Editorial

Special Issue "Bacteriophage Genomics": Editorial

Igor V. Babkin * and Nina V. Tikunova

Laboratory of Molecular Microbiology, Institute of Chemical Biology and Fundamental Medicine,
630090 Novosibirsk, Russia

* Correspondence: i_babkin@mail.ru or babkin@niboch.nsc.ru

Virus genomics as a separate branch of biology has emerged relatively recently. Although the term "genomics" appeared only in 1986 [1], the history of genomics began in 1976, when the first ever complete RNA genome of bacteriophage MS2 (3569 nucleotides) had been sequenced [2]. During 1977, F. Sanger and coauthors first used their sequencing method to determine the first complete DNA genome of PhiX174 phage that infected *E. coli* [3]. Sequencing technologies have been further developed and automated, leading to the creation of hardware platforms for next generation sequencing (NGS) in the early 2000s [4]. So, the entire history of sequencing technology is closely related to phage genomes. PhiX174 is still one of the most commonly used positive controls for sequencing, including NGS.

The widespread use of NGS, both in metagenomic studies and in the sequencing of specific prokaryotic viruses revealed a striking diversity of bacteriophages. The development of sequencing techniques has led to the discovery of previously unknown bacteriophages of uncultivable prokaryotes. The size of newly discovered phage genomes varies from 3.5 kb in the case of microphages [5] to 735 kb in megaphages [6]. Bacteriophages have either DNA or RNA as their genome material. The striking diversity of bacteriophage genomes has led to the current predominance of the concept of their polyphyletic origin [7–9]. A large community of huge phages whose genome is more than 200 kb has recently been discovered [6,7,10]. These phages are found to be widespread in various ecosystems. In most cases, phages with extended genomes formed a separate clade according to phylogenetic analysis based on their specific genes [6]. Some of these genomes have a gene synteny that is different from most known phages; in some genomes, alternative genetic code is used. Some members of this phage community have even longer genomes than small bacteria [7,10].

The progress made in sequencing of the phage genomes has given a new look at their taxonomy. In 2008–2015, the *Autographiviridae*, *Myoviridae* and *Siphoviridae* families were revised based on their genomics [11]. Thereafter, the taxonomy of bacteriophages has been continuously modified and developed due to the appearance of new data on their genomes. Thus, while in 1999 the tailed phages had been divided into three families, which included 16 genera, and 30 species, in 2018 they were grouped into five families, 26 subfamilies, 363 genera, and 1320 species [7].Currently, phage taxonomy includes 47 families, 98 subfamilies, 1197 genera, and 3601 species (https://ictv.global/taxonomy; accessed on 20 February 2023). The process of phage taxonomy modification is not over yet.

The main result achieved in the bacteriophages genomics in recent years is the discovery of a significant mosaicism of their genomes. This indicates a widespread exchange of genetic information between phages by both homologous and non-homologous recombination [7,8,11]. This often makes the use of conventional methods of phylogenetic analysis unreliable, creating considerable difficulties in the taxonomy of bacteriophages and the study of their evolution. Network-based approaches, which are much more informative [7], should be used to solve such a problem.

Progress in expanding the bacteriophage genome databases has led to significant advances in both comparative and functional genomics of the viruses [7,8]. However, there

are difficulties in the comparative genomics of phages. This is due to the lack of standard guidelines for the NGS phage genomes data submission; therefore, even closely related genomes are often submitted with different orientations and/or starting points. Such difficulties occur when linear genomes with terminal repeats are submitted into databases. To solve the above problems, various special softwares have been created for comparative analysis of bacteriophage genomes–ViPTree server [12], Gegenees [13], VIRIDIC [14], and other programs. These programs are based on matching phage ORFs and/or BLAST comparison. Another difficulty in studying phage genomes is the presence of modified nucleotide bases in some genomes [15].

In comparative genomics of temperate bacteriophages, a significant part of their genetic information may be extracted from databases of bacterial genomes. In this case, viral sequences are prophages integrated into the host genome [16]. When analyzing the prophage sequences, it should be taken into account that some of them belong to the cryptic prophages that never form active phage particles [17]. However, even such genomes are of interest for studying the evolution of viruses.

