Korean Journal of Microbiology (2018) Vol. 54, No. 3, pp. 289-292
DOI https://doi.org/10.7845/kjm.2018.8034

pISSN 0440-2413
eISSN 2383-9902

Complete genome sequence of the polycyclic aromatic hydrocarbons biodegrading bacterium *Idiomarina piscisalsi* strain 10PY1A isolated from oil-contaminated soil

Alexis Nzila[1†*], Byung Kwon Jung[2†], Min-Chul Kim[2], Jerald Conrad Ibal[2], Fitri Budiyanto[1], Musa M. Musa[3], Assad Thukair[1], Sang-Jun Kim[4], and Jae-Ho Shin[2*]

[1]*Department of Life Sciences, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*
[2]*School of Applied Biosciences, College of Agriculture and Life Sciences, Kyungpook National University, Daegu 41566, Republic of Korea*
[3]*Department of Chemistry, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*
[4]*Department of Natural Sciences, Republic of Korea Naval Academy, Changwon 51704, Republic of Korea*

기름으로 오염된 토양에서 분리된 다환방향족탄화수소 분해 세균 *Idiomarina piscisalsi* 10PY1A의 유전체 염기서열 해독

Alexis Nzila[1†*] · 정병권[2†] · 김민철[2] · Jerald Conrad Ibal[2] · Fitri Budiyanto[1] · Musa M. Musa[3] · Assad Thukair[1] · 김상준[4] · 신재호[2*]

[1]킹파드석유광물대학교 생명과학부, [2]경북대학교 농업생명과학대학 응용생명과학부, [3]킹파드석유광물대학교 화학부, [4]해군사관학교 자연과학부

(Received June 1, 2018; Revised June 24, 2018; Accepted July 3, 2018)

Using pyrene as the enrichment nutrient, a bacterial strain 10PY1A, was isolated by enrichment culture from oil-contaminated sea sand of Arabian Gulf in Saudi Arabia, and this strain belongs to the species *Idiomarina piscisalsi*, based on 16S RNA gene sequence analysis. The genome of *I. piscisalsi* strain 10PY1A contains 2,346 protein-coding sequences and an average GC content of 47.4% in its chromosome (2.59 Mbp). Genes encoding proteins related to the degradation of pyrene were existed in the strain 10PY1A genome, indicating that this strain can be used to degrade polycyclic aromatic hydrocarbons in oil-contaminated marine flora and soil.

Keywords: *Idiomarina piscisalsi* strain 10PY1A, complete genome sequence, polycyclic aromatic hydrocarbon, pyrene

[†]These authors contributed equally to this work.
***For correspondence.** (J.H. Shin) E-mail: jhshin@knu.ac.kr;
Tel.: +82-53-950-5716; Fax: +82-53-953-7233/
(A. Nzila) E-mail: alexisnzila@kfupm.edu.sa;
Tel.: +966-03-860-7716; Fax: +966-03-860-7031

The contamination of the environment by polycyclic aromatic hydrocarbons (PAHs) is a growing concern since these compounds are associated with toxic effect in marine flora and human health (Bostrom *et al*., 2002). These compounds are found in environments with high salinity, and one of such environments is “produced water”, which results from oil extraction (Fakhru'l-Razi *et al*., 2009). Produced water contains contaminant PAHs and has high salinity, reaching values as high as 30% (w/v) NaCl (Fakhru'l-Razi *et al*., 2009). Thus, the use of bioremediation to remove these contaminants requires halophilic bacteria that can biodegrade these pollutants. *Idiomarina piscisalsi* strain 10PY1A was isolated from oil-contaminated

sea sand of Arabian Gulf in Saudi Arabia, its ability to biodegrade pyrene was assessed in various conditions such as pH, temperature, and salinity (unpublished data). Here, we present the complete genome sequence and annotation result of *I. piscisalsi* strain 10PY1A.

The Wizard Genomic DNA Purification Kit (Promega) was used to extract genomic DNA from *I. piscisalsi* strain 10PY1A. A single molecule real-time (SMRT) sequencing platform on a PacBio RS II instrument with P6-C4 chemistry (Pacific Biosciences) (Eid *et al*., 2009) was used to obtain the whole genome sequence. Using a SMRT cell with a 180 min movie, raw sequence data were obtained as 83,385 reads of 754,626,799 bp. The reads were *de novo*-assembled using the protocol RS HGAP Assembly 2 in SMRT analysis version 2.3 (Chin *et al*., 2013). Prokaryotic Genome Annotation Pipeline version 4.2 software on the NCBI website (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) was used to predict the coding DNA sequences. Additional functional annotation was performed with Rapid Annotation using Subsystem Technology server (Aziz *et al.*, 2008).

The complete genome features of *I. piscisalsi* strain 10PY1A are summarized in Table 1 and Fig. 1. The genome was composed of a circular chromosome. The total size of the chromosome was 2,587,498 bp with 244 × average coverage.

