Characterization and Complete Genome Analysis of a Novel *Escherichia* Phage, vB_EcoM-RPN242

Napakhwan Imklin
Silpakorn University Faculty of Engineering and Industrial Technology
Pattaraporn Sriprasong
Silpakorn University Faculty of Engineering and Industrial Technology
Narut Thanantong
Kasetsart University Kamphaeng Saen Campus
Porntippa Lekcharoensuk
Kasetsart University - Bangkhen Campus: Kasetsart University
Rujikan Nasanit (nasanit_r@su.ac.th)
Silpakorn University Faculty of Engineering and Industrial Technology https://orcid.org/0000-0002-7388-9168

Research Article

Keywords: Escherichia phage vB_EcoM-RPN242, genome analysis, novel, Agtrevirus

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1207197/v1

Abstract

The novel *Escherichia* phage vB_EcoM-RPN242 was isolated using a strain of *Escherichia coli* host originated from a diarrheal piglet. The phage was able to form plaques on the *E. coli* lawn at 15–45ºC. Moreover, it was stable over a wide pH (4–10) and temperature (4–70ºC) range. The vB_EcoM-RPN242 genome was found to be a linear, double-stranded DNA consisting of 154,840 base pairs. There were 195 protein-encoding genes and 2 tRNAs detected in the genome, however no unfavorable gene was found. According to the overall nucleotide sequence comparison, the vB_EcoM-RPN242 possibly represents a new phage species in the genus *Agtrevirus*.

Main Text

A well-known gram-negative bacterium, *Escherichia coli*, is one of the most serious problems for the swine industry. Depending on the pathotype, it can cause severe diseases such as edema disease and neonatal diarrhea. Furthermore, antibiotic resistance has become the major concern for therapeutic approaches [1]. This bacterial infection problem has a direct detrimental impact on ranchers, such as product loss, increased expenses, and health concerns. As a result, medical solutions for both prevention and therapy must be developed in order to address the issue.

Bacteriophages (phages) are naturally ubiquitous and are widely acknowledged to be the most diverse natural materials in the biosphere [2]. Generally, they are classified according to their life cycles which are chronic, lytic and lysogenic [3]. These useful viruses have been employed in various fields such as food safety, therapeutic applications, and biological studies. Focusing on therapeutic purposes, lytic phages are preferable because they strictly cause bacterial death directly. In contrast, lysogenic phages integrate their genetic elements into the host genome and, in some cases, do not immediately kill bacteria [4-6]

In this study, a novel virulent phage, vB_EcoM-RPN242, was isolated using *E. coli* M242, which originated from the intestine of a diarrheal piglet as a host. The phage was purified and propagated to investigate its morphology, efficiency of plating (EOP), thermal and pH stabilities, adsorption rate, single-step growth, and genome characteristics.

To observe phage morphology, the purified phage suspension was deposited on a formvar coated copper grid and stained with 2% uranyl acetate. The object was examined using a Hitachi HT7700 transmission electron microscope. The micrograph of the vB_EcoM-RPN242 particles revealed icosahedral heads with approximately 87±5 nm in diameter and long-contractile tails with 133±6 nm in length (Fig. 1).

To obtain information about phage characteristics, the aforementioned experiments were conducted. The EOP test was performed using the agar overlay technique. The plates were incubated at different temperatures following the method described by Seeley and Primrose [7]. The result indicated that the vB_EcoM-RPN242 was able to produce clear plaques against *E. coli* M242 at 15–45ºC. The optimal temperature for plating was 20ºC (supplementary Fig. S1). This suggests the phage is a mid-temperature phage [7]. For thermal stability, the phage was tested at 4, 28, 37, 50, 60 and 70°C for 1 h in SM buffer.

