· Article · Advances in Polar Science
doi: 10.13679/j.advps.2020.0034 March 2021 Vol. 32 No. 1: 20-27

Complete genome analysis of bacteriochlorophyll *a*-containing *Roseicitreum antarcticum* ZS2-28^T reveals its adaptation to Antarctic intertidal environment

ZENG Yinxin*, YU Yong, LI Huirong, LUO Wei & DING Haitao

MNR Key Laboratory for Polar Science, Polar Research Institute of China, Shanghai 200136, China

Received 12 November 2020; accepted 8 March 2021; published online 18 March 2021

Abstract Aerobic anoxygenic phototrophic bacteria (AAPB) are photoheterotrophic prokaryotes able to use both light and dissolved organic matter as energy sources. *Roseicitreum antarcticum* ZS2-28^T was isolated from intertidal sediment in the Larsemann Hills, Princess Elizabeth Land, Antarctica, and was able to produce bacteriochlorophyll *a*. It is the type strain of the sole species within the genus *Roseicitreum*. The complete genome sequence of the bacterium was determined using Illumina HiSeq X and PacBio RSII systems. The genome of *R. antarcticum* ZS2-28^T was 4253095 bp and consisted of one chromosome and four plasmids. A number of genes related to the bacteriochlorophyll *a* production, photosynthetic reaction, cold adaptation, salt adaptation, ultra-violet resistance and DNA damage repairing were found in the genome. In addition to genomic islands and type IV secretion systems, genes related to gene transfer agents were detected in the genome of *R. antarcticum* ZS2-28^T, suggesting that this bacterium can adapt to its environment by acquiring exogenous nucleic acids. The annotated complete genome sequence provides genetic insights into the environmental adaptation and ecological function of *R. antarcticum* ZS2-28^T in Antarctic coastal area.

Keywords ***Roseicitreum*, complete genome, adaptation, gene transfer, strain, intertidal sediment, Antarctica**

Citation: Zeng Y X, Yu Y, Li H R, et al. Complete genome analysis of bacteriochlorophyll *a*-containing *Roseicitreum antarcticum* ZS2-28^T reveals its adaptation to Antarctic intertidal environment. Adv Polar Sci, 2021, 32(1): 20-27, doi: 10.13679/j.advps.2020.0034

1 Introduction

Aerobic anoxygenic phototrophic bacteria (AAPB) are bacteriochlorophyll *a*-containing bacteria with the capability of photoheterotrophy; they appear to play a unique role in the ocean's carbon cycle (Swingley et al., 2007; Tang et al., 2010; Graham et al., 2018). Heterotrophy usually is the main system of energy gain of AAPB (Beatty, 2002), and phototrophy is minimal (Ferrera et al., 2017). However, light can enhance the growth rates of AAPB (Ferrera et al., 2017; Piwosz et al., 2018). AAPB are widely distributed in open and coastal oceans (Jiao et al., 2007), including Arctic and Antarctic marine environments (Boeuf et al., 2013; Zeng et al., 2016). AAPB are phylogenetically diverse and include members of the *Alpha-*, *Beta-* and *Gammaproteobacteria* (Imhoff et al., 2017; Lehours et al., 2018; Auladell et al., 2019). Physiological constraints play an important role in structuring AAPB assemblages at a global scale, and salinity seems to favor lineage-specific adaptations of AAPB (Lehours et al., 2018). AAPB belonging to *Alphaproteobacteria* and *Betaproteobacteria* dominate the offshore and river-influenced surface waters in the western Arctic Ocean, respectively (Boeuf et al., 2013). Represented by the orders *Rhodospirillales, Rhizobiales*

* Corresponding author, ORCID: 0000-0002-3689-7855, E-mail: yxzeng@yahoo.com

and *Rhodobacterales* (Imhoff et al., 2017), members of the *Alphaproteobacteria* AAPB are widespread and abundant in Arctic and Antarctic environments (Koh et al., 2011; Boeuf et al., 2013; Lehours and Jeanthon, 2015; Zeng et al., 2016).

The genus *Roseicitreum* of the family *Rhodobacteriaceae* within the class *Alphaproteobacteria* was proposed by Yu et al. (2011) with one species, *Roseicitreum antarcticum* ZS2-28^T. The bacterium was isolated from sandy intertidal sediment samples collected from the coastal regions of the Chinese Antarctic Zhongshan Station in the Larsemann Hills, Princess Elizabeth Land, Antarctica, during the 23rd Chinese National Antarctic Research Expedition in March 2007 (Yu et al., 2011). Phylogenetic analysis based on 16S rRNA gene sequences indicated that strain ZS2-28^T formed a distinct evolutionary lineage within the clade containing members of the genera *Roseibaca*, *Roseinatronobacter*, and *Rhodobaca* of the class *Alphaproteobacteria* (Yu et al., 2011). *R. antarcticum* ZS2-28^T contained bacteriochlorophyll (Bchl) *a*, and was obligately heterotrophic, strictly aerobic, non-motile, and moderately halophilic (Yu et al., 2011). The phenotypic and chemotaxonomic characteristics of *R. antarcticum* ZS2-28^T were reported by Yu et al. (2011). However, the complete genomic information of *R. antarcticum* ZS2-28^T, which would be helpful for us to understand the adaptation of the bacterium to Antarctic environment and provide further insight into its ecological functions, was not reported. Thus, we performed genome sequencing and present the complete genome sequence of *R. antarcticum* ZS2-28^T in this study by using Illumina HiSeq X and PacBio RSII systems. At the time of writing, *R. antarcticum* ZS2-28^T is the type strain of the sole species with a valid published name within the genus *Roseicitreum*. Therefore, this is also the first complete genome sequence of a *Roseicitreum* strain.

