RESEARCH ARTICLE



# ISOLATION, CHARACTERIZATION AND GENOME ANALYSIS OF A NOVEL VIRULENT ESCHERICHIA COLI BACTERIOPHAGE VB\_ECOM\_011D4

Xinwu Wang <sup>[D]</sup>, Hao Zhang <sup>[D]</sup>, Lanting Bi <sup>[D]</sup>, Hengyu Xi<sup>1</sup>, Zijing Wang <sup>[D]</sup>, Yalu Ji <sup>[D]</sup>, Xiao Liu <sup>[D]</sup>, Rihong Zhao <sup>[D]</sup>, Ming Liu<sup>1</sup>, Meimei Zhang<sup>1</sup>, Chunzheng Huang<sup>1</sup>, Hongda Ou<sup>1</sup>, Wenyu Han<sup>1,2</sup>, Jianbao Dong <sup>[D]</sup><sup>3</sup>, Jingmin Gu<sup>1,4</sup> and Wei Zhu <sup>[D]</sup><sup>3</sup>\*

<sup>1</sup>Key Laboratory of Zoonosis Research, Ministry of Education, College of Veterinary Medicine, Jilin University, Changchun 130062, P.R. China; <sup>2</sup>Jiangsu Co-Innovation Center for the Prevention and Control of Important Animal Infectious Disease and Zoonose, Yangzhou University, Yangzhou 225009, P.R. China; <sup>3</sup>Department of Veterinary Medical Science, Shandong Vocational Animal Science and Veterinary College, Weifang 261061, P.R. China;

<sup>4</sup>Shanghai Public Health Clinical Center, Fudan University, Shanghai 201508, P.R. China

\*Corresponding author: niccole522@hotmail.com

## ABSTRACT

Phage therapy is a promising alternative therapy for treating E. coli infection. Although the total number of phages on the earth is as high as 1031, the reported phages and thoroughly studied are very limited. Therefore, the continuous discovery of new phages and in-depth research will provide materials for the wide application of phage therapy in the future. In this study, a novel E. coli phage vB\_EcoM\_011D4 was isolated from sewage samples, and the biological characteristics were studied. Electron microscopy and homology analysis results showed that vB EcoM 011D4 belongs to the family Myoviridae. One-step growth curve showing the latent period of vB EcoM 011D4 was 10min, with 115 PFU/cell burst size. Phage vB\_EcoM\_011D4 was highly stabled under different temperatures (range 4-70°C) and pH conditions (range 6-10). At the same time, its genome was subjected to highthroughput sequencing and compared with the reported phages. The results of high-throughput sequencing assembly showed that vB\_EcoM\_011D4 is a linear, double-stranded DNA virus containing 163764 bp, with an average GC content of 40.50%, and a total of 273 open reading frames (ORFs). Genomic comparison analysis revealed that most ORFs were similar to Enterobacteria phage Phi1 and RB49. However, ORF147 and ORF148 putative DNA methylase family protein is less than 67% homology with already published phages. In addition, the phylogenetic analysis of terminates large subunit showed that it belongs to a new branch and shows less than 50 similarities to reported phages. No lysogenic, toxin, or antibiotic-resistant related gene was found in the genome of vB\_EcoM\_011D4. In summary, vB EcoM 011D4 is a newly discovered phage that can be further studied to elucidate the phage diversity and its benefits for the wide application of phage therapy.

*Keywords: Escherichia coli*; Opportunistic pathogen; Phage; Genome analysis

Article History (2021-0771) || Received: 04 Jul 2021 || Revised: 20 Jul 2021 || Accepted: 23 Jul 2021 || Published Online: 25 Jul 2021

## **1. INTRODUCTION**

*Escherichia coli* (*E. coli*) is a pathogenic bacterium, mainly parasitic in the intestines of warm-blooded animals and cause a variety of animal and human diseases (Jang et al. 2017; Dusek et al. 2018), resulting in more than 2 million deaths each year (Kaper et al. 2004). In poultry and humans, *E. coli* infection predominantly colonized in the lower digestive tract after 24 h of hatching (Ballou et al. 2016) or at birth (Bettelheim and Lennox-King 1976). However, increasingly number of antibiotic-resistant strains were discovered and caused adverse effects on animal and human health, so alternative antimicrobial strategies are essential (Endersen et al. 2014).

Bacteriophages (phages) was thought to be an antimicrobial agent against bacterial infection has been around for a century, but the widespread application of phage therapy has been put on ice by the emergence of antibiotics (Bardina et al. 2016). In recent years, the general decline in the efficacy of exist antibiotics and the deficiency of newly effective antibiotics has made phage therapy regain attention (Bloom et al. 2018) and there are increasing reports on the application of phages in bacterial infected animal models (Cheng et al. 2017; Cafora et al. 2019; Anand et al. 2020), human (Kutter et al. 2010) and food safety (Kang et al. 2013; Han et al. 2017; Huang et al. 2018) since 2006 (Sulakvelidze 2011).



