Genomics Data 7 (2016) 237–239

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

Draft genome sequence of *Thalassobius mediterraneus* CECT 5383^T, a poly-beta-hydroxybutyrate producer

Lidia Rodrigo-Torres, María J. Pujalte, David R. Arahal *

Departamento de Microbiología y Ecología and Colección Española de Cultivos Tipo (CECT), Universidad de Valencia, Valencia, Spain

ARTICLE INFO

Article history:
Received 23 December 2015
Received in revised form 8 January 2016
Accepted 14 January 2016
Available online 15 January 2016

Keywords:
Rhodobacteraceae
Roseobacter clade
PHB
Aromatic compounds

ABSTRACT

Thalassobius mediterraneus is the type species of the genus *Thalassobius* and a member of the *Roseobacter* clade, an abundant representative of marine bacteria. *T. mediterraneus* XSM19^T (= CECT 5383^T) was isolated from the Western Mediterranean coast near Valencia (Spain) in 1989. We present here the draft genome sequence and annotation of this strain (ENA/DDBJ/NCBI accession number CYSF00000000), which is comprised of 3,431,658 bp distributed in 19 contigs and encodes 10 rRNA genes, 51 tRNA genes and 3276 protein coding genes. Relevant findings are commented, including the complete set of genes required for poly-beta-hydroxybutyrate (PHB) synthesis and genes related to degradation of aromatic compounds.

1. Direct link to deposited data

Specifications	
Organism/cell line/tissue	*Thalassobius mediterraneus*
Strain	CECT 5383^T
Sequencer or array type	Illumina MiSeq
Data format	Processed
Experimental factors	Bacteria cells cultured in Marine Agar, DNA genomic extraction and sequencing
Experimental features	Draft genome sequence of *Thalassobius mediterraneus* CECT 5383^T, assembly and annotation
Consent	Reads, contig sequences and annotated features are publicly available
Sample source location	2–3 km off Mediterranean coast near Valencia, Spain. 39° 26′ 24″N 0°18′26″W

http://www.ebi.ac.uk/ena/data/view/CYSF00000000

Thalassobius is a genus of the family *Rhodobacteraceae*, order *Rhodobacterales* within the class *Alphaproteobacteria* [1]. It also belongs to the so-called *Roseobacter* group, which contains several of the more ubiquitous and predominant components of marine bacteria and plays an important role in biochemical global processes [2]. Currently, the genus *Thalassobius* contains four species with validly published names: *T. mediterraneus* [1], *T. gelatinovorus* [1], *T. aestuarii* [3] and *T. maritimus* [4]. '*T. aquaeponti*' was reported recently [5], although it has not been yet validated, and even more recently *T. abysii* has been also proposed [6].

2. Experimental design, materials and methods

T. mediterraneus CECT 5383^T is a Gram-negative, coccoid-rod shaped, strictly aerobic, mesophilic, non-pigmented, non-motile and chemoorganotrophic bacterium isolated 2–3 km off Mediterranean coast near Valencia, Spain. It usually forms 2–3 cell chains but not rosettes and utilizes organic acids and amino acids as carbon source but few carbohydrates. It does not reduce nitrate to nitrite [1].

T. mediterraneus CECT 5383^T was cultured in marine agar (MA; Difco) at 26 °C under aerobic conditions during three days. Genomic DNA was isolated using Real Pure Spin kit (Durviz) following the standard protocol recommended by the manufacturer. The integrity of the extracted DNA was checked by visualization in a 2.0% agarose gel electrophoresis. Its purity and quantity was checked by measuring the absorbance at 260 and 280 nm with a spectrophotometer Nanodrop2000c (Thermo Scientific) and calculating the ratio A260/A280.

Genomic DNA was sequenced at Central Service of Support to Experimental Research (SCSIE) of the University of Valencia (Valencia, Spain) using an Illumina MiSeq platform with 2 × 250 paired-end reads. A total of 1,707,192 reads were obtained with 426,425,241 bp, which resulted in a sequencing coverage of 118×.

