Genomic analysis of Bacillus subtilis lytic bacteriophage ONIT1 capable of obstructing natto fermentation carrying genes for the capsule-lytic soluble enzymes poly-β-glutamate hydrolase and levanase.

Ozaki Tatsuro, Abe Naoki, Kimura Keitarou, Suzuki Atsuto, Kaneko Jun

Bioscience, Biotechnology, and Biochemistry
Volume 81, Number 1, Page 135-146, 2016-09-22
URL: http://hdl.handle.net/10097/00122398

Genomic analysis of *Bacillus subtilis* lytic bacteriophage φNIT1 capable of obstructing natto fermentation carrying genes for the capsule-lytic soluble enzymes poly-γ-glutamate hydrolase and levanase.

Tatsuro Ozaki¹, Naoki Abe¹, Keitarou Kimura², Atsuto Suzuki¹ and Jun Kaneko¹†

¹ Department of Microbial Biotechnology, Graduate School of Agricultural Science, Tohoku University, Tsutsumi-dori Amamiya-machi 1-1, Aoba-ku, Sendai 981-8555, Japan
² Laboratory of Applied Microbiology, Food Research Institute-National Agriculture and Food Research Organization (NFRRI-NARO), Tsukuba 305-8642, Japan
† Corresponding author, E-mail: Jun.kaneko.b6@tohoku.ac.jp

Abstract

*Bacillus subtilis* strains including the fermented soybean (natto) starter produce capsular polymers consisting of poly-γ-glutamate and levan. Capsular polymers may protect the cells from phage infection. However, bacteriophage φNIT1 carries a γ-PGA hydrolase gene (*pghP*) that help it to counteract the host cell’s protection strategy. φNIT had a linear double stranded DNA genome of 155,631-bp with a terminal redundancy of 5,103-bp, containing a gene encoding an active levan hydrolase. These capsule-lytic enzyme genes were located in the possible foreign gene cluster regions between central core and terminal redundant regions, and were expressed at the late phase of the phage lytic cycle. All tested natto origin *Spounavirinae* phages carried both genes for capsule degrading enzymes similar to φNIT1. A comparative genomic analysis revealed the diversity among φNIT1 and *Bacillus* phages carrying *pghP*-like and levan-hydrolase genes, and provides novel understanding on the acquisition mechanism of these enzymatic genes.

Introduction

Phage infection starts by adsorption of phage particles into the bacterial cells through the recognition of phage receptors on the host cell surface. Production of extracellular polymers such as polysaccharides has been known as one of the strategy of bacteria to avoid phage adsorption by covering cell surface phage receptors. For instance, capsular polysaccharides act as a barrier to phage infection in several bacteria.¹⁻⁴

However, phages have evolved strategies to infect bacteria covered with capsules and biofilms.⁵ Numerous phages possess capsular polysaccharide degrading enzymes associated with phage virion.⁶⁻⁹ The *Escherichia coli* K1 strain produces the polysialic acid capsule known as the K1 antigen, on the surface of the host cell. The K1 antigen covers the LPS which is directly recognized by bacteriophage T7, hence the capsule of strain K1 acts as a barrier to T7 infection.¹⁰ In contrast, *E. coli* T7-like phages K1E and K1F recognize K1 antigen via their tail structural protein, and the endosialidase associated with the tail structural protein of the phages helps the phage approach the cell surface by degrading the K1 antigen.⁷ The genomic analysis of K1E suggested that acquisition of an endosialidase domain in the tail structural protein allowed K1E and other K1 phages.⁹

Some *Bacillus* species, including *Bacillus anthracis* and the *Bacillus subtilis* strains for natto (fermented soybeans) starter, produce a capsular polymer, poly-γ-glutamate (γ-PGA; poly glutamate with a γ-linkage), and levan.¹¹ The composition of γ-PGA is different among the species, e.g., *Bacillus anthracis* has only D-glutamate polymer,¹² while γ-PGA of *Bacillus subtilis* consists of a mixture of a polymer of glutamate enantiomers.⁸ In *B.
Japan

in the production, enzymes, the host analysis of phage Miyagino, Takahashi, Naruse, and Asahikawa were used.

Bacillus natto products with suspected phage contamination. P

Bacteria

suspected phage contamination was also investigated. The

located in one of

locate regions (TR), and possible foreign gene clusters (FGCs)

sequence of pghP

allow eliminating random cleaves both D

non

PghP

overcome the barrier.

γ

because of absence of Anthracis,

- -

PGA is anchored to the host cell

PGA

is

not covalently

PGA

by

membrane associated CapD

produce γ-PGA by PgsABCE, but their

γ-PGA is not covalently-associated with peptidoglycan,

of absence of a CapD ortholog.

Despite whether γ-PGA is anchored to the host cell surface, the γ-PGA

capsid plays an important role in prevention of phage

infection of host cells.