2 Materials and methods

R. antarcticum ZS2-28^T was isolated from sandy intertidal sediment samples in the Larsemann Hills, Princess Elizabeth Land, Antarctica (Yu et al., 2011). The bacterium was deposited into the China General Microbiological Culture Collection Center (CGMCC) with the accession number CGMCC 1.8894^T and the Belgian Coordinated Collections of Micro-organisms (BCCM) with the accession number LMG 24863^T. The complete chromosome sequence and four plasmid sequences of *R. antarcticum* ZS2-28^T have been deposited in GenBank under the accession numbers CP061498, CP061499, CP061500, CP061501 and CP061502, respectively.

The general features of *R. antarcticum* ZS2-28^T and MIGS mandatory information are listed in Table 1. Genomic DNA was extracted from overnight cultures using a MagAttract HMW DNA Kit (Qiagen, Germany) according to the manufacturer's instructions. The harvested DNA was visualized on 1% (w/v) agarose gels, and DNA concentration and purity were measured with a Qubit 2.0 Fluorometer (Life Technologies, USA). Purified DNA was used to construct an Illumina standard shotgun library with an insert size of 300–400 bp followed the NEB Next Ultra DNA Library Prep Kit for Illumina (New England Biolabs, USA), and then was sequenced using the Illumina HiSeq X platform using the PE150 model. A 10-kb DNA library was constructed by the PacBio SMRTbell 10 kb Library preparation Kit (Pacific Biosciences, USA) according to the manufacturer's instructions. Library construction and sequencing were performed at Sangon Biotech Co. Ltd (Shanghai, China). The whole genome sequencing was performed using Illumina HiSeq X (Illumina, USA) and PacBio RSII (Pacific Biosciences, USA) systems.

Table 1 General features of *Roseicitreum antarcticum* ZS2-28^T and MIGS mandatory information

	Items	Description
General feature [a]	Classification	Domain *Bacteria*; Phylum *Proteobacteria*; Class *Alphaproteobacteria*; Order *Rhodobacterales*; Family *Rhodobacteraceae*; Genus *Roseicitreum*; Species *antarcticum*; Strain ZS2-28^T
	Gram-staining	Negative
	Cell shape	Lemon-shaped
	Motility	Non-motile
	Temperature	0–33 ℃(optimum 25–27 ℃)
	Salinity	0–15% (optimum 7%–8%)
	pH	5.0–9.5 (optimum 7.0)
MIGS data	Submitted to INSDC	CP061498-CP061502
	Investigation type	Bacteria
	Project name	Complete genome sequence of *Roseicitreum antarcticum* ZS2-28^T
	Collection date [a]	Mar-2007
	Geographic location [a]	Antarctica: Larsemann Hills, Princess Elizabeth Land
	Environment (biome)	Marine biome (ENVO:00000447)
	Environment (feature)	Cold environment (ENVO:01000309)

Continued

Items		Description
MIGS data	Environment (material)	Sandy sediment (ENVO:01000118)
	Depth	Surficial sediment
	Relationship to oxygen [a]	Strictly aerobic
	Trophic level [a]	Obligately heterotrophic
	Source material identifier [a]	ZS2-28^T = CGMCC 1.8894^T = LMG 24863^T
	Sample collection device	Small sterilized shovel
	Number of replicons	5
	Sequencing platform	Illumina HiSeq and PacBio RSII
	Assembly	Canu (version 1.3) and Pilon (version 1.23)
	Finishing strategy	Complete; 722 × coverage

Note: [a] Data from reference (Yu et al., 2011).

De novo genome assembly was performed using continuous long reads following the Canu workflow v1.3 (Koren et al., 2017), and then Pilon v1.23 was engaged to correct assembled contigs with Illumina reads (Walker et al., 2014). Annotation of the genome was generated by using Prokka v1.10 (Seemann, 2014) to predict coding sequences, ribosomal RNA genes, and transfer RNA genes. A whole genome Blast (v2.2.28) search was performed against the databases CDD, PFAM, COG, NR, Swiss-Prot, and TrEMBL. KEGG ontology was identified by submitting predicted peptides to the KAAS server (http://www.genome.jp/tools/kaas/). GO was detected from Swiss-Prot and TrEMBL annotation results. Four additional databases PHI, VFDB, CARD, and CAZy, were used to annotate peptides. Signal peptides were detected on the genome assembly by SignalP v4.1 (Petersen et al., 2011). Transmembrane proteins were detected by TMHMM v2.0 (Möller et al., 2001). Lipoproteins were detected with LipoP v1.0 (Juncker et al., 2003). Repeat regions within the genome were detected with RepeatModeler (http://www.repeatmasker.org/RepeatModeler/) and RepeatMasker (http://repeatmasker.org), and CRISPRs were detected with CRISPRCasFinder (https://crisprcas.i2bc.paris-saclay.fr). Genomic islands were predicted using IslandPath-DIOMB (Bertelli and Brinkman, 2018). The prophage regions were detected with PhiSpy (Akhter et al., 2012). Architecture of the photosynthesis gene cluster was produced using gggenes v0.4.1 (https://wilkox.org/gggenes/). Multiple sequence alignments (MSAs) were produced with ClustalW algorithm implemented in the MEGA 5.05 software package (https://www.megasoftwa re.net). DNA sequences of *pufL* and *pufM* genes were obtained from GenBank. *PufL* and *pufM* MSAs were concatenated to form a single *pufLM* MSA. A neighbor-joining phylogenetic tree was constructed based on *pufLM* gene sequences using MEGA 5.05.

