The draft genome sequencing for *Flavobacteriaceae* strain EG-1, isolated from coastal seawater in Korea, was performed using Illumina HiSeq platform. As a result, the draft genome was comprised of a total length of approximately 4.45 Mbp with G + C content of 33.5%, and included a total of 3,719 protein-coding, 7 rRNA, 37 tRNA, 4 non-coding RNA, and 14 pseudo genes. The strain EG-1 had various biopolymer degradation-related genes and carotenoid biosynthesis related genes, and is assumed to contribute to the cycling of organic matters in the marine environment.

**Keywords:** *Flavobacteriaceae*, draft genome sequence, Illumina HiSeq

The family *Flavobacteriaceae* is the largest bacterial group (comprising 143 genera) of the phylum *Bacteroidetes*. It is found in a wide variety of marine, freshwater, and soil habitats (McBride, 2014). Most of members were characterized as Gram-negative, rod-shaped, chemoheterotrophic, and carotenoid-or flexirubin-pigmented bacteria possessing menaquinone 6 as the sole respiratory quinone. The family *Flavobacteriaceae* was subdivided based on the 16S rRNA gene sequence-based phylogenetic analysis, among which species of “marine clade” were known to contribute greatly to the mineralization of primary produced organic matter in various marine habitats (Kirchman, 2002). Several members of the marine clade were also known to be opportunistic pathogens such as *Tenacibaculum maritimum* (Suzuki *et al*., 2001) or potentially exacerbate existing disease (Bowman and Nowak, 2004; McBride, 2014).

In this report, the draft genome sequence and annotation of a marine *Flavobacteriaceae* strain EG-1 is described.

An orange-pigmented marine bacterium, designated strain EG-1, was isolated from coastal seawater, Guryongpo (35° 59' 23.3'' N, 129° 33' 56.3'' E) of Korea, using a standard dilution plating method on 10 fold-diluted marine agar 2216 (MA; Difco). The strain EG-1 was routinely cultured on MA at 25°C for 3 days. Phylogenetic analysis based on 16S rRNA gene sequence showed that strain EG-1 belonged to the family *Flavobacteriaceae* and was closely related to *Aurantibacter crassamenti* HG732^T^ (98.4% sequence similarity), followed by *Kriegella aquimaris* KMM 3665^T^ (94.3% sequence similarity) and *Zobellia galactanivorans* DsI^T^ (93.6% sequence similarity).

The genomic DNA was extracted from cells grown in marine broth 2216 (Difco) at 25°C for 2 days using MagAttract® HMW DNA kit (Qiagen) according to the manufacturer’s instruction. The sequencing of the draft genome was performed on the Illumina HiSeq platform with TruSeq Nano DNA (350 bp insert
The draft genome of Flavobacteriaceae strain EG-1 was sequenced using the Illumina Hiseq platform. The de novo assembly was performed by SPAdes (version 3.10.0). The estimated completeness and contamination of genome were verified with CheckM (version 1.0.18) (Parks et al., 2015). Genome annotation was performed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016) and additional function of the predicted genes were conducted by EggNOG 5.0, BlastKOALA with KEGG database, and Rapid Annotations using Subsystems Technology (RAST) server with SEED (Aziz et al., 2008).

The genome statistics are shown in Table 1. The draft genome sequence of strain EG-1 consisted of 4,449,445 bp (10 contigs; genome coverage, 150.3X) with a DNA G + C content of 33.5%. The results of CheckM estimation indicated that genome completeness at 96.6% with 0.61% contamination and 0.0% strain heterogeneity. The genome was shown to contain 3,719 protein-coding, 7 rRNA (5S, 16S, 23S), 37 tRNA, 4 non-coding RNA, and 14 pseudo genes.

