Current status: the morphology diversity of *Bacillus subtilis* phages and their genome size

Y A Handoko1,2, A K Wardani1, A Sutrisno1 and S B Widjanarko1

1 Department of Agroindustry Technology, Faculty of Agricultural Technology, Universitas Brawijaya, Malang, Indonesia
2 Department of Agrotechnology, Faculty Agriculture and Business, Satya Wacana Christian University, Salatiga, Indonesia

E-mail: yoga.handoko@uksw.edu

Abstract. *Bacillus subtilis* phages are viruses that specifically infect *Bacillus subtilis*. The exploration of *Bacillus subtilis* phages are very important related to their application in the agroindustry, especially as biocontrol in contamination of the raw materials and food products. The comparative study of morphology and genome size of *Bacillus subtilis* phages were carried out to understand phage interactions with their hosts. Until this article is presented, there are 18 species of *Bacillus subtilis* phages from various region of the world. They are 9 species in the *Myoviridae* family, 4 species in the *Siphoviridae* family, and 5 species in the *Podoviridae* family. The range of genome size of *B. subtilis* phages were known from 18,000 to 160,000 bp with different ranges of ORFs 15–250. Their protein function can be identified between 20–60%.

1. Introduction

*Bacillus subtilis* is a species of gram-positive bacteria which are distributed widely around the world [1]. This species has an important role in agroindustry activities. It is reported that *B. subtilis* caused contamination in food products in the Netherlands, including flour, pasta products, cocoa, chocolate, milk, herbs, spices, bakery and meat products [2]. *B. subtilis* also contaminated dough and bread in South Africa [3]. Moreover, their spores contaminated gelatin, ropy bread, and wheat flour [4]. As a gram-positive bacterium, *B. subtilis* produced endospores when they are in unfavorable environments, such as: high temperature, chemical agents, and also antibiotic treatments [1]. This condition makes us to find alternative strategies for controlling these bacteria as contaminant agent in the agricultural industries.

In the 21st century, we know that pesticides aren’t an eco-friendly agent [5] and many antibiotics are ineffective and resistant for some bacteria [6, 7]. Looking at this phenomenon, a breakthrough in utilizing bacteriophages as a biocontrol for unfavorable bacteria could be a good opportunity. Bacteriophage (s) which is also called phage (s) are viruses that infect bacteria [8]. Phages infect and destroy bacterial cells with high specificity in the short time [9]. In line with *B. subtilis* as a contamination agent, the study of morphology diversity and genome size of *B. subtilis* phages have become an important for the initial consideration before utilizing phages as biocontrol [10].

2. The Diversity of *B. subtilis* Phages Morphology

The comparison of *Bacillus phages* morphology was shown in Figure 1. Figure 1 shows that variation on *Bacillus phages* is divided into 3 families, namely: *Myoviridae*, *Siphoviridae*, and *Podoviridae*. The similarity of those families was based on the icosahedral capsid form and the difference of the neck and tail characterization. The *Myoviridae* has a neck with a fairly long contractile tail such as: SP82G (c), PBS1 (e), Phi25 (g), SP01 (h), SP02 (i), BSP10 (m), SPO (n), SIoΦ (o), and SUB (p). The *Siphoviridae* has a longer and non-contractile tail include: β22 (a), SP3 (b), PBSX (d), SPP1 (j), and
ARπ (l). *Podoviridae* has a characteristic of a very short tail, as seen in Phi29 (f) and vB_BsuP_Goe1 (k).

![Phage Images](image)

**Figure 1.** Comparison on the morphology of *B. subtilis* phages; *B. subtilis* phage β22 (a), *B. subtilis* phage SP3 (b), *B. subtilis* phage SP82G (c), *B. subtilis* phage PBSX (d), *B. subtilis* phage PBS1 (e), *B. subtilis* phage Phi29 (f), *B. subtilis* phage Phi25 (g), *B. subtilis* phage SP01 (h), *B. subtilis* phage SP02 (i) [11]; *B. subtilis* phage SPP1 (j) [12]; *Bacillus* phage vB_BsuP_Goe1 (k) [13]; *B. subtilis* phage BSP10 (m) [2]; *B. subtilis* B3 phage ARπ (l), *B. subtilis* SWV215 phage SPOσ (n), *B. subtilis* 10 phage SIOΦ(o), *B. subtilis* ATCC 6633 phage SUBω (p) [14].

