Accepted Manuscript

Title: Complete genome sequence of *Planococcus donghaensis* JH1^T, a pectin-degrading bacterium

Authors: Wah-Seng See-Too, Kah-Ooi Chua, Yan-Lue Lim, Jian-Woon Chen, Peter Convey, Taznim Begam Mohd Mohidin, Wai-Fong Yin, Kok-Gan Chan

PII: S0168-1656(17)30207-9
DOI: http://dx.doi.org/doi:10.1016/j.jbiotec.2017.05.005
Reference: BIOTEC 7887

To appear in: *Journal of Biotechnology*

Received date: 6-1-2017
Revised date: 4-5-2017
Accepted date: 5-5-2017

Please cite this article as: See-Too, Wah-Seng, Chua, Kah-Ooi, Lim, Yan-Lue, Chen, Jian-Woon, Convey, Peter, Mohidin, Taznim Begam Mohd, Yin, Wai-Fong, Chan, Kok-Gan, Complete genome sequence of Planococcus donghaensis JH1T, a pectin-degrading bacterium.Journal of Biotechnology http://dx.doi.org/10.1016/j.jbiotec.2017.05.005

Short Genome Communications

Complete genome sequence of *Planococcus donghaensis* JH1^T, a pectin-degrading bacterium

Wah-Seng See-Too[a, b], Kah-Ooi Chua[a], Yan-Lue Lim[a], Jian-Woon Chen [a,d] , Peter Convey [b,c], Taznim Begam Mohd Mohidin[e], Wai-Fong Yin[a], Kok-Gan Chan [a,d]

[a] Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia

[b] National Antarctic Research Centre (NARC), Institute of Postgraduate Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia

[c] British Antarctic Survey, NERC, High Cross, Madingley Road, Cambridge CB3 OET, UK

[d] UM Omics Centre, University of Malaya, Kuala Lumpur, Malaysia

[e] Division of Microbiology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia

Address:

*****Corresponding author: Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia. Email: kokgan@um.edu.my (K-G, Chan).

HIGHLIGHTS

- The first complete genome sequence of *Planococcus donghaensis* JH1^T type strain, a psychrotolerant and halotolerant bacterium.
- Strain JH1^T shows ability to degrade complex polysaccharides, including pectin and D-galacturonic acid.
- Genes encoding for pectin-degrading enzymes were determined from the complete genome.
- Biosynthetic pathways for biotechnologically important compounds including carotenoids and butanol were identified in the genome.

ABSTRACT

The type strain *Planococcus donghaensis* JH1^T is a psychrotolerant and halotolerant bacterium with starch-degrading ability. Here, we determine the carbon utilization profile of *P.*

donghaensis JH1^T and report the first complete genome of the strain. This study revealed the strain's ability to utilize pectin and $_D$-galacturonic acid, and identified genes responsible for degradation of the polysaccharides. The genomic information provided may serve as a fundamental resource for full exploration of the biotechnological potential of *P. donghaensis* JH1^T.

Keywords: Psychrotolerant, Halotolerant, Pectinase enzymes, Butanol, Carotenoid

Pectin is a complex polysaccharide that, together with cellulose and hemicellulose, constitutes the middle lamella and primary cell wall of terrestrial plants (Abbott and Boraston, 2008; Kashyap et al., 2001; Soriano et al., 2006). The pectinaceous nature of plants presents major challenges for industries that use plants as raw materials, such as the juice beverage industry (Jayani et al., 2005). For example, increased viscosity of fruit juices caused by fruit pectin renders subsequent processing steps such as filtration and concentration of juice less efficient (Sagu et al., 2014). To resolve this issue, microbial pectinase enzymes are employed in these industries to degrade pectin, in order to speed up the fruit juice extraction process and to increase juice yields (Kashyap et al., 2001; Sagu et al., 2014). Furthermore, pectinase enzymes are also being increasingly utilised in an expanding range of industrial applications, for instance in clarification of fruit juices (Ceci and Lozano, 1998), pre-treatment of pectic wastewater (Hoondal et al., 2002), and fermentation of tea and coffee (Couto and Sanromán, 2006). Hence, there is growing demand and strong industrial interest in discovery of microbial isolates which produce novel pectinase enzymes with potent catalytic properties.