Some Bacillus phages possess unique strategies to overcome the barrier. Lytic bacteriophage φNIT1 has been isolated from natto product with suspected phage

contamination. φNIT1 carries a gene for the soluble γ-PGA hydrolase, pghP, on a 3.2-kb EcoRV fragment of its genome. PghP participates in the phage infection, in spite of non-constituent of with phage virion. PghP internally and randomly cleaves both D- and L-γ-PGA substrates, eliminating the γ-PGA capsid from the cells, thereby allowing the progeny phages to approach the receptors on the host cell surface. Because PghP is not homologous with poly-γ-glutamate hydrolases of B. subtilis and other bacteria, the origin and acquisition mechanism of the pghP gene is of interest. A possible transposase gene is found downstream of pghP, but further information is required before its function can be clarified.

Herein, we describe the determination of whole genome sequence of φNIT1. It comprised a linear double stranded DNA with a central core region (CR), terminal redundancy regions (TR), and possible foreign gene clusters (FGCs) located between CR and both TRs. The pghP gene was located in one of the FGCs, and the gene for the novel levan-degrading enzyme LevP was found in another FGC. The existence of the pghP and levP genes among the nine phages isolated from independent natto products with suspected phage contamination was also investigated.

Materials and methods

Bacteria and phage strains, plasmids and media. Lytic bacteriophages, φNIT1, FSG, KKP, MOP, ONPC, ONPB, P-1, SUP, SS2P and THP, were isolated from independent natto products with suspected phage contamination. Bacillus typing phage SP50 was obtained from Ackermann et al. Bacillus subtilis (natto) NAFM5 was used as a propagation and experimental phage indicator strain, and strains Miyagino, Takahashi, Naruse, and Asahikawa were used for the host analysis of phages. B. subtilis strains 1046, 1149, 1174, 1243, 1346, 1395, and 1417 were also used for host analysis of phages. For the expression of recombinant enzymes, E. coli BL21 (DE3) cells harboring appropriate plasmids were used. Unless otherwise noted, B. subtilis and E. coli strains were cultured in LB medium. For the capsule production, B. subtilis (natto) strain Takahashi was cultured in the GSP medium (1.5% L-glutamic acid mono potassium salt (Nacalai, Japan), 1.5% sucrose, and 1.5% polyepeton-S (a polypeptone derived from soybean; Nihon Pharmaceutical, Japan)).

Phage techniques. Phages isolated from a single plaque on the lawn of the indicator strain were amplified and used for the experiments. The phage particles were purified from liquid culture by the method described previously.

Briefly, 10⁶ pfu of phages (100μl) were added to 100 ml of host cell culture at early log phase (OD₆₅₀ = 0.2 – 0.4), and were incubated at 37°C with vigorous shaking. When OD₆₅₀ decreased to 0.1, culture was centrifuged at 5,000 x g for 10 min at 4°C to remove cell debris. Sodium chloride and polyethylene glycol 6000 were added to the supernatant at a final concentration of 0.7 M and 10% (w/v), respectively. After incubation at 4°C for 12h, the reaction mixture was centrifuged at 5,000 x g for 20 min. at 4°C, and the pellet containing phage particles was suspended into 2 ml of SMC buffer (50 mM Tris-HCL buffer (pH 7.0), containing 0.1 M NaCl, 8 mM MgSO₄, 1 mM CaCl₂, and 0.01% gelatin), then filtered through 0.20 μm cellulose acetate (Advantec, Japan). For further purification, the phage suspension was treated with DNase I and RNase A at room temperature for 1 h, and then, subjected to CsCl density gradient centrifugation. A white and opaque phage band was collected, and pelleted by centrifugation at 60,000 x g for 3h. The obtained pellet was resuspended in SMC buffer, and the genomic DNA from the phage particles was prepared according to the method described previously.

A spot test method was used for the host range analysis. Indicator strains cultured in LB medium until the OD at 660 nm reached 0.4 were mixed with LB medium containing 0.8% soft agar, and then overlaid on the LB plate. A total of 10 μl of diluted phage particles were spotted on the soft agar, and incubated at 37°C.

For the one-step growth analysis, a log-phase culture of NAFM5 at 37°C in LB medium was used as a host, and phages were added at multiplicity of infection (MOI) of approximately 4. Cells were separated from free phages by centrifugation, and then cultured in new LB media at 37°C with vigorous shaking. Aliquots were taken sequentially from the culture and filtered through 0.45 μm filter, and then the plaque forming units were counted by the soft agar overlay method.

Electron microscopic examination. Phage particles were negatively stained with 2% sodium phosphotungstic acid (pH = 7.4) and observed by using a transmission electron microscope (H-8100, Hitachi).

