

Complete genome sequence of a clinical isolate of *Clostridioides difficile* bacteriophage phiCDKH01 of the family Siphoviridae

Krzysztof Hinc (✉ krzysztof.hinc@gumed.edu.pl)

Intercollegiate Faculty of Biotechnology of the University of Gdańsk and the Medical University of Gdańsk: Miedzyuczelniany Wydział Biotechnologii Uniwersytetu Gdanskiego i Gdanskiego Uniwersytetu Medycznego <https://orcid.org/0000-0002-6067-7852>

Monika Kabała

Medical University of Silesia: Slaski Uniwersytet Medyczny w Katowicach

Adam Iwanicki

Intercollegiate Faculty of Biotechnology of the University of Gdańsk and the Medical University of Gdańsk: Miedzyuczelniany Wydział Biotechnologii Uniwersytetu Gdanskiego i Gdanskiego Uniwersytetu Medycznego

Gajane Martirosian

Medical University of Silesia: Slaski Uniwersytet Medyczny w Katowicach

Alessandro Negri

Intercollegiate Faculty of Biotechnology of the University of Gdańsk and the Medical University of Gdańsk: Miedzyuczelniany Wydział Biotechnologii Uniwersytetu Gdanskiego i Gdanskiego Uniwersytetu Medycznego

Michał Obuchowski

Intercollegiate Faculty of Biotechnology of the University of Gdańsk and the Medical University of Gdańsk: Miedzyuczelniany Wydział Biotechnologii Uniwersytetu Gdanskiego i Gdanskiego Uniwersytetu Medycznego

Research Article

Keywords: Novel Siphophage, Genomic and Phylogenetic Analysis, Genetic Diversity, Tailed Phages

Posted Date: February 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-186488/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

A new temperate phiCDKH01 siphophage was obtained from clinical isolate of *Clostridioides difficile*. The phage genome is a 45,089 bp linear double-stranded DNA molecule with an average G + C content of 28.7%. It shows low similarity to known phage genomes except for phiCD24-1. Genomic and phylogenetic analysis revealed that phiCDKH01 is a novel phage. 66 putative ORFs were predicted in the genome, 37 of which code for proteins with predicted functions. The phiCDKH01 prophage has been localized in the host genome. Results of this study increases genetic diversity of known tailed phages.

Introduction

Clostridioides difficile is a pathogen with great epidemiological potential and a serious threat to human health [1]. In the CDC's latest report on the risk of drug resistance, *C. difficile* was classified as the leading cause of nosocomial infections [2]. *C. difficile* infection (CDI) is closely related to the weakening of the function of the intestinal microbiome as a side effect of antibiotic therapy [3, 4]. CDI clinical picture is complex and most often manifested with mild, moderate, or severe diarrhea. The development of CDI infection can turn into life-threatening pseudomembranous colitis or toxic megacolon [5, 6, 7]. Currently, acute *C. difficile* infection is treated with antibiotics, i.e. metronidazole, vancomycin and fidaxomicin [8]. The use of antibiotics in the treatment of CDI increases the risk of exacerbation of microflora dysbiosis causing a reduction or removal of normal intestinal commensals. Consequently, *C. difficile* may colonize this niche [9]. Moreover, in the case of this infection, antibiotic therapy promotes the recurrence of the disease and increases the chance of emergence of antibiotic resistance [10].

In the last decade, interest in bacteriophages that infect the pathogenic *C. difficile* has increased due to their possible contribution to virulence, host biology and their potential as alternative therapeutic agents [11]. All so far described phages infecting *C. difficile* are temperate. In most cases they have been isolated from bacterial cells after induction of prophages [12, 13, 14]. Described *C. difficile* phages belong to the Myoviridae or Siphoviridae families of the order Caudovirales, i.e. phages with contractile or non-contractile tails, respectively [12, 15]. Myoviridae phages are the most numerous and their genomes show significant DNA homology with a tendency to create phylogenetically related clusters. On the contrary, a limited number of Siphoviridae phages have been described and sequenced and these phages have been shown to be genetically more distant [16].

