

SHORT GENOME REPORT

Open Access



Genome sequence of *Acuticoccus yangtzensis* JL1095^T (DSM 28604^T) isolated from the Yangtze Estuary

Lei Hou^{1,2}, Jia Sun^{1,2}, Xiabing Xie^{1,2}, Nianzhi Jiao^{1,2} and Yao Zhang^{1,2*}

Abstract

Acuticoccus yangtzensis JL1095^T is a proteobacterium from a genus belonging to the family *Rhodobacteraceae*; it was isolated from surface waters of the Yangtze Estuary, China. This strain displays the capability to utilize aromatic and simple carbon compounds. Here, we present the genome sequence, annotations, and features of *A. yangtzensis* JL1095^T. This strain has a genome size of 5,043,263 bp with a G + C content of 68.63%. The genome contains 4286 protein-coding genes, 56 RNA genes, and 83 pseudo genes. Many of the protein-coding genes were predicted to encode proteins involved in carbon metabolism pathways, such as aromatic degradation and methane metabolism. Notably, a total of 31 genes were predicted to encode form II carbon monoxide dehydrogenases, suggesting potential for carbon monoxide oxidation. The genome analysis helps better understand the major carbon metabolic pathways of this strain and its role in carbon cycling in coastal marine ecosystems.

Keywords: *Acuticoccus yangtzensis* JL1095^T, Aromatic compounds degradation, Methane metabolism, Form II CODH, Aerobic CO oxidation, Yangtze estuary

Introduction

We isolated a member in the family *Rhodobacteraceae*, *Acuticoccus yangtzensis* JL1095^T (= CGMCC 1.12795 = DSM 28604), from surface waters of the Yangtze Estuary, China (31° N, 122° E) [1, 2]. The physiological properties of members in the family *Rhodobacteraceae* suggest that they may be important in regulating the carbon cycle in terrestrial and marine ecosystems. For instance, many members of this family can degrade aromatic compounds [3] and metabolize one-carbon compounds [4]. Physiological tests of *A. yangtzensis* JL1095^T have shown that strain JL1095^T was able to degrade naphthol-AS-Bi-phosphate, and utilize acetic acid and glycerol [1]. In addition, many members of the family *Rhodobacteraceae* examined to date have the ability to oxidize CO.

CO is an important atmospheric trace gas that contributes to climate change despite its low concentrations (0.05–0.12 ppm) in air [5]. Although CO is toxic

for many organisms, a number of microbes can consume CO. Marine microbial CO oxidation represents an important CO sink in the oceans. CODHs, key enzymes for CO oxidation, have been classified into two major types based on their cofactor composition, structure, and stability in the presence of dioxygen [6]. Ni- and Fe-containing CODHs are found in anaerobic bacteria and archaea, while Cu- and Mo-containing CODHs are found in aerobic bacteria [7]. Compared with the relatively hypoxic and high CO concentrations in the early Earth environment [8], the ecological significance of aerobic CO oxidation has become increasingly critical in the relatively aerobic and low CO concentrations in modern environments. Aerobic CO oxidation is carried out by phylogenetically and physiologically diverse aerobic bacteria and certain newly identified archaea that are distributed in a variety of habitats, including terrestrial, sedimentary, freshwater, and marine ecosystems [9]. The most active CO oxidizers belong to various genera, such as *Ruegeria*, *Roseobacter*, *Stappia* and *Silicibacter*, mostly from the family *Rhodobacteraceae* [10, 11]. Based on phylogenetic analysis of 16S rRNA sequences and physiological characteristics, *A. yangtzensis* JL1095^T is most closely related to

* Correspondence: yaozhang@xmu.edu.cn

¹State Key Laboratory of Marine Environmental Sciences, Xiamen University, Xiamen 361102, People's Republic of China

²Institute of Marine Microbes and Ecospheres, Xiamen University, Xiamen 361102, People's Republic of China

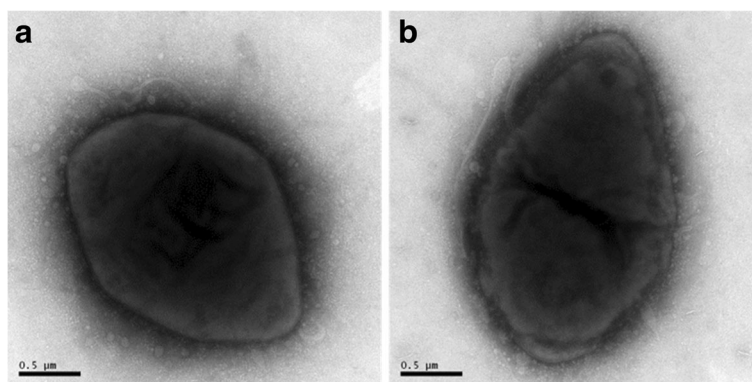


Fig. 1 Transmission electron micrographs of *Acuticoccus yangtzensis* JL1095^T cultured on marine agar 2216 (MA; Difco) medium. **a** Oval-shaped cells with one peak end; **b** a cell divided by binary fission. Scale bar, 0.5 μm

