**Standards in
Genomic Sciences**

SHORT GENOME REPORT

Open Access

CrossMark

Complete genome of a novel virulent phage ST0 lysing *Escherichia coli* H8

Honghui Liu, Xinchun Liu[*] and Jinqing Li

Abstract

Phage ST0 lysing *Escherichia coli* H8 was isolated from wastewater and sequenced using an Illumina HiSeq system. Genomic analyses revealed that it was virulent phages and contained a circular double-stranded DNA genome, consisting of 170,496 nucleotides with an average G + C content of 37.67%. This study may provide possible alternative materials for phage therapy.

Keywords: *Escherichia coli*, Virulent phage, Complete genome, Phage therapy, Antimicrobial agents

Introduction

A large number of antibiotics were produced and widely used in medical and agricultural areas. These substances in the environment didn't tend to be biodegradable, and were easily stored and accumulated in water and soil environment and even in the atmospheric environment [1–3]. Recently, antibiotics had been recognized as the emerging environmental pollutants, because of their potential undesirable effects on the ecosystem and human health [4–6], such as antibiotic resistance. The resistance of bacteria to current antibiotics increased the difficulty in medical treatment, which accounted for 23,000 deaths annually in the USA. The spread of antibiotic resistant bacteria and antibiotic resistance genes in the environment was a major public health issue. Obviously, strict control of the use of antibiotics and the development of a possible alternative medicine seemed extremely urgent.

Compared with antibiotics, phage therapy had the advantages of high specificity, few side effects and capacity for low-dosage use and so on [7]. In particular, it was alternative and effective to adopt phage therapy to treat diseases caused by antibiotics-resistant bacteria strains [8]. Currently phage therapy mainly had single phage treatment, multiple phage treatment and combined therapy of phage and antibiotics. More recently, bacteriophages had been intensively studied and potential application for the control of *Escherichia coli* in livestock, aquaculture and food products [9–11].
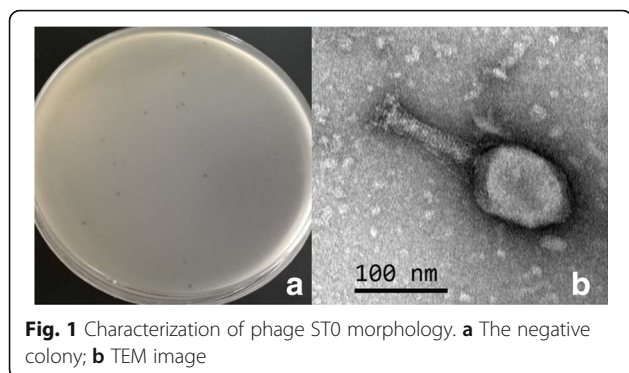
In this work, phage ST0 against *Escherichia coli* H8 was isolated from industrial wastewater in China. Its morphology, complete genome sequence and bioinformatics analysis were explored. This could provide a better understanding to the development of a possible alternative medicines and biocontrol agents.

Organism information
Classification and features

Escherichia coli H8 (ST100), the host to isolate virulent phages, carrying shiga toxin genes (*stx1*, *stx2*) was obtained from the Chinese Center for Disease Control and Prevention. Phage ST0 was isolated from a sewage treatment plant of wastewater in Beijing. The isolation, propagation and titration of phage was done according to the methods described previously [12]. Phage ST0 generated clear plaques on double-layer plate (Fig. 1a), indicating that it was a virulent phage. The diameter of plaques was 1–2 mm. The transmission electron microscopy image (Fig. 1b) showed that phage ST0 had an icosahedral head approximately 120 nm in long diameter and 80 nm in short diameter. Its long tail was about 120 nm in length and 20 nm in diameter. Phylogenetic analysis based on complete genome sequences revealed that phage ST0 was closely related to *Enterobacteria* phage HX01 (Accession JX536493.1), whereas the score was relatively low (Fig. 2). A summary of the isolation

* Correspondence: xcliu@ucas.ac.cn
College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

Liu *et al. Standards in Genomic Sciences* (2017) 12:85

Page 2 of 6

Fig. 1 Characterization of phage ST0 morphology. **a** The negative colony; **b** TEM image

and general phylogenetic features of phage ST0 are shown in Table 1.

Genome sequencing information
Genome project history
Phage ST0 infecting *Escherichia coli* was isolated and sequenced because of its potential for use in phage therapy. The genome sequence and annotation are available in GenBank (MF044457). These data were summarized in Table 2.

