



SHORT GENOME REPORT

Open Access



Complete genome sequence of bacteriophage P26218 infecting *Rhodoferrax* sp. strain IMCC26218

Kira Moon¹, Ilnam Kang², Suhyun Kim², Jang-Cheon Cho^{2*} and Sang-Jong Kim^{1*}

Abstract

Bacteriophage P26218 is a virus that thrives in freshwater and infects *Rhodoferrax* sp. strain IMCC26218, both of which were isolated from Soyang Lake, Korea. The bacterial host, IMCC26218, belongs to the genus *Rhodoferrax* and is closely related to *R. saidenbachensis*, with 98.7 % 16S rRNA gene sequence similarity. Bacteriophage P26218 has an icosahedral head structure with a diameter of ~52 nm and short tail of ~9 nm, which is a typical morphology of the *Podoviridae* family. Its complete dsDNA genome was 36,315 bp with 56.7 % G + C content. This is the first genome sequence reported for a lytic phage of the genus *Rhodoferrax*.

Keywords: Bacteriophage, *Rhodoferrax*, Freshwater, *Podoviridae*, Genome

Introduction

Bacteriophages, which are obligate parasites of bacterial cells, are the most abundant biological entities that can be found in all biospheres [1–3]. Considering the fact that phages heavily influence the bacterial community structure [4] and various biochemical cycles such as the carbon cycle [5], understanding the genetic potential and diversity of phages would be important in the study of microbial community dynamics. Due to the lack of a universal phylogenetic marker gene to help understand phage diversity, several studies have been reported that include a survey of entire phage populations via metagenomics, from various environments including seawater, hot springs, soil, and freshwater [6–8]. These viral metagenomic studies demonstrate the extremely diverse nature and novel genetic repertoire of viruses, but the limited number of phage genomes poses a challenge for interpretation of virome data. Difficulty in phage isolation and genome sequencing is simply due to the lack of available bacterial hosts, since many of bacteria in natural environments are yet to be cultured [2]. Therefore, isolation of phages infecting major groups of bacteria

and unveiling their genomic information are required to provide detailed information about each phage and enable meaningful interpretation of virome data.

The class *Betaproteobacteria* is often the most abundant group in freshwater environments, though less abundant in marine environments [9, 10]. Metagenomic studies on several freshwater bacteria revealed that the family *Comamonadaceae*, arbitrarily named betI [9], is the most frequently found family [11] within this class. The genus *Rhodoferrax* [12], belonging to the family *Comamonadaceae*, is found in diverse habitats including ditch water, activated sludge, Antarctic microbial mats, and water reservoirs [10, 12–14]. Additionally, this is one of the most abundant genera within the 16S rRNA gene database [15]. Therefore, understanding the ecology of the genus *Rhodoferrax* and its lytic phage will contribute to the understanding of freshwater microbial dynamics and help in further freshwater phage genomic studies. To isolate bacteriophages infecting *Rhodoferrax* spp., we successfully isolated phage P26218, which infects *Rhodoferrax* sp. IMCC26218 and further details of its genome features and annotations are described below.

* Correspondence: chojc@inha.ac.kr; sjkimm@snu.ac.kr

²Department of Biological Sciences, Inha University, Incheon 402-751, Republic of Korea

¹School of Biological Sciences, Seoul National University, Seoul 151-747, Republic of Korea



Virus information

Classification and features

A bacteriophage, designated P26218 that infects the bacterial strain IMCC26218 was isolated from Soyang Lake, located inland of Gangwon-do, Korea, in October 2014. A bacterial strain, IMCC26218, was also isolated from the same site, using standard dilution plating technique on R2A agar (Becton, Dickenson and Company, Franklin Lakes, NJ, USA) in April 2014. Based on a comparative 16S rRNA gene sequence analyses, strain IMCC26218 was found to belong to the genus *Rhodoferrax* with 98.7 % sequence similarity to *R. saidenbachensis* ED16^T. To screen a representative lytic phage infecting representatives of the class *Betaproteobacteria*, *Rhodoferrax* sp. IMCC26218 was used as the bacterial host.