The phage was also subjected for pH stability trial. The experiment was performed in varied pH levels (2–11) of tryptic soy broth (TSB) at 37°C. The results illustrated that this phage was stable at 4–70°C and at a wide pH range (4–10) (supplementary Fig. S2a and S2b). It was unable to survive in the medium below pH 4 and above pH 10. The adsorption test was carried out following the procedure of Merabishvili and colleagues with some modification [8]. An aliquot (400 µL) of the phage-host mixture in the TSB at MOI 0.001 was taken every 3 min for 30 min. On the other hand, the control sample was collected from the phage-TSB mixture at the last minute (30 min) of the test. All samples were treated with chloroform and used for phage count. The single-step growth experiment was conducted in accordance with the protocol described by Imklin and Nasanit [9]. The results demonstrated that approximately 97% of phage particles adsorbed to host cells within 3 min. At 15 min, 99% of them attached to the cells and began lysing during 21–24 min (supplementary Fig. S3). The latent and burst periods were about 25 and 50 min, respectively. The burst size was 1074±27 PFU/infected cell (supplementary Fig. S4). These biological characteristics implied that this phage was able to lyse its host at a wide range of temperatures and survive in environmental conditions, which are favorable for therapeutic purposes [10]. Thus, vB_EcoM-RPN242 is a promising candidate for veterinary medicine application.

For the phage genome study, the phage genomic DNA was extracted by the phenol-chloroform procedure as previously described [9]. Quality control was performed on the genomic DNA before the DNA library construction. The sequencing was conducted with the Illumina HiSeq 2500 sequencing system (Macrogen, Korea). Downstream processes were carried out to validate the obtained raw reads. The reads underwent FastQC (v0.11.5) (http://www.bioinformatics.babraham. ac.uk/projects/fastqc) and Trimmomatic (v0.36) [11] to receive the qualified reads. Continuously, the filtered reads were subjected to the SPAdes algorithm [12] for *de novo* assembly to achieve a single contig of vB_EcoM-RPN242 phage genome. These processes were conducted by the Macrogen bioinformatics team.

The genome analyses began with the contig end verification using the PCR technique. The PCR reactions were performed using specifically designed primers, RPN242-F 5’-GGTGATGTTTGATGACCGACG-3’ and RPN242-R 5’-GGTAGATGCACAACAACCTGCT-3’. Open reading frames (ORFs) were predicted by Glimmer 3 [13] and GeneMarkS [14]. The translational sequences were compared to online bioinformation on BLASTp [15] conserved domain database [16] and HHpred [17] to predict their functions. tRNA genes in the phage genome were scanned using Aragorn (v1.2.41) [18]. Phylogenetic trees based on amino acid sequence alignments were constructed by MEGA 7 using the neighbor-joining (NJ) method with 1000 bootstrap replicates [19]. The whole genome sequence of the vB_EcoM-RPN242 phage was applied to PAirwise Sequence Comparison (PASC) analysis [20] to identify overall nucleotide similarity compared to its relatives. The genetically closest phage identified by PASC was selected to perform a linear comparison against the vB_EcoM-RPN242 genome and generate a genome visual representation using Easyfig [21].

The vB_EcoM-RPN242 phage has a linear, double-stranded DNA with a length of 154,840 bp and 48.8% GC content. In the genome, 195 ORFs and 2 tRNA genes were discovered. Most of them (129 ORFs) were annotated on the same strand. In contrast, the other 66 ORFs and tRNAs were located on the opposite