The difference of neck and tail structure pose of phages has implications for the genome transfer mechanism from the phages to the host cytoplasm. Jakutytė et al. [15] reported that *Siphoviridae* SPP1 leads to depolarize the cytoplasmic membrane faster than *Podoviridae* Phi29 or *Myoviridae* SP01 during *B. subtilis* infection. The structure of virus particles also determines the characteristic of phage resistance SIOΦ and SPOσ to high temperatures, acidity, and alkaline pH [14].

Kurtboke [16] stated that phage structures obtained by electron microscopy do not typically provide detailed information on the atomic components owing to methods used for visualization of
particles. However, electron microscopy has allowed visualisation of these minuscule particles and morphological analysis. Each virion has a polyhedral, predominantly icosahedral, capsid (head) that covers the genome. Each phage species can vary both in length, width, and genome size.

The capsids are composed of many copies of one or several different proteins and have a very stable organization. A phage tail is attached to the capsid through a connector which serves as an adaptor between these two crucial components of the phage. The tail and its related structures are indispensable phage elements securing the entry of the viral nucleic acid into the host bacterium during the infectivity process. The tail serves both as a signal transmitter and subsequently as a pipeline through which DNA is delivered into the host cell during infection [16].

All types of tails have outer appendages attached to the distant end of the tail and often include a baseplate with several fibers and a tip or a needle that has specificity to the membrane receptors of the bacterium [17]. As soon the receptor has been found by the tail needle, which happens during multiple short-living reversible attachments to the bacterium, the baseplate, and tail fibers are involved in the binding of the phage to the bacterial outer membrane that makes the attachment irreversible [18].

3. The Genome Size of B. subtilis Phages
The genome size of phage describes the complexity of the structural and functional proteins. These proteins are expressed in Open Reading Frames (ORFs). Each ORFs represents a gene product (gp) that encodes structural and functional proteins. In the phage, these genes are divided into roles in the process of DNA metabolism, virion structures, packaging, lysis, and special functions that are usually called additional functions, including genes that carry pathogenic traits.

Table 1 shows that the genome size of Podoviridae is in the range of 18,000 to 20,000 bp. It is shorter than Myoviridae and Siphoviridae. Table 1 also shows that ORFs of Podoviridae have no more than 30 gene products, while the gene products of Siphoviridae is in the range of 50 to 190, and at least 200 gene products for Myoviridae. The Myoviridae family is also unique because in addition has more than 200 product genes, this family also has a genome size more than 130,000 bp, except SPOσ which not show the characteristics of protein functions through ORFs. In fact, the finding of protein functions is important for describing and knowing the specific genes that each phage has. Revisions and reconfirmations of phage genome size also required to obtain the detail of protein functions, such as SPP1 [9, 30].

The product genes belonging to DNA metabolism are functional proteins that are responsible for replication, transcription, and translation. These genes such as: recombinase, exonuclease, exonuclease, ribonucleoside-diphosphate reductase alpha subunit and beta subunit, DNA primase, DNA helicase, DNA-binding protein, helicase, ssDNA binding protein, DNA modification protein, DnaJ-like protein, DnaB replication protein. HTH binding protein, ATP-binding protein, HNH endonuclease, RNA polymerase sigma factor, tRNA-His guanylytransferase, dUTPase, XRE family transcription regulator, nucleotidytransferase, Dephospho-CoA kinase. Whereas the genes that encode the formation of phage structures include: major capsid protein, tape measure protein, head-tail adapter protein, tail protein, major tail protein, tail fiber, tail lysin, tail sheath, tail spike protein, tail tube protein tail assembly chaperone, adsorption associated tail protein, baseplate protein, baseplate J, scaffolding protein, collagen-like protein, membrane protein, coat protein.