Phage P26218 is a lytic phage that forms plaques of 1 to 2 mm in diameter, on *Rhodoferrax* sp. IMCC26218 culture plates. Transmission electron microscopy of purified phage particles revealed its icosahedral-shaped head (52.1 nm in diameter) with a 9.4-nm long short tail (Fig. 1). The capsid encapsulates a linear dsDNA with length of 36,315 bp with 56.7 % G + C content. The morphology of the viral particle, including a uniform, icosahedral-shaped head with a short tail indicated that this phage belonged to the family *Podoviridae* of the order *Caudovirales* [16]. However, when its genomic characteristics were considered, no similar genomic architecture was found among the known viral genera, leaving P26218 without an assigned genus. The amino acid sequence of DNA polymerase I (encoded by *polA*) of P26218, one of the widely used viral phylogenetic markers [17, 18], was aligned with that of representative strains of the families *Podoviridae* and *Siphoviridae* and the aligned sequences were used for phylogenetic analysis. The phylogenetic tree based on DNA polymerase I revealed that P26218 formed a clade with a marine metagenome sequence, parted from previously known type species, confirming limitations in its assignment to

a known genus (Fig. 2). A summary of the general phylogenetic features and isolation information are shown in Table 1.

Genome sequencing information

Genome project history

Compared to phage genomics and viromics in marine environments, relatively fewer studies with phage isolation and viral metagenome have been conducted in freshwater environments. Bacteriophage P26218 is the first lytic phage identified that infects the genus *Rhodoferrax*, one of the representatives of the class *Betaproteobacteria* in freshwater environments. In this study, both virus and host were isolated from Soyang Lake in Korea. This phage was selected for genome sequencing as an initial approach to understand phages infecting members of the *Betaproteobacteria* isolated from surface freshwater in Korea. Genomic DNA was sequenced by the ChunLab Inc. Genome assembly, annotation, and submission to GenBank were performed at the Department of Biological Sciences, Inha University. This genome project was registered in Genomes Online Database, with accession ID, Gp0111341 as well as GenBank, with an accession ID of KP792623. A summary of the project information is described in Table 2.

Growth conditions and genomic DNA preparation

The bacterial host, *Rhodoferrax* sp. IMCC26218, was routinely cultured and maintained on R2A agar or in R2A broth (MB Cell, Los Angeles, CA, USA) at 20 °C. To screen lytic phages infecting this bacterial host, 10 l of water sample was collected from Soyang Lake at a depth of 1 m. The water sample was initially filtered using a 0.2- μ m polyvinylidene difluoride membrane filter (Merck Millipore, Darmstadt, Germany) to remove bacterial-sized particles. To 400 ml filtrate, 100 ml of 5 \times R2A broth and 20 ml of IMCC26218 culture in the exponential phase were added, followed by incubation at 20 °C for 2 weeks for

