Marine Genomics 35 (2017) 39–42

Contents lists available at ScienceDirect

Marine Genomics

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

Genomics/technical resources

Complete genome of a metabolically-diverse marine bacterium *Shewanella japonica* KCTC 22435^T

Kyung Mo Kim [a,b,*], Hanna Choe [b], Byung Kwon Kim [c], Arshan Nasir [d,e]

[a] Division of Polar Life Sciences, Korea Polar Research Institute, Incheon, Republic of Korea
[b] Biological Resource Center, Korea Research Institute of Bioscience and Biotechnology, Jeongeup, Republic of Korea
[c] Omicspia Inc., Daejeon, Republic of Korea
[d] Department of Crop Sciences, University of Illinois, Urbana, IL, USA
[e] Department of Biosciences, COMSATS Institute of Information Technology, Tarlai Kalan, Islamabad, Pakistan

ARTICLE INFO

Article history:
Received 9 April 2017
Received in revised form 10 May 2017
Accepted 10 May 2017
Available online 15 May 2017

Keywords:
Shewanella japonica
Microbial fuel cell
Cytochrome *c*
Complete genome

ABSTRACT

Shewanella japonica KCTC 22435^T is a facultatively anaerobic, Gram-negative, mesophilic, rod-shaped bacterium isolated from sea water at the Pacific Institute of Bio-organic Chemistry of the Marine Experimental Station, Troitza Bay, Gulf of Peter the Great, Russia. Here, we report the complete genome of *S. japonica* KCTC 22435^T, which consists of 4,975,677 bp (G + C content of 40.80%) with a single chromosome, 4036 protein-coding genes, 97 tRNAs and 8 rRNA operons. Genes detected in the genome reveal that the strain possesses a type II secretion system, cytochrome *c* family proteins with various numbers of heme-binding motifs, and metabolic pathways for utilizing diverse carbon sources, supporting the potential of KCTC 22435^T to generate electricity in salinity culture conditions.

1. Introduction

The bacterial genus *Shewanella* (family: *Shewanellaceae*, order: *Alteromonadales*, Class: *Gammaproteobacteria*) was first described by MacDonell and Colwell (1985) and currently consists of >60 species. *Shewanella* species can utilize a wide range of carbon sources and a number of organic and inorganic compounds as terminal electron acceptors under anaerobic conditions (Hau and Gralnick, 2007) allowing these species to survive in diverse habitats including seawater, freshwater, activated sludge, sediment and spoiled food (Liu et al., 2015). Thanks to their remarkable metabolic versatility, *Shewanella* species have been regarded as promising model organisms in the bioremediation of environmental pollutants (e.g., radionuclides, halogenated organics, petroleum, etc.) and microbial fuel cell (MFC) research (Hau and Gralnick, 2007).

MFC devices utilize microorganisms as biocatalysts to convert waste organic matter and biomass into electricity. The process involves production of electrons from carbon sources, transfer of electrons to the extracellular electron acceptors, and generation of electricity by reduction of acceptors (Rabaey and Verstraete, 2005). The choice of the biocatalyst depends upon the metabolic versatility of microorganisms, their ability to transfer electrons outside the cell, and the potential to operate in diverse environmental conditions (e.g. freshwater vs. marine, aerobic vs, anaerobic). *Geobacter*- and *Shewanella*-containing MFCs have recently become popular because their metabolic pathways are well-described in literature (Lovely, 2006; Fredrickson et al., 2008). *Geobacter* species however operate under strict anaerobic conditions and thus their utility is rather limited. In turn, *Shewanella* species can operate well in air-exposed conditions, can utilize a wide range of substrates, and have become attractive models for utilization in MFCs. For example, *Shewanella oneidensis* MR-1 is a fresh-water microbe that is widely used in MFCs. However, freshwater accounts for only ~2.5% of total Earth's water (https://water.usgs.gov/edu/earthwherewater.html), necessitating the need to discover novel microorganisms capable of generating electricity in high salinity environments. Recently *Shewanella marisflavi* EP1 was shown to produce electricity at high ionic strengths (up to 1488 mM or 8% NaCl) (Huang et al., 2010). Unfortunately, respiration of EP1 can apparently be conducted under limited carbon sources (e.g., lactate). Therefore, search for microorganisms capable of transforming wide range of organic substrates into electrical energy and under saline conditions remains an active area of research.