Reads were assembled with SeqMan NGen 12.0.1 included as a free application in BaseSpace Genomics Cloud Computing (https://basespace.illumina.com). The final assembly is comprised of 19 contigs with genome coverage of 93×.

* Corresponding author.
E-mail addresses: lidia.rodrigo@uv.es (L. Rodrigo-Torres), maria.j.pujalte@uv.es (M.J. Pujalte), arahal@uv.es (D.R. Arahal).

http://dx.doi.org/10.1016/j.gdata.2016.01.013

Table 1
Genome assembly and annotation (by Prokka) data.

Parameter	Value
Finishing quality	Draft
Sequencing platform	Illumina MiSeq
Type of reads	Paired-end, 2 × 250 bp
Number of reads	1,707,192
Fold coverage	118×
Assembler	SeqMan Ngen 12.0.1 (BaseSpace Free App)
Final assembly coverage	93×
ENA/DDBJ/NCBI accession number	CYSF00000000
Genome size	3,431,658 bp
G + C content	58.7%
Sequences in final assembly	19
Total genes	3337
Protein coding genes	3276
rRNA genes	10
tRNA genes	51

This draft genome was annotated with Prokka [7], within Galaxy Orione Server, and RAST v.2.0 [8] using default parameters. A total of 3276 protein coding genes, 51 tRNA genes and 10 rRNA genes were predicted by Prokka and 3270 protein coding genes, 43 tRNA genes and 7 rRNA genes by RAST. Genome features are summarized in Table 1 and a graphical representation of the genome sequence is given in Fig. 1.

T. mediterraneus CECT 5383^T draft genome codes for all the enzymes of Tricarboxilic Acids Cycle. Glycolysis pathway is incomplete as gene coding for 6-phosphofructokinase is not present however Entner–Doudoroff pathway is complete. Pentose–Phosphate pathway is lacking of 6-phosphogluconate dehydrogenase enzyme gene in the oxidative route, suggesting also an incomplete pathway. This could explain why this strain utilizes few carbohydrates as carbon and energy sources [1]. Glyoxylate pathway is also incomplete, without gene encoding isocitrate lyase, however Ethyl–Malonyl–CoA pathway can supply glyoxilate for serine cycle as all genes involved in this route are present.

Although gene encoding for nitrate transporter precursor (NrtA) was encoded, Module Reconstruction by KEGG Mapper (http://www.genome.jp/kegg/tool/map_module.html) showed that assimilatory and dissimilatory nitrate reduction were incomplete with 1 block missing in both cases (nitrite reductase in assimilatory pathway and respiratory nitrate reductase in dissimilatory pathway). This finding is in agreement with the negative result of nitrate to nitrite conversion [1].

The genus *Thalassobius* was reported to be a poly-beta-hydroxybutyrate producer and granules of this bioplastic were found in strain XSM19^T when it was described [1]. To support these finding *phaABCZPR* genes, coding for beta-ketothiolase, acetoacetylCoA reductase, PHB synthase, depolymerase C, phasin protein granule-associated and polyhydroxyalcanoate synthesis repressor respectively, were explored. Six *phaA* genes were found and one copy of the rest of *pha* genes with the exception of *phaP*, suggesting the granule associated protein is accessory.

