Marine Genomics 34 (2017) 11–14

Contents lists available at ScienceDirect

Marine Genomics

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

Complete genome sequence of *Altererythrobacter dongtanensis* KCTC 22672^T, isolated from a tidal flat

Hong Cheng, Yue-Hong Wu, Ying-Yi Huo, Peng Zhou, Qian Liu, Chun-Sheng Wang, Xue-Wei Xu *

Key Laboratory of Marine Ecosystem and Biogeochemistry, Second Institute of Oceanography, State Oceanic Administration, Hangzhou, PR China

ARTICLE INFO

Article history:
Received 22 December 2016
Received in revised form 10 February 2017
Accepted 10 February 2017
Available online 24 February 2017

Keywords:
Altererythrobacter dongtanensis
Complete genome
Arsenate reductase
Aromatic hydrocarbon degradation
Antibiotic resistance

ABSTRACT

Altererythrobacter dongtanensis KCTC 22672^T is a Gram-stain-negative, yellow-pigmented, and aerobic bacterium isolated from a tidal flat of Dongtan Wetland, Chongming Island, China. Here we describe its complete genome sequence, annotation and features. The complete genome of *A. dongtanensis* KCTC 22672^T consists of one chromosome (3,009,495 bp) with the G + C content of 65.75%. The genome contains 2844 protein-coding genes, 47 tRNAs and 3 rRNAs. According to the genome annotation, *A. dongtanensis* KCTC 22672^T possesses dozens of genes related to arsenate reduction, toxic ion resistance, aromatic hydrocarbon degradation and drug resistance, including the genes encoding arsenate reductases, metal transporters, a series of short-chain dehydrogenases and multidrug efflux transporters. It suggests that the strain might have the mechanism to adapt to the polluted coastal tidal flat. This study will facilitate further study on the adaptation of *A. dongtanensis* KCTC 22672^T to contaminated environments near human settlements.

1. Introduction

The genus *Altererythrobacter* belongs to the family *Erythrobacteraceae* of the phylum *Proteobacteria* (Kwon et al., 2007). So far, 21 *Altererythrobacter* species have been proposed and published with valid names (http://www.bacterio.net/altererythrobacter.html). The *Altererythrobacter* strains have been isolated from a variety of environments, such as marine water, deep-sea sediment, mountain soil, mudstone core, tidal flat, fresh water, sea urchin and even air (Rosenberg et al., 2014).

Six complete and two draft genomes of the *Altererythrobacter* strains have been deposited into the NCBI genome database and released up till December 2016. The genomic features of these *Altererythrobacter* spp. show that they are capable of producing novel enzymes with multiple functions and proteins related to antioxidative metabolism or heavy metal resistance, indicating that *Altererythrobacter* strains can be potentially applied in industry, pharmaceutical manufacture and environmental bioremediation (Li et al., 2016; Wu et al., 2015; Shi et al., 2016; Cheng et al., 2016; Zhou et al., 2016). As the number of newly characterized *Altererythrobacter* strains increases, more information regarding to their habitat, physiology and ecology is available. In addition, the genome sequencing of these strains also provides enormous data of protein-coding genes with potential functions. These studies make it feasible to analyze the potential relationship between the environmental features and functional genes in genomes of *Altererythrobacter* strains.

Here we report the complete genome sequence of *Altererythrobacter dongtanensis* KCTC 22672^T, which was isolated from a tidal flat of Dongtan Wetland, Chongming Island, a coastal environment located in the suburb of Shanghai, China (Fan et al., 2011). This area was deeply influenced by anthropogenic activities. The genomic features of *A. dongtanensis* KCTC 22672^T may differ from those of the strains isolated far from human settlements, facilitating further study concerning the adaptation of strain KCTC 22672^T to the depositional island at the estuary.