3 Results and discussion

3.1 General description of the genome

The genome sequence of *R. antarcticum* ZS2-28^T was obtained from a total of 3.0 Gb of high-quality data, which comprised 1.653 Gb Illumina HiSeq X data and 1.396 Gb PacBio RSII data. These data respectively represented 391- and 331-fold coverage of the genome. The genome of *R. antarcticum* ZS2-28^T consists of one chromosome (3537072 bp, 63.45 mol% G + C), and four plasmids, named as pRA01 (425208 bp, 62.70 mol% G + C), pRA02 (131922 bp, 59.56 mol% G + C), pRA03 (103727 bp, 60.17 mol% G + C) and pRA04 (55166 bp, 58.51 mol% G + C), respectively (Table 2). Graphic circular maps of *R. antarcticum* ZS2-28^T are shown in Figure 1.

Table 2 Genomic features of *Roseicitreum antarcticum* ZS2-28^T

Attribute	Chromosome	Plasmid	Plasmid	Plasmid	Plasmid
		pRA01	pRA02	pRA03	pRA04
Size/bp	3537072	425208	131922	103727	55166
(G + C) content/mol%	63.45	62.7	59.56	60.17	58.51
Total genes	3457	412	145	113	54
Protein coding genes	3407	411	145	113	54
rRNA genes	6	0	0	0	0
tRNA genes	44	1	0	0	0
Genes assigned to COG	2471	311	75	52	20
Genes assigned to PFAM	2582	305	85	63	28
Genes assigned to KEGG	1439	203	26	7	3
Transmembrane proteins	739	94	28	12	5
Signal peptides	241	32	10	3	1

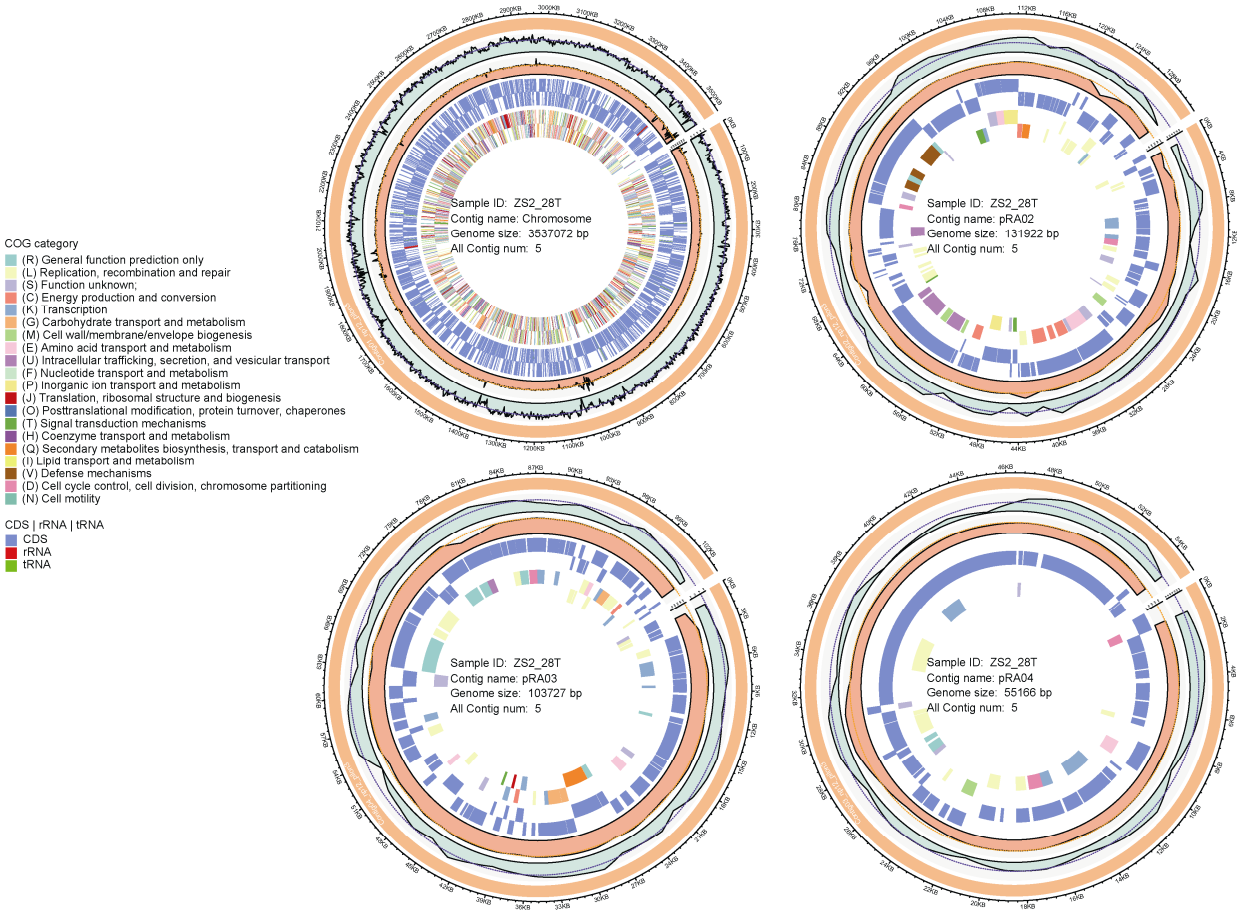

Figure 1 Circular map of one chromosome and three plasmids in the genome of *Roseicitreum antarcticum* ZS2-28^{T}. From outside to the center: the number of bases, chromosome or plasmid, G+C content, coverage, CDS/rRNA/tRNA on the forward strand, CDS/rRNA/tRNA on the reverse strand, COG on the forward strand, and COG on the reverse strand (colored by COG categories).