The genus *Thalassobius* was associated with bacterial communities in dissolved organic matter enriched water [10] and oil and/or dispersant mixed water [11] and an isolate was identified to be able to degrade phthalates [12]. In view of these data, genes involved in aromatic or aliphatic compound degradation were investigated. Although the KEGG Reconstruction module did not recognize complete aromatic degradation pathways, many genes have been found encoding genes related: two toluene efflux pumps and toluene-4-sulfonate monooxygenase system, iron–sulfur subunit TsaM1, *bbsGF* genes involved in anaerobic toluene metabolic pathway, 3-chloro-4-hydroxyphenylacetate reductive dehalogenase precursor and two haloalkane dehalogenases coding genes involved in chloroalkane and chlorobenzene degradation, a 3-hydroxybenzoate

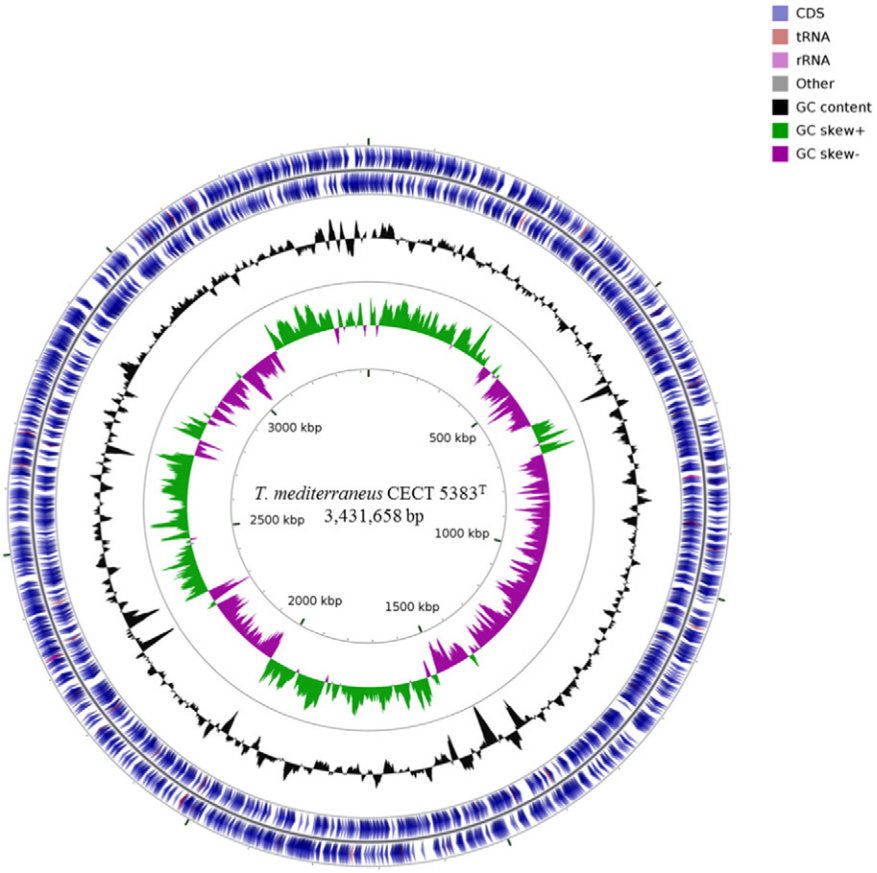

Fig. 1. Genome map of *T. mediterraneus* CECT 5383^T (concatenated gbk file was used in CG view server [9]).

6-hydroxylase 1 and two naphthalene 1,2-dioxygenase subunits genes involved in naphthalene degradation, an anaerobic benzoate catabolism transcriptional regulator and anthranilate 1,2-dioxygenase large subunit genes, two alkane monooxygenase genes, p-hydroxyphenylacetate 3-hydroxylase, reductase component, phenylacetate-CoA ligase and two *paaG* genes involved in phenylacetate degradation, a phenylacetaldehyde dehydrogenase *feaB* gene involved in styrene degradation and two 3-oxoadipate enol-lactonase 2 encoding genes involved in catechol cleavage to oxoadipate. These findings show the large number of genes related to a potential of degrading such polluting compounds however it is necessary to continue working on closing the genome sequence and on experimentation to see if it has this ability or has lost it, as predicted by KEGG.

3. Nucleotide sequence accession number

The Whole Genome Shotgun project is deposited at DDBJ/EMBL/GenBank under accession number CYSF00000000.