Growth conditions and genomic DNA preparation
Phage ST0 was isolated from a wastewater sample that was filtered through a 0.22-μm polycarbonate membrane filter (Millipore, Bedford, MA, USA). The host strain *Escherichia coli* H8 was cultured at 37 °C using LB medium [13]. Phage DNA was extracted as described by Sambrook and Russell [14]. The phage lysates were concentrated in polyethylene glycol 8000 and bacterial nucleic acids were removed from phage lysates by DNase I (Sigma-Aldrich, Oakville, Canada) and RNaseA (Sigma-Aldrich). Then the phage particles were amplified and stored in SM buffer (100 mM NaCl, 8 mM MgSO4, 50 mM Tris-HCl [pH 7.5]) at 4 °C.

Genome sequencing and assembly
DNA was sequenced using the Illumina HiSeq 2500 platform in Beijing Fixgene Tech Co., Ltd. More than 5000-fold coverage of the phage genome is generated by sequencing the cloned fragments. The paired-end reads were assembled using the abyss v. 1.3.6. Possible tRNAs in the genome were determined using tRNAscan-SE. These data were summarized in Table 3.

Genome annotation
The potential ORFs were predicted using PHASTER [15]. Putative protein function of ORFs was annotated by BLASTp against NCBI database and HMMER search against the COG database [16] (These data were summarized in Table 4). The map of a circular representation of phage ST0 genome was generated using CGView Server. Neighbor joining tree was drawn by MEGA 5.05 [17].

Genome properties
The complete genome sequence of phage ST0 had been deposited in GenBank with the accession number MF044457. The complete genome sequence of phage ST0 consisted of 170,496 bp and was circular double-stranded DNA with an average GC content of 37.67%. There were ten tRNAs detected in this genome indicating that phage ST0 could be reliant on its tRNAs after entering into the hosts.

A total of 269 ORFs were predicted in this complete genome, compared with those from the NCBI database (Fig. 3; Additional file 1: Table S1). These ORFs showed more than 94% identity with 18 different phage strains. Of those, 41 ORFs were predicted in the minus strand and others were in the plus strand. Eighty nine putative ORFs were predicted to have unkown functions.

Insights from the genome sequence
Phage ST0 possessed replication-related genes encoding DNA polymerase (ORF64), DNA primase/helicase
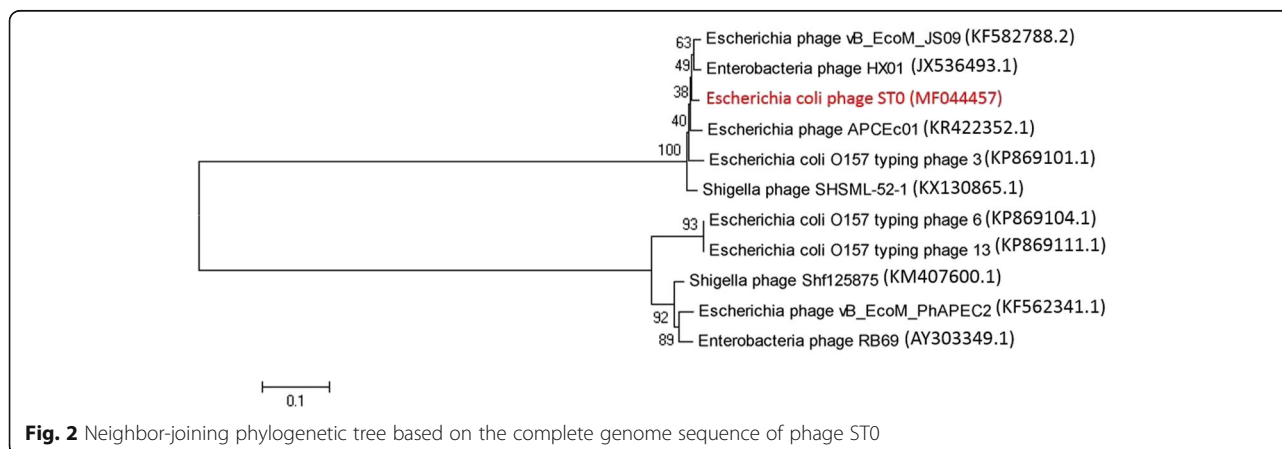

Fig. 2 Neighbor-joining phylogenetic tree based on the complete genome sequence of phage ST0

