Complete genome sequence of Microbacterium aurum strain KACC 15219T, a carbohydrate-degrading bacterium

YeonGyun Jung1, Byung Kwon Jung1, Chang Eon Park1, Jerald Conrad Ibal1, Sang-Jun Kim2, and Jae-Ho Shin1* 1
1School of Applied Biosciences, College of Agriculture and Life Sciences, Kyungpook National University, Daegu 41566, Republic of Korea
2Department of Natural Sciences, Republic of Korea Naval Academy, Changwon 51704, Republic of Korea

The complete genomic information of Microbacterium aurum KACC 15219T (= IFO 15204T = DSM 8600T) is described. The genome of M. aurum KACC 15219T contains 3,096 protein coding genes and an average G+C content of 69.9% in its chromosome (3.42 Mbp). This strain can use various carbon sources for growth, including quinic acid. Quinic acid is used as a substrate for the synthesis of aromatic amino acids via the shikimate pathway which are useful in the industry. M. aurum KACC 15219T will provide basis to improve our understanding of this organism and allow more efficient application of the strain to industry.

Keywