Arch Virol (2013) 158:2179–2183
DOI 10.1007/s00705-013-1700-0

ANNOTATED SEQUENCE RECORD

Complete genome sequence analysis of bacterial-flagellum-targeting bacteriophage chi

Ju-Hoon Lee · Hakdong Shin · Younho Choi · Sangryeol Ryu

Received: 14 February 2013 / Accepted: 19 March 2013 / Published online: 19 April 2013

Abstract Bacteriophage chi is a well-known phage that infects pathogens such as *E. coli*, *Salmonella*, and *Serratia* via bacterial flagella. To further understand its host-phage interaction and infection mechanism via host flagella, the genome was completely sequenced and analyzed. The phage genome contains 59,407-bp-length DNA with a GC content of 56.51 %, containing 75 open reading frames (ORFs) with no tRNA genes. Its annotation and functional analysis revealed that chi is evolutionarily very closely related to *Enterobacter* phage Enc34 and *Providencia* phage Redjac. However, most of the annotated genes encode hypothetical proteins, indicating that further genomic study of phage chi is required to elucidate the bacterial-flagellum-targeting infection mechanism of phage chi.

Intake of food-borne pathogens such as *E. coli* and *Salmonella* via contaminated foods causes food poisoning accompanied by high fever, diarrhea and vomiting [2, 14, 19]. Although many food preservatives have been developed and used to control these food-borne pathogens, the number of food poisoning outbreaks is increasing every year [10, 11, 20]. Therefore, effective and safe novel biocontrol agents should be developed to control food-borne pathogens.

Bacteriophages are bacterial viruses that infect and lyse specific bacterial host cells, suggesting their bactericidal activity [6]. In addition, they infect only specific host bacteria without affecting other bacteria in the same habitat [7]. Recently, human feeding trials showed efficient inhibition of specific bacterial hosts without side effects, suggesting that phage treatment should be safe for human applications [4]. Therefore, phage applications have been reconsidered and tested as alternative approaches to inhibit food-borne pathogens in foods [5, 12, 22].

Bacteriophage chi, which infects major food-borne pathogens such as *E. coli*, *Salmonella*, and *Serratia*, was first isolated and characterized in 1930s [25]. While other bacteriophages generally infect host strains via extracellular membrane receptors such as lipopolysaccharide (LPS) and outer membrane proteins (like BtuB, FhuA, and OmpC) [17], phage chi is the first reported bacteriophage to infect host strains via flagella [21]. However, the infection mechanism of chi phage via the host flagellum is not yet fully understood at the genomic level. In this study, to further understand this receptor specificity and host-phage interaction, the genome of phage chi was completely sequenced and analyzed.

The chi phage (ATCC 9842-B1^{TM}) was obtained from the American Type Culture Collection (ATCC). For propagation of phage chi, it was added to a culture of *Salmonella enterica* serovar Typhimurium SJW1103 [27] at a multiplicity of infection (MOI) of 1 when the optical density (OD) of the culture at a wavelength of 600 nm reached 1.0. The mixture was incubated at 37 °C for 4 h with vigorous shaking, and phage particles were recovered

Electronic supplementary material The online version of this article (doi:10.1007/s00705-013-1700-0) contains supplementary material, which is available to authorized users.

J.-H. Lee
Department of Food Science and Biotechnology, Kyung Hee University, Yongin 446-701, Korea

H. Shin · Y. Choi · S. Ryu (✉)
Department of Food and Animal Biotechnology, Department of Agricultural Biotechnology and Center for Food and Bioconvergence, Seoul National University, Seoul 151-921, Korea
e-mail: sangryu@snu.ac.kr

by centrifugation at 6,000×*g* for 10 min and subsequent filtration using 0.22-μm-pore-size filters (Millipore, Billerica, MA, USA). To purify the phage particles, precipitation with polyethylene glycol (PEG) 6,000 (Sigma, St. Louis, MO, USA) was carried out, followed by ultracentrifugation (Himac CP 100β, Hitachi, Japan) in a 1.3 to 1.7 g/ml $CsCl_2$ density gradient at 25,000×*g* and 4 °C for 2 h.