Planococcus donghaensis strain JH1^T (=DSM 22276^T =LMG 23779^T) is a psychrotolerant and halotolerant bacterium isolated from deep-sea sediment of the East Sea, South Korea, and was reported to have the ability to degrade starch which suggested an industrial application potential (Choi et al., 2007). Members of the genus *Planococcus* have mostly been isolated from the polar regions or marine environments (Shivaji et al., 2014) and one such isolate, *P. halocryophilus*, has been reported to be metabolically active at -25 °C (Mykytczuk et al., 2013). However, information about the metabolic capabilities of this genus is limited and, to date, complete genomes of only *P. kocurii* (See-Too et al., 2016a), *P. massiliensis* (Seck et al., 2016), *P. rifietoensis* (See-Too et al., 2016b) and *P. versutus* (See-Too et al., 2016c; See-Too et al., 2017) have been reported. In this study, we determined the

carbon utilization profile of strain JH1^T using the GEN III MicroPlate^{TM} (Biolog), in which utilisation of various substrates with industrial significance such as pectin and D-galacturonic acid was observed (Fig. 1). We then obtained the strain's genomic data via whole genome sequencing and report the first complete genome of the species, central information required for its future application in relevant industries.

To characterize the carbon utilization profile of *P. donghaensis* JH1^T, a single colony of fresh culture was suspended into inoculating fluid B (Biolog) and adjusted to a specific transmittance (98% T) using a turbidimeter following the manufacturer's protocol. Aliquots (100 µl) of cell suspension were inoculated into each well of a GEN III MicroPlate^{TM} (Biolog). The plate was immediately placed in an OmniLog (Biolog, Inc.) incubator and was incubated at 26°C. The colour change of each well was recorded spectrophotometrically every 15 min over a period of 60 h. Visible colour changes due to reduction of tetrazolium violet dye are indicative of active bacterial respiration and thus utilization of the specific substrate in each well (Line et al., 2010; Wragg et al., 2014).

Whole genome sequencing of *P. donghaensis* JH1^T was performed as described previously (See-Too et al., 2016a). In brief, the genomic DNA of the strain was extracted from an overnight culture in trypticase soy yeast extract medium using the MasterPure^{TM} Gram positive DNA purification kit (Epicentre Technologies). Extracted genomic DNA was then constructed into a 20-kb SMRTbell^{TM} template library. A PacBio RSII sequencing platform (Pacific Biosciences) was employed for whole genome sequencing using P6-C4 chemistry in one single molecule real time (SMRT) cell (Pacific Biosciences).

From the GEN III MicroPlate^{TM}, *P. donghaensis* JH1^T was revealed to have the ability to degrade dextrin, sucrose, β-methyl-D-glucoside, D-salicin, *N*-acetyl-D-glucosamine, D-fructose, inosine, D-mannitol, glycerol, D-glucose-6-phosphate, D-fructose-6-phosphate, glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-pyroglutamic acid, L-serine, pectin, D-galacturonic acid, D-gluconic acid, glucuronamide, D-lactic acid methyl ester, tween 40, β-hydroxy-D,L-butyric acid, acetoacetic acid, acetic acid and formic acid.

A total of 45,497 reads with a mean read length of 12,385 bp were generated from whole genome sequencing. The reads were *de novo* assembled using the hierarchical genome assembly process (HGAP) algorithm version 2 (Chin et al., 2013) to generate the complete genome of *P. donghaensis* JH1^T, which consists of 2 polished contigs with an average genome

coverage of 145.5 fold. The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 2.10 (Tatusova et al., 2016), Rapid Annotation using Subsystem Technology (RAST) version 3.0 (Aziz et al., 2008; Overbeek et al., 2014) and IMG ER pipeline (Markowitz et al., 2009). The genome project and the complete genome sequence were deposited in the Genomes On Line Database (Liolios et al., 2010) and GenBank.

The genome of *P. donghaensis* JH1^T is 3,319,331 bp in length, consisting of a circular chromosome and a plasmid, with G + C contents of 40.1 and 46.1 mol%, respectively. A total of 3,099 protein coding genes, 33 rRNA and 79 tRNA were predicted from the chromosome and plasmid by PGAP analysis (Table 1, Fig. 2). Furthermore, there were 2696 genes assigned to different function categories based on the clusters of orthologous genes (COG) designation (Table 2) (Tatusov et al., 2003).

Functional annotation of the genome revealed the presence of genes encoding pectate lyase (NCBI locus tag: BCM40_09360). Pectate lyase and pectin lyase, which catalyse degradation of pectate and pectin respectively, are categorized as pectic transeliminases. This group of pectinase enzymes act by cleaving the glycosidic bonds of either pectate or pectin by beta-elimination and generate unsaturated end products. A hypothetical protein that belongs to the polygalacturonase family (NCBI locus tag: BCM40_03605) was also detected. Polygalacturonases are enzymes that catalyse hydrolysis of the polygalacturonic acid chain in pectin (Gummadi and Kumar, 2005; Jayani et al., 2005).

