Virus Research 259 (2019) 77–84

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Comparative analysis and characterization of Enterobacteria phage SSL-2009a and '*HK578likevirus*' bacteriophages

Shu Li[a], Shuguang Lu[a], Hui Huang[b], Li Tan[a], Qingshan Ni[a], Weilong Shang[a], Yi Yang[a], Zhen Hu[a], Junmin Zhu[a], Ming Li[a], Xiancai Rao[a], Fuquan Hu[a], Xiaomei Hu[a,*]

[a] Department of Microbiology, Army Medical University, Chongqing, China
[b] Department of Clinical Laboratory Science, Army Medical University, Chongqing, China

ARTICLE INFO

Keywords:
Bacteriophage
HK578likevirus
Enterobacteria phage SSL-2009a
Genome annotation
Comparative analysis

ABSTRACT

Enterobacteria phage SSL-2009a is a virulent bacteriophage with strong and abroad lytic ability against lots of engineering *E. coli* strains. In this study, we re-sequenced its whole genome and made a detail analysis on its genomic and proteomic characteristics according to the updated genomic sequence. The genome of SSL-2009a is a circular double-stranded DNA of 44,899 base pairs in length, with a 54.67% G + C content. A total of 67 open reading frames were predicted as protein coding sequences, 24 of which encode products highly homologous to known phage proteins. There are 10 promoters and 22 terminators identified in the genome of SSL-2009a, but no tRNA is found. SSL-2009a belongs to the '*HK578likevirus*' genus of *Siphoviridae*. Comparative analyses indicated that other twelve phages share high homology with SSL-2009a at nucleotide and amino acid levels and also should be clustered into the same genus. In-depth analysis was performed to reveal the genomic, proteomic, and morphological features of these '*HK578likevirus*' phages, which may promote our understanding of Enterobacteria phage SSL-2009a and the '*HK578likevirus*' genus, even the biodiversity and evolution of bacteriophages.

1. Introduction

Bacteriophages (phages) are the natural infectious agents of bacteria. They can infect and kill bacteria with high specificity and effectiveness, and are considered an important biological tool for the typing, treatment, and decontamination of bacteria. Phages are also associated with the genetic variation of bacteria through the horizontal transfer of genetic material as specific virulence or drug-resistance genes (Touchon et al., 2017; Brown-Jaque et al., 2015; Penadés et al., 2015), and contribute to the evolution of bacteria. Recently, increasing attention has been focused on phages because of their potential contributions to control infectious diseases caused by refractory bacterial pathogens, such as multidrug-resistant bacteria (Cisek et al., 2017; Malik et al., 2017; Lin et al., 2017).

Bacteriophages are distributed ubiquitously in nature, and are presumed to be the most abundant biological entity on earth. It is estimated that there are about 10^{30} to 10^{32} phages in the biosphere, approximately ten-fold higher than the number of bacteria (Hendrix, 2003). The giant bacteriophage community has abundant biological diversity, but our understanding of them is extremely deficient. Since the discoveries of phages were reported by Frederick William Twort in 1915 and Felix d` Herelle in 1917 (Salmond and Fineran, 2015), about 6300 different phages have been examined using electron microscopy (Ackermann and Prangishvili, 2012). Based on data from the National Center for Biotechnology Information (Bethesda, MA, USA), only about 2000 phage genomes have been completely sequenced up to February 2016, even fewer than those of their host (more than 5000 genomes of bacteria have been examined). Therefore, it is necessary to isolate and identify novel bacteriophages, to sequence and perform in-deep analyses of their genomes, which will improve our understanding of the biodiversity and evolution of bacteriophages and bacteria.

Enterobacteria phage SSL-2009a is a virulent bacteriophage with strong and abroad lytic ability against many engineering *E. coli* strains. It is able to infect and lyse 10 different engineered *E. coli* strains, including BL21 (DE3), DH10B, DH5α, JM109, M15, Rosetta (DE3), Rosetta-gami (DE3), S17-1, S17-1 λ pir+, and TOP10. SSL-2009a also displays powerful tolerance to terrible environments, and can survive in 75% or 100% ethanol, high temperatures as 90 °C and even in strong

* Corresponding author at: Department of Microbiology, Army Medical University, Chongqing, 400038, China.
E-mail addresses: 1416234121@qq.com (S. Li), shulang88@126.com (S. Lu), huanghui@tmmu.edu.cn (H. Huang), 540899224@qq.com (L. Tan), qsni@tmmu.edu.cn (Q. Ni), shangwl0414@163.com (W. Shang), yangyi19871230@126.com (Y. Yang), 35090772@qq.com (Z. Hu), zhujunmincn@126.com (J. Zhu), sotx7080@163.com (M. Li), raoxiancai@126.com (X. Rao), hufuquan2012@aliyun.cn (F. Hu), huxiaomei@tmmu.edu.cn (X. Hu).

https://doi.org/10.1016/j.virusres.2018.10.019
Received 4 June 2018; Received in revised form 3 September 2018; Accepted 23 October 2018
Available online 25 October 2018

acid at pH 4 (Li et al., 2010). Therefore, SSL-2009a shows powerful potential threat to scientific research and industrialized fermentation processes. The genome of SSL-2009a had been sequenced in 2009. However, some errors were found in this previous sequence during the subsequent researches, confusing our understanding of the phage. So, the genome of SSL-2009a was re-sequenced. In the new sequence, 5107 missing base pairs are added, the inversion and staggered arrangement of genomic fragments are also corrected (Fig. S1). In the present study, the comparative analysis on the updated sequence was performed, which indicated that SSL-2009a should be clustered into the '*HK578likevirus*' of *Siphoviridae*. Other 12 bacteriophages that share high similarity with SSL-2009a at nucleotide and amino acid levels also should be classed into this genus. Based on the in-depth analysis of these 13 phages, some genomic, proteomic, and biological characteristics of '*HK578likevirus*' phages were also revealed in this paper.