Phage genomic DNA sequencing. The DNA from φNIT1 particles was prepared as described previously. The shotgun approach was used to determine the nucleotide sequence of the whole genomic DNA of φNIT1. Phage genomic DNA was digested by sonication, and the resulting 1.5 to 3.0 kb fragments were cloned into pUC119 to make the shotgun genomic library. Gaps of contiguous fragments and flanking regions were amplified by PCR and inverse PCR, respectively. The cycle sequencing reaction and sequencing were done by the method described previously using an ABI 310 (Applied Biosystems, CA) automated sequencer.
Computer analysis. The obtained sequences were assembled using Genetyx Mac ATSQ software (Genetyx, Tokyo, Japan). ORFs of more than 150 bp, with ATG, GTG, and TTG as the initiation codons, were predicted for all six frames by the Genetyx Mac program package (Genetyx). A homology search was performed using FASTA and BLAST programs against sequences present in the DDBJ/EMBL/GenBank nucleotide sequence database and the SWISSPROT/ NBRF-PIR protein database. Transfer RNAs were identified by tRNAscan SE 1.21 (http://lowelab.ucsc.edu/tRNAscan-SE/). Promoters corresponding to bacterial sigma 70 were searched using BPROM programs available at the SoftBerry web site (http://linux1.softberry.com/berry.phtml). A possible promoter consensus sequence (TGGGAC/t – 17 to 19-bp – (t/c/aATAAT) for host sigma A was also investigated by using the Genetyx program. Possible ϕ-independent terminators were searched by using ARNold programs available at the University Paris-Nord 11 site (http://rna.igmours.u-psud.fr/toolbox/arnold/index.php). Comparison analysis of genomes among ϕNIT1 and related phages was performed by using the GenomeMatcher program.  

Northern blotting and determination of the transcription initiation site. Strain NAFM5 cells were grown in LB medium. At early exponential phase (OD660 = 0.4), ϕNIT1 was infected with a moi of approximately 4. Aliquots of culture were collected sequentially, and phages in the supernatant and host cells were separated by centrifugation. Plaque forming units from the supernatant were counted by the overlay method. The total RNA fraction was prepared from the cells by phenol extraction (equilibrated in 20 mM sodium acetate (pH 5.5), 0.5% SDS, and 1 mM EDTA) followed by treatment with DNase I. A total of 3 µg of total RNA preparation was separated on a denaturing 1.5% agarose gel and blotted onto a nylon filter. Northern blotting and hybridization were done by standard methods using the PCR products of each gene as probes described in the legend of Fig. 6. Probe labeling and detection were performed using the DIG system (Roche, Basel, Switzerland).

Determination of the transcription initiation sites of *pgbP* and *levP* was carried out by using the 5'-Full RACE core kit (Takara Bio, Shiga, Japan). As a template, the total RNA purified from infected cells by using the RNeasy Protect Bacteria Mini Kit (Qiagen, Venlo, Netherlands) was treated twice with DNase to remove the remaining DNA. Reverse transcription and 5'-Race were performed according to the manufacturer’s instructions.

Cloning of the levanase gene and measurement of the enzyme activity of the recombinant enzyme.

The putative levanase gene was amplified by PCR by using the ϕNIT1 genome as a template. The amplified DNA fragment was digested with *Nde*I and *Bam*HI, and cloned into the corresponding sites of the pET15b expression vector. The resulting plasmid was then introduced into *E. coli* BL21(DE3), and the transformant was cultured in 100 ml of LB medium containing 100 µg/ml ampicillin at 37°C. When the optimal density at 660 nm reached 0.8, isopropyl-b-D-thiogalactopyranoside was added to a final concentration of 1 mM. After incubation for an additional 3h at 37 °C, the cells were harvested by centrifugation at 4 °C, and resuspended in 50 ml of 20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl and 20 mM imidazole (binding buffer). Cells were disrupted by sonication, and unbroken cells and large debris were removed by centrifugation at 16,000 x g for 15 min at 4 °C. The supernatant was dialyzed twice against binding buffer, and filtered by 0.20 µm cellulose acetate filter. The resulting recombinant His6-tagged levanase was loaded onto a HisTrap HP (GE Healthcare, UK), and eluted with a linear gradient of 20 to 500 mM of imidazole. The eluate was dialyzed twice against more than a 100-fold volume of 20 mM potassium phosphate buffer (pH 5.5).

The levanase activity of purified recombinant LevP (rLevP) was evaluated by the analysis of product oligo-saccharides using thin-layer chromatography. The purified rLevP preparation (final concentration of 200 ng/µl) was mixed with an equal volume of 1% levans (from Serratia levanicum, Wako) in 20 mM potassium phosphate buffer (pH 5.5) and incubated at 37°C for 10, 30, 60 and 120 min. After denaturing the enzyme by heating at 98°C for 5 min, 5 µl of the reaction supernatant was applied onto a silica gel 60 F 254 TLC plate (Merck). The TLC plate was developed with a solvent system of chloroform : methanol : acetic acid : water (5 : 12 : 1 : 3, v/v/v/v) and air-dried. After spraying 50% (v/v) sulfuric acid, the plate was dried and placed in a 110°C oven for 10 min to visualize the reaction spots.

