Marine Genomics 55 (2021) 100790

Contents lists available at ScienceDirect

Marine Genomics

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

Complete genome of *Polaromonas vacuolata* KCTC 22033^T isolated from beneath Antarctic Sea ice

Kyuin Hwang [a], Hanna Choe [b], Arshan Nasir [c], Kyung Mo Kim [a,*]

[a] Division of Polar Life Sciences, Korea Polar Research Institute, 26 Songdomirae-ro, Yeonsu-gu, Incheon 21990, Republic of Korea
[b] Biological Resource Center, Korea Research Institute of Bioscience and Biotechnology, 181 Ipsin-gil, Jeongeup 56212, Republic of Korea
[c] Theoretical Biology & Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, USA

ARTICLE INFO

Keywords:
Betaine
Genome reduction
Horizontal gene transfer
Psychrophile
Rhodopsin

ABSTRACT

Polaromonas vacuolata KCTC 22033^T is an obligate aerobic, Gram-negative, psychrophilic and rod-shaped bacterium isolated from beneath the sea ice off the coast of the Palmer Peninsula, Anvers Islands, Antarctica. *P. vacuolata* is the type species of *Polaromonas* genus and the first example of gas vacuolate *Betaproteobacteria* isolated from marine habitats. Here, we report a complete genome of *P. vacuolata* KCTC 22033^T, which consists of 3,837,686 bp (G + C content of 52.07%) with a single chromosome, 3461 protein-coding genes, 56 tRNAs and 6 rRNA operons. Genomic analysis revealed the presence of genes involved in bacterial adaptation under saline conditions, cold adaptation via the production of gas vesicles and cell adhesion proteins, and a photoheterotrophic lifestyle when challenged by starvation. Intriguingly, several of these genes were likely acquired from species outside the *Polaromonas* genus. The genomic information therefore describes the unique evolution and adaptation of *P. vacuolata* to its extraordinary habitat, i.e., beneath the Antarctic sea ice.

1. Introduction

The bacterial genus *Polaromonas* was first described by Irgens and colleagues in Irgens et al., 1996 (Irgens et al., 1996) and currently consists of nine recognized species (Choi et al., 2018). The genus is closely related to genera *Acidovorax*, *Rhodoferax* and *Variovorax* and is classified into family *Comamonadaceae* of *Betaproteobacteria* (Choi et al., 2018). Members of *Polaromonas* have been isolated from a variety of habitats, including seawater, ground water, tap water, sediment, soil, alpine glacier cryoconite and ice core (Mattes et al., 2008; Choi et al., 2018; and refs therein), and are widely distributed from temperate to polar regions. Among these, *Polaromonas vacuolata* KCTC 22033^T (= 34-P^T = ATCC 51984^T) is an obligate aerobic, Gram-negative, psychrophilic (optimally growing at 4 °C), rod-shaped and vacuolate bacterium (Irgens et al., 1996; Irgens et al., 1989). Interestingly, the strain was isolated from beneath sea ice off the Palmer Peninsula near the U.S. Palmer Station, Anvers Island, Antarctica (Irgens et al., 1996; Irgens et al., 1989). The bacterium produces gas vacuoles that allow it to remain buoyant in the stratified nutrient-rich layer beneath the Antarctic sea ice (Irgens et al., 1989), where it can directly utilize nutrients produced by the sea ice bottom-dwelling primary producers (i.e. photosynthetic algae) (Guo et al., 2017).

At this extraordinary habitat, previous studies of *P. vacuolata* have been limited to conventional phenotypic characterizations including utilization of carbon sources, fatty acid composition, salinity range, growth temperature and colony morphology (Irgens et al., 1989; Irgens et al., 1996). Although four complete genomes (*P. naphthalenivorans* CJ2, and strains JS666, SP1 and Pch-P of *Polaromonas* spp.), four high-quality genomes with small numbers of contigs (*P. glacialis* R3–9, and *Polaromonas* strains EUR3 1.2.1, CG9_12 and YR568), seven low-quality genomes and 14 metagenome-assembled genomes have been described in *Polaromonas* genus, the genome of the type species *P. vacuolata* has not been sequenced. Previous genomic analysis revealed the versatile metabolic potential of *Polaromonas* members that possibly results from extensive horizontal gene transfer (HGT) (Yagi et al., 2009). Therefore, the *P. vacuolata* genome, described in this study, can further enhance our understanding of habitat adaptation and evolution of *Polaromonas* members.