Considering the large number of phages on earth and the lack of understanding of their diversity and mechanism, discovering of new phage and exploring their basic characteristics and genomic diversity is necessary to reveal the evolutionary relationship between phages and the expansion of phage libraries. In this study, a newly lytic E. coli phage vB\_EcoM\_011D4 was isolated, the characteristics and genome information were evaluated, it is maybe helpful to understand co-evolutionary relationship between bacteria and phages, and provide candidate materials for phage therapy in the future.

# 2. MATERIALS AND METHODS

## 2.1. Isolation and Identification of E. coli 011D

*E. coli 011D* was isolated from clinical animal stools collected from abortive donkey in Shandong, China. All the suspected strains were confirmed by biochemical tests (VITEK2 Compact, France bio,) and 16s rDNA sequencing using the universal primers F (5'-AGAGTTTGATCMTGGCTCAG-3') and R (5TACGGYTACCTTGTTACGACTT-3'). The purified strain was cultured in Luria-Bertani (LB) medium (tryptone 10 g/l; NaCl 10g/l; yeast 5 g/l. Becton, Dickinson and Company, USA), then stored in 30% glycerol, at  $-80^{\circ}$ C.

## 2.2. Isolation and purification of vB\_EcoM\_011D4

*E. coli 011D* was used as a host to isolate phage from sewage samples collected from Xizang, China. The double layer agar plate method (Gu et al. 2012) was performed to isolate and purify of phages. In brief, 1 mL of *E. coli 011D* was added to 100mL L.B. liquid medium which was prepared with sewage samples. Cultures were incubated at 37°C with shaking of 180 rpm for 12h. 1mL of cultures were filtered with 0.22µm pore-size filter. 100µL filtered culture and 100µL overnight cultured *E. coli 011D* were added into 7mL 0.75% agar L.B. medium, mixed then poured onto 1.5% agar L.B. plate. The single plaque was picked up and added to 100µL *E. coli* contained L.B. medium, then poured onto 1.5% agar L.B. plate again. This process was repeated three times to obtain the purified phage particles. Finally, the purified phages were stored both at 4°C and -80 °C in glycerol (3:1 [vol/vol]).

#### 2.3. Electron microscopy of vB\_EcoM\_011D4

Large-scale amplification of vB\_EcoM\_011D4 was performed to obtain concentrated phage particles (Cheng et al. 2017). The 800mL of lysates were centrifuged at  $12,000 \times g$  for 10 min at 4°C to remove cell debris. 10% Olyethylene glycol 8,000 and 1M NaCl were added into supernatant and then placed on ice overnight to precipitate phage particles. The resulting pellet suspended in 2mL SM buffer (0.01% gelatin,100mM/L NaCl, 50mmol/L Tris-HCl, and 10mM/L MgSO4), then the suspension was extracted three times with an equal volume of chloroform. The concentrated phage particles were placed on carbon-coated copper grids and absorbed for 15 min and then negatively stained with phosphotungstic acid (PTA,2%w/v). A transmission electron microscope (TEM) (JEOL JEM-1200EXII, Japan Electronics and Optics Laboratory, Tokyo, Japan) at an acceleration voltage of 80 kV was used to observe the morphology.

## 2.4. Thermal and pH stability of vB\_EcoM\_011D4

To analysis the thermal stability, the suspension of vB\_EcoM\_011D4 was incubated at different temperatures (4, 30, 40, 50, 60, 70 and 80°C), and the suspension were collected after at 40 min and 80 min of incubation, then the double-layer agar method was used to detect the phage titers (Gu et al. 2012). To analyze the pH stability, the suspension of vB\_EcoM\_011D4 was mixed with PBS which was pre-adjustified to different pH levels (2, 4, 6, 8, 10, 12 and 13) and incubated at 37°C for an hour, then the double-layer agar method was used to detect the phage titers (Gu et al. 2012).

## 2.5. Multiplicity of Infection (MOI)

Multiplicity of infection (MOI) was defined as the ratio of virus particles to host cells (Chen et al. 2016). Before the MOI test preformed, the colony formation units (CFU) of *E. coli* 011D and plaque formation units of vB\_EcoM\_011D4 were determined, then mixed phage and bacteria to different MOI (10, 1, 0.1, 0.01, 0.001, 0.00001, 0.000001, 0.0000001 and 0.00000001) and incubated at  $37^{\circ}$ C for 3 h. Then the phage titers were evaluated by the double-layer agar plate method.

## 2.6. One-step growth curve of vB\_EcoM\_011D4

The one-step growth process of the vB\_EcoM\_011D4 were estimated as previously described with some modifications (Pajunen et al. 2000). Briefly, a mid-exponential growth culture of *E. coli 011D* was diluted to  $1 \times 10^8$  CFU/mL, then mixed with vB\_EcoM\_011D4 at MOI of 0.1. After incubated at 37°C for 10 min, the



mixture was added to 10mL LB medium and incubated at 37°C with shaking of 180 rpm. At different time intervals, the suspensions were collected, then double-layer method was performed to detect the phage titers.