The genome contains 45 tRNA genes and 2 rRNA operons. The numbers of 5S, 16S and 23S rRNA genes were all two. Among the 4181 predicted protein-coding genes, 3251 (78.93%), 2929 (71.11%), 3918 (95.12%), 3063 (74.36%), 2773 (67.32%), 3904 (94.78%), 2692 (65.36%), and 1856 (45.06%) were annotated by querying the CDD, COG, NR, PFAM, Swiss-Prot, TrEMBL, GO, and KEGG databases, respectively. There were 257 (6.14%) genes that failed to annotate in at least one database. Based on KEGG pathway classification (Figure 2), 74.77% of the annotated genes were found to be involved in metabolisms, including amino acid metabolism (17.75%), carbohydrate metabolism (14.11%), overview (11.61%), energy metabolism (6.76%), metabolism of cofactors and vitamins (6.75%), nucleotide metabolism (5.14%), xenobiotics biodegradation and metabolism (4.26%), lipid metabolism (4.26%), biosynthesis of other secondary metabolites (1.68%), metabolism of terpenoids and polyketides (1.58%), and glycan biosynthesis and metabolism (0.99%).

3.2 Genes related to photoheterotrophic lifestyle

Strain ZS2-28^{T} is obligately heterotrophic (Yu et al., 2011), utilizing L-arabinose, cellobiose, D-galactose, gentiobiose, D-glucose, maltose, D-mannose, L-rhamnose, D-ribose, sucrose, trehalose, turanose, D-xylose, D-mannitol, adipic acid, gluconate, malic acid, glycerin, amygdalin, pyruvate, casein hydrolysate and yeast extract as sole carbon and energy sources. This strain is also positive for arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, urease, β-galactosidase and β-glucosidase activities (Yu et al., 2011). The major COG categories of the genome were amino acid transport and metabolism (10.17%), carbohydrate transport and metabolism (8.71%), transcription (7.17%), replication, recombination and repair (7.10%), energy production and conversion (6.21%), inorganic ion transport and metabolism (5.84%), and translation, ribosomal structure and biogenesis (5.71%). Annotation based on the CAZy database indicates that strain ZS2-28^{T} genome contains a large number of carbohydrate-active genes, including 18 auxiliary activities (AA), 6 carbohydrate-binding modules (CBM), 36 carbohydrate esterases (CE), 50 glycoside hydrolases (GH), 36 glycosyl transferases (GT), and 2 polysaccharide lyase (PL). In addition, annotation based on the PFAM database reveals a total of 93 peptidases and 23 proteases in the genome.

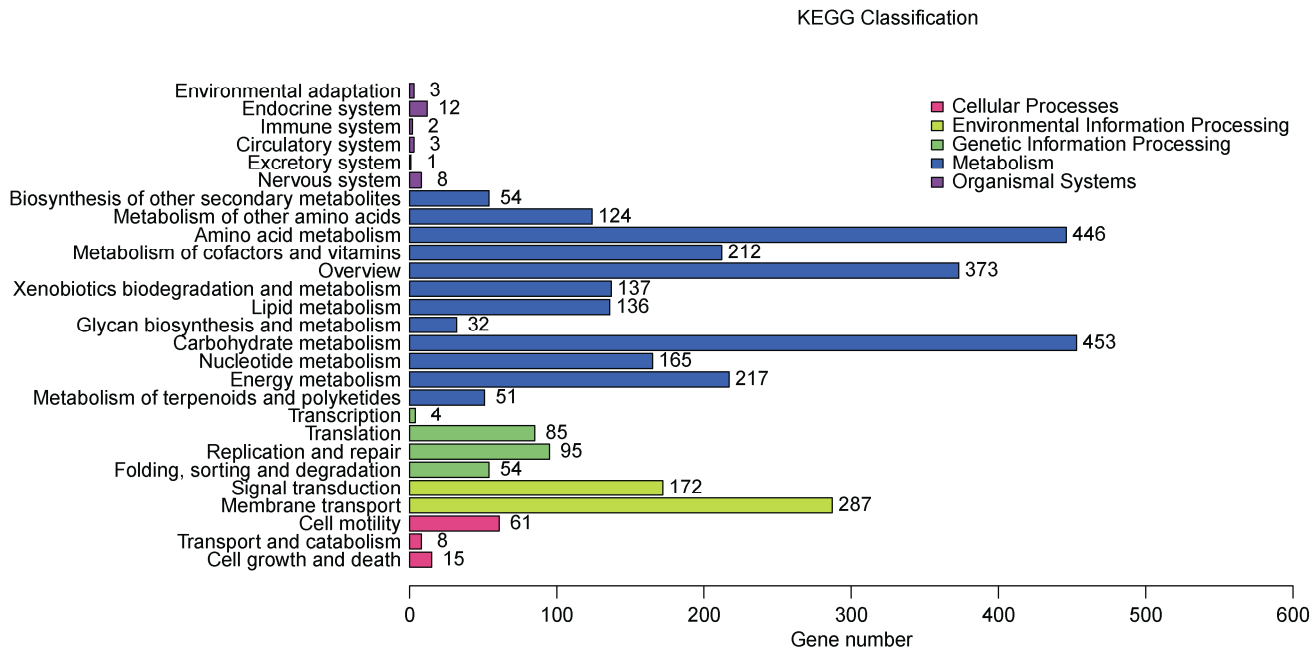

Figure 2 KEGG pathway classification of the annotated genes.

