
	dfrA17		100	100	Trimethoprim resistance	FJ460238
	mph(A)		100	100	Macrolide resistance	D16251
	qacE		100	84.68	Disinfectant resistance	X68232
	sitABCD		98.84	99.74	Disinfectant resistance	AY598030
	sul1		100	100	Sulphonamide resistance	U12338
		gyrA p.D87N			Quinolone resistance	
		gyrA p.S83L			Quinolone resistance	
		parC p.E84V			Quinolone resistance	
		parC p.S80I			Quinolone resistance	
		parE p.I529L			Quinolone resistance	
EC_9.2 (20-7-1)	sitABCD		97.22	99.74	Disinfectant resistance	AY598030

(4) Prophage regions in the *E. coli* sets' strains

In all the 17 *E. coli* strains from both sets were identified at least one intact prophage region (Table 4). Among the 17 EC strains, in the **EC_1** strain, nine prophage regions have been identified (of which 3 regions are intact, 3 regions are questionable, and 3 regions are incomplete). In **EC_2** strain, ten prophages regions were identified (4 intact, 1 questionable, and 5 incomplete). in **EC_3** strain, ten prophage regions of which 2 are intact, 1 questionable, and 7 incomplete.

In **EC_4.1** strain, nine prophages have been identified (3 intact, 2 questionable, and 4 incomplete). In **EC_5.1** strain, sixteen prophage regions were identified (6 intact, 4 questionable, 6 incomplete). The strain **EC_6.1** has thirteen prophage regions of which 6 are intact, 3 questionable, and 4 incomplete. In **EC_7.1** strain, eleven prophages have been identified (5 intact, 1 questionable, and 5 incomplete). In **EC_8.1** strain, eleven prophage regions were identified (2 intact, 2 questionable, 7 incomplete). The strain **EC_9.1** has seven prophage regions of which 1 is intact, 0 questionable, and 6 incomplete.

In EC_10 strain, eleven prophages have been identified (2 intact, 0 questionable, and 9 incomplete). **In EC_11** strain, seven prophage regions were identified (4 intact, 1 questionable, 2 incomplete).

The strain **EC_4.2** has seventeen prophage regions of which 3 are intact, 2 questionable, and 12 incomplete. In **EC_5.2** strain, also seventeen prophages have been identified (8 intact, 3 questionable, and 6 incomplete). In **EC_6.2** strain, eleven prophage regions were identified (1 intact, 1 questionable, 9 incomplete). The strain **EC_7.2** has fifteen prophage regions of which 5 are intact, 3 questionable, and 7 incomplete. **In EC_8.2** strain, also fifteen prophages have been identified (6 intact, 2 questionable, and 7 incomplete). In **EC_9.2** strain, five prophage regions were identified (2 intact, 1 questionable, 2 incomplete).

Table 4. Prophage regions in the genome of the 17 *E. coli* strains

Set	Bacteria code	Prophage regions				Most common prophage [intacts]
		Intact	Questionable	Incomplete	Total	
Set 1&2	EC_1 (3697)	**3**	3	3	**9**	PHAGE_Entero_BP_4795_NC_004813(20) PHAGE_Escher_TL_2011b_NC_019445(40) PHAGE_Escher_500465_1_NC_049342(12)
	EC_2 (1005)	**4**	1	5	**10**	PHAGE_Klebsi_4LV2017_NC_047818(30) PHAGE_Entero_lambda_NC_001416(23) PHAGE_Entero_mEp460_NC_019716(10) PHAGE_Salmon_SJ46_NC_031129(3)
	EC_3 (3415)	**2**	1	7	**10**	PHAGE_Entero_P4_NC_001609(10) PHAGE_Entero_lambda_NC_001416(17)
Set 1	EC_4.1 (6-5-1_φR)	**3**	2	4	**9**	PHAGE_Escher_SH2026Stx1_NC_049919(4) PHAGE_Klebsi_4LV2017_NC_047818(30) PHAGE_Entero_lambda_NC_001416(24)
	EC_5.1 (13-24-2_φR)	**6**	4	6	**16**	PHAGE_Vibrio_X29_NC_024369(16) PHAGE_Entero_lambda_NC_001416(17) PHAGE_Entero_HK630_NC_019723(6) PHAGE_Escher_phiV10_NC_007804(41) PHAGE_Salmon_SEN34_NC_028699(23) PHAGE_Shigel_SfII_NC_021857(24)
	EC_6.1 (21-16-1_φR)	**6**	3	4	**13**	PHAGE_Entero_fiAA91_ss_NC_022750(25) PHAGE_Entero_DE3_NC_042057(17) PHAGE_Entero_BP_4795_NC_004813(18) PHAGE_Escher_500465_1_NC_049342(12) PHAGE_Salmon_118970_sal3_NC_031940(4) PHAGE_Escher_TL_2011b_NC_019445(39)
	EC_7.1 (1005_φR)	**5**	1	5	**11**	PHAGE_Klebsi_4LV2017_NC_047818(30) PHAGE_Entero_mEp460_NC_019716(9) PHAGE_Entero_lambda_NC_001416(22) PHAGE_Salmon_118970_sal3_NC_031940(13) PHAGE_Escher_RCS47_NC_042128(3)
	EC_8.1 (3415_φR)	**2**	2	7	**11**	PHAGE_Entero_lambda_NC_001416(17) PHAGE_Entero_P4_NC_001609(10)
	EC_9.1 (3959_φR)	**1**	0	6	**7**	PHAGE_Escher_pro147_NC_028896(28)
Set 1&2	EC_10 (28-1-1)	**2**	0	9	**11**	PHAGE_Shigel_SfII_NC_021857(39) PHAGE_Klebsi_4LV2017_NC_047818(29)
	EC_11 (4178)	**4**	1	2	**7**	PHAGE_Burkho_phiE255_NC_009237(31) PHAGE_Entero_BP_4795_NC_004813(23) PHAGE_Pectob_ZF40_NC_019522(12) PHAGE_Stx2_c_1717_NC_011357(4)
Set 2	EC_4.2 (2015-2)	**3**	2	12	**17**	PHAGE_Entero_DE3_NC_042057(20) PHAGE_Escher_HK639_NC_016158(7) PHAGE_Entero_P2_NC_001895(33)
	EC_5.2 (2016-6)	**8**	3	6	**17**	PHAGE_Yersin_L_413C_NC_004745(15) PHAGE_Entero_P88_NC_026014(42) PHAGE_Entero_cdtI_NC_009514(6) PHAGE_Pectob_ZF40_NC_019522(12) PHAGE_Entero_BP_4795_NC_004813(22) PHAGE_Yersin_L_413C_NC_004745(23) PHAGE_Entero_mEp460_NC_019716(32) PHAGE_Stx2_c_Stx2a_F451_NC_049924(3)
	EC_6.2 (2016-9)	**1**	1	9	**11**	PHAGE_Stx2_c_Stx2a_F451_NC_049924(3)
	EC_7.2 (2017-6)	**5**	3	7	**15**	PHAGE_Entero_P88_NC_026014(31) PHAGE_Burkho_BcepMu_NC_005882(31) PHAGE_Entero_BP_4795_NC_004813(23) PHAGE_Pectob_ZF40_NC_019522(12) PHAGE_Entero_BP_4795_NC_004813(8)
	EC_8.2 (2017-10)	**6**	2	7	**15**	PHAGE_Entero_BP_4795_NC_004813(24) PHAGE_Entero_mEp460_NC_019716(11) PHAGE_Pectob_ZF40_NC_019522(12) PHAGE_Burkho_phiE255_NC_009237(22) PHAGE_Entero_P88_NC_026014(31) PHAGE_Escher_500465_1_NC_049342(12)
	EC_9.2 (20-7-1)	**2**	1	2	**5**	PHAGE_Entero_DE3_NC_042057(24) PHAGE_Entero_cdtI_NC_009514(6)

B. Original phage

(1) Genome similarity between the phages of the original cocktail

For this study, we employed four different *E. coli* phages provided by the phage bank of our laboratory. These 4 phages are the phage EC_1Φ (3697.1Φ), EC_2 Φ (1005Φ), EC_3Φ (3415Φ), and the phage EC_4Φ (3697.2Φ). The genome alignment of the 4 bacteriophages showed that the phages EC_1Φ and EC_2Φ had significant DNA homology (Coverage 95% and identity 98.22%). The phages EC_3Φ and EC_4Φ did not show significant genetic similarity between them and nor with the phages EC_1Φ and EC_2Φ (Figs. 3 and 4). The sequence length of the phages EC_1Φ, EC_2Φ, EC_3Φ and EC_4Φ are 151,549, 150,969, 40,584, and 39,358 base pairs, respectively.

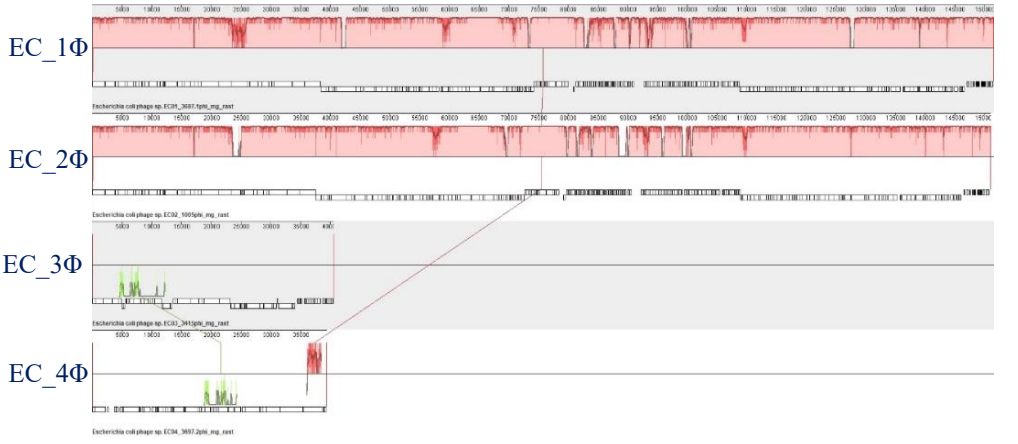


Figure 3. Multiple genome alignment between original phages using progressive MAUVE.

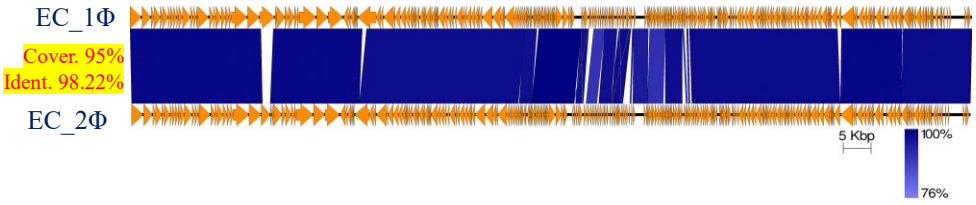


Figure 4. Pairwise genome alignment between the original phages EC_1Φ and EC_2Φ using Easyfig ver. 2.2.5.

(2) Host range of the original phages among the 17 E. coli strains from both sets

At the beginning of the experiment, the host range of the original 4 phages among the 2 sets of *E. coli* strains used for the phage evolution experiment is shown in Table 5. The phage **EC_1Φ** infects the 3 strains of the phage-cocktail's hosts (EC_1, EC_2, and EC_3). The phage **EC_2Φ** infects only the strains EC_2 and EC_3. The phage **EC_3Φ** infects only the strain EC_3. The phage **EC_4Φ** infects only the strain EC_1. And the 4-phage cocktail infects the strains EC_1, EC_2, and EC_3.

(3) phage-resistant *E. coli* mutant strains

As it was mentioned in methods, the Set-1 included phage-resistant mutants to the original phages. The strain **EC_4.1** was initially susceptible to the phage EC_1Φ, but after the prolonged co-culture with this phage, the strain become resistant to that phage. The strains **EC_5.1** and **EC_6.1** were originally susceptible to the phages EC_1Φ and EC_4Φ, but after the co-culture with the two phages, became resistant to both. The strains **EC_7.1** and **EC_9.1** were initially susceptible to the phages EC_1Φ and EC_2Φ, and after the prolonged co-culture with the phage EC_1φ, both strains become EC_1Φ -resistant and had

cross-resistance to the phage EC_2Φ. The strain **EC_8.1** was originally susceptible to the phages EC_1Φ, EC_2Φ, and EC_3Φ, and after the co-culture with the EC_1Φ, the strain become resistant to this phage and had cross-resistance with the phages EC_2Φ and EC_3Φ. (Tables 5 & 7).