Table 1 Classification and general features of *Acuticoccus yangtzensis* strain JL1095^T [16]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain <i>Bacteria</i>	TAS [30]
		Phylum <i>Proteobacteria</i>	TAS [31]
		Class <i>Alphaproteobacteria</i>	TAS [32]
		Order <i>Rhodobacterales</i>	TAS [33]
		Family <i>Rhodobacteraceae</i>	TAS [33]
		Genus <i>Acuticoccus</i>	TAS [1, 2]
		Species <i>Acuticoccus yangtzensis</i>	TAS [1, 2]
		Type strain: JL1095 ^T (= CGMCC 1.12795 = DSM 28604)	
	Gram stain	Negative	TAS [1]
	Cell shape	Oval-shaped with one peak end	TAS [1]
	Motility	Motile	TAS [1]
	Sporulation	Not reported	NAS
	Temperature range	15–50 °C	TAS [1]
	Optimum temperature	35 °C	TAS [1]
	pH range; Optimum	6.0–9.0; 7.6	TAS [1]
	Carbon source	Tween 40, Tween 80, L-arabinose, methyl-pyruvate, D,L-Lactic acid, acetic acid, urocanic acid, α-hydroxy butyric acid, β-hydroxy butyric acid and γ-hydroxy butyric acid	TAS [1]
MIGS-6	Habitat	Estuary	TAS [1]
MIGS-6.3	Salinity	2–10% NaCl (w/v)	TAS [1]
MIGS-22	Oxygen requirement	Aerobic	TAS [1]
MIGS-15	Biotic relationship	free-living	NAS
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Yangtze Estuary, China	TAS [1]
MIGS-5	Sample collection	January 2006	IDA
MIGS-4.1	Latitude	31° N	TAS [1]
MIGS-4.2	Longitude	122° E	TAS [1]
MIGS-4.4	Altitude	Sea level	TAS [1]

^aEvidence codes - IDA Inferred from Direct Assay, TAS Traceable Author Statement (i.e., a direct report exists in the literature), NAS Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [22]

the genus *Stappia* [1], in which all known and examined to date have the ability to oxidize CO, containing form I and II *cox* gene operons [12–14].

In this study, we describe the classification and features of *A. yangtzensis* JL1095^T, report its first draft genome sequence, and explore its major carbon metabolic pathways and potential capability to oxidize CO.

Organism information

Classification and features

A. yangtzensis JL1095^T (= CGMCC 1.12795 = DSM 28604), as the type strain of *A. yangtzensis* in the family *Rhodobacteraceae*, is a Gram-negative, aerobic, motile (possibly through gliding), oval-shaped with one peak end bacterium (Fig. 1). The detailed classification and features were previously reported [1, 2]. Briefly, the solo-carbon-source utilization test indicated that Tween 40, Tween 80, L-arabinose, methyl-pyruvate, β -hydroxy butyric acid, D,L-lactic acid, acetic acid, urocanic acid, α -hydroxy butyric acid, γ -hydroxy butyric acid, L-proline, glycerol, α -keto butyric acid, D-fructose, L-fucose, D-galactose, α -D-glucose, D-mannose, L-serine, D-sorbitol, D-gluconic acid, α -keto glutaric acid, succinamic acid, L-glutamic acid, pyruvate, and gelatin were utilized by this strain. In addition, strain JL1095^T produces various enzymes for the degradation of organic matter, including urease, protease, alkaline phosphatase enzyme, esterase (C4), leucine arylamidase, valine arylamidase, trypsin and naphthol-AS-BI-phosphate hydrolase [1]. The current classification and general features of *A. yangtzensis* JL1095^T are listed in Table 1.

