Marine Genomics 29 (2016) 35–38

Contents lists available at ScienceDirect

Marine Genomics

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

Genomics/technical resources

Complete genome sequence of bacteriophage P2559Y, a marine phage that infects *Croceibacter atlanticus* HTCC2559

Ilnam Kang, Dongmin Kang [1], Jang-Cheon Cho *

Department of Biological Sciences, Inha University, Incheon 22212, Republic of Korea

ARTICLE INFO

Article history:
Received 26 May 2016
Received in revised form 3 July 2016
Accepted 4 July 2016
Available online 7 July 2016

Keywords:
Marine phage
Siphoviridae
Genome sequencing
Flavobacteriaceae

ABSTRACT

The genus *Croceibacter* of the family *Flavobacteriaceae* represents a ubiquitous component of marine bacterioplankton, and is known to be involved in the interaction with phytoplankton. Here, we report the isolation and genome sequencing of a lytic siphovirus P2559Y that infects *Croceibacter atlanticus* HTCC2559, the type strain of the genus *Croceibacter*. The complete genome of P2559Y was 43,153 bp in length, with a GC content of 38.9%. Functional annotation of 51 genes predicted in the genome showed that the P2559Y genome had a modular architecture. Comparison to the genome of P2559S, another phage that infects the same host strain, revealed an interesting feature in the genetic diversity of phages infecting the genus *Croceibacter*. The two phage genomes had a synteny in the structure module and shared many structural genes, while little similarity was found in the DNA replication/metabolism module.

1. Introduction

The bacterial species *Croceibacter atlanticus* was established based on the polyphasic description of strain HTCC2559 isolated from the Sargasso Sea using the dilution-to-extinction high-throughput culturing method (Cho and Giovannoni, 2003). Since its formal description, *Croceibacter atlanticus* has remained the only species of the genus *Croceibacter* within the family *Flavobacteriaceae*, which suggests the phylogenetic uniqueness of the species. A flavobacterial group represented by HTCC2559 seems to be widespread in the marine water column. A BLAST search of the GenBank nr/nt database using the 16S rRNA gene sequence of strain HTCC2559 as a query showed that strains or sequences highly similar to strain HTCC2559 have been reported from diverse marine environments, including the hydrothermal plume of the Southwest Indian Ridge (GenBank accession; KJ549260), Arctic cyanobacterial mat (JF312941) (Prasad et al., 2012), Antarctic sea ice (JQ753219), the Mediterranean Sea (KP887607), the Arctic Ocean (KJ365325), and the northern Bering Sea (GQ984357). In the analysis based on fragment recruitment of GOS sequencing reads to representative marine reference prokaryotic genomes, the HTCC2599 genome recruited a minor fraction of metagenomic reads, but with similar proportions across all GOS stations (Yooseph et al., 2010). Recently, the genus *Croceibacter* was revealed to be important in terms of interaction with marine phytoplankton. The genus *Croceibacter* was one of the phylogenetic groups that increased in abundance during the iron fertilization-induced phytoplankton bloom in an experiment conducted in the Southern Ocean, where many sequences nearly identical (>99.7%) to strain HTCC2559 were reported by cloning (Singh et al., 2015). Moreover, strains that have 16S rRNA gene sequences identical to that of HTCC2559 were isolated from cultures of marine diatom *Pseudo-nitzschia multiseries* and one of the strains was shown to have algicidal activity (Amin et al., 2015). Taken together, these findings indicate that HTCC2559 represents a phylogenetically unique and ecologically relevant flavobacterial group in marine environments.

Bacteriophages are the most abundant biological entities in the marine water column and exert substantial effects on bacterial communities (Breitbart, 2012; Wommack et al., 2015). Together with metagenomic approaches, characterization of bacteriophages that infect diverse marine bacterial groups has been shown to be essential to understanding the dynamics and diversity of both bacterial and viral communities (Brum et al., 2015; Kang et al., 2013; Zhao et al., 2013). However, marine bacteriophages that infect the family *Flavobacteriaceae* have not been studied actively when considering the abundance and ecological importance of the family (Buchan et al., 2014), and the number of flavobacterial genera for which lytic phages have been reported has increased only at a slow pace (Holmfeldt et al., 2013; Kang et al., 2015).