Table 5. Initial original phages host range and *E. coli* bacterial sets list

	Bacterial SET-1↓		Phages↓ EC_1Φ (3697.1Φ)	EC_2Φ (1005Φ)	EC_3Φ (3415Φ)	EC_4Φ (3697.2Φ)	cocktail
Cocktail's hosts	3697	EC_1	+	-	-	+	+
	1005	EC_2	+	+	-	-	+
	3415	EC_3	+	+	+	-	+
Φ-R.Mut	6-5-1_ΦR	EC_4.1	-	-	-	-	-
	13-24-2_ΦR	EC_5.1	-	-	-	-	-
	21-16-1_ΦR	EC_6.1	-	-	-	-	-
	1005_ΦR	EC_7.1	-	-	-	-	-
	3415_ΦR	EC_8.1	-	-	-	-	-
	3959_ΦR	EC_9.1	-	-	-	-	-
Nat. Φ-R	28-1-1	EC_10	-	-	-	-	-
	4178	EC_11	-	-	-	-	-

	Bacterial SET-2↓		Phages↓ EC_1Φ (3697.1Φ)	EC_2Φ (1005Φ)	EC_3Φ (3415Φ)	EC_4Φ (3697.2Φ)	cocktail
Cocktail's hosts	3697	EC_1	+	-	-	+	+
	1005	EC_2	+	+	-	-	+
	3415	EC_3	+	+	+	-	+
Nat.Φ-R	2015-2	EC_4.2	-	-	-	-	-
	2016-6	EC_5.2	-	-	-	-	-
	2016-9	EC_6.2	-	-	-	-	-
	2017-6	EC_7.2	-	-	-	-	-
	2017-10	EC_8.2	-	-	-	-	-
	20-7-1	EC_9.2	-	-	-	-	-
	28-1-1	EC_10	-	-	-	-	-
	4178	EC_11	-	-	-	-	-

\+ visible lysis
\- no lysis

Φ-R.Mut phage-resistant mutant strains
Nat.Φ-R naturally phage-resistant strains (no mutants)

2. Bacteriophage evolution outcome

A. Evolved phages

From **Set-1-Round-30 cocktail**, two evolved phages (**EC_7.1Φ** and **EC_9.1Φ**) that acquired infectivity to phage-resistant mutant *E. coli* strains were isolated. **EC_7.1Φ phage** infected the 3 cocktail's hosts and the Φ-R.Mut EC_7.1 strain. **EC_9.1Φ phage** infected two of the cocktail's hosts and Φ-R.Mut EC_8.1 and EC_9.1 strains. Notice that no visible lysis on the EC_8.1 strain was observed from the Set-1-rounds' cocktails, but after isolation and propagation, the EC9.1Φ phage was able to infect the EC_8.1 strain. (Table 6A).

From **Set-2-Round-30 cocktail**, three phages (**EC_6.2Φ, EC_8.2Φ** and **EC_11Φ**) infecting naturally phage-resistant *E. coli* strains were isolated. **EC_6.2Φ phage** infected only the strain EC_6.2. The **phages EC_8.2Φ** and **EC_11Φ** each infected the strains EC_8.2 and EC11. (Table 6B). EC_6.2Φ phage did not produce enough titer for further analysis. EC_8.2Φ and EC_11Φ resulted to be 99.99% identical, with just one nucleotide difference. EC_8.2Φ was chosen for further analysis.

Table 6. Original/evolved phages and rounds' cocktails host range in the 2 sets of *E. coli* strains

A

Bacterial SET-1		Original Φs				Set-1 rounds' cocktail				Set-1 evolved Φs	
		EC_1Φ	EC_2Φ	EC_3Φ	EC_4Φ	R0	R10	R20	R30	EC_7.1Φ	EC_9.1Φ
Cocktail's hosts	3697 EC_1	+	-	-	+	+	+	+	+	+	-
	1005 EC_2	+	+	-	-	+	+	+	+	+	+
	3415 EC_3	+	+	+	-	+	+	+	+	+	+
Φ-R.Mut	6-5-1_φR EC_4.1	-	-	-	-	-	-	-	-	-	-
	13-24-2_φR EC_5.1	-	-	-	-	-	-	-	-	-	-
	21-16-1_φR EC_6.1	-	-	-	-	-	-	-	-	-	-
	1005_φR EC_7.1	-	-	-	-	-	+	+	+	+	-
	3415_φR EC_8.1	-	-	-	-	-	-	-	-	-	+
	3959_φR EC_9.1	-	-	-	-	-	-	+	+	-	+
Nat.Φ-R	28-1-1 EC_10	-	-	-	-	-	-	-	-	-	-
	4178 EC_11	-	-	-	-	-	-	-	-	-	-

B

Bacterial SET-2		Original Φs				Set-2 rounds' cocktail				Set-2 evolved Φs		
		EC_1Φ	EC_2Φ	EC_3Φ	EC_4Φ	R0	R10	R20	R30	EC_6.2Φ	EC_8.2Φ	EC_11Φ
Cocktail's hosts	3697 EC_1	+	-	-	+	+	+	+	+	-	-	-
	1005 EC_2	+	+	-	-	+	+	+	+	-	-	-
	3415 EC_3	+	+	+	-	+	+	+	+	-	-	-
Nat.Φ-R	2015-2 EC_4.2	-	-	-	-	-	-	-	-	-	-	-
	2016-6 EC_5.2	-	-	-	-	-	-	-	-	-	-	-
	2016-9 EC_6.2	-	-	-	-	-	+	+	+	+	-	-
	2017-6 EC_7.2	-	-	-	-	-	-	-	-	-	-	-
	2017-10 EC_8.2	-	-	-	-	-	+	+	+	-	+	+
	20-7-1 EC_9.2	-	-	-	-	-	-	-	-	-	-	-
	28-1-1 EC_10	-	-	-	-	-	-	-	-	-	-	-
	4178 EC_11	-	-	-	-	-	+	+	+	-	+	+

\+ visible lysis

\- no lysis

Φ-R.Mut phage-resistant mutant strains

Nat.Φ-R naturally phage-resistant strains (no mutants)

B. Original vs evolved phages host range

A Second host range screening of the original and evolved phages was performed against a panel of 40 *E. coli* strains, including the 17 strains from the two sets. (Table 7)

The evolved phage **EC_7.1Φ** from the Set-1 displayed a host range similar to the original phage EC_1, except for its newly acquired ability to infect the phage-resistant mutant EC_7.1 strain. Likewise, the evolved phage **EC_9.1Φ** from the set-1 demonstrated a host range similar to the original phage EC_2Φ, apart from their newfound capability to infect the phage-resistant mutants EC_8.1 strain and EC_9.1 strain.

The phage **EC_8.2Φ** derived from Set-2 exhibited a distinct spectrum of hosts (3 out 40), compared to the host range of any of the four original phages.

Table 7. Host range of the original and evolved phages against a panel of 40 *Escherichia coli* strains

Phage	Bacteria code																																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	EC_1 (3697)	EC_2 (1005)	EC_3 (3415)	EC_4.1 (6-5-1_ΦR)	EC_5.1 (13-24-2_ΦR)	EC_6.1 (21-16-1_ΦR)	EC_7.1 (1005_ΦR)	EC_8.1 (3415_ΦR)	EC_9.1 (3959_ΦR)	EC_10 (28-1-1)	EC_11 (4178)	EC_4.2 (2015-2)	EC_5.2 (2016-6)	EC_6.2 (2016-9)	EC_7.2 (2017-6)	EC_8.2 (2017-10)	EC_9.2 (2017-10)	6-5-1	13-24-2	21-16-1	3959	1272	1273	3982	4134	2018-77	2018-79	2018--8	2017-8	2017-4	2016-2	2016-4	2014-10	2014-1	2014-2	13-24-3	2018-1 (1)	2018-1 (527)	22-13-1	2018-60
OrP EC_1Φ	+	+	+	-	·-^{t}	-	-	-	-	-	-	-	-	·-^{t}	-	-	-	+	+	+	+	-	·-^{t}	+	+	+	+	+	-	-	-	+	·-^{t}	·-^{t}	·-^{t}	·-^{t}	-	-	·-^{t}	-
EvP EC_7.1φ	+	+	+	-	·-^{t}	-	+	-	-	-	-	-	-	·-^{t}	-	-	-	+	+	+	+	-	·-^{t}	+	+	+	+	+	-	-	-	+	·-^{t}	·-^{t}	·-^{t}	·-^{t}	-	-	·-^{t}	-
OrP EC_2Φ	-	+	+	-	-	-	-	-	-	-	-	-	-	·-^{t}	-	-	-	-	-	-	+	-	-	+	+	-	·-^{t}	+	-	-	-	+	·-^{t}	·-^{t}	·-^{t}	·-^{t}	-	-	·-^{t}	-
EvP EC_9.1Φ	-	+	+	-	-	-	-	+	+	-	-	-	-	·-^{t}	-	-	-	-	-	-	+	-	-	+	+	-	·-^{t}	+	-	-	-	+	·-^{t}	·-^{t}	·-^{t}	·-^{t}	-	-	·-^{t}	-
OrP EC_3Φ	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-
OrP EC_4Φ	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	·-^{c}	+	+	-	-	-	-	+	+	+	-	-	-	-	+	-	-	-	-	-	-	-	-
EvP EC_8.2Φ	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+

OrP Original phage
EvP Evolved phage

\+ lysis and plaque-forming units (PFUs) visualization in dilutions (infection)
·-^{t} turbid spot but no plaques in dilutions (no infection)
·-^{c} clear spot but no plaques in dilutions (no infection)
\- no lysis nor plaques (no infection)

C. Original vs evolved phages characterization

The characterization tests were made between the original phage (OrP) vs the evolved phage (EvP) that had similar host range. The OrP **EC_1Φ** vs the EvP **EC_7.1Φ** on the host EC_1. The OrΦ **EC_2Φ** vs the EvΦ **EC_9.1Φ** on the host EC_2. And the EvΦ **EC_8.2Φ** alone on its host EC_8.2.

(1) Adsorption rate

In the adsorption test between the OrP **EC_1Φ** vs the EvP **EC_7.1Φ** to the strain EC_1 demonstrated that at 5 min the OrΦ **EC_1Φ** vs EvΦ **EC_7.1Φ** had absorption rate of 79.2% vs 62.7%, respectively. 93.4% vs 80.2% at 10 min, 99.1% vs 95.5% at 20 min. And the absorption ration at 30 min was 99.7% vs 99.1%, respectively. (Fig. 5A)

In the adsorption test between the OrP **EC_2Φ** vs the EvP **EC_9.1Φ** to the strain EC_2 showed that the adsorption rate between the **EC_2Φ** vs **EC_9.1Φ** was 18.1% vs 66.5% at 20 min, and 29.2% vs 77.8% at 30 min, respectively. (Fig. 5B).

The adsorption rate of the evolved phage **EC_8.2Φ** to the strain EC_8.2 was 79.2% at 5 min, 93.4% at 10 min, 99.1% at 20 min, and 99.7% at 30 min. (Fig. 5C)

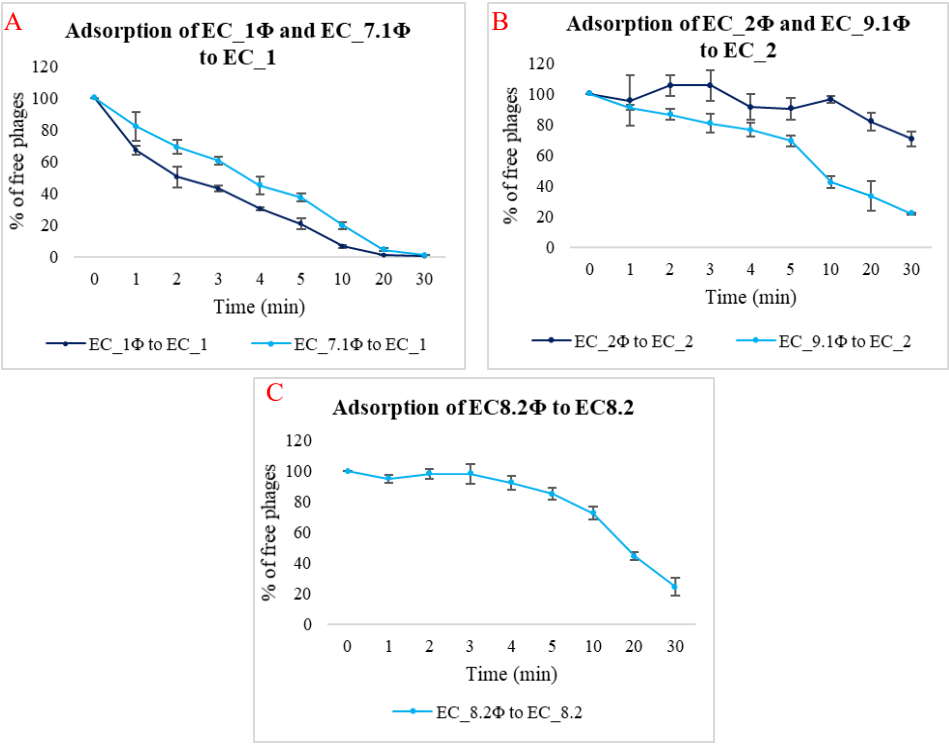


Figure 5. Adsorption rate between original phages and evolved phages. Mean standard ± deviation data are from triplicate experiments.

(2) One-step growth

In the one-step growth test, the original phage **EC_1Φ** and the evolved phage **EC_7. 1Φ** with the *E. coli* strain EC_1 had both a latent period of 35 min and a burst size of 54 and 71 virions per infected cell, respectively. (Fig. 6A).

The original phage **EC_2Φ** and the evolved phage **EC_9.1Φ** using the *E. coli* strain EC_2 had both a latent time of 25 min and a burst size of 20 and 41 virions per infected cell, respectively. (Fig. 6B).