The D-galacturonic acid and pectin-degrading potential of strain JH1^T indicates the strain's potential value for industrial use. From the genome mining analysis, various genes which encode other products with wide industrial interest were also detected. These included genes encoding phytoene desaturase (NCBI locus tag: BCM40_01765 & BCM40_01745), phytoene synthase (NCBI locus tag: BCM40_01755) and diapolycopene oxygenase (NCBI locus tag: BCM40_12610), enzymes involved in the biosynthesis of carotenoids. Genes responsible for the biosynthesis of butanol through the acetone/isopropanol butanol–ethanol (A/IBE) pathway, including acetyl-CoA acetyltransferase (NCBI locus tag: BCM40_02320), 3-hydroxybutyryl-coA-dehydrogenase (NCBI locus tag: BCM40_02290 & BCM40_02285), Enoyl-CoA hydratase (NCBI locus tag: BCM40_02280) and acetaldehyde dehydrogenase (NCBI locus tag: BCM40_03710) were also detected. The availability of this genome provides the basis for in-depth exploration of biotechnological potential in *P. donghaensis* and other related species.

Nucleotide sequence accession number

The complete chromosome sequence has been deposited in GenBank under the accession numbers **CP016543.2** (chromosome) and **CP016544.1** (plasmid).

Acknowledgements

This work was supported by the University of Malaya High Impact Research Grants (UM-MOHE HIR GrantUM.C/625/1/HIR/MOHE/CHAN/14/1, Grant No. H-50001-A000027;UM-MOHE HIR Grant UM.C/625/1/HIR/MOHE/CHAN/01, Grant No. A-000001-50001) awarded to Kok-Gan Chan and PPP Grant (Grant No. PG089-2015B) awarded to Kah-Ooi Chua. Wah-Seng See-Too thanks the Bright Sparks Program of the University of Malaya for scholarship support. Peter Convey is supported by NERC core funding to the British Antarctic Survey's 'Biodiversity, Evolution and Adaptation' Team, and by a University of Malaya Visiting Icon Professorship.

References

Abbott, D.W., Boraston, A.B., 2008. Structural biology of pectin degradation by *Enterobacteriaceae*. Microbiol. Mol. Biol. Rev. 72, 301-316.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L.,Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9, 75.

Ceci, L., Lozano, J., 1998. Determination of enzymatic activities of commercial pectinases for the clarification of apple juice. Food Chem. 61, 237-241.

Chin, C.-S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W., Korlach, J, 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature Methods 10, 563-569.

Choi, J.-H., Im, W.-T., Liu, Q.-M., Yoo, J.-S., Shin, J.-H., Rhee, S.-K., Roh, D.-H., 2007. *Planococcus donghaensis* sp. nov., a starch-degrading bacterium isolated from the East Sea, South Korea. Int. J. Syst. Evol. Microbiol. 57, 2645-2650.

Couto, S.R., Sanromán, M.A., 2006. Application of solid-state fermentation to food industry-a review. J. Food Eng. 76, 291-302.

Gummadi, S.N., Kumar, D.S., 2005. Microbial pectic transeliminases. Biotechnol. Lett. 27, 451-458.

Hoondal, G., Tiwari, R., Tewari, R., Dahiya N., Beg, Q., 2002. Microbial alkaline pectinases and their industrial applications: a review. Appl. Microbiol. Biotechnol. 59, 409-418.

Jayani, R.S., Saxena, S., Gupta, R., 2005. Microbial pectinolytic enzymes: a review. Process Biochem. 40, 2931-2944.

Kashyap, D.R., Vohra, P.K., Chopra, S., Tewari, R., 2001. Applications of pectinases in the commercial sector: a review. Bioresour. Technol. 77, 215-227.

Line, J.E., Hiett, K.L., Guard-Bouldin, J., Seal, B.S., 2010. Differential carbon source utilization by *Campylobacter jejuni* 11168 in response to growth temperature variation. J. Microbiol. Methods 80, 198-202.

Liolios, K., Chen, I-M.A., Mavromatis, K., Tavernarakis N., Hugenholtz, P., Markowitz, V.M., Kyrpides, N.C., 2010. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. 38, D346-D354.

Markowitz, V.M., Mavromatis, K., Ivanova, N.N., Chen, I-M.A., Chu, K., Kyrpides, N.C., 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25, 2271-2278.