The strain ZS2-28^T can produce Bchl *a* (Yu et al., 2011). In the biosynthesis of BChl *a*, conversion of chlorin to bacteriochlorin ring is known to be catalyzed by chlorophyllide *a* oxidoreductase (COR), which is a nitrogenase-like enzyme and a three-subunit complex consisting of BchX, BchY and BchZ (Tsukatani et al., 2013). Product of this reaction is further catalyzed by BchF and BchC (Tsukatani et al., 2013). According to genome information, the *bchXYZ* gene set, and *bchC* and *bchF* genes were present in strain ZS2-28^T (Figure 3a), and showed more than 72% sequence similarities to those of *Rhodobacteriaceae* based on BlastX searching in NCBI database. There were 23 genes related to the formation of Bchl found in the chromosome based on GO annotation. Two proteins (Pfam05398 and Pfam07284) involved in Bchl biosynthesis pathway were observed. At the same time, similar to the oxygen regulated puf operon of purple photosynthetic bacterium *Rhodobacter sphaeroides* (Chidgey et al., 2017), *pufQ*, *pufB* and *pufA* genes encoding the light-harvesting (LH)1 α and β polypeptides, and *pufL*, *pufM* and *puhA* genes encoding the type-II photosynthetic reaction center (RC) L, M and H submits were observed in the genome of strain ZS2-28^T (Lee et al., 1989; Hunter et al., 1991; Imhoff et al., 2019). In addition, a photosynthetic reaction center cytochrome C encoding gene *pufC* was found in the genome. The cytochrome associated with the photosynthetic reaction center is an important component in many of the PS-II type photosynthetic bacteria (Imhoff et al., 2019). However, different from *Rhodobacter sphaeroides*, *pufX* gene encoding the PufX polypeptide was absent from the genome of strain ZS2-28^T. Three genes encoding proteins (Pfam02276 for photosynthetic reaction center cytochrome C subunit, Pfam03073 for sensory protein and Pfam04940 for blue light sensor protein) involved in photosynthesis were found in the chromosome. Phylogenetic analysis (Figure 3b) based on *pufL* and *pufM* genes indicated that strain ZS2-28^T fell into the *Roseobacter* clade containing the genera *Jannaschia*, *Mameliella*, *Roseovarius*, *Tateyamaria* and *Thalassococcus*. The *pufL* and *pufM* genes of strain ZS2-28^T showed close relationship (77.6% sequence similarity) to *Jannaschia* sp. CCS1.

The *acsF* gene encoding Mg-protoporphyrin IX monomethyl ester cyclase was observed in strain ZS2-28^T, showing more than 80% sequence similarities to those of *Rhodobacteriaceae* based on BlastX searching in NCBI database. AcsF activity exists only under aerobic growth conditions in a *Betaproteobacterium* (Pinta et al., 2002). As an analogue of the oxygen-dependant cyclase encoded by *acsF* gene, *bchE* gene encoding the oxygen-independent Mg-protoporphyrin monomethylester cyclase was not detected in the genome of strain ZS2-28^T. The *bchE* gene seems to be unessential for phototrophy in *Roseobacter* species (Koblížek et al., 2013). Results support that strain ZS2-28^T is an aerobic photosynthetic purple bacterium, and are consistent with previous finding that strain ZS2-28^T is photoheterotrophic (Yu et al., 2011).

3.3 Genes related to genetic exchange

Horizontal gene transfer (HGT) has been regarded to play an important role in the adaptation of the microbes to environment by providing the tools that are necessary to face the adversity and survive in a harsh environment (Springael and Top, 2003; Paquola et al., 2018). Consistent with the finding of a gene transfer agent (GTA) capsid