The connector structure is built as a component of the DNA-filled capsid but remains attached to the phage tail when the viral particle is disrupted, probably due to the symmetry match between the 6-fold symmetric tail [31] and the 12-fold symmetric connector. The connector-tail interface is a critical requirement for viral DNA to find its way out of the capsid at the beginning of infection. The connector is a heterooligomer composed of several proteins [32]. Connectors carry out several functions during the phage life cycle. They participate in the packaging of dsDNA into the capsid, and later they perform the function of a gatekeeper: locking the capsid exit of the phage, preventing leakage of DNA which is under high pressure and later, after a signal transmitted by the tail indicating that the phage is attached to the bacterium, the connector will be open allowing the release of DNA into the bacterium [31].
Table 1. Genome size and characterization of *B. subtilis* phages

| Phage name      | Size (bp) | GC (%) | ORFs | tRNA | Accession Number | Family          | Reference |
|-----------------|-----------|--------|------|------|------------------|-----------------|-----------|
| Phi29           | 19,282    | 39.99  | 27   | 0    | EU771092.1       | *Podoviridae*   | [19]      |
| S10phi          | 146,698   | 39.02  | 206  | 0    | KC699836         | *Myoviridae*    | [19]      |
| PZA             | 19,366    | 39.66  | 27   | 0    | M11813           | *Podoviridae*   | [20]      |
| SPBc2           | 134,416   | 34.64  | 185  | 0    | NC_001884        | *Siphoviridae*  | [21]      |
| PhiNIT1         | 155,631   | 42.12  | 215  | 4    | NC_021856        | *Podoviridae*   | [22]      |
| SPI1            | 132,562   | 39.97  | 204  | 5    | NC_011421        | *Myoviridae*    | [23]      |
| Nf              | 18,753    | 37.32  | 27   | 0    | EU622808         | *Podoviridae*   | [23]      |
| SP10            | 143,986   | 40.49  | 236  | 0    | NC_019487        | *Myoviridae*    | [25]      |
| B103            | 18,630    | 37.66  | 17   | 0    | NC_004165        | *Podoviridae*   | [26]      |
| CampHawk        | 146,193   | 40.20  | 229  | 2    | NC_022761        | *Myoviridae*    | [27]      |
| Grass           | 156,648   | 42.25  | 242  | 3    | NC_022771        | *Myoviridae*    | [28]      |
| PM1             | 50,861    | 41.29  | 86   | 0    | NC_020883        | *Siphoviridae*  | [29]      |
| SIOΦ*           | 154,000   |        |      |      |                  | *Myoviridae*    | [14]      |
| SUBΦ*           | 154,000   |        |      |      |                  | *Myoviridae*    | [14]      |
| SPΟΦ*           | 25,000    |        |      |      |                  | *Myoviridae*    | [14]      |
| ARπ*            | 40,000    |        |      |      |                  | *Siphoviridae*  | [14]      |
| vB_BsuP_Goe1    | 18,379    | -      | 24   | 0    | KU831549         | *Podoviridae*   | [13]      |
| SPP1            | 44,016    | 43.7   | 53   | 0    | X97918.3         | *Siphoviridae*  | [30]      |
| BSP10           | 153,767   | 42.10  | 236  | 5    | MF422185         | *Myoviridae*    | [8]       |

Note: *) Needed to be confirm to the author for GC content, ORFs, and accession number

When the proteins formed the virion structure, at the same time the genes that encode the packaging process also work to do the phage assembly. These genes include: head morphogenesis protein, protease prohead, small subunit terminase, large subunit terminase, and protein portal. Kurtboke [16] stated the mechanism of packaging requires a sensor that measures the amount of DNA head filling and a nuclease that will cleave DNA as soon the head is full. Termination of the DNA packaging is coordinated with the closure of the portal system to avoid leakage of the viral genome. In tailed phages, this is most frequently achieved through the binding of head completion proteins (or adaptor proteins). After termination of the first packaging cycle initiated at pac (initiation cycle), a second packaging event is initiated at the non-encapsidated DNA end created by the headful cleavage and additional cycles of encapsidation follow.