Mykytczuk, N.C.S., Foote, S.J., Omelon, C.R., Southam, G., Greer, C.W., Whyte, L.G., 2013. Bacterial growth at −15°C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. ISME J. 1211-1226.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 42, D206–D214.

Sagu, S.T., Nso, E.J., Karmakar, S., De, S., 2014. Optimisation of low temperature extraction of banana juice using commercial pectinase. Food Chem. 151, 182-190.

Seck, E.H., Sankar, S.A., Khelaifia, S., Croce, O., Robert, C., Couderc, C., Di Pinto, F., Sokhna, C., Fournier, P.-E., Raoult, D., Lagier, J.-C., 2016. Noncontiguous finished genome sequence and description of *Planococcus massiliensis* sp. nov., a moderately halophilic bacterium isolated from the human gut. New Microbes New Infect. 10, 36-46.

See-Too, W.-S., Tan, J.Y., Ee, R., Lim, Y.-L., Convey, P., Pearce, D.A., Yin, W.-F., Chan, K.G., 2016a. *De novo* assembly of complete genome sequence of *Planococcus kocurii* ATCC 43650^{T}, a potential plant growth promoting bacterium. Mar. Genomics 28, 33-35.

See-Too, W.-S., Convey, P., Pearce, D.A., Lim, Y.L., Ee, R., Yin, W.-F., Chan, K.-G., 2016b. Complete genome of *Planococcus rifietoensis* M8^{T}, a halotolerant and potentially plant growth promoting bacterium. J. Biotechnol. 221, 114-115.

See-Too, W.S., Ee, R., Madhaiyan, M., Kwon, S.-W., Tan, J.Y., Lim, Y.L., Convey, P., Pearce, D.A., Yin, W.-F., Chan, K.G., 2016c. *Planococcus versutus* sp. nov., isolated from Antarctic soil. Int. J. Syst. Evol. Microbiol. doi: 10.1099/ijsem.0.001721.

See-Too, W.S., Ee, R., Lim, Y.-L., Convey, P., Pearce, D.A., Yin, W.-F., Chan, K.-G., 2017. AidP, a novel *N*-acyl homoserine lactonase gene from Antarctic *Planococcus* sp. Sci. Rep. 7, 42968.

Shivaji, S., Srinivas, T.N.R., Reddy, G.S.N., 2014. The family *Planococcaceae*, in: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), The Prokaryotes: Firmicutes and Tenericutes. fourth ed. Springer-Verlag, Berlin, Heidelberg, pp. 303–352.

Soriano, M., Diaz, P., Pastor, F.I.J., 2006. Pectate lyase C from *Bacillus subtilis*: a novel endo-cleaving enzyme with activity on highly methylated pectin. Microbiology 152, 617-625.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2003. The COG database: and updated version includes eukaryotes. BMC Bioinformatics 4, 41.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze A., Pruitt, K.D., Borodovsky, M., Ostell, J., 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614-6624.

Wragg, P., Randall, L., Whatmore, A.M., 2014. Comparison of Biolog GEN III MicroStation semi-automated bacterial identification system with matrix-assisted laser desorption

ionization-time of flight mass spectrometry and 16S ribosomal RNA gene sequencing for the identification of bacteria of veterinary interest. J. Microbiol. Methods 105, 16-21.

