The modern personalized approach to phage therapy is based on the detailed assessment of the interaction between phages and bacterial cells. Bacterial carbohydrates exposed on the cell surface, O-antigens and capsular polysaccharides, are one of the most important specificity determinants in the phage–cell interaction. Capsular polysaccharides of K. pneumoniae, being the virulence factors [1], are highly diverse in their structure. Currently, bioinformatics databases indicate the existence of at least 134 genetic variants [2]. To meet the challenges of clinical practice it is necessary to create the collection of phages with...
The KPPK108.1 bacteriophage forms clear plaques, 5 mm in diameter, surrounded by the translucent halos, in the bacterial cultures grown on the agar plates (Fig. 1). The presence of a halo typically indicate the presence of phage-derived depolymerase, which has been confirmed by further research. The one-step growth curve showed a latent period of 15 min and burst size of 46 phage particles per one infected cell.

**RESULTS**

Genome sequences of bacteriophages to be used for comparison with the KPPK108.1 phage were downloaded from the NCBI Genome database [11]. The average nucleotide identity was calculated using the VIRIDIC online service [12] and the orthoANIu software [13]. Phylogenetic analysis was performed by the maximum likelihood method implemented in the RaxML program [14] with the use of the GAMMA LG amino acid substitution model [15] and the concatenated amino acid sequence alignment of the major capsid protein, terminase large subunit, DNA polymerase, and RNA polymerase. The sequences were aligned with the MAFFT program [16] and concatenated with the Geneious Prime software [10]. The intergenomic comparison diagram was created with the Easyfig application [17] using the TBLASTX tool [7] to find the homologous regions within genomes.

The models of the gene 8 product tertiary structure and the tailspike protein quaternary structure for the *Klebsiella* bacteriophage KPPK108.1 were constructed with the AlphaFold-Multimer application [18, 19]. The tailspike protein structure of the *Enterobacteria* phage ~92 was downloaded from the PDB database [20]. Electrostatic surface charge of the tailspike protein was calculated using the APBS program [21]. The UCSF Chimera program was used for structure alignment and visualization [22].

**METHODS**

The P224 (1732) and P225 (1333) clinical strains of *K. pneumoniae* with the K108 type capsular polysaccharide were obtained from the collection of the Institute of Epidemiology (Moscow, Russia). The wastewater samples, collected from the wastewater treatment facilities in Moscow, were used for bacteriophage isolation. Dry components of the bacterial culture medium (trypton, 10 g/L, yeast extract, 5 g/L, NaCl, 5 g/L) were added to the wastewater samples previously clarified by centrifugation, then the media were inoculated with the bacterial cells culture being in the phase of exponential growth. Cultivation was carried out at 37 °C for 16 h. The bacterial culture was subsequently inactivated with chloroform, and the samples were clarified by centrifugation. Phages were detected by titration using the double-layer agar plate method. The isolated phage was titrated twice in a row in order to obtain single phage plaques. Preparative bacteriophage growth was performed in 1 L of the P224 strain culture at 37 °C. Bacteriophage was precipitated with polyethylene glycol and purified by caesium chloride density gradient ultracentrifugation [5].

Genomic DNA of the phage was extracted from the purified phage preparation by incubation with the solution, containing 100 mM Tris-HCl (pH 7.5), 25 mM EDTA, 1.5 M NaCl, 2% (w/v) CTAB buffer, 0.3% (v/v) β- mercaptoethanol, and 50 mg/mL of proteinase K, at 50 °C for 30 min, with subsequent chloroform DNA extraction, and precipitated by adding 0.6 volume of isopropyl alcohol. Genome sequencing was performed on the MiSeq platform using the Nextera DNA library preparation kit (Illumina; USA). A single contig was assembled from the resulting sequences using v. 3.13 of the SPAdes software [4].

The experiment aimed at assessing the phage particle production was performed in accordance with the previously reported protocol [5]. Negative contrast electron microscopy was utilized to visualize the phage particles. Phage preparation with a volume of 3 μL was applied to a carbon-coated 400 mesh grid. The negatively contrasted preparation was obtained by 1% uranyl acetate staining for 30 s. Imaging was performed with the JEOL JEM-2100 200 kV transmission electron microscope (JEOL USA Inc.; USA) at 30,000x magnification.