Accession number. The genomic sequence of ϕNIT1 was assigned accession number AP013029 in the DDBJ/GenBank/EMBL nucleotide sequence libraries.

Results and Discussion

**General features of ϕNIT1, and its genomic structure**

ϕNIT1 has typical morphology of *Myoviridae* phage with an isometric head of about 100 nm in diameter and a contractile tail 250 nm long (Fig. 1A and B). The complete nucleotide sequence of the ϕNIT1 genome encased in the phage virion was determined to consist of a linear double

![A](image.png) ![B](image.png)

Fig. 1. Electron micrographs of phage ϕNIT1 in a native state (A) and contracted state (B), negatively stained with phosphotungstic acid.
Fig. 2. Gene organization of φNIT1. Predicted orfs are indicated by an arrow showing the direction of transcription. Open reading frames (orfs) are shown in arrows. Genes predicted as major functional proteins are shaded, and genes homologous to those of Listeria phage A511 are surrounded by heavy line. tRNAs are shown as small arrows. Upper and lower lines show the transcription units of sense and antisense strands, respectively. Possible promoters and putative rho-independent terminators are indicated by bent arrow and a lollipop, respectively. Terminal redundancy (TR) sequences in both ends are indicated as rounded squares.
stranded DNA of 155,631-bp, with a terminal redundancy (TR) of 5,103-bp, similar to that found in SPO1-related group of Myoviridae of Ackermann’s classification of tailed bacteriophages (Fig. 2). A part of the genome containing 6,230-bp a gene for poly-gamma-glutamate hydrolase of a phage (pghP) determined previously (AB469692) was confirmed. The total G+C content of the φNIT1 genome was 42.1%, which was similar to that of the host B. subtilis (natto) chromosomal DNA, as well as other B. subtilis bacteriophages.

To confirm whether the structure of the genome in the capsid is circular or linear, the electrophoretic and in silico restriction maps of the φNIT1 genome were compared. Among the 10 restriction enzymes investigated, AccII and EcoRI provided digested fragments having lengths different from that of the in silico digestion of the circular genome. Addititionally, ligation of the genome before digestion and heat treatment of the digested DNA fragments before electrophoresis did not alter the fragment patterns, which suggested that φNIT1 has a linear genome without cohesive ends. Therefore, the 1.6 and 3.5-kb EcoRI fragments that might contain the ends of genome were extracted from the agarose gel, and both ends of these fragments were determined by run-off sequencing using the corresponding internal primers. As expected, the TR sequences of 5,103-bp in both ends of the linear genome were determined, and an EcoRI site was located at 1,625-bp and 3,478-bp from the left and right ends of the TR, respectively. Bacillus myophage SPO1 (25) and Listeria myophage A511 (26) are known to possess the TRs of 13,185-bp and 3,125-bp in their genome respectively, while the TR of φNIT1 was different from them in terms of its length and nucleotide sequence. Although a number of orfs of Staphylococcus myophage Twort (27) and Enterococcus myophage φEF24C (28) showed homology and synteny with those of phages A511 and φNIT1, they have been reported as circular genomes.

Seventy-one ORFs in the φNIT1 genome, which belong to head, tail morphogenesis and DNA replication modules, shared synteny and gene similarity with Listeria phage A511 (29) indicating φNIT1 belongs to subfamily
\textit{Spounavirinae of Myoviridae}. Similar to other \textit{Spounavirinae} phages, four putative tRNA genes were found upstream of the head morphological gene cluster. The central approximately 100-kbp region consists of 116 \textit{orf}s for morphogenesis, DNA replication and transcription genes, suggesting that this region is involving the phage life-cycle. In addition, all genes homologous to phage A511 genes were located within this region. Therefore we designated it as a core region (Fig. 2).

After the \textit{φNIT1} genome was submitted to the DDBJ /GenBank/EMBL databases, the genome of two \textit{φNIT1} related \textit{Bacillus} myophages, phage Grass (KF669652) and \textit{φSPG24} (AB930182) carrying \textit{pghP} became available. The genome organization of \textit{φSPG24} was almost the same as that of \textit{φNIT1}, and almost all open reading frames for the core region were conserved among these \textit{pghP}-carrying phages. However, phage Grass did not possess the \textit{levP} gene (Fig. 3A and B). Phage Grass has a TR sequence of 4208 bp, whereas the existence of a TR in \textit{φSPG24} has not been investigated.

