Journal of
Microbial & Biochemical Technology

Gontijo et al., J Microb Biochem Technol 2017, 9:4
DOI: 10.4172/1948-5948.1000362

Short Communication Open Access

Bacteriophage Genome Sequencing: A New Alternative to Understand Biochemical Interactions between Prokaryotic Cells and Phages

Marco Túlio Pardini G[1], Laís Silva B[2], Luiz Augusto Aguiar A[3] and Maryoris Elisa Soto L[4]*

[1]*Departamento de Microbiologia, Universidade Federal de Viçosa, Brazil*
[2]*Departamento de Tecnologia de Alimentos, Universidade Federal de Viçosa, Brazil*
[3]*Departamento de Tecnologia de Alimentos, Universidade Federal Rural do Rio de Janeiro, Brazil*
[4]*Departamento de Ingeniería de Alimentos, Universidad de Córdoba, Colombia*

Abstract

Bacteriophages are viruses or semi-autonomous genetic entities that depend on prokaryotic cell's metabolism to multiply. The use of lytic phages as biocontrol agents in many fields such as food preservation, disease control, agriculture production signifies the need of most appropriate and standard methods to insure application safety. Bacteriophage full genetic material sequencing is a new alternative to better understand phage encoded proteins and biomolecules (especially phage lytic enzymes) involved in the process of bacterial cell lysis and death. Thus, this short review aimed to discuss sequencing importance applied to bacteria-phage interactions in society commercial sectors, focusing in the food industry.

Introduction

Bacteriophages were independently discovered by Frederick Twort in 1915 and Felix d'Herelle in 1917 in cell cultures of *Staphylococcus aureus* and Shigella, respectively [1]. Phages were first used in medical field by d'Herelle in 1919. Despite promising results, antibiotic discovery in 1940 decreased researcher's interest in the use of bacteriophages for human disease treatment due controversial public opinion [2,3]. Nowadays, researches concerning bacteriophage application have risen.

Until last decade, the main form to control bacteria human pathogenesis and prevent bacterial multiplication in many sectors was using chemical agents, such as antibiotics and sanitizers. However, indiscriminate use of these chemical agents resulted in bacterial multiresistance. In this way, bacteriophage technology and phage therapy emerged [4,5]. Phages are able to infect only specific species of prokaryote, or even strains within the same species [5]. Bacteriophages are able to adhere on the microbial cell surface, intermediated by proteins and other accessory structures, and inject their genetic material into the host, and may present different cycles of infection: lytic cycle, lysogenic cycle, pseudo-lysogenic cycle or chronic infection. Regarding phage application, preferentially, phages who present lytic cycle of infection could be used as biocontrol tools. In this cycle, bacteriophages inject their genetic material into the cell to produce new viral particles, causing cellular lysis [5,6].

The first complete genome ever sequenced was the one of the bacteriophage øX174 in 1977. This advance was a breakthrough for science. The information contained in the genetic material, which varies widely among bacteriophages, is able to predict the proteins synthesized by the polynucleotide sequence genes and is used to taxonomically classify viruses in orders, families, subfamilies, genus and species [7,8]. Therefore, this mini review paper aimed to discuss bacteriophage sequencing importance encompassing different fields of application, focusing on phages specific for foodborne pathogens.

Phage genome sequencing

Ten families of bacteriophages have been already reported, according to the International committee on taxonomy of viruses (ICTV), phage classification in families and genera are based on capsid morphology, conserved genomic synteny and homology in amino acid sequences of phage genetic material encoded proteins [9]. There are basically four types of genetic material comprising bacteriophages genome, single-stranded and double-stranded DNA and RNA (ssDNA, dsDNA, ssRNA, dsRNA) [10,11]. Genetic material varies widely among phages: genome length ranges from 3405 bp to 497513 bp, gene density ranges from 0.29 to 1.36 and number of encoded proteins ranges from 1 to 675. Until 2017 first semester, there were 7163 viruses' genetic material sequences available online on NCBI genome bank (Table 1). Bacteriophages genome, which includes viruses' specific for archaea and bacteria domains, represents about 31.6% of the total.

Bacteriophage DNA sequencing still presents difficulties even with new sequencing techniques development, which is the bottleneck for phage functional genomic studies [12,13]. The main obstacles are: i) obtaining pure phage genomic material, ii) PCR amplification, and iii) complex nature of its genetic material due to intrinsic characteristics, such as methylated bases and repetition zones, which are intrinsically difficult to sequence and organize [14].