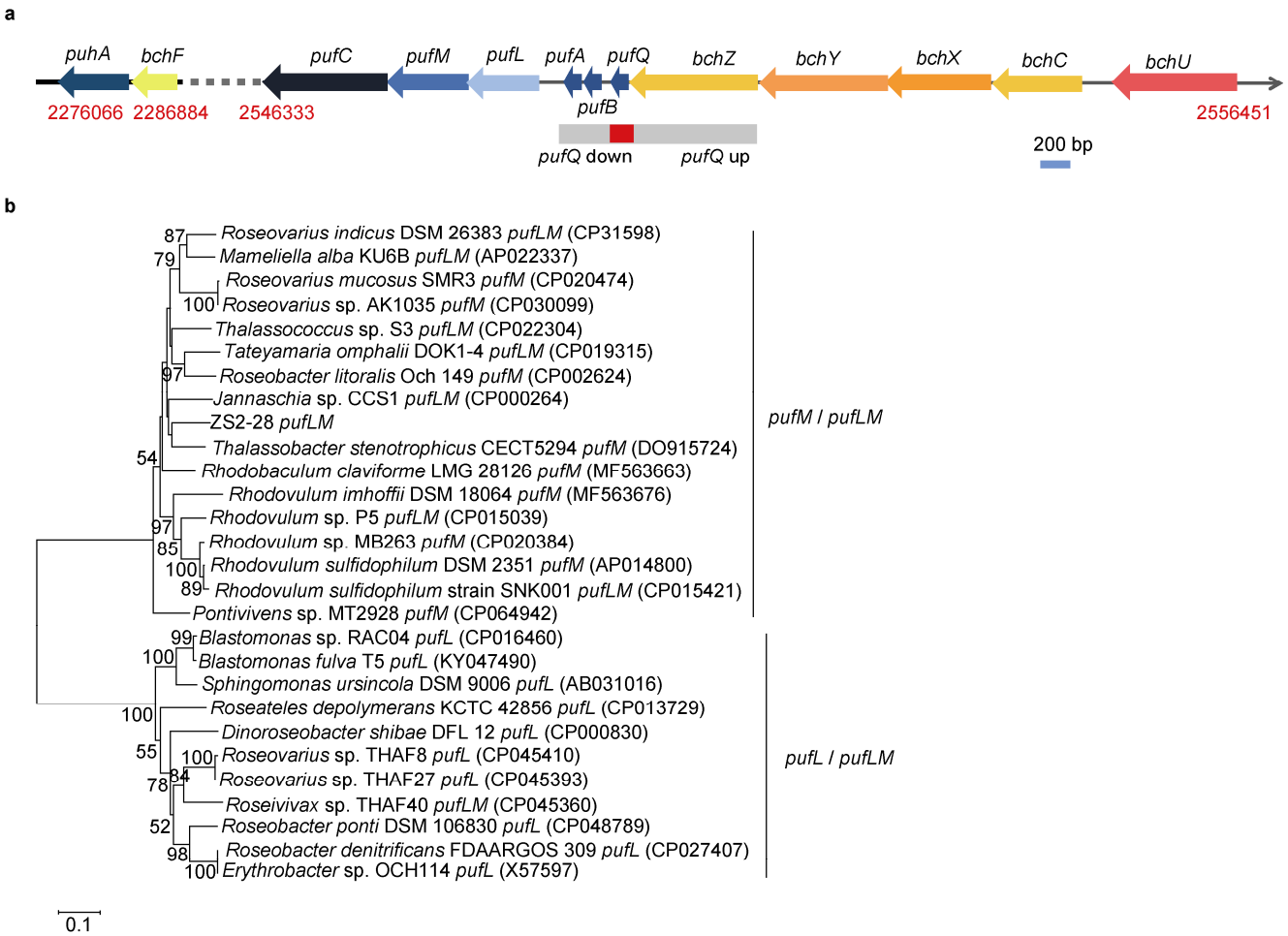

Figure 3 Photosynthesis genes in the genome of *Roseicitreum antarcticum* ZS2-28^{T}. **a**, Genes in the potential photosynthesis gene cluster of *R. antarcticum* ZS2-28^{T}; **b**, Neighbor-joining phylogenetic tree based on *pufL* and *pufM* gene sequences showing the relationship between strain ZS2-28^{T} and representatives of some other related taxa. Numbers at nodes indicate percentages of 1000 bootstrap re-samplings, only values above 50% are shown. Bar, 0.1 substitutions per site.

protein gene (*g*5) in strain ZS2-28^{T} (Zeng, 2019), a total of six genes related to GTA were observed in the chromosome based on NR annotation. GTAs have evolved from prophages that have lost the ability to target their own DNA for packaging (Lang et al., 2012). However, no prophage region was found in the genome of strain ZS2-28^{T}. Neither repeated regions nor the CRISPR-Cas system were detected in the genome.

A total of 16 genomic islands were predicted in the chromosome. At the same time, five type IV secretion system genes were observed in the genome based on Swiss-Prot annotation. Type IV secretion systems and genomic islands-mediated horizontal gene transfer have been reported in *Pseudomonas* and *Haemophilus* (Juhas, 2015). Transposases and integrases can mediate the movement of DNA sequences within or between genomes (Rice and Baker, 2001). Strain ZS2-28^{T} contains 74 transposase and 42 integrase genes, and annotation based on the Pfam database indicated that diverse transposases (Pfam00872, Pfam01527, Pfam01548, Pfam01695, Pfam02371, Pfam03050, Pfam03400, Pfam04986, Pfam05598, Pfam13005, Pfam13340, Pfam13586, Pfam13610, Pfam13737, and Pfam13751) and integrases (Pfam00589, Pfam00665, and Pfam13683) are present in the genome. These results suggest that *R. antarcticum* ZS2-28^{T} may adapt to the environment by acquiring exogenous nucleic acids.

3.4 Genes related to adaptation to Antarctic intertidal environment

The combination of seasonal scouring and encasement in ice, high UV irradiation, and high levels of salinity and temperature fluctuations make the Antarctic intertidal zone possibly the world's most physically disturbed environment (Peck et al., 2006). Cells of strain ZS2-28^{T} were surrounded with slime (Yu et al., 2011). Slime is actually a kind of exopolysaccharides (EPS) produced by bacteria, and is helpful for bacteria surviving in extreme environments (Flemming, 2016). Four capsule polysaccharide biosynthesis proteins (Pfam05159) and 21 glycosyl

transferases (Pfam00534, Pfam00535, Pfam00591, Pfam00953, Pfam00982, Pfam03808, Pfam04101, Pfam13632, Pfam13641 and Pfam13704) were detected in the genome.

One ultra-violet resistance protein (Pfam12344), showing 85% similarity to excinuclease ABC subunit UvrB of *Rhizobium loti* strain MAFF303099 based on Swiss-Prot database, was detected in the chromosome. At the same time, KEGG pathway annotation reveals that 8 genes encoding single-strand DNA-binding protein Ssb and ATP-dependent DNA helicase RecG are present in the genome. Those proteins may play role in repairing of DNA damage caused by ultra-violet irradiation (Trgovcević et al., 1989; Xu et al., 2020). It is interesting to find that all three genes assigned to KEGG in plasmid pRA04 (Table 2) are associated with single-strand DNA-binding protein Ssb, ATP-dependent DNA helicase RecG and ATP-dependent RNA helicase DeaD, suggesting that the smallest plasmid may play a role in genetic information processing in strain ZS2-28^T.

RNA helicase controls RNA folding and degradation in bacteria under low temperatures (Médigue et al., 2005). Annotation based on the Pfam database indicates that diverse RNA helicases, including 4 DEAD/DEAH box helicases (Pfam00270), 1 DEAD/H associated (Pfam08494), 1 SecA DEAD-like domain (Pfam07517), and 3 helicase conserved C-terminal domains (Pfam00271) are present in the genome of strain ZS2-28^T. Four genes related to cold-shock DNA-binding domain (Pfam00313) were also observed in the genome. At the same time, five salt adaptation-related genes, including *ectABC*, *trkA* and *trkH* (Kraegeloh et al., 2005; Sadeghi et al., 2014), were detected in the genome based on Pfam annotation. These results suggest that strain ZS2-28^T is adapted to cold intertidal environment.