#### 2.7. Sequencing and Bioinformatics Analysis of the vB\_EcoM\_011D4 Genome

After large-scale amplified and concentrated of phage particles, the genome of vB\_EcoM\_011D4 was extracted using a viral genome extraction kit (Omega Bio-Tek Inc., Norcross, GA, USA). Whole genome was sequenced using IlluminaHiSeq 2500 platform at Wuhan GENEWIZ Biotechnology Co. Ltd and assembled using SPAdes with 1000-fold coverage. The potential function of ORFs were predicted and identified by GeneMarkS and BLAST (Besemer et al. 2001). The potential tRNAs were detected by tRNA scanner (Schattner et al. 2005). G.C. skew and content of genome were observed using CGView (Grant and Stothard 2008). A co-linearity relationship with reported phages was analyzed via Mauve 5.05. The neighbor join-tree was constructed using MEGA version 5.05 by maximum likelihood method with 1000 bootstrap replicates.

## 3. RESULTS

#### 3.1. Isolation and Morphology of the vB\_EcoM\_011D4

The vB\_EcoM\_011D4 was isolated and restored from sewage samples. The plaque on L.B. plate was clear and with a proximate diameter of 0.1cm (Fig. 1A). The Electron microscopy showed that vB\_EcoM\_011D4 belongs to family *Myoviridae* with a proximate head diameter of  $90\pm0.5$ nm (n=5) (Fig. 1B).



**Fig. 1:** Morphological characteristics of the *Escherichia coli* phage vB\_EcoM\_011D4. (A) single Plaque of the phage vB\_EcoM\_011D4 on double-layer L.B. plates; (B) Transmission electron micrograph of the phage vB EcoM 011D4.

#### 3.2. The temperature and pH stability of the vB\_EcoM\_011D4

The vB\_EcoM\_011D4 was cultured under different temperatures and pH conditions to evaluate its stability. The titers of vB\_EcoM\_011D4 was stabled with when incubated at 30°C, 40°C, 50°C, 60°C and 70°C for 80 min, but no plaque was detected after 80 min of incubation at 80°C (Fig. 2A).

As shown in Fig. 2B, compared without treated phages, the titers of vB\_EcoM\_011D4 was stabled when incubated under pH value of 6 to 10. However, no plaque was observed when it was incubated in other extreme acidic or alkaline conditions.

#### 3.3. MOI of the vB\_EcoM\_011D4

The optimal MOI of vB\_EcoM\_011D4 was detected by mixed phage and bacteria in different proportions. The titers of the vB\_EcoM\_011D4 reached the highest peak when phage and bacteria were mixed at MOI of  $10^{-5}$ . This result showed the optimal MOI of vB\_Eco\_D226 was  $10^{-5}$  (Fig. 2C).

## 3.4. One-step growth curve of vB\_EcoM\_011D4

The intracellular proliferation process of vB\_EcoM\_011D4 was examined by the double-layer plate methods (Gu et al. 2012). The first 10 min was the latent period, which was the intracellular increment process, then 10 to 20 min was the rapid replication period, in this period the outer membrane of the bacterial cells were destroyed and progeny phages were released. Then the secondary proliferation process was initiated, this process was finished after 50 min of phage infection with an average burst size of 115 PFU/cell (Fig. 2D).

#### 3.5. Sequencing and bioinformatics analysis of the vB\_EcoM\_011D4

Sequencing and assembling results showed vB\_EcoM\_011D4 was a linear, double-stranded DNA virus, which consist of 163764 bp with an average G.C. content of 40.50% (Fig. 3). When compared its sequence to those already released in NCBI database showed different degrees of similarities to *Enterobacteria* phage





Phi1 (EF437941.1) and RB49 (AY343333.1), *Escherichia* phage kaaroe (MN850574.1), *Shigella* phage Sf20 (MF327006.1), vB\_EcoM\_PHB13 (MK573636.1), *Enterobacteria* phage GEC-3S (HE978309.1), *Escherichia* phage E26 (MN655998.1), *Escherichia* virus Ec\_Makalu\_002 (MN709127.1), Ec\_Makalu\_003 (MN882349.1), and Ec\_Makalu\_001(MN894885.1).



**Fig. 2:** Biological features of the *Escherichia coli* phage vB\_EcoM\_011D4. (A) Stability of the vB\_EcoM\_011D4 in different temperature condition; (B) Stability of the vB\_EcoM\_011D4 in different pH condition; (C) The optimal multiplicity of infection; (D) One-step growth curve of the vB\_EcoM\_011D4 in *E. coli* 011D.

**Fig. 3:** Genetic and physical diagram of the vB\_EcoM\_011D genome. The circle map of the vB\_EcoM\_011D was constructed using GCviw online web (http://wishart.biology.ualberta.ca/cgvie w) (Grant and Stothard 2008). The total 273 putative ORFs were depicted and their transcript direction are shown by arrow. The G.C. content and GC skew are visualized also.