In the current study, a newly discovered phage named phiCDKH01 was isolated and characterized. The phage genome was sequenced, annotated and phylogenetic analysis indicated that phiCDKH01 is a member of the Siphoviridae and might belong to a new phage lineage. We also showed the location of the newly discovered siphophage in the genome of its host.

Results

Bacterial strain isolation

Clostridioides difficile CD34-Sr was isolated from the hospital environment, in a 600-bed clinical hospital of Medical University of Silesia, Katowice, Poland. The strain comes from a Nephrology Ward, from the bed frame in one of patients' room. The material was collected using a selective broth enabling *C. difficile* spores' germination: C diff Banana Broth (Hardy Diagnostics, Santa Maria USA). After incubation one loop broth was replated on selective *C. difficile* media chromID *C. difficile* (bioMérieux, Marcy L'Etoile, France) and incubated for 48 hours under anaerobic conditions. Colonies with a characteristic horse odor and yellow-green fluorescence under UV light, microscopically recognized as a cylindrical Gram-positive bacilli, were identified in automatic system - VITEK 2 Compact (bioMérieux, Marcy L'Etoile, France) as *C. difficile*.

Prophage induction and phage isolation

To determine if strain CD34-Sr contained a functional prophage, we used the mitomycin C high-throughput induction method described previously [14]. In this method, the inducible phage DNA in the heated lysate was PCR-confirmed using specific phage primers targeting-holin genes of myovirus and siphovirus [17]. The results confirmed that the amplified PCR product was holin gene of siphovirus. Finally, the strain of *C. difficile* CD34-Sr containing inducible temperate phage, was chosen for large scale phage induction. Mitomycin C induction was performed on 500 ml of log phase bacteria cultured in BHI broth (Sigma-Aldrich, USA). Following the overnight incubation, phage lysate was collected, filtered, and concentrated using PEG precipitation [18]. We analyzed the concentrated phages lysate under the electron microscope and found only one type of phage particle. Phage fraction was then further purified using CsCl gradient as previously described [18]. The isolated phage was named after its discoverer's initials phiCDKH01.

Genomes sequencing and annotation

Phage genomic DNA was purified using a Phage DNA Isolation Kit (Norgen Biotek Corp., Canada) following the manufacturer's instructions. Whole-genome sequencing was performed by the Genomed S.A. (Poland) on the Illumina MiSeq platform with 764-fold coverage. High-quality paired-end reads were assembled *de novo* using SPAdes v. 3.13.0 (<https://github.com/ablab/spades>). Obtained consensus sequence was annotated with myRAST v. 36 (<https://rast.nmpdr.org/>) [20] and deposited in the GenBank under accession number MN718463.

The genomic features of phiCDKH01

The genome of phage phiCDKH01 is 45,089 bp in length with a G+C content of 28.7%, similar to that of its host *C. difficile*. In the initial annotation a total of 66 ORFs were identified as probable protein-coding genes. 53 were located on the positive strand, while only 13 ORFs were located on the negative strand. No rRNA or tRNA genes were identified. Thirty-seven genes were assigned a predicted function. The complete phage genome could be divided into functional clusters that encode proteins involved in DNA packaging, head and tail morphogenesis, host cell lysis, and replication (Fig 1). We identified genes for the terminase large subunit (phiCDKH01_44), terminase small subunit (phiCDKH01_43), tail tape-measure protein

(phiCDKH01_43), two tail family proteins (phiCDKH01_62/63), pre-neck appendage-like protein (phiCDKH01_65), portal protein (phiCDKH01_45), scaffolding protein (phiCDKH01_51) and capsid protein (phiCDKH01_52).