From a technological point of view, bacteriophage sequencing is still essential for any study of functional genomics, as well as for approval and release of bacteriophages use or derived products by regulatory agencies such as the Food and Drug Administration (FDA). As biocontrol tools in the food industry and in the medical field, genetic studies are necessary, once it is known that some viruses are capable of enhancing bacteria pathogenicity [14,15]. In this way, a few companies already market products based on bacteriophages, including products used to control of foodborne pathogens. These products are considered safe for the consumer and are approved by FDA, such as ListShield™ and SalmFres™, used to control *Listeria monocytogenes* and Salmonella on food processing surfaces, respectively [16].

***Corresponding author:** Maryoris Elisa Soto L, Departamento de Ingeniería de Alimentos, Universidad de Córdoba, Colombia, Tel: 573104690204; E-mail: mesoto@correo.unicordoba.edu.co

Received June 01, 2017; **Accepted** July 13, 2017; **Published** July 20, 2017

Citation: Gontijo MTP, Batalha LS, Lopez MES, Albino LAA (2017) Bacteriophage Genome Sequencing: A New Alternative to Understand Biochemical Interactions between Prokaryotic Cells and Phages. J Microb Biochem Technol 9:169-173. doi: 10.4172/1948-5948.1000362

Copyright: © 2017 Gontijo MTP, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Genome sequencing presents one of the most complete technique to study phage encoded proteins, however, genetic material only shows potential predicted proteins, not showing each one of these proteins are actually expressed in the process of host infection [17-20]. In this way, other omic approaches like transcriptomics, proteomics and metabolomics could be used in combination with phage genome sequencing to completely understand phage-bacteria interactions [20-23].

Phage genome sequencing focusing on economic applications in the food industry

In Table 2 is presented some of the phage genome sequences focusing on phage application in the food industry. As shown, phage genetics vary among phages of different hosts, even in an interlinked supply chain like the industry of food processing [24-29]. The main bottleneck for functional overall genomics, including phage genomics, is the low number of available genomes and described genes (open reading frames or ORFs). Many of phage predicted proteins represent "hypothetical proteins" with homology among phages, but with none described function. This scenario shows how low we know about genome and demonstrates the need of much more studies concerning genome sequencing [30,31].

Host	Number of available genomes *	Oldest deposited Genome	Most recent deposited genome	Minimal length (bp)	Maximal length (bp)	Minimal encoded proteins	Maximal encoded proteins
Algae	52	2001	2016	1901	473558	1	886
Archaea	79	1988	2016	9082	77670	7	281
Bacteria	2183	1982	2017	3405	497513	1	675
Diatom	3	2017	2017	4576	4742	4	5
Environment	171	2007	2016	838	31314	1	47
Fungi	1933	1993	2017	1705	24899	1	12
Humans	441	1982	2017	1682	235646	1	233
Invertebrates	1671	1987	2017	647	567670	1	468
Invertebrates and plants	66	1993	2017	3164	29339	1	14
Invertebrates and vertebrates	11	1987	2013	11088	15867	3	12
Invertebrates, vertebrates and humans	7	1987	2005	11088	11703	3	5
Plants	1492	1982	2017	220	231621	1	113
Protozoa	56	1993	2016	497	2473870	1	2541
Vertebrates	1504	1982	2017	859	359853	1	328
Vertebrates and humans	310	1982	2017	1682	235646	1	233
Vertebrates and invertebrates	157	1993	2017	4401	170101	1	152
Vertebrates, invertebrates and humans	124	1993	2017	6391	29210	1	13

*The total does not equal 7163 because some genomes appear more than once in the table when infecting more than one specific host
Source: Adapted from NCBI (https://www.ncbi.nlm.nih.gov/genome)