Liu *et al. Standards in Genomic Sciences* (2017) 12:85

Page 3 of 6

Table 1 Classification and general features of *Genusspecies* strain designation[T] [19]

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain: Viruses, dsDNA viruses	TAS [21]
		Phylum: unassigned	
		Class: unassigned	
		Order: *Caudovirales*	TAS [21]
		Family: *Siphoviridae*	TAS [21]
		Genus: unassigned	
		Species: unassigned	
		(Type) strain: unassigned	
	Gram stain	N/A	
	Cell shape	N/A	
	Motility	N/A	
	Sporulation	N/A	
	Temperature range	N/A	
	Optimum temperature	N/A	
	pH range; Optimum	N/A	
	Carbon source	N/A	
MIGS-6	Habitat	Water	IDA
MIGS-6.3	Salinity	N/A	
MIGS-22	Oxygen requirement	N/A	
MIGS-15	Biotic relationship	Intracellular parasite of *Escherichia coli* H8	IDA
MIGS-14	Pathogenicity	Lytic phage of *Escherichia coli* H8	IDA
MIGS-4	Geographic location	China	IDA
MIGS-5	Sample collection	April, 2017	IDA
MIGS-4.1	Latitude	40°N	IDA
MIGS-4.2	Longitude	116°E	IDA
MIGS-4.4	Altitude	Unknown	

[a]Evidence codes - *IDA* inferred from direct assay, *TAS* traceable author statement (i.e., a direct report exists in the literature). These evidence codes are from the Gene Ontology project [20]

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Complete
MIGS-28	Libraries used	Illumina library
MIGS 29	Sequencing platforms	Illumina HiSeq 2500
MIGS 31.2	Fold coverage	5000-folds
MIGS 30	Assemblers	Abyss 1.3.6
MIGS 32	Gene calling method	PHASTER
	Locus Tag	N/A[a]
	Genbank ID	MF044457
	GenBank Date of Release	28-JUN-2017
	GOLD ID	N/A[a]
	BIOPROJECT	N/A[a]
MIGS 13	Source Material Identifier	N/A[a]
	Project relevance	Isolation and application of phages infecting *Escherichia coli*

[a]Not available

Liu *et al. Standards in Genomic Sciences* (2017) 12:85

Page 4 of 6

Table 3 Genome statistics

Attribute	Value	% of Total
Genome size (bp)	170,496	100.00
DNA coding (bp)	157,610	92.44
DNA G + C (bp)	64,218	37.67
DNA scaffolds	0	100.00
Total genes	269	100.00
Protein coding genes	269	100.00
RNA genes	12	0.55
Pseudo genes	0	0.00
Genes in internal clusters	0	0.00
Genes with function prediction	89	33.09
Genes assigned to COGs	121	44.98
Genes with Pfam domains	0	0.00
Genes with signal peptides	0	0.00
Genes with transmembrane helices	0	0.00
CRISPR repeats	0	0.00

(ORF76, ORF87, ORF101 and ORF195), DNA ligase (ORF179), topoisomerase (ORF110, ORF121), DNA binding protein (ORF139, ORF142), terminase (ORF211 and ORF212) and other related proteins (ORF7, ORF18, ORF33, ORF34, ORF60-62, ORF89, ORF138, ORF146-147, ORF149-151, ORF159, ORF166, ORF231 and ORF259). Terminase was observed in the phage ST0 genome, and plays the essential role in the double-stranded DNA packaging process. Terminase generally composed of two subunits identifies the pre-capsid protein and the specific packaging sites, providing energy to packaging process through hydrolysis of ATP [17, 18]. This showed that phage ST0 possibly depended on its own terminase to obtain these kinds of function, while many phages lacked this enzyme.

Phage ST0 may be dependent on its own gene transcription and translation, because it possessed RNA-related enzymes such as tRNA synthetase modifier (ORF2), putative thioredoxin (ORF28), RNA polymerase sigma factor (ORF49), RNA ligase (ORF152 and ORF200) and other related proteins (ORF59, ORF63,