The Klebsiella bacteriophage KPPK108.1 was annotated with the Prokka tool [6] using the embedded databases. The functions of the genome protein products were predicted with the BLAST search tool [7] based on the known homologs, and by the HMM-HMM comparison, performed with the HHpred [8] and Phyre2 [9] web-based tools using the SCOPe70_2.07, ECOD_ECOD_F70, and UniProt-SwissProt-viral70 databases. E value < 10⁻⁵ was used as a criterion of significant similarity in BLAST analysis; the Phyre2 “confidence” and HHpred “probability” values exceeding 95% were used as the criteria of significant similarity for the HMM-HMM comparison. Genetic mapping was carried out with the Geneious Prime software [10].

![Electron microscopy of Klebsiella bacteriophage KPPK108.1](image-url)
CP003200.1). The search for coding sequences revealed a total of 56 protein-coding genes and no tRNA-encoding genes in the genome (Fig. 2). The search for homologous and similar sequences using the BLAST algorithm and the HMM-HMM comparison in public databases and web servers made it possible to predict the functions of 29 proteins, encoded in the genome. It was not possible to identify the functions of 27 proteins. No genes, encoding integrases or other proteins specific to temperate phages, were found in the genome.

Comparison of the average nucleotide identity (ANI), involving all 14,923 genes of tailed bacteriophages, deposited in the NCBI Genome database, revealed a group of Klebsiella bacteriophages of the genus Drulisvirus, being the most close to phage KPPK108.1 based on this parameter (Fig. 3). The ANI values of phage KPPK108.1 and a typical phage of the genus Drulisvirus, Klebsiella phage KP34, are 73.0%. Phylogenetic analysis, performed with the use of the concatenated amino acid sequences of the major capsid protein, large terminase subunit, DNA polymerase, and RNA polymerase, shows that Drulisvirus bacteriophages and KPPK108.1 phage form a monophyletic group (Fig. 4). The genetic makeup and genomic organization of the phage KPPK108.1 are generally similar to those of the other Autographiviridae phages (Fig. 5), and are almost identical to those of other members of the genus Drulisvirus. An interesting feature of the gene 8 product was found. Protein structural modeling revealed unusual L-shape of the protein with a tubular C-terminal region (Fig. 6.1). This region was characterized by the number of positively charged amino acid residues above the average. Electrostatic field simulation showed that the C-terminal region of the gene 8 product had a significant negative surface charge (up to ~5) (Fig. 6.2).

Bioinformatic analysis of the KPPK108.1 phage genome revealed genes, encoding the head-tail connector and tailspike proteins. Modeling and analysis of the tailspike protein structure was performed (Fig. 6). The search for similar structures revealed a high degree of similarity between the tailspike of the phage KPPK108.1 and the tailspike of the Enterobacteria phage -92 (PDB entry 6E0V) (Fig. 6) exhibiting colanidase activity confirmed by experimental data [23].

DISCUSSION

The genome-wide similarity score of the phages KPPK108.1 and KP34 exceeding 70% of the genus boundary, together with the results of the phylogenetic analysis performed based on the concatenated sequences of conservative genes, show that the Klebsiella phage KPPK108.1 belongs to the genus Drulisvirus, subfamily Slopekvirinae, family Autographiviridae. Intergenomic comparisons support this finding. Minor differences in genome organization can be explained by the recombination events that took place during the Klebsiella bacteriophages’ evolution, as confirmed by the presence of NHN endonuclease genes in the genomes of KPPK108.1 and other related bacteriophages. The genome structure of the KPPK108.1 phage is typical for bacteriophages of the Autographivirinae family and is characterized by the presence of the early gene region [24], comprising gene B, encoding a hypothetical protein with an unusual L-shaped tertiary structure. Regardless of the fact that

---

Fig. 2. Genetic map of Klebsiella bacteriophage KPPK108.1. Genes are colored in accordance with the functions of their products (see caption). Arrows indicate gene directions in accordance with their encoding functions.
Fig. 3. Average nucleotide identity (ANI) distance matrix of the *Klebsiella* phage KPPK108.1 generated with the VIRIDIC web-based tool using the genomes of various *Autographiviridae* family members.

the function of this protein has not been defined by searching for homologues using BLAST, or searching for similar proteins by HMM-HMM comparison, the surface charge distribution makes it possible to assume that this protein mimics nucleic acid, like Ocr proteins, which are also located within the early gene regions in the genomes of other *Autographiviridae*, and are capable of DNA mimicking [25, 26]. It has been shown that Ocr protein effectively inhibits the BREX restriction modification system, facilitating phage infection [27].