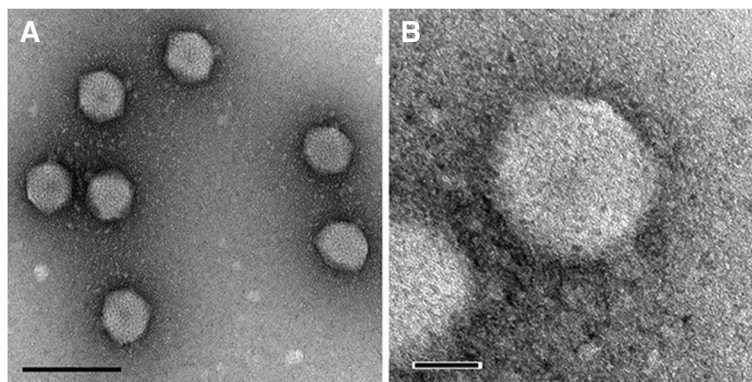
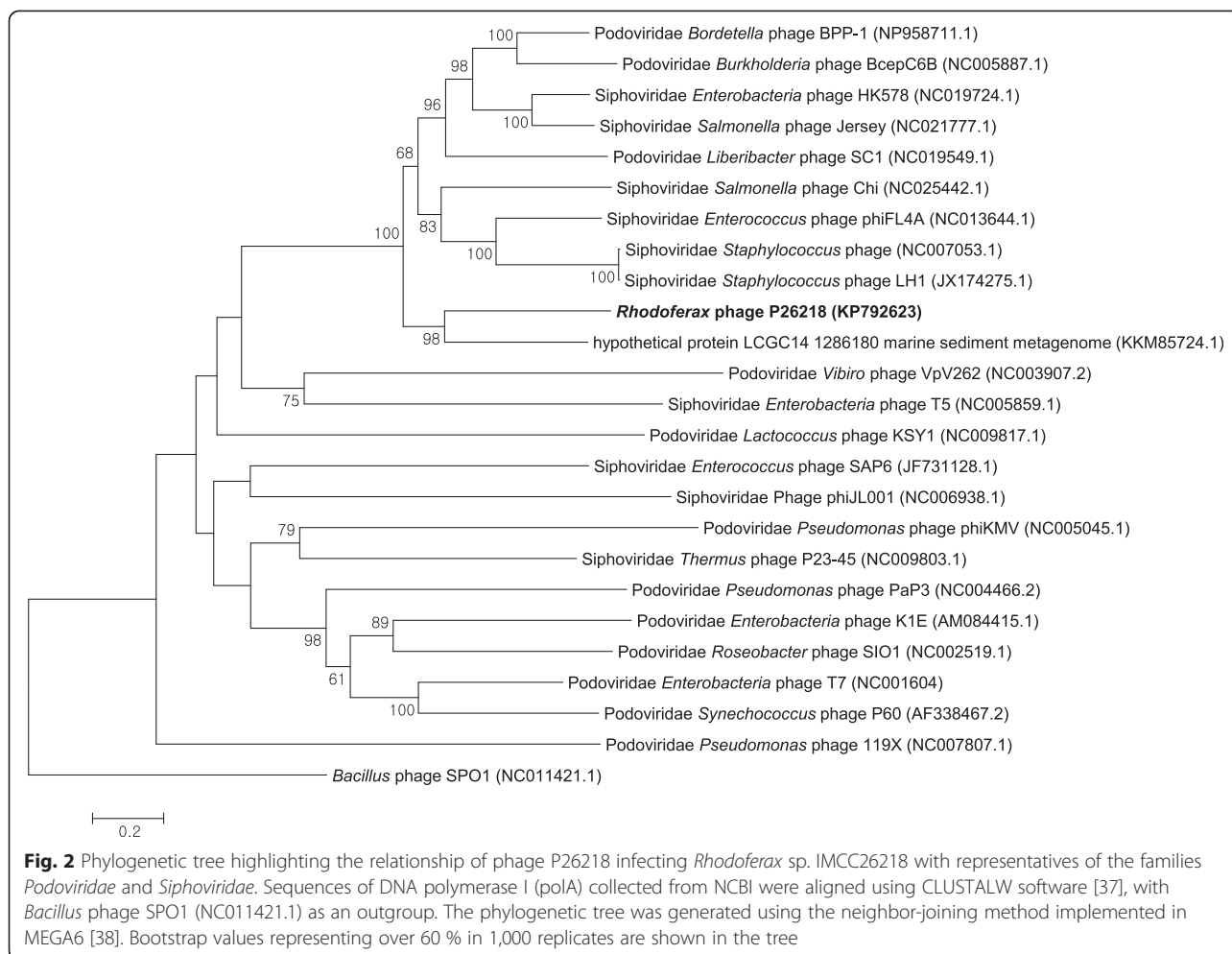


Fig. 1 Transmission electron micrographs of phage P26218 particles infecting *Rhodoferrax* sp. IMCC26218. The TEM images were obtained using Philips CM200 electron microscope. Scale bars represent 100 nm in (a) and 20 nm in (b)



enrichment of bacteriophages that specifically infect *Rhodofera* sp. IMCC26218. During the incubation period, 10 ml of the enrichment culture was sub-sampled 5 times at a 3-day interval. Each sub-sample was treated with approximately 3 ml chloroform to inactivate the bacterial cells. The treated samples were used for spot-double agar layer plaque assay on a *Rhodofera* sp. IMCC26218 lawn plate for phage screening via appearance of plaques [19], resulting in the isolation of phage P26218.