Table 4 Number of genes associated with general COG functional categories

Code	Value	%age	Description
J	2	0.74	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	7	2.60	Transcription
L	15	5.58	Replication, recombination and repair
B	0	0.00	Chromatin structure and dynamics
D	2	0.74	Cell cycle control, Cell division, chromosome partitioning
V	2	0.74	Defense mechanisms
T	3	1.12	Signal transduction mechanisms
M	5	1.86	Cell wall/membrane biogenesis
N	0	0.00	Cell motility
U	0	0.00	Intracellular trafficking and secretion
O	6	2.23	Posttranslational modification, protein turnover, chaperones
C	2	0.74	Energy production and conversion
G	4	1.49	Carbohydrate transport and metabolism
E	6	2.23	Amino acid transport and metabolism
F	9	3.35	Nucleotide transport and metabolism
H	2	0.74	Coenzyme transport and metabolism
I	0	0.00	Lipid transport and metabolism
P	2	0.74	Inorganic ion transport and metabolism
Q	0	0.00	Secondary metabolites biosynthesis, transport and catabolism
R	0	0.00	General function prediction only
S	54	20.07	Function unknown
–	148	55.02	Not in COGs

The total is based on the total number of protein coding genes in the genome

Liu *et al. Standards in Genomic Sciences* (2017) 12:85

Page 5 of 6

Fig. 3 CGView Server map of phage ST0 complete genome
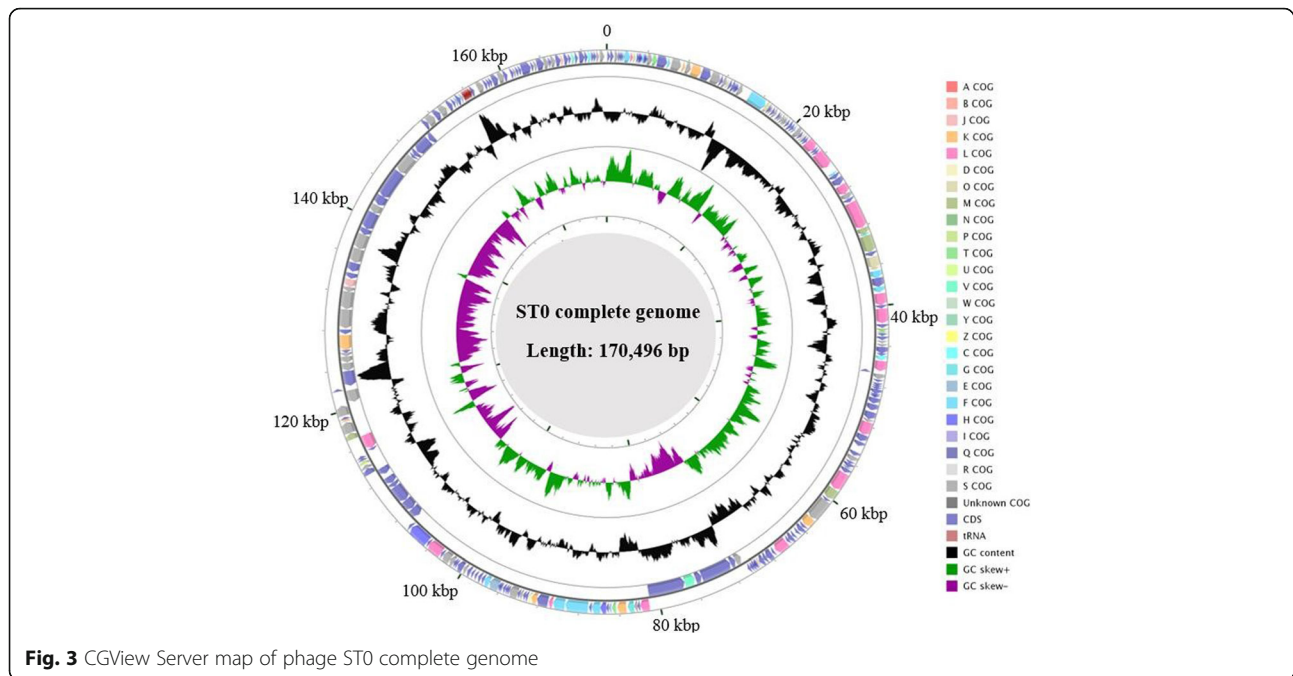
ORF123, ORF140, ORF165, ORF181) to regulate gene expression.

Forty one genes that were related to the structure of phage ST0 were identified in its genome, including phage virion protein and tail tube. The ORFs related to structure were mainly concentrated between 106,926 bp and 153,860 bp, which facilitated the rapid assembly of phages and also reflected the lowest energy principle in nature.

The presence of lysozyme (ORF257) and holing (ORF132) indicated that phage ST0 was a lytic phage. Moreover, other several active factors played a quite important role in host cell lysis and inhibition of host cell growth, such as exonuclease (ORF102), inhibitor of host transcription (ORF153), nudix hydrolase (ORF256) and rIIA protein (ORF113, ORF114 and ORF169).