All 273 putative ORFs were found, and 59.3% (162/273) were annotated as hypothetical proteins, and one of them showed no significant similarity. Although most ORFs of vB\_EcoM\_011D4 are highly homologous to reported phages, some sequences such as DNA methylase proteins (OR147, ORF148) showed less than (57.48% ~ 67.26%) sequence similarity to the reported phages. No tRNAs were found in the genome of vB\_EcoM\_011D4 revealing it relies on the host tRNA for transcription. There were no lysogenic-related gene, toxic coding gene, antibiotic resistance genes or phage virulence factors were found in its genome. So, vB\_EcoM\_011D4 is a lytic phage could be further applied in biocontrol and phage therapy.



The co-linearity analysis of the genomic of vB\_EcoM\_011D4 revealed that it contains similar genetic components to these ten phages, which could be divided into three functional modules: DNA packing and morphology, replication and nucleotide metabolism related, and host lysis (Fig. 4).

	40000	60000	80000	100000	120000	140000	16000
		Y Y Y		W. W. W. W. W.	AND DESIGNATION OF A DE		
tarabastaria ni	Dago Phi			>			
20000	40000	60000	80000	100000	120000	140000	16000
1111111111		And Alasta	10000	100000	The second	ALL	TWO NET
nterobacteria pl	hade RB49						
20000	40000	60000	80000	100000	120000	140000	16000
		A States				All a sub-	NW T
		12					
scherichia phag	e Kaaroe	$\leftarrow$					
20000	40000	60000	80000	100000	120000	140000	16000
V				an a		Under the second second	
nigella phage St	40000	60000	80000	100000	120000	140000	16000
20000	10000			AVM THE	The second	MULTIN (MALIN	CT WIN
B_ECOM_PHB13	40000	60000	80000	100000	170000	140000	16000
20000	40000	60000	80000	100000	120000	140000	10000
Variation and and and and	Contraction of the local distance	and a state of the		Versien in the second	terret and the second in		Very services
ntorobactoria nk					<		
ILEI UUALLEI A							
20000 /	40000	60000	80000	100000	120000	140000	16000
20000	40000	60000	80000	100'000	120'000	140000	16000
20000	40000	5 69000	80000	100000	120'000	140000	16000
20000	age GEC-3: 40000	60000	80000	100'000	120000	140000	16000
20000 scherichia phag	e E26	5 60000 60000	80000	100'000	120'000	140000	16000
20000 scherichia phag 20000	e E26 40000	5 50000 60000	80000 80000	100000	120000 120000 120000	140000	16000
20000 scherichia phag 20000	e E26 40000	5 50000 60000	80000 80000	100000	120'000 120'000 120'000	140000	16000 16000
20000 scherichia phag 20000	e E26	5 60000 60000	80000	100'000	120'000 120'000 120'000	140000	16000 16000
20000 scherichia phag 20000 c_Makalu_002	e E26 40000 40000	5 50000 50000	80000 80000	100'000	120000 120000 120000	140000	16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000	age GEC-33 40000 e E26 40000 40000	5 60000 60000 60000	80000 80000 80000	100'000	120000 120000 110000	140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000	age GEC-33 40000 e E26 40000 40000	5 60000 60000 60000	80000 80000 80000	100'000	120000 120000 110000 120000	140000 140000 140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000	age GEC-33 40000 40000 40000 40000	5 60000 60000 60000	80000 80000 80000	100'000	120000 120000 110000 120000	140000 140000 140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000 c_Makalu_003	age GEC-33 40000 e E26 40000 40000	5 60000 60000 60000	80000 80000 80000	100'000 100'000 100'000	120000 120000 110000 120000	140000 140000 140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000 c_Makalu_003 20000	age GEC-33 40000 e E26 40000 40000 40000	S SOÓOO SOÓOO SOÓOO SOÓOO SOÓOO	80000 80000 80000 80000	100'000 100'000 100'000	120000 120000 120000 120000 120000	140000 140000 140000 140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000 c_Makalu_003 20000	age GEC-33 40000 e E26 40000 40000 40000	S SOÓDO SOÓDO SOÓDO SOÓDO SOÓDO SOÓDO SOÓDO SOÓDO	80000 80000 80000 80000	100'000 100'000 100'000	120000 120000 120000 120000 120000	140000 140000 140000 140000	16000 16000 16000
20000 scherichia phag 20000 c_Makalu_002 20000 c_Makalu_003 20000	age GEC-33 40000 e E26 40000 40000 40000	S 60000 60000 60000 50000	80000 80000 80000 80000	100'000 100'000 100'000 100'000	120000 120000 120000 120000 120000	140000 140000 140000 140000	16000 16000 16000
20000 20000 c_Makalu_002 20000 c_Makalu_003 20000 c_Makalu_001 20001	age GEC-33 40000 e E26 40000 40000 40000	5 60000 60000 60000 60000	80000 80000 80000 80000	100'000 100'000 100'000 100'000	120000 120000 120000 120000 120000	140000 140000 140000 140000	16000 16000 16000 16000

**Fig. 4:** A comparison analysis of the genomic organization of homologous open reading frames (ORFs) of vB\_EcoM\_011D4 and other reported phages. The genome of ten phages which shows highly similarity to vB\_EcoM\_011D4 were selected and downloaded by BLASTN analysis on NCBI. The co-linearity figure was generated via Mauve 5.05. The different color represents different functional modules.