Structural proteins of the purified \textit{φNIT1} phage virion were separated by SDS-PAGE (Fig. 4A). The N-terminal amino acid sequence of the 48 kDa protein was determined as SFTTG, which corresponded to the 31st to 35th residues of the deduced amino acid sequence of a putative capsid gene (orf481), indicating that the peptide bond between K30 and S31 was digested during maturation. The molecular mass of resulting mature capsid protein was calculated as 48,805 Da, which was in agreement with the electrophoresis results. Parker and Eiserling reported that an approximately 2-kDa N-terminal fragment of a major capsid protein of phage SPO1 was proteolytically cleaved during phage head maturation. Among the \textit{Spounavirinae} phages, the pro-capsid proteins of \textit{Listeria} phage A511, \textit{Staphylococcus} phages K and Twort, and \textit{Lactobacillus} phage LP65 are cleaved between K and S, the same cleavage site as \textit{φNIT1}, whereas the cleavage site of \textit{Bacillus} phage SPO1 pro-capsid is K-A. The N-terminal amino acid sequence of the 60-kDa protein was AQSYG corresponding to the 2nd to 6th amino acids of ORF569, which has 49% and 44% identity with the tail sheath proteins of phage A511 (gp93) and phage K (ORF49), respectively. Phylogenetic analyses of the capsid and tail sheath proteins among sequenced \textit{Spounavirinae} phages showed that \textit{φNIT1} (and related phages Grass and \textit{φSPG24}) were placed independently from the phages of the genus \textit{Twortlikevirus} (Fig. 4B).

Two genes for capsule-lytic enzymes located at possible foreign gene clusters in the phage genome

Between the core region and both TRs of the \textit{φNIT1} genome, the regions containing nonfunctional gene fragments and some atypical phage genes, were found (Fig. 2). Between the core region and both TRs of the \textit{φNIT1} genome, the two regions containing some atypical phage genes, were found (Fig. 2). The average length of ORFs in these regions (143.17 amino acids) was significantly shorter than that in core region (248.4 amino acids), and a number of genes were presumed to be defective fragments. BLAST analysis of the deduced amino acid sequences of ORFs in these regions revealed that 50% (18/36) and 54% (29/54) ORFs were homologous to \textit{Bacillus} species and other Gram-positive bacterial genes, whereas 73% (85/116) ORFs in core region were typical phage genes. These facts suggest that genes or gene fragments in these regions have been acquired from host organisms during phage infection cycles. Therefore we classified these regions as possible foreign gene clusters (FGCs). Possible FGCs of several dozen-kbp in size were also observed in the genomes of other \textit{Spounavirinae} phages, SPO1, A511, \textit{φEA24C}, \textit{φTwort}, \textit{φParker}, and \textit{φLP65}. Only a few possible functional genes are found in the FGCs, suggesting genes encoding products unnecessary to phage growth might be disrupted.

Interestingly, \textit{pghP} and a putative levanase gene (orf477) encoding a glycosyl hydrolase (GH) family 32 protein, which was designated as \textit{LevP}, were found independently in different FGCs (Figs 2 and 3). Two \textit{φNIT1} related phages Gross (KF669652) and \textit{φSPG24} (AB930182) possess \textit{pghP} in the same position of their FGCs as \textit{φNIT1} (Fig. 3A and B). A possible SI3141-like insertion sequence (IS) sandwiched between two 7-bp direct repeats of AATAGG, was found at downstream of \textit{pghP} of \textit{φNIT1} (Figs. 2 and 3). However, this IS was absent in other
Expression of *pghP* and *levP* genes during the phage lytic cycle.

To analyze the timing of expression of the enzyme genes in the phage lytic cycle, a one-step growth experiment of *φNIT1* was performed. Complete cell lysis occurred within 45 min after infection (Fig. 6A). The burst size of *φNIT1* was calculated as approximately 50, and this value was almost the same as that of *Listeria* phage A511.36

Therefore, the expression of *pghP*, *levP*, capsid, and holin genes was then analyzed sequentially from 5 to 35 min after phage infection by northern blotting. In phage A511, the mRNA for the capsid gene, one of the phage structural proteins expressed in late phase of the lytic cycle, was detected at 20 min after infection to *L. monocytogenes.*36 The 1.5-Kb transcript for the *φNIT1* capsid protein was initially detected 20 min after infection, then it peaked after 25 min (Fig. 6B). In *φNIT1*, putative genes for a lytic enzyme and holin, a small protein that accumulates at programmed times in the phage lytic cycle to permeabilize the folded endolysin, were located distantly in its genome similar to those of *Myoviridae* phage T4*35* and *Enterococcus* phage *φEF24C,*27 and expressed at late phase (from 25 min after infection). If the *pghP* and *levP* genes are bacterial in origin, and were controlled under bacterial housekeeping sigma A, their expression might be observed immediately after infection. However, *pghP* (approximately 2.9 kb) and *levP* (approximately 1.5 kb) were expressed synchronously, and their expression patterns were similar to that of holin (Fig. 6B).

---

φNIT1-like phages, where an AAATAGG sequence remained at the corresponding site instead of the IS (Fig. 3B). In addition, this IS was found in other position of the *φSPG24* genome. These findings suggest that this IS is irrelevant to *pghP* acquisition.