2. Data description

General features of strain KCTC 22033^T are summarized in Table 1. Cells of KCTC 22033^T were grown for three weeks at 4 °C on Marine agar (MB, BD). Isolated colonies were picked using a sterile toothpick and

* Corresponding author.
E-mail address: kmkim@kopri.re.kr (K.M. Kim).

https://doi.org/10.1016/j.margen.2020.100790
Received 6 May 2020; Received in revised form 1 June 2020; Accepted 5 June 2020
Available online 17 June 2020

Table 1
General features of *Polaromonas vacuolata* KCTC 22033^T and MIGS mandatory information.

Item	Description
General features	
Classification	Domain *Bacteria* Phylum *Proteobacteria* Class *Betaproteobacteria* Order *Burkholderiales* Family *Comamonadaceae* Genus *Polaromonas*
Type strain	KCTC 22033^T
Gram strain	Negative
Cell shape	Rods
Motility	Motile
Temperature range	0–12 °C, optimally at 4 °C
Salinity range	NaCl, 0 to 6%
pH range	6.0 to 9.5
MIGS data	
Submitted to INSDC	Accession number CP051461
Investigation type	Bacteria
Project name	Complete genome sequence of *Polaromonas vacuolata* KCTC 22033
Geographic location	Anvers Island, Antarctica
Lat_lon	64.55° S 63.58° W
Depth	NA
Collection date	1986–12
Environment (biome)	Polar biome (ENVO:01000339)
Environment (feature)	Cold temperature habitat (ENVO:00002026)
Environment (material)	Coastal sea water (ENVO:00002150)
Environment (package)	Sea water (ENVO:00002149)
Isolation and growth conditions	PMID: 8782696
Sequencing	
Sequencing platform	PacBio RS II with P6-C4 chemistry
Fold coverage	302.87×
Assembler	SMRT Analysis v2.3.0
Annotation source	Prokka v1.13

genomic DNA was extracted using the i-genomic BYF mini kit (iNtRON Biotechnology, Republic of Korea) following manufacturer's protocols. Genome sequencing was performed using PacBio RS II single-molecule real-time (SMRT) sequencing technology (Pacific Biosciences, USA). A 20-kb insert SMRTbell library was constructed and sequenced, yielding a 302-fold average genome coverage. De novo assembly of 119,996 subreads with 11,199 nucleotides on average (1,343,943,408 bp in total) was conducted using the hierarchical genome-assembly process pipeline of the SMRT Analysis v2.3.0 (Chin et al., 2013). According to the procedure of (Kim et al., 2017), raw PacBio reads were re-mapped into a genome in order to correct sequencing errors.

Gene calling, annotation, and CRISPRs detection were performed using Prokka v1.13 (*E-value* < 0.001) (Seemann, 2014). Pseudogenes were predicted by NCBI Microbial Genome Submission Check available at https://www.ncbi.nlm.nih.gov/genomes/frameshifts/frameshifts.cgi . All predicted CDSs were BLAST-searched against COG (Tatusov et al., 2000) and KEGG (Kanehisa and Goto, 2000) databases (>30% sequence identity, *E-value* < 0.0001) to gain insights into their molecular functions. Giant proteins were additionally analyzed at protein domain and superfamily levels using InterProScan v5.42 (Jones et al., 2014). Signal peptides, transmembrane helices, genomic islands (GIs), and prophages were predicted using SignalP v5 (Armenteros et al., 2019), TMHMM v2.0 (Krogh et al., 2001), IslandViewer v4 (Bertelli et al., 2017), and PHASTER (Arndt et al., 2016), respectively. The graphic circular map of genome was constructed and visualized using CGView server v1.0 (Grant and Stothard, 2008).