In conclusion, the genomic analysis of *Roseicitreum antarcticum* ZS2-28^T isolated from Antarctic intertidal sediment has revealed that its genome contains various genes involved in the bacterium's Bchl *a* production, photosynthetic reaction, cold adaptation, salt adaptation, ultra-violet resistance, and horizontal gene transfer events. The genome sequence will improve our understanding of the environmental adaptations and ecological functions of AAPB in the Antarctic marine environment.

Acknowledgments This work was supported by the National Key R & D Program of China (Grant no. 2018YFC1406903) and the National Natural Science Foundation of China (Grant nos. 91851201 and 41476171). We would like to thank two anonymous reviewers and Associate Editor, Dr Steve Coulson for their valuable suggestions and comments on further improvement of this article.

References

Akhter S, Aziz R K, Edwards R A. 2012. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. Nucleic Acids Res, 40(16): e126, doi: 10.1093/nar/gks406.

Auladell A, Sánchez P, Sánchez O, et al. 2019. Long-term seasonal and interannual variability of marine aerobic anoxygenic photoheterotrophic bacteria. ISME J, 13(8): 1975-1987, doi: 10.1038/s41396-019-0401-4.

Beatty J T. 2002. On the natural selection and evolution of the aerobic phototrophic bacteria. Photosynth Res, 73(1-3): 109-114, doi: 10.1023/A:1020493518379.

Bertelli C, Brinkman F S L. 2018. Improved genomic island predictions with IslandPath-DIMOB. Bioinformatics, 34(13): 2161-2167, doi: 10.1093/bioinformatics/bty095.

Boeuf D, Cottrell M T, Kirchman D L, et al. 2013. Summer community structure of aerobic anoxygenic phototrophic bacteria in the western Arctic Ocean. FEMS Microbiol Ecol, 85(3): 417-432, doi: 10.1111/1574-6941.12130.

Chidgey J W, Jackson P J, Dickman M J, et al. 2017. *PufQ* regulates porphyrin flux at the haem/bacteriochlorophyll branchpoint of tetrapyrrole biosynthesis via interactions with ferrochelatase. Mol Microbiol, 106(6): 961-975, doi: 10.1111/mmi.13861.

Ferrera I, Sánchez O, Kolářová E, et al. 2017. Light enhances the growth rates of natural populations of aerobic anoxygenic phototrophic bacteria. ISME J, 11(10): 2391-2393, doi: 10.1038/ismej.2017.79.

Flemming H C. 2016. EPS-then and now. Microorganisms, 4(4): E41, doi: 10.3390/microorganisms4040041.

Graham E D, Heidelberg J F, Tully B J. 2018. Potential for primary productivity in a globally-distributed bacterial phototroph. ISME J, 12(7): 1861-1866, doi: 10.1038/s41396-018-0091-3.

Hunter C N, McGlynn P, Ashby M K, et al. 1991. DNA sequencing and complementation/deletion analysis of the *bchA - puf* operon region of *Rhodobacter sphaeroides*: *in vivo* mapping of the oxygen-regulated *puf* promoter. Mol Microbiol, 5(11): 2649-2661, doi: 10.1111/j.1365-2958.1991.tb01974.x.

Imhoff J F, Rahn T, Künzel S, et al. 2017. Photosynthesis is widely distributed among Proteobacteria as demonstrated by the phylogeny of PufLM reaction center proteins. Front Microbiol, 8: 2679, doi: 10.3389/fmicb.2017.02679.

Imhoff J F, Rahn T, Künzel S, et al. 2019. Phylogeny of anoxygenic photosynthesis based on sequences of photosynthetic reaction center proteins and a key enzyme in bacteriochlorophyll biosynthesis, the chlorophyllide reductase. Microorganisms, 7(11): 576, doi: 10.3390/microorganisms7110576.

Jiao N, Zhang Y, Zeng Y, et al. 2007. Distinct distribution pattern of abundance and diversity of aerobic anoxygenic phototrophic bacteria in the global ocean. Environ Microbiol, 9(12): 3091-3099, doi: 10.1111/j.1462-2920.2007.01419.x.

Juhas M. 2015. Type IV secretion systems and genomic islands-mediated horizontal gene transfer in *Pseudomonas* and *Haemophilus*. Microbiol Res, 170: 10-17, doi: 10.1016/j.micres.2014.06.007.

Juncker A S, Willenbrock H, von Heijne G, et al. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci, 12(8): 1652-1662, doi: 10.1110/ps.0303703.

Koblížek M, Zeng Y, Horák A, et al. 2013. Regressive evolution of photosynthesis in the *Roseobacter* clade//Beatty J T. Advances in botanical research, Volume 66. Amsterdam: Elsevier, Academic Press, 385-405, doi: 10.1016/b978-0-12-397923-0.00013-8.

Koh E Y, Phua W, Ryan K G, 2011. Aerobic anoxygenic phototrophic

bacteria in Antarctic sea ice and seawater. Environ Microbiol Rep, 3(6): 710-716, doi: 10.1111/j.1758-2229.2011.00286.x.

Koren S, Walenz B P, Berlin K, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res, 27(5): 722-736, doi: 10.1101/gr.215087.116.

Kraegeloh A, Amendt B, Kunte H J. 2005. Potassium transport in a halophilic member of the bacteria domain: identification and characterization of the K^+ uptake systems TrkH and TrkI from *Halomonas elongata* DSM 2581^T. J Bacteriol, 187(3): 1036-1043, doi: 10.1128/jb.187.3.1036-1043.2005.