#### 3.6. DNA packing and Morphology modules

The DNA packing modules of vB\_EcoM\_011D4 consist of ORF65 (terminase small subunit ) and ORF66 (terminase large subunit), which showed more than 98% similarity to *Shigella* phage Sf20 (AUV63840.1) and *Escherichia* phage RB49 (NP\_891724.1). Additionally, there were twenty-two ORFs encoding structural proteins of the vB\_EcoM\_011D4 including head protein (ORF46, ORF61-63, ORF70-76, ORF81), and tail protein (ORF41-43, ORF57-59, ORF60-68, ORF 92 and 94).

## 3.7. Replication and Nucleotide metabolism related module

Twenty ORFs encoding phage replication-associated proteins mainly including ATP-dependent helicase (ORF83, ORF84, ORF156 and ORF200), DNA primase (ORF205, ORF211), DNA and RNA binding protein



(ORF158 and ORF221), all of them showed more than 98% similarity to reported phages. Based on these findings, the replication of vB\_EcoM\_011D4 dependent on nucleotide excision repair pathway. ORF2, ORF7, ORF9-10, ORF14, ORF37, ORF42, ORF96, ORF97, ORF100, ORF102.ORF129, ORF131, ORF135, ORF142, ORF146-149, ORF157, ORF183, ORF196, ORF203, ORF223, ORF228, ORF215, ORF217, ORF223, ORF228 and ORF248 are enzymes in the vB\_EcoM\_011D4 genome associated with nucleotide metabolism, all of them mainly similar to *Escherichia* phage E26 (MN655998.1) and *Escherichia* phage Phi1 (EF437941.1). ORF155 encoded a putative transcriptional regulator, showing 98% similarity to *Escherichia* phage E26 (QHB48837.1).

### 3.8. Lysis module

Lytic section of vB\_EcoM\_011D4 consist of ORF18, ORF48 and ORF167. ORF18 encoded a putative endolysin, which showed 97% homology to that of *Escherichia* phage vB\_EcoM\_PHB13 (MK573636.1). ORF167 encoded a putative holin, which showed 100% identity to *Escherichia* phage vB\_EcoM\_G2248 (MK327932.1). Notably, ORF48 is encoded a putative gp5 baseplate hub subunit and tail lysozyme protein, which 100% homologous to *Escherichia* phage Phi1 (EF437941.1).

### 3.9. Evolutionary relationship analysis

Using the terminase large subunit and the major capsid protein sequences to construct an evolutionary tree and analyze the evolutionary relationship between vB\_EcoM\_011D4 and other reported phages. The results show that the terminase large subunit of the vB\_EcoM\_011D4 showed less than 50% similarity to reported phage and belongs to a new branch. However, the topological structure of the major capsid protein was similar to that of the vB\_EcoM\_G2494 and belong to the same evolutionary branch (Fig. 5).



**Fig. 5:** The neighbor joining phylogenetic tree of the terminase large subunit (A) and the major capsid protein (B). Trees were constructed by MEGA 5.05 with 1000 bootstrap replicates to analysis the evolutionary relationship of vB\_EcoM\_011D4 and other reported phages.

## 4. **DISCUSSION**

In this study, we used *Escherichia coli* 011D isolated from a donkey farm in Shandong as the host to isolate an *E. coli* phage vB\_EcoM\_011D4, which was highly similar to previously reported *Enterobacteria* phage Phi1 (Capparelli et al. 2010) and RB49 (Desplats et al. 2002) and Escherichia phage Kaaroe (MN850574.1) and *Shigella* phage Sf20 (MF327006.1). Moreover, when the MOI is  $10^{-5}$ , the phage titer reaches the maximum, indicating that it has not only lytic activity but also very high activity. Additionally, vB\_EcoM\_011D4 remained stable after incubated at (4 °C~70°C) and (pH value of 4~10) for 80 min. Addition to, there is no lysogen-related gene, virulence factor or antibiotic-resistant gene were found in its genome. All of these results allowed vB\_EcoM\_011D4 to be a candidate phage for biocontrol and therapeutic application (Niu et al. 2012).