The draft genome sequence of *A. yangtzensis* JL1095^T has one full-length 16S rRNA gene sequence (1450 bp; BIX52_RS22260) that was consistent with the partial 16S rRNA gene sequence from the original species description (1397 bp; KF741873) [1]. Strain JL1095^T showed the highest 16S rRNA gene sequence similarity with *Stappia indica* B106^T (92.7%) followed by *Stappia stellata* IAM 12621^T (92.6%) and *Labrenzia suaedae* DSM 22153^T (92.3%). The phylogenetic tree was constructed to assess the evolutionary relationships between strain JL1095^T and other related strains with the MEGA 5.05 software by using a neighbor-joining algorithm with the Jukes-Cantor model. The phylogeny of the strain JL1095^T illustrated that one monophyletic branch is formed at the periphery of the evolutionary radiation occupied by the various genera in the family *Rhodobacteraceae* (Fig. 2).

Genome sequencing information

Genome project history

This strain was selected for sequencing on the basis of its important evolutionary position, the degradation of aromatic and simple hydrocarbon compounds via metabolism [1], and its potential CO oxidation ability. The

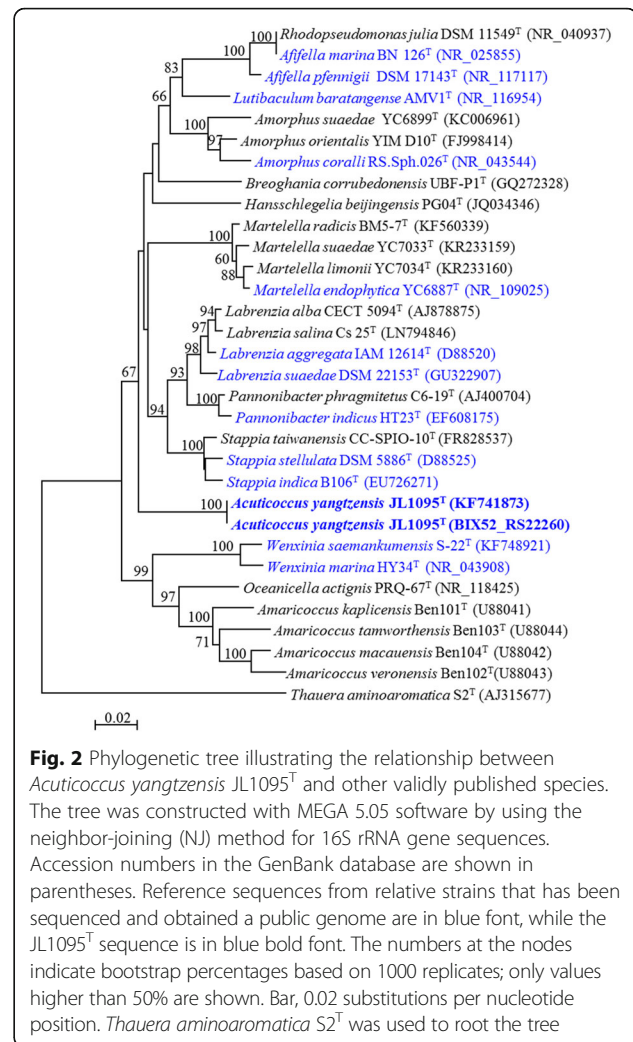


Fig. 2 Phylogenetic tree illustrating the relationship between *Aceticoccus yangtzensis* JL1095^T and other validly published species. The tree was constructed with MEGA 5.05 software by using the neighbor-joining (NJ) method for 16S rRNA gene sequences. Accession numbers in the GenBank database are shown in parentheses. Reference sequences from relative strains that has been sequenced and obtained a public genome are in blue font, while the JL1095^T sequence is in blue bold font. The numbers at the nodes indicate bootstrap percentages based on 1000 replicates; only values higher than 50% are shown. Bar, 0.02 substitutions per nucleotide position. *Thauera aminoaromatica* S2^T was used to root the tree

sequencing of the *A. yangtzensis* JL1095^T genome was carried out at Beijing Novogene Bioinformatics Technology Co., Ltd. The genome sequence of *A. yangtzensis* JL1095^T has been deposited in the GOLD [15] and DDBJ/EMBL/GenBank under accession number MJUX00000000. A summary for the genome sequencing information of *A. yangtzensis* JL1095^T is listed in Table 2, in compliance with MIGS version 2.0 [16].

Growth conditions and genomic DNA preparation

A. yangtzensis JL1095^T (= CGMCC 1.12795 = DSM 28604) was cultivated aerobically in MB (Difco) medium. The genomic DNA of strain JL1095^T was extracted using the Tguide Bacteria Genomic DNA Kit (OSR-M502, TIANGEN Biotech Co. Ltd., Beijing, China) in accordance with the instruction manual. After this strain was cultivated in MB medium in the shaker at 35 °C for 2–3 days, the total DNA obtained was subjected to quality control by agarose gel electrophoresis and quantified by Qubit 2.0 fluorometer (Life Technologies, MA, USA).