Bioinformatic analysis of the KPPK108.1 phage genome makes it possible to predict the organizational structure of the adsorption apparatus comprised of the head-tail connector and the tailspike protein, possessing enzymatic properties. The tailspike protein seems to be the receptor-binding protein (RBP), which determines host specificity and the host spectrum of the phage [28]. The tailspike protein structure analysis indicates the presence of depolymerizing activity against the polysaccharide, presumed to be related to the *E. coli* colanic acid. Colanic acid, the extracellular polysaccharide, consisting of several types of carbohydrate residues (such as L-fucose, D-glucose, D-galactose, and D-glucuronic acid), which is released into the extracellular environment by bacteria of the *Enterobacteriaceae* family, is the colanidase substrate [29]. Colanidases have been relatively recently discovered in phage RBPs [30]. Colanidases are present in a number of virulent bacteriophages of the evolutionarily distant groups, such as podoviruses and myoviruses [23, 30], some of which have proven to be effective when used in phage cocktails for phage therapy [30]. It is essential to define the structure of the type K108 *K. pneumoniae* capsular polysaccharide to clarify the question of the similarity of this polymer to colanic acid.

**CONCLUSIONS**

*Klebsiella* bacteriophage KPPK108.1 is a virulent bacteriophage of the genus *Drulisvirus*, family *Autographiviridae*. Thorough
Bioinformatic analysis has revealed the lytic nature of the phage infection cycle. The analysis makes it possible to predict the structure of the phage adsorption apparatus comprised of the head-tail connector and the tailspike protein, exhibiting colanidase activity. The predicted characteristics of KPPK108.1 bacteriophage indicate the feasibility of using KPPK108.1 in phage cocktails for phage therapy. As far as we know, KPPK108.1 is the first fully described phage specific for capsular type KL108.
Fig. 5. Intergenomic comparison diagram created with EasyFig and TBLASTX using the genomes of Klebsiella phage KPPK108.1 and other Autographiviridae phages. The vertical lines are colored in accordance with the color scale showing the degree of similarity.

Fig. 6. Predicted tertiary structure of the KPPK108.1 phage gene 8 product painted with rainbow colors, where blue indicates N-terminal region, and red indicates C-terminal region of the protein (1). Predicted tertiary structure of the gene 8 product painted in accordance with the charge of the protein surface electrostatic field (2). Predicted tertiary structure of the KPPK108.1 phage tailspike trimer painted with rainbow colors, where blue indicates N-terminal region, and red indicates C-terminal region of the protein, longitudinal view (3). Predicted tertiary structure of the KPPK108.1 phage tailspike trimer with monomers painted with different colors, view along transverse axis (4).
References

1. Paczosa MK, Mecasias J. Klebsiella pneumoniae: Going on the Offensive with a Strong Defense. Microbiol Mol Biol Rev. 2016; 80 (3): 629–81.
2. Wyres KL, Wick RR, Gorrie C, Jenney A, Follador R, Thomson NR, et al. Identification of Klebsiella capsule synthesis loci from whole genome data. Microb Genom. 2016; 2 (12): e000102.
3. Colombo J, Robin A, Lavie L, Bettarel Y, Cauchie HM, Sime-Ngando T. Viroplastkaton “pegylation”: use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic ecosystems. J Microbiol Methods. 2007; 71 (3): 212–9.
4. Bankevich A, Nurf S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012; 19 (5): 455–77.
5. Druis-Kawa Z, Mackiewicz P, Kęsik-Szeloch A, Maciaszczyk-Dziubinska E, Weber-Dąbrowska B, Dorotkiewicz-Jach A, et al. Isolation and characterisation of KP34 — a novel ΔkMV-like bacteriophage for Klebsiella pneumoniae. Appl Microbiol Biotechnol. 2011; 90 (4): 1333–45.
6. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014; 30 (14): 2068–9.
7. Achtshul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215 (3): 403–10.
8. Gabler F, Nam S-Z, Till S, Mirdita M, Steinegger M, Söding J, et al. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Current Protocols in Bioinformatics. 2020; 72 (1): e108.
9. Kelley LA, Mezulis S, Yates CM, Waas MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015; 10 (6): 845–58.
10. Geneious. Bioinformatics Software for Sequence Data Analysis. Geneious [cited 2021 Nov 11]. Available from: https://www.geneious.com/.
11. Home — Genome — NCBI [cited 2021 Nov 11]. Available from: https://www.ncbi.nlm.nih.gov/genome.
12. Berman HM, Westbrook J, Feng Z, Gilland G, Bhat TN, Weisg H, et al. The Protein Data Bank. Nucleic Acids Research. 2000; 28 (1): 235–42.
13. Jurrus E, Engel D, Star K, Monson K, Brand J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. Protein Science. 2018; 27 (1): 112–28.
14. Schwarzer D, Buettner FFR, Browning C, Nazarov S, Rabsch W, Bethe A, et al. A Multivalent Adsorption Apparatus Enables the Broad Host Range of Phage phi62: a Comprehensive Genomic Atlas. Structural Analysis. Journal of Virology. 2012; 86 (19): 10384–98.
15. Zavilgelsky GB, Rastorguev SM. Antirestriction proteins ArDA and Ocr as efficient inhibitors of type I restriction-modification enzymes. Mol Biol. 2009; 43 (2): 241.
16. Isaev A, Drobizko A, Sierro N, Gordieva J, Yosef I, Qimron U, et al. Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defense. Nucleic Acids Res. 2020; 48 (10): 5397–406.
17. Zampara A, Sørensen MCH, Grimon D, Antenucci F, Vitt AR, Bortolaia V, et al. Exploring phage receptor binding proteins to enable endolysins to kill Gram-negative bacteria. Sci Rep. 2020; 10 (1): 12087.
18. Stevenson G, Andrianopoulos K, Hobbs M, Reeves PR. Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol. 1996; 178 (16): 4885–93.
19. Kim H, Kim M, Bai J, Lim J-A, Heu S, Ryu S. Colanic Acid Is a Novel Phage Receptor of Pectobacterium carotovorum subsp. carotovorum Phage PCP72. Frontiers in Microbiology. 2019; 10: 143.
12. Moraru C, Varsani A, Kropinski AM. VIRIDIC — A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. Viruses. 2020; 12 (11): 1268.