2. Materials and methods

2.1. Bacteriophage propagation and purification

Enterobacteria phage SSL-2009a was primarily isolated from a culture of the *E. coli* BL21 (DE3) strain, and BL21 (DE3) was used as the host cell for the subsequent propagation of SSL-2009a (Li et al., 2010). In detail, the log phase liquid cultures of the BL21 (DE3) strain were infected with the phage at a MOI of 0.01 and incubated for 12–16 h at 37°C with shaking. The overnight culture was centrifuged (10,000 *g* for 15 min at 4°C), and then the supernatant was collected and filtered with a 0.22 μm filter membrane. The crude phage suspension was concentrated with 10% polyethylene glycol 8000 and 1 M NaCl according to the method described by Govind et al. (Govind et al., 2006). Further purification of SSL-2009a particles was performed using CsCl gradient ultracentrifugation (Casas and Rohwer, 2007).

2.2. Genomic DNA extraction and sequencing

The genomic DNA of SSL-2009a was extracted from the purified phage particles using the SDS-proteinase K protocol according to the methods described previously (Lu et al., 2013), and was further purified by phenol extraction and ethanol precipitation. The genomic DNA of SSL-2009a was subjected to high-throughput sequencing at the State Key Laboratory of Pathogens and Biosecurity at the Beijing Institute of Microbiology and Epidemiology using a Life Technologies Ion Personal Genome Machine Ion Torrent sequencer (San Francisco, CA). The complete genome sequence was assembled using the Velvet (Zerbino and Birney, 2008) and CLC Bio (Aarhus, Denmark) software programs.

2.3. Sequence analysis and annotation

The basic features of SSL-2009a genome sequence were analyzed using the software packages DNAStar and DNAMAN (http://www.lynnon.com/). The GC skew and GC content analysis of the SSL-2009a genome was identified with CGView program (http://stothard.afns.ualberta.ca/cgview_server) (Stothard and Wishart, 2005). Open reading frames (ORFs) were scanned using ORF Finder at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) (Wheeler et al., 2003). Putative genes were predicted using the GeneMark.HMM software (http://opal.biology.gatech.edu/GeneMark/genemark24.cgi) with a length threshold of 100 bp (Besemer and Borodovsky, 1999). The homology searches of the nucleotide sequences and protein sequences were conducted using the BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1997). Both the whole genomic sequence of SSL-2009a and its proteomic annotation have been deposited in GenBank with the accession number of FJ750948. The transfer RNA (tRNA) genes in genomic DNA were screened using the online tool, tRNAscan-SE 1.21 (http://lowelab.ucsc.edu/tRNAscan-SE), with a cover score cutoff of 20 (Schattner et al., 2005). Bacteriophage-specific promoters were identified by neural network promoter prediction with the minimum score set at 0.99 (http://www.fruitfly.org/seq_tools/promoter.html) (Reese, 2001), and with PHIRE (Lavigne et al., 2004). The Rho-independent Transcriptional terminators were determined using the web tool, FindTerm (http://linux1.softberry.com/berry.phtml?topic=findterm&group=programs&subgroup=gfindb).

2.4. Comparative genome analysis

Using the BLASTn online tool on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi), thirteen bacteriophages were found with high nucleotide homology with SSL-2009a: Enterobacteria phage JL1 (NC_019419), EK99P-1 (NC_024783), HK578 (NC_019724), slur05 (LN881730), slur06 (LN881731), Envy (KX534335), Gluttony (KX534336), Lust (KX534338), Sloth (KX534339, partial genome), Pride (KX534341), and Shigella phage EP23 (NC_016566), Sodalis phage SO-1 (NC_013600), Siphoviridae phage YD-2008.s (KM896878). Except Enterobacteria phage Sloth whose genome sequence is incomplete, the genomes of other 12 bacteriophages were compared with that of SSL-2009a using BLAST Ring Image Generator (BRIG) with the percentage identity cut off value of 80% (http://brig.sourceforge.net/) (Alikhan et al., 2011). Paired comparisons between each two of these phages were conducted using BLASTn. A dot plot analysis was also performed for the whole genome comparison of the thirteen bacteriophages and the results were displayed using the Gepard program (Krumsiek et al., 2007).

The homologies among the 13 bacteriophages at amino acid level were analyzed and graphed with Easyfig (http://easyfig.sourceforge.net/) according to tBLASTx (Sullivan et al., 2011). CoreGenes3.5 was also used to evaluate the protein homologies of these phages with a BLASTp threshold score of 75 (http://binf.gmu.edu:8080/CoreGenes3.5) (Turner et al., 2013). Multiple sequence alignment of the terminase large subunits was conducted by a ClustalW algorithm using default parameters (Chenna et al., 2003; Li, 2003), and related phylogenetic tree was also constructed and displayed using the MEGA6 program (Tamura et al., 2013) via the neighbor-joining method (Som and Fuellen, 2009).

3. Results and discussion

3.1. General characteristics of the re-sequenced SSL-2009a genome

SSL-2009a possesses a circular dsDNA genome (Fig. S2) of 44,899 bp with an average G + C content of 54.67% which is slightly higher than that of its host *E. coli* (50.6%) (Blattner *et al.*, 1997). A total of 257 ORFs (> 100 bp) were predicted using the ORF Finder software program (Wheeler et al., 2003) with ATG, GTG, or TTG as start codons. Among these, 67 ORFs were identified as protein coding sequences (CDs) by the GeneMark.HMM tool. Sixty-three CDs are initiated with the start codon ATG, one is start with TTG, and the other three with GTG. The average length of a gene is about 644 bp. Approximately 96.08% of the SSL-2009a genome (about 43,140 bp) is predicted to be involved in proteins encoding, which is in concordance with the highly efficient and economical utilization of genomic sequences by bacteriophages. Among the 67 putative genes, 42 were in the forward strand and 25 were in complementary strand (Fig. 1). Ten putative promoters (Table S1) and 22 rho-dependent terminators (Table S2) were identified from the SSL-2009a genome, while no tRNA gene was found with tRNAScan software program. The general genome features of Enterobacteria phage SSL-2009a were collected in Table S3.