The purification of phage P26218 genomic DNA was performed as per the 'Molecular Cloning: A Laboratory Manual' [20] with minor modifications. To 200 ml of phage lysate prepared for DNA purification, 1 $\mu\text{g ml}^{-1}$ of DNase I and RNase A were added, followed by 11.7 g of NaCl. The obtained mixture was transferred to centrifuge bottles, to which PEG 8000 (Sigma-Aldrich, St. Louis, MO, USA) was added to attain a concentration of 10 % (w/v). After overnight incubation at 4 °C, the mixture was centrifuged at 11,000 $\times g$ for 40 min, supernatant was discarded by gentle inversion of the bottle, and the pellet was resuspended in 3–5 ml of SM buffer (50 mM Tris-HCl, pH 7.5; 100 mM NaCl; 10 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; 0.01 %

gelatin). PEG was removed from the liquid by treating it with equal volume of chloroform. The aqueous phase was then collected and further concentrated by ultracentrifugation at 246,000 $\times g$ for 2 h using a Beckman Coulter L-90 K ultracentrifuge with a SW 50 Ti swinging-bucket rotor. The phage pellet was resuspended in 100 μl SM buffer and used for genomic DNA extraction, using Qiagen DNeasy Blood and Tissue Kit, according to the manufacturer instructions.

Genome sequencing and assembly

The genome of phage P26218 was sequenced at ChunLab Inc. using Illumina MiSeq system with 2 \times 300-bp paired-end reads. The Illumina platform produced a total of 2 \times 798,245 reads. The initial total reads were split by 2 \times 50,000 reads into 16 sets [21] to facilitate the assembly process. Each set of sequence reads was independently assembled using SPAdes-3.1.1 [22], yielding a single contig but with different start points. Gap-closing PCR was performed with primers designed within the end region of a contig, which resulted in the circularization of the genome sequence. Circular

Table 1 Classification and general features of phage P26218 infecting *Rhodofex* sp. IMCC26218

MIGS ID	Property	Term	Evidence code ^a
	Classification	Genome group: dsDNA viruses, no RNA stage Phylum: unassigned Class: unassigned Order: <i>Caudovirales</i> Family: <i>Podoviridae</i> Genus: unassigned Species: unassigned Strain: P26218	IDA TAS [16] TAS [16]
	Particle shape	Icosahedral head with a short tail	IDA
MIGS-6	Habitat	Freshwater lake, surface	IDA
MIGS-15	Biotic relationship	Intracellular parasite of <i>Rhodofex</i> sp. IMCC26218	IDA
MIGS-14	Pathogenicity	Virulent phage of <i>Rhodofex</i> sp. IMCC26218	IDA
MIGS-4	Geographic location	Soyang Lake, Gangwon-do, South Korea	IDA
MIGS-5	Sample collection	Oct. 17, 2014	IDA
MIGS-4.1	Latitude	37°57'11" N	IDA
MIGS-4.2	Longitude	127°49'02" E	IDA
MIGS-4.3	Depth	1 m	IDA
MIGS-4.4	Altitude		

^aIDA: Inferred from Direct Assay and TAS: Traceable Author Statement. The evidence codes are from the Gene Ontology project [39]

assembly of the genome sequence suggested that the phage genome is terminally redundant or circularly permuted. This procedure for genome sequencing and assembly finally produced 36,315 bp with approximately 2,500× fold-coverage of the genome.

Genome annotation

The ORFs were predicted using 3 gene prediction programs: GeneMark.hmm version 3.25 [23], Rapid Annotation using Subsystem Technology server version 2.0 [24], and NCBI Gene Locator and Interpolated Markov ModelER version 3.02 [25]. Only the ORFs that were identified by 2 of the 3 gene-prediction programs were included in the annotation. Each predicted ORF was translated and used to search for its homologous proteins and predict its domains using the NCBI BLASTP [17, 26], HHpred server [27] and HMMER [28] upon NCBI non-redundant database [26], the Conserved Domain Database [29], Pfam database [30], COG [31], PRK [29], and TIGRFam [32]. Then, TMHMM [33] and SignalP [34] were used to predict transmembrane helices and signal peptides.

