Marine Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

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

Genomic insights into *Photobacterium damselae* subsp. *damselae* strain KC-Na-1, isolated from the finless porpoise (*Neophocaena asiaeorientalis*)

Kyunglee Lee[a,1], Hye Kwon Kim[b,1], Hawsun Sohn[a], Yuna Cho[a], Young-Min Choi[a], Dae Gwin Jeong[b,*], Ji Hyung Kim[b,*]

[a] Cetacean Research Institute (CRI), National Institute of Fisheries Science (NIFS), Ulsan 44780, Republic of Korea
[b] Infectious Disease Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Republic of Korea

ARTICLE INFO

Keywords:
Photobacterium damselae ssp. *damselae*
KC-Na-1
Cetaceans
Pathogens

ABSTRACT

Photobacterium damselae subsp. *damselae* (PDD) is a marine bacterium that can infect a variety of marine animals and humans. Although this bacterium has been isolated from several stranded dolphins and whales, its pathogenic role in cetaceans is still unclear. In this study, we report the complete genome of PDD strain KC-Na-1 isolated from a finless porpoise (*Neophocaena asiaeorientalis*) rescued from the South Sea (Republic of Korea). The sequenced genome comprised two chromosomes and four plasmids. Among the recently identified major virulence factors in PDD, only phospholipase (*plpV*) was found in strain KC-Na-1. Interestingly, two genes homologous to *Vibrio* thermostable direct hemolysin (*tdh*) and its transcriptional regulator *toxR*, which are known virulence factors associated with *Vibrio parahaemolyticus*, were encoded on the plasmid pPDD-Na-1-3. Based on these results, strain KC-Na-1 may have potential pathogenicity in humans and other marine animals and also could act as a potential virulent strain. To the best of our knowledge, this is the first report of the complete genome sequence of *P. damselae*.

1. Introduction

The genus *Photobacterium* comprises 23 valid species in *Vibrionaceae* (*Proteobacteria*: *Gammaproteobacteria*) and is ubiquitous in coastal, open-ocean, and deep-sea environments (Moi et al., 2017). Members of the genus are typically found in seawater and in association with marine animals as saprophytes and enteric commensals; several luminescent species form highly specific bacterial endosymbionts with fish and squid (Urbanczyk et al., 2011). However, two subspecies of *P. damselae* (ssp. *damselae* and ssp. *piscicida*) have been found to be associated with mortality in fish, and *P. damselae* ssp. *damselae* (hereinafter referred as PDD), which was originally described as *Vibrio damsela* (Love et al., 1981), is now considered a bacterial pathogen that can cause infections in a variety of marine ectotherms, including fish, sea turtles, mollusks, and crustaceans (Moi et al., 2017). Furthermore, PDD has been isolated as the causative agent of human infections, including some fatal cases (Rivas et al., 2013a, 2013b).

Several bacterial species, including *Brucella* spp., *Mycobacterium* spp., and *Erysipelothrix rhusiopathiae*, were recently recognized as causative agents of emerging infectious diseases in cetaceans (Van Bressem et al., 2009). PDD has also been isolated from stranded dolphins and whales (Fujioka et al., 1988; Buck et al., 1991; Casalone et al., 2014; Di Francesco et al., 2016); however, the pathogenic role of PDD in cetaceans still remains unclear due to the limitations of post-mortem analyses of stranded individuals (Casalone et al., 2014). Moreover, no studies of porpoises have described PDD isolation or detection, so far. Although the mechanisms of infection and virulence in PDD have not been thoroughly investigated in marine animals (particularly cetaceans and chelonians), a recent genotyping approach revealed that a clonal PDD group might be related to the mortality events among cetaceans (Alba et al., 2016).