strand (Fig. 2). Among all ORFs, 130 genes were predicted to be hypothetical proteins, whilst 65 genes encoded functional gene products. The 65 ORFs were assigned to five modules based on their biological product properties provided by bioinformatics analysis. The group of DNA replication, recombination, repair, and packaging associated genes contained 30 ORFs, which were HNH endonuclease, DNA topoisomerase, DNA helicase, DNA ligase, RecA-like recombination protein, ssDNA binding protein, crossover junction endodeoxyribonuclease, DNA primase, endonuclease subunit, RNase H, clamp loader, sliding clamp, nuclease, DNA repair/recombination protein UvsY, homing endonuclease, terminase large subunit, terminase small subunit, and DNA polymerase. Other 14 genes encoding proteins associated with transcription, translation, and nucleotide metabolism were predicted. They were deoxycytidylate deaminase, 2-Deoxyuridine 5-triphosphate nucleotidohydrolase, thymidylate synthase, phage late-transcription coactivator, glutaredoxin, ribonucleoside-diphosphate reductase, PhoH-like protein, RNA polymerase sigma factor, nucleoside triphosphate pyrophosphohydrolase, translation repressor protein, 5'(3')-deoxyribonucleotidase, P-loop nucleotide kinase 2, and alpha-glutamyl/putrescinyl thymine pyrophosphorylase clade 2. Twelve structural coding genes of this phage were baseplate wedge subunits, tail fiber proteins, tail spike protein, neck proteins, tail sheath protein, tail tube protein, major capsid protein, and baseplate protein. There were 6 morphogenesis-related genes presented in the vB_EcoM-RPN242 genome, which were tail sheath stabilizer, portal vertex protein, prohead core scaffolding protein and protease, prohead core protein, tail completion and sheath stabilizer protein, and head completion protein. In addition, N-acetylmuramidase was mentioned as a lysozyme [22]. The baseplate hub subunit and tail lysozyme gene function as structural protein and lysozyme activity [23]. The DNA end protector protein, gp2, was identified as a protein involving host-virus interaction as its role was to prevent DNA degradation by exonuclease V [24]. Therefore, these 3 genes were ordered as genes associated with phage-host interaction and lysis activity. The BLAST search results of these proteins are presented in supplementary Table. S1. Noticeably, phages carrying virulence, antimicrobial resistance, and toxin genes are unsatisfactory for therapeutic purposes [25]. Thus, vB_EcoM-RPN242 was able to employ for veterinary work because neither the aforementioned genes nor lysogenic associated genes were captured throughout its genome.

To investigate the evolutional relationship between vB_EcoM-RPN242 and other phages, phylogenetic trees were constructed based on the alignment of amino acid sequences of the terminase large subunit (TerL) and major capsid protein (MCP). The other phage TerL and MCP sequences were retrieved from BLASTp homology search results. Both TerL and MCP trees obviously displayed the closest neighbor of vB_EcoM-RPN242, which was the *Escherichia* phage vB_EcoM-ZQ1 (supplementary Fig. S5a and S5b). Moreover, the PASC analysis acknowledged the vB_EcoM-ZQ1 phage to show the greatest similarity, with approximately 87% overall nucleotide identity, while the other 9 relatives were in a range of 79–73% similarity (supplementary Table. S2). The visualization of BLASTn-based genome comparison demonstrated the reciprocally genealogical relationship between vB_EcoM-RPN242 and vB_EcoM-ZQ1 phages with 65-100% nucleotide homology (Fig. 3). According to the National Center for Biotechnology Information (NCBI) database, the vB_EcoM-ZQ1 was unassigned to any genus as it was an unclassified *Aglimvirinae* virus. This probably implied that vB_EcoM-RPN242 could not be ranked at the genus level.

Considering the cut-off criteria for genera agreed by the International Committee on Taxonomy of Viruses (ICTV) Bacterial and Archaeal Viruses Subcommittee, phages are recognized as belonging to the same genus if their whole genome sequences share more than 70% nucleotide identity [26]. Eventually, vB_EcoM-RPN242 was classified as a member of the genus *Agtrevirus* as well as those mentioned phages in PASC result. Nevertheless, it was unassigned to any species in genus *Agtrevirus* recognized by ICTV since they are less than 95% identical. This suggests the discovery of a new virulent phage.

To evaluate whether phages are suitable for therapeutic purposes and food safety, they must be genetically characterized [27]. In addition, biological properties provide information that can be used to determine the efficacy of phages. In conclusion, vB_EcoM-RPN242 had the ability to infect *E. coli* associated with diarrheal piglets. It remained stable under detrimental conditions and lacked any undesirable genes. Thus, the vB_EcoM-RPN242 phage was qualified to be used as a biotherapeutic agent.