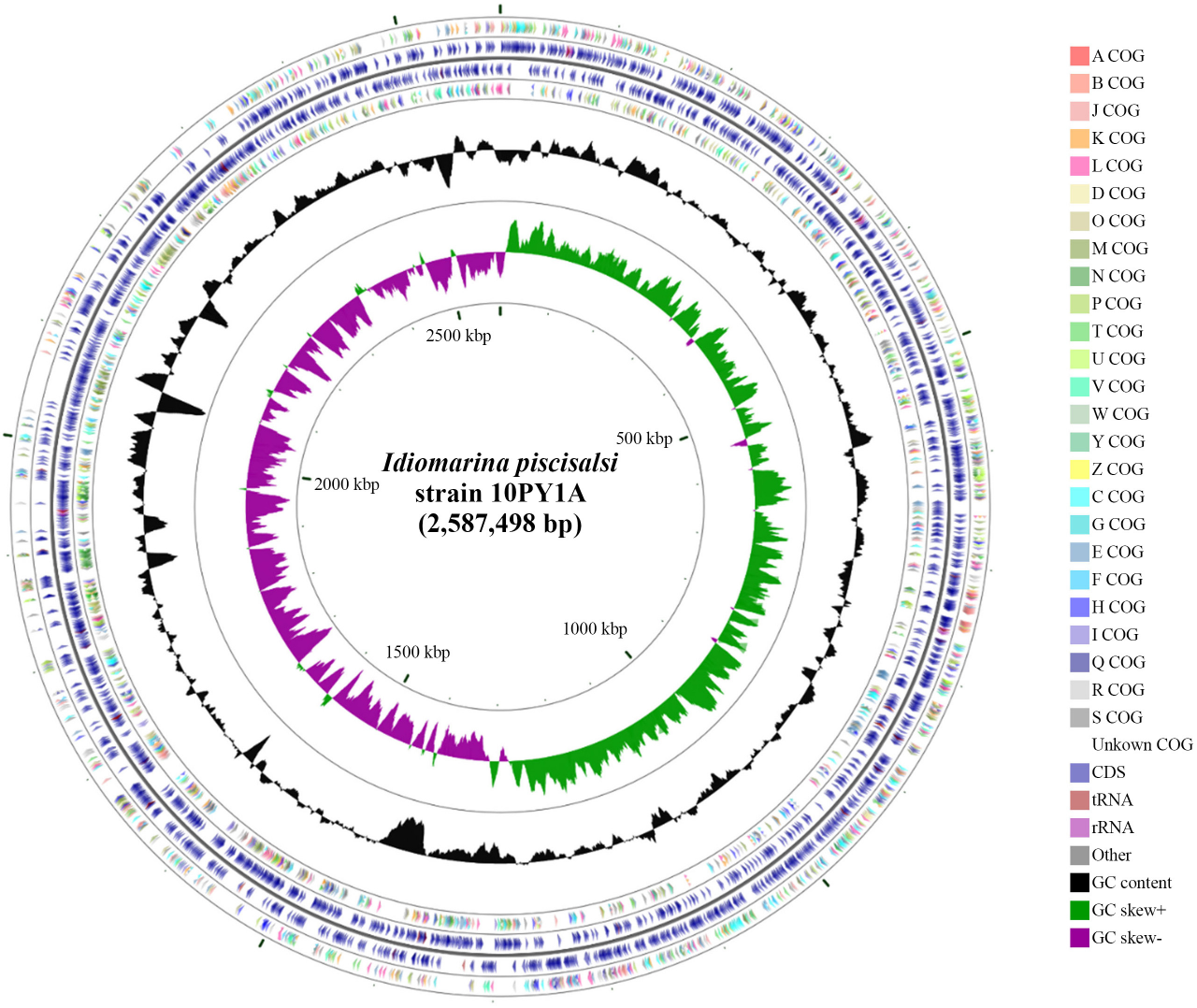

Fig. 1. Circular map of the chromosome of *Idiomarina piscisalsi* strain 10PY1A. From the outside to the center: genes on the forward strand (colored by clusters of orthologous groups categories, COGs), coding DNA sequence (CDS) on the forward strand, CDS on the reverse strand, genes on the reverse strand (colored by COG categories), GC content, and GC skew.