The genomic DNA of phage chi was isolated as described previously by Wilcox et al. [26]. Prior to isolation of phage genomic DNA, phage particles were treated with DNase I and RNase A at 37 °C for 1 h to remove bacterial host DNA and RNA, respectively. The phage particles were then lysed with standard lysis buffer (50 μg of proteinase K per ml, 0.5 % sodium dodecyl sufate (SDS), and 20 mM EDTA) for 2 h at 56 °C. In the final step, phenol–chloroform treatment and ethanol precipitation of genomic DNA were conducted as described by Sambrook et al. [23].

Purified genomic DNA of phage chi was sheared and randomly sequenced using a Genome Sequencer FLX (GS-FLX) (Roche, Mannheim, Germany), and the qualified filtered reads were assembled using the Newbler 2.3 program (Roche) at Macrogen, Inc. (Seoul, South Korea). Open reading frames (ORFs) were predicted using gene prediction programs such as Glimmer3 [13], GeneMarkS [3], and FgenesB (Softberry, Inc. Mount Kisco, NY, USA) and confirmed using the RBSFinder program (J. Craig Venter Institute, Rockville, MD, USA). Their annotation and functional analysis were performed using the BLASTP [1] and InterProScan [28] databases. Genomic DNA and annotation data were handled and edited using Artemis14 [8]. Phylogenetic analysis of major capsid proteins (MCPs) of bacteriophages, including phage chi, was conducted using MEGA5 based on the neighbor-joining method with *P*-distance values [15]. The lifestyle of phage chi was predicted using PHACTS program [18].

Bacteriophage chi genome contains 59,407-bp-length DNA with a GC content of 56.51 %, containing 75 ORFs with no tRNA genes (Fig. 1). Annotated functions of all predicted ORFs in phage chi are listed in Table S1. The average gene length is 748 bp, and the gene coding percentage is 94.5 %. The predicted functions of ORFs in phage chi were classified into five functional groups: structure (head-tail joining protein [chi_053], decorator protein [chi_0056], major capsid protein [chi_057], tape measure protein [chi_065], tail assembly proteins 1 and 2 [chi_067 and chi_068], tail fiber protein [chi_071], and prohead protease [chi_055]), packaging (terminase small and large subunits [chi_051 and 052, respectively], phage portal protein [chi_054]), host lysis (lysis protein A [chi_003] and B [chi_002], endolysin-like protein [chi_004], and Rz1 protein [chi_005]), DNA manipulation (recombination associated protein [chi_023], primase [chi_042], DNA polymerase I [chi_048], and helicase [chi_050]), and additional function (N-6-adenine-methyltransferase [chi_017]).

BLASTP analysis of the functional ORFs showed that this phage genome is very similar to those of *Enterobacter* phage Enc34 and *Providencia* phage Redjac (Table 1). Interestingly, phage head proteins are very similar to those of *Enterobacter* phage Enc34, with 66 to 92 % protein sequence identity. Furthermore, host lysis proteins are also similar to those of *Enterobacter* phage Enc34, with 57 to 76 % protein sequence identity. However, phage tail proteins are very similar to those of *Providencia* phage Redjac, with 70 to 90 % protein sequence identity, suggesting that phage chi structural genes may be derived from a common ancestor. BLASTP best matches of DNA manipulation genes are mixed with those of two different bacteriophages, supporting this hypothesis (Table 1). In addition, further phylogenetic analysis of phage chi and other bacteriophages based on major capsid proteins (MCPs) revealed that phage chi is evolutionarily very closely related to these phages, *Enterobacter* phage Enc34 and *Providencia* phage Redjac, confirming their close evolutionary relationship (Fig. 2). To further elucidate the type of phage chi, additional phylogenetic analysis of