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Finished
	Number of contigs	1
MIGS-28	Libraries used	One paired-end Illumina library
MIGS 29	Sequencing platforms	Illumina Miseq
MIGS 31.2	Fold coverage	2,551×
MIGS 30	Assemblers	SPAdes version 3.1.1
MIGS 32	Gene calling method	RAST version 2.0, GeneMark.hmm version 3.25 and GLIMMER version 3.02
	GenBank ID	KP792623
	GenBank Date of Release	April, 2015
	GOLD ID	Gp0111341
	BIOPROJECT	NA ^a
MIGS 13	Source Material Identifier	NA ^a
	Project relevance	Diversity of freshwater bacteriophage

^aNot available

Genome properties

The properties and statistics of P26218 genome are summarized in Tables 3 and 4. The total length of the P26218 dsDNA genome was found to be 36,315 bp with 56.7 % G + C content. All 44 predicted ORFs were protein-coding sequences. However, only 15 of them were assigned to putative protein functions, while 29

Table 3 Nucleotide content and gene-count levels of the genome

Attribute	Value	% of Total ^a
Genome size (bp)	36,315	100.00
DNA coding (bp)	33,796	93.06
DNA G + C (bp)	20,589	56.70
DNA scaffolds	1	100.00
Total genes	44	100.00
Protein coding genes	44	100.00
RNA genes	0	0.00
Pseudo genes	0	0.00
Genes in internal clusters	0	0.00
Genes with function prediction	15	34.09
Genes assigned to COGs	11	25.00
Genes with Pfam domains	21	36.36
Genes with signal peptides	1	2.27
Genes with transmembrane helices	0	0.00
CRISPR repeats	0	0.00

^aThe total percentage is based on the total number of protein coding genes in the genome

Table 4 Number of genes associated with general COG functional categories

Code	Value	% of total ^a	Description
J	1	2.22	Translation, ribosomal structure and biogenesis
A	0		RNA processing and modification
K	1	2.22	Transcription
L	3	9.09	Replication, recombination and repair
B	0		Chromatin structure and dynamics
D	1	2.22	Cell cycle control, Cell division, chromosome partitioning
V	0		Defense mechanisms
T	0		Signal transduction mechanisms
M	0		Cell wall/membrane biogenesis
N	0		Cell motility
U	0		Intracellular trafficking and secretion
O	1	2.22	Posttranslational modification, protein turnover, chaperones
C	0		Energy production and conversion
G	0		Carbohydrate transport and metabolism
E	1	2.22	Amino acid transport and metabolism
F	2	4.44	Nucleotide transport and metabolism
H	0		Coenzyme transport and metabolism
I	0		Lipid transport and metabolism
P	0		Inorganic ion transport and metabolism
Q	0		Secondary metabolite biosynthesis, transport, and catabolism
R	2	4.44	General function prediction only
S	0		Function unknown
X	1	2.22	Mobilome: prophages, transposons
-	33	75.00	Not in COGs

^aThe percentage is based on the total number of protein-coding genes in the genome

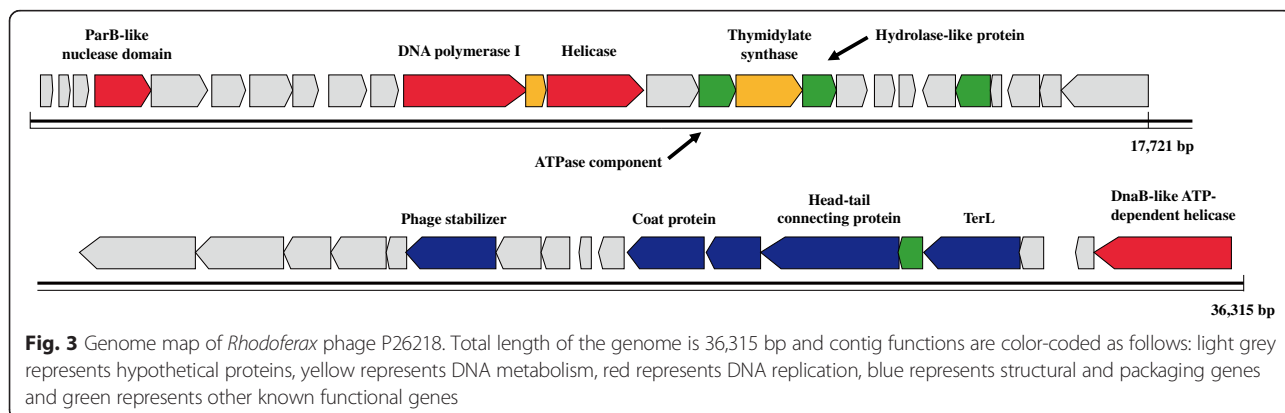
were assigned to hypothetical proteins. One gene with a signal peptide was identified but none were found to have transmembrane helices.