One of the remarkable features of vB\_EcoM\_011D4 is the presence of a putative gp5 baseplate hub subunit and tail lysozyme protein (ORF48), which has been shown involved in the locally digests the peptidoglycan layer upon infection, the lysozyme activity can be inhibited by spackle protein, result in phage resistant to later infection by related phages (Kanamaru et al. 2020). Furthermore, vB\_EcoM\_011D4 encoding two typical lysis



proteins, endolysin (ORF18) and holin (ORF167), which indicating holin-endolysin lysis system was used by vB\_EcoM\_011D4 to kill bacteria at first, the inner membrane was destroyed by holin and then the peptidoglycan of the out membrane were digested by lysin and tail lysozyme protein (Cahill and Young 2019).

Notably, putative DNA methylase proteins (OR147, ORF148) showed less than (57.48% ~ 67.26%) has been found in its genome, which has been found associated with single methyl group present in the viral DNA serves as a recognition site for a specific endonuclease, and responsible for the excision of the single-stranded one-genome long phage DNA, before final maturation of the phage occurs, which indicated that the different way of the DNA synthesis among vB\_EcoM\_011D4 and other reported phages (Razin et al. 1975; Friedman and Razin 1976).

The most ORFs of the vB\_EcoM\_011D4 showed highly similarity to reported phages isolated worldwide, which indicating the horizontal gene transfer may occur. Studies have revealed that there are gene transfers between phages or large phage gene clusters when infecting different bacterial species. This gene transfer between phages increases the lateral spread of these genes (Ravin et al. 2000; Desiere et al. 2001). Therefore, horizontal gene transfer between phages is the main component of phage evolution, and the diversity of phage comes from non-homologous and homologous recombination between genes (Juhala et al. 2000; Clark et al. 2001).

Studies have shown that the genetic hybridization between phages of different species and different families may occurred (Recktenwald and Schmidt 2002; Canchaya et al. 2003). As shown in Fig. 4, all selected phages in co-linearity analysis were isolated from different location of the world at different times. As shown in Fig. 5, They were distributed in different branches of the evolutionary tree in evolution, and their kinships were different. Some phages genes with the same function were found in different phage species infected with different bacteria. For example, the terminase large subunits (Fig. 5A) and major capsid proteins (Fig. 5B) of vB\_EcoM\_011D4 are closely related to the evolution of *Enterobacteria* phage Phi1 and *E. coli* phage vB\_EcoM\_G2494, which showed that the gene exchange between phages through horizontal gene transfer is an integral part of evolution. However, the exchange between the main clade and the main phage functional genes has not co-evolved with a specific host (Dekel-Bird et al. 2013).

#### Nucleotide sequence accession numbers

The accession number of the 16SrDNA sequence of the *E. coli 011D* is MN015021 in the GenBank database. The accession numbers of the vB\_EcoM\_011D4 is MT478991.1 in the GenBank database.

**Acknowledgements:** This work was financially supported through grants from the National Key Research and Development Program of China (no. 2017YFD0501000), the National Natural Science Foundation of China (no. 31572553 and 31872505), the Jilin Province Science Foundation for Youths (Changchun, China; no. 20190103106JH), the Achievement Transformation Project of the First Hospital of Jilin University (no. JDYYZH-1902025), and the Fundamental Research Funds for the Central Universities.

**Author contributions:** W.H., J.G. conceived and designed the experiments. X.W., H.Z., L.B. and H.X. performed the experiments. X.W., H.Z., L.B. and H.X. wrote the manuscript. X.W., H.Z., L.B., X.L., R.Z., Y.J, M.L., C.H. and H.O. conducted data collection and analysis and drafted the manuscript. J.D., S.R., W.H. and J.G. read and revised the manuscript.

**Conclusion:** In this study, a *E. coli* phage vB\_EcoM\_011D4 was isolated from sewage, which belongs to family *Myoviridae*. Our data show that vB\_EcoM\_011D4 displayed a rapid and strong cell lysis pattern only to *Escherichia coli spp.* Additionally, it shows good tolerance to extreme temperature and pH condition. Sequencing analysis of the phage genome is helpful for in-depth exploration of phage gene transfer, evolutionary relationship between phages and co-evolutionary relationship between phages and their host bacteria. In addition, vB\_EcoM\_011D4 has no lysogen-related genes in its genome and can be used as a candidate phage for phage therapy in the future.

## ORCID

Xinwu Wang	https://orcid.org/0000-0002-7196-3280
Hao Zhang	https://orcid.org/0000-0001-9774-808X
Lanting Bi	https://orcid.org/0000-0003-0830-006X
Zijing Wang	https://orcid.org/0000-0002-1256-6955
Yalu Ji	https://orcid.org/0000-0003-2566-2564



Xiao Liu Rihong Zhao Jianbao Dong Wei Zhu

https://orcid.org/0000-0001-9006-1437 https://orcid.org/0000-0003-0549-0865 https://orcid.org/0000-0003-0424-1525 https://orcid.org/0000-0001-8625-3751