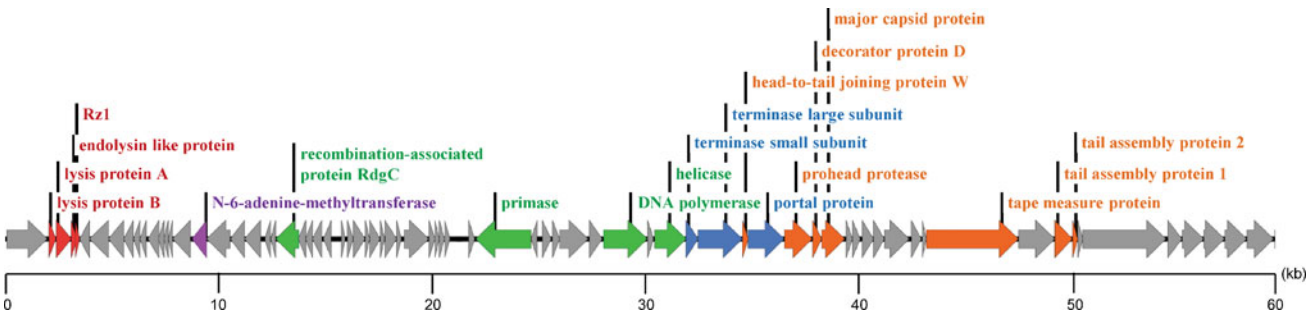

Fig. 1 Genome map of bacteriophage chi. Functional ORFs were classified into five groups. *Red*, *purple*, *green*, *blue*, and *orange arrows* indicate host lysis, additional function, DNA manipulation, packaging, and structure-related ORFs, respectively. The scale unit is kilobase pairs (kb) (color figure online)

Table 1 Comparative analysis of predicted ORFs using BLASTP

Locus_tag	Predicted function	Length[a]	BLASTP best match	Length[a]	Identity (%)[b]	GenBank accession no.
chi_002	Lysis protein B	112	Holin [Enterobacter phage Enc34]	114	64/112 (57.1)	YP_007007037.1
chi_003	Lysis protein A	237	Endolysin [Enterobacter phage Enc34]	242	181/237 (76.4)	YP_007007038.1
chi_004	Endolysin like protein	84	Putative Rz protein [Enterobacter phage Enc34]	84	56/84 (66.7)	YP_007007039.1
chi_005	Possible Rz1 protein	67	Putative Rz1 protein [Enterobacter phage Enc34]	66	51/67 (76.1)	YP_007007040.1
chi_017	Possible N-6-adenine-methyltransferase	228	DNA methyltransferase [Enterobacter phage Enc34]	234	173/228 (75.9)	YP_007007051.1
chi_023	Possible recombination associated protein RdgC	357	Recombination-associated protein [Enterobacter phage Enc34]	355	208/357 (58.3)	YP_007007059.1
chi_041	Helix-turn-helix domain-containing protein	95	Hypothetical protein PaP1_gp024 [Pseudomonas phage PaP1]	103	23/95 (24.2)	YP_007236435.1
chi_042	Possible primase	861	DNA primase [Enterobacter phage Enc34]	864	624/861 (72.5)	YP_007007000.1
chi_048	Putative DNA polymerase I	679	DNA polymerase I [Providencia phage Redjac]	678	529/679 (77.9)	YP_006906013.1
chi_049	VRR-NUC domain-containing protein	95	VRR-NUC domain protein [Enterobacter phage Enc34]	94	62/95 (65.3)	YP_007007007.1
chi_050	Possible helicase	491	DNA helicase [Providencia phage Redjac]	502	382/491 (77.8)	YP_006906014.1
chi_051	Possible terminase small subunit	189	Terminase small subunit [Enterobacter phage Enc34]	191	161/189 (85.2)	YP_007007009.1
chi_052	Putative terminase large subunit	691	Terminase large subunit [Providencia phage Redjac]	691	624/691 (90.3)	YP_006906015.1
chi_053	Possible head-tail joining protein Lambda W	84	Head-to-tail joining protein [Enterobacter phage Enc34]	83	56/84 (66.7)	YP_007007011.1
chi_054	Phage portal protein, lambda family	560	Phage portal protein [Providencia phage Redjac]	588	509/560 (90.9)	YP_006906016.1
chi_055	Putative prohead protease ClpP	428	ClpP [Providencia phage Redjac]	431	297/428 (69.4)	YP_006906017.1
chi_056	Possible decorator protein	139	Phage structural protein [Enterobacter phage Enc34]	139	109/139 (78.4)	YP_007007014.1
chi_057	Putative major capsid protein	354	Major capsid protein E [Enterobacter phage Enc34]	354	326/354 (92.1)	YP_007007015.1
chi_062	Bacterial Ig-like domain-containing protein	381	Phage structural protein [Providencia phage Redjac]	379	304/381 (79.8)	YP_006906023.1
chi_065	Putative tape measure protein	1431	Tape measure protein [Providencia phage Redjac]	1435	1002/1431 (70.0)	YP_006906025.1
chi_067	Putative conserved tail assembly protein 1	272	Conserved tail assembly protein [Providencia phage Redjac]	272	247/272 (90.8)	YP_006905985.1
chi_068	Putative conserved tail assembly protein 2	76	Tail assembly protein [Enterobacter phage Enc34]	78	60/76 (78.9)	YP_007007027.1
chi_071	Possible tail fiber protein	221	Tail fiber protein [Providencia phage Redjac]	245	162/221 (73.3)	YP_006905987.1