During preparation of this manuscript, existence of a possible levanesan gene corresponding to *levP* in *φSPG24* was reported.35 Masrufi and Levesque suggested that the *φSPG24* gene corresponding to *orf477* encodes a possible levansucrase by phylogenetic analysis, molecular docking and secretability predictions. However, its enzymatic activity has not been investigated. The *levP* gene was located at the same position in *φNIT1* and *φSPG24* phages, whereas two unknown genes were found instead of *levP* at the same position of the phage Grass genome (Fig. 3A and B). The levan degrading activity was detected in the supernatant of a *φNIT1*-infected NAFM5 lysate, while a cell lysate without phage infection had no levanesan activity. BLAST and domain analyses of the deduced amino acid sequence of *orf477* showed that ORF447 belongs to GH family 32 β-fructosidase (from the conserved domain database of the National Center for Biotechnology Information) with homology to putative levanesans of Gram positive bacteria. All three active site amino acids for beta-fructofuranosidase activity, two aspartic acid and a glutamic acid residues, were conserved as D19, D150 and E198 of ORF447.

We assessed the levan hydrolyzing activity of ORF477 using recombinant protein. ORF447 was expressed in *E. coli* BL21(DE3) as a His6-tagged fusion protein, and a recombinant His6-ORF447 protein was purified homogenously (Fig. 5A). The levan hydrolyzing activity of His6-ORF447 was then confirmed by thin-layer chromatographic analysis of the reaction products from levan (Fig. 5B). After 60 min incubation, the spots of fructose and various lengths of fructo-oligosaccharides were observed, indicating that ORF477 acts as an endo-levanase. Thus, we designated the product of *orf477* as LevP, a levanase of phage.

![Fig. 5. Recombinant LevP preparation and its levan degradation activity.](image)

Expression of *pghP* and *levP* genes during the phage lytic cycle.

To analyze the timing of expression of the enzyme genes in the phage lytic cycle, a one-step growth experiment of *φNIT1* was performed. Complete cell lysis occurred within 45 min after infection (Fig. 6A). The burst size of *φNIT1* was calculated as approximately 50, and this value was almost the same as that of *Listeria* phage A511.36

Therefore, the expression of *pghP*, *levP*, capsid, and holin genes was then analyzed sequentially from 5 to 35 min after phage infection by northern blotting. In phage A511, the mRNA for the capsid gene, one of the phage structural proteins expressed in late phase of the lytic cycle, was detected at 20 min after infection to *L. monocytogenes.*36 The 1.5-Kb transcript for the *φNIT1* capsid protein was initially detected 20 min after infection, then it peaked after 25 min (Fig. 6B). In *φNIT1*, putative genes for a lytic enzyme and holin, a small protein that accumulates at programmed times in the phage lytic cycle to permeabilize the folded endolysin, were located distantly in its genome similar to those of *Myoviridae* phage T4*35* and *Enterococcus* phage *φEF24C,*27 and expressed at late phase (from 25 min after infection). If the *pghP* and *levP* genes are bacterial in origin, and were controlled under bacterial housekeeping sigma A, their expression might be observed immediately after infection. However, *pghP* (approximately 2.9 kb) and *levP* (approximately 1.5 kb) were expressed synchronously, and their expression patterns were similar to that of holin (Fig. 6B).

![Fig. 6. Expression of *pghP* and *levP* during the phage cycle.](image)
Hence, the transcription initiation sites of pghP and levP were determined from analysis of the promoter regions. The 2.9-kb mRNA for pghP was polycistronic, and levP was transcribed as a monocistronic mRNA of 1.5 kb. These results were supported by northern analysis data. No significant sigma A promoter consensus sequences were found in the promoter regions of pghP and levP. A common possible consensus sequence AAaGa/tAT was found in the promoter regions of pghP and levP, suggesting that a phage dependent expression control mechanism for initiation sites, phage infection plays an important role for abnormal natto. Nagai and Yamasaki suggested that phages isolated from abnormal natto belong to Myoviridae. Transmission electron microscopy revealed that nine B. subtilis (natto) Myoviridae phages isolated from independent abnormal natto products in Japan, FSG, KKP, MOP, ONPC, ONPB, P-1, SUP, SS2P, and THP have φNIT1-like hexagonal head and contractile tail. These phages isolated from B. subtilis (natto) starter strains, whereas they did not form plaque on the lawn of B. subtilis laboratory strains. Preliminary Southern blot analysis showed that all the phages carry both pghP and levP genes (data not shown). Sequence analysis and the polycistronic structure of pghP and the flanking regions including upstream of orf587 were highly conserved (more than 97% identity at the nucleotide level) among all of these 10 B. subtilis (natto) origin myophages indicating φNIT1. The nucleotide sequences of the 2.3-kb PCR fragments including the levP gene of B. subtilis (natto) phages were also highly conserved (more than 99% identity at the nucleotide level). Additionally, restriction mapping and major capsid sequencing of these phages indicated that φNIT1 and these pghP/levP-carrying Japanese phages were closely related to each other. We found that Bacillus typing phage SP50 infected both B. subtilis indicator and natto starter strains, and carried only the pghP gene at the same position of φNIT1 (Table 1 and Fig. 3B). Sequence analyses of the orfs for the capsid and tail sheath proteins suggest that these B. subtilis (natto) phages and SP50 are closely related to φNIT1.