Since 2016, we have investigated the bacterial diversity in cetacean species present in coastal waters in the Republic of Korea in order to identify the potential pathogens that can colonize and establish infection in marine mammals for conservation. In this study, we present the complete genome of PDD strain KC-Na-1, which was isolated from a finless porpoise (*Neophocaena asiaeorientalis*) found bycaught in 2016 along the South Sea (Republic of Korea). We aimed to provide insights

Abbreviations: PDD, *Photobacterium damselae* subsp. *damselae*; COG, Clusters of Orthologous Groups; *plpV*, *phospholipase*; *colP*, *collagenase*; *nusB*, N utilization substance protein B; *tdh*, thermostable direct hemolysin

* Corresponding authors.
E-mail addresses: dgjeong@kribb.re.kr (D.G. Jeong), kzh81@kribb.re.kr (J.H. Kim).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.margen.2017.09.004
Received 1 August 2017; Received in revised form 12 September 2017; Accepted 13 September 2017
1874-7787/

Table 1
General features of PDD strain KC-Na-1 and MIGS mandatory information.

Items	Description
Classification	Domain *Bacteria* Phylum *Proteobacteria* Order *Vibrionales* Family *Vibrionaceae* Genus *Photobacterium* Species *damselae* Subspecies *damselae* Strain: KC-Na-1
General features	
Gram stain	Gram negative
Cell shape	Curved rod
Motility	Motile with polar flagella
Temperature	4–42 °C
Pigmentation	Non-pigmented
MIGS data	
Investigation_type	Bacteria_archaea
Project_name	Genome sequence of *P. damselae* subsp. *damselae* KC-Na-01
Lat_lon	34.5 N, 128.4 E
Geo_loc_name	South Korea: South sea
Collection_date	Jan-2017
Env_biome	Ocean [ENVO:00,000,015]
Env_feature	Environmental material [ENVO:00,010,483]
Env_material	Body fluid [ENVO:02,000,019]
Num_replicons	6
Extrachrom_elements	4
Estimated_size	4,544,586
Ref_biomaterial	None
Source_mat_id	KCTC 52975
biotic_relationship	Commensal (or Infectious)
Host	Finless porpoise (*Neophocaena asiaeorientalis*)
Rel_to_oxygen	Facultative anaerobic
Isol_growth_condt	PMID: 21875966
Seq_meth	Illumina Hiseq 2000, PacBio RSII sequencing
Annot_source	GenBank
Finishing_strategy	Complete; 218 × coverage, 6 contigs
Genome assembly data	
Assembly method	HGAP
Assembly name	HGAP algorithm ver. 3
Genome coverage	218 ×
Sequencing technology	Illumina; PacBio

into the biodiversity of the genus *Photobacterium* and obtain useful information for the study of potential virulence factors and antibiotic resistance in PDD.

2. Data description

The general features and MIxS mandatory information for PDD strain KC-Na-1 are summarized in Table 1. The bacterial strain was originally isolated from an anal swab of the rescued juvenile male finless porpoise (*N. asiaeorientalis*, voucher no. CRI007079) found bycaught from net fisheries in December 2016 along the South Sea (Republic of Korea). The non-luminescent, gram-negative, rod-shaped isolate was oxidase and catalase positive and showed weak β-hemolysis on 5% sheep blood agar (Hanil Komed, Republic of Korea) after 24 h of incubation at 37 °C. The 16S rRNA of the isolate (MF099892) showed a match of > 99% with other *P. damselae* strains in the GenBank database, and due to its growth ability at 37 °C (a temperature inhibitory for ssp. *piscicida*) and hemolysis on sheep blood agar (Rivas et al., 2013a,b), the isolate was classified in the subspecies *damselae* and finally designated as PDD strain KC-Na-1.

Genomic DNA was isolated using a DNeasy blood and tissue kit (Qiagen Korea Ltd., South Korea) following the manufacturer's protocols. Sequencing of the strain KC-Na-1 was performed at Macrogen Inc. (South Korea) using the hybrid approach (Koren et al., 2012) with a PacBio RS II system (Pacific Biosciences, USA) by constructing a 20-kb SMRTbell template library and with paired-end Illumina short read data using a HiSeq 2000 instrument (Illumina, USA). The PacBio long read data (1,464,887,777 bp, 167,668 reads) were de novo assembled by the Hierarchical Genome Assembly Process program (ver. 3.0), and the Illumina pair end reads (1,532,556,320 bp, 15,183,058 reads) were mapped to the assembled contigs to improve the accuracy of the genome sequences. Genome annotation was carried out using the NCBI's Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/books/NBK174280/). Bacterial tRNAs and rRNAs were analyzed using tRNAscan-SE 1.21 (Lowe and Eddy, 1997) and RNAmmer 1.2 (Lagesen et al., 2007), respectively. Functional categories of open reading frames were analyzed by BLASTP search against Clusters of Orthologous Groups (COG) database (Tatusov et al., 2001) with an E-value cutoff of 1E-4 and an identity cutoff of 20%.