Lang A S, Zhaxybayeva O, Beatty J T. 2012. Gene transfer agents: phage-like elements of genetic exchange. Nat Rev Microbiol, 10(7): 472-482, doi: 10.1038/nrmicro2802.

Lee J K, DeHoff B S, Donohue T J, et al. 1989. Transcriptional analysis of *puf* operon expression in *Rhodobacter sphaeroides* 2.4.1 and an intercistronic transcription terminator mutant. J Biol Chem, 264(32): 19354-19365, doi: 10.1016/S0021-9258(19)47309-2.

Lehours A C, Enault F, Boeuf D, et al. 2018. Biogeographic patterns of aerobic anoxygenic phototrophic bacteria reveal an ecological consistency of phylogenetic clades in different oceanic biomes. Sci Rep, 8(1): 4105, doi: 10.1038/s41598-018-22413-7.

Lehours A C, Jeanthon C. 2015. The hydrological context determines the beta-diversity of aerobic anoxygenic phototrophic bacteria in European Arctic seas but does not favor endemism. Front Microbiol, 6: 638, doi: 10.3389/fmicb.2015.00638.

Médigue C, Krin E, Pascal G, et al. 2005. Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. Genome Res, 15(10): 1325-1335, doi: 10.1101/gr.4126905.

Möller S, Croning M D R, Apweiler R. 2001. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics, 17(7): 646-653, doi: 10.1093/bioinformatics/17.7.646.

Paquola A C M, Asif H, Pereira C A D B, et al. 2018. Horizontal gene transfer building prokaryote genomes: genes related to exchange between cell and environment are frequently transferred. J Mol Evol, 86(3-4): 190-203, doi: 10.1007/s00239-018-9836-x.

Peck L S, Convey P, Barnes D K A. 2006. Environmental constraints on life histories in Antarctic ecosystems: tempos, timings and predictability. Biol Rev, 81(1): 75-109, doi: 10.1017/S146479310 5006871.

Petersen T N, Brunak S, von Heijne G, et al. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods, 8(10): 785-786, doi: 10.1038/nmeth.1701.

Pinta V, Picaud M, Reiss-Husson F, et al. 2002. Rubrivivax gelatinosus acsF (previously orf358) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. J Bacteriol, 184(3): 746-753, doi: 10.1128/jb.184.3.746-753.2002.

Piwosz K, Kaftan D, Dean J, et al. 2018. Nonlinear effect of irradiance on photoheterotrophic activity and growth of the aerobic anoxygenic phototrophic bacterium *Dinoroseobacter shibae*. Environ Microbiol, 20(2): 724-733, doi: 10.1111/1462-2920.14003.

Rice P A, Baker T A. 2001. Comparative architecture of transposase and integrase complexes. Nat Struct Biol, 8(4): 302-307, doi: 10.1038/ 86166.

Sadeghi A, Soltani B M, Nekouei M K, et al. 2014. Diversity of the ectoines biosynthesis genes in the salt tolerant *Streptomyces* and evidence for inductive effect of ectoines on their accumulation. Microbiol Res, 169(9-10): 699-708, doi: 10.1016/j.micres.2014.02. 005.

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics, 30(14): 2068-2069, doi: 10.1093/bioinformatics/ btu153.

Springael D, Top E M. 2003. Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. Trends Microbiol, 12(2): 53-58, doi: 10.1016/j.tim.2003.12.010.

Swingley W D, Sadekar S, Mastrian S D, et al. 2007. The complete genome sequence of *Roseobacter denitrificans* reveals a mixotrophic rather than photosynthetic metabolism. J Bacteriol, 189(3): 683-690, doi: 10.1128/jb.01390-06.

Tang K, Zong R, Zhang F, et al. 2010. Characterization of the photosynthetic apparatus and proteome of *Roseobacter denitrificans*. Curr Microbiol, 60(2): 124-133, doi: 10.1007/s00284-009-9515-7.

Trgovcević Z, Lers N, Brcić-Kostić K, et al. 1989. Post-ultraviolet DNA synthesis in the absence of repair: role of the single-strand DNA-binding protein. Int J Radiat Biol, 55(5): 739-745, doi: 10.1080/09553008914550791.

Tsukatani Y, Yamamoto H, Harada J, et al. 2013. An unexpectedly branched biosynthetic pathway for bacteriochlorophyll *b* capable of absorbing near-infrared light. Sci Rep, 3(1): 1-7, doi: 10.1038/srep 01217.

Walker B J, Abeel T, Shea T, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One, 9(11): e112963, doi: 10.1371/journal. pone.0112963.

Xu J, Wu C X, Yang Z H, et al. 2020. For: Pesticide biochemistry and physiology recG is involved with the resistance of Bt to UV. Pestic Biochem Physiol, 167: 104599, doi: 10.1016/j.pestbp.2020.104599.

Yu Y, Yan S L, Li H R, et al. 2011. *Roseicitreum antarcticum* gen. nov., sp. nov., an aerobic bacteriochlorophyll *a*-containing alphaproteobacterium isolated from Antarctic sandy intertidal sediment. Int J Syst Evol Microbiol, 61: 2173-2179, doi: 10.1099/ijs.0.024885-0.

Zeng Y X, Dong P Y, Qiao Z Y, et al. 2016. Diversity of the aerobic anoxygenic phototrophy gene *pufM* in Arctic and Antarctic coastal seawaters. Acta Oceanol Sin, 35(6): 68-77, doi: 10.1007/s13131-016-0877-y.

Zeng Y X. 2019. Phylogenetic diversity of dimethylsulfoniopropionate-dependent demethylase gene *dmdA* in distantly related bacteria isolated from Arctic and Antarctic marine environments. Acta Oceanol Sin, 38(8): 64-71, doi: 10.1007/s13131-019-1393-7.