The complete genome of *P. vacuolata* KCTC 22033^T is composed of a circular chromosome of 3,837,686 bp with 52.1% G + C content (Table 2). No plasmid or CRISPR sequences were detected. In total, 3461 protein coding genes, 6 rRNA operons, 56 tRNAs and one tmRNA were found (Fig. 1; Table 2). A total of 50 out of 3461 protein coding genes (1.44%) were annotated as pseudogenes. Signal peptides and transmembrane helices were detected in 462 (13.35%) and 791 proteins (22.85%), respectively. Nine GIs as well as one intact prophage (34,961 bp in length) were also detected (see the blue and red boxes on the innermost circle of Fig. 1, respectively).

Table 2
Genomic features of *Polaromonas vacuolata* KCTC 22033^T.

Attribute	Value
Genome features	
Genome size (bp)	3,837,686
Protein coding region (bp)	3,371,552
G + C content (%)	52.1
Chromosome	1
Total genes	3536
Protein coding genes	3461
rRNA operons	6 (23S–5S-16S)
tRNA	56
Pseudogenes	50
Genes assigned to COGs	3017
Genes with signal peptides	462
Genes with transmembrane helices	791

The genome of *P. vacuolata* is significantly shorter (3.84 Mb) and has a lower G + C content (52.1%) compared to other available *Polaromonas* genomes (4.40–5.9 Mb, 60–63% G + C). The decrease in G + C content typically leads to an increase in DNA flexibility (Liu et al., 2020), which can facilitate bacterial growth and DNA replication under low temperatures. This might explain how *P. vacuolata* has adapted to the cold habitat beneath the Antarctic sea ice. In turn, the shorter genome may indicate past or ongoing reductive evolution in the genome of *P. vacuolata*. Intriguingly, and despite harboring a relatively reduced genome, *P. vacuolata* encodes 6 rRNA operons compared to 1–3 previously detected in *Polaromonas* genomes or between 3.1 and 3.6 detected on average at higher taxonomic levels (e.g., *Comamonadaceae, Burkholderiales,* and *Betaproteobacteria* (Stoddard et al., 2015)). This is an atypical observation since the number of rRNA operons usually positively correlate with genome size (Roller et al., 2016).

A total of nine GIs representing 5.66% of total genes (200 out of 3536) were sparsely distributed in the *P. vacuolata* genome (blue boxes in Fig. 1). The majority of these genes exhibited significant sequence homologies to genes outside *Polaromonas* species indicating diverse and complex evolutionary origins. For example, the first half of GI-1 (locus tag HC248_00182 to _00197) encodes the subunits of ATP synthase complex shared with *Polaromonas* relatives while the second half (HC248_00200 to _00213) had top BLAST matches to genes of *Acidovorax* (*Comamonadaceae*) species. The genes of GI-2 (HC248_00483 to _00494) were predicted to be involved in cell wall modification (e.g. S-layer protein, *bam*B, *oat*A genes) and had top BLAST-hits with genes of *Azovibrio restrictus* (*Rhodocyclales*, *Betaproteobacteria*) while the last four GI-3 genes (HC248_01313 to _01320) were assigned as hypothetical proteins and had top BLAST matches to genes of *Azotobacter beijerinckii* (*Gammaproteobtacteria*). Interestingly, two of these four GI-3 genes encoded giant proteins (HC248_01317 with 7533 amino acids, HC248_01318 with 3441 amino acids) that possess calcium-dependent cell adhesion domains/superfamilies necessary either for adhesion to host cells or for bacterial ice-binding (Guo et al., 2017) (see below for ecological relevance of these proteins). The GI-4 (HC248_01424 to _01445) partially overlapped with the intact prophage (HC248_01388 to _01432) while GI-5 (HC248_01907, _01909, _01918) harbored three gas vesicle proteins supporting the existence of gas vacuoles in this bacterium (Irgens et al., 1996). The genomic fragment (HC248_1902 to _01919, except on HC248_1914) including the three vacuole-related genes exhibited top BLAST-hits with genes of *Marinobacter* (*Gammaproteobacteria*). Gene sequences of GI-6 (HC248_02546 and _02547), albeit predicted as a genomic island, were nearly identical to those of other *Polaromonas* strains. The first half of GI-7 (HC248_03177 to _03185) was not found in other *Polaromonas* species, but detected in *Rhodoferax ferrireducens* that belongs to the same family