Insights from the genome sequence

According to the genome annotation, bacteriophage P26218 is a unique phage, with no closely related phages. Therefore, this phage could only be classified based on its morphological characteristics, which attributed it to the family *Podoviridae*. Out of 44 predicted ORFs, only 15 (34 %) were assigned with a known function. As shown in Fig. 3, four ORFs were predicted to be related to DNA replication, 2 to DNA metabolism, 5 to packaging and structural functions, and 4 to other known functions (Additional file 1). BLASTP analyses showed that each ORF with an identified function was homologous to ORFs from different phages belonging to different viral families. All ORFs encoding viral packaging function were closely related to those of other viruses in the family *Podoviridae*. The ORFs encoding DNA polymerase I, ATPase component, thymidylate synthase, and hydrolase-like protein were similar to those of the family *Siphoviridae*, while the genes for DnaB-like ATP-dependent helicase and ParB-like nuclease domain showed a high degree of homology to those of the family *Myoviridae*. This genomic architecture of P26218 confirmed the mosaic genome structure, known to be a result of lateral gene transfer usually predicted in viral genomes in attempts to enhance their genetic diversity [35, 36] and often observed in species of the order *Caudovirales* such as phages P22 and lambda.

Conclusions

Lytic bacteriophage P26218, isolated from a freshwater lake is the first virus identified that infects the genus *Rhodospirillum rubrum*. Based on its morphology, this phage was identified to be a member of the family *Podoviridae*, with an icosahedral-shaped head and short tail. All predicted ORFs from this phage genome were protein-



coding, with 3 specifically coding for DNA replication, 7 for DNA metabolism, and 5 for packaging and structural proteins. The group of ORFs with similar function was postulated to originate from different groups of viral families (*Podoviridae*, *Siphoviridae*, and *Myoviridae*), which was indicative of the mosaic property of the P26218 genome. It is expected that phage P26218 isolated in this research and its genome sequence would be further used to study bacteria-phage interactions in freshwater environments, to reveal the evolutionary role of phage lateral gene transfer and to interpret freshwater virome data.

Additional file

Additional file 1: Table S1. Description of data: Gene annotation table of bacteriophage P26218. (PDF 183 kb)

Abbreviations

PEG: Polyethylene glycol.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

IK and J-CC designed the study. SK isolated the bacterial host and collected freshwater samples. KM performed laboratory experiments, analyzed the data, and drafted the manuscript. KM, IK, J-CC, and S-JK wrote the manuscript together and finalized the study. We all authors have read and approved the final manuscript.

Acknowledgements

This work was supported by Mid-Career Research Program through National Research Foundation (NRF) funded by the Ministry of Science, ICT and Future Planning (to J-CC, NRF-2013R1A2A2A01068004) and partially by the General Research Program through NRF (to S-JK; 2012R1A1B3003609), Korea.