# **REFERENCES**

- Anand T, Virmani N, Kumar S, Mohanty AK, Pavulraj S, Bera BC, Vaid RK, Ahlawat U and Tripathi BN, 2020. Phage therapy for treatment of virulent Klebsiella pneumoniae infection in a mouse model. Journal of Global Antimicrobial Resistance 21: 34–41. <u>https://doi.org/10.1016/j.jgar.2019.09.018</u>
- Ballou AL, Ali RA, Mendoza MA, Ellis JC, Hassan HM, Croom WJ and Koci MD, 2016. Development of the chick microbiome: how early exposure influences future microbial diversity. Frontiers in Veterinary Science 3: 2. https://doi.org/10.3389/fvets.2016.00002
- Bardina C, Colom J, Spricigo DA, Otero J, Sánchez-Osuna M, Cortés P and Llagostera M, 2016. Genomics of three new bacteriophages useful in the biocontrol of salmonella. Frontiers in Microbiology 7: 545. https://doi.org/10.3389/fmicb.2016.00545
- Besemer J, Lomsadze A and Borodovsky M, 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Research 29: 2607-2618. <u>https://doi/org/10.1093/nar/29.12.2607</u>
- Bettelheim KA and Lennox-King SM, 1976. The acquisition of Escherichia coli by new-born babies. Infection 4: 174-179. https://doi.org/10.1007/BF01638945
- Bloom DE, Black S, Salisbury D and Rappuoli R, 2018. Antimicrobial resistance and the role of vaccines. Proceedings of National Academy of Sciences of the United States of America 115: 12868-12871. https://doi.org/10.1073/pnas.1717157115
- Cafora M, Deflorian G, Forti F, Ferrari L, Binelli G, Briani F, Ghisotti D and Pistocchi A, 2019. Phage therapy against *Pseudomonas aeruginosa* infections in a cystic fibrosis zebrafish model. Scientific Reports 9: 1527. <u>https://doi.org/10.1038/s41598-018-37636-x</u>
- Cahill J and Young R, 2019. Phage lysis: Multiple genes for multiple barriers. Advances in Virus Research 103: 33-70. https://doi.org/10.1016/bs.aivir.2018.09.003
- Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML and Brussow H, 2003. Phage as agents of lateral gene transfer. Current Opinion in Microbiology 6: 417–424. <u>https://doi.org/10.1016/S1369-5274(03)00086-9</u>
- Capparelli R, Nocerino N, Iannaccone M, Ercolini D, Parlato M, Chiara M and Iannelli D, 2010. Bacteriophage therapy of Salmonella enterica: a fresh appraisal of bacteriophage therapy. The Journal of Infectious Diseases 201: 52–61. https://doi.org/10.1086/648478
- Chen M, Xu J, Yao H, Lu C and Zhang W, 2016. Isolation, genome sequencing and functional analysis of two T7-like coliphages of avian pathogenic Escherichia coli. Gene 582: 47-58. <u>https://doi/org/10.1016/j.gene.2016.01.049</u>
- Cheng M, Liang J, Zhang Y, Hu L, Gong P, Cai R, Zhang L, Zhang H, Ge J, Ji Y, Guo Z, Feng X, Sun C, Yang Y, Lei L, Han W and Gu J, 2017. The bacteriophage EF-P29 efficiently protects against lethal vancomycin-resistant enterococcus faecalis and alleviates gut microbiota imbalance in a murine bacteremia model. Frontiers in Microbiology 8: 837. <u>https://doi.org/10.3389/fmicb.2017.00837</u>
- Clark AJ, Inwood W, Cloutier T and Dhillon TS, 2001. Nucleotide sequence of coliphageHK620 and the evolution of lambdoid phages. Journal of Molecular Biology 311: 657–679. <u>https://doi.org/10.1006/jmbi.2001.4868</u>
- Dekel-Bird NP, Avrani S, Sabehi G, Pekarsky I, Marston MF, Kirzner S and Lindell D, 2013. Diversity and evolutionary relationships of T7-like podoviruses infecting marine cyanobacteria. Environmental Microbiology 15: 1476–1491. https://doi.org/10.1111/1462-2920.12103
- Desiere F, McShan WM, van Sinderen D, Ferretti JJ and Brussow H, 2001. Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic streptococci: evolutionary implications for prophage-host interactions. Virology 288: 325–341. <u>https://doi.org/10.1006/viro.2001.1085</u>
- Desplats C, Dez C, Tétart F, Eleaume H and Krisch HM, 2002. Snapshot of the genome of the pseudo-T-even bacteriophage RB49. Journal of Bacteriology 184: 2789–2804. <u>https://doi.org/10.1128/JB.184.10.2789-2804.2002</u>
- Dusek N, Hewitt AJ, Schmidt KN, Bergholz PW, 2018. Landscape-scale factors affecting the prevalence of *Escherichia coli* in surface soil include land cover type, edge interactions, and soil pH. Applied Environmental Microbiology 84: e02714-17. <u>https://doi.org/10.1128/AEM.02714-17</u>
- Endersen L, O'Mahony J, Hill C, Ross RP, McAuliffe O and Coffey A, 2014. Phage therapy in the food industry. Annual Reviews in Food Science and Technology 5: 327–349. <u>https://doi.org/10.1146/annurev-food-030713- 092415</u>
- Friedman J and Razin A, 1976. Studies on the biological role of DNA methylation. II. Role of phiX174 DNA methylation in the process of viral progeny DNA synthesis. Nucleic Acids Research 3: 2665–2675. https://doi.org/10.1093/nar/3.10.2665
- Grant JR and Stothard P, 2008. The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Research 36 (Web Server issue): W181-184. <u>https://doi/org/10.1093/nar/gkn179</u>
- Gu J, Liu X, Li Y, Han W, Lei L, Yang Y, Zhao H, Gao Y, Song J, Lu R, Sun C and Feng X, 2012. A method for generation phage cocktail with great therapeutic potential. PloS One 7: e31698. <u>https://doi/org/10.1371/journal.pone.0031698</u>