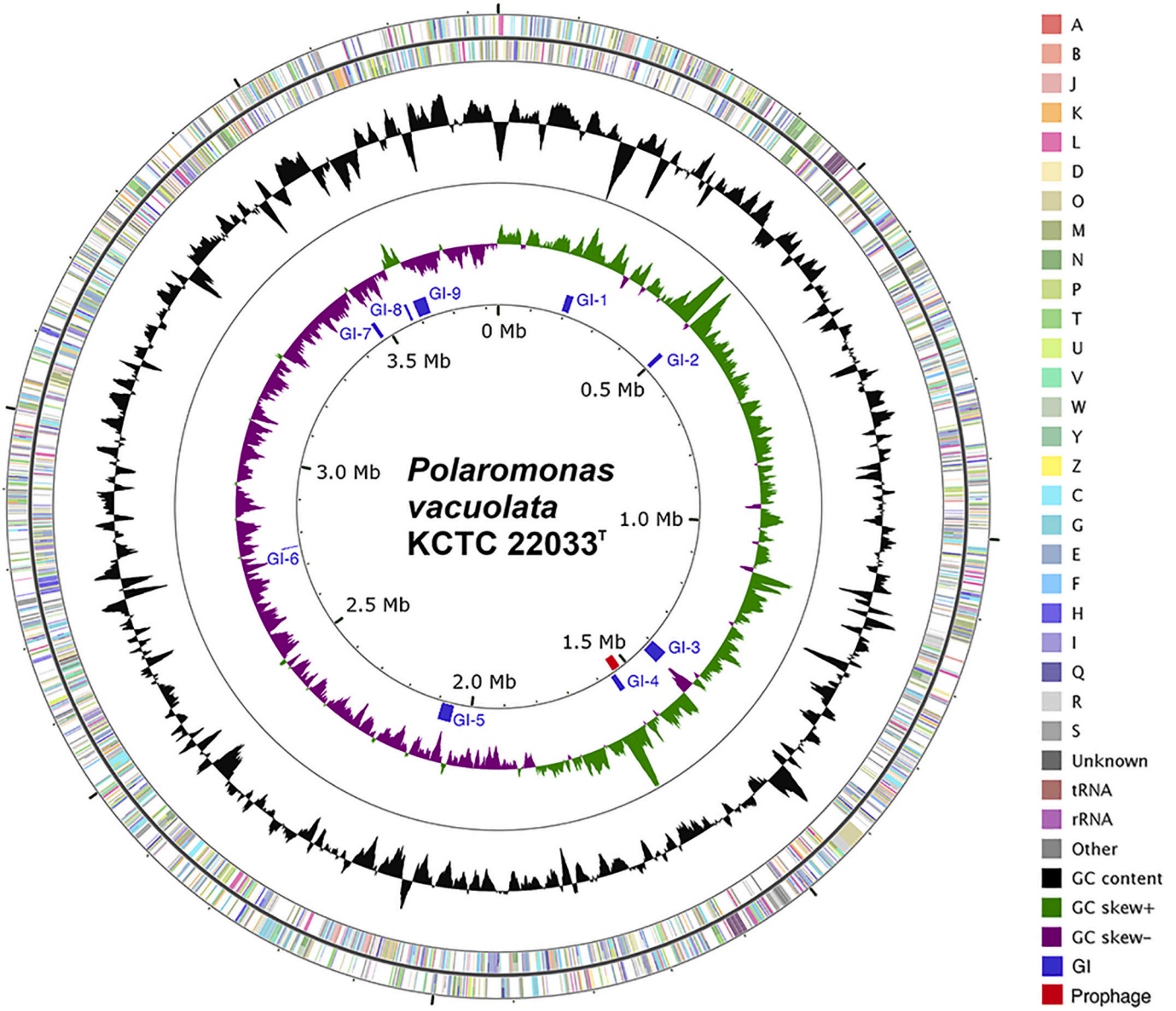

Fig. 1. Circular map of the *Polaromonas vacuolata* KCTC 22033^T genome**.** From outside to the center: Genes on forward strand, Genes on reverse strand, GC content (black), GC skew (green/violet), GI (Genomic islands, blue), and prophage (red). Individual genes are colored by COG categories: A (RNA processing and modification), B (Chromatin structure and dynamics), J (Translation, ribosomal structure and biogenesis), K (Transcription), L (Replication, recombination and repair), D (Cell cycle control, cell division, chromosome partitioning), O (Posttranslational modification, protein turnover, chaperones), M (Cell wall/membrane/envelop biogenesis), N (Cell motility), P (Inorganic ion transport and metabolism), T (Signal transduction mechanisms), U (Intracellular trafficking, secretion and vesicular transport), V (Defense mechanisms), W (Extracellular structures), Y (Nuclear structure), Z (Cytoskeleton), C (energy production and conversion), G (Carbohydrate transport and metabolism), E (Amino acid transport and metabolism), F (Nucleotide transport and metabolism), H (Coenzyme transport and metabolism), I (Lipid transport and metabolism), Q (Secondary metabolites biosynthesis, transport and catabolism), R (General functional prediction only), and S (Function unknown). Genes lacking significant matches to COG were considered Unknown. tRNAs shown in pale brown, rRNAs in light violet, other RNAs in grey, genomic islands in blue, and prophage in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Comamonadaceae. The second half of GI-7 island was highly homologous to ribosomal proteins of other *Polaromonas* strains. The majority of GI-8 genes (HC248_03282 to _03294) were homologous to many different non-*Polaromonas* species in *Betaproteobacteria*. Finally, the first half genes of GI-9 (HC248_03316 to _03324) had top BLAST matches to genes of *Variovorax* (*Comamonadaceae*) while the second half genes (HC248_03326 to _03349) were closest to proteins of other *Polaromonas* strains.