Additionally, we detected genes encoding proteins whose presence confirms the temperate nature of phiCDKH01, including a recombinase (phiCDKH01_31), integrase (phiCDKH01_12), antirepressors (phiCDKH01_20/24), and 5 putative transcriptional regulators (phiCDKH01_07/17/19/22/27), suggesting that the prophage could affect some bacterial functions.

We identified the gene cluster for host cell lysis containing an *N*-acetylmuramoyl-L-alanine amidase (phiCDKH01_06), putative holin protein (phiCDKH01_05) and ImmA/IrrE family metallo-endopeptidase (phiCDKH01_13).

We also found genes involved in DNA replication encoding: a dnaD domain protein (phiCDKH01_32), single-stranded DNA-binding protein (phiCDKH01_33) and two putative PcmI proteins (phiCDKH01_10/42). These proteins have been shown to be essential for the autonomous replication of natural plasmids with a low copy number, *i.e.*, R100 [21].

Finally, we identified several additional interesting genes that encode proteins with different functions *e.g.*, an ADP-ribosyltransferase exoenzyme family protein (phiCDKH01_48) that might covalently modify cell actin to modify physiology of eukaryotic cells, similarly to *Clostridium botulinum* C2 or *Clostridium perfringens* E iota toxins do [22]. Gene coding for a putative lipoprotein (phiCDKH01_60) might play a role in cortex modification and thus spore germination [23]. Another one is the HicB antitoxin (phiCDKH01_23) that is a member of a type II toxin-antitoxin system family found in bacteria and archaea and has been shown to be involved in the stress response, virulence and persistence [24] (Tab.S1).

Among other interesting features of the phiCDKH01 genome, a putative CRISPR (clustered regularly interspaced short palindromic repeats) and a nearby CRISPR array comprising 5 spacers of 35, 36 or 37 bp (Fig 1, Tab.S2) were identified. Analysis of the CRISPR array revealed that all spacers do not target known *C. difficile* phages. Spacers 2 (100% identity) and 5 (97,14% identity) were, however, detected in several other *C. difficile* genomes, but spacers 1, 3 and 4 did not match known sequences (Table S2). Of note, no other phages were detected in the strain carrying phiCDKH01, supporting that the CRISPR array could phiCDKH01 be active and prevents further infection by phages.

Phylogenetic analysis

The entire genome sequence of phiCDKH01 was used in multiple genome alignment along with genomic sequences of other 10 *Clostridioides difficile* siphophages available in the GenBank. Alignment of genomes was performed with Mauve v. 2.3.1 (<http://darlinglab.org/mauve/mauve.html>) [25] using progressive Mauve method. Obtained results were visualised with FigTree v. 1.4.4 software (<https://github.com/rambaut/figtree>) (Fig.2A). The most closely related phage turned out to be phiCD24-

1, which was originally isolated from clinical isolate exhibiting 078 PCR ribotype [13, 26]. The sequences of phiCDKH01 and phiCD24-1 share 89% identity (Fig.2B).

Location of phiCDKH01 phage in the genome of *C. difficile*

Bacterial genomic DNA of CD34-Sr strain was isolated using E.Z.N.A. Bacterial DNA Kit (OMEGA bio-tek, USA). Whole-genome sequencing was performed on the Illumina MiSeq platform (Genomed S.A.) with 72-fold coverage. Upon quality check reads were assembled *de novo* using SPAdes v. 3.13.0 into 70 contigs. Obtained sequences were deposited in GenBank under accession number JACSDL000000000 and subjected to automatic annotation. The sequence of phiCDKH01 was found in the contig JACSDL010000003.1 in position 288,650 to 333,698. The prophage is integrated between loci H7706_07450 and H7706_07755. H7706_07450 shares homology to the manganese catalase family protein (GenBank accession MBC6710325.1). H7706_07755 is annotated as *ilvB* gene coding for the biosynthetic-type acetolactate synthase large subunit (GenBank accession MBC6710385.1).

Declarations

Funding The research was supported by the National Science Centre of Poland MINIATURA Programme (2018/02/X/NZ6/01360).