[a] Base pairs (bp)

[b] Amino acid sequence identity

phage chi was performed based on the terminase large subunit following the method of Casjens and Gilcrease's [9], and the result of this analysis showed that phage chi belongs to "λ-like 5′-extended COS ends" group (Fig. S1). To predict the lifestyle of phage chi, PHACTS analysis was conducted with amino acid sequences of all predicted ORFs. However, a clear lifestyle prediction was not possible for phage chi, probably due to the extremely low

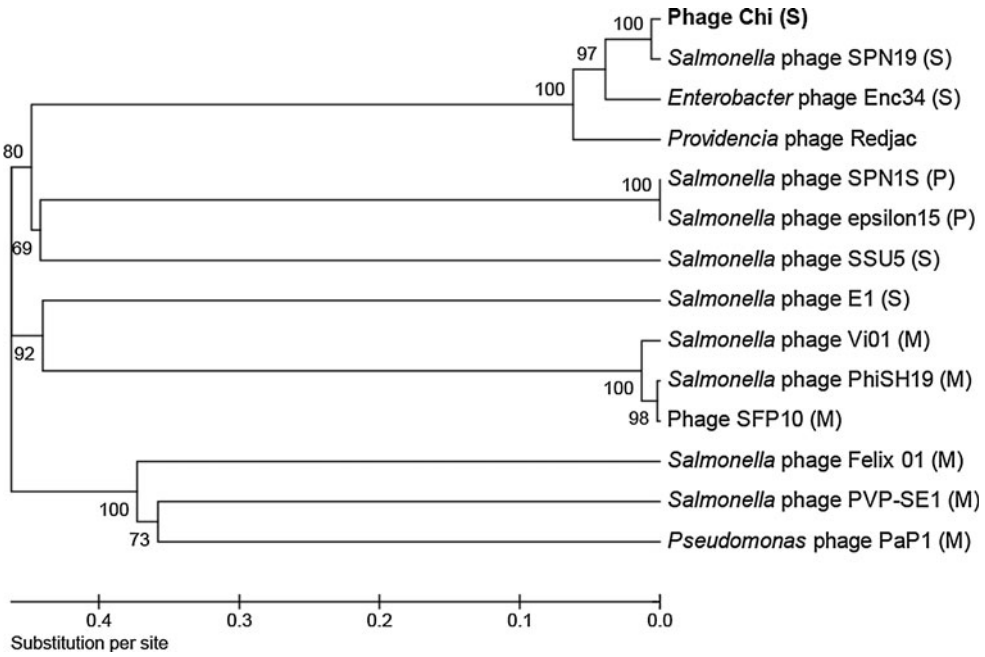

Fig. 2 Phylogenetic analysis of MCPs in phage chi and various other bacteriophages. MCPs were compared by ClustalW alignment [16], and the phylogenetic tree was generated by the neighbor-joining method with *P*-distance values using MEGA5 [15]. S, *Siphoviridae*; P, *Podoviridae*; M, *Myoviridae*

amino acid sequence identities of predicted ORFs in chi phage to those of other phages (data not shown).