Shewanella japonica KCTC 22435^T is a facultatively anaerobic, Gram-negative, mesophilic, rod-shaped bacterium isolated from sea water at the Pacific Institute of Bio-organic Chemistry of the Marine Experimental Station, Troitza Bay, Gulf of Peter the Great, Russia (Table 1). This type strain was originally reported as agar-digesting bacteria (Ivanova et al., 2001). However, a recent study has revealed that *S. japonica* strain ATCC: BAA-316 (= KCTC 22435^T) can respire diverse carbon sources

* Corresponding author at: Division of Polar Life Sciences, Korea Polar Research Institute, Incheon, Republic of Korea.
E-mail address: kmkim@kopri.re.kr (K.M. Kim).

http://dx.doi.org/10.1016/j.margen.2017.05.004
1874-7787/

Table 1
General features of *Shewanella japonica* KCTC 22435^T and MIGS mandatory information.

Item	Description
General features	
Classification	Domain *Bacteria* Phylum *Proteobacteria* Class *Gammaproteobacteria* Order *Alteromonadales* Family *Shewanellaceae* Genus *Shewanella*
Type strain	KCTC 22435^T
Gram strain	Negative
Cell shape	Straight rod
Motility	Not reported
Sporulation	Non-spore-forming
Temperature range	10–37 °C, optimally at 20–25 °C
Salinity range	NaCl, 0–3%
pH range	6–9, and optimally at 7.5
Investigation	
Submitted to INSDC	Accession number CP020472
Investigation type	bacteria_archaea
Project name	Genome sequence of *Shewanella japonica* KCTC 22435
Environment	
Geographic location	Gulf of Peter the Great, Russia
Depth	1 m
Collection date	1994-01
Environment (biome)	Temperate marginal sea biome (ENVO:01000856)
Environment (feature)	Bay (ENVO:00000032)
Environment (material)	Water (ENVO:00002006)
Environment (package)	Sea water (ENVO:00002149)
Isolation and growth conditions	PMID: 11411670
Sequencing	
Sequencing platform	PacBio RS II with P6-C4 chemistry
Fold coverage	204.89×
Assembler	SMRT Analysis v2.3.0
Annotation source	Prodigal v2.6.3

ranging from monosaccharides to sucrose to agar and generate electricity in marine environments, consequently supplementing the weaknesses of *S. oneidensis* and *S. marisflavi* and thus becoming a promising model for utilization in MFC devices (Biffinger et al., 2011). Despite the industrial relevance of strain KCTC 22435^T, its genomic information was hitherto unavailable. Therefore, we report the complete genome of the strain KCTC 22435^T and also evaluate its potential to be utilized as biocatalyst in MFC devices.

2. Data description

S. japonica KCTC 22435^T was grown for three days at 25 °C on marine agar. Isolated colonies were picked using a sterile toothpick and genomic DNA was extracted using the i-genomic BYF mini kit (iNtRON Biotechnology, Seongnam, Republic of Korea) following manufacturer's protocols. Genome sequencing was performed using PacBio RS II SMRT sequencing technology (Pacific Biosciences, Menlo Park, CA, USA). A 20-kb insert SMRTbell library was constructed and sequenced yielding >204× average genome coverage. De novo assembly of 138,026 subreads with 8909 nucleotides on average (1,229,796,822 bp in total) was conducted using the hierarchical genome-assembly process (HGAP) pipeline of the SMRT Analysis v2.3.0 (Chin et al., 2013). The overlapping regions at both ends of a contig were manually identified and trimmed to generate a unique stretch on both ends. Then, a new version for the contig was generated by cutting the contig into two halves and switching the first half with the second. The newly generated contig served as reference to which raw PacBio reads were mapped using the resequencing module of the SMRT Analysis. This enabled us to correct possible sequencing errors at contig ends, where mapping coverage is relatively lower.