### Table 1. Host ranges of Bacillus subtilis (natto) infective phages

<table>
<thead>
<tr>
<th>Strains</th>
<th>Phages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φNIT1</td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>NAFM5</td>
<td>+++</td>
</tr>
<tr>
<td>Asahikawa</td>
<td>+++</td>
</tr>
<tr>
<td>Miyagino</td>
<td>+++</td>
</tr>
<tr>
<td>Naruse</td>
<td>+++</td>
</tr>
<tr>
<td>Takahashi</td>
<td>+</td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
<tr>
<td>1046</td>
<td></td>
</tr>
<tr>
<td>1149</td>
<td></td>
</tr>
<tr>
<td>1174</td>
<td></td>
</tr>
<tr>
<td>1243</td>
<td></td>
</tr>
<tr>
<td>1346</td>
<td></td>
</tr>
<tr>
<td>1395</td>
<td>+</td>
</tr>
<tr>
<td>1417</td>
<td>+</td>
</tr>
</tbody>
</table>

Plaque formation of each phage was tested by spotting phage suspensions of different concentrations on the soft agar with indicator strains. +, ++, and +++ indicates that plaque formation was observed when 10³, 10⁴, and 10⁵ pfu of phage suspension was spotted. Nd: not done.
Fig. 8. Analyses of the γ-PGA hydrolases encoded in the genome of *Bacillus* phages.

All protein sequences were obtained from the DDBJ/GenBank/EMBL databases. Alignment and tree construction were performed using ClustalW. (A) Comparison of the deduced amino acid sequences of γ-PGA hydrolases encoded in the genome of *Bacillus* phages and bacteria. The zinc binding motif and phosphate binding motif are indicated as black and open boxes, respectively. Amino acids identical with PghP are shaded. Bsu ; *Bacillus subtilis* strain 168, 45) Bpu ; *Bacillus pumilus* hypothetic protein (accession No. WP_012009647), Bli ; *Bacillus Licheniformis* hypothetic protein (accession No. AAU24062), Sha ; *Staphylococcus haemolyticus* hypothetic protein (accession No. WP_011276332).

(B) Phylogenetic tree of γ-PGA hydrolases of *Bacillus* phages and bacteria.
Fig. 8. (Continued)

Fig. 9. Outline of the genome of *Bacillus* phages carrying genes for capsule-hydrolyzing enzymes.  
(A) Comparative organization of genome structures of *Bacillus* phages carrying genes for capsule-hydrolyzing enzymes. Predicted modules are indicated by boxes. Positions of *pghP*-like and *levP*-like genes are indicated as arrows.  
(B) Details of gene structure around the *pghP*-like and *levP*-like genes. Genes are indicated as arrows, and conserved regions are connected by gray shading. Identity values (%) of the predicted gene products are indicated.
In addition to φNIT1 and its related phages, other phages carrying a possible poly-γ-glutamate hydrodase gene on their genomes have been reported. Siphoviridae bacteriophage PM1, which has been isolated from an abnormal natto product, is carrying a gene coding for a possible poly-γ-glutamate hydrodase. BLAST analysis revealed that a related Bacillus siphophage SPP1 (NC_004166) also carries a different gene encoding a PghP-like protein. The deduced amino acid sequence of the PghP-like proteins of PM1 and SPP1 showed 51% and 42% identity with PghP, respectively (Fig. 8A). Most recently, pghP-like genes in the genomes of some Bacillus Myoviridae phages, such as AGATE (NC_020081) and Bp8p-c (NC_029121) were reported, and a pghP-like gene was also found in the related phage Bobb (NC_024792) by BLAST analysis. The crystal structure of PghP indicated that a phosphate binding motif (His78-Thr9-Ser81), zinc binding motif (His40-Glu45-His103) and Glu249 are important for its enzymatic catalytic properties. These residues were completely conserved among all of these PghP-like proteins (Fig. 8A). It is worthwhile to note that the deduced amino acid sequences of PghP-like proteins in these myophages showed around 50% identity with PghP, and those of the Siphoviridae phages (Fig. 8A). As the result, phage encoded PghP and PghP-like proteins were classified into four groups (Fig. 8B). Thus, the groups of pghP-carrying myophages were divided into two groups, coincident with the grouping of morphological protein groups indicated in Fig. 3B.