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	High-quality draft
MIGS-28	Libraries used	500 bp Paired-end
MIGS 29	Sequencing platforms	Illumina HiSeq 2500
MIGS 31.2	Fold coverage	331X
MIGS 30	Assemblers	SOAPdenovo version 2.04
MIGS 32	Gene calling method	GeneMarkS version 4.17
	Locus Tag	BIX52
	Genbank ID	MJUX00000000
	GenBank Date of Release	December 31th, 2016
	GOLD ID	Gp0206530
	BIOPROJECT	PRJNA343888
MIGS 13	Source Material Identifier	CGMCC 1.12795=DSM 28604
	Project relevance	Environmental, microbes

Genome sequencing and assembly

The genome sequencing of this strain was conducted using Illumina HiSeq 2500 paired-end sequencing technology under the PE 150 strategy. A total filtered read size of 1674 Mbp was obtained. The filtered reads were assembled by SOAPdenovo version 2.04 software and 29 contigs were generated [17, 18]. Gene prediction was performed on the genome assembly using GeneMarkS version 4.17 [19].

Genome annotation

Functional annotation of the coding sequences was performed by searching various databases (KEGG [20], NR, COG [21], and GO [22]). The rRNA genes of strain JL1095^T were predicted using rRNAmmer software [23], tRNA genes were identified using tRNAscan-SE [24], and sRNA were predicted by BLAST searches against the Rfam database [25]. The online CRISPRFinder program was used for CRISPR identification [26].

Genome properties

The *A. yangtzensis* JL1095^T genome was composed of 5,043,263 bp with a G + C content of 68.63%. A total of 4286 protein-coding genes were predicted with an average length of 994 bp, occupying 87.01% of the genome. The genome also contained 56 RNA genes and 83 pseudo genes. Detailed genome statistical information is shown in Table 3. COG categories were assigned to 2522 of the protein-coding genes which were classified into 21 functional groups. The most dominant COG categories were “amino acid transport and metabolism” followed by “general function prediction only”, “function unknown”, and “energy production and conversion”. Detailed gene numbers and percentages related with the COG categories are shown in Table 4. In total, 2470

Table 3 Genome statistics

Attribute	Value	% of Total
Genome size (bp)	5,043,263	100.00
DNA coding (bp)	4,388,143	87.01
DNA G + C (bp)	3,461,191	68.63
DNA scaffolds	28	100.00
Total genes	4425	100.00
Protein coding genes	4286	96.86
RNA genes	56	1.27
Pseudo genes	83	1.88
Genes in internal clusters	NA	NA
Genes with function prediction	3781	85.45
Genes assigned to COGs	2522	56.99
Genes with Pfam domains	3139	70.94
Genes with signal peptides	348	7.86
Genes with transmembrane helices	1043	23.57
CRISPR repeats	3	0.07

NA, no analysis

Table 4 Number of genes associated with general COG functional categories

Code	Value	%age	Description
J	162	3.78	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	139	3.24	Transcription
L	111	2.59	Replication, recombination and repair
B	3	0.07	Chromatin structure and dynamics
D	19	0.44	Cell cycle control, Cell division, chromosome partitioning
V	20	0.47	Defense mechanisms
T	93	2.17	Signal transduction mechanisms
M	126	2.94	Cell wall/membrane biogenesis
N	30	0.70	Cell motility
U	43	1.00	Intracellular trafficking and secretion
O	111	2.59	Posttranslational modification, protein turnover, chaperones
C	223	5.20	Energy production and conversion
G	198	4.62	Carbohydrate transport and metabolism
E	388	9.05	Amino acid transport and metabolism
F	63	1.47	Nucleotide transport and metabolism
H	122	2.85	Coenzyme transport and metabolism
I	138	3.22	Lipid transport and metabolism
P	187	4.36	Inorganic ion transport and metabolism
Q	109	2.54	Secondary metabolites biosynthesis, transport and catabolism
R	378	8.82	General function prediction only
S	232	5.41	Function unknown
–	1764	41.16	Not in COGs