Next, we identified protein-coding genes using Prodigal v2.6.3 (Hyatt et al., 2010). The predicted CDSs were BLAST-searched against UniProt (Wu et al., 2006), Pfam (Punta et al., 2011) and COG (Tatusov et al., 2003) databases to gain insights about the molecular functions and family classifications of predicted genes. Signal peptides and transmembrane helices were predicted using SignalP v4.1 (Petersen et al., 2011) and TMHMM v2.0 (Krogh et al., 2001). rRNA, tRNA and other miscellaneous features were predicted using RNAmmer v1.2 (Lagesen et al., 2007), tRNAscan-SE v1.21 (Lowe and Eddy, 1997) and Rfam v12.0 (Griffiths-Jones et al., 2005). The graphic circular map of genome was constructed and visualized using Circos v0.67 (Krzywinski et al., 2009). Automatic detection of clustered regularly interspaced palindromic repeats (CRISPRs) was performed using MinCED v.0.2.0 (Bland et al., 2007).

The complete genome of *S. japonica* KCTC 22435^T is composed of a circular chromosome of 4,975,677 bp with G + C content of 40.8% (Fig. 1; Table 2). The protein coding regions cover 84.19% of the genome (4,189,134 bp) and encode 4036 proteins (Table 2). The genome also encodes eight rRNAs (seven operons of 5S, 16S and 23S, in that order, and one operon of two copied 5S, 16S and 23S in that order), 97 tRNAs and three other RNAs (one tmRNA and two ncRNAs). Signal peptides and transmembrane helices were found in 606 and 1086 protein coding genes, respectively (Table 2). Plasmids and CRISPR repeats were not detected.

Molecular mechanisms of electron transfer in *Shewanella* species have not been fully explored. However, both a type II secretory pathway and outer membrane cytochromes are likely to play a critical role in transferring electrons to the extracellular electron acceptors (Hau and Gralnick, 2007). The genome of *S. japonica* KCTC 22435^T encodes a complete gene cluster of the type II secretion system (locus tags SJ2017_4026 to SJ2017_4037), in addition to genes encoding Type VI secretion system apparatus and secretion proteins (SJ2017_2446 to SJ2017_2448, SJ2017_2456 to SJ2017_2463, respectively). The type VI secretion system confers bacteria the abilities of cell adherence and biofilm formation (Linares et al., 2016). Consequently, the presence of type VI secretion system suggests the possibility that the strain KCTC 22435^T can be attached to an electrode surface and form biofilms on the surface. The genome of *S. japonica* KCTC 22435^T also encodes diverse *c*-type cytochrome family proteins that play an important role as electron carriers. The numbers of CXXCH heme-binding motifs of the cytochrome *c* family proteins ranged from one to eleven, but hexaheme cytochrome *c* was not detected in the genome. As the most abundant type, 14 out of the 34 family proteins (e.g., SJ2017_3429) possessed tetraheme-binding motifs. The existence of cytochrome *c* family proteins with diverse numbers of heme-binding motifs support that the strain KCTC 22435^T is able to efficiently attach electrons to the family proteins (Kranz et al., 2009).

Besides, the genome of *S. japonica* KCTC 22435^T indicates that the strain can utilize a diverse range of carbohydrates as carbon source. The genome encodes four beta-agarases (e.g., SJ2017_1905 for degrading agarose), three beta-xylanases (e.g., SJ2017_2021 for xylane), a glycogen debranching protein GlgX (SJ2017_2159 for starch), and an endoglucanase (SJ2017_3472 for cellulose). In addition, the strain may degrade hexoses (e.g., glucose or fructose) using both the pentose phosphate pathway (e.g., SJ2017_3279, SJ2017_3945) and the Entner-Doudoroff pathway (SJ2017_1967 to SJ2017_1970), instead of using the Embden-Meyerhof pathway. The presence of a UDP-galactose-4-epimerase (SJ2017_2010) and galactose-1-phosphate uridylyltransferase (SJ2017_2306) supports that the strain may directly use galactose as a carbon source. The genome also encodes all enzymes (SJ2017_1933, SJ2017_1940, and SJ2017_1941) involved in De Ley-Doudoroff pathway, indicating that galactonate can be degraded by the strain. Finally, the strain may also use carbon sources gluconolactone and gluconate since the genome encodes a 2-dehydro-3-deoxygluconokinase (SJ2017_1903), 2-dehydro-3-deoxyphosphogluconate aldolase (SJ2017_1904), and gluconolactonase (SJ2017_1906).