As a result of genome annotation and analysis of functional characterization, the draft genome sequence of strain EG-1 contained various biopolymer degradation-related genes; xylanase Xyl, xylose isomerase XylA, xylose kinase XylB, N-acetylglucosamine-6-phosphatase deacetylase NagA, glucosamine-6-phosphate deaminase NagB, type I pullulanase PulA, α-amylase AmyA and exodeoxyribonuclease III XthA. It also contained carotenoid biosynthesis related genes; 15-cis-phytoene synthase CrtB, phytoene desaturase CrtI and β-carotene 3-hydroxylase CrtZ. The genome sequence contained antibiotic-resistance related genes; tetracycline resistance protein TetA, penicillin-binding protein MrcA/MrdA, penicillin repressor BlaI, virginiamycin A acetyltransferase Vat, multiple antibiotic resistance protein MarC, and multidrug efflux pump subunit AcrB. It contained bacterial secretion system-related genes such as outer membrane protein TolC, preprotein translocase subunit SecA/SecD/SecE/SecG/SecY/YajC, YidC/Oxa1 family membrane protein insertion YidC, fused signal recognition particle receptor FtsY/SRP54, signal recognition particle subunit SRP54 and sec-independent translocase protein TatA/TatC. Nitrogen metabolism-related genes nitrite reductase large subunit NirB, ferredoxin-nitrate reductase NarB, nitrous oxide reductase maturation protein NosD/NosF/NosL, nitrous oxide reductase transmembrane maturation protein NosY and nitrous-oxide reductase NosZ were also found in the genome sequence. Based on the genome analysis, strain EG-1 is thought to play an important role of the degradation of various biopolymers and the cycling of organic matters.

Table 1. Genome features of Flavobacteriaceae strain EG-1

<table>
<thead>
<tr>
<th>Genome features</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of contigs</td>
<td>10</td>
</tr>
<tr>
<td>Depth (X)</td>
<td>150.3</td>
</tr>
<tr>
<td>Genome size (bp)</td>
<td>4,449,445</td>
</tr>
<tr>
<td>G + C content (%)</td>
<td>33.5</td>
</tr>
<tr>
<td>Protein-coding genes</td>
<td>3,719</td>
</tr>
<tr>
<td>tRNA genes</td>
<td>37</td>
</tr>
<tr>
<td>rRNA genes (5S, 16S, 23S)</td>
<td>7 (3, 2, 2)</td>
</tr>
<tr>
<td>Non-coding RNA genes</td>
<td>4</td>
</tr>
<tr>
<td>Pseudo genes</td>
<td>14</td>
</tr>
</tbody>
</table>

The strain EG-1 has been deposited in Korean Culture Center of Microorganisms (KCCM) and Japan Collection of Microorganisms (JCM) under the accession number KCCM 43340 and JCM 33636, respectively. The draft genome sequence is accessible in GenBank under the accession number WJPK 00000000. The version described in this paper is version WJPK01000000.

Nucleotide sequence accession numbers

The strain EG-1 has been deposited in Korean Culture Center of Microorganisms (KCCM) and Japan Collection of Microorganisms (JCM) under the accession number KCCM 43340 and JCM 33636, respectively. The draft genome sequence is accessible in GenBank under the accession number WJPK 00000000. The version described in this paper is version WJPK01000000.

적요

이 연구에서는 Illumina Hiseq platform을 사용하여 연안 해수로부터 분리된 Flavobacteriaceae 균주 EG-1의 유전체 염기서열 해독을 수행하였다. 그 결과, 유전체는 대략 4.45 Mbp의 길이 및 33.5%의 G + C 함량으로 구성되었고, 전체 3,719개의 단백질인 암호 유전자, 7개의 rRNA 유전자, 37개의 tRNA 유전자, 4개의 non-coding RNA 유전자 및 14개의 위유전자 (pseudo gene) 가 확인되었다. EG-1 균주는 다양한 생물 고분자 물질 분해 관련 유전자 및 카로테인과 관련 유전자들을 가지고 있었고, 해양 환경에서 유기물 순환에 기여할 것으로 추정된다.
Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A3B04033871).

References