This Special Issue "Bacteriophage Genomics" focuses on the latest achievements and problems of bacteriophage genomics. To date, 14 articles have been published on comparative and functional genomics of these viruses [18–31]. Most articles published in the Special Issue deal with the functional and structural genomics of various phages, including the identification of phage endolysins, potentially useful for therapy. Phages and prophages of *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* have been described in two articles [24,26]. These bacteria pose the greatest threat to human health, according to WHO (https://www.who.int/). In addition, viruses of pathogenic *Escherichia coli* strains have been analyzed in three studies [20,22,31] and in one article, phages of the *Campylobacter* genus are characterized [30]. *E. coli* and *Campylobacter* spp. are among the most common human and animal pathogens causing bacterial diarrhea. The published studies are in the trend of general interest in phage therapy of patients with infections caused by bacterial pathogens with multidrug resistance. It can be noted that five of the 14 publications in the Special Issue are devoted to prophages found in bacterial genomes, making an important contribution to the diversity of their hosts and to the exchange of genetic information between bacteria [19,25,26,28,30].

Importantly, some studies published in the Special Issue have proposed improved bioinformatic methods for annotation and analysis of phage genomes. Katelyn McNair and co-authors have developed a reliable method for in silico predicting protein-coding genes [18]. The developed workflow is important for comprehensive and accurate phage genome annotation. The method identifies both new genes and variants of existing genes. Certainly, this area of phage genomics requires further development since there are phages with extended genomes using alternative genetic code for stopping codons. This makes it difficult to correctly identify some genes when using traditional stop codons [6,10,32].

Moreover, prediction of host ranges of bacteriophages can be difficult in some cases. Currently, there is a growing interest in the therapeutic use of phage cocktails, which are combinations of phages able to lyse target bacteria. The use of such mixtures makes it possible to control infections for a wide range of pathogenic bacteria. Versoza and Pfeifer have performed computational prediction of the bacteriophage host range using modern bioinformatics tools [27]. Computer prediction of bacteriophage host ranges could significantly simplify the selection of the composition for such cocktails.

Although significant progress has been made in describing some viromes, such as those from the human intestine, the role of phages in many biotic communities remains poorly understood. In addition, it should be noted that the phage community is a dynamic system and genetic information is constantly exchanged both between phages and their hosts and between different phages, which is the main driving force of their evolution. The exponential increase in the number of phage genomes in databases leads to the need for constant revision of their relationships and taxonomy. This process will continue for the foreseeable future.