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

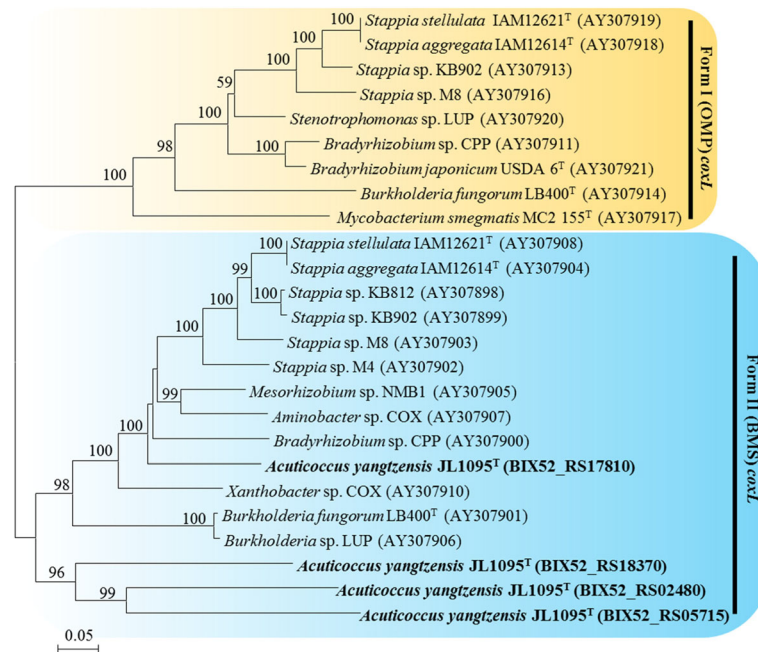


Fig. 3 Unrooted phylogenetic tree showing the *coxL* genotype of *Aceticoccus yangtzensis* JL1095^T. The tree was constructed with MEGA 5.05 software by using the neighbor-joining (NJ) method based on the form I *coxL* and form II putative *coxL* genes from CO-oxidizing microbes. Accession numbers in the GenBank database are shown in parentheses. The *coxL* genes encoded in the *Aceticoccus yangtzensis* JL1095^T genome are shown in bold. Sequences in orange and blue shades represent form I and II *coxL* genes, respectively. The numbers at the nodes indicate bootstrap percentages based on 1000 replicates; only values higher than 50% are shown. Bar, 0.05 substitutions per nucleotide position

protein-coding genes were assigned to 153 KEGG metabolic pathways, including key genes involved in carbon metabolism processes such as gluconeogenesis, polycyclic aromatic hydrocarbon degradation, and methane metabolism. In addition, based on the GO database, 1992 protein-coding genes were assigned to molecular function, 1394 genes were assigned to cellular components, and 2646 genes were assigned to biological processes.

Insights from the genome sequence

We performed a systematic analysis of the protein-coding genes with functional predictions by BLAST searches against the four databases (KEGG, NR, COG, and GO), with E-value $<1e-5$ and minimal alignment length of $>40\%$.

Strain JL1095^T was predicted to contain most of the genes central to carbon metabolism, including those related to glycolysis/gluconeogenesis, the tricarboxylic acid cycle, and the pentose phosphate pathway. Furthermore, about 198 genes were assigned to COG categories related to carbohydrate transport and metabolism, including fructose, mannose, and galactose metabolism. These carbohydrate metabolic characteristics are generally coincident with those obtained from a sole-carbon-source utilization experiment [1]. The capacity of this strain to degrade aromatic compounds such as naphthol-AS-BI-phosphate has been identified. Approximately 236 genes

were involved in 13 KEGG metabolic pathways related to aromatic compounds degradation, such as polycyclic aromatic hydrocarbon, bisphenol, and naphthalene. Aromatic compounds are important environmental organic pollutants because of their persistence in environments, toxicity, and carcinogenic characteristics [27]. Furthermore, strain JL1095^T was annotated to contain 48 genes related to methane metabolism.

Based on results from the four functional annotation databases, the *A. yangtzensis* JL1095^T genome contained a total of 31 genes predicted to encode aerobic-type CODHs (Additional file 1: Table S1). The *cox* gene clusters that encode aerobic CODHs have been classified into two major forms based on genome analysis [9]. Form I genes are mainly from *Oligotropha*, *Mycobacterium* and *Pseudomonas*, and form II putative genes are mainly from *Bradyrhizobium*, *Mesorhizobium*, and *Sinorhizobium* [13]. Form I and II *cox* gene operons consisted of three conserved structural genes that were transcribed as *coxMSL* and *coxSLM*, respectively [28, 29]. For strain JL1095^T, three structural genes containing *coxS* (small subunit), *coxM* (medium subunit) and *coxL* (large subunit) were all sequenced. Form I *coxS* and *coxM* gene sequences were similar to form II *coxS* and *coxM* gene sequences, but the form II putative *coxL* gene sequence was approximately 40–50% similar to the form I *coxL* gene sequence [9]. Therefore, the *coxL* gene has been used as a molecular