Nucleotide sequence accession number

The complete genome sequence of *Escherichia* phage vB_EcoM-RPN242 is available from GenBank under accession number OL656110.

Declarations

Acknowledgements

This work was supported by Kasetsart University Research and Development Institute, KURDI Grant no. FF(KU)17.64. Laurence Crouch is credited with grammar and language proofreading.

Funding

This work was supported by Kasetsart University Research and Development Institute, KURDI Grant no. FF(KU)17.64.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

The study conception, design and supervision were contributed by Rujikan Nasanit. Material preparation was performed by Napakhwan Imklin, Pattaraporn Sriprasong and Narut Thanantong. Data collection and analysis were performed by Napakhwan Imklin and Pattaraporn Sriprasong. Funding acquisition and research program management were contributed by Porntippa Lekcharoensuk. The first draft of the manuscript was written by Napakhwan Imklin. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Luppi A (2017) Swine enteric colibacillosis: diagnosis, therapy and antimicrobial resistance. Porcine Health Manag 3:16
2. Dion MB, Oechslin F, Moineau S (2020) Phage diversity, genomics and phylogeny. Nat Rev Microbiol 18:125–138
3. Chevallereau A, Pons BJ, van Houte S, Westra ER (2022) Interactions between bacterial and phage communities in natural environments. Nat Rev Microbiol 20:49–62
4. Sohail HA, Coffey A, Debrowska K, Meyer IM, Middelboe M, Sohail M, Clokie MRJ (2020) Bacteriophages: emerging applications in medicine, food, and biotechnology. Phage 1:75–82
5. Rohde C, Wittmann J (2020) Phage diversity for research and application. Antibiotics (Basel) 9:734
6. Principi N, Silvestri E, Esposito S (2019) Advantages and limitations of bacteriophages for the treatment of bacterial infections. Front Pharmacol 10:513
7. Seeley ND, Primrose SB (1980) The effect of temperature on the ecology of aquatic bacteriophages. J Gen Virol 46:87–95
8. Merabishvili M, Vandenheuvel D, Kropinski AM, Mast J, De Vos D, Verbeken G, Noben JP, Lavigne R, Vaneechoutte M, Pirnay JP (2014) Characterization of newly isolated lytic bacteriophages active against *Acinetobacter baumannii*. PLoS One 9:e104853
9. Imklin N, Nasanit R (2020) Characterization of *Salmonella* bacteriophages and their potential use in dishwashing materials. J Appl Microbiol 129:266–277
10. Loponte R, Pagnini U, Iovane G, Pisanelli G (2021) Phage Therapy in Veterinary Medicine. Antibiotics (Basel) 10:421
11. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120
12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477
13. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679
14. Besemer J, Lomsadze A, 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 Res 29:2607–2618
15. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL (2008) NCBI BLAST: a better web interface. Nucleic Acids Res 36:W5–W9

16. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226
17. Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–W248
18. Laslett D, Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16
19. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
20. Bao Y, Chetvernin V, Tatusova T (2014) Improvements to pairwise sequence comparison (PASC): a genome-based web tool for virus classification. Arch Virol 159:3293–3304
21. Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: a genome comparison visualizer. Bioinformatics 27:1009–1010
22. Broendum SS, Buckle AM, McGowan S (2018) Catalytic diversity and cell wall binding repeats in the phage-encoded endolysins. Mol Microbiol 110:879–896
23. Kanamaru S, Gassner NC, Ye N, Takeda S, Arisaka F (1999) The C-terminal fragment of the precursor tail lysozyme of bacteriophage T4 stays as a structural component of the baseplate after cleavage. J Bacteriol 181:2739–2744
24. Wang GR, Vianelli A, Goldberg EB (2000) Bacteriophage T4 self-assembly: *in vitro* reconstitution of recombinant gp2 into infectious phage. J Bacteriol 182:672-679
25. Philipson CW, Voegtly LJ, Lueder MR, Long KA, Rice GK, Frey KG, Biswas B, Cer RZ, Hamilton T, Bishop-Lilly KA (2018) Characterizing phage genomes for therapeutic applications. Viruses 10:188
26. Turner D, Kropinski AM, Adriaenssens EM (2021) A Roadmap for genome-based phage taxonomy. Viruses 13:506
27. Endersen L, O'Mahony J, Hill C, Ross RP, McAuliffe O, Coffey A (2014) Phage therapy in the food industry. Annu Rev Food Sci Technol 5:327–349