Conflict of interests The authors declare that they have no conflict of interests.

Ethical approval This article does not contain any study performed with human participants or animals.

References

1. Evans CT, Safdar N (2015) Current trends in the epidemiology and outcomes of *Clostridium difficile* Clin Infect Dis 60 (Suppl 2): S66-S71. <https://doi.org/10.1093/cid/civ140>
2. Centers for Disease Control and Prevention : ANTIBIOTIC RESISTANCE THREATS in the United States, 2019
3. Kwok CS, Arthur AK, Anibueze CI, Singh S, Cavallazzi R, Loke YK (2012) Risk of *Clostridium difficile* infection with acid suppressing drugs and antibiotics: meta-analysis. Am J Gastroenterol. 107:1011-1019. <https://doi.org/1038/ajg.2012.108>
4. Slimings C, Riley TV. (2014) Antibiotics and hospital-acquired *Clostridium difficile* infection: update of systematic review and meta-analysis. J Antimicrob Chemother. 69:881-891. <https://doi.org/10.1093/jac/dkt477>
5. George RH, Symonds JM, Dimock F, Brown JD, Arabi Y, Shinagawa N, Keighley MR, Alexander-Williams J, Burdon DW (1978) Identification of *Clostridium difficile* as a cause of *pseudomembranous colitis*. Br Med J. 1:695. <https://doi.org/10.1136/bmj.1.6114.695>
6. Dobson G, Hickey C, Trinder J. (2003) *Clostridium difficile* colitis causing toxic megacolon, severe sepsis and multiple organ dysfunction syndrome. Intensive Care Med. 29:1030.

<https://doi.org/10.1007/s00134-003-1754-7>

7. Leffler DA, LaMont JT (2015) *Clostridium difficile* N Engl J Med. 372:1539-1548. <https://doi.org/10.1056/NEJMra1403772>
8. Aslam S, Hamill RJ, Musher DM (2005) Treatment of *Clostridium difficile*- associated disease: old therapies and new strategies. Lancet Infect Dis 5:549-557. [https://doi.org/10.1016/S1473-3099\(05\)70215-2](https://doi.org/10.1016/S1473-3099(05)70215-2)
9. Zucca M, Scutera S, Savoia D (2013) Novel avenues for *Clostridium difficile* infection drug discovery. Expert Opin Drug Discov 8:459-477. <https://doi.org/10.1517/17460441.2013.770466>
10. Freeman J, Baines SD, Jabes D, Wilcox MH (2005) Comparison of the efficacy of ramoplanin and vancomycin in both in vitro and in vivo models of clindamycin-induced *Clostridium difficile* J Antimicrob Chemother 56:717-725. <https://doi.org/10.1093/jac/dki321>
11. Sekulovic O, Fortier L-C (2015) Global transcriptional response of *Clostridium difficile* carrying the CD38 prophage. Appl Environ Microbiol. 81:1364-1374. <https://doi.org/10.1128/AEM.03656-14>
12. Hargreaves KR, Clokie MRJ (2015) A taxonomic review of *Clostridium difficile* phages and proposal of a novel genus, "Phimmp04likevirus." Viruses 7:2534-2541. <https://doi.org/10.3390/v7052534>
13. Sekulovic O, Garnea JR, Néron A, Fortier L-C (2014) Characterization of temperate phages infecting *Clostridium difficile* isolates of human and animal origins. Appl Environ Microbiol. 80:2555-63. <https://doi.org/1128/AEM.00237-14>
14. Phothichaisri W, Ounjai P, Phetruen T, Janvilisri T, Khunrae P, Singhakaew S, Wangroongsarb P, Chankhamhaengdecha S (2018) Characterization of Bacteriophages Infecting Clinical Isolates of *Clostridium difficile*. Front Microbiol. 9:1701. <https://doi.org/10.3389/fmicb.2018.01701>
15. Ackermann HW (2009) Phage classification and characterization. Methods Mol Biol 501:127-140. https://doi.org/10.1007/978-1-60327-164-6_13
16. Rashid SJ, Barylski J, Hargreaves KR, Millard AA, Vinner GK, Clokie MRJ (2016) Two novel myoviruses from the north of Iraq reveal insights into *Clostridium difficile* phage diversity and biology. Viruses 8:310. <https://doi.org/10.3390/v8110310>
17. Shan J, Patel KV, Hickenbotham PT, Nale JY, Hargreaves KR, Clokie MRJ (2012) Prophage carriage and diversity within clinically relevant strains of *Clostridium difficile*. Appl. Environ. Microbiol. 78: 6027-6034. <https://doi.org/1128/AEM.01311-12>
18. Sambrook J, and Russell D. (2001) Molecular Cloning: A Laboratory Manual. New York, NY: Cold Spring Harbor Laboratory Press.
19. Ackermann HW, Prangishvili D (2012) Prokaryote viruses studied by electron microscopy. Arch Virol 157:1843-1849. <https://doi.org/10.1007/s00705-012-1383-y>
20. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T et al (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genomics. 9:75. <https://doi.org/10.1186/1471-2164-9-75>
21. Tsuchimoto S, Ohtsubo H, Ohtsubo E (1988) Two genes, pemK and pemL, responsible for stable maintenance of resistance plasmid R100. J Bacteriol 170:1461-6.