Conflict of interest

The authors declare that there is no conflict of interests with respect to the work published in this paper.

Acknowledgments

This work was financially supported by projects TAXPROMAR 2010 (CGL2010-18134/BOS) of the Spanish Ministerio de Economía y Competitividad and GV PROMETEO/2012/040 of the Generalitat Valenciana. M.A. Ruvira is acknowledged for cultivation and DNA extraction.

References

[1] D.R. Arahal, M.C. Macián, E. Garay, M.J. Pujalte, *Thalassobius mediterraneus* gen. nov., sp. nov., and reclassification of *Ruegeria gelatinovorans* as *Thalassobius gelatinovorus* comb. nov. Int. J. Syst. Evol. Microbiol. 55 (2005) 2371–2376, http://dx.doi.org/10.1099/ijs.0.63842–0.

[2] M.J. Pujalte, T. Lucena, M.A. Ruvira, D.R. Arahal, M.C. Macián, The Family Rhodobacteraceae. in: E. Rosenberg, E.F. DeLong, E. Stackebrandt, S. Lory, F. Thompson (Eds.), The Prokaryotes—Alphaproteobacteria and Betaproteobacteria, fourth ed., 8, SpringerVerlag, Berlin 2014, pp. 439–512, http://dx.doi.org/10.1007/978-3-642-30197-1_377.

[3] H. Yi, J. Chun, *Thalassobius aestuarii* sp. nov., isolated from tidal flat sediment. J. Microbiol. 44 (2006) 171–176, http://dx.doi.org/10.1099/ijs.0.027458–0.

[4] S. Park, M.-H. Lee, J.-S. Lee, T.-K. Oh, J.-H. Yoon, *Thalassobius maritimus* sp. nov., isolated from seawater. Int. J. Syst. Evol. Microbiol. 62 (2012) 8–12, http://dx.doi.org/10.1099/ijs.0.029199–0.

[5] S. Park, Y.T. Jung, S.M. Won, J.M. Park, J.H. Yoon, *Thalassobius aquaeponti* sp. nov., an alphaproteobacterium isolated from seawater. Antonie Van Leeuwenhoek 106 (2014) 535–542, http://dx.doi.org/10.1007/s10482–014-0225-1.

[6] Y. Nogi, K. Mori, H. Makita, Y. Hatada, *Thalassobius abyssi* sp. nov., a marine bacterium isolated from the cold-seep sediment. Int. J. Syst. Evol. Microbiol. (2015) http://dx.doi.org/10.1099/ijsem.0.000758 (in press).

[7] T. Seemann, Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (2014) 2068–2069, http://dx.doi.org/10.1093/bioinformatics/btu153.

[8] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, et al., The RAST server: rapid annotations using subsystems technology. BMC Genomics 9 (2008) 75, http://dx.doi.org/10.1186/1471-2164-9-75.

[9] P. Stothard, D.S. Wishart, Circular genome visualization and exploration using CGView. Bioinformatics 21 (2005) 537–539, http://dx.doi.org/10.1093/bioinformatics/bti054.

[10] J. McCarren, J.W. Becker, D.J. Repeta, Y. Shi, C.R. Young, R.R. Malmstrom, et al., Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 16420–16427, http://dx.doi.org/10.1073/pnas.1010732107.

[11] H.P. Bacosa, Z. Liu, D.L. Erdner, Natural sunlight shapes crude oil-degrading bacterial communities in Northern Gulf of Mexico surface waters. Front. Microbiol. 6 (2015) 1325, http://dx.doi.org/10.3389/fmicb.2015.01325.

[12] H. Iwaki, A. Nishimura, Y. Hasegawa, Isolation and characterization of marine bacteria capable of utilizing phthalate. World J. Microbiol. Biotechnol. 28 (2012) 1321–1325, http://dx.doi.org/10.1007/s11274-011-0925-x.