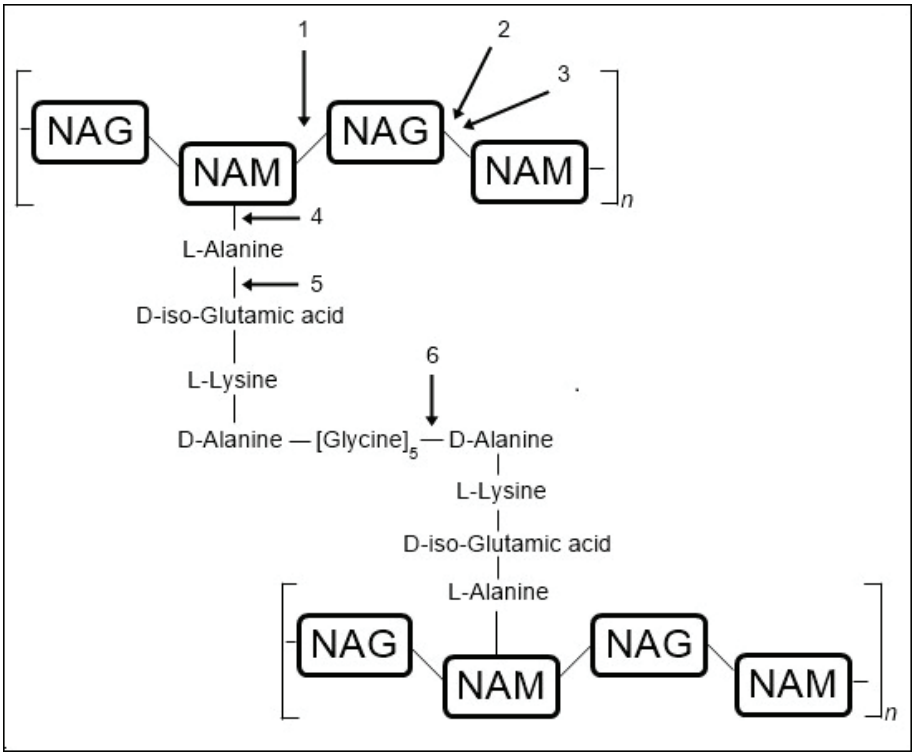
Table 1: Division of host's virus's genomic material available on NCBI genome bank.

Phage identification	Host species	Accession number	Main founds and highlights	References
UFV-AREG1	*Escherichia coli* O157:H7	KX009778	Genome size: 170788 bp ORFs: 274 ORFs encoding hypothetical proteins: 134 Gene density: 1.60 G+C content: 35.3% Phage specific to control *E. coli* causing food poisoning and more severe symptoms. Similar genome organization and encoded proteins among other enteric phages, providing further information especially for *E. coli* specific phages. Presence of both holin and endolysin conserved genes.	[12]
UFV-P2	*Pseudomonas fluorescens*	JX863101	Genome size: 45517 bp ORFs: 75 ORFs encoding hypothetical proteins: 51 Gene density: 1.65 G+C content: 51.5% Pioneer study regarding food safety applications to control spoilage microorganisms in the food industry. Presence of holin independent endolysins as lytic enzymes to rupture host membrane.	[16]
AR9	*Bacillus subtilis*	KU878088	Genome size: 251042 bp ORFs: 292 ORFs encoding hypothetical proteins: 150 Gene density: 1.16 G+C content: 27.8% Presence of an endolysin, with the conserved N-terminal glycoside hydrolase domain and a C-terminal cell wall binding domain. In this study it was proposed some holing candidates possessing transmembrane domains.	[17]