In summary, the evolutionary origins of GIs seem complicated and exhibit sequence homologies to diverse and distant taxonomic groups. These observations likely support a significant role of HGT in the evolution and adaptation of *Polaromonas* genomes. The GIs 3 and 5 together encode genes involved in cell adherence and gas vesicle formation, which are involved in bacterial adaptation to cold environments. As mentioned above, buoyancy produced by gas vesicles allows marine bacteria to reach up to the surface seawater, where they can interact with sea-ice algae and acquire carbon sources directly from the primary producers (Irgens et al., 1989; Guo et al., 2017).

Several other genes detected outside the GIs provided additional clues about *P. vacuolata* lifestyle. For example, genes encoding betaine synthesis pathway and betaine transporters had a higher copy number in *P. vacuolata* relative to other *Polaromonas* genomes. The *P. vacuolata* genome harbored four betaine aldehyde dehydrogenases, eleven betaine transporters, and five betaine-binding proteins (representatively from

HC248_01970 to _01982) outnumbering other *Polaromonas* genomes, which harbored one to two copies of betaine aldehyde dehydrogenase, one to six copies of betaine transporters, and zero to three copies of betaine-binding proteins. The choline dehydrogenase enzyme, which functions at the first step of betaine biosynthesis, had five copies in the *P. vacuolata* genome but was not detected in other *Polaromonas* genomes. Betaine is among the most effective osmoprotectants (Kappes et al., 1996). The unique ability to synthesize betaine and a higher number of betaine transporters therefore likely supports *P. vacuolata* growth under saline environments (0 to 6.0% NaCl) better than other *Polaromonas* species (0 to 2.0%) (Choi et al., 2018; and refs therein).