Received: 8 April 2015 Accepted: 29 October 2015

Published online: 24 November 2015

References

- Casas V, Rohwer F. Phage metagenomics. *Methods Enzymol.* 2007;421:259–68.
- Edwards RA, Rohwer F. Viral metagenomics. *Nat Rev Microbiol.* 2005;3:504–10.
- Rohwer F. Global phage diversity. *Cell.* 2003;113:141.
- Rodríguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, et al. Viral and microbial community dynamics in four aquatic environments. *ISME J.* 2010;4:739–51.
- Wommack KE, Nasko DJ, Chopyk J, Sakowski EG. Counts and sequences, observations that continue to change our understanding of viruses in nature. *J Microbiol.* 2015;53:181–92.
- Suttle CA. Viruses in the sea. *Nature.* 2005;437:356–61.
- Breitbart M, Wegley L, Leeds S, Schoenfeld T, Rohwer F. Phage community dynamics in hot springs. *Appl Environ Microbiol.* 2004;70:1633–40.
- Srinivasiah S, Bhavsar J, Thapar K, Liles M, Schoenfeld T, Wommack KE. Phages across the biosphere: contrasts of viruses in soil and aquatic environments. *Res Microbiol.* 2008;159:349–57.
- Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol R.* 2011;75:14–49.
- Zwart G, Crump BC, Kamst-van Agterveld MP, Hagen F, Han S-K. Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microbiol Ecol.* 2002;28:141–55.
- Cottrell MT, Waidner LA, Yu L, Kirchman DL. Bacterial diversity of metagenomic and PCR libraries from the Delaware River. *Environ Microbiol.* 2005;7:1883–95.
- Kaden R, Spröer C, Beyer D, Krolla-Sidenstein P. *Rhodoferrax saidenbachensis* sp. nov., a psychrotolerant, very slowly growing bacterium within the family Comamonadaceae, proposal of appropriate taxonomic position of *Albidiferrax ferrireducens* strain T118T in the genus *Rhodoferrax* and emended description of the genus *Rhodoferrax*. *Int J Syst Evol Microbiol.* 2014;64:1186–93.
- Madigan MT, Jung DO, Woese CR, Achenbach LA. *Rhodoferrax antarcticus* sp. nov., a moderately psychrophilic purple nonsulfur bacterium isolated from an Antarctic microbial mat. *Arch Microbiol.* 2000;173:269–77.
- Hiraishi A, Hoshino Y, Satoh T. *Rhodoferrax fermentans* gen. nov., sp. nov., a phototrophic purple nonsulfur bacterium previously referred to as the "*Rhodocyclus gelatinosus*-like" group. *Arch Microbiol.* 1991;155:330–6.
- Salka I, Čuperová Z, Mašín M, Koblížek M, Grossart HP. *Rhodoferrax*-related *pufM* gene cluster dominates the aerobic anoxygenic phototrophic communities in German freshwater lakes. *Environ Microbiol.* 2011;13:2865–75.
- King AM, Adams MJ, Carstens EB, Lefkowitz EJ. The double stranded DNA viruses. In: *Virus Taxonomy: Classification and Nomenclature of Viruses; Ninth Report of the International Committee on Taxonomy of Viruses.* San Diego: Elsevier Academic Press; 2012. p. 63–85.
- Adriaenssens EM, Cowan DA. Using signature genes as tools to assess environmental viral ecology and diversity. *Appl Environ Microbiol.* 2014;80:4470–80.
- Breitbart M, Miyake JH, Rohwer F. Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiol Lett.* 2004;236:249–56.
- Grabow W. Bacteriophages: update on application as models for viruses in water. *Water Sa.* 2004;27:251–68.
- Green MR, Sambrook J. *Molecular cloning: a laboratory manual.* New York: Cold Spring Harbor Laboratory Press; 2012.
- Lonardi S, Mirebrahim H, Wanamaker S, Alpert M, Ciardo G, Duma D, et al. When less is more: "slicing" sequencing data improves read decoding accuracy and *de novo* assembly quality. *bioRxiv.* 2015; doi:10.1101/013425
- Bankevich A, Nurk 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;19:455–77.
- Lukashin AV, Borodovsky M. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 1998;26:1107–15.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genomics.* 2008;9:75.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 1999;27:4636–41.
- Lavigne R, Seto D, Mahadevan P, Ackermann H-W, Kropinski AM. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. *Res Microbiol.* 2008;159:406–14.
- Söding J. Protein homology detection by HMM–HMM comparison. *Bioinformatics.* 2005;21:951–60.
- Eddy SR. Profile hidden Markov models. *Bioinformatics.* 1998;14:755–63.
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 2011;39:D225–9.
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res.* 2014;42:D222–30.
- Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28:33–6.
- Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. *Nucleic Acids Res.* 2003;31:371–3.
- Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001;305:567–80.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8:785–6.
- Yoshida M, Yoshida-Takashima Y, Nunoura T, Takai K. Genomic characterization of a temperate phage of the psychrotolerant deep-sea bacterium *Aurantimonas* sp. *Extremophiles.* 2015;19:49–58.

36. Swanson MM, Reavy B, Makarova KS, Cock PJ, Hopkins DW, Torrance L, et al. Novel bacteriophages containing a genome of another bacteriophage within their genomes. *PLoS ONE*. 2012;7:e40683.
37. Thompson JD, Gibson T, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*. 2002;Chapter 2:Unit 2.3.
38. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–9.
39. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25:25–9.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