clP1	*Pediococcus damnosus*	JN051154	Genome size: 38013 bp ORFs: 57 ORFs encoding hypothetical proteins: 35 Gene density: 1.50 G+C content: 47.6% Phage specific to control spoilage Pediococcus in beer. This phage possesses putative endolysin and holing, conserved in Lactobacillus phages.	[18]
PLgT-1	*Lactococcus garvieae*	KU892558	Genome size: 40273 bp ORFs: 66 ORFs encoding hypothetical proteins: 42 Gene density: 1.63 G+C content: 35.4% This genome showed high similarity among other Lactococcus genomes. This phage presented two probable holin genes but no endolysin or similar was predicted, which could explain the lysogenic behavior of this phage.	[20]
ΦCD6356	*Clostridium difficile*	GU949551	Genome size: 37664 bp ORFs: 59 ORFs encoding hypothetical proteins: 38 Gene density: 1.57 G+C content: 28.4% This bacteriophage presented a genome more similar to others Siphoviridae phages than among *C. difficile* previously described phages. The endolysin described in this study was similar to the ones of *C. difficile* strains, but no conserved domain of holing was observed.	[21]
LDG	*Leuconostoc* sp.	KX555527	Genome size: 26500 to 28900 bp ORFs: 35 to 50 G+C content: 36 to 39% All four phages presented similar genome size, G+C content and number of ORFs. However, phylogenetic clustering showed more similarities among previously sequenced phages. Endolysins encoded by these presented two different versions of putative domains. Theses bacteriophages presented conserved holin.	[22]
CHA		KX578044		
CHB		KX578043		
Ln-7		KX578042		
vB_YenP_AP5	*Yersinia enterocolitica*	KM253764	Genome size: 38646 bp ORFs: 45 ORFs encoding hypothetical proteins: 16 Gene density: 1.16 G+C content: 50.7% This phage genome was similar to other T7like phages, which includes other bacteriophages specific for enteric pathogens like Salmonella. This phage DNA also encoded both endolysins and holing genes, with putative domains conserved and observed in some Yersinia previously sequenced phages.	[23]
SA11	*Staphylococcus aureus*	JX194239	Genome size: 136326 bp ORFs: 186 ORFs encoding hypothetical proteins: 171 Gene density: 1.36 G+C content: 30.0% Bioinformatics analysis showed that this phage genome showed low similarity to other *S. aureus* phages found in literature, once only 19 of the predicted 186 ORFs were identified. This phage genome also encoded conserved endolysins with putative catalytic and binding domains, but not holing genes were predicted.	[24]
SS3e	Salmonella	AY730274	Genome size: 40793 bp ORFs: 59 ORFs encoding hypothetical proteins: 56 Gene density: 1.45 G+C content: 50.0% SS3 bacteriophage showed high similarity to other Siphoviridae salmonella phages. It was not observed any endolysin or holin conserved genes in this phage genome, which shows the need of further studies to identify proteins responsible for Salmonella cell lysis.	[25]
phiE142	*Escherichia coli* O157:H7 and *Salmonella enterica*	KU255730	Genome size: 121442 bp ORFs: 194 ORFs encoding hypothetical proteins: 115 Gene density: 1.65 G+C content: 37.4% This phage genome showed high similarity to *E. coli*, Enterobacteria and Shigella phages, including endolysin and holin genes.	[26]
vB_LmoS_188	*Listeria monocytogenes*	KP399677	Genome size: 38392 bp and 40759 bp ORFs: 60 and 72 ORFs encoding hypothetical proteins: 32 and 39 Gene density: 1.56 and 1.77 G+C content: 35.9% and 36.9% Genome organization and clustering of both of these phages were similar to other previously described Listeria bacteriophages, but genome identity was slightly low in the group of Listeria phages. Both phages presented conserved domains of both holin and endolysin genes.	[27]
vB_LmoS_293		KP399678		

pSf-1	Shigella	KC710998	Genome size: 51821 bp ORFs: 94 ORFs encoding hypothetical proteins: 94 Gene density: 1.81 G+C content: 44.0% pSf-1 genome sequence showed high similarity with other Shigella phages, but showing low gene homology, especially regarding some genome translocation a no identification of lytic enzyme genes.	[28]
CP21	*Campylobacter jejuni* and *Campylobacter coli*	HE815464	Genome size: 182833 bp ORFs: 259 ORFs encoding hypothetical proteins: 124 Gene density: 1.42 CP21 genome showed high homology with other three Campylobacter phage genomes, in the genus T4-like of bacteriophages. No obvious lytic enzymes were observed in theses phage genome.	[29]

Table 2: Comparative analysis of some sequenced bacteriophages specific for foodborne bacteria and main results found in which research.



NAG: N-Acetylglucosamine; NAM: N-Acetylmuramic Acid
1) N-acetyl-β-D-glucosaminidase; 2) N-acetyl-β-D-muramidase; 3) lytic transglycosylase; 4) N-acetylmuramoyl-L-alanine amidase; 5) L-alanoyl-D-glutamate endopeptidase; and 6) D-alanyl-glycyl endopeptidase

Figure 1: Schematic representation of the peptidoglycan layer on the cell wall of bacteria and target chemical bonds by phage lytic enzymes.

However, the main focus of the phage genomic studies target the identification of lytic enzymes, and these enzymes show some homology and putative domains among species of bacteriophages [32]. Studies regarding these enzymes description have grown exponentially over the years compared to other genetic fields.