While the bacterial-flagellum-mediated infection mechanism of phage chi has been suggested to follow a "nut and bolt" model, infecting through counterclockwise rotating flagella [24], the infection mechanism of phage chi via host flagella based on its complete genome sequence analysis is not clearly understood yet, probably due to insufficient database information on the bacterial-flagellum-targeting bacteriophages. The genome annotation result showed that 52 of the 75 predicted ORFs encode hypothetical proteins, supporting this. Therefore, further functional genome studies of phage chi will be needed to explain the mechanism of infection via host flagella.

Nucleotide sequence accession number. The complete genome sequence of bacteriophage chi is available in the GenBank database under accession number JX094499.

Acknowledgments This research was supported by WCU (World Class University) program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology (R32-2008-000-10183-0) and the R&D Convergence Center Support Program of the Ministry for Food, Agriculture, Forestry and Fisheries, Republic of Korea.

References

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
2. Barbara G, Stanghellini V, Berti-Ceroni C, De Giorgio R, Salvioli B, Corradi F, Cremon C, Corinaldesi R (2000) Role of antibiotic therapy on long-term germ excretion in faeces and digestive symptoms after *Salmonella* infection. Aliment Pharmacol Ther 14:1127–1131
3. Besemer J, Lomsadze A, Borodovsky M (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618
4. Bruttin A, Brussow H (2005) Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy. Antimicrob Agents Chemother 49:2874–2878
5. Cairns BJ, Payne RJH (2008) Bacteriophage therapy and the mutant selection window. Antimicrob Agents Chemother 52:4344–4350
6. Calendar R (2006) The bacteriophages, 2nd edn. Oxford University Press, Oxford
7. Campbell A (2003) The future of bacteriophage biology. Nat Rev Genet 4:471–477
8. Carver T, Berriman M, Tivey A, Patel C, Bohme U, Barrell BG, Parkhill J, Rajandream MA (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24:2672–2676
9. Casjens SR, Gilcrease EB (2009) Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol Biol 502:91–111
10. CDC (2007) Bacterial foodborne and diarrheal disease national case surveillance: annual report, 2005. Centers for Disease Control and Prevention, Atlanta
11. CDC (2008) *Salmonella* surveillance: annual summary, 2006. Centers for Disease Control and Prevention, Atlanta
12. Coffey B, Mills S, Coffey A, McAuliffe O, Ross RP (2010) Phage and their lysins as biocontrol agents for food safety applications. Ann Rev Food Sci Technol 1:449–468
13. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679
14. Karmali MA, Steele BT, Petric M, Lim C (1983) Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. Lancet 1:619–620

15. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9:299–306
16. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948
17. Lindberg AA (1973) Bacteriophage receptors. Ann Rev Microbiol 27:205–241
18. McNair K, Bailey BA, Edwards RA (2012) PHACTS, a computational approach to classifying the lifestyle of phages. Bioinformatics 28:614–618
19. Mead PS, Griffin PM (1998) *Escherichia coli* O157:H7. Lancet 352:1207–1212
20. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV (1999) Food-related illness and death in the United States. Emerg Infect Dis 5:607–625
21. Meynell EW (1961) A phage, øχ, which attacks motile bacteria. J Gen Microbiol 25:253–290
22. Payne RJH, Jansen VAA (2000) Phage therapy: The peculiar kinetics of self-replicating pharmaceuticals. Clin Pharmacol Ther 68:225–230
23. Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, NY
24. Samuel ADT, Pitta TP, Ryu WS, Danese PN, Leung ECW, Berg HC (1999) Flagellar determinants of bacterial sensitivity to χ-phage. P Natl A Sci USA 96:9863–9866
25. Sertic V, Boulgakov N-A (1936) Bactnriophages specifique pour des varites bacteriennes flagellees. C R Soc Biol Paris 113: 105–113
26. Wilcox SA, Toder R, Foster JW (1996) Rapid isolation of recombinant lambda phage DNA for use in fluorescence in situ hybridization. Chromosome Res 4:397–398
27. Yamaguchi S, Fujita H, Sugata K, Taira T, Iino T (1984) Genetic analysis of H2, the structural gene for phase-2 flagellin in *Salmonella*. J Gen Microbiol 130:255–265
28. Zdobnov EM, Apweiler R (2001) InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847–848