And finally, the evolved phage **EC_8.2Φ** from the Set-2 using the *E. coli* strain had a latent period of 45 min and a burst size of 22 virions per infected cell. (Fig. 6C)

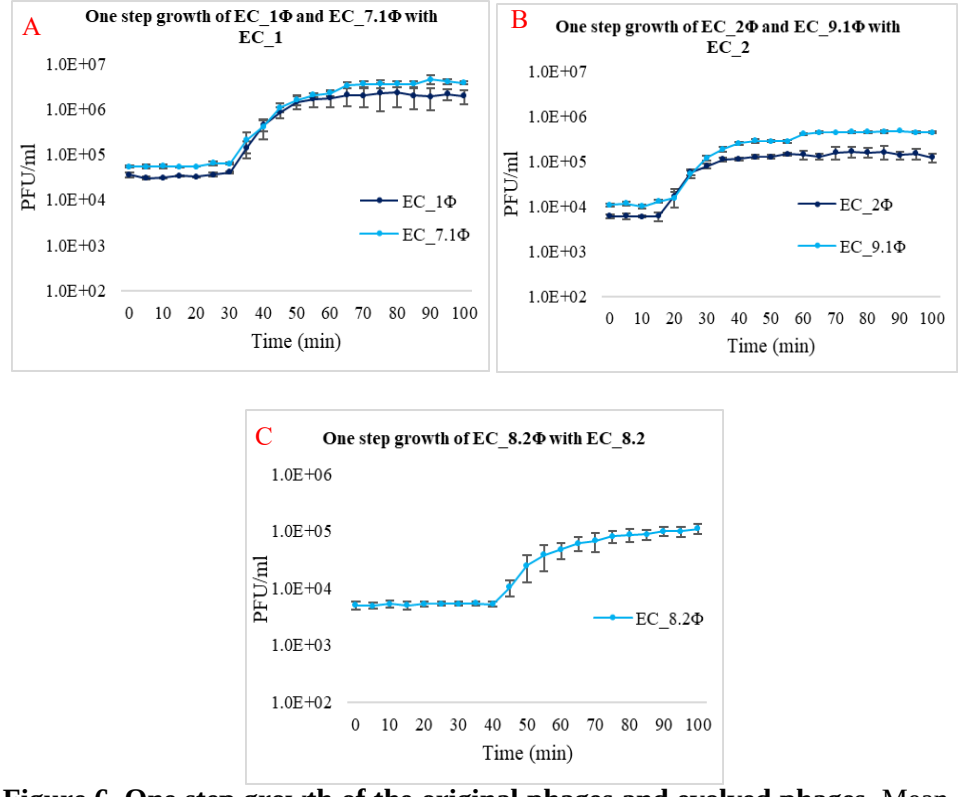


Figure 6. One step growth of the original phages and evolved phages. Mean standard ± deviation data are from triplicate experiments.

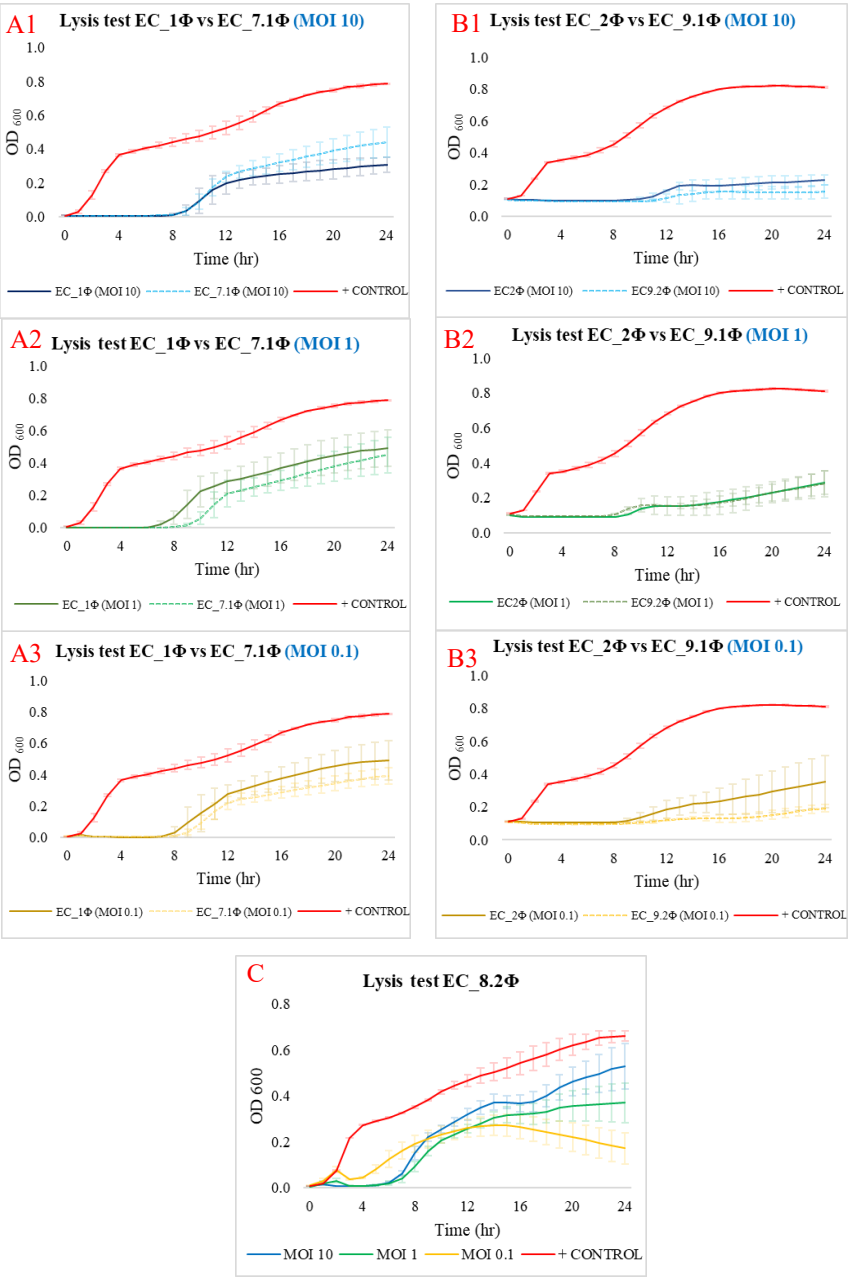


Figure 7. Lysis test of the original phages and evolved phages. EC_1Φ and EC_7.1Φ against the strain EC_1. EC_2Φ and EC_9.1Φ against the strain EC_2, and EC_8.1Φ and against the strain EC_8.2. At different MOIs 10, 1 and 0.1. Mean standard ± deviation data are from triplicate experiments.

(3) Lysis test

In the lysis test between the original phage (OrP) **EC_1Φ** and the evolved phage (EvP) **EC_7.1Φ** against *E. coli* strain EC_1, at a multiplicity of infection **(MOI) of 10**, both phages inhibited bacterial growth for the first 8 hours. By the end of the 24-hour test, **EC_1Φ** exhibited slightly stronger growth inhibition compared to **EC_7.1Φ** (Fig. 7A1). At **MOI 1**, **EC_1Φ** inhibited bacterial growth for the first 6 hours, whereas **EC_7.1Φ** suppressed growth for the initial 8 hours. Additionally, **EC_7.1Φ** exhibited a slightly stronger reduction in bacterial growth compared to **EC_1Φ** by the end of the 24-hour test (Fig. 7A2). At an **MOI of 0.1**, **EC_1Φ** inhibited bacterial growth within the first 7 hours, while **EC_7.1Φ** achieved this within the first 8 hours. Over the 24-hour duration of the experiment, the EvP slightly beat the OrP in reducing bacterial growth (Fig. 7A3).

In the lysis test conducted between the OrP **EC_2Φ** and the EvP **EC_9.2Φ** against *E. coli* strain EC_2 at an **MOI of 10**, **EC_2Φ** inhibited bacterial growth within the first 10 hours, whereas **EC_9.2Φ** achieved this within the initial 12 hours of the test. Additionally, the **EC_9.2Φ** reduced the bacterial growth lightly more compared with the **EC_2Φ** throughout the test (Fig. 7B1). At **MOI 1**, **EC_2Φ** and **EC_9.1Φ** inhibited bacterial growth within the initial 8 and 7 hours, respectively. From the 11th hour until the end of the 24-hour experiment, both phages reduced bacterial growth at a similar rate (Fig. 7B2). At **MOI 0.1**, **EC_2Φ** inhibited bacterial growth within the first 9 hours, whereas **EC_9.2Φ** achieved this within the initial 10 hours. Notably, throughout the test, the **EC_9.1Φ** reduces the bacterial growth significantly compared with the **EC_2Φ** (Fig. 7B3).

The lysis test of the evolved phage **EC_8.2Φ** against the strain EC_8.2 showed that at **MOIs of 10 and 1**, the phage inhibited the bacterial growth within the first 5 hours. However, at **MOI of 0.1** the bacteria grew in parallel with the

positive control for the first 3 hours and then the phage paradoxically reduced bacterial growth more compared to MOIs of 10 and 1 by the end of the experiment (Fig. 7C).

D. Genetic characterization of the evolved phages

(1) Evolved phage EC_7.1Φ (From Set-1)

The genome alignment between the evolved phage EC_7.1Φ with the four original phages exhibited that the evolved phage EC_7.1Φ had significant DNA homology with the original phage EC_1Φ (cover 98% and ident. 99.06%) as well as with the original phage EC_2Φ (cover 96%, ident. 99.84%) (Figs. 8 and 9), and did not show significant genetic similarity with the original phages EC_3Φ and EC_4Φ (Fig. 8). The sequence length of the evolved phage EC_7.1Φ is 152,915 base pairs.

The alignment between annotated genomes was performed to find the possible genetic mechanisms involved in the genome of the evolved phage EC_7.1Φ from their ancestors the original phages EC_1Φ and EC_2Φ. Genome analysis revealed recombination events and mutations occurring between and from the original phages EC_1Φ and EC_2Φ, resulting in the formation of the evolved phage EC_7.1Φ. Among the 280 ORFs, (269 CDSs and 11 tRNAs) found in the evolved phage EC_7.1Φ genome, 157 ORFs were identical to ORFs of EC_2Φ, 77 ORFs to EC_1Φ, 38 ORFs same to both EC_1Φ and EC_2Φ, 6 ORFs were a recombination of EC_1Φ and EC_2Φ (4 unique for EC_7.1Φ and 2 shared with EC_9.1Φ), one ORF from EC_1Φ with an insertion mutation, and one ORF from EC_2Φ with point mutation (shared with EC_9.1Φ) (Fig. 10A).

The **ORF #25** coding the tail fiber protein came from the phage EC_1Φ (Fig. 10B). Among the **four unique ORFs** resulting from recombination between

EC_1Φ and EC_2Φ phages, we have the **ORF #24** coding the non-contractile tail sheath, the **ORF #41** coding the DNA polymerase I, the **ORF #42** coding the putative DNA N6-adenine methyltransferase, the **ORF #48** coding glycosyltransferase (Figs. 10C-10F). And the two recombinant ORFs shared with EC_9.1Φ were the **ORFs #229** and **#231** both coding ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1) (Figs. 10G-10H). Then the ORF with an insertion mutation with the addition of an adenine (A) was the **ORF #113** that came from the EC_ 1Φ phage and codes a hypothetical protein (Fig. 10J). Also the missense mutation was found in the **ORF #33** that comes from the EC_ 1Φ phage and codes the tail fiber protein (Fig. 10I), changing cytosine (C) to thymine (T), leading the substitution of the amino acid proline (P) with serine (S), and this mutated ORF is shared with the another evolved phage, EC_9.1Φ, which is also described in the next point.

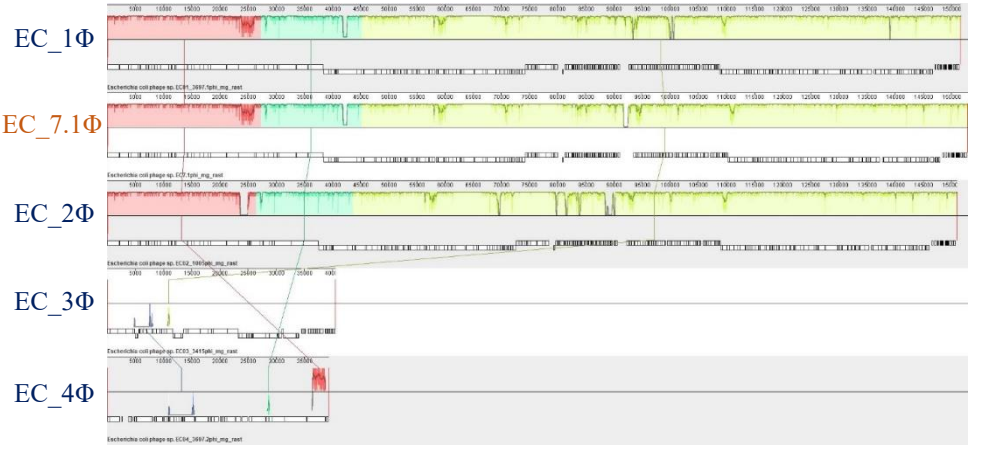


Figure 8. Multiple genome alignment between the evolved phage EC_7.1Φ with the four original phages using progressive MAUVE.

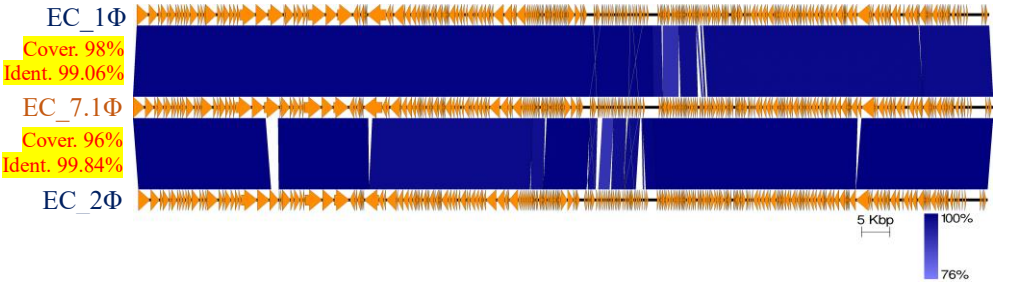


Figure 9. Multiple genome alignment between the evolved phage EC_7.1Φ and the original phages EC_1Φ and EC_2Φ using Easyfig ver. 2.2.5.