Depending on specificity of the catalytic domain of the lytic enzyme, phage endolysins can be classified in putative groups, showing similar mode of action. Until further discoveries, there are six classes of lytic endolysins described in literature: i) N-acetyl-β-D-glucosaminidase; ii) N-acetyl-β-D-muramidase and iii) lytic transglycosylase; all responsible for breaking the sugar bonds of β-(1,4) N-acetylglucosamine and N-acetylmuramic acid; iv) N-acetylmuramoyl-L-alanine amidase, responsible for cleave the bond between L-alanine and N-acetylmuramic acid; v) L-alanoyl-D-glutamate endopeptidase, which cleaves the chemical bond between L-alanine and D-glutamic acid; and vi) D-alanyl-glycyl endopeptidase, which breaks the peptide bond between the 5-glycine inter-bridge between glycine and L-alanine in Gram-positive bacteria (Figure 1) [32,33].

In this way, phages and phage enzymes have advantages of usage related to bacterial resistance. Once endolysins present some specificity among different genera or species, the probability of the bacteria acquiring mechanisms of resistance is low. This is explained by the theory of phage-bacteria co-evolution, which postulates that to ensure phage multiplication and survival in the environment, phage and its endolysins were naturally selected, which difficult bacterial resistance.

Final Considerations

Genome sequencing is one of the most effective technologies to

understand phage-bacteria interactions during phage infection cycle, however, compared to the number of isolated phages, phage genomes available online are limited. In this way, it is widely necessary more researches concerning genomics and phage sequencing, including other omic approaches, which in combination can provide additional information to complement bacteriophage studies in the biological fields and insure phage application and diffusion in many society sectors.