<https://doi.org/10.1128/jb.170.4.1461-1466.1988>

22. Popoff MR, Rubin EJ, Gill DM, Boquet P (1988) Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* *Infect Immun.* 56:2299-306. <https://doi.org/10.1128/IAI.56.9.2299-2306.1988>.
23. Diaz OR, Sayer CV, Popham DL, Shen A (2018) *Clostridium difficile* Lipoprotein GerS Is Required for Cortex Modification and Thus Spore Germination. *mSphere.* 3:e00205-18 <https://doi.org/doi:10.1128/mSphere.00205-18>.
24. Li G, Shen M, Lu S, Le S, Tan Y, Wang J, Zhao X, Shen W, Guo K, Yang Y, et al. (2016) Identification and characterization of the HicAB toxin-antitoxin system in the opportunistic pathogen *Pseudomonas aeruginosa*. *Toxins.*8:113. <https://doi.org/doi:3390/toxins8040113>
25. Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14:1394-403. <https://doi.org/1101/gr.2289704>
26. Fortier L-C, Moineau S (2007) Morphological and genetic diversity of temperate phages in *Clostridium difficile*. *Appl Environ Microbiol.* 73:7358-66. <https://doi.org/1128/AEM.00582-07>
27. Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009-1010. <https://doi.org/10.1093/bioinformatics/btr039>