In addition, a gene cluster that encodes both rhodopsin and complete retinal biosynthesis pathways (*crt*B, *crt*E, *crt*I, *crt*Y, *blh*) was detected in the *P. vacuolata* genome (HC248_03361 to _03366) but not found in any of the eight complete or high-quality *Polaromonas* genomes. Interestingly, a nearly identical gene cluster was detected in the genome of *Variovorax* sp. PAMC 28711 that was isolated from lichen of King George Island, Antarctica. Rhodopsins supply energy when bacteria encounter lack of nutrients in marine ecosystems (Gomez-Consarnau et al., 2010). Presence of rhodopsins and retinal biosynthesis genes therefore likely supports the possible photoheterotrophic lifestyle when *P. vacuolata* undergoes starvation in surface seawaters. These findings further imply that HGT has greatly shaped the evolution of *P. vacuolata* gene and is likely responsible for the acquisition of rhodopsin and chromophore genes and its photoheterotrophic lifestyle.

3. Nucleotide sequence accession numbers

The complete genome sequence of *Polaromonas vacuolata* KCTC 22033^T has been deposited in GenBank under the accession number CP051461. This strain is available from the Korean Collection for Type Cultures (Jeongeup, Republic of Korea).

Acknowledgements

This work was supported by Korea Polar Research Institute (grant number PE20130). AN is supported by the U.S. Department of Energy LDRD program at Los Alamos National Laboratory (20180751PRD3).

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., et al., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37, 420–423.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., et al., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–W21.

Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing Group, Lau, B.Y., et al., 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 45, W30–W35.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., et al., 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods 10, 563–569.

Choi, K.D., Siddiqi, M.Z., Liu, Q., Jo, J.H., Chun, S.Y., et al., 2018. Polaromonas ginsengisoli sp. nov., isolated from ginseng field soil. Int. J. Syst. Evol. Microbiol. 68, 1436–1441.

Gomez-Consarnau, L., Akram, N., Lindell, K., Pedersen, A., Neutze, R., et al., 2010. Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. PLoS Biol. 8, e1000358.

Grant, J.R., Stothard, P., 2008. The CGView server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 36, W181–W184.

Guo, S., Stevens, C.A., Vance, T.D.R., Olijve, L.L.C., Graham, L.A., et al., 2017. Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. Sci. Adv. 3, e1701440.

Irgens, R.L., Suzuki, I., Staley, J.T., 1989. Gas vacuolate bacteria obtained from marine waters of Antarctica. Curr. Microbiol. 18, 261–265.

Irgens, R.L., Gosink, J.J., Staley, J.T., 1996. Polaromonas vacuolata gen. Nov., sp. nov., a psychrophilic, marine, gas vacuolate bacterium from Antarctica. Int. J. Syst. Bacteriol. 46, 822–826.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., et al., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.

Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.

Kappes, R.M., Kempf, B., Bremer, E., 1996. Three transport systems for the osmoprotectant glycine betaine operate in Bacillus subtilis: characterization of OpuD. J. Bacteriol. 178, 5071–5079.

Kim, K.M., Choe, H., Kim, B.K., Nasir, A., 2017. Complete genome of a metabolically-diverse marine bacterium Shewanella japonica KCTC 22435(T). Mar. Genomics 35, 39–42.

Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Liu, Q., Song, W.Z., Zhou, Y.G., Dong, X.Z., Xin, Y.H., 2020. Phenotypic divergence of thermotolerance: molecular basis and cold adaptive evolution related to intrinsic DNA flexibility of glacier-inhabiting Cryobacterium strains. Environ. Microbiol. 22, 1409–1420.

Mattes, T.E., Alexander, A.K., Richardson, P.M., Munk, A.C., Han, C.S., et al., 2008. The genome of Polaromonas sp. strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. Appl. Environ. Microbiol. 74, 6405–6416.

Roller, B.R., Stoddard, S.F., Schmidt, T.M., 2016. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nat. Microbiol. 1, 16160.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069.

Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R., Schmidt, T.M., 2015. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic Acids Res. 43, D593–D598.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28, 33–36.

Yagi, J.M., Sims, D., Brettin, T., Bruce, D., Madsen, E.L., 2009. The genome of Polaromonas naphthalenivorans strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. Environ. Microbiol. 11, 2253–2270.