References

1. Negash A, Ejo M (2016) Review on bacteriophages and its antimicrobial uses. Amer-Eurournal Sci Res 11: 199-208.
2. Taylor MW (2014) Viruses and man: A history of interactions. (1stedn), Springer International Publishing Switzerland.
3. Wittebole X, De Roock S, Opal SM (2014) A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. Virulence 5: 226-235.
4. Golkar Z, Bagasra O, Pace DG (2014) Bacteriophage therapy: a potential solution for the antibiotic resistance crisis. J Infect Dev Ctries 8: 129-136.
5. Muñoz SLD, Koskella B (2014) Bacteria-phage interactions in natural environments. Adv Appl Microbiol 89: 135-183.
6. Jassim SAA, Limoges RG (2014) Natural solution to antibiotic resistance: bacteriophages: The living drugs. World J Microbiol Biotechnol 30: 2153-2170.
7. Klumpp J, Fouts DE, Sozhamannan S (2012) Next generation sequencing technologies and the changing landscape of phage genomics. Bacteriophage 2: 190-199.
8. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR et al. (1977) Nucleotide sequence of bacteriophage fX174 DNA. Nature 265: 687-695.
9. Kulikov EE, Golomidova AK, Letarova MA, Kostryukova ES, Zelenin AS et al. (2014) Genomic sequencing and biological characteristics of a novel Escherichia coli bacteriophage 9g, a putative representative of a new Siphoviridae genus. Viruses 6: 5077-5092.
10. Hungaro HM, Lopez MES, Albino LAA, Mendonça RCS (2014) Bacteriophage: The viruses infecting bacteria and their multiple applications. In: Jull A T. Reference Module in Earth Systems and Environmental Sciences. (1stedn), Elsevier Inc.
11. Ackermann H W (1998) Tailed bacteriophages: The order Caudovirales. Adv Virus Res 51: 135-201.
12. Lopez MES, Batalha LS, Vidigal PMP, Albino LAA, Boggione DMG et al. (2016) Genome sequence of the enterohemorrhagic *Escherichia coli* bacteriophage UFV-AREG1. Genome Announc 4: e00412-16.
13. Petrov VM, Ratnayaka S, Nolan JM, Miller ES, Karam JD (2010) Genomes of the T4-related bacteriophages as windows on microbial genome evolution. Virol J 7: 292.
14. Klumpp J, Fouts DE, Sozhamannan S (2013) Bacteriophage functional genomics and its role in bacterial pathogen detection. Brief Funct Genomics elt009.
15. Endersen L, O'Mahony J, Hill C, Ross RP, McAuliffe O, et al. (2014) Phage therapy in the food industry. Annu Rev Food Sci Technol 5: 327-349.
16. Eller MR, Vidigal PM, Salgado RL, Alves MP, Dias RS, et al. (2014) UFV-P2 as a member of the Luz24likevirus genus: A new overview on comparative functional genome analyses of the LUZ24-like phages. BMC Genomics 15: 7.
17. Lavysh D, Sokolova M, Minakhin L, Yakunina M, Artamonova T, et al. (2016) The genome of AR9, a giant transducing Bacillus phage encoding two multi-subunit RNA polymerases. Virology 495: 185-196.
18. Kelly D, O'Sullivan O, Mills S, McAuliffe O, Ross RP, et al. (2012) Genome sequence of the phage clP1, which infects the beer spoilage bacterium *Pediococcus damnosus*. Gene 504: 53-63.
19. Lal T M, Sano M, Hatai K, Ransangan J (2016) Complete genome sequence of a giant Vibrio phage ValKK3 infecting *Vibrio alginolyticus*. Genom Data 8: 37-38.
20. Kelly D, O'Sullivan O, Mills S, McAuliffe O, Ross R P, et al. (2012) Genome sequence of the phage clP1, which infects the beer spoilage bacterium Pediococcus damnosus. Gene 504: 53-63.
21. Horgan M, O'Sullivan O, Coffey A, Fitzgerald GF, Van Sinderen D, et al. (2010) Genome analysis of the *Clostridium difficile* phage FCD6356, a temperate phage of the Siphoviridae family. Gene 462: 34-43.
22. Pujato SA, Guglielmotti DM, Martínez-García M, Quiberoni A, Mojica FJ (2017) *Leuconostoc mesenteroides* and *Leuconostoc pseudomesenteroides* bacteriophages: Genomics and cross-species host ranges. Int J Food Microbiol 257: 128-137.
23. Leon-Velarde CG, Kropinski AM, Chen S, Abbasifar A, Griffiths MW, et al. (2014) Complete genome sequence of bacteriophage vB_YenP_AP5 which infects *Yersinia enterocolitica* of serotype O: 3. Virology J 11: 188.
24. Kim MS, Myung H (2012) Complete genome of *Staphylococcus aureus* phage SA11. J Virol 86: 10232-10232.
25. Kim SH, Park JH, Lee BK, Kwon HJ, Shin JH, et al. (2012) Complete genome sequence of Salmonella bacteriophage SS3e. J Virol 86: 10253-10254.
26. Amarillas L, Chaidez C, González-Robles A, León-Félix J (2016) Complete genome sequence of new bacteriophage phiE142, which causes simultaneously lysis of multidrug-resistant *Escherichia coli* O157: H7 and *Salmonella enterica*. Stand Genomic Sci 11: 89.
27. Casey A, Jordan K, Coffey A, McAuliffe O (2015) Complete genome sequences of vB_LmoS_188 and vB_LmoS_293, two bacteriophages with specificity for *Listeria monocytogenes* strains of serotypes 4b and 4e. Genom Announc 3: e00040-15.
28. Jun JW, Kim JH, Shin SP, Han JE, Chai JY, et al. (2013) Characterization and complete genome sequence of the Shigella bacteriophage pSf-1. Res Microbiol 164: 979-986.
29. Hammerl JA, Jäckel C, Reetz J, Hertwig S (2012) The complete genome sequence of bacteriophage CP21 reveals modular shuffling in Campylobacter group II phages. J Virol 86: 8896-8896.
30. Oliveira H, Azeredo J, Lavigne R, Kluskens LD (2012) Bacteriophage endolysins as a response to emerging foodborne pathogens. Trends Food Sci Technol 28: 103-115.
31. Garcia P, Rodriguez L, Rodriguez A, Martinez B (2010) Food biopreservation: Promising strategies using bacteriocins, bacteriophages and endolysins. Trends Food Sci Technol 21: 373-382.
32. Schmelcher M, Donovan DM, Loessner MJ (2012) Bacteriophage endolysins as novel antimicrobials. Future Microbiol 7: 1147-1171.
33. Schmelcher M, Shen Y, Nelson DC, Eugster MR, Eichenseher F, et al. (2015) Evolutionarily distinct bacteriophage endolysins featuring conserved peptidoglycan cleavage sites protect mice from MRSA infection. J Antimicrob Chemother 70: 1453-1465.