References

1. Ten Have, H.; Neves, M.D.C.P. Genomics. In *Dictionary of Global Bioethics*; Ten Have, H., Patrão Neves, M., Eds.; Springer: Cham, Switzerland, 2021; p. 571. [CrossRef]
2. Fiers, W.; Contreras, R.; Duerinck, F.; Haegeman, G.; Iserentant, D.; Merregaert, J.; Jou, W.M.; Molemans, F.; Raeymaekers, A.; Berghe, A.V.D.; et al. Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature* **1976**, *260*, 500–507. [CrossRef]
3. Sanger, F.; Air, G.M.; Barrell, B.G.; Brown, N.L.; Coulson, A.R.; Fiddes, J.C.; Hutchison, C.A.; Slocombe, P.M.; Smith, M. Nucleotide sequence of bacteriophage φX174 DNA. *Nature* **1977**, *265*, 687–695. [CrossRef] [PubMed]
4. Behjati, S.; Tarpey, P.S. What is next generation sequencing? *Arch. Dis. Child.-Educ. Pract. Ed.* **2013**, *98*, 236–238. [CrossRef]
5. Kirchberger, P.C.; Martinez, Z.A.; Ochman, H. Organizing the Global Diversity of Microviruses. *Mbio* **2022**, *13*, e0058822. [CrossRef] [PubMed]
6. Al-Shayeb, B.; Sachdeva, R.; Chen, L.-X.; Ward, F.; Munk, P.; Devoto, A.; Castelle, C.J.; Olm, M.R.; Bouma-Gregson, K.; Amano, Y.; et al. Clades of huge phages from across Earth's ecosystems. *Nature* **2020**, *578*, 425–431. [CrossRef]
7. Dion, M.B.; Oechslin, F.; Moineau, S. Phage diversity, genomics and phylogeny. *Nat. Rev. Microbiol.* **2020**, *18*, 125–138. [CrossRef] [PubMed]
8. Casey, A.; Coffey, A.; McAuliffe, O. Genetics and Genomics of Bacteriophages. In *Bacteriophages*; Harper, D.R., Abedon, S.T., Burrowes, B.H., McConville, M.L., Eds.; Springer: Cham, Switzerland, 2017; pp. 193–218. [CrossRef]
9. Koonin, E.V.; Dolja, V.V.; Krupovic, M.; Varsani, A.; Wolf, Y.I.; Yutin, N.; Zerbini, F.M.; Kuhn, J.H. Global Organization and Proposed Megataxonomy of the Virus World. *Microbiol. Mol. Biol. Rev.* **2020**, *84*, e00061-19. [CrossRef] [PubMed]
10. Devoto, A.E.; Santini, J.M.; Olm, M.R.; Anantharaman, K.; Munk, P.; Tung, J.; Archie, E.A.; Turnbaugh, P.J.; Seed, K.D.; Blekhman, R.; et al. Megaphages infect Prevotella and variants are widespread in gut microbiomes. *Nat. Microbiol.* **2019**, *4*, 693–700. [CrossRef]
11. Tolstoy, I.; Kropinski, A.M.; Brister, J.R. Bacteriophage Taxonomy: An Evolving Discipline. In *Bacteriophage Therapy. Methods in Molecular Biology*; Azeredo, J., Sillankorva, S., Eds.; Humana Press: New York, NY, USA, 2018; Volume 1693, pp. 57–71. [CrossRef]
12. Nishimura, Y.; Yoshida, T.; Kuronishi, M.; Uehara, H.; Ogata, H.; Goto, S. ViPTree: The viral proteomic tree server. *Bioinformatics* **2017**, *33*, 2379–2380. [CrossRef]
13. Ågren, J.; Sundström, A.; Håfström, T.; Segerman, B. Gegenees: Fragmented Alignment of Multiple Genomes for Determining Phylogenomic Distances and Genetic Signatures Unique for Specified Target Groups. *PLoS ONE* **2012**, *7*, e39107. [CrossRef]
14. Moraru, C.; Varsani, A.; Kropinski, A.M. VIRIDIC—A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses* **2020**, *12*, 1268. [CrossRef] [PubMed]
15. Weigele, P.; Raleigh, E.A. Biosynthesis and Function of Modified Bases in Bacteria and Their Viruses. *Chem. Rev.* **2016**, *116*, 12655–12687. [CrossRef]
16. Andrade-Martínez, J.S.; Valera, L.C.C.; Cárdenas, L.A.C.; Forero-Junco, L.; López-Leal, G.; Moreno-Gallego, J.L.; Rangel-Pineros, G.; Reyes, A. Computational Tools for the Analysis of Uncultivated Phage Genomes. *Microbiol. Mol. Biol. Rev.* **2022**, *86*, e0000421. [CrossRef]
17. Wang, X.; Wood, T.K. Cryptic prophages as targets for drug development. *Drug Resist. Updat.* **2016**, *27*, 30–38. [CrossRef]
18. McNair, K.; Zhou, C.E.; Souza, B.; Malfatti, S.; Edwards, R. Utilizing Amino Acid Composition and Entropy of Potential Open Reading Frames to Identify Protein-Coding Genes. *Microorganisms* **2021**, *9*, 129. [CrossRef] [PubMed]
19. Gillis, A.; Hock, L.; Mahillon, J. Comparative Genomics of Prophages Sato and Sole Expands the Genetic Diversity Found in the Genus *Betatectivirus*. *Microorganisms* **2021**, *9*, 1335. [CrossRef] [PubMed]
20. Zhang, Y.; Liao, Y.-T.; Salvador, A.; Lavenburg, V.; Wu, V. Characterization of Two New Shiga Toxin-Producing *Escherichia coli* O103-Infecting Phages Isolated from an Organic Farm. *Microorganisms* **2021**, *9*, 1527. [CrossRef]