Recently, existence of four genes similar to pghp in the genome of B. subtilis strain 168 was reported. Two of them are located near the phage-related region 5, and others are flanked by PBSX and SPβ prophage elements, suggesting ancestral temperate phages of these elements may carry pghp-like gene. PghP related proteins were also found in the genome of several bacteria, but these PghP-like proteins were classified into a different group from the PghP and PghP-like proteins of myophages except for that of temperate phage SP10 (Fig. 8A and B).

The genomic structure of the Bacillus Myoviridae phages phiAGATE and Bp8p-c carrying a PghP-like gene was similar to that of φNIT1, but the existence of TRs in their genomes was not investigated (Fig. 9A). The upstream orf587-like gene and pgP-like gene were thought to be transcribed into poly cistronic mRNA similar to φNIT1 (Fig. 8B). In contrast, the genome structure of another myophage SP10 was different from those of φNIT1 and Bp8p-c (Fig. 9A). Genes of pghP-like and levP-like proteins were found in the possible FGCs, but the structure of their transcriptional units differed from those of φNIT1 (Fig. 9B). The orf587-like orf was absent at the flanking region of the pghP-like gene in SP1, and this structure was rather close to that of Siphoviridae PM1 (Fig. 9B).

The variations of primary and genetic structures among phage-carrying PghP, LevP and related proteins suggest that the genes encoding these proteins originated from different ancestor and/or evolved independently after acquisition by phages. B. subtilis produces different types of biofilms, such as colony biofilms, floating biofilms (robust pellicles) and submerged surface-adhered biofilms, depending on the culture condition. ɣ-PGA is one of the components in the floating biofilms, but its production is not required for floating biofilms formation. Conversely, ɣ-PGA production is involved in mucoid colony morphology, and enhances submerged biofilm formation. Recently, it was shown that levan can be a structural and possibly stabilizing component of B. subtilis floating biofilms. Therefore, under certain conditions, having ɣ-PGA in the biofilm may be enough on its own to defend against phage infection, and levan may play an auxiliary role as a barrier. B. subtilis (natto) produces ɣ-polyglutamate and levan as the extracellular capsule polymers, and they are thought to act as a barrier to phage adsorption to the cell. In fact, removal of capsular ɣ-PGA from the surface of the host cell by the ɣ-polyglutamate hydrodase activity of PghP in a supernatant of φNIT1-infected NAFM5 lysate is important for the effective infection of encapsulated cell with φNIT1. Namely, elimination of ɣ-PGA by PghP allows progeny phages approach to their receptors on the host cell. Existence of phages carrying only pghP (phages Grass and SP50) and pghP-like (phages AGATE and Bp8p-c) genes suggests their ɣ-PGA hydrolyzing enzymes are beneficial for the phage infection.

Unlike phages carrying a polysaccharide depolymerase on their virions, φNIT1 requires the initial infection of the host cells to produce the capsule-hydrolyzing enzymes. As no signal sequence was contained in the deduced amino acid sequences of PghP and LevP, these enzymes are intracellularly produced by infected cells, and should be released by cell lysis concomitantly with the release of phage progenies to depolymerize the capsule polymers. In contrast to polyglutamate capsule of B. anthrasis, the capsule of B. subtilis (natto) is not anchored to the peptidoglycan such some B. subtilis (natto) cells in the population may remain as naked cells allowing phage infection. Once φNIT1 successfully infects the host cell, PghP and LevP should be expressed in parallel with the production of phage proteins before host cell lysis.

We tried to assess the importance of LevP in the phage infection using Bacillus phage SP50 that only has the pgph gene. The Bacillus subtilis (natto) strain Takahashi is known as a ɣ-PGA and levan producer in the presence of sucrose in the media. One-step growth analysis was used to evaluate the effect of PghP and LevP on the phage infection. When φNIT1 and SP50 were added to the encapsulated stationary-phase cells of the Takahashi strain, phage progenies were detected at 120 min after phage addition, and treatment of recombinant PghP (rPghP) shortened the eclipse period to 60 min. These results support the previous observation that removal of ɣ-PGA from the surface of the cells by rPghP enabled the phage particles to easily access the host receptors. This apparently long eclipse of phage to the encapsulated cells might be related to the time required to produce the capsule depolymerase by the fraction of cellular population that were infected. Unexpectedly, recombinant LevP (rLevP) treatment of encapsulated host
cells did not show significant shortening of the eclipse time, and treatment of the cells with both rPgL and rLcvL did not show any synergistic or additive effects for the phage growth. The role of the phage-encoded levanase in the phage life cycle remains to be investigated.

Author contributions

J. K. and N. A. designed this study, and T. O. and A. S. performed the experiment. T. O., N. A., K. K., A. S. and J. K. analyzed the data and discussed the results. J. K. wrote the manuscript. All authors have read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Sciences (20580069 to J. K.). J. K. was the recipient of a support from Takano Life Science Research Foundation (took over from Yoshifumi Itoh).

Acknowledgments

We are grateful to K. Itoh and Y. Wakabayashi for technical assistance with electron microscopy.

REFERENCES