Conclusions

The morphology, complete genome sequence and bio-informatics analysis showed phage ST0 was a novel virulent phage infecting and lysing *Escherichia coli* H8, which may provide a better understanding to the development of a possible alternative medicines and biocontrol agents.

Additional file

Additional file 1: Table S1. Predicted protein function of phage ST0. (DOCX 53 kb)

Abbreviations
TEM: Transmission electron microscopy

Authors' contributions
XL designed, guided the experiment and wrote this manuscript. HL was mainly responsible for the implementation of the experiment and writing this manuscript. JL provided the wastewater sample for isolating phage and assisted the experiment. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Daghrir R, Drogui P. Tetracycline antibiotics in the environment: a review. Environ Chem Lett. 2013;11:209–27.
2. Gandolfi I, Franzetti A, Bertolini V, Gaspari E, Bestetti G. Antibiotic resistance in bacteria associated with coarse atmospheric particulate matter in an urban area antibiotic resistance in atmosphere. J Appl Microbiol. 2011; 110(6):1612–20.
3. Tasho RP, Cho JY. Veterinary antibiotics in animal waste, its distribution in soil and uptake by plants: a review. Sci Total Environ. 2016;563-564:366–76.

Liu *et al. Standards in Genomic Sciences* (2017) 12:85

Page 6 of 6

4. Ding C, He J. Effect of antibiotics in the environment on microbial populations. Appl Microbiol Biotechnol. 2010;87:925–41.
5. Jechalke S, Heuer H, Siemens J, Amelung W, Smalla K. Fate and effects of veterinary antibiotics in soil. Trends Microbiol. 2014;22:536–45.
6. Coutinho FH, Silveira CB, Pinto LH, Salloto GRB, Cardoso AM, Martins OB, Vieira RP, Clementino MM. Antibiotic resistance is widespread in urban aquatic environments of Rio de Janeiro, Brazil. Microbiol Ecol. 2014;68:441–52.
7. Loc-Carrilloa C, Abedon ST. Pros and cons of phage therapy. Bacteriophage. 2011;1(2):111–4.
8. Kutter E, Vos DD, Gvasalia G, Alavidze Z, Gogokhia L, Kuhl S, Abedon ST. Phage therapy in clinical practice: treatment of human infections. Curr Pharm Biotechnol. 2010;11:69–86.
9. Mateus L, Costa L, Silva YJ, Pereira C, Cunha A, Almeida A. Efficiency of phage cocktails in the inactivation of *Vibrio* in aquaculture. Aquaculture. 2014;424-425:167–73.
10. Raul R, Rebecca AO, Ben MM, Wieland S, Callaway TR, Kutter E, Brabban AD. Naturally resident and exogenously applied T4-like and T5-like bacteriophages can reduce *E.coli* O157:H7 levels in sheep guts. Bacteriophage. 2011;1(1):15–24.
11. Hudson JA, Billington C, Cornelius AJ, Wilson T, On SLW, Premaratne A, King NJ. Use of a bacteriophage to inactivate *E.coli* O157:H7 on beef. Food Microbiol. 2013;36:14–21.
12. Adams MH. Bacteriophage. London, New York: Interscience Publishers, Inc.; 1959.
13. Bertani G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol. 1951;62(3):293–300.
14. Sambrook J, Russell D. Molecular cloning: a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor; 2001. p. 543–54.
15. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016. doi:10.1093/nar/gkw387.
16. Bryan MJ, Burroughs NJ, Spence EM, Clokie MRJ, Mann NH, Bryan SJ. Evidence for the intense exchange of MazG in marine cyanophages by horizontal gene transfer. PLoS One. 2008;3:e2048.
17. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28(10):2731–9.
18. Field D, Garrity GM, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol. 2008;26:541–7.
19. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000;25:25–9.
20. Fauquet CM, Mayo MA, Maniloff J, et al. Virus taxonomy: VIIIth report of the international committee on taxonomy of viruses. San Diego: Elsevier Academic Press; 2005.
21. Niehol S, Beaty B, Elliott R. Bunyaviridae//Fauquet CM, Mayo MA, Maniloff J, et al. (eds). Virus Taxonomy. VIIIth Report of the International Committee on Taxonomy of Viruses. Amsterdam: Elsevier Academic Press; 2005.