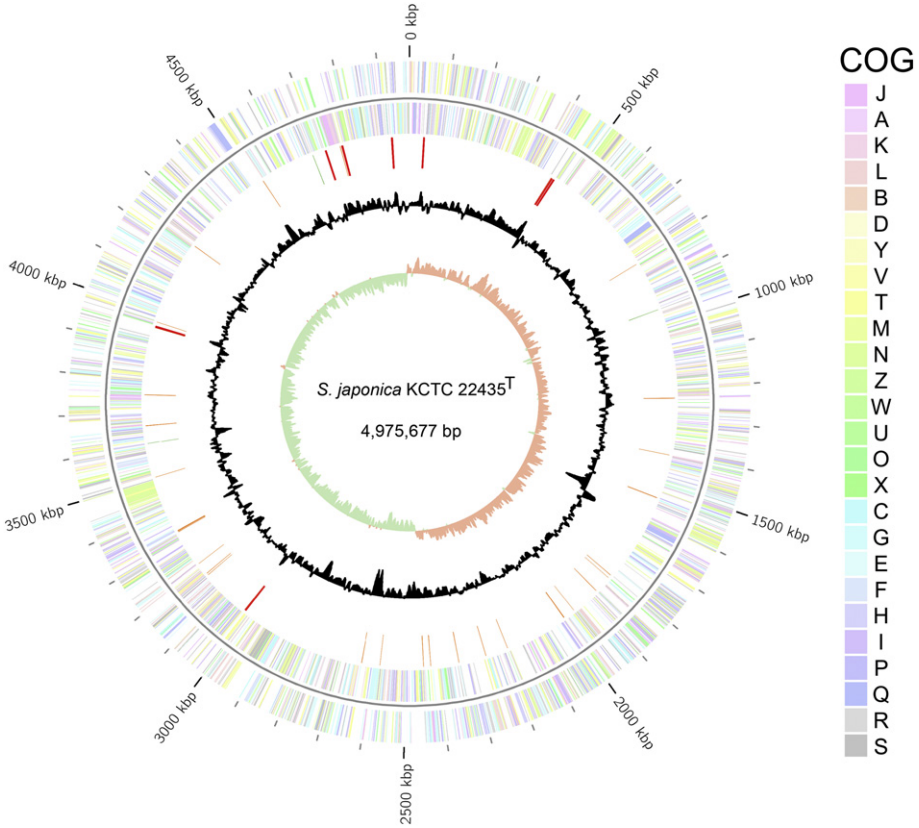

Fig. 1. Circular map of the *S. japonica* KCTC 22435^T genome. From outside to the center: Genes on forward strand, Genes on reverse strand, RNA genes (tRNAs orange, rRNAs red, other RNAs green), GC content (black), and GC skew (light green/orange). Individual genes are colored by COG categories: J (Translation, ribosomal structure and biogenesis), A (RNA processing and modification), K (Transcription), L (Replication, recombination and repair), B (Chromatin structure and dynamics), D (Cell cycle control, cell division, chromosome partitioning), Y (Nuclear structure), V (Defense mechanisms), T (Signal transduction mechanisms), M (Cell wall/membrane/envelop biogenesis), N (Cell motility), Z (Cytoskeleton), W (Extracellular structures), U (Intracellular trafficking, secretion and vesicular transport), O (Posttranslational modification, protein turnover, chaperones), X (Mobilome: prophages and transposons), 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), P (Inorganic ion transport and metabolism), Q (Secondary metabolites biosynthesis, transport and catabolism), R (General functional prediction only), and S (Function unknown).

In conclusion, the availability of the *S. japonica* KCTC 22435^T genome sequence highlights the metabolic versatility of *Shewanella* species and provides molecular insights into its applicability in MFC research (Biffinger et al., 2011).