In a previous study, we reported the characteristics and genome sequence of P2559S, a marine bacteriophage isolated from the South Sea of Korea that infects HTCC2559 (Kang et al., 2012). Motivated by the wide distribution of the genus *Croceibacter* and a study of the number and diversity of mycobacteriophages infecting a single mycobacterium

* Corresponding author at: Department of Biological Sciences, Inha University, Inharo 100, Nam-gu, Incheon 22212, Republic of Korea.
E-mail address: chojc@inha.ac.kr (J.-C. Cho).
[1] Present address: Intron Biotechnology, 1013 Jungang Induspia, Seongnam, Gyeonggi-do, Republic of Korea.

http://dx.doi.org/10.1016/j.margen.2016.07.001
1874-7787/

(Pope et al., 2015), we searched for more bacteriophages that infect strain HTCC2559 in coastal seas of Korea. Here, we report the isolation and genomic characterization of P2559Y that was isolated from the Yellow Sea of Korea and showed lytic activity toward strain HTCC2559. A brief comparison between P2559S and P2559Y, two phages infecting the same flavobacterial strain HTCC2559, is also presented.

2. Data description

Bacteriophage P2559Y was isolated from a surface seawater sample collected at June 2011 at Incheon Harbor located in the Yellow Sea using enrichment culture followed by plaque assay (Table 1). The seawater sample was filtered using 0.2-μm pore size membrane filter before being used for enrichment culture. After initial isolation, the purity of P2559Y was established by three rounds of plaque selection. P2559Y formed clear plaques on the bacterial lawn of the host strain, *Croceibacter atlanticus* HTCC2559, suggesting the lytic ability of the phage. However, it was not demonstrated clearly if P2559Y is exclusively lytic or can establish a lysogenic infection. For morphological and genomic characterization, viral particles of P2559Y were concentrated and purified using CsCl equilibrium gradient ultracentrifugation starting with a large volume of lysate sample. Throughout the experiments, strain HTCC2559 was routinely grown at 20 °C in R2A media (BD Difco) prepared using aged seawater diluted with distilled water in a 4:1 ratio (seawater:DW *v*/*v*).

For morphological characterization of P2559Y, phage particles were adsorbed onto formvar- and carbon-coated 200-mesh copper grids, stained with uranyl acetate solution (2%, *w*/*v*), and examined by transmission electron microscope. Morphological examination showed that phage P2559Y belonged to the family *Siphoviridae*, a group of dsDNA bacteriophages characterized by icosahedral heads and long non-contractile tails. The head diameter and tail length of P2559Y were approximately 45 nm and 130 nm, respectively (Fig. 1). Phage P2559S, a previously reported *Croceibacter* phage, was also a siphovirus with a slightly larger head and longer tail than P2559Y. P2559Y did not infect bacterial strain IMCC1106, which was isolated from the East Sea of Korea and showed 99.8% 16S rRNA gene sequence similarity to HTCC2559. This finding suggested that P2559Y has strain specificity in the infection of the *Croceibacter atlanticus* species. The general features of P2559Y are summarized in Table 1.

Genomic DNA of phage P2559Y was extracted from concentrated phage particles using DNeasy Blood & Tissue kits (Qiagen) and subsequently used for genome sequencing by the Illumina HiSeq platform. One paired-end library was generated from genomic DNA and sequenced with a 2 × 101 bp read length, resulting in a total of approximately 7.5 million paired reads. Assembly by SOAPdenovo using subsampled raw sequencing reads (Desai et al., 2013) produced a single contig, which was subsequently circularized by PCR targeting the end regions of the contig (Table 1). Circular closing of the genome suggested that the P2559Y genome is circularly permuted or terminally redundant. The P2559Y genome was 43,153 bp in length, with a 38.9% GC content.