The presumptive functions of the 67 putative genes were predicted by comparing with the non-redundant protein sequences in NCBI database. Twenty-four of the 67 predicted genes (about 35.82%) show high sequence homology to known phage-related genes with identified functions, 40 genes (about 56.9%) code for conserved hypothetical

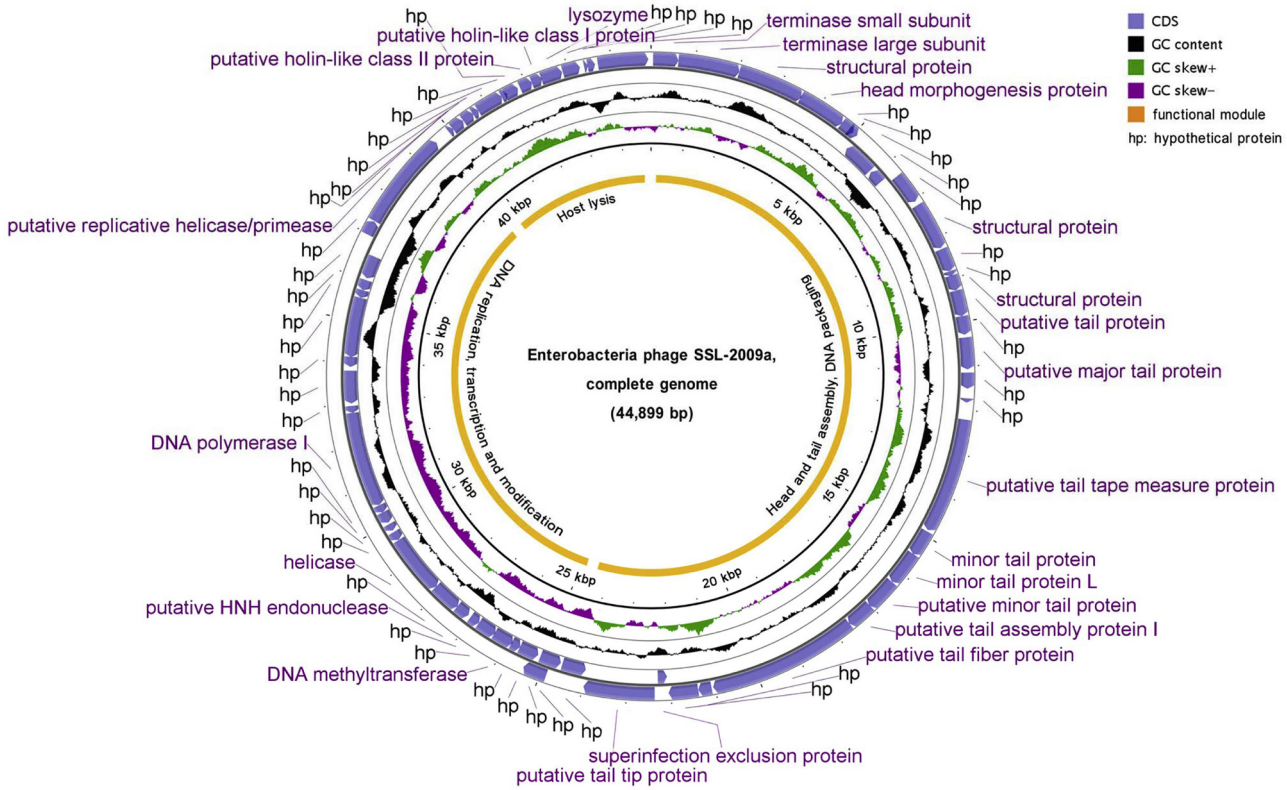

Fig. 1. BRIG and CGView image combined to represent the updated SSL-2009a genome. Genome is represented by a circular visualization and coding regions are represented by arrows. The putative function of each CDS is labeled at the outermost ring and the organization of the function modules of 67 CDS are represented at the innermost rings. The third ring shows GC skew with pink for GC skew- and green for GC skew+. The GC content appears in the black ring.

phage proteins with unknown functions, while no significant homologues of gp31, gp65 and gp66 were found in the databases. The basic characteristics of the 67 predicted genes are listed in Table S4 and the physical map of these genes is illustrated in Fig. 1.

3.2. The evolutionary relationship of phage SSL-2009a

3.2.1. Comparative analysis at nucleotide level

When compared its whole genome sequence against the nucleotide sequence database in NCBI using BLASTn, SSL-2009a showed high homology with thirteen bacteriophages: Enterobacteria phage JL1 (NC_019419, 43,757 bp), EK99P-1 (NC_024783, 44,332 bp), HK578 (NC_019724, 43,741 bp), slur05 (LN881730, 43,900 bp), slur06 (LN881731, 44,515 bp), Envy (KX534335, 45,206 bp), Gluttony (KX534336, 44,513 bp), Lust (KX534338, 41,942 bp), Sloth (KX534339, 45,206 bp, partial genome), Pride (KX534341, 44,899 bp), and Shigella phage EP23 (NC_016566, 44,077 bp), Sodalis phage SO-1 (NC_013600, 45,169 bp), Siphoviridae phage YD-2008.s (KM896878, 44,613 bp). Except the Enterobacteria phage Sloth with imcomplete genome, other 13 phages were included in comparative analysis. The detailed genomic features of the 13 phages were listed in Table S3. As showed in Table S3, the genomic sequence of SSL-2009a is 96% identical to that of JL1, 94% to YD-2008.s, 93% to EK99P-1, 92% to both HK578 and EP23, 91% to slur05, slur06, Envy, Gluttony, Lust and Pride, 90% to SO-1. Meanwhile, high query coverages at 85%–95% and E values of 0.0 were also found in this comparison (Table S3). Moreover, the nucleotide identities between each two phage of the thirteen phages were as high as 90%–100% (Table S5). It is notable that almost the whole genome sequence of phage Lust can be found in Pride genome and covered about 94% of Pride genomic sequence (marked by bold in Table S5). The high identities among the 13 genomes were also revealed by the genomic comparison drawn with BLAST Ring image Generator (BRIG) (Fig. 2). Gepard dotplots analysis was employed to compare the full genome sequences of the 13 bacteriophages against each other. Diagonal lines indicated that the nucleotide sequences of these phages are organized similarly and matched well with each other (Fig. S3). Among these phages, phage JL1, YD-2008.s, slur05, Envy and Lust showed reverse matching with the other phages, revealing that the genomes of these 5 phages were organized in the opposite direction against others.