biomarker to explore the distribution of aerobic CO bacteria in ecosystems [29]. We constructed the *coxL* phylogenetic tree for strain JL1095^T and confirmed that four predicted *coxL* genes (Locus tag: BIX52_RS02480, BIX52_RS05715, BIX52_RS17810 and BIX52_RS18370) were recognized as form II *coxL* genes (Fig. 3). Additionally, the accessory genes were also essential for CO oxidation to take place. The accessory genes in forms I and II varied substantially, and even within the same form, the order and sub-unit types varied among isolates [9]. Form I *cox* accessory genes, including *coxB*, *C*, *G*, *H*, *I*, and *K*, were distributed flexibly around the structural genes. Among the form II *cox* accessory genes, *coxG* was usually an indispensable gene compared with other accessory genes, such as *coxD*, *E*, and *F* [28]. For this

strain, the accessory gene *coxG* was detected. Form I CODH has been specifically characterized for its ability to oxidize CO, while form II is a putative CODH and its ability to oxidize CO remains uncertain. For the *Roseobacter* clade, both *coxL* forms were present, which enables them to oxidize CO [11]. Phylogenetic analysis using the 16S rRNA gene sequences of *A. yangtzensis* JL1095^T and *Roseobacter* clade bacteria indicates that JL1095^T does not belong to the *Roseobacter* clade (Fig. 4). However, many other bacteria containing only form II *cox* genes have been shown by molecular and culture-based methods to oxidize CO, including *Mesorhizobium* sp. strain NMB1, *Mesorhizobium loti*, *Aminobacter* sp. strain COX, *Xanthobacter* sp. strain COX, and *Burkholderia* sp. strain LUP [13]. According to the phylogenetic tree (Fig. 3), the *coxL* genes of JL1095^T

