ISOLATION, CHARACTERIZATION AND GENOME ANALYSIS OF A NOVEL VIRULENT ESCHERICHIA COLI BACTERIOPHAGE VB_ECOM_011D4

Xinxu Wang 1, Hao Zhang 1, Lanting Bi 1, Hengyu Xi 1, Zijing Wang 1, Yalu Ji 1, Xiao Liu 1, Rihong Zhao 1, Ming Liu 1, Meimei Zhang 1, Chunzheng Huang 1, Hongda Ou 1, Wenyu Han 1,2, Jianbao Dong 1, Jingmin Gu 1 and Wei Zhu 1,2*

1Key Laboratory of Zoonosis Research, Ministry of Education, College of Veterinary Medicine, Jilin University, Changchun 130062, P.R. China; 2Jiangsu Co-Innovation Center for the Prevention and Control of Important Animal Infectious Disease and Zoonose, Yangzhou University, Yangzhou 225009, P.R. China; 3Department of Veterinary Medical Science, Shandong Vocational Animal Science and Veterinary College, Weifang 261061, P.R. China; 4Shanghai 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 the treatment of E. coli infection. Although the total number of phages on the earth is as high as 1011, 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 10 min, with the burst size of 115 PFU/cell. Additionally, Phage vB_Ecom_011D4 was highly stable under different temperatures (range 4 ~ 70°C) and pH conditions (range 6 ~ 10). At the same time, its genome was subjected to high-throughput 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 of the 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. There is 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, which can be further studied for elucidating the phage diversity and it is benefits for the wide application of phage therapy.

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

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 (5'TACGYYTACCTTGTTACGACTT-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°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 × g for 10 min at 4°C to remove cell debris. 10% Oleyethylene 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 analyze 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-adjusted 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, 0.0001, 0.00001, 0.000001, 0.0000001) and incubated at 37°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 1×10⁸ 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 replications.

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±0.5nm (n=5) (Fig. 1B).

![Morphological characteristics of the Escherichia coli phage vB_EcoM_011D4.](image)

Fig. 1: Morphological characteristics of the Escherichia coli phage vB_EcoM_011D4. (A) single Plaque of the phage vB_EcoM_011D4 on double-layer LB. 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_EcoM_011D4 was 10⁻⁵(Fig. 2C).

3.4. One-step growth curve of vB_EcoM_011D4

The intracellular proliferation process ofvB_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.

**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/cgview) (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).

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 (AY63840.1) and Escherichia phage RB49 (NP_891724.1). Additionally, there were 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, ORF92 and 94).

3.7. Replication and Nucleotide metabolism related module

Twenty ORFs encoding phage replication-associated proteins mainly including ATP-dependent helicase (ORF8, ORF14, 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, 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 phages 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 SI20 (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 outer 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 https://orcid.org/0000-0001-9006-1437
Rihong Zhao https://orcid.org/0000-0003-0549-0865
Jianbao Dong https://orcid.org/0000-0003-0424-1525
Wei Zhu https://orcid.org/0000-0001-8625-3751

**REFERENCES**