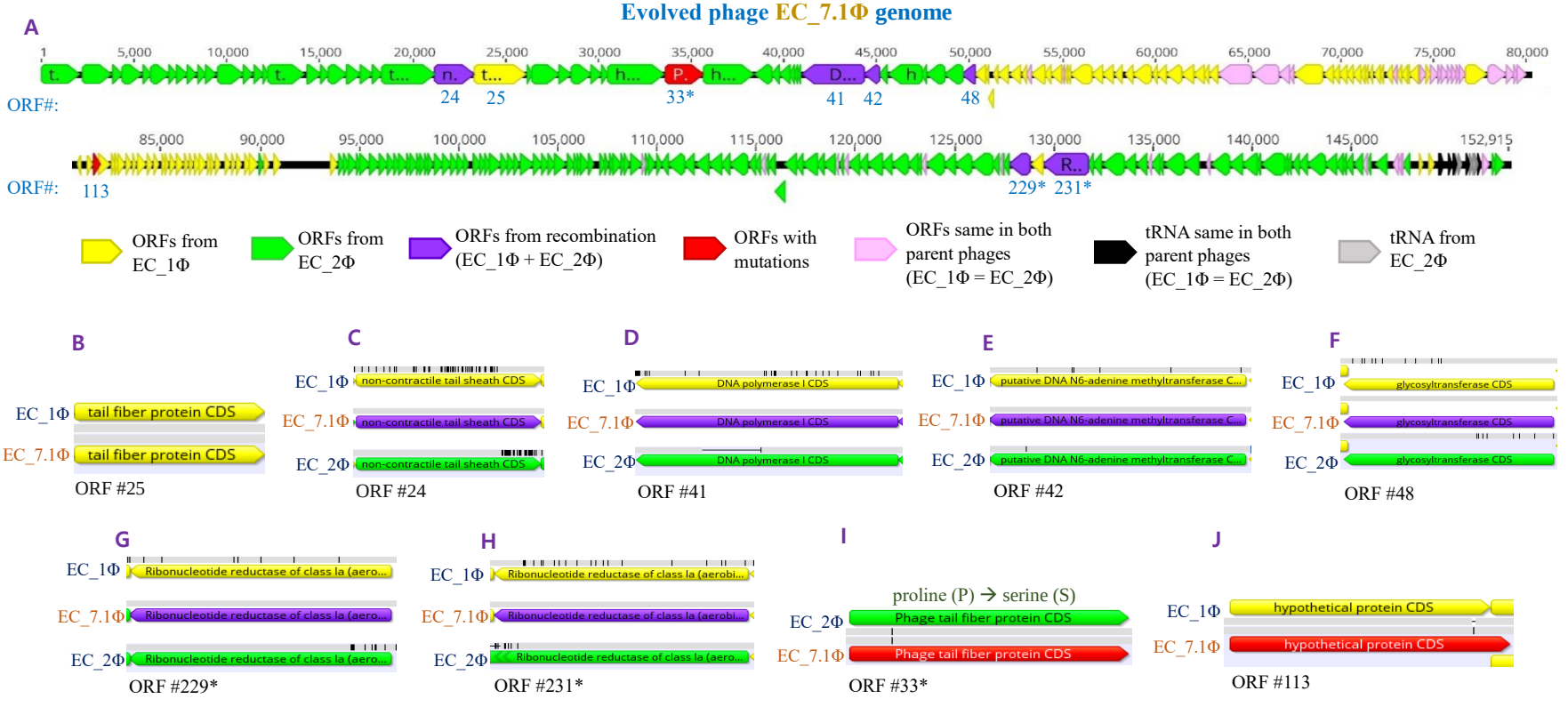


Figure 10. Recombination events and mutations in the evolved phage EC_7.1Φ genome. **A**: shows the origin of the different ORFs: CDSs (colored arrows) and tRNA (black and gray arrows) of the EC_7.1Φ, where yellow arrows are the ORFs identical with EC_1Φ, green arrows are ORFs identical to EC_2Φ, purple arrows are ORFs with recombination of EC_1Φ and EC_2Φ, red arrows are ORF with mutations, and pink arrows are ORFs that are same in both EC_1Φ and EC_2Φ. **B:** shows the match of the ORF with EC_1Φ that codes phage tail structure which could explain why the phage EC_7.1Φ kept similar host range with the EC_1Φ. **C-H:** show the recombinant EC_7.1Φ's ORFs from EC_1Φ and EC_2Φ's ORFs. **I**: shows the ORF that came from EC_2Φ with missense mutations (this ORF is shared with EC_9.1Φ). And **J** shows the ORF that comes from the EC_1Φ with a nucleotide insertion. *Same in EC_9.1Φ. In **B-J** grey and black parts indicate the agreement and disagreement of nucleotide matching, respectively.

(2) Evolved phage EC_9.1Φ (From Set-1)

Similarly, the genome alignment between the evolved phage EC_9.1Φ with the four original phages showed that the evolved phage EC_9.1Φ had significant DNA homology with the original phages EC_1Φ (cover 98% and ident. 99.47%) and EC_2Φ (cover 97%, ident. 99.86%) (Figs. 11 and 12), and did not exhibit significant genetic similarity with the original phages EC_3Φ and EC_4Φ (Fig. 11). The sequence length of the evolved phage EC_7.1Φ is 150,342 base pairs.

The alignment between annotated genomes showed the possible genetic mechanisms involved in the new genome of the evolved phage EC_9.1Φ from their ancestors the original phages EC_1Φ and EC_2Φ. The genome analysis of the evolved phage EC_9.1Φ, much like what was observed with EC_7.1Φ, revealed recombination events and mutations originating from both the original phages EC_1Φ and EC_2Φ. Among the 281 ORFs (270 CDSs and 11 tRNAs sequences) found in the evolved phage EC_9.1Φ genome, 133 ORFs were identical to CDSs of EC_2Φ, 95 ORFs identical to EC_1Φ, 38 ORFs were same to both EC_1Φ and EC_2Φ, nine ORFs were a recombination of ORFs of EC_1Φ and EC_2Φ (7 unique in EC_9.1Φ and 2 shared with EC_7.1Φ), five ORFs that came from one from EC_2Φ with missense mutations (one shared with EC_7.1Φ), and one ORF with a deletion mutation. (Fig. 13A).

Among the 7 unique ORFs in EC_9.1Φ resulted by recombination between the original phages EC_1Φ and EC_2Φ are: the **ORF #42** coding the DNA polymerase I, the **ORF #50** coding glycosyltransferase, the **ORF #58** that codes a hypothetical protein, the **ORF #64** which codes the tellurite resistance gene, and the **ORFs #153, #171, #171** all coding hypothetical protein. (Figs. 13B to 13H). And the two recombinant ORFs shared with EC_7.1Φ were the **ORFs**

#230 and **#232** both coding ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1) (Figs. 10I-10J).

Among the five ORFs in EC_9.1Φ with missense mutations that came from the EC_2Φ, they are: the **ORF #19** coding the baseplate hub that had a change in the nucleotide cytosine (C) for thymine (T) which led the substitution of the amino acid alanine (A) with valine (V) (Fig. 13K). The **ORF #24** coding the non-contractile tail sheath by changing the nucleotide guanine (G) for thymine (T) leading the substitution of the amino acid glycine (G) to tryptophan (W) (Fig. 13L). The **ORF #28** coding the baseplate wedge subunit with the change of the nucleotide guanine (G) for adenine (A) which led the substitution of the amino acid alanine (A) for threonine (T) (Fig. 13M). In the **ORF #30** coding for the tail fiber protein there was the substitution of 2 nucleotides, thymine (T) for cytosine (C), and guanine (G) for adenine (A) thus leading the substitution of 2 amino acids, isoleucine (I) for threonine (T) and serine (S) for asparagine (N) (Fig. 13N). The **ORF #34** (shared with EC_7.1Φ) codes the tail fiber protein, changing cytosine (C) to thymine (T), leading the substitution of the amino acid proline (P) with serine (S) (Fig 13O). There was a deletion mutation [of a guanine (G)] from the ORF in EC_2Φ leading a shorter ORF (**ORF #25**) in the evolved phage EC_9.1Φ, which codes for hypothetical protein. (Fig. 13P)

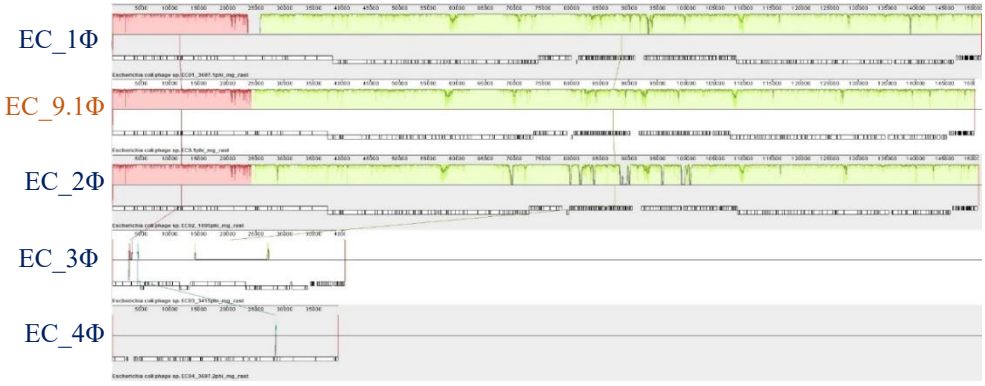


Figure 11. Multiple genome alignment between the evolved phage EC_9.1Φ with the four original phages using progressive MAUVE.

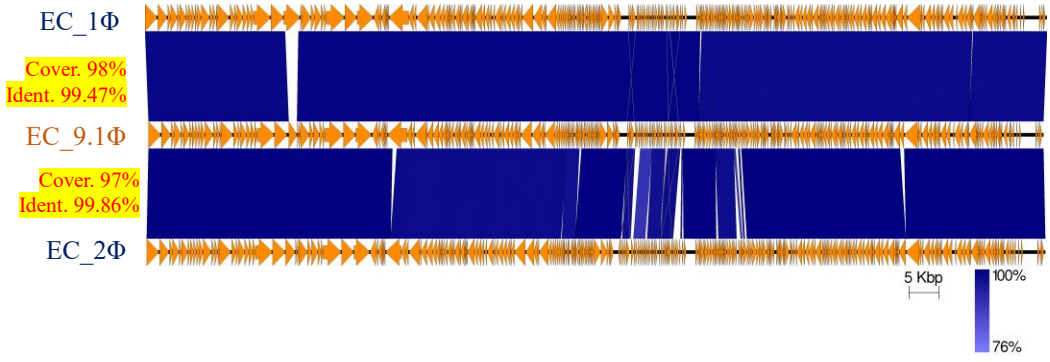


Figure 12. Multiple genome alignment between the evolved phage EC_9.1Φ and the original phages EC_1Φ and EC_2Φ using Easyfig ver. 2.2.5.

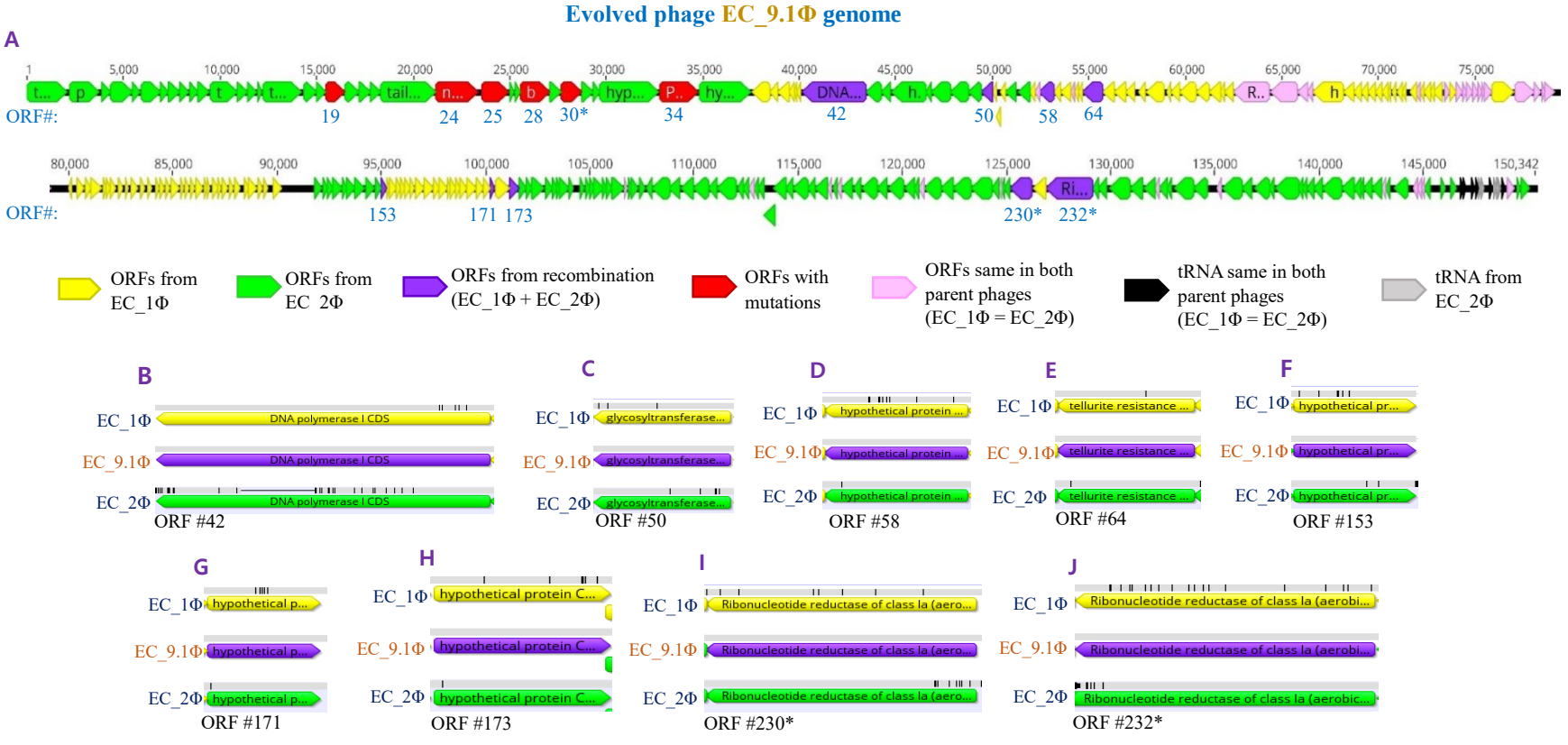


Figure 13. Recombination events and mutations in the evolved phage EC_9.1Φ genome. **A**: shows the origin of the different ORFs: CDSs (colored arrows) and tRNA (black and gray arrows) of the EC_9.1Φ, where yellow arrows are the ORFs identical with EC_1Φ, green arrows are ORFs identical to EC_2Φ, purple arrows are ORFs with recombination of EC_1Φ + EC_2Φ, red arrows are ORF with mutations, and pink arrows are ORFs that are same in both EC_1Φ and EC_2Φ. **B-J:** show the recombination of EC_9.1Φ's ORFs between EC_1Φ and EC_2Φ's ORFs. **K-O**: show EC_9.1Φ's ORFs that came from EC_2Φ with missense mutations. And **P** shows the ORF that come from the EC_2Φ with a nucleotide deletion. *Same in EC_7.1Φ. In **B-P** grey and black parts indicate the agreement and disagreement of nucleotide matching, respectively.