The genes which encode holin and endolysin are the responsible genes for the lysing process of bacterial cells, both the entry process of viral genetic material and the releasing process of virus particles (Figure 2). Tailed phages achieve correctly-timed lysis by the consecutive use of endolysins and holins. Holins are small hydrophobic proteins that are encoded by the phage and inserted into cytoplasmic membrane to form membrane lesions or holes for endolysin passage. Whereas endolysins are phage-coded enzymes that break down bacterial peptidoglycan at the terminal stage of the phage reproduction cycle [33].
Figure 2. Schematic representation of how phage endolysins gain access to the PG through the most common holin-endolysin lytic system. A generalized PG structure illustrates all cleavage sites found: 1, N-acetyl-β-D-muramidase (LYSO, MURA, SLT, TRANG, GH19, GH25, and GH108); 2, N-acetyl-β-D-glucosaminidase (GH19 and GLUCO); 3, N-acetylmuramoyl-L-alanine amidase (AMI-2, AMI-3, AMI02-C, and CHAP); 4, L-alanoyl-D-glutamate (LD) endopeptidase (VANY); 5, c-D-glutamyl-m-diaminopimelic acid (DL) peptidase (PET-M23 and PET-M15-2); 6, D-Ala-m-DAP (DD) endopeptidase (PET-M23 and PET-M15-4); 7, m-DAP-m-DAP (LD) endopeptidase (PET-M23 and PET-M15-4); 8, D-alanyl-glycyld endopeptidase (CHAP); 9, D-alanyl-D-alanine peptidase (CHAP); 10, m-diaminopimelic acid-D-alanine (PET-U40). PET-M15-3, YKUD, and NLPD are peptidases with an unknown cleavage site [34].

Based on Figure 2, endolysins (or lysins) are phage-encoded lytic enzymes that break down the peptidoglycan of the bacterial cell wall during the terminal stage of the bacteriophage reproduction cycle. These enzymes compromise the mechanical strength and resistance of the cell wall that is needed to withstand the internal cytoplasmic turgor (osmotic) pressure, causing bacteriolysis and the subsequent release of the bacteriophage progeny. Based on phage lysis specificity and host recognition, several food applications have been stipulated to treat or prevent spoilage [35].

Some of the genes that encode specific proteins in phages exhibit special characteristics possessed by certain phage species, such as: glycosyltransferase, sporulation protein, acetylmuramoyl-L-alanine amidase, thymidylate synthase, dihydrofolate reductase, metal-dependent hydrolase, plasmid segregation protein, calcineurin-like phosphoesterase, exopolyphosphatase, synaptosomal complex protein, thioredoxin, flavodoxin, holiday junction resolvase, transcriptional regulator, cysteine peptidase, cytosolic protein, murein transglycosylase A.

Analysis of protein functions and identification of gene products in each species are known from 20 to 60% (data not shown). Some genes were known for their role in DNA metabolism, holin, cell lysis, virion structure and assembly processes. The protein characterization and their gene products play key roles in the bacteriophage’s application as biocontrol. As an example, the DNA encapsidation of SPP1 phage follows a processive unidirectional headful-mechanism and initiates at a unique genomic location (pac) [36]. They also cloned a fragment of SPP1 DNA containing the pac site flanked by reporter genes into B. subtilis chromosome. Infection of such cells with SPP1 will lead to a highly efficient packaging, initiated at the inserted pac site of chromosomal DNA.

Comparative studies demonstrate that bacteriophages have many common features on the molecular level and common principle of interaction with a bacterium cell, although components that
trigger adsorption of phages to the host cell and the genome release are host-dependent. Phage infection also depends on the availability of specific receptors on the cell surface, and investigation of the structure and biosynthesis of the bacterial cell membrane may be undertaken using phage-resistant mutants. Therefore, there is a need to carry out further studies on phages, identifying receptors of targeted bacteria and environmental features that affect phage activity [37]. The growing interest of the agricultural industries in phages requires more information on phage interactions, survivability and methods of their preservation.

A full knowledge of phage genome sequences is important to ensure that the phage does not carry obvious virulence factors, resistance or lysogeny genes. The identification of gene homologies requires detailed bioinformatics analysis. The latter is essential to evaluate possible complications that might arise during biocontrol and phage therapy. It was suggested that data from phage genome sequences could be used to establish a bank of “safe” therapeutic phages increasing the availability, safety, and efficacy of biocontrol [38].

4. Conclusions

There are 18 species of Bacillus subtilis phages from various region of the world until presents. They are 9 species in the Myoviridae family, 4 species in the Siphoviridae family, and 5 species in the Podoviridae family. The genome size of B. subtilis phages were known from 18,000 to 160,000 bp with different ranges of ORFs 15–250, which can be identified their protein functions between 20–60%. Based on these data, identifying all proteins should be done to ensure the phage does not bring pathogenic genes, virulence factors, resistance, or lysogeny genes.