3.2.2. Comparative analysis at amino acid level

Among the 67 predicted genes of SSL-2009a, the amino acid sequences of 64 genes (about 95.5%) share great similarity with those of the other 12 phages, with high identities ranging from 75% to 100% (Table S4). Of note, more than 95% sequence identities in the amino acid composition were reported between 50 putative SSL-2009a proteins and those of JL1, EK99P-1, HK578, EP23, SO-1, and YD-2008.s. For example, the amino acid sequences of gp17, gp40, gp61 and gp62 of SSL-2009a were 100% homologous with those of EP23, HK578 and JL1. The major head protein of thirteen phages shared 96%–99% sequence identity at amino acid level, while the large subunit of terminase has 97%–99% protein sequence similarity among the phages. As the conserved proteins of bacteriophage, the high identity of major head protein and terminase large subunit at the amino acid level reveal that these phages may be descended from a common ancestor and should be grouped into the same genus.

CoreGenes 3.5 was used to further analyze the relationship of the 13 phages at protein level with the default BLASTp threshold of 75. It was found that 40 proteins were encoded in common among these bacteriophages, which covered most predicted proteins of these phages, and the coverage is about 59.70% of the putative proteins of SSL-2009a, 66.66% of JL1, 64.52% of EK99P-1, 66.66% of HK578, 70.18% of EP23, 67.80% of SO-1, 64.52% of YD-2008.s, 68.97% of slur05, 68.97% of slur06, 64.52% of Envy, 65.57% of Gluttony, 70.18% of Lust, and 63.49% of Pride respectively. The core proteins comparisons between

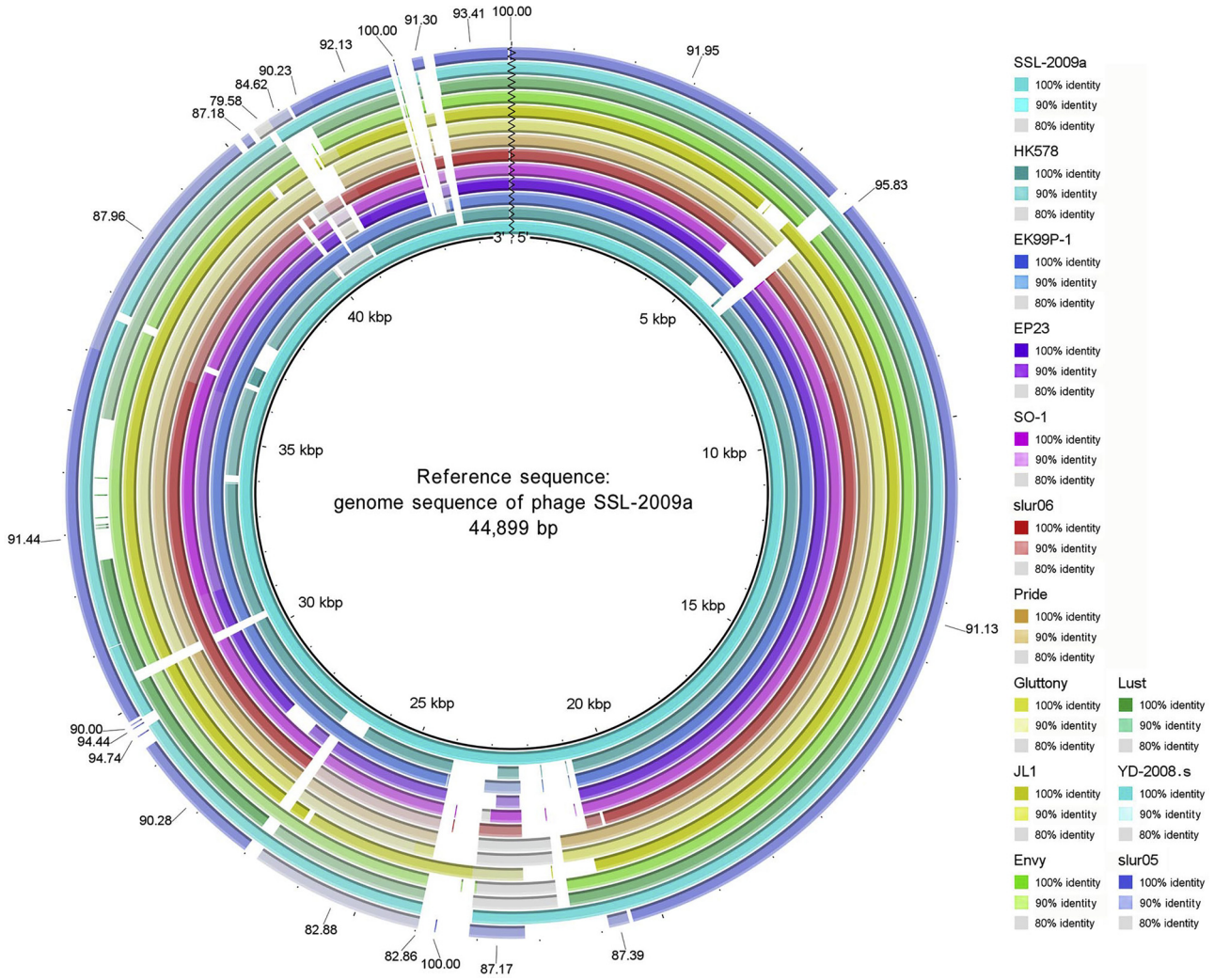

Fig. 2. BRIG output image of the genomic similarities of the 13 '*HK578likevirus*' phages. Genomes of the 13 '*HK578likevirus*' phages are shown as concentric rings, and each ring represents the similarity with a single phage. Central ring is multifasta of all 13 genomes. The colored regions indicate high pairwise genomic sequence similarity according to BLASTn, and the missing color means absence of relative genomic region. Legend indicates which colors correspond with which phages and the shade of that color indicates the level of similarity is observed.

each two genomes of the 13 phages were showed in Table S6, which indicted that these 13 bacteriophages share high homology with each other at protein level. Even the lowest percent of core proteins presented between YD-2008.s and HK578 is more than 80% (80.65%). Moreover, all the predicted proteins of phage Lust can be found from those of phages Envy and Gluttony, and the proteins of Gluttony are also 100% identity to those of phage Envy. The high identity at proteomic level can also be visualized from the Easyfig CDs comparison of the 13 phages (Fig. 3). It had been previously demonstrated that phages of the Podoviridae and Myoviridae families can be grouped together when they share 40% of core proteins with each other (Lavigne et al., 2008; 2009). According to this cutoff value, these 13 bacteriophages may be grouped in the same genus. From Fig. 3, we also found that the genome organizations and transcription directions of YD-2008.s, slur05, JL1, Lust and Envy are inverted against the other 9 bacteriophages, which is consistent with the discovery in dot plot analysis (Fig. S3).

