Korean Journal of Microbiology (2018) Vol. 54, No. 3, pp. 283-285
DOI https://doi.org/10.7845/kjm.2018.8031

pISSN 0440-2413
eISSN 2383-9902

Complete genome sequences of *Azoarcus* sp. TSPY31 and TSNA42 potentially having biosynthetic ability to produce indigo

Hae-Seon Kim[1], Sun Ho Cha[2], Ho Young Suk[3], Nyun-Ho Park[1], and Jung-Hee Woo[1]*

[1]*Gyeongbuk Institute for Marine Bio-Industry (GIMB), Uljin 36315, Republic of Korea*
[2]*GenoTech Corporation, Daejeon 34113, Republic of Korea*
[3]*Department of Life Sciences, Yeungnam University, Gyeongsan 38541, Republic of Korea*

인디고 생산능을 가진 *Azoarcus* sp. TSPY31과 TSNA42의 유전체 분석

김해선[1] · 차선호[2] · 석호영[3] · 박년호[1] · 우정희[1]*

[1]경북해양바이오산업연구원, [2]㈜제노텍, [3]영남대학교 생명과학과

(Received May 30, 2018; Revised June 17, 2018; Accepted June 18, 2018)

Azoarcus are known to contain bacterial strains usually found in contaminated areas. Two strains of *Azoarcus* sp., TSPY31 and TSNA42, were isolated from oil-contaminated marine tidal flats, and their genomic structures were analyzed. The genomes of both TSPY31 and TSNA42 were composed of a single complete chromosome of 4,572,082 bp (G + C content: 63.2%) and 4,886,934 bp (G + C content: 62.8%), respectively. Both genomes were found to contain two copies of styrene monooxygenases that are predicted to be responsible for converting indole to indigo.

Keywords: *Azoarcus*, complete genome, indigo, styrene monooxygenase

Two strains of *Azoarcus* sp., TSPY31 and TSNA42, were isolated from oil-contaminated tidal flats around city of Taean in western South Korea, and their genomic structures were analyzed to examine their biosynthesis potential. The samples collected were subjected to cultural enrichment method on the minimal medium containing polycyclic aromatic hydrocarbons (PAHs). The PAH-degrading strains could be isolated by sublimation test in which PAHs was supplied as a vapor (Kwon *et al*., 2015). The PAH-degrading strains also degraded the indole, an aromatic compound, and then converted it into indigo through a spontaneous oxidation reaction (Dagher *et al*., 1997). Therefore, we expected these strains to potentially contain bioconversion activity producing indigo. Indigo, a raw material for blue dye, can be produced from indole via an intermediate, indoxyl, by the activity of various oxygenases commonly found in microorganisms (Kim *et al*., 2003; Han *et al*., 2008).

The extracted DNA samples of the two strains were used to construct the whole-genome libraries using SMRTbell™ template prep kit, and the sequencing was performed on a PacBio RSII platform (Pacific Biosciences; Pearson *et al*., 2006). The subreads obtained were *de novo* assembled using Hierarchical Genome Assembly Process (HGAP, Version 2.3) workflow and the draft assemblies were polished using Quiver (Chin *et al*., 2013). Our sequencing run produced approximately 53X genomic coverage, yielding 150,007,652 bp high-quality sequence. Since bacterial genomes are typically circular, all contigs were checked using MUMmer 3.5 (Kurtz *et al*., 2004) and one of the self-similar ends was trimmed for manual genome closure. Genome annotation was performed using NCBI's Prokaryotic Genome Annotation Pipeline (PGAP).