Acknowledgement

The first author thanks the Faculty of Agriculture and Business, Satya Wacana Christian University for technical support during the writing process this paper.

References

[1] Earl A M, Losick R, Kolter R 2008 Ecology and genomics of Bacillus subtilis Trends Microbiol. 16 269–275.
[2] Te Giffel M C, Beumer R R, Leijendekkers S, Rombouts F M 1996 Incidence of Bacillus cereus and Bacillus subtilis in foods in the Netherlands Food Microbiol. 13 53–58.
[3] Bailey C P, von Holy A 1993 Bacillus spore contamination associated with commercial bread manufacture Food Microbiol. 10 287–294.
[4] Sorokulova I B, Reva O N, Smirnov V V, Pinchuk I V, Lapa S V, Urdaci M C 2003 Genetic diversity and involvement in bread spoilage of Bacillus strains isolated from flour and ropy bread Lett. Appl. Microbiol. 37 169–173.
[5] Huang Y, Xiao L, Li F, Xiao M, Lin D, Long X, Wu Z 2018 Microbial degradation of pesticide residues and an emphasis on the degradation of cypermethrin and 3-phenoxy benzoic acid: a review Mol. 23 2313 1-23.
[6] Singer A C, Shaw H, Rhodes V, Hart A 2016 Review of antimicrobial resistance in the environment and its relevance to environmental regulators Front. Microbiol. 7 1728 1-22.
[7] Grenni P, Ancona V, Caracciolo A B 2018 Ecological effects of antibiotics on natural ecosystems: A review Microchem. J. 136 25–39.
[8] Ghosh K, Seneviratne A, Kang H S, Hyun W B, Kim J E, Kim K P 2018 Complete nucleotide sequence analysis of a Novel Bacillus subtilis-infecting Bacteriophage BSP10 and its effect on poly-gamma-glutamic acid degradation Viruses 10 1–27.
[9] Gual A, Camacho A G, Alonso J C 2000 Functional Analysis of the Terminase Large Subunit, G2P, of Bacillus subtilis Bacteriophage SPP1 J. Biol. Chem. 275 35311–35319.
[10] Alonso J C, Luder G, Stiege A C, Chai S, Weise F, Trautner T A 1997 The complete nucleotide sequence and functional organization of Bacillus subtilis bacteriophage SPP1 Gene 204 201–212.
[11] Hemphil H E, Whiteley H R 1975 Bacteriophages of *Bacillus subtilis* *Bacteriol. Rev.* **39** 257–315.

[12] Baptista C D G 2009 Recognition of *Bacillus subtilis* receptors by bacteriophages SPP1–A genetic and biochemical approach Dissertation Universidade de Lisboa. Lisabon.

[13] Willms I M, Hertel R 2016 Phage vB BsuP-Goel: the smallest identified lytic phage of *Bacillus subtilis* *FEMS Microbiol. Lett.* **363** 1-7.

[14] Krasowska A, Biegalska A, Augustyniak D, Los M, Richert M, Lukaszewicz M 2015 Isolation and characterization of phages infecting *Bacillus subtilis* *BioMed Res. Int.* ID 179597 1–10.

[15] Jakutytė L, Lurz R, Baptista C, Carballido-Lopez R, São-José C, Tavares P, Daugelavičius R 2012 First steps of bacteriophage SPP1 entry into *Bacillus subtilis* *Virol.* **422** 425–434.

[16] Kurtboke I 2012 Bacteriophages InTech Rijeka.

[17] Leiman P G, Arisaka F, van Raaij M J, Kanamaru S, Itoh H R 1975 Bacteriophages *Bacteriol. Rev.* 39.

[18] Grose J H, Jensen G L, Burnett S H, Breakwell D P 2014 Correction: Genomic comparison of 93 *Bacillus* phages reveals 12 clusters, 14 singletons and remarkable diversity *BMC Genomics* **15** 1184.