3. Nucleotide sequence accession number

The complete genome sequence of *S. japonica* KCTC 22435^T has been deposited to GenBank/EMBL/DDBJ under the accession number CP020472. This strain is available from the Korean Collection for Type Cultures (Jeongeup, Republic of Korea).

Table 2
Genomic features of the *S. japonica* KCTC 22435^T.

Attribute	Value
Genome size (bp)	4,975,677
Protein coding region (bp)	4,189,134
G + C content (%)	40.8
Chromosome	1
Total genes	4274
Protein coding genes	4036
rRNA (operon)	25 (8)
tRNA	97
Pseudogenes	112
Genes with functional prediction	2873
Genes assigned to COGs	3033
Genes with Pfam domains	3322
Genes with signal peptides	606
Genes with transmembrane helices	1086

Acknowledgements

This research was supported by a grant from the Marine Biotechnology Program (PJT200620) funded by the Ministry of Oceans and Fisheries, Republic of Korea.

References

Biffinger, J.C., Fitzgerald, L.A., Ray, R., Little, B.J., Lizewski, S.E., et al., 2011. The utility of *Shewanella japonica* for microbial fuel cells. Bioresour. Technol. 102, 290–297.
Bland, C., Ramsey, T.L., Sabree, F., Lowe, M., Brown, K., et al., 2007. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 8, 209.
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.
Fredrickson, J.K., Romine, M.F., Beliaev, A.S., Auchtung, J.M., Driscoll, M.E., et al., 2008. Towards environmental systems biology of *Shewanella*. Nat. Rev. Microbiol. 6, 592–603.
Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., et al., 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 33, D121–D124.
Hau, H.H., Gralnick, J.A., 2007. Ecology and biotechnology of the genus *Shewanella*. Annu. Rev. Microbiol. 61, 237–258.
Huang, J., Sun, B., Zhang, X., 2010. Electricity generation at high ionic strength in microbial fuel cell by a newly isolated *Shewanella marisflavi* EP1. Appl. Microbiol. Biotechnol. 85, 1141–1149.
Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119.

Ivanova, E.P., Sawabe, T., Gorshkova, N.M., Svetashev, V.I., Mikhailov, V.V., Nicolau, D.V., Christen, R., 2001. *Shewanella japonica* sp. nov. Int. J. Syst. Ecol. Microbiol. 51, 1027–1033.

Kranz, R.G., Richard-Fogal, C., Taylor, J.S., Frawley, E.R., 2009. Cytochrome c biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. Microbiol. Mol. Biol. Rev. 73, 510–528.

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.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., et al., 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639–1645.

Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.H., Rognes, T., Ussery, D.W., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108.

Linares, D., Jean, N., Overtvelt, P., Ouidir, T., Hardouin, J., Blache, Y., Molmeret, M., 2016. The marine bacteria *Shewanella frigidimarina* NCIMB400 upregulates the type VI secretion system during early biofilm formation. Environ. Microbiol. Rep. 8, 110–121.

Liu, Y., Shang, X.X., Yi, Z.W., Gu, L., Zeng, R.Y., 2015. *Shewanella mangrovi* sp. nov., an acetaldehyde-degrading bacterium isolated from mangrove sediment. Int. J. Syst. Ecol. Microbiol. 65, 2630–2634.

Lovely, D.R., 2006. Bug juice: harvesting electricity with microorganisms. Nat. Rev. Microbiol. 4, 497–508.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

MacDonell, M.T., Colwell, R.R., 1985. Phylogeny of the Vibrionaceae, and recommendation for two new genera, *Listonella* and *Shewanella*. Syst. Appl. Microbiol. 6, 171–182.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., et al., 2011. The Pfam protein families database. Nucleic Acids Res. 40, D290–D301.

Rabaey, K., Verstraete, W., 2005. Microbial fuel cells: novel biotechnology for energy generation. Trends Biotechnol. 23, 291–298.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., et al., 2003. The COG database: an updated version includes eukaryotes. BMC bioinformatics 4, 41.

Wu, C.H., Apweiler, R., Bairoch, A., Natale, D.A., Barker, W.C., et al., 2006. The Universal Protein Resource (UniProt): an expanding universe of protein information. Nucleic Acids Res. 34, D187–D191.