Based on the comparative analyses at nucleotide and protein levels, we presumed that the thirteen bacteriophages are closely related phylogenetically with each other and may be descended from a single common ancestor.

It was previously reported that most phages (about 96%) investigated so far are tailed, double-stranded DNA viruses that belong to the order *Caudovirales* (Ackermann, 1998, 2009). Due to their tail morphology and nucleic acid properties, the tailed bacteriophages are divided into three families: *Myoviridae*, *Siphoviridae* and *Podoviridae*. The phages in the *Myoviridae* family always have a long contractile tail, while *Siphoviridae* have a long non-contractile tail and *Podoviridae* have a short non-contractile tail. Among the three families, the *Siphoviridae* is the most abundant, and about 61% of observed phages and most of the sequenced phages belong to this family (Ackermann, 2009). According to the report of ICTV on virus taxonomy (http://www.ictvonline.org), the *Siphoviridae* family is classified into 11 subfamilies, 100 genera. Because of their typical morphology (Li et al., 2010; Pan et al., 2013; Chang and Kim, 2011) and high homology to HK578 both at the nucleotide and amino acid level, the other twelve phages should be classified into the same genus, *HK578likevirus* genus of *Siphoviridae*. Besides the seven bacteriophages (JL1, SSL-2009a, HK578, EP23, EK99P-1, YD-2008.s and SO-1) which had been clustered into this genus, the additional joining of slur05, slur06, Envy, Gluttory, Lust and Pride extend this genus and will promote our understanding of '*HK578likevirus*' bacteriophages.

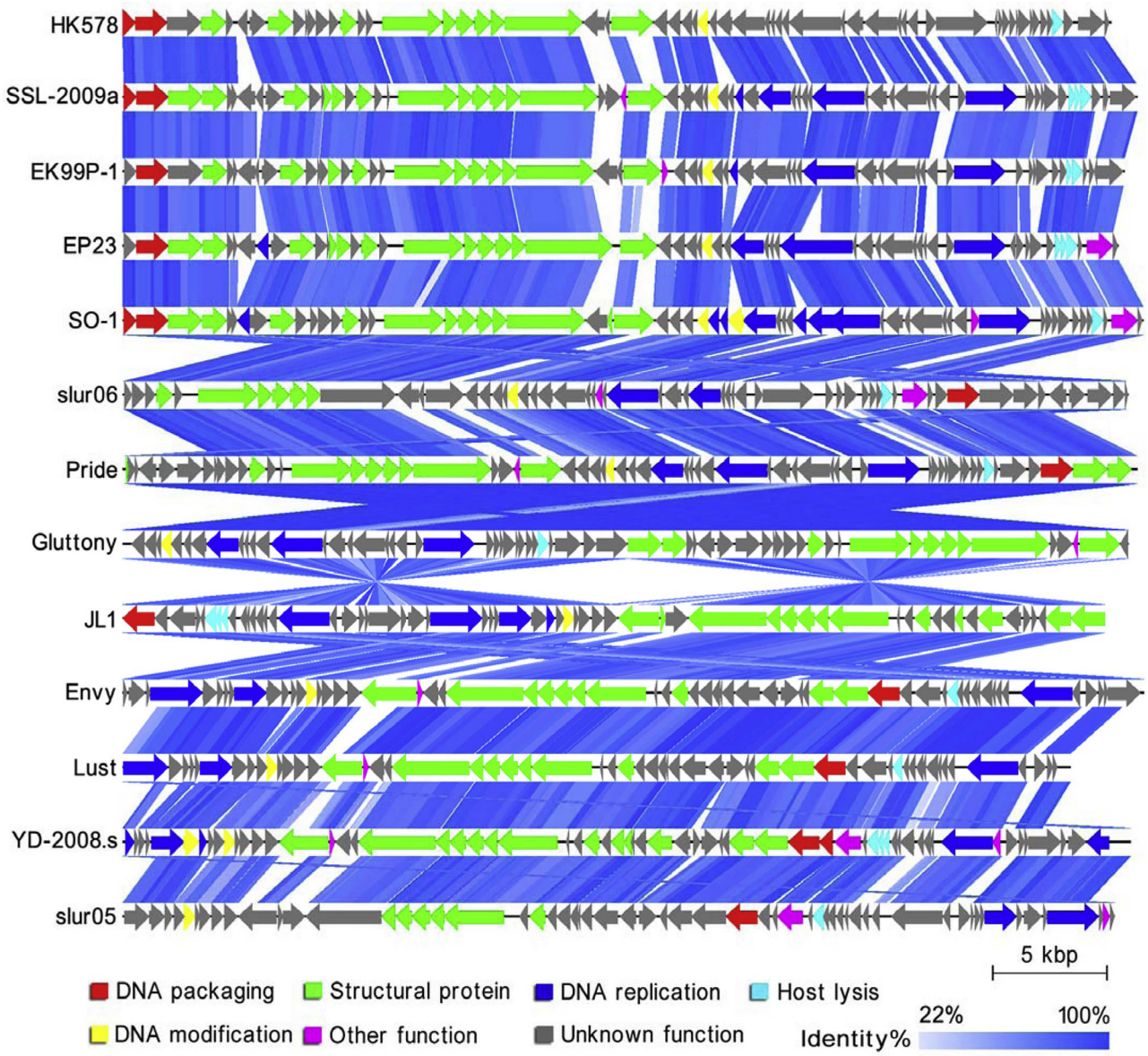

Fig. 3. Easyfig image of the genomic comparison among the 13 '*HK578likevirus*' phages. Genomes of each phage is represented by linear visualization and coding regions represented by arrows. The direction of arrow indicated the transcription direction of each CDS. CDS are colored according to their functions. The blue vertical blocks between sequences indicate regions of shared similarity shaded according to BLASTn, and the degree of sequence similarity is indicated by the intensity of blue.