Table 1. Genome features of *Idiomarina piscisalsi* strain 10PY1A

Genomic features	Value
Genome size (bp)	2,587,498
GC content (%)	47.4
Total genes	2,454
Protein coding genes	2,346
rRNAs (5S, 16S, 23S)	12
tRNAs	56
Pseudogenes	36

The genome contained 2,346 protein-coding genes, 12 rRNA genes, 56 tRNA genes, and 36 pseudogenes with a 47.4% GC content. Among the protein-coding genes, several open-reading frames were identified as degradation of aromatic compounds and hydrocarbons-related genes: homogentisate 1,2-dioxygenase (ASG65912.1), 4-hydroxyphenylpyruvate dioxygenase (ASG65913.1), maleylacetoacetate isomerase (ASG66932.1), fumarate hydratase (ASG5443.1), 4-hydroxyphenylpyruvate dioxygenase (ASG65913.1), phenylalanine-4-hydroxylase (ASG65916.1), alcohol dehydrogenase (ASG66282.1), and aldehyde dehydrogenase (ASG66409.1). For the detection of presence of pyrene degradation genes in *Idiomarina piscisalsi* 10PY1A, BLASTX was done with all the hypothetical proteins (482) in 10PY1A with the annotated genes involved in the pyrene degradation pathway present in *Mycobacterium vanbaalenii* PYR-1. Rieske nonheme iron aromatic ring-hydroxylating oxygenase, dihydrodiol dehydrogenase, ferrodin reductase, γ-carboxymuconolactone decarboxylase/β-ketoadipate enol-lactone hydrolase, aldehyde dehydrogenase, 1-hydroxy-2-naphthoate hydroxylase, phthalate 3,4-dioxygenase, and protocatechuate 3,4-dioxygenase were all compared to the hypothetical proteins in 10PY1A which showed 57,14% similarity (ASG66445.1), 44.64% (ASG64867.1), 53.6% (ASG66070.1), 54.29% (ASG65749.1), 51.35% (ASG66406.1), 50% (ASG66814.1), 52.63% (ASG65577.1), 51.53% (ASG64827.1), respectively. The substrate specificity of alcohol dehydrogenase is not restricted to aliphatic alcohols; xenobiotic aromatic and alicyclic hydroxyls are also metabolized through similar pathways, highlighting the physiological importance of this enzyme system (Oppermann and Maser, 2000). Additionally, aldehyde dehydrogenases are widely distributed in living organisms and are involved in the detoxification of the toxic aldehydes produced by several cellular metabolic pathways, being recognized as one of the essential enzymes for the degradation of many hydrocarbon compounds (Sierra-García *et al*., 2014). The complete genome sequence shows the existence of various genes of enzymes that mediate the biodegradation of aromatic compounds, supporting further the usefulness of this strain in bioremediation strategies in halophilic conditions.

Availability of the sequence data and strain

The complete genome sequence of *I. piscisalsi* strain 10PY1A has been deposited in DDBJ/EMBL/GenBank under the accession number CP022133. The genome project for this strain is listed in the JGI GOLD under project Gp0256221. The strain is available from Korean Collection for Type Cultures with the accession number KCTC 62123.

적 요

석유로 오염된 사우디아라비아의 바다 모래에 pyrene을 영양물로 첨가하는 집식배양을 통해 우점하는 박테리아 균주 10PY1A가 분리되었다. 이 균주는 16S rRNA 유전자 분석을 통해 *Idiomarina piscisalsi*로 동정되었다. 이 균주는 G + C 비율이 47.4%이며 2.69 Mbp 크기의 원형 염색체를 보유하고 있으며, 해당 염색체는 다환방향족탄화수소 분해와 관련 있는 유전자를 포함한 2,346개의 단백질 코딩 유전자로 구성되어 있다. 이는 이 균주가 석유로 오염된 해양과 토양에서 방향족 탄화수소를 분해하는 데 사용될 수 있음을 보여준다.

Acknowledgements

The work was supported by King Abdulaziz City for Science and Technology (KACST) through the Science and Technology Unit at King Fahd University of Petroleum and Minerals (KFUPM), for funding this work through project No. 13-ENV1628-04, as part of the National Science, Technology and Innovation Plan.

References

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, *et al.* 2008. The

RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75.

Bostrom CE, Gerde P, Hanberg A, Jernstrom B, Johansson C, Kyrklund T, Rannug A, Tornqvist M, Victorin K, and Westerholm R. 2002. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ. Health Perspect*. **110**, 451-488.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, *et al.* 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* **10**, 563-569.

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, *et al.* 2009. Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133-138.

Fakhru'l-Razi A, Pendashteh A, Abdullah LC, Biak DRA, Madaeni SS, and Abidin ZZ. 2009. Review of technologies for oil and gas produced water treatment. *J. Hazard Mater*. **170**, 530-551.

Oppermann UCT and Maser E. 2000. Molecular and structural aspects of xenobiotic carbonyl metabolizing enzymes. Role of reductases and dehydrogenases in xenobiotic phase I reaction. *Toxicology* **144**, 71-81.

Sierra-García IN, Alvarez JC, de Vasconcellos SP, de Souza AP, dos Santos Neto EV, and de Oliveira VM. 2014. New hydrocarbon degradation pathways in the microbial metagenome from Brazilian petroleum reservoirs. *PLoS One* **9**, e90087.