- Han H, Wei X, Wei Y, Zhang X, Li X, Jiang J and Wang R, 2017. Isolation, characterization, and bioinformatic analyses of lytic Salmonella enteritidis phages and tests of their antibacterial activity in food. Current Microbiology 74: 175-183. https://doi.org/10.1007/s00284-016-1169-7
- Huang C, Shi J, Ma W, Li Z, Wang J, Li J and Wang X, 2018. Isolation, characterization, and application of a novel specific Salmonella bacteriophage in different food matrices. Food Research International (Ottawa, Ont) 111: 631–641. <u>https://doi.org/10.1016/j.foodres.2018.05.071</u>
- Jang J, Hur HG, Sadowsky MJ, Byappanahalli MN, Yan T and Ishii S, 2017. Environmental *Escherichia coli*: ecology and public health implications—a review. Journal of Applied Microbiology 123: 570-581. <u>https://doi.org/10.1111/jam.13468</u>
- Juhala RJ, Ford ME, Duda RL, Youlton A, Hatfull GF and Hendrix RW, 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in thelambdoid bacteriophages. Journal of Molecular Biology 299: 27–51. https://doi.org/10.1006/jmbi.2000.3729
- Kanamaru S, Uchida K, Nemoto M, Fraser A, Arisaka F and Leiman PG, 2020. Structure and Function of the T4 Spackle Protein Gp61.3. Viruses 12: 1070. https://doi/org/10.3390/v12101070
- Kang HW, Kim JW, Jung TS and Woo GJ, 2013. wksl3, a new biocontrol agent for Salmonella enterica serovars enteritidis and typhimurium in foods: characterization, application, sequence analysis, and oral acute toxicity study. Applied Environmental Microbiology 79: 1956-1968. <u>https://doi.org/10.1128/AEM.02793-12</u>
- Kaper JB, Nataro JP and Mobley HLT, 2004. Pathogenic Escherichia coli. Nature Reviews Microbiology 2: 123–140. https://doi.org/10.1038/nrmicro818
- Kutter E, De Vos D, Gvasalia G, Alavidze Z, Gogokhia L, Kuhl S and Abedon ST, 2010. Phage therapy in clinical practice: treatment of human infections. Current Pharmaceutical Biotechnology 11: 69–86. https://doi.org/10.2174/138920110790725401
- Niu YD, Stanford K, Kropinski AM, Ackermann HW, Johnson RP, She YM, Ahmed R, Villegas A and McAllister TA, 2012. Genomic, proteomic and physiological characterization of a T5-like bacteriophage for control of Shiga toxin-producing *Escherichia coli* O157:H7. PLoS One 7: e34585. <u>https://doi.org/10.1371/journal.pone.0034585</u>
- Pajunen M, Kiljunen S and Skurnik M, 2000. Bacteriophage phiYeO3-12, specific for Yersinia enterocolitica serotype O:3, is related to coliphages T3 and T7. Journal of Bacteriology 182: 5114-5120. <u>https://doi/org/10.1128/jb.182.18.5114-5120.2000</u>
- Ravin V, Ravin N, Casjens S, Ford ME, Hatfull GF and Hendrix RW, 2000. Genomic sequence and analysis of the atypical temperate bacteriophage N15. Journal of Molecular Biology 299: 53–73. <u>https://doi.org/10.1006/jmbi.2000.3731</u>
- Razin A, Goren D and Friedman J, 1975. Studies on the biological role of DNA methylation: inhibition of methylation and maturation of the bacteriophage phichi174 by nicotinamide. Nucleic Acids Research 2: 1967–1974. https://doi.org/10.1093/nar/2.10.1967
- Recktenwald J and Schmidt H, 2002. The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage P27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. Infection and Immunity 70: 1896–1908. https://doi.org/10.1128/IAI.70.4.1896-1908.2002
- Schattner P, Brooks AN and Lowe TM, 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Research 33 (Web Server issue): W686-689. <u>https://doi/org/10.1093/nar/gki366</u>
- Sulakvelidze A, 2011. Safety by nature: potential bacteriophage applications. Microbe 6: 122–126. https://doi/org/10.1128/microbe.6.122.1