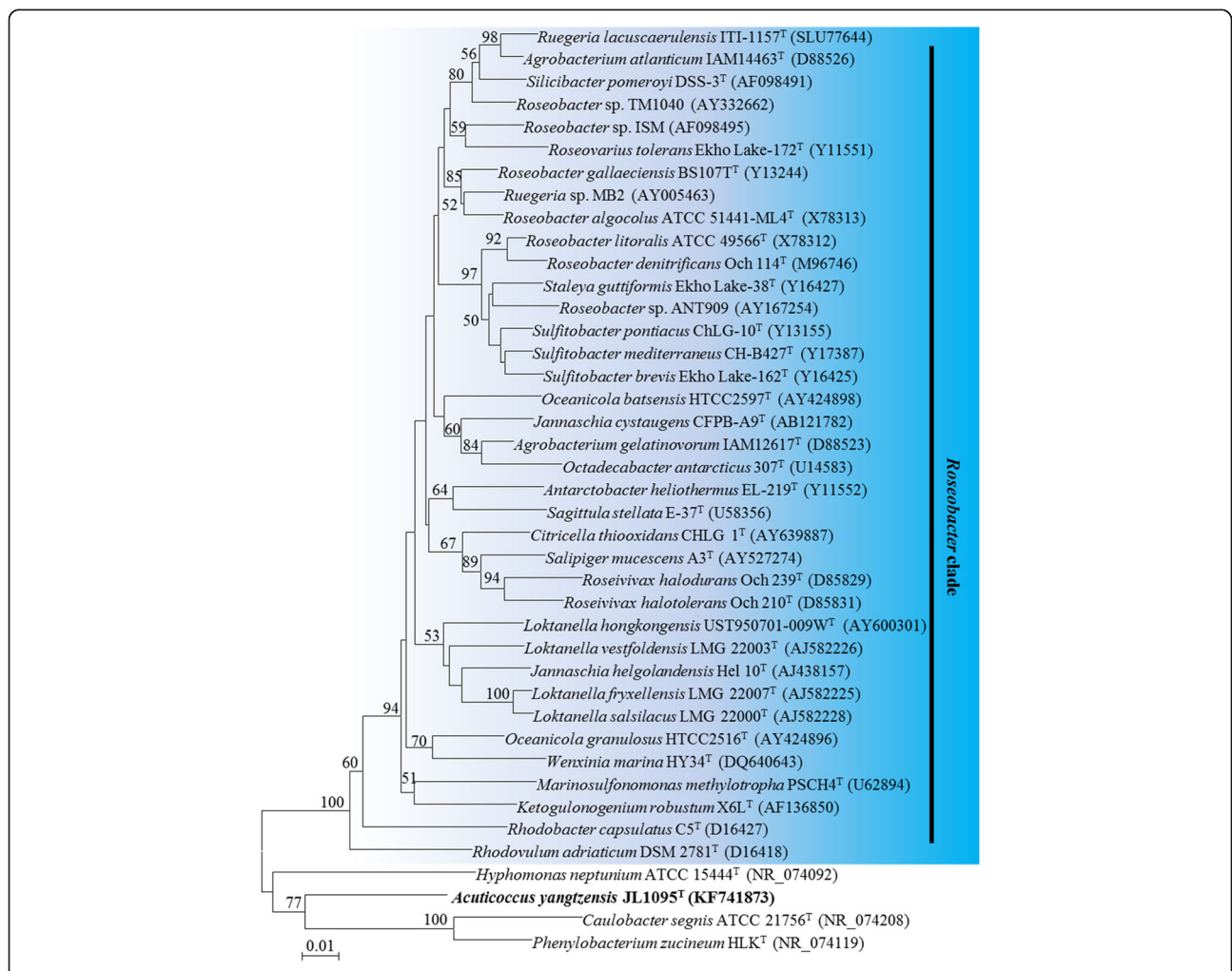


Fig. 4 Unrooted phylogenetic tree displaying the relationship between *Aceticoccus yangtzensis* JL1095^T and *Roseobacter* clade bacteria. The tree was constructed with MEGA 5.05 software by using the neighbor-joining (NJ) method based on 16S rRNA gene sequences. Accession numbers in the GenBank database are shown in parentheses. The 16S rRNA gene encoded in the *Aceticoccus yangtzensis* JL1095^T genome is shown in bold. The numbers at the nodes indicate bootstrap percentages based on 1000 replicates; only values higher than 50% are shown. Bar, 0.01 substitutions per nucleotide position

clustered tightly with these bacterial isolates. Thus, we speculate that JL1095^T is capable of oxidizing CO. Future studies are needed to determine its function in CO oxidation.

Conclusions

In the present study, the genome of *A. yangtzensis* JL1095^T, the type strain of *A. yangtzensis*, was characterized. It contains numerous genes involved in carbohydrate transport and metabolism, aromatic compounds degradation, and methane metabolism. Knowledge of the genome sequence of *A. yangtzensis* JL1095^T lays a foundation for better understanding the carbon metabolism of this strain. Based on genome analysis, we speculate that JL1095^T is capable of oxidizing CO. Future studies are needed to determine its function in CO oxidation. These genomic data provide insight into the carbon metabolic characteristics of *A. yangtzensis* JL1095^T and its role in alleviating coastal water pollution and effects on the marine carbon cycle.

Additional file

Additional file 1: Table S1. Aerobic-type CODH-encoding genes of *Aceticoccus yangtzensis* JL1095^T predicted using four different databases. (DOCX 29 kb)

Abbreviations

CGMCC: China General Microbiological Culture Collection Center; CO: Carbon monoxide; CODHs: CO dehydrogenases; CRISPR: Clustered regularly interspaced short palindromic repeats; DSMZ: Leibniz-Institut DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH; GOLD: Genomes OnLine Database; MA: marine agar 2216; MB: marine broth 2216; MIGS: Minimum information on the genome sequence

Funding

This research was supported by the SOA projects GASI-03-01-02-03, the national key research program 2016YFA0601400, the NSFC projects 41,422,603, 41,676,125, and 91,428,308.

Authors' contributions

This project was founded by YZ and NJ. The main tasks, including experiments, data analysis and manuscript writing, were performed by LH and YZ. JS was associated with this bacteria isolation. XX provided technical support for this research. 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.