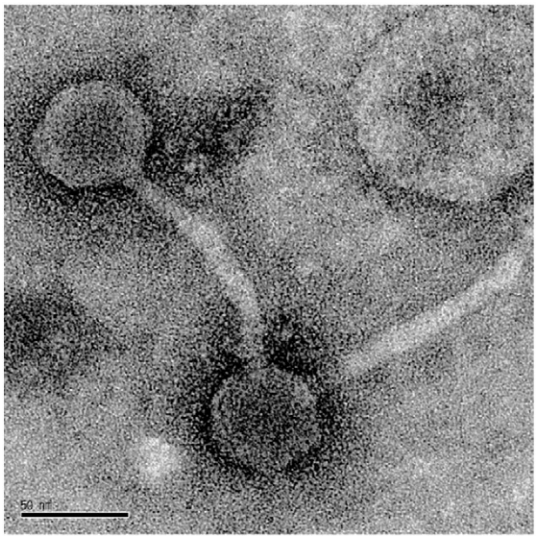

Fig. 1. Transmission electron micrograph of *Croceibacter* phage P2559Y. Scale bar, 50 nm.

Gene calling and annotation using the RAST server predicted 51 protein-encoding genes (Aziz et al., 2008). No tRNA gene was found in the genome. Assignment of functions to protein encoding genes was helped by BLASTp searches of the nr database and RPS-BLAST searches of the Conserved Domain Database of NCBI. Among the 51 protein-coding genes, approximately 19 were functionally annotated, while the other 32 genes were predicted to encode hypothetical proteins (Supplementary Table S1). Functional annotation of the proteins predicted in the P2559Y genome showed that the genome had a modular structure typical of dsDNA bacteriophages. When the genome was permuted to begin at the terminase small subunit, structural proteins clustered immediately downstream in a structural gene module and were followed by a module of genes encoding proteins related to DNA replication and metabolism (Fig. 2a).

Proteins encoded by DNA replication/metabolism genes included DNA polymerase, putative ATPase, helicase, dUTP diphosphatase, DNA methylases, and exonuclease (Fig. 2a). Functional assignment of these proteins was mostly based on the existence of conserved functional domains in each protein, in addition to sequence similarities to other proteins in public databases (Supplementary Table S1). Functional annotation of structural proteins depended largely on the sequence similarity to proteins of other phages that were detected in respective viral particles. ORF4, ORF6, and ORF10 of P2559Y showed approximately 46%, 42%, and 44% sequence identities with major protein 2 (MP2), major protein 3 (MP3), and major protein 1 (MP1) of *Bacteroides* phage B40-8 that were characterized as a tail protein (MP2) and capsid proteins (MP3 and MP1) by immunogold electron microscopy (Puig and Gironés, 1999). ORF8 of P2559Y showed approximately 29% homology with ORF39 of *Bacteroides fragilis* phage B124-14, which was found to be present in viral particles by proteome analysis based on mass spectrometry (Ogilvie et al., 2012). ORF14 showed approximately 26% sequence identity with ORF38 of *Flavobacterium* phage 1/32, which was annotated as a tail assembly chaperone (Senčilo et al., 2015). As in the phage 1/32 genome, a tail length tape measure protein (ORF15) was located downstream of the tail assembly chaperone. ORF17 showed similarity to putative tail fiber proteins of *Bacteroides* phages ATCC 51477-B1 and B124-14 (Hawkins et al., 2008; Ogilvie et al., 2012), and was

Table 1
Classification, general features, and genome sequencing information for phage P2559Y according to the MIxS recommendations.

Item	Description
Classification	Domain: unassigned (ds DNA viruses) Order *Caudovirales* Family *Siphoviridae*
Particle shape	Isometric capsid with a long non-contractile tail
Submitted to INSDC	KC688701 (GenBank)
Investigation type	Virus
Geographic location	Yellow Sea, Incheon, South Korea
Latitude and longitude	37.4973 N 126.6405 E
Depth	0.3 m
Collection date	June 2011
Environment (biome)	Temperate shelf and sea biome (EnvO: 00000895)
Environment (feature)	Coastal water body (EnvO: 02000049)
Environment (material)	Sea water (EnvO: 00002149)
Sequencing method	Illumina HiSeq platform
Number of contigs	1
Assembly method	SOAPdenovo
Finishing quality	Finished (complete)
Assembly coverage	~32,000×