Figures

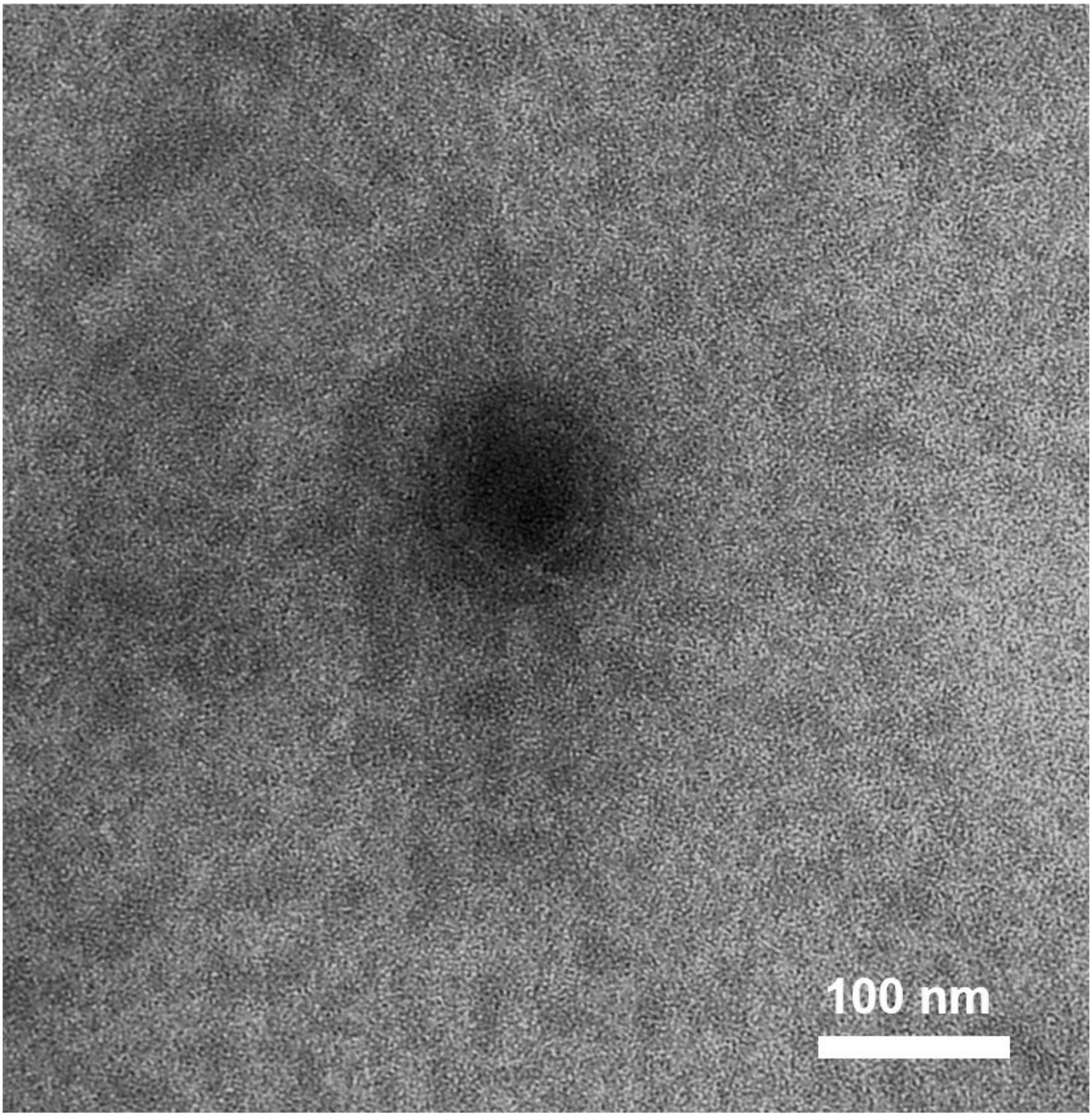

Figure 1

Morphology of vB_EcoM-RPN242

Figure 2

The annotated genome map of vB_EcoM-RPN242. The arrows indicate the transcriptional direction. The numbers represent the nucleotide positions. Each ORF is colored based on its predicted protein function. The illustration was created using Geneious Prime® version 2022.0.1.

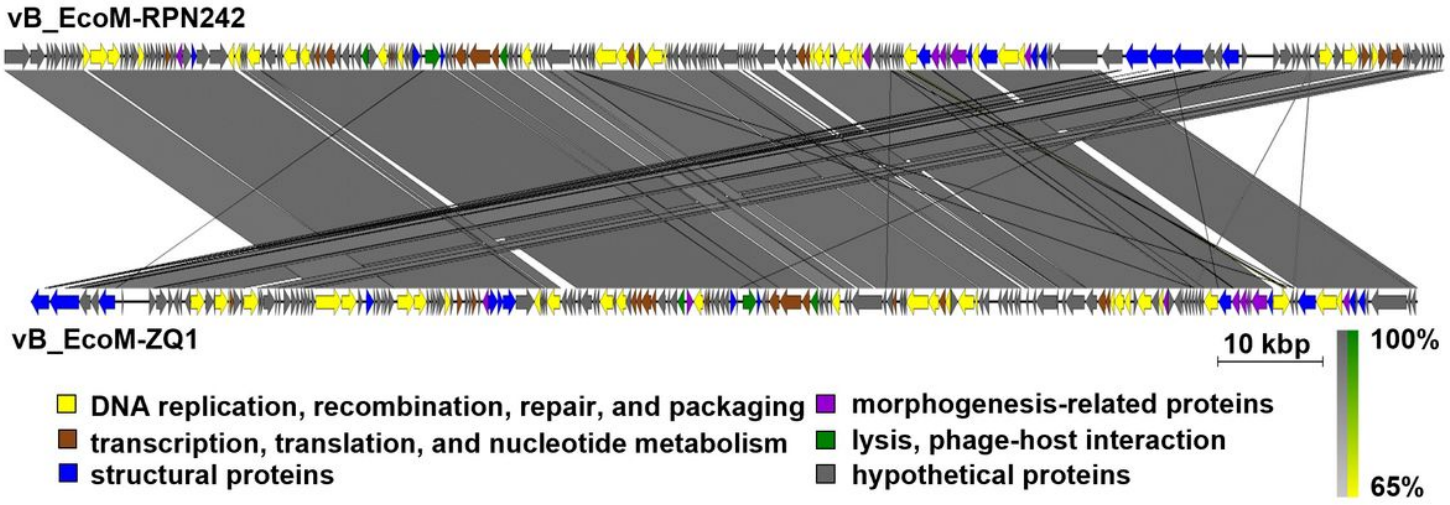

Figure 3

Comparative genome between vB_EcoM-RPN242 and *Escherichia* phage vB_EcoM-ZQ1 (MW650886). The dark-light gray and green-yellow gradient showed similarity of nucleotide sequences in the same and reverse matches, respectively. The colored arrows indicate categories of protein coding regions depending on their putative functions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ElectronicSupplementaryMaterialAchofVirol.pdf
- MCPaminoacidsequencesvBEcoMRPN242.txt
- TerLaminoacidsequencesvBEcoMRPN242.txt
- vBEcoMRPN242wholegenomesequence.fasta