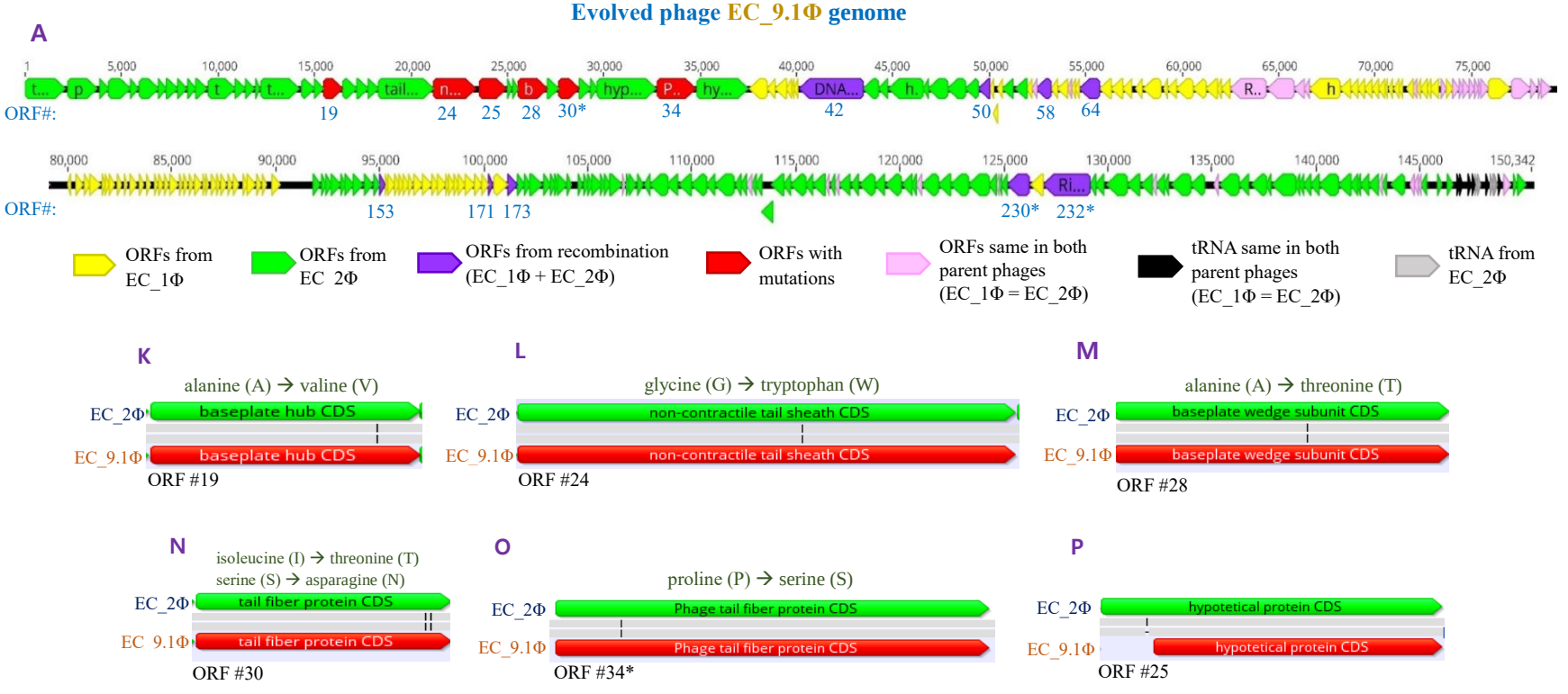


Figure 13 (Cont.). Recombination events and mutations in the evolved phage EC_9.1Φ genome. **A**: shows the origin of the different ORFs: CDSs (colored arrows) and tRNA (black and gray arrows) of the EC_9.1Φ, where yellow arrows are the ORFs identical with EC_1Φ, green arrows are ORFs identical to EC_2Φ, purple arrows are ORFs with recombination of EC_1Φ + EC_2Φ, red arrows are ORF with mutations, and pink arrows are ORFs that are same in both EC_1Φ and EC_2Φ. **B-J:** show the recombination of EC_9.1Φ's ORFs between EC_1Φ and EC_2Φ's ORFs. **K-O**: show EC_9.1Φ's ORFs that came from EC_2Φ with missense mutations. And **P** shows the ORF that come from the EC_2Φ with a nucleotide deletion. *Same in EC_7.1Φ. In **B-P** grey and black parts indicate the agreement and disagreement of nucleotide matching, respectively.

(3) Evolved phage EC_8.2Φ (From Set-2)

The evolved EC_8.2Φ phage genome, obtained from the Set-2 experiment, did not exhibit any significant DNA similarity with any of the four original phages (Fig. 14). However, it did show DNA homology with prophage regions of the Set-2 *E. coli* EC_5.2 (Fig. 15). The sequence length of the evolved phage EC_8.2Φ is 32,064 base pairs.

The screening of the presence of prophage regions of the genome of the *E. coli* strain EC_5.2 used in the Set-2 of experiment for phage evolution, identified 17 prophage regions showed in the Table 8, and after performing the genome alignment of the evolved phage EC_8.2Φ with each of the 17 prophages regions, three prophages regions (region 1, 9 and 12) matched with the genome of the EC_8.2Φ (Fig. 15).

In Table 9 are the most common phages name of the 3 prophage regions that originated the phage EC_8.2. In the table, the phage names that had more than 4 hit genes matching are highlighted with colorful letters. Phage names sharing the same color within the list indicate a commonality, where the 3 regions share common phage names. Among them, the 3 most in common phage names between the 3 regions are PHAGE_Yersin_L_413C_NC_004745, PHAGE_Entero_fiAA91_ss_NC_022750, and PHAGE_Escher_vB_EcoM_ 12474III_NC_ 049457.

The alignment of the annotated genomes of the evolved phage EC_8.2Φ with the 3 prophage regions of the *E. coli* strain EC_5.2 allowed us to know the origin of the ORFs of the EC_8.2Φ (Fig. 16A). Among of the 43 ORFs of the phage EC_8.2Φ, twenty ORFs came from the EC_5.2 prophage region 9, ten ORFs came from the region 1, ten ORFs came from the region 12. Two ORFs were a recombination between region 9 and 12 genes, the **ORF #1** coding the phage terminase, ATPase subunit GpP (Fig. 16B), and the **ORF #5** coding the

phage replication protein GpA, endonuclease (Fig. 16C). One **ORF (#13)** originated from prophage region 12 but contained a single nucleotide difference leading a missense mutation affecting the phage immunity repressor protein GpC. This mutation involved the substitution of a guanine (G) with an adenine (A), resulting in the replacement of the amino acid glycine (G) with glutamic acid (E) (Fig. 17D).

Table 8. Prophage regions in the *Escherichia coli* strain EC_5.2

Region	Region length	Completeness (score)	Region Position	# Total Proteins	Fisrt Most Common Phage Name (hit genes count)	GC %
1	18.2Kb	intact(150)	350-18563	24	PHAGE_Yersin_L_413C_NC_004745(15),	52.72%
2	49.8Kb	intact(150)	1858644-1908485	58	PHAGE_Entero_P88_NC_026014(42),	52.00%
3	36.8Kb	questionable(90)	1921582-1958385	53	PHAGE_Burkho_BcepMu_NC_005882(31)	54.93%
4	33.5Kb	incomplete(20)	2158312-2191860	19	PHAGE_Escher_SH2026Stx1_NC_049919(5)	49.42%
5	25.5Kb	intact(150)	2199466-2224993	32	PHAGE_Entero_cdtI_NC_009514(6)	51.64%
6	36Kb	incomplete(30)	2362456-2398500	32	PHAGE_Pectob_ZF40_NC_019522(5)	50.38%
7	49.4Kb	intact(150)	2835764-2885234	65	PHAGE_Pectob_ZF40_NC_019522(12)	48.17%
8	52.3Kb	intact(150)	2926872-2979175	58	PHAGE_Entero_BP_4795_NC_004813(22)	52.89%
9	27.8Kb	intact(120)	3092797-3120632	38	PHAGE_Yersin_L_413C_NC_004745(23)	50.60%
10	87.2Kb	intact(150)	3664212-3751473	107	PHAGE_Entero_mEp460_NC_019716(32)	51.04%
11	9.9Kb	incomplete(50)	4143976-4153944	15	PHAGE_Escher_SH2026Stx1_NC_049919(3)	48.68%
12	39.3Kb	questionable(70)	5062915-5102313	18	PHAGE_Entero_fiAA91_ss_NC_022750(9)	48.72%
13	6.6Kb	incomplete(50)	5099167-5105859	12	PHAGE_Entero_P1_NC_005856(2)	47.80%
14	14.2Kb	questionable(70)	5114116-5128405	18	PHAGE_Escher_SH2026Stx1_NC_049919(4)	55.99%
15	30.1Kb	incomplete(30)	5146973-5177098	17	PHAGE_Entero_933W_NC_000924(2)	49.92%
16	6.5Kb	incomplete(30)	5233794-5240318	13	PHAGE_Escher_phi191_NC_028660(1)	55.34%
17	24.8Kb	intact(150)	5291330-5316173	38	PHAGE_Stx2_c_Stx2a_F451_NC_049924(3)	49.87%

Table 9. List of the most common phage names of the prophage regions 1, 9 and 12 of the *Escherichia coli* EC_5.2 strains

Most Common Phage Names (hit genes count)		
Region 1	**Region 9**	**Region 12**
PHAGE_Yersin_L_413C_NC_004745(15)	PHAGE_Yersin_L_413C_NC_004745(23)	PHAGE_Entero_fiAA91_ss_NC_022750(9)
PHAGE_Entero_fiAA91_ss_NC_022750(15)	PHAGE_Entero_fiAA91_ss_NC_022750(19)	PHAGE_Escher_P2_NC_041848(9)
PHAGE_Escher_vB_EcoM_12474III_NC_049457(15)	PHAGE_Entero_WPhi_NC_005056(18)	PHAGE_Yersin_L_413C_NC_004745(8),
PHAGE_Escher_pro483_NC_028943(13)	PHAGE_Escher_vB_EcoM_12474III_NC_049457(18)	PHAGE_Salmon_SP_004_NC_021774(7)
PHAGE_Escher_P2_NC_041848(12)	PHAGE_Entero_P2_NC_001895(17)	PHAGE_Escher_pro147_NC_028896(7)
PHAGE_Entero_WPhi_NC_005056(11)	PHAGE_Escher_pro483_NC_028943(17)	PHAGE_Escher_vB_EcoM_12474III_NC_049457(5)
PHAGE_Entero_P2_NC_001895(10)	PHAGE_Escher_pro147_NC_028896(16)	PHAGE_Entero_WPhi_NC_005056(5)
PHAGE_Escher_pro147_NC_028896(7)	PHAGE_Escher_P2_NC_041848(12)	PHAGE_Entero_P2_NC_001895(4)
PHAGE_Yersin_PST_NC_027404(1)	PHAGE_Salmon_SP_004_NC_021774(5)	PHAGE_Escher_pro483_NC_028943(4)
PHAGE_Entero_HK630_NC_019723(1)	PHAGE_Klebsi_ST147_VIM1phi7.1_NC_049451(2)	PHAGE_Escher_500465_1_NC_049342(3)
PHAGE_Salmon_RE_2010_NC_019488(1)	PHAGE_Klebsi_ST16_OXA48phi5.4_NC_049450(2)	PHAGE_Salmon_118970_sal3_NC_031940(2)
PHAGE_Entero_HK629_NC_019711(1)	PHAGE_Entero_lambda_NC_001416(1)	PHAGE_Salmon_SI7_NC_049460(2)
PHAGE_Entero_lambda_NC_001416(1)	PHAGE_Escher_ECML_134_NC_025449(1)	PHAGE_Escher_503458_NC_049341(2)
PHAGE_Escher_ECML_134_NC_025449(1)	PHAGE_Escher_RCS47_NC_042128(1)	PHAGE_Escher_520873_NC_049344(2)
	PHAGE_Entero_HK629_NC_019711(1)	PHAGE_Vibrio_VP58.5_NC_027981(1)
	PHAGE_Yersin_PST_NC_027404(1)	PHAGE_Altero_vB_AmeM_PT11_V22_NC_048847(1)
	PHAGE_Klebsi_3LV2017_NC_047817(1)	PHAGE_Altero_vB_AspP_H4/4_NC_047834(1)
	PHAGE_Clostr_c_st_NC_007581(1)	PHAGE_Entero_phiT5282H_NC_049429(1)
	PHAGE_Salmon_SW9_NC_049459(1)	PHAGE_Entero_mEp460_NC_019716(1)
	PHAGE_Entero_HK630_NC_019723(1)	
	PHAGE_Salmon_SEN5_NC_028701(1)	
	PHAGE_Geobac_GBSV1_NC_008376(1)	

