doi 10.18699/bgrs2024-5.2-17

A novel environmental pseudomonas jumbo phage PseuM_296: genome analysis and putative taxonomy

Morozova V.V.[1]*, Mogileva A.A.[1, 2], Kozlova Yu.N.[1], Tikunov A.Yu.[1], Fedorets V.A.[1, 2], Yakubovskij V.I.[1, 2], Zhirakovskaya E.V.[1], Tikunova N.V.[1]

[1] *Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia*

[2] *Faculty of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia*

* *morozova@niboch.nsc.ru*

Key words: jumbo phage; Pseudomonas mandelii; phiKZ; Caudoviricetes; myovirus

Motivation and Aim: The genus Pseudomonas is a very large taxonomic group, it contains more than three hundred species. They represent a significant part of microbial communities, are both pathogens of plants and animals and their commensals. These bacteria encode a huge variety of metabolic pathways, can produce a number of organic compounds, and are able to disrupt various xenobiotics, therefore, the members of the genus are extensively studied. In contrary, environmental Pseudomonas bacteriophages have been poorly investigated, the vast majority of known Pseudomonas phages are specific to the nosocomial pathogen Pseudomonas aeruginosa. As the environmental Pseudomonas phages can affect pseudomonads, which are used in biotechnology or agriculture, so these phages need to be studied. The phage PseuM_296 and its bacterial host were isolated from the water sample from Kazanka river in Tatarstan Republic, RF. Based on 16S rRNA analysis, bacterial host strain was identified as Pseudomonas mandelii (P. fluorescens lineage) and deposited in Collection of Extremophilic Microorganisms and Type Cultures of ICBFM SB RAS as strain P. mandelii CEMTC 6978. Phage PseuM_296 genome was sequenced, analyzed and its preliminary taxonomic position was identified.

Methods and Algorithms: A negative staining electron microscopy technique was applied to examine phage PseuM_296 particles morphology. The phage suspension (109 pfu/ml) was adsorbed on a copper grid and contrasted using 1 % uranyl acetate; afterwards, the grid was examined for phage particles with a JEM 1400 transmission electron microscope (JEOL, Tokyo, Japan). To determine the sequence of the PseuM_296 genome, DNA was extracted from the phage preparation as described previously [1]. Briefly, phage particles suspension was incubated with RNase and DNase (Thermo Fisher Scientific, USA) for 1 h at 37 ºC. Then, the phage suspension was supplemented with EDTA, proteinase K (Thermo Fisher Scientific, USA) and SDS to final concentrations of 20 mM, 100–200 mkg/ml, and 0.5 %, respectively, and the mixture was incubated for 3 h at 55 °C. After that, phage DNA was purified by phenol/chloroform extraction and subsequent ethanol precipitation. Paired-end library of the PseuM_296 genome was done using the Nextera DNA Sample Preparation Kit (Illumina, Inc, USA). Sequencing was carried out using the MiSeq Benchtop Sequencer and MiSeq Reagent Kit v.1 (2 × 250 base reads) (Illumina Inc, USA). Genome sequence was de novo assembled using the SPAdes genome assembler v.3.15.2 (http://cab.spbu.ru/software/spades). Rapid Annotation Subsystem Technology (RAST) v.2.0 (https://rast.nmpdr.org) was used to predict putative open reading frames (ORFs). Next, obtained ORFs were checked manually against the NCBI GenBank protein

database (https://www.ncbi.nlm.nih.gov) using BLASTX and DELTA-BLAST algorithms. In addition, InterProScan and HHPred software [2, 3] were used to analyze ORFs encoding proteins without homology with the sequences deposited in the GenBank database. A comparative proteomic phylogenetic analysis was performed using Viral Proteomic tree server (ViPTree server) (https://www.genome.jp/viptree). Sequence identity matrix was calculated using Virus Intergenomic Distance Calculator, VIRIDIC (http://rhea.icbm.uni-oldenburg.de/VIRIDIC).

Results: Electron microscopy revealed that the PseuM_296 particle consists of a large capsid (Ø130 nm) connected to a long contractile tail (L = 230 nm). Therefore, the virion morphology corresponds to the myovirus morphotype. The PseuM_296 genomic characteristics were as follows: the genome length was 273 174 bp; it contained 332 putative genes and three of them correspond to tRNAs. Only eighty-one genes encode proteins with predicted functions that were determined based on their amino acid sequences and domain structure similarity. The remaining 248 genes were defined as hypothetical. Putative multisubunit RNAPs (both virion RNA polymerase and nonvirion RNA polymerase), SbcCD complex ATPase, DnaB-like replicative helicase, RNA helicase, and family B DNA polymerase were revealed in the PseuM_296 genome. These genes have been defined previously as the core ones specific to the phiKZ-like group of jumbo phages [4, 5]. In addition, genes encoding chimallin (a major component of the nucleus-like compartment) and PhuZ (tubulin-like protein) were detected, which provide a unique lifestyle of jumbo phages. Phylogenetic proteomic analysis confirmed that phage PseuM_296 belongs to phiKZ-like phages and the most similar phage was Pseudomonas phage pPa_SNUABM_DT01 (Fig. 1).