3.3. The characterization of 'HK578likevirus' genus

3.3.1. The general genomic features of 'HK578likevirus' phages

According to the analysis to the genomes of these 13 phages, it were found that the genome sizes for '*HK578likevirus*' phages range from 41.9 kb to 45.2 kb (with an average GC content of 54.42% to 54.77%). Fifty-seven to 67 predicted genes, 9 to 16 promoters as well as 17 to 24 terminators were identified in the genomes of '*HK578likevirus*' phages, but no tRNA was found in all the 13 genomes (Table S3).

3.3.2. The features of 'HK578likevirus' functional modulars

In general, the genome of bacteriophage is organized in a mosaic modular structure where each module contains gene clusters with specific function (Brussow and Hendrix, 2002), which suggest that the evolution of phage genomes is always pushed by combining modules from different species (Hendrix et al., 1999). For the 13 '*HK578likevirus*' phages, the putative genes with predicted function were also organized in a modular format, and can be classified into six functional modules based on their functions: DNA packaging, head assembly, tail assembly, DNA modification, DNA replication and host lysis (Table S7).

3.3.3. The morphological characterization of 'HK578likevirus' phages

As most of bacteriophages, morphogenesis module, consisted of head assembly module and tail assembly module, is the largest one for all the 13 phages. More than 10 genes with predicted function are included in this module (Table S7). The high amino acid homologies among these structural associated genes indicated the 13 phages are nearly morphologically identical to each other (Table S7). According to the revealed morphological features of phages SSL-2009a (Li et al., 2010) and JL1 (Pan et al., 2013), we presumed that the '*HK578likevirus*' phages have an icosahedric head and a long flexible noncontractile tail.

3.3.4. The packaging strategy of 'HK578likevirus' phages

It was well known that the tailed dsDNA bacteriophages always has similar packaging machine. After the replication of genomic DNA, the replicating concatemeric or circular chromosome DNA molecules will

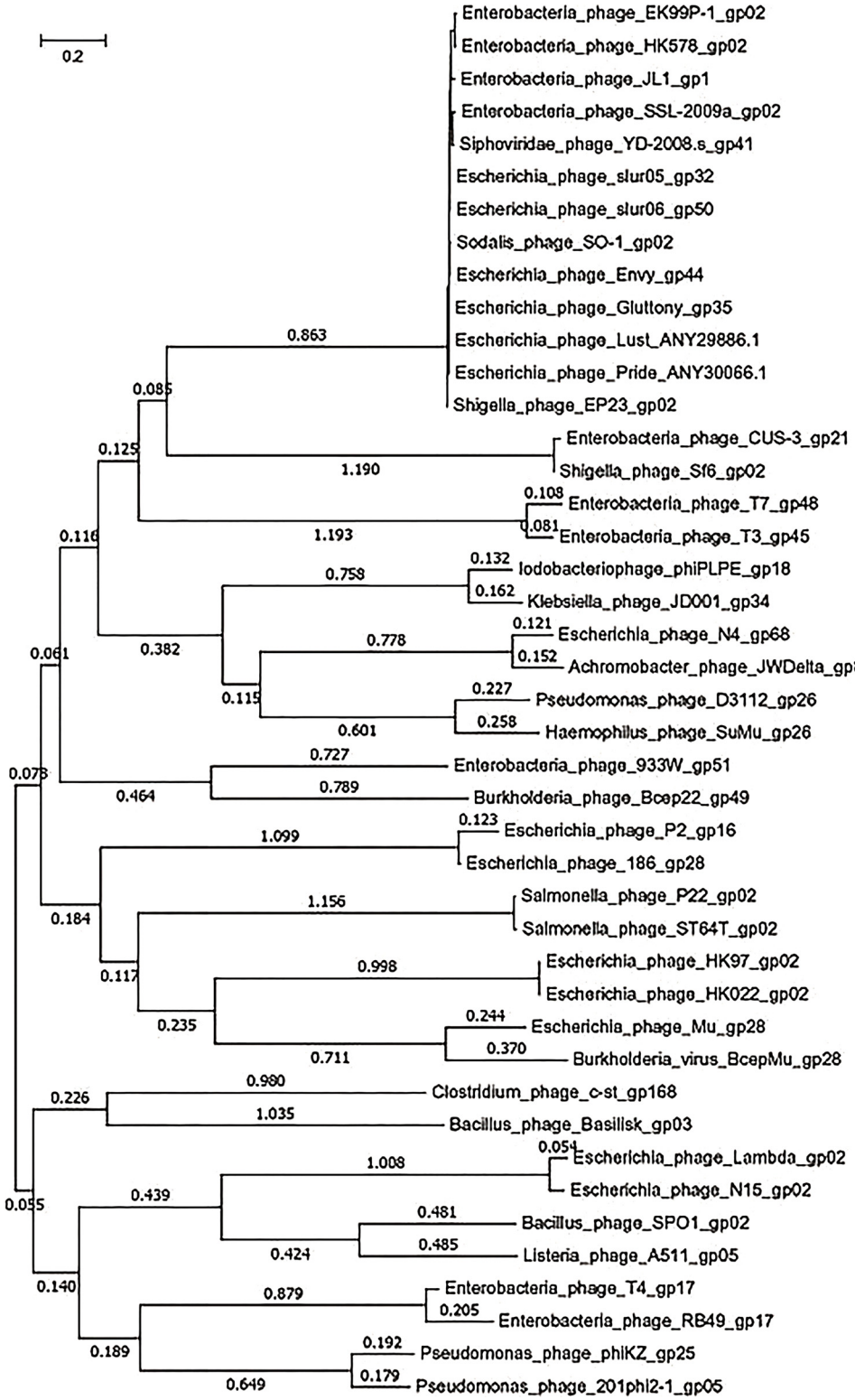

Fig. 4. The phylogenetic analysis of the large terminase subunits of the 13 '*HK578likevirus*' phages. The diagram was constructed using the MEGA6 program, and the evolutionary relationship was inferred using the neighbor-joining method. The tree was drawn to scale, and the length of each branch is associated with the relative evolutionary distance.