2. Data description

A. dongtanensis KCTC 22672^T was cultivated in R2A (Difco) liquid medium at 30 °C (Fan et al., 2011). The genomic DNA was prepared with the AxyPrep™ Bacterial Genomic DNA Miniprep Kit (Corning Life Sciences, USA). High-throughput sequencing was carried out on an Illumina HiSeq2000 platform with PE125 strategy. One paired-end library was constructed with the insert size of 500 bp. Reads were assembled de novo into contigs and subsequently joined into scaffolds via ABySS 1.5.2 software (Simpson et al., 2009). The assembly k-mer was tested from k = 53 to 64 for seeking the optimal value of k = 61 by *abyss-pe* command. The clean reads were assembled into only 1 contig. Genome completeness was confirmed by polymerase chain reaction (PCR) with TaKaRa GXL polymerase. The forward and reverse primers located at the start and end of the contig sequence, respectively, were designed via Primer Premier 5 (PREMIER Biosoft, USA). The sequence

* Corresponding author at: 36th North BaoChu Road, Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012, PR China.
E-mail address: xuxw@sio.org.cn (X.-W. Xu).

http://dx.doi.org/10.1016/j.margen.2017.02.003
1874-7787/

Table 1
General features of *Altererythrobacter dongtanensis* KCTC 22672^T and MIGS mandatory information.

Items	Term
Classification	Domain *Bacteria* Phylum *Proteobacteria* Class *Alphaproteobacteria* Order *Sphingomonadales* Family *Erythrobacteraceae* Genus *Altererythrobacter* Species *Altererythrobacter dongtanensis*
Type strain	JM27^T = CCTCC AB 209199^T = KCTC 22672^T
Gram stain	Negative
Cell shape	Rod
Motility	Motile
Sporulation	Non-spore-forming
Temperature range; Optimum (°C)	10–37 °C; 30–37 °C
pH range; Optimum	6.0–10.0; 7.0–9.0
NaCl concentration range; Optimum (%)	0–1%; 0%
Investigation	
Submitted to INSDC	Accession number CP016591
Investigation type	MIMARKS specimen
Project name	Genus *Altererythrobacter* genome sequencing and assembly
Environment	
Geographic location (latitude and longitude)	NR
Geographic location (depth)	NR
Geographic location (country)	China
Collection date	NR
Environment (biome)	Estuarine biome (ENVO:01000020)
Environment (feature)	Depositional island (ENVO:1000750)
Environment (material)	Wetland silt (ENVO_01000016)
MIGS/MIMS/MIMARKS extension	
Environmental package	Isolation sample collected from a tidal flat of Dongtan Wetland, Chongming Island, China
Nucleic acid sequence source	
Isolation and growth conditions	PMID: 20851911
Sequencing	
Target gene or locus	Complete genome sequence of *Altererythrobacter dongtanensis* KCTC 22672
Sequencing method	Sequencing-by-synthesis (Illumina HiSeq 2500 platform); ABI 3730
Assembly	
Assemble method	Abyss 1.5.2; DNAStar seqman v. 8.1.3
Assembly name	KCTC22672_abyss_K61
Genome coverage	400×
Finishing strategy	Primer design, PCR and Sanger sequencing

of PCR products was obtained by Sanger sequencing. The ribosomal RNAs, transfer RNAs and genes on the chromosome sequence were predicted and annotated using Prokka package v1.11 (Seemann, 2014) and NCBI prokaryotic genome annotation pipeline (PGAP) (Tatusova et al., 2016). CRISPRFinder server (http://crispr.u-psud.fr/Server/) was used for screening CRISPR elements (clustered regularly interspaced short palindromic repeat elements) (Grissa et al., 2007).