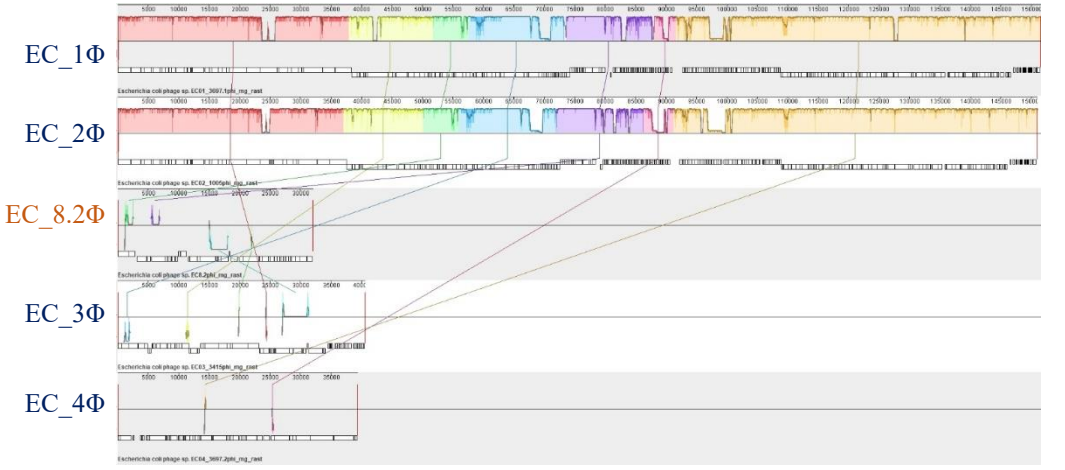


Figure 14. Multiple genome alignment between the evolved phage EC_8.2Φ with the four original phages using progressive MAUVE.

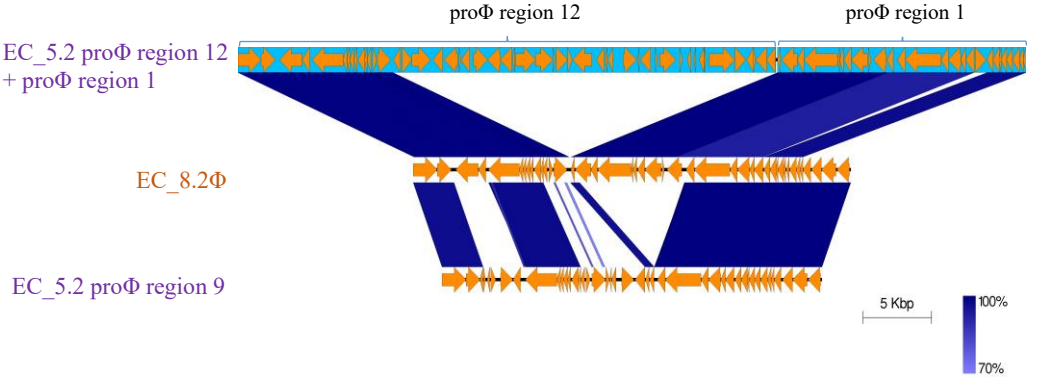


Figure 15. Multiple genome alignment between the evolved phage EC_8.2Φ and the prophage regions 1, 9 and 12 of the strain EC_5.2 using Easyfig ver. 2.2.5.

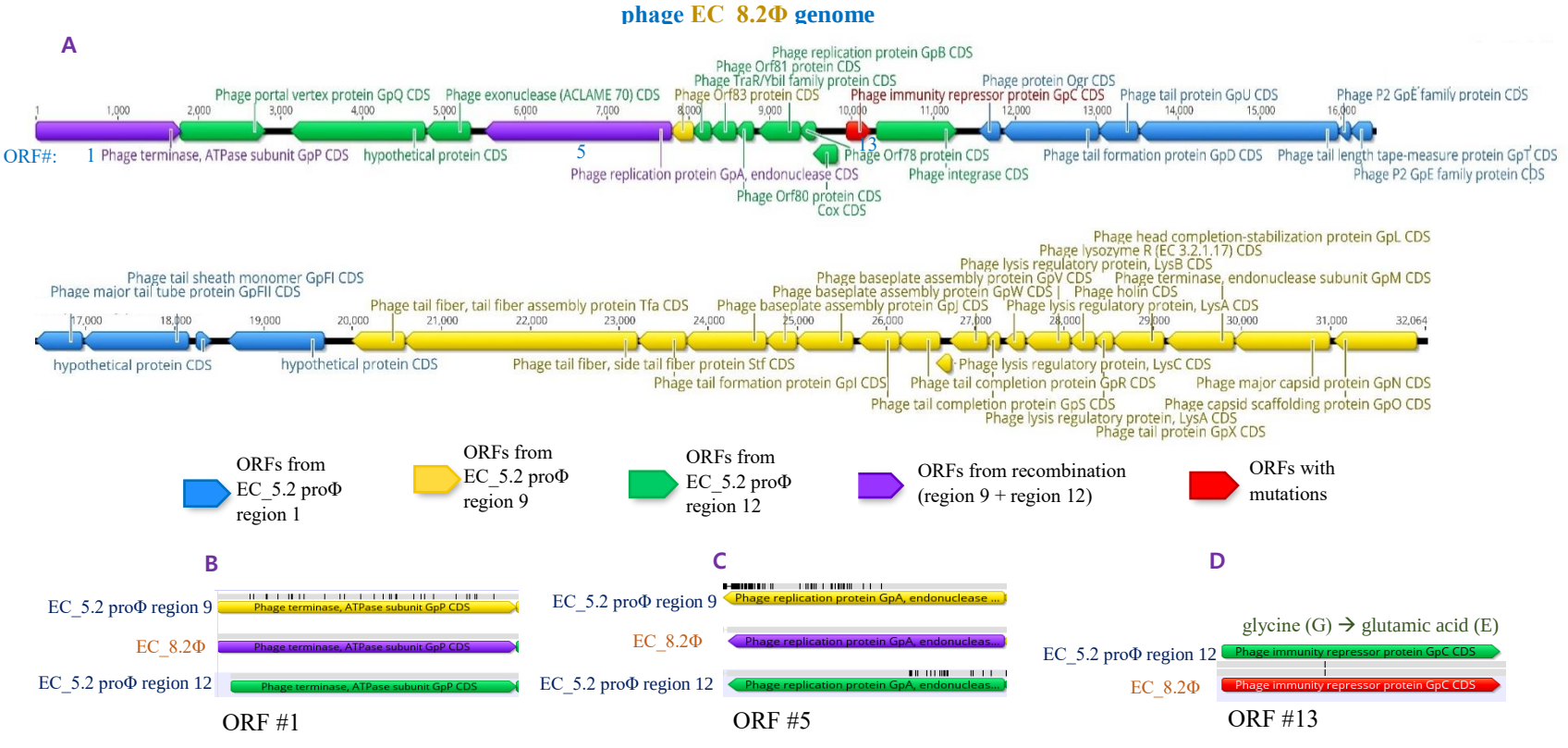


Figure 16. Phage EC_8.2 origin from the *E. coli* EC_5.2 prophage regions. **A:** shows the origin of the different ORFs of the EC_8.2Φ where blue arrows represent ORFs identical with EC_5.2 prophage region 1's proteins, yellow arrows are the ORFs identical with the EC_5.2 prophage region 9, arrows of green color are the ORFs identical with the EC_5.2 prophage region 12, purple arrows are the ORFs resulted of recombination between EC_5.2 prophage regions 9 and 12's ORFs, and the red arrow is the ORF that came from the prophage region 12 with missense mutation. **B and C:** show the recombination of EC_8.2Φ's ORFs between EC_5.2 prophage regions 9 and 12. **D:** shows EC_8.2Φ's ORF that came from the prophage region 12 with missense mutation. In **B-D** grey and black parts indicate the agreement and disagreement of nucleotides matching, respectively.

E. Prophage induction experiment in the 17 *E. coli* strains from both sets

Having obtained a phage (EC_8.2Φ) that did not come from any of the original phages but rather came from prophages regions of one of the *E. coli* bacteria used in the phage evolution experiment, we conducted an additional experiment to induce prophages in the genome of the 17 tested *E. coli* strains using mitomycin C (mitC), which is well known to induce prophage to come out**.** A control group, without mitomycin C, was included to observe any spontaneous induction of prophages during the experiment.

During the experiment, each bacterium was cultured for a short period with or without the addition of mitomycin C. The cultures were then centrifuged, and the resulting supernatant was sterile-filtered and analyzed to detect the presence of inducible prophages capable of producing plaques. For detailed methodology, please refer to the method section. After the prophage induction experiment, filtered supernatants from each *E. coli* strain were tested for phage plaque formation on their respective strain (e.g., EC_3 strain supernatant on EC_3 strain). However, none of the bacterial supernatants, whether induced with mitomycin C or not, produced plaques on their own bacteria. (Table 10A). Later we tested each bacterial supernatant against the *E. coli* strain EC_8.2 (the host of phage EC_8.2Φ) and coincidentally only the supernatants (+/- mitomycin C) of the *E. coli* strain EC_5.2 produced plaques on the *E. coli* strain EC_8.2 (Table 10B). We then tested the supernatants from strain EC_5.2, both with and without mitomycin C induction, against all 17 *E. coli* strains. However, plaque formation was only observed on *E. coli* strain EC_8.2 (Table 10C). Finally, 2 plaques with different morphology were isolated from each EC_5 strain supernatant (with and without mitC) on the host EC_8.2 to be later purified and propagated. Later a host range test of the propagated plaques was performed against the 17 *E. coli* strains, where the plaque #1 (mitC) and plaque #1 (control) infected only the strains

EC_8.2 and EC_11 coinciding with the host range of the evolved phage EC_8.2Φ among the 17 *E. coli* strains, and the plaque #2 (mitC) and plaque #2 (control) only infected the strain EC_8.2 (Table 10D).

Table 10. Prophage induction with and without mitomycin C in the 17 *E. coli* strains used for the phage evolution experiment

A

Prophage induction in the 17 *E. coli* strains against itself			
E. coli supernatant vs	*E. coli* strain	Prophage induction	
		with mitomycin C	control (without mitomycin C)
EC_1	EC_1	No plaques	No plaques
EC_2	EC_2	No plaques	No plaques
EC_3	EC_3	No plaques	No plaques
EC_4.1	EC_4.1	No plaques	No plaques
EC_5.1	EC_5.1	No plaques	No plaques
EC_6.1	EC_6.1	No plaques	No plaques
EC_7.1	EC_7.1	No plaques	No plaques
EC_8.1	EC_8.1	No plaques	No plaques
EC_9.1	EC_9.1	No plaques	No plaques
EC_10	EC_10	No plaques	No plaques
EC_11	EC_11	No plaques	No plaques
EC_4.2	EC_4.2	No plaques	No plaques
EC_5.2	EC_5.2	No plaques	No plaques
EC_6.2	EC_6.2	No plaques	No plaques
EC_7.2	EC_7.2	No plaques	No plaques
EC_8.2	EC_8.2	No plaques	No plaques
EC_9.2	EC_9.2	No plaques	No plaques

B

Prophage induction in the 17 *E. coli* strains against the EC_8.2 strain			
E. coli supernatant vs	*E. coli* strain	Prophage induction	
		with mitomycin C	control (without mitomycin C)
EC_1	EC_8.2	No plaques	No plaques
EC_2	EC_8.2	No plaques	No plaques
EC_3	EC_8.2	No plaques	No plaques
EC_4.1	EC_8.2	No plaques	No plaques
EC_5.1	EC_8.2	No plaques	No plaques
EC_6.1	EC_8.2	No plaques	No plaques
EC_7.1	EC_8.2	No plaques	No plaques
EC_8.1	EC_8.2	No plaques	No plaques
EC_9.1	EC_8.2	No plaques	No plaques
EC_10	EC_8.2	No plaques	No plaques
EC_11	EC_8.2	No plaques	No plaques
EC_4.2	EC_8.2	No plaques	No plaques
EC_5.2	EC_8.2	Plaques	Plaques
EC_6.2	EC_8.2	No plaques	No plaques
EC_7.2	EC_8.2	No plaques	No plaques
EC_8.2	EC_8.2	No plaques	No plaques
EC_9.2	EC_8.2	No plaques	No plaques