be packaged into a procapsid and cut into genome-sized lengths rely on the ATPase and endonuclease activities of terminase proteins (Black, 1989). In general, the terminase proteins of dsDNA viruses are composed of a small subunit and a large subunit. The DNA packaging modules of '*HK578likevirus*' phages are also composed of two terminase subunits with high identity in amino acid sequence (Table S7). Two conserved terminase domains (PF03237, Terminase_6 and CL12054 Terminase_3) were found in the sequences of all 13 terminase large subunits. The DNA packaging strategies of tailed dsDNA bacteriophages can be further classified into 15 subtypes according to the chromosomal termini of bacteriophages: 5′cos (Lambda, P2); 3′ cos (HK97), headful (P2, P22, Sf6, T4, 933 W, phiPLPE, phiKZ), host ends (Mu, D3112),

short DTRs (T7, N4, C-st) and long DTRs (SPO1) (Merrill et al., 2016). Generally, the packaging strategy of a tailed bacteriophage can be predicted according to the amino acid sequences of its large terminase subunit (Merrill et al., 2016). The phages who possess terminases with similar amino acid sequences usually package DNA using similar mechanism (Merrill et al., 2016; Casjens and Gilcrease, 2009). In order to determine the probable packaging strategy employed by the '*HK578likevirus*' bacteriophages, the 13 large terminase subunits were compared, in amino acid sequence, to those of bacteriophages with predicted packaging strategies using phylogenetic analysis. As showed in Fig. 4, the large terminase subunits of 13 bacteriophages showed high identities with each other in the amino acid sequences, which suggest that they package their replicated chromosomal DNA with similar strategy. However, the 13 phages are separated from the other known phage groups and uniquely positioned in the phylogenetic tree (Fig. 4), indicating that the '*HK578likevirus*' phages may use a novel genome packaging strategy that differs from those known strategies.

3.3.5. The lytic strategy of 'HK578likevirus' phages

At the end of a phage lytic cycle, it is necessary for the progenies to lyze the cell wall, especially the rigid peptidoglycan (PG) layer, to release from infected host cell. To destroy host cell wall, phages usually use a holin-endolysin strategy: holins form holes in bacterial cytoplasmic membrane and disrupt it, and then lysis begins. Endolysins target the integrity of the PG layer (Pimentel, 2014). Holins are small membrane proteins with at least one transmembrane domain (TMD) and a hydrophilic C-terminus. According to the number of TMDs and sequence similarity, holins can be classified into three classes (Class I-III) (Wang et al., 2000). Some phages also encode an anti-holin, a negative regulator of holin function contributing to the precise regulation of lysis (Wang et al., 2000). In some cases the anti-holin is encoded by holin gene, while sometimes it is encoded by a separated gene (Pimentel, 2014). Endolysins, small enzymes that degrade cell wall peptidoglycan, can be classed into four groups according to their activity against the three covalent linkages that maintain the integrity of cell wall: (i) glycosylase, (ii) transglycosylase targeting the glycosidic linkages, (iii) amidase, and (iv) endopeptidase targeting the oligopeptide cross-linkages (Young et al., 2000). For Gram-negative bacteria, Spanins, a third functional class of lysis proteins, are responsible for elimination of the final outer membrane barrier (Pimentel, 2014).

The protein annotation indicated that all 13 bacteriophages use the holin-endolysin strategy to lyse host cells. Three adjacent genes are included in their lysis cassettes: two separately encoded for holin-like class Iand class II proteins, and the other encoded for lysozyme. Holins disrupt cytoplasmic membrane by forming holes in it, while lysozyme degrade cell wall PG layer. However, although all the host cells of these phages belong to Gram-negative bacteria, no gene with predicted function was found to encode spanins which destroy the final outer membrane of Gram-negative bacteria. Furthermore, no typical dual start motifs were found in the amino acid sequence of the 13 holin proteins, which indicated that the holin inhibitors of these phages may be encoded by separate genes as those of phages P1 and P2 (Ziermann et al., 1994; Schmidt et al., 1996). The regulation of '*HK578likevirus*' phages cellular lysis awaits further analysis and research.

4. Conclusion

In this study, the re-squenced genome of Enterobacteria phage SSL-2009a was detail analysised at nucleotide and amino acid levels. SSL-2009a is a virulent bacteriophage with a circular double-stranded DNA of 44,899 base pairs in length. Twenty-four of 67 predicted protein coding sequences encode products highly homologous to known phage proteins. Ten promoters and 22 terminators are identified in its genome, but no tRNA is found. SSL-2009a belongs to the '*HK578likevirus*' genus of *Siphoviridae*. Other twelve phages showed high homology with it at genomic and proteomic levels, and also should be grouped into the same genus. According to the analysis of these 13 bacteriophages, it was declared that '*HK578likevirus*' phage has an icosahedric head and a long flexible noncontractile tail. Their genome sizes range from 41.9 kb to 45.2 kb (with an average GC content of 54.42% ~ 54.77%), encoding 57 ~ 67 putative CDSs. There are 9 ~ 16 promoters, and 17 ~ 24 terminators identified in these genomes, but no tRNA is found. '*HK578likevirus*' phages use the holin-endolysin strategy to lyse host cells, and the lysis cassette is composed of a holin-like class I, a holin-like class II protein, and a lysozyme. The packaging strategies of '*HK578likevirus*' phages are found to be different from the known strategies. Although there are only 13 reported '*HK578likevirus*' phages up to now, more and more phages will be identified and grouped into this genus in the future. In-deep analysis on these phages would promote our understanding of this *Siphoviridae* genus, even the high diversity of the phage community and the genetic traits involved in phage evolution.

Additional information

No competing financial and non-financial interests exist in the submission of this manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grants 31470241, 31570127, 31170159, and 31400163) and the Municipal Natural Science Foundation of Chongqing (Grant CSTC2011jjA10070).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2018.10.019.