Received: 12 July 2017 Accepted: 5 December 2017

Published online: 29 December 2017

References

- Hou L, Zhang Y, Sun J, Xie X. *Aceticoccus yangtzensis* gen. nov., sp. nov., a novel member in the family *Rhodobacteraceae*, isolated from the surface water of the Yangtze estuary. *Curr Microbiol.* 2015;70:176–82.
- Oren A, Garrity GM. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol.* 2017;67:1095–8.
- Buchan A, Neidle EL, Moran MA. Diversity of the ring-cleaving dioxygenase gene *pcpA*H in a salt marsh bacterial community. *Appl Environ Microbiol.* 2001;67:5801–9.
- Doronina NV, Trotsenko YA, Tourova TP. *Methylarcula marina* gen. nov., sp. nov. and *Methylarcula terricola* sp. nov.: novel aerobic, moderately halophilic, facultatively methylotrophic bacteria from coastal saline environments. *Int J Syst Evol Microbiol.* 2000;50:1849–59.
- Air quality guidelines for Europe (Second edition). 2000. <http://hdl.handle.net/20.500.11822/8681>. Accessed 2000.
- Jeoung JH, Dobbek H. Carbon dioxide activation at the Ni, Fe-cluster of anaerobic carbon monoxide dehydrogenase. *Science.* 2007;318:1461–4.
- Ragsdale SW. Life with carbon monoxide. *Crit Rev Biochem Mol Biol.* 2004;39:165–95.
- Miyakawa S, Yamanashi H, Kobayashi K, Cleaves HJ, Miller SL. Prebiotic synthesis from CO atmospheres: implications for the origins of life. *Proc Natl Acad Sci.* 2002;99:14628–31.
- King GM, Weber CF. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol.* 2007;5:107–18.
- Tolli JD, Sievert SM, Taylor CD. Unexpected diversity of bacteria capable of carbon monoxide oxidation in a coastal marine environment, and contribution of the *Roseobacter*-associated clade to total CO oxidation. *Appl Environ Microbiol.* 2006;72:1966–73.
- Zhang Y, Sun Y, Jiao N, Stepanauskas R, Luo H. Ecological genomics of the uncultivated marine *Roseobacter* lineage CHAB-I-5. *Appl Environ Microbiol.* 2016;82:2100–11.
- Kim BC, Park JR, Bae JW, Rhee SK, Kim KH, Oh JW, et al. *Stappia marina* sp. nov., a marine bacterium isolated from the Yellow Sea. *Int J Syst Evol Microbiol.* 2006;56:75–9.
- King GM. Molecular and culture-based analyses of aerobic carbon monoxide oxidizer diversity. *Appl Environ Microbiol.* 2003;69:7257–65.
- Weber CF, King GM. Physiological, ecological, and phylogenetic characterization of *Stappia*, a marine CO-oxidizing bacterial genus. *Appl Environ Microbiol.* 2007;73:1266–76.
- Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 2008;36(Suppl 1):475–9.
- Field D, Garrity G, 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.
- Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. *Bioinformatics.* 2008;24:713–4.
- Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 2010;20:265–72.
- Besemer J, Lomsadze A, Borodovsky M. GeneMark: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic acids Res.* 2001;29:2607–18.
- Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 2004;32:277–80.
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. *BMC bioinformatics.* 2003;4:1–14.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet.* 2000;25:25–9.
- Lagesen K, Hallin PF, Rødland E, Stærfeldt HH, Rognes T, Ussery DW. RNAMmer: consistent annotation of rRNA genes in genomic sequences. *Nucleic Acids Res.* 2007;35:3100–8.
- Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic acids Res.* 1997;25:955–4.
- Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, et al. Rfam: updates to the RNA families database. *Nucleic acids Res.* 2009;37:136–40.
- Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic acids research.* 2007;35(Suppl 2):52–7.
- Liu Y, Chen L, Jianfu Z, Qinghui H, Zhiliang Z, Hongwen G. Distribution and sources of polycyclic aromatic hydrocarbons in surface sediments of rivers and an estuary in Shanghai, China. *Environ Pollut.* 2008;154:298–305.

28. Santiago B, Schübel U, Egelseer C, Meyer O. Sequence analysis, characterization and CO-specific transcription of the *cox* gene cluster on the megaplasmid *pHCG3* of oligotropha carboxidovorans. *Gene*. 1999;236:115–24.
29. Yang J, Zhou E, Jiang H, Li W, Wu G, Huang L, et al. Distribution and diversity of aerobic carbon monoxide-oxidizing bacteria in geothermal springs of China, the Philippines, and the United States. *Geomicrobiol J*. 2015;32:903–13.
30. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci*. 1990;87:4576–9.
31. Stackebrandt E, Murray RGE, Truper HG. *Proteobacteria* classis nov., a name for the phylogenetic taxon that includes the “purple bacteria and their relatives”. *Int J Syst Bacteriol*. 1988;38:321–5.
32. Garrity GM, Bell JA, Lilburn T. Class I. *Alphaproteobacteria* class. nov. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. *Bergeys Manual of Systematic Bacteriology*, second edition, vol. 2 (The *Proteobacteria*), part C (The Alpha-, Beta-, Delta-, and *Epsilonproteobacteria*). New York: Springer; 2005. p. 1.
33. Garrity GM, Bell JA, Lilburn T. Family I. *Rhodobacteraceae* fam. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, second edition, vol. 2, part C. New York: Springer; 2005. p. 161

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