The fully assembled and closed PDD strain KC-Na-1 genome contained 4,544,586 bp consisting of two chromosomes, designated Chr I (3,134,662 bp) and Chr II (1,105,401 bp), and a total of four plasmids, designated pPDD-Na-1-1 (105,771 bp), pPDD-Na-1-2 (81,002 bp), pPDD-Na-1-3 (72,321 bp), and pPDD-Na-1-4 (45,429 bp), as shown in Fig. 1a. The two chromosomes showed similar G + C contents (41.6% and 39.3%) and percentages of coding regions (86.2% and 84.9%). Moreover, most of the predicted tRNAs (n = 182), rRNA (n = 47), and ncRNA (n = 4) genes were encoded on Chr I, except 11 and one tRNA genes on Chr II and pPDD-Na-1-4, respectively (Table 2).

The COG functional category analysis of PDD strain KC-Na-1 revealed that Chr I had higher percentages of genes related to basic cellular functions than Chr II (Supplementary Fig. 1a). Functional genes encoded on Chr I were mainly involved in COG categories of J (translation, ribosomal structure, and biogenesis), L (replication, recombination, and repair), D (cell cycle control, cell division, and chromosome partitioning), T (signal transduction mechanisms), M (cell wall/membrane/envelope biogenesis), N (cell motility), U (intracellular trafficking, secretion, and vesicular transport), O (post-translational modification, protein turnover, chaperones), C (energy production and conversion), E (amino acid transport and metabolism), F (nucleotide transport and metabolism), H (coenzyme transport and metabolism), and I (lipid transport and metabolism). In contrast, Chr II possessed higher percentages of genes involved in K (transcription), V (defense), X (mobilome: prophages, transposons), G (carbohydrate transport and metabolism), P (inorganic ion transport and metabolism), and Q (secondary metabolites biosynthesis, transport, and catabolism). However, both chromosomes contained genes involved in S (function unknown in COG database), and 6.8% and 15.1% of the predicted genes on Chr I and Chr II, respectively, failed to find a match in the database. As expected, most of the functional genes encoded on plasmids did not have matches in the COG database, and the remaining genes were mainly involved in K, L, U, and X (Supplementary Fig. 1b).

Currently, the complete genome of *P. damselae* is not available in the GenBank database, except that for strain KC-Na-1. Therefore, the OrthoANI algorithm (Lee et al., 2016) was used to assess overall genome similarity between *P. damselae* and other related strains in the family *Vibrionaceae*. OrthoANI values were obtained, and a phylogenetic tree was constructed based on OrthoANI analysis of the three species of *Photobacterium* (*P. damselae*, *P. profundum*, and *P. gaetbulicola*) and the other seven related *Vibrio* species using the orthologous average nucleotide identity tool. The resulting phylogenetic trees based on OrthoANI values for strain KC-Na-1 and other related strains indicated that strain KC-Na-1 was most related to *P. profundum* and that *Vibrio* species were close relatives of the bacterium. However, the other species in the genus *Photobacterium*, i.e., *P. gaetbulicola* strain Group47, showed the lowest ANI value, thus suggesting that its taxonomical position in *Vibrionaceae* needs to be re-examined (Fig. 1b).

Members of *Photobacterium* species, including *P. angustum*, *P. aquimaris*, *P. kishitanii*, *P. leiognathi*, *P. mandapamensis*, and *P. phosphoreum*, are known to produce luminescence, and certain strains of *P. damselae* are also reported to be bioluminescent (Urbanczyk et al., 2011).