References

Ackermann, H.W., 1998. Tailed bacteriophage: the order caudovirales. Adv. Virus Res. 51, 135–201.
Ackermann, H.W., 2009. Phage classification and characterization. Methods Mol. Biol. 501, 127–140.
Ackermann, H.W., Prangishvili, D., 2012. Prokaryote viruses studied by electron microscopy. Arch. Virol. 157, 1843–1849.
Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., Beatson, S.A., 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12, 402.
Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
Besemer, J., Borodovsky, M., 1999. Heuristic approach to deriving models for gene finding. Nucleic Acids Res. 27, 3911–3920.
Black, L.W., 1989. DNA packaging in dsDNA bacteriophages. Annu. Rev. Microbiol. 43, 267–292.
Blattner, F.R., Plunkett 3rd., G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453–1462.
Brown-Jaque, M., Calero-Cáceres, W., Muniesa, M., 2015. Transfer of antibiotic-resistance genes via phage-related mobile elements. Plasmid 79 (C), 1–7.
Brussow, H., Hendrix, R.W., 2002. Phage genomics: small is beautiful. Cell 108, 13–16.
Casas, V., Rohwer, F., 2007. Phage metagenomics. Methods Enzymol. 421, 259–268.
Casjens, S.R., Gilcrease, E.B., 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol. Biol. 502, 91–111.
Chang, H.W., Kim, K.H., 2011. Comparative genomic analysis of bacteriophage EP23 infecting *Shigellasonnei* and *Escherichia coli*. J. Microbiol. 49, 927–934.
Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31, 3497–3500.
Cisek, A.A., Dąbrowska, I., Gregorczyk, K.P., Wyżewski, Z., 2017. Phage therapy in bacterial infections treatment: one hundred years after the discovery of bacteriophages. Curr. Microbiol. 74 (2), 277–283.
Govind, R., Fralick, J.A., Rolfe, R.D., 2006. Genomic organization and molecular characterization of *Clostridium difficile* bacteriophage φCD119. J. Bacteriol. 188, 2568–2577.
Hendrix, R.W., Smith, M.C., Burns, R.N., Ford, M.E., Hatfull, G.F., 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage.

Proc. Natl. Acad. Sci. U. S. A. 96, 2192–2197.
Hendrix, R.W., 2003. Bacteriophage genomics. Curr. Opin. Microbiol. 6, 506–511.
Krumsiek, J., Arnold, R., Rattei, T., 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics 23, 1026–1028.
Lavigne, R., Sun, W.D., Volckaert, G., 2004. PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. Bioinformatics 20, 629–635.
Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H.W., Kropinski, A.M., 2008. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. Res. Microbiol. 159, 406–414.
Lavigne, R., Darius, P., Summer, E.J., Seto, D., Mahadevan, P., Nilsson, A.S., Ackermann, H.W., Kropinski, A.M., 2009. Classification of *Myoviridae* bacteriophages using protein sequence similarity. BMC Microbiol. 9, 224.
Li, K.B., 2003. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics 19, 1585–1586.
Li, S., Liu, L., Zhu, J., Zou, L., Li, M., Cong, Y., Rao, X., Hu, X., Zhou, Y., Chen, Z., Hu, F., 2010. Characterization and genome sequencing of a novel coliphage isolated from engineered Escherichia coli. Intervirology 53, 211–220.
Lin, D.M., Koskella, B., Lin, H.C., 2017. Phage therapy: an alternative to antibiotics in the age of multi-drug resistance. World J. Gastrointest. Pharmacol. Ther. 8 (3), 162–173.
Lu, S., Le, S., Tan, Y., Zhu, J., Li, M., Rao, X., Zou, L., Li, S., Wang, J., Jin, X., Huang, G., Zhang, L., Zhao, X., Hu, F., 2013. Genomic and proteomic analyses of the terminally redundant genome of the *Pseudomonas aeruginosa* phage PaP1: establishment of genus PaP1-like phages. PLoS One 8, e62933.
Malik, D.J., Sokolov, I.J., Vinner, G.K., Mancuso, F., Cinquerrui, S., Vladisavljevic, G.T., Clokie, M.R.J., Garton, N.J., Stapley, A.G.F., Kirpichnikova, A., 2017. Formulation, stabilisation and encapsulation of bacteriophage for phage therapy. Adv. Colloid Interface Sci. 249, 100–133.
Merrill, B.D., Ward, A.T., Grose, J.H., Hope, S., 2016. Software-based analysis of bacteriophage genomes, physical ends, and packaging strategies. BMC Genomics 17, 679.
Pan, F., Wu, H., Liu, J., Ai, Y., Meng, X., Meng, R., Meng, Q., 2013. Complete genome sequence of *Escherichia coli* O157: H7 lytic phage JL1. Arch. Virol. 158, 2429–2432.
Penadés, J.R., Chen, J., Quiles-Puchalt, N., Carpena, N., Novick, R.P., 2015. Bacteriophage-mediated spread of bacterial virulence genes. Curr. Opin. Microbiol. 23, 171–178.
Pimentel, M., 2014. Genetics of phage lysis. Microbiol. Spectr. 2 (1) MGM2-0017-2013.
Reese, M.G., 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput. Chem. 26, 51–56.
Salmond, G.P.C., Fineran, P.C., 2015. A century of the phage: past, present and future. Nat. Rev. Microbiol. 13 (12), 777–786.
Schattner, P., Brooks, A.N., Lowe, T.M., 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 33, W686–689.
Schmidt, C., Velleman, M., Arber, W., 1996. Three functions of bacteriophage P1 involved in cell lysis. J. Bacteriol. 178, 1099–1104.
Som, A., Fuellen, G., 2009. The effect of heterotachy in multigene analysis using the neighbor joining method. Mol. Phylogenet. Evol. 52, 846–851.
Stothard, P., Wishart, D.S., 2005. Circular genome visualization and exploration using CGView. Bioinformatics 21 (4), 537–539.
Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27 (7), 1009–1010.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Touchon, M., Moura de Sousa, J.A., Rocha, E.P., 2017. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. Curr. Opin. Microbiol. 38, 66–73.
Turner, D., Reynolds, D., Seto, D., Mahadevan, P., 2013. CoreGenes3.5: a webserver for the determination of core genes from sets of viral and small bacterial genomes. BMC Res. Notes 6, 140.
Wang, I.N., Smith, D.L., Young, R., 2000. Holins: the protein clocks of bacteriophage infections. Annu. Rev. Microbiol. 54, 799–825.
Wheeler, D.L., Church, D.M., Federhen, S., Lash, A.E., Madden, T.L., Pontius, J.U., Schuler, G.D., Schriml, L.M., Sequeira, E., Tatusova, T.A., Wagner, L., 2003. Database resources of the national center for biotechnology. Nucleic Acids Res. 31, 28–33.
Young, R., Wang, I.N., Roof, W.D., 2000. Phages will out: strategies of host cell lysis. Trends Microbiol. 8 (3), 120–128.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.
Ziermann, R., Bartlett, B., Calendar, R., Christie, G.E., 1994. Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. J. Bacteriol. 176, 4974–4984.