C

Prophage induction in the EC_5.2 strain against all 17 *E. coli* strains			
E. coli supernatant vs	*E. coli* strain	Prophage induction	
		with mitomycin C	control (without mitomycin C)
EC_5.2	EC_1	No plaques	No plaques
EC_5.2	EC_2	No plaques	No plaques
EC_5.2	EC_3	No plaques	No plaques
EC_5.2	EC_4.	No plaques	No plaques
EC_5.2	EC_5.1	No plaques	No plaques
EC_5.2	EC_6.1	No plaques	No plaques
EC_5.2	EC_7.1	No plaques	No plaques
EC_5.2	EC_8.1	No plaques	No plaques
EC_5.2	EC_9.1	No plaques	No plaques
EC_5.2	EC_10	No plaques	No plaques
EC_5.2	EC_11	No plaques	No plaques
EC_5.2	EC_4.2	No plaques	No plaques
EC_5.2	EC_5.2	No plaques	No plaques
EC_5.2	EC_6.2	No plaques	No plaques
EC_5.2	EC_7.2	No plaques	No plaques
EC_5.2	EC_8.2	Plaques	Plaques
EC_5.2	EC_9.2	No plaques	No plaques

D

Host Range of induced EC_5.2 strain's prophages propagated in EC_8.2 strain against all 17 *E. coli* strains				
E. coli strain	Propagated prophage			
	with mitomycin C		control (without mitomycin C)	
	Plaque #1	Plaque #2	Plaque #1	Plaque #2
EC_1	no lysis	no lysis	no lysis	no lysis
EC_2	no lysis	no lysis	no lysis	no lysis
EC_3	no lysis	no lysis	no lysis	no lysis
EC_4.	no lysis	no lysis	no lysis	no lysis
EC_5.1	no lysis	no lysis	no lysis	no lysis
EC_6.1	no lysis	no lysis	no lysis	no lysis
EC_7.1	no lysis	no lysis	no lysis	no lysis
EC_8.1	no lysis	no lysis	no lysis	no lysis
EC_9.1	no lysis	no lysis	no lysis	no lysis
EC_10	no lysis	no lysis	no lysis	no lysis
EC_11	lysis	no lysis	lysis	no lysis
EC_4.2	no lysis	no lysis	no lysis	no lysis
EC_5.2	no lysis	no lysis	no lysis	no lysis
EC_6.2	no lysis	no lysis	no lysis	no lysis
EC_7.2	no lysis	no lysis	no lysis	no lysis
EC_8.2	lysis	lysis	lysis	lysis
EC_9.2	no lysis	no lysis	no lysis	no lysis

IV. DISCUSSION

Phage therapy is a promising candidate to combat multi-drug resistant bacterial infections. However, there is still much to study and understand about the interactions between phages and their hosts, which can provide valuable insights for improving phage applications. Given that bacteriophages exhibit high specificity and often have a narrow host spectrum, and considering bacteria's ability to adapt and acquire phage resistance through various mechanisms, there is an imperative need to discover and develop phages that can effectively address these challenges.

In our study, 4 phages were co-cultured with sets of *E. coli* strains (hosts, previous hosts and non-previous hosts) to observe how the phages can evolve after a period of 30 days/rounds of consecutive co-culture and propagation. We considered two different sets of *E. coli* strains to use for the phage evolution experiment, where in set-1 we included phage-resistant mutants (previous hosts of some of the cocktail's phages), while in set-2 were included naturally phage-resistant strains (non-previous hosts). Those sets were designed to observe how the phages can evolve depending of the kind/set of bacterial strains used for the experiment.

Set-1

- From the **set-1** phage evolution experiment, we obtained phages that effectively evolved to re-infect phage-resistant mutant strains, showing the capability of the phage to readapt quickly to previous hosts as co-evolutionary arms race, where the phages evolve new strategies of regain infection when faced with those bacterial strains. The two evolved phages (**EC_7.1Φ and EC_9.1Φ**) isolated from the set-1 experiment were product of recombination event + some mutations between two phages (EC_1Φ and EC_2Φ) of the original cocktail that had significant DNA homology. In another study[17] using

Pseudomonas aeruginosa phages, they showed also that the evolved phages were product of recombination events between the phages present in the cocktail, generating more genetic diversity that would contribute to the expanded host range, arguing that those recombination events have more opportunity to occur between phages with high genetic homology in short-term laboratory experiment.

- The **evolved EC_7.1Φ phage** had most of the ORFs/genes inherited from the EC_2Φ phage. However, the EC_7.1Φ phage demonstrated a host range similar to the EC_1Φ phage, except for its newly acquired ability to infect one of the phage-resistant mutant strains, suggesting that this phage acquired specific genes from EC_1Φ related to host recognition. Indeed, the **ORF #25** of the EC_7.1Φ phage encoding part of the tail fiber protein structure came from the EC_1Φ phage (Fig. 10B), and this protein shares significant similarity (82% coverage and 96.44% identity) with a protein previously published[36] called proteolytically matured endosialidase 92* (endoN92*), mature tail spike protein, endo-alpha-2,8-sialidase (Escherichia phage phi92) which is known for its role in recognizing and removing host cell surface receptors. This suggests that this protein could play a critical role in host recognition by EC_7.1Φ phage. Another ORF, which is unique for phage EC_7.1Φ, derived from the recombination of EC_1Φ and EC_2Φ, is the **ORF #24** that encodes the non-contractile tail sheath (Fig. 10C) which is crucial component of phages for the binding to bacterial cells, penetrate the host, and deliver the genetic material for replication, which may contribute to the phage's ability to recognize and infect new hosts, in this case the phage-resistant mutant strain. Also this ORF shared significant similarity (100% coverage and 76% identity) with the protein Phi92_gp142 which functions as a potential tail fiber protein with glycosidase activity or carbohydrate binding module[37].

- Similarly, the another **evolved EC_9.1Φ phage** also inhered most of the ORFs/genes from the EC_2Φ phage and kept similar host range to EC_2Φ besides of its recently acquired capability to infect two of the phage-resistant mutant strains, suggestion changes in genes related to host receptor recognition. Precisely, the EC_9.1Φ presented some mutations in genes (derived from EC_2Φ) coding tail structure proteins responsible for the host recognition. Among those changes in EC_9.1Φ are: the **ORF #19** coding the baseplate hub, the **ORF #24** coding the non-contractile tail sheath, in the **ORF #28** coding the baseplate wedge subunit, and in the **ORF #30** coding tail fiber protein. These mutations resulted in missense mutations, leading to the substitution of specific amino acids (Figs. 13K – 13N). Furthermore, a deletion mutation was identified in the **ORF #25** (Fig. 13P), which encodes a hypothetical protein. It was hypothesized that the function of this protein might be associated also with the phage tail, as it exhibited a significant similarity of 83.85% with a known *E. coli* phage tail protein, as indicated by its accession number WKV168171. As mentioned before the non-contractile tail sheath is an essential component of phages as it plays a critical role in binding to bacterial cells, penetrating the host, and delivering the genetic material required for replication; and the baseplate is a complex structure located at the tail end of the phage and plays also a role in recognizing and binding to particular receptors on the surface of the bacterial cell[38]. We propose the all these multiple mutations in genes coding complex structures related to the bacterial receptor recognition and the host specificity have equipped the evolved phage with the capacity to recognize and re-infect the phages-resistant mutant strains.

- In the case of other recombined or mutated ORFs in **EC_7.1Φ** and **EC_9.1Φ**, which encode DNA polymerase I, methyltransferase, glucosyltransferase,

ribonucleotide reductase or hypothetical proteins, some of them are associated with DNA replication and metabolism. However, it is our belief that these changes do not significantly impact the phages' ability to acquire new capabilities, such as re-infecting phage-resistant strains. In some instances, the functions of these proteins remain unknown.

Set-2

- On the other hand, the phage isolated from **set-2**, the **EC_8.2Φ phage**, was not originated from any of the 4 original/parent phages of the input cocktail. However, it was derived from prophages regions of the ***E. coli* strain EC_5.2**. During our phage evolution experiment was possible to lead to the activation and release of prophages from the bacterial genome. The induced prophage from *E. coli* strain EC_5.2 isolated through the host EC_8.2 revealed a unique pattern of infectivity. Contrary to expectations, this prophage did not infect the original *E. coli* strain EC_5.2 from which it was derived. Instead, it exhibited the capability to infect other bacterial strains (EC_8.2, EC_11, 2018-60) (Table 7), shedding light on the complex dynamics of prophage-host interactions. It suggests that prophages may not always have a straightforward relationship with their host strains, and their infectivity might be influenced by a variety of factors, including the genetic makeup of the host. The mechanism underlying this phenomenon warrants further investigation. However, it appears that the prophages derived from one strain can infect closely genetically related strains, as evidenced by strains EC_5.2, EC_8.2, and EC_11, were found to be closely related in phylogenetic analysis (Fig. 2). We do not have the genetic information of the 2018-60 strain.

- Following the phage evolution experiment in set-2, which resulted in the isolation of the phage **EC_8.2Φ** derived from bacterial prophage, we conducted a prophage induction with mitomycin C and a control group

without mitomycin C (for spontaneous prophage induction) of all the 17 *E. coli* strains utilized for the phage evolution experiment. Even though prophage regions were found in the genome of all the 17 strains, no induced prophages (with or without mitomycin C) were observed to infect their respective own strains (Table 10A). However, since we knew that the **EC_8.2Φ** phage genome came from prophage regions of the EC_5.2 strain, we tested if the supernatant of this strain (EC_5.2) after the prophage induction experiment could infect or not the strain EC_8.2. Remarkably, only the supernatants of the strain EC_5.2 (+/- mitomycin C) produced visible plaques on the strain EC_8.2 (Tables 10B & 10C). Subsequently, plaques isolated and propagated from the EC_5.2 strain supernatants (+/- mitomycin C) were found to exclusively infect the EC_8.2 and EC_11 strains among the 17 tested strains (Table 10D). This mirrored the host range observed for the **EC_8.2Φ** phage across the same **17 *E. coli* strains**, as outlined in Tables 6 and 7. These findings strongly suggest the prophage isolated from the EC_5.2 strain supernatant and the **EC_8.2Φ** phage are likely the same or ancestors.

- Furthermore, the prophage induction in the *E. coli* strain EC_5.2 appeared to be spontaneous, as it did not require mitomycin C for induction. An important yet frequently overlooked occurrence is the spontaneous activation of these elements within individual bacterial cells, even when there's no external stimulus present, a phenomenon referred as "spontaneous prophage induction" (SPI)[39]. In one study with *Mycobacterium abscessus* phages, authors[40] isolated the phage phiT46-1 in the strain *M. abscessus* BWH-C from the culture supernatant of another strain *M. abscessus* Taiwan-46. In a different study, the authors[41] suggested that inducible prophage mutants of *E. coli* could lyse new hosts, where among the 54 *E. coli* strains tested, the lysates (supernatant) of *E. coli* strains K88 and DE147 generated phages capable of lysing clinical isolates of avian pathogenic *E. coli* (APEC) strain

DE048. These studies corroborate our findings, demonstrating that the supernatant of one strain can contain inducible prophages capable of infecting other hosts.

- With the result of the prophage induction experiment, we can deduce that the ancestor (prophage) of the **EC_8.2Φ** phage was originated at the beginning of the phage evolve experiment from the strain EC_5.2 and was continuously propagated on its new hosts (EC_8.2 and EC_11) of the set-2 for the 30 days/round when the **EC_8.2Φ** was finally isolated. Genome analysis of **EC_8.2Φ** phage reveled that two ORFs of this phage were product of recombination of 2 prophages regions of the EC_5.2 strain. First, the **ORF #1** that codes for the phage terminase, ATPase subunit GpP (Fig. 16B), a vital component involved in supplying energy for translocating phage DNA into the capsid during viral replication[42]. The incorporation of this gene highlights the adaptive nature of EC_8.2Φ, as this enzyme is crucial for efficient replication cycles. Second, the **ORF #5** that encodes the phage replication protein GpA, endonuclease (Fig. 16C), a key player in phage DNA replication processes. Furthermore, a missense mutation was observed in **ORF #13**, originating from prophage region 12 of the EC_5.2 strain (Fig. 16D). This ORF encodes for the phage immunity repressor protein GpC, which plays a vital role in regulating the lysogenic/lytic cycle switch[43]. The mutation in this regulatory protein hints at potential alterations in the phage's lifecycle dynamics, possibly affecting its ability to enter lysogenic or lytic pathways in response to environmental signals.

Our findings suggest that the outcome of phage evolution experiments may be influenced by the choice of bacterial strains used in the experiment, as well as their genetic background and the genetic similarity of the phages present in the cocktail.

In the set-1 experiment, phages with significant DNA homology within the input cocktail underwent recombination and acquired mutations that facilitated their adaptation to and reinfection of phage-resistant mutant strains. The presence of these phage-resistant mutant strains in the set-1 experiment, which were formerly hosts, played a pivotal role in driving rapid phage evolution for re-infection. Indeed, bacteriophages can evolve and again infect the phage-resistant mutant strains[13-15]. Coevolution between bacteriophage with their respective host, is a common phenomenon that occur naturally in microbial communities[44]. Additionally, bacteriophages can enhance their host range efficacy through recombination with other phages[17] or by acquiring genes from prophages within host bacterial genomes[14]. This last mechanism suggests that one route to the improvement of evolved phages involves recombination with prophage genes present in the host bacteria's genome[14]. However, in our study, original phages did not engage in recombination with prophage genes from our tested bacterial strains. In contrast, in a separate study, an evolved phage rapidly emerged independently, employing directed evolution to re-infect phage-resistant *E. coli* variants[15].

On the other hand, in the set-2 experiment, the four original phages were unable to evolve to infect any of the tested strains, as these strains were naturally resistant to the four phages and were phylogenetically distant from the phages' usual hosts. Nevertheless, we succeeded in obtaining a phage derived from the prophage regions of the EC_5.2 *E. coli* strain. This suggests that the ability to induce prophages depends on the specific strain included in the experiment; without the EC_5.2 strain, our results would have been different. Notably, the EC_5.2 strain carried a larger number of intact prophage regions compared to the other tested strains (Table 4), which likely increased the probability of obtaining inducible prophages. It is worth mentioning that even though the EC_11 strain was present in both sets (Table 6), induction of prophages from the Set-2 strain

(EC_5.2) resulted in a phage capable of infecting the EC_11 strain, which was not observed in Set-1.