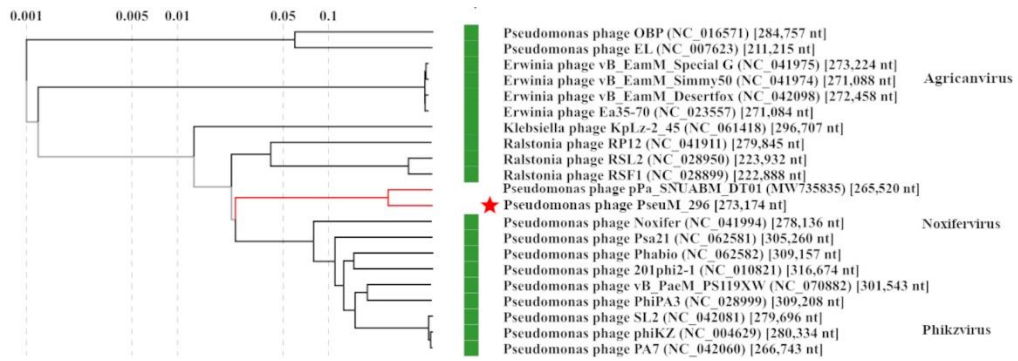

Fig. 1. Phylogenetic proteomic analysis of the jumbo phage PseuM_296. PseuM_296 genome was marked with asterisk. Group of pPa_SNUABM_DT01 and PseuM_296 phages is marked with red

The analysis of intergenomic similarity for PseuM_296 and fourteen most similar phages was carried out using VIRIDIC tool. It has been shown that pPa_SNUABM_DT01 and PseuM_296 genomes have a limited level of similarity (52.4 %), this level is significantly lower than the threshold value for grouping phages into one genus (70 %). Therefore, these phages may be prototype species for two new genera among phiKZ-like phages (Fig. 2).

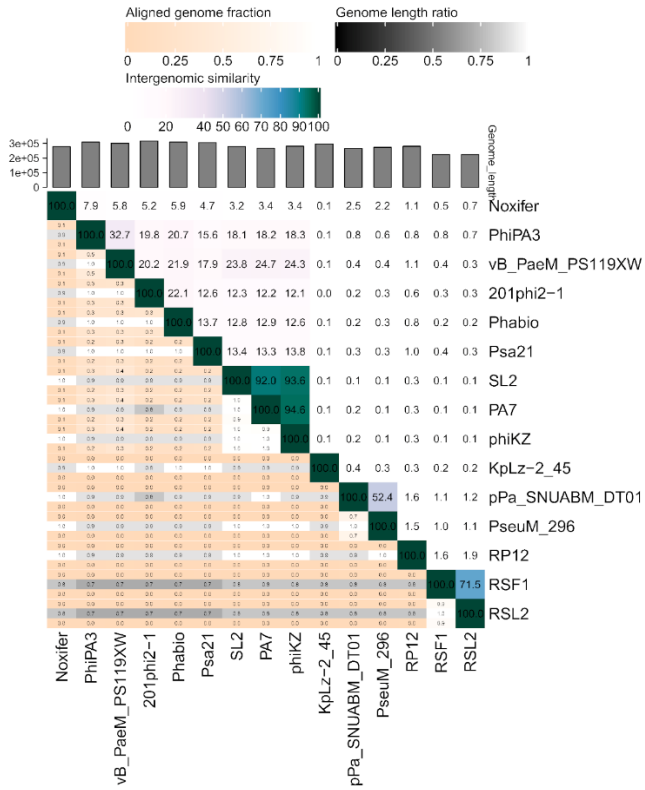

Fig. 2. An intergenomic similarity matrix was calculated for the studied phage PseuM_296 and the phiKZ-like phages closest to it

Conclusion: A novel jumbo phage, PseuM_296, has been isolated, specific to the *P.mandelii* inhabiting in the environment. The particles of phage PseuM_296 had the morphotype of myovirus. Analysis of the PseuM_296 genome has shown that PseuM_296 is a new representative of the phiKZ-like group of giant phages and may be a prototype species of a new genus.

Funding: Funding: This study was supported by the Ministry of Science and Higher Education of the Russian Federation under project grant No. 075-15-2021-1085.

References

1. O'Flaherty S., Coffey A., Edwards R. et al. Genome of staphylococcal phage K: a new lineage of Myoviridae infecting gram-positive bacteria with a low GC content. *J Bacteriol.* 2004;186:2862-2871
2. Quevillon E., Silventoinen V., Pillai S., Harte N., Mulder N., Apweiler R., Lopez R. InterProScan: protein domains identifier. *Nucleic Acids Res.* 2005;33(Web Server issue):W116-W120. doi 10.1093/nar/gki442
3. Söding J., Biegert A., Lupas A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005;33(Web Server issue):W244-W248. doi 10.1093/nar/gki408
4. Sokolova M.L., Misovetc I.V., Severinov K. Multisubunit RNA polymerases of Jumbo bacteriophages. *Viruses.* 2020;12(10):1064. doi 10.3390/v12101064
5. Lavysh D., Sokolova M., Minakhin L., Yakunina M., Artamonova T., Kozyavkin S., Makarova K.S., Koonin E.V., Severinov K. The genome of AR9, a giant transducing *Bacillus* phage encoding two multisubunit RNA polymerases. *Virology.* 2016;495:185-196. doi 10.1016/j.virol.2016.04.030