Figures

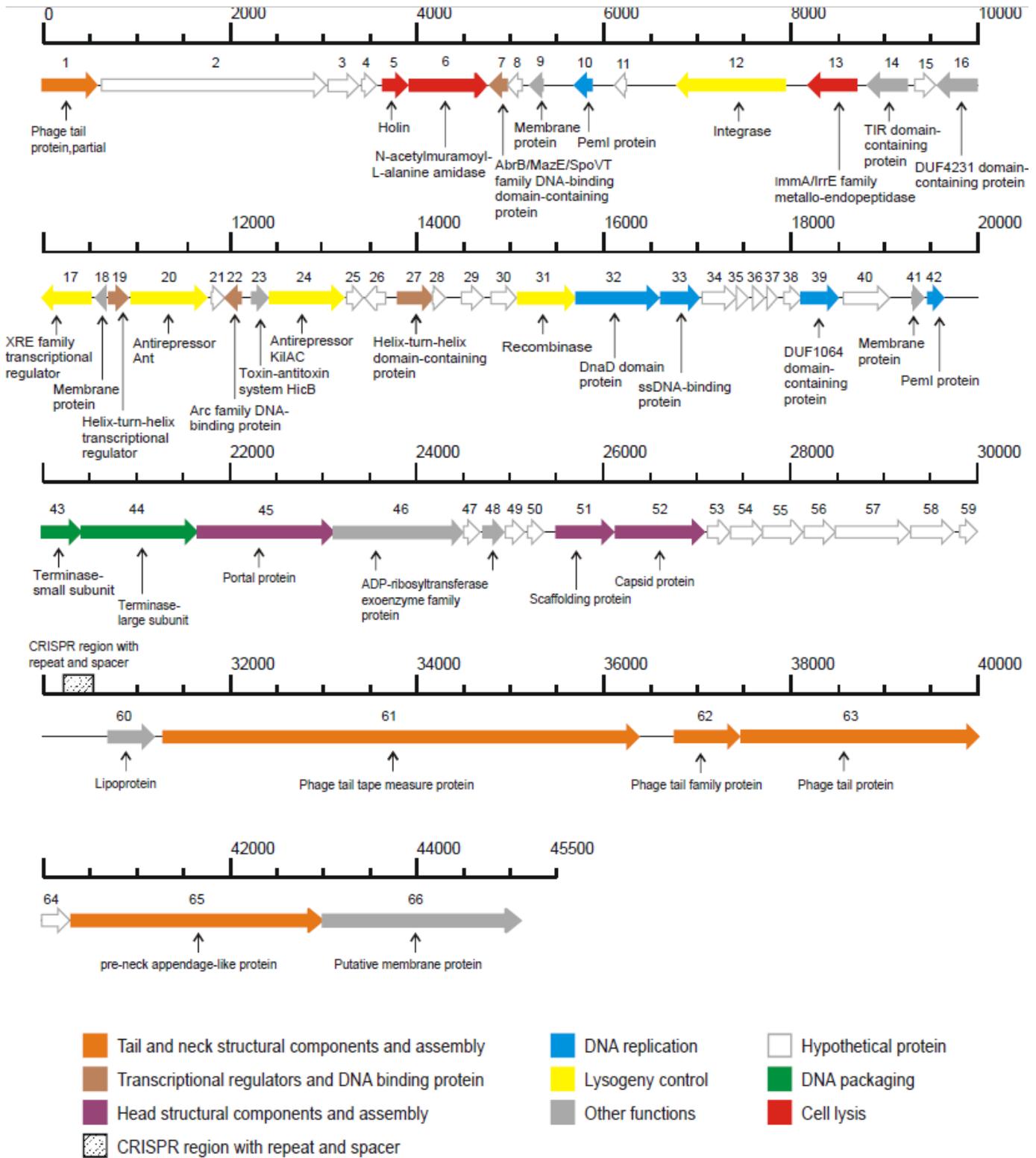


Figure 1

Genome features of phage phiCDKH01. The predicted ORFs and their orientations are represented by arrows. The putative functional assignments are indicated below the ORFs. The functional modules were assigned based on gene annotation and genomic organization and are shown in different colors. The position of CRISPR region with repeats and spacers is indicated and are shown in obliquely hatched pattern.