The Appelmans protocol employed in the phage evolution experiment effectively facilitated the evolution of the parent phage, enabling it to re-infect phage-resistant mutant strains and inducing prophages spontaneously from the bacterial genome.

V. CONCLUSIONS

- Our research indicates that phages exhibit a heightened capacity for evolutionary adaptation when confronted with hosts to which they were previously susceptible. This adaptation is facilitated by specific genetic mechanisms, primarily recombination events and missense mutations in key phage tail structure proteins.
- The presence of prophages within a bacterial genome reveals a remarkable capacity for these genetic elements to 'jump' and infect other bacterial strains.
- The phage evolution seemed to depend on the bacterial sets, with phages in Set-1 readapting to infect phage-resistant mutants while the emergence of the phages in Set-2 appeared to be independent of the original phages but rather driven by prophages present in the bacterial set.
- Studying phage-host interactions can help in developing better phage-based therapies for MDR infections by understanding how phages evolve and adapt to new host strains.

REFERENCES

1. Centers for Disease Control and Prevention (CDC), National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), Division of Foodborne, Waterborne, and Environmental Diseases (DFWED). *E. coli (Escherichia coli);* [reviewed 2022 Aug 17]. Available from https://www.cdc.gov/ecoli/index.html
2. Doi Y, Park YS, Rivera JI, Adams-Haduch JM, Hingwe A, Sordillo, EM, *et al.* (2012). Community-Associated Extended-Spectrum β-Lactamase–Producing Escherichia coli Infection in the United States. *Clinical Infectious Diseases*, *56*(5), 641–648.
3. Khan HA, Ahmad A, & Mehboob R (2015). Nosocomial infections and their control strategies. *Asian Pacific Journal of Tropical Biomedicine*, *5*(7), 509–514.
4. Ibrahim M, Bilal N, & Hamid, M. (2013). Increased multi-drug resistant Escherichia coli from hospitals in Khartoum state, Sudan. *African Health Sciences*, *12*(3).
5. Petty NK, ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde B, *et al.* (2014). Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proceedings of the National Academy of Sciences*, *111*(15), 5694–5699.
6. Bao D, Xu X, Wang Y & Zhu F. (2022). Emergence of a Multidrug-Resistant Escherichia coli Co-Carrying a New mcr-1.33 Variant and blaNDM-5 Genes Recovered from a Urinary Tract Infection. *Infection and Drug Resistance*, *Volume 15*, 1499–1503.
7. Clokie MRJ, Millard AD, Letaro, AV & Heaphy S. (2011). Phages in nature. *Bacteriophage*, *1*(1), 31–45.
8. Hatfull GF & Hendrix RW. (2011). Bacteriophages and their genomes. *Current Opinion in Virology*, *1*(4), 298–303.

9. Harper DR, Abedon, ST, Burrowes BH & McConville ML (2021). *Bacteriophages: Biology, Technology, Therapy* (1st ed. 2021 ed.). Springer. bundle).
10. Lamy-Besnier Q, Chaffringeon L, Lourenço M, Payne RB, Trinh JT, Schwartz JA, *et al.* (2021). Prophylactic Administration of a Bacteriophage Cocktail Is Safe and Effective in Reducing Salmonella enterica Serovar Typhimurium Burden in Vivo. *Microbiology Spectrum*, *9*(1).
11. Stone E, Campbell K, Grant I & McAuliffe O. (2019). Understanding and Exploiting Phage–Host Interactions. *Viruses*, *11*(6), 567.
12. Oechslin F. (2018). Resistance Development to Bacteriophages Occurring during Bacteriophage Therapy. *Viruses*, *10*(7), 351.
13. Mizoguchi K, Morita M, Fischer CR, Yoichi M, Tanji Y & Unno H. (2003). Coevolution of Bacteriophage PP01 and *Escherichia coli* O157:H7 in Continuous Culture. *Applied and Environmental Microbiology*, *69*(1), 170–176.
14. Borin JM, Avrani S, Barrick JE, Petrie KL, & Meyer JR. (2021). Coevolutionary phage training leads to greater bacterial suppression and delays the evolution of phage resistance. *Proceedings of the National Academy of Sciences*, *118*(23).
15. Salazar KC, Ma L, Green SI, Zulk JJ, Trautner BW, Ramig RF, *et al.* (2021). Antiviral Resistance and Phage Counter Adaptation to Antibiotic-Resistant Extraintestinal Pathogenic *Escherichia coli*. *mBio*, *12*(2).
16. Międzybrodzki R, Borysowski J, Weber-Dąbrowska B, Fortuna W, Letkiewicz S, Szufnarowski K, *et al.* (2012). Clinical Aspects of Phage Therapy. *Advances in Virus Research*, 73–121.
17. Burrowes B, Molineux I & Fralick J. (2019). Directed in Vitro Evolution of Therapeutic Bacteriophages: The Appelmans Protocol. *Viruses*, *11*(3), 241.

18. Pires DP, Cleto S, Sillankorva S, Azeredo J, Lu TK. (2016). Genetically Engineered Phages: a Review of Advances over the Last Decade. *Microbiology and Molecular Biology Reviews*, *80*(3), 523–543.
19. Ryan EM, Alkawareek MY, Donnelly RF, Gilmore BF. (2012). Synergistic phage-antibiotic combinations for the control of*Escherichia coli*biofilms*in vitro*. *FEMS Immunology & Medical Microbiology*, *65*(2), 395–398.
20. Wang L, Tkhilaishvili T, Bernal Andres B, Trampuz A, Gonzalez Moreno M (2020). Bacteriophage–antibiotic combinations against ciprofloxacin/ceftriaxone-resistant Escherichia coli in vitro and in an experimental Galleria mellonella model. *International Journal of Antimicrobial Agents*, *56*(6), 106200.
21. Magiorakos AP, Srinivasan A, Carey R, Carmeli Y, Falagas M, Giske C, *et al.* (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clinical Microbiology and Infection*, *18*(3), 268–281.
22. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing*. 32nd ed. CLSI supplement M100. Clinical and Laboratory Standards Institute; 2022.
23. Clokie MRJ, Kropinski A. (2010). *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions (Methods in Molecular Biology, 501)* (Softcover reprint of hardcover 1st ed. 2009 ed.). Humana.
24. Bourdin G, Schmitt B, Guy LM, Germond J, Zuber S, Michot L, *et al*. (2014). Amplification and Purification of T4-Like Escherichia coli Phages for Phage Therapy: from Laboratory to Pilot Scale. *Applied and Environmental Microbiology*, *80*(4), 1469–1476.
25. Jeon J, Yong D. (2019). Two Novel Bacteriophages Improve Survival in *Galleria mellonella* Infection and Mouse Acute Pneumonia Models Infected

with Extensively Drug-Resistant *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, *85*(9).

26. Raza T, Andleeb S, Ullah SR, Jamal M, Mehmood K, Ali M. (2018). Isolation and Characterization of a Phage to Control Vancomycin Resistant Enterococcus faecium. Open Life Sciences, 13(1), 553–560.
27. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, *et al.* (2012). Multilocus sequence typing of Total-Genome-Sequenced bacteria. *Journal of Clinical Microbiology*, *50*(4), 1355–1361.
28. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. (2014). Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. *PLOS ONE*, *9*(8), e104984.
29. Bortolaia V, Kaas RS, Ruppé É, Roberts MC, Schwarz Š, Cattoir V, *et al.* (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *Journal of Antimicrobial Chemotherapy*, *75*(12), 3491–3500.
30. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos JS, Bealer K, *et al.* (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*(1).
31. Zankari E, Allesøe RL, Joensen KG, Cavaco L, Lund O, Aarestrup, FM. (2017). PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *Journal of Antimicrobial Chemotherapy*, *72*(10), 2764–2768.
32. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, *et al.* (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*(W1), W16–W21.
33. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, *et al*. (2013). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, *42*(D1), D206–D214.

34. Katoh K, Standley DM. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. Molecular Biology and Evolution, 30(4), 772–780.
35. Filipiak M, Łoś J, Łoś M. (2019). Efficiency of induction of Shiga-toxin lambdoid prophages in Escherichia coli due to oxidative and antibiotic stress depends on the combination of prophage and the bacterial strain. *Journal of Applied Genetics*, *61*(1), 131–140.
36. Schwarzer D, Browning CH, Stummeyer K, Oberbeck A, Mühlenhoff M, Gerardy-Schahn R, *et al.* (2015). Structure and biochemical characterization of bacteriophage phi92 endosialidase. *Virology*, *477*, 133–143.
37. Schwarzer D, Buettner FFR, Browning CH, Nazarov S, Rabsch W, Bethe A, *et al.* (2012). A Multivalent Adsorption Apparatus Explains the Broad Host Range of Phage phi92: a Comprehensive Genomic and Structural Analysis. *Journal of Virology*, *86*(19), 10384–10398.
38. Yap ML, Klose T, Arisaka F, Speir JA, Veesler D, Fokine A, *et al.* (2016). Role of bacteriophage T4 baseplate in regulating assembly and infection. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2654–2659.
39. Nanda A, Thormann KM, Frunzke J. (2014). Impact of spontaneous prophage induction on the fitness of bacterial populations and Host-Microbe interactions. *Journal of Bacteriology*, *197*(3), 410–419.
40. Amarh ED, Dedrick RM, Garlena RA, Russell DA, Jacobs-Sera D, Hatfull G. F. (2021). Genome Sequence of Mycobacterium abscessus Phage phiT46-1. *Microbiology Resource Announcements*, *10*(2).
41. Chen M, Zhang L, Xin S, Yao H, Chen L, Zhang W. (2017). Inducible Prophage Mutant of Escherichia coli Can Lyse New Host and the Key Sites of Receptor Recognition Identification. *Frontiers in Microbiology*, *8*.

42. Shen X, Li M, Zeng Y, Hu X, Tan Y, Rao X, *et al.* (2012). Functional identification of the DNA packaging terminase from Pseudomonas aeruginosa phage PaP3. *Archives of Virology, 157*(11), 2133–2141.
43. Häuser R, Blasche S, Dokland T, Haggård-Ljungquist E, Von Brunn A, Salas M, *et al.* (2012). Bacteriophage Protein–Protein interactions. In *Elsevier eBooks* (pp. 219–298).
44. Koskella B, Brockhurst, MA. (2014). Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiology Reviews, 38*(5), 916–931.

ABSTRACT (IN KOREAN)

다제내성 *Escherichia coli*를 이용한 박테리오파지 진화의 유전 기작 규명

<지도교수 용동은>

연세대학교 대학원 의과학과

아바디에 사엔즈 리카르도 엔리케

다제내성균은 국제 공중보건에 대한 주요 위협 중 하나이다. 특히 카바페넴 내성 *Escherichia coli* (EC)는 심각한 감염을 유발하여 혁신적인 치료 전략이 필요한 가장 긴급한 위협 중 하나이다. 세균를 죽일 수 있는 바이러스인 박테리오파지(파지)는 다제내성균 감염에 대응하기 위한 대안으로 연구되고 있다. 그러나 세균의 파지 내성 획득과 파지의 제한된 숙주 감염 범위는 극복해야 할 요인이다. 본 연구에서 다제내성 EC를 표적으로 하는 파지를 진화시키기 위해 단기 실험을 수행하였다. 숙주 감염 범위를 넓히고 관련된 유전적 메커니즘을 이해하기 위한 목적으로 4개의 파지로 구성된 칵테일과 11종의 EC 균주로 구성된 2세트의 균주를 각각 30회 동안 공동 배양하였다. 세트 1은 파지 내성 변이체 EC 균주(ΦR-Mut)를 포함하였고, 세트 2는 자연적으로 파지 내성을 갖는 EC 균주(Nat-ΦR)를 포함하였다. 30회의 선택 과정 후, 세트 1에서는 ΦR-Mut 균주를 감염할 수 있는 두 개의 진화된 파지(EC_7.1Φ 및 EC_9.1Φ)가 분리되었다. 이들은 꼬리 구조 단백질 유전자에 유전자 재조합 및 돌연변이를 나타내었다. 이러한 변화가 파지가 ΦR-Mut 균주에 재적응하는 데 잠재적으로 기여할 수 있음을 시사한다. 세트

2에서는 Nat-ΦR 균주를 감염할 수 있는 세 개의 파지(EC_6.2Φ, EC_8.2Φ 및 EC_11Φ)를 얻었다. EC_6.2Φ는 추가 분석을 위한 충분한 정보를 얻지 못했지만, EC_8.2Φ와 EC_11Φ는 거의 유전적으로 동일했다. EC_8.2Φ는 4개의 원래 파지와 유전적으로 다른 특징을 나타냈지만, 하나의 Nat-ΦR 균주(EC_5.2)의 프로파지 유전자와 상동성을 나타냈다. 이러한 결과는 파지가 이전에 파지에 민감했던 숙주에 대항하기 위해 더 빠르게 진화한다는 것을 시사하며, 세균 유전체의 프로파지가 다른 균주에 감염하기 위해 "점프"할 수 있음을 제안한다. 본 연구는 세균의 파지 내성 변이에 대응할 수 있는 파지를 개발하고 세균 유전체의 프로파지를 유도하여 더 효과적인 파지 적용 위한 숙주 감염 범위를 확장하는 잠재력을 제공한다.

핵심되는 말: 박테리오파지, 다제 내성 대장균, 박테리오파지 내성 변이체, 박테리오파지 진화, 프로페이지 유도.

본 논문작성자는 한국정부초청장학금(Global Korea Scholarship)을 지원받은 장학생임