Azoarcus sp. TSPY31 genome contained only a single

***For correspondence.** E-mail: wjh1991@naver.com;
Tel.: +82-54-780-3454; Fax: +82-54-780-3469

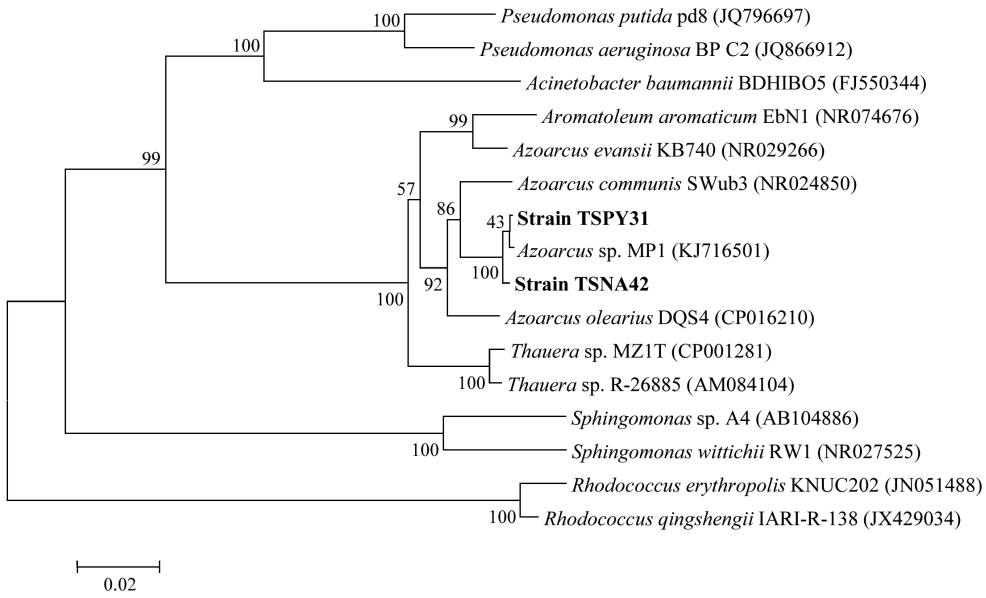

Fig. 1. Neighbor-joining tree reconstructed using MEGA v7 to identify the phylogenetic placement of *Azoarcus* sp. TSPY31 and TSNA42.

Table 1. General feature of *Azoarcus* sp. TSPY31 and TSNA42

Genome feature	Value	
	TSPY31	TSNA42
Genome size (bp)	4,572,082	4,886,934
No. of contig	1	1
GC content (%)	63.2	62.8
Genes	4,182	4,449
Protein coding genes	4,230	4,516
rRNA	12	12
tRNA	58	58
Protein coding genes with function prediction	3,557	3,747
Pseudo genes	49	44

chromosome of 4,572,082 bp with G + C content of 63.2%. A total of 4,230 protein coding genes were predicted to exist in the genome, 3,557 genes of which were completely annotated. In addition, 12 rRNA and 58 tRNA were discovered. *Azoarcus* sp. TSNA42 genome also consisted of a similar-sized single chromosome (4,886,934 bp with G + C content of 62.8%). Our analysis annotated 12 rRNAs, 58 tRNAs and 3,747 functional protein-coding genes (out of 4,516 predicted genes) from this strain.

Both strains were identified to be highly similar in genomic structure (Fig. 1) and to contain various oxygenases potentially responsible for the bioconversion of indole to indigo. In both strains, there were two copies of an oxygenase gene highly similar to styrene monooxygenase (SMO, accession no. ABX24519) derived from *Pseudomonas putida* that is already well known to play a role in indigo synthesis (O'Connor *et al*., 1997; Cheng *et al*., 2016).

Nucleotide sequence accession number

The complete genome sequences of *Azoarcus* sp. TSPY31 and TSNA42 were deposited in NCBI GenBank under the accession numbers of CP022187 and CP022188, respectively. Both strains are available from the Korean Culture Center of Microorganisms with the accession number of KCCM 11873P and KCCM11872P, respectively.

적 요

유류 오염된 해양 갯벌에서 분리한 다환방향수소족(PAHs)을 분해하는 균주들로부터 인디고로 생물전환 활성을 가진 것으로 예측되는 *Azoarcus* sp. TSPY31과 TSNA42 균주를 동정하였다. 이 두 균주의 유전체 분석을 실시한 결과, 모두 하나의 완전한 chromosome으로 구성되며, TSPY31은 총 4,572,082 bp에 G + C 함량은 63.2%로 이루어져 있고, TSNA42는 4,886,934 bp에 G + C 함량은 62.8%이었다. 이 두 균주 모두 인돌을 인디

고로 전환하는 효소인 styrene monooxygenase를 각각 2 copy 씩 보유하고 있는 것으로 확인되었다.

Acknowledgements

This work (Grants No. C0563841) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2017 and Gyeongsangbuk-do R&D Program.

References

Cheng L, Yin S, Chen M, Sun B, Hao S, and Wang C. 2016. Enhancing indigo production by over-expression of the styrene monooxygenase in *Pseudomonas putida*. *Curr. Microbiol*. **73**, 248-254.

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.

Dagher F, Déziel E, Lirette P, Paquette G, Bisaillon JG, and Villemur R. 1997. Comparative study of five polycyclic aromatic hydrocarbon degrading bacterial strains isolated from contaminated soils. *Can. J. Microbiol.* **43**, 368-377.

Han GH, Shin HJ, and Kim SW. 2008. Optimization of bio-indigo production by recombinant *E. coli* harboring fmo gene. *Enzyme Microb. Technol.* **42**, 617-623.

Kim JY, Lee K, Kim Y, Kim CK, and Lee K. 2003. Production of dyestuffs from indole derivatives by naphthalene dioxygenase and toluene dioxygenase. *Lett. Appl. Microbiol.* **36**, 343-348.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, and Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12.

Kwon TH, Woo JH, Park NH, and Kim JS. 2015. Characterization of PAH (Polycyclic aromatic hydrocarbon)-degrading bacteria isolated from commercial gasoline. *Korean J. Environ. Agric.* **34**, 244-251.

O'Connor KE, Dobson ADW, and Hartmans S. 1997. Indigo formation by microorganisms expressing styrene monooxygenase. *Appl. Environ. Microbiol.* **63**, 4287-4291.

Pearson GA, Serrao EA, Procaccini G, Duarte C, and Marba N. 2006. Genomic DNA isolation from green and brown algae (caulerpales and fucales) for microsatellite library construction. *J. Phycol.* **42**, 741-745.