The general features of strain KCTC 22672^T and MIGS mandatory information are shown in Table 1. A total of 1,324,222,500 clean bases were obtained with approx. 440× sequencing depth. After assembling and gap finishing, the complete genome of the strain KCTC 22672^T yields one chromosome with a total length of 3,009,495 bp and the G + C content of 65.75%. The genome contains 2844 CDSs, 47 tRNAs, one operon of 16S-23S-5S rRNA genes. No CRISPR gene was found as predicted (Table 2). The classification of protein families and clusters of orthologous groups (COGs) of predicted genes were analyzed using Pfam (Finn et al., 2016) and COG database (Galperin et al., 2015). The pathways of enzymes were analyzed via KEGG database (Kanehisa et al., 2008). The distribution of protein-coding sequences, RNAs and COG functional categories is given in Fig. 1.

Table 2
Genome statistics of *Altererythrobacter dongtanensis* KCTC 22672^T.

Attribute	Value
Genome size (bp)	3,009,495
DNA coding region (bp)	2,754,240
DNA G + C content (mol%)	65.75
Chromosome	1
Total genes	2915
Protein-coding genes	2844
Pseudo genes	17
rRNAs	3
tRNAs	47
ncRNAs	4
Genes with Pfam domains	1350
Genes with signal peptides	347
Genes with transmembrane helices	625
Genes assigned to COGs	2427
1 or more conserved domains	2165
2 or more conserved domains	221
3 or more conserved domains	38
4 or more conserved domains	3
CRISPR repeats	0

Due to the rapid urbanization and industrialization, coastal areas are suffering severe contaminations from precipitation of heavy metals (Wang et al., 2013), persistent input of organic pollutants (Liu et al., 2012) and the overdosing of antibiotics (Zhang et al., 2015). The ubiquity of toxic or harmful chemicals in the environment would lead to the evolution of enzymes to detoxify the pollutants (Zhang et al., 2015). The genome of strain KCTC 22672^T harbors 40 protein-coding genes related to arsenate reduction, aromatic hydrocarbon degradation and antibiotic resistance (Supplementary Table 1). Among them, two genes encode arsenate reductase (glutaredoxin) (WP_067676552 and WP_067676931), which belong to the family of oxidoreductase. The arsenate reductase participates in the first step of arsenic metabolism pathway, reducing arsenate (As^{5+}) to arsenite (As^{3+}) (Mukhopadhyay and Rosen, 2002). The energy yielded during this process could support the growth of bacteria (Vanden Hoven and Santini, 2004). Some other genes of strain KCTC 22672^T may encode proteins involved in the process of metal ion resistance or tolerance, such as magnesium/cobalt efflux protein (WP_067678887), Co^{2+}/Mg^{2+} efflux protein ApaG (WP_067677451), divalent metal cation transporter FieF (WP_067675576), metal ABC transporter permease (WP_067681663), toxic anion resistance protein (WP_067676868), cation transporter (WP_067682048) and cation-binding protein (WP_067678267).

It has been shown that several *Altererythrobacter* strains have abilities to degrade polycyclic aromatic hydrocarbons (PAHs), chloronitrobenzene, pesticides and petroleum, suggesting their potentials of petrochemical degradation in wastewater (Ding et al., 2015).

Fig. 1. Graphical circular map of the genome of *Altererythrobacter dongtanensis* KCTC 22672^T. The map was visualized using the local CGView application (Stothard and Wishart, 2005) with adjusted parameters (-title '*Altererythrobacter dongtanensis* KCTC 22672^T' -draw_divider_rings T -gene_decoration arrow -linear circular). Labelling from outside to the center: circle 1, RNA genes on the forward strand (tRNAs red and rRNAs blue); circle 2, CDSs on the forward strand (coloured by COG categories); circle 3, CDSs on the reverse strand (coloured by COG categories); circle 4, RNA genes on the reverse strand (tRNAs red and rRNAs blue); circle 5, G + C content (peaks out/inside the circle indicate values higher or lower than the average G + C content 65.75%, respectively); circle 6, GC skew (calculated as (G − C) / (G + C) using a window size of 10,000 and step of 100, green/purple peaks out/inside the circle indicates values higher or lower than average GC skew value (0.0004), respectively); and circle 7, genome size.