21. Zrelovs, N.; Dislers, A.; Kazaks, A. Genome Characterization of Nocturne116, Novel *Lactococcus lactis*-Infecting Phage Isolated from Moth. *Microorganisms* **2021**, *9*, 1540. [CrossRef]

22. Koonjan, S.; Cooper, C.; Nilsson, A. Complete Genome Sequence of vB_EcoP_SU7, a *Podoviridae* Coliphage with the Rare C3 Morphotype. *Microorganisms* **2021**, *9*, 1576. [CrossRef]

23. Nguyen, T.-T.; Xuan, T.T.T.; Ngoc, T.H.; Duyen, L.T.M.; Vinh, T.Q.; My, P.D.T.; Hoang, H.A.; Nga, L.P. Diverse Bacteriophages Infecting the Bacterial Striped Catfish Pathogen *Edwardsiella ictaluri*. *Microorganisms* **2021**, *9*, 1830. [CrossRef]

24. Kwon, J.; Kim, S.W.; Kim, S.G.; Kang, J.W.; Jung, W.J.; Bin Lee, S.; Lee, Y.M.; Giri, S.S.; Chi, C.; Park, S.C. The Characterization of a Novel Phage, pPa_SNUABM_DT01, Infecting *Pseudomonas aeruginosa*. *Microorganisms* **2021**, *9*, 2040. [CrossRef] [PubMed]

25. Olonade, I.; van Zyl, L.J.; Trindade, M. Genomic Characterization of a Prophage, Smhb1, That Infects *Salinivibrio kushneri* BNH Isolated from a Namib Desert Saline Spring. *Microorganisms* **2021**, *9*, 2043. [CrossRef]

26. Marques, A.T.; Tanoeiro, L.; Duarte, A.; Gonçalves, L.; Vítor, J.M.B.; Vale, F.F. Genomic Analysis of Prophages from *Klebsiella pneumoniae* Clinical Isolates. *Microorganisms* **2021**, *9*, 2252. [CrossRef] [PubMed]

27. Versoza, C.J.; Pfeifer, S.P. Computational Prediction of Bacteriophage Host Ranges. *Microorganisms* **2022**, *10*, 149. [CrossRef]

28. Happel, A.-U.; Kullin, B.R.; Gamieldien, H.; Jaspan, H.B.; Varsani, A.; Martin, D.; Passmore, J.-A.S.; Froissart, R. In Silico Characterisation of Putative Prophages in *Lactobacillaceae* Used in Probiotics for Vaginal Health. *Microorganisms* **2022**, *10*, 214. [CrossRef]

29. Willenbücher, K.; Wibberg, D.; Huang, L.; Conrady, M.; Ramm, P.; Gätcke, J.; Busche, T.; Brandt, C.; Szewzyk, U.; Schlüter, A.; et al. Phage Genome Diversity in a Biogas-Producing Microbiome Analyzed by Illumina and Nanopore GridION Sequencing. *Microorganisms* **2022**, *10*, 368. [CrossRef]

30. Tanoeiro, L.; Oleastro, M.; Nunes, A.; Marques, A.T.; Duarte, S.V.; Gomes, J.P.; Matos, A.P.A.; Vítor, J.M.B.; Vale, F.F. Cryptic Prophages Contribution for *Campylobacter jejuni* and *Campylobacter coli* Introgression. *Microorganisms* **2022**, *10*, 516. [CrossRef] [PubMed]

31. Abdelrahman, F.; Rezk, N.; Fayez, M.S.; Abdelmoteleb, M.; Atteya, R.; Elhadidy, M.; El-Shibiny, A. Isolation, Characterization, and Genomic Analysis of Three Novel *E. coli* Bacteriophages That Effectively Infect *E. coli* O18. *Microorganisms* **2022**, *10*, 589. [CrossRef] [PubMed]

32. Ivanova, N.N.; Schwientek, P.; Tripp, H.J.; Rinke, C.; Pati, A.; Huntemann, M.; Visel, A.; Woyke, T.; Kyrpides, N.C.; Rubin, E.M. Stop codon reassignments in the wild. *Science* **2014**, *344*, 909–913. [CrossRef] [PubMed]