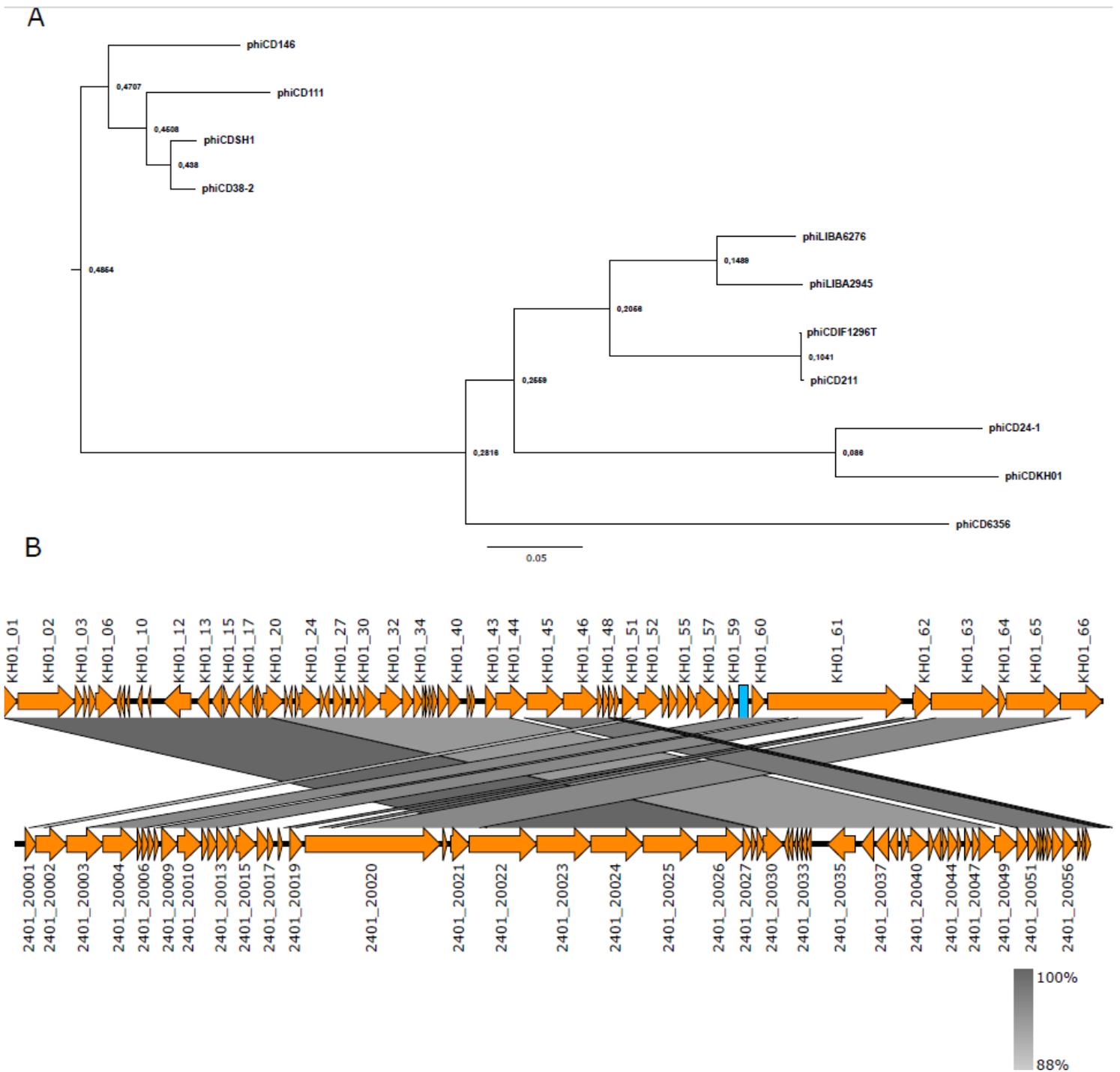


Figure 2

(A) Comparative phylogenetic analysis of entire genomes of *C. difficile* siphophages available in the GenBank. The figure represents the guide tree calculated by the progressive Mauve algorithm. Numbers associated with each branch represent node ages. (B) Comparison of the genome sequence of phage phiCDKH01 (top) with phiCD24-1 (bottom). Predicted ORFs and direction of transcription are indicated by block arrows. Blue box represents a putative CRISPR. Conserved regions are shaded with grey. Colour intensity corresponds to identity level (89% to 100%). Genomic comparisons were performed with

BLASTN. Similarities with E values lower than $1e-100$ are plotted. The figure was produced with Easyfig 2.2.5 [27].

Supplementary Files

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

- [TableS1.doc](#)
- [TableS2.doc](#)