13. Lee I, Ouk Kim Y, Park S-C, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol. 2016; 66 (2): 1100–3.

14. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30 (9): 1312–3.

15. Le SQ, Gauczuk O. An improved general amino acid replacement matrix. Mol Biol Evol. 2006; 25 (7): 1307–20.

16. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research. 2002; 30 (14): 3059–66.

17. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011; 27 (7): 1009–10.

18. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ponneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596 (7873): 583–9.

19. Evans R, O’Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. 2021 [cited 2021 Dec 6]. 2021.10.04.463034. Available from: https://www.biorxiv.org/content/10.1101/2021.10.04.463034v1.

20. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Research. 2000; 28 (1): 235–42.

21. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. Protein Science. 2018; 27 (1): 112–28.

22. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25 (13): 1605–12.

23. Schwarzer D, Buettner FFR, Browning C, Nazarov S, Rabisch W, Bethe A, et al. A Multivalent Adsorption Apparatus Explains the Broad Host Range of Phage phi92: a Comprehensive Genomic and Structural Analysis. Journal of Virology. 2012; 86 (19): 10384–98.

24. Evseev PV, Lukianova AA, Shneider MM, Korzhenkov AA, Bugaeva EN, Kabanova AP, et al. Origin and Evolution of Studiervirinae Bacteriophages Infecting Pectobacterium: Horizontal Transfer Assists Adaptation to New Niches. Microorganisms. 2020; 8 (11): 1707.

25. Roberts GA, Stephanie AS, Kanwar N, Dawson A, Cooper LP, Chen K, et al. Exploring the DNA mimicry of the Ocr protein of phage T7. Nucleic Acids Res. 2012; 40 (16): 8129–43.

26. Zavilgelsky GB, Rastorguev SM. Antirestriction proteins ArdA and Ocr as efficient inhibitors of type I restriction-modification enzymes. Mol Biol. 2009; 43 (2): 241.

27. Isaev A, Drobiazko A, Siemo N, Gordeeva J, Yosef I, Qimron U, et al. Phage T7 DNA mimic protein Ocr is a potent inhibitor of BIREX defence. Nucleic Acids Res. 2020; 48 (10): 5397–406.

28. Zampara A, Serensen MCH, Grimol D, Antenucci F, Vitt AR, Bortolaia V, et al. Exploiting phage receptor binding proteins to enable endolysins to kill Gram-negative bacteria. Sci Rep. 2020; 10 (1): 12087.

29. Stevenson G, Andrianopoulos K, Hobbs M, Reeves PR, Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol. 1996; 178 (16): 4885–93.

30. Kim H, Kim M, Bai J, Lim J-A, Heu S, Ryu S. Colanic Acid Is a Novel Phage Receptor of Pectobacterium carotovorum subsp. carotovorum Phage POP72. Frontiers in Microbiology. 2019; 10: 143.