[19] Paces V, Vlcek C, Urbanek P, Hostomsky Z 1986. Nucleotide sequence of the right early region of *Bacillus subtilis* phage PZA completes the 19366-bp sequence of PZA genome. Comparison with the homologous sequence of phage phi 29 *Gene* **44** 115–120.

[20] Lazarевич V, Dusterhoft A, Soldo B, Hilbert H, Mauel’ C, Karamatal D 1999 Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPpc2 *Microbiol.* **145** 1055–1067.

[21] Kimura K, Itoh Y 2003 Characterization of poly-gamma-glutamate hydrolase encoded by a bacteriophage genome: possible role in phage infection of *Bacillus subtilis* encapsulated with poly-gamma-glutamate *Appl Environ. Microbiol.* **69** 2491–2497.

[22] Kropinski A M, Hayward M, Agnew M D, Jarrell K F 2005 The genome of BCJAlc: a bacteriophage active against the alkaliphilic bacterium, *Bacillus clarkii* *Extremophiles* **9** 99–109.

[23] Castilla-Llorente V, Salas M, Meijer W J 2009 Different responses to Spo0A mediated suppression of the related *Bacillus subtilis* phages Nf and phi29 *Environ. Microbiol.* **11** 1137–1149.

[24] Yee L M, Matsumoto T, Yano K, Matsuoka S, Sadaie Y, Yoshikawa H, Asai K 2011 The genome of *Bacillus subtilis* phage SP10: a comparative analysis with phage SPO1 *Biosci. Biotechnol. Biochem.* **75** 944–952.

[25] Parent K N, Gilcrease E B, Casjens S R, Baker T S 2012 Structural evolution of the P22-like phages: comparison of Sf6 and P22 procapsid and virion architectures *Virol.* **427** 177–188.

[26] Ritz M P, Perl A L, Colquhoun J M, Chamakura K R, Kuty Everett G F 2013 Complete genome of *Bacillus subtilis* myophage camphawk *Genome Announc.* **1** 6 1-2.

[27] Miller S Y, Colquhoun J M, Perl A L, Chamakura K R, Kuty Everett G F 2013 Complete genome of *Bacillus subtilis* myophage grass *Genome Announc.* **1** 6.

[28] Umene K, Shiraishi A 2013 Complete nucleotide sequence of *Bacillus subtilis* (natto) bacteriophage PM1, a phage associated with disruption of food production *Virus Genes* **46** 524–534.

[29] Godinho L M, Fadel M E S, Monniot C, Jakutytė L, Auzat I, Labarde A, Djacem K, Oliveira L, Carballido- Lopez R, Ayora S, Tavares P 2018 The revisited genome of *Bacillus subtilis* bacteriophage SPP1 *Viruses* **10** 705 1-28.

[30] Plisson C, White H E, Auzat I, Zafarani A, São-José C, Lhuillier S, Tavares P, Orlova E V 2007 Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection *The EMBO J.* **26** 3720–3728.
[32] Lurz R, Orlova E V, Günther D, Dube P, Dröge A, Weise F, van Heel M, Tavares P 2001 Structural organization of the head-to-tail interface of a bacterial virus J. Mol. Biol. 310 1027–1037.

[33] Moak M, Molineux I J 2004 Peptidoglycan hydrolytic activities associated with bacteriophage virions Mol Microbiol. 51 1169–1183.

[34] Oliveira H, Melo L D R, Santos S B, Nóbrega F L, Ferreira E C, Cerca N, Azeredo J, Kluskens L D 2013 Molecular Aspects and Comparative Genomics of Bacteriophage Endolysins J. Virol. 87 4558–4570.

[35] Oliveira H, Lavigne R, Azeredo J, Kluskens L D 2012 Bacteriophage endolysins as a response to emerging foodborne pathogens Trends Food Sci. Technol. 28 103–115.

[36] Bravo A, Alonso J C, Trautner T A 1990 Functional analysis of the Bacillus subtilis bacteriophage SPP1 pac site Nucleic Acids Res. 18 2881–2886.

[37] Jończyk E, Klak M, Międzybrodzki R, Górski A 2011 The influence of external factors on bacteriophages Folia Microbiol. 56 191–200.

[38] Petty N K, Evans T J, Fineran P C, Salmond G P C 2007 Biotechnological exploitation of bacteriophage research Trends Biotechnol. 25 1 7-15.