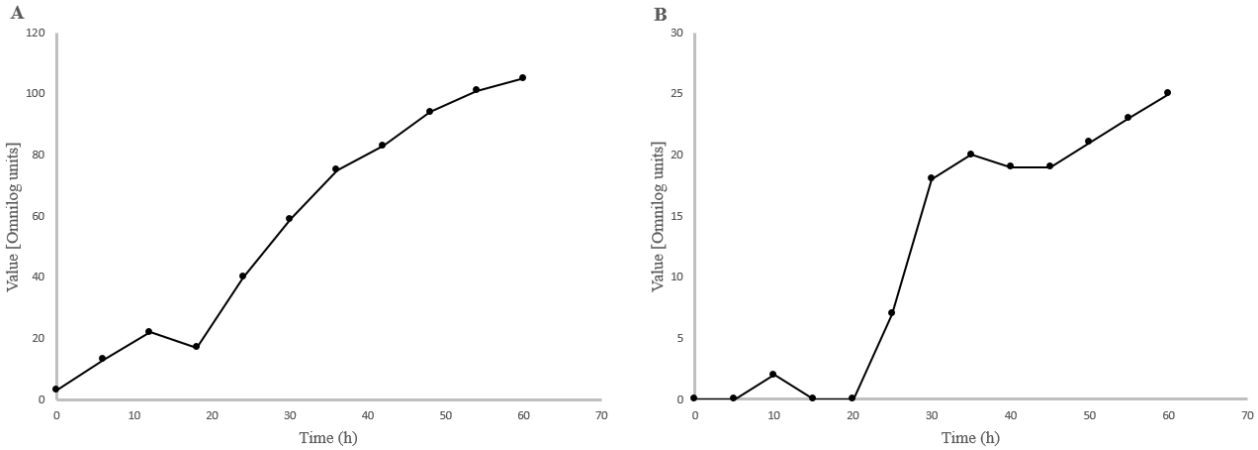

Fig. 1. Carbon utilization assay using GEN III MicroPlate^{TM} (Biolog) for (A) pectin and (B) D-galaturonic acid. Omnilog units (arbitrary colour units due to dye reduction) were plotted against time (0 to 60 h).

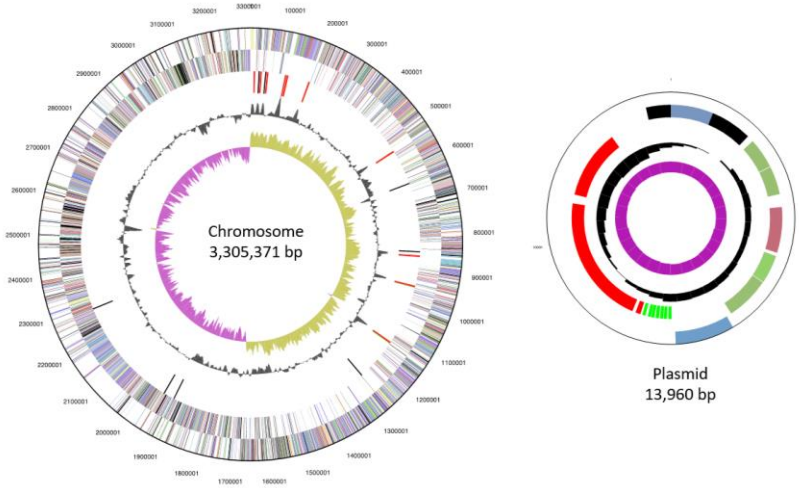

Fig. 2. Circular map of the chromosome and plasmid of *P. donghaensis* JH1^{T}. From the outside to the centre: genes on forward strand (coloured by COG categories), genes on reverse strand (coloured by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content and GC skew.

Table 1

Genome features of *P. donghaensis* JH1^{T}.

Features	Chromosome (**CP016543.2**)	Plasmid (**CP016544.1**)
Genome size (bp)	3,305,371	13,960
Contig	1	1
GC content mol%	40.1	46.1
Total number of genes	3224	18
Protein coding genes (CDS)	3092	7
Pseudogenes	25	1
rRNA genes (5S, 16S, 23S)	30 (10, 10, 10)	3 (1, 1, 1)
tRNA genes	72	7

Table 2

COG functional categories of *P. donghaensis* JH1^{T}.

COG code	Descriptions	Count	Proportion
E	Amino acid transport and metabolism	254	9.42%
G	Carbohydrate transport and metabolism	166	6.16%
D	Cell cycle control, cell division, chromosome partitioning	45	1.67%
N	Cell motility	43	1.59%
M	Cell wall/membrane/envelope biogenesis	132	4.90%
B	Chromatin structure and dynamics	1	0.04%
H	Coenzyme transport and metabolism	160	5.93%
V	Defense mechanisms	53	1.97%
C	Energy production and conversion	147	5.45%
W	Extracellular structures	7	0.26%
S	Function unknown	179	6.64%
R	General function prediction only	268	9.94%
P	Inorganic ion transport and metabolism	145	5.38%
U	Intracellular trafficking, secretion, and vesicular transport	24	0.89%
I	Lipid transport and metabolism	152	5.64%
X	Mobilome: prophages, transposons	7	0.26%
F	Nucleotide transport and metabolism	95	3.52%
O	Posttranslational modification, protein turnover, chaperones	116	4.30%
L	Replication, recombination and repair	105	3.89%
Q	Secondary metabolites biosynthesis, transport and catabolism	74	2.74%
T	Signal transduction mechanisms	130	4.82%
K	Transcription	178	6.60%
J	Translation, ribosomal structure and biogenesis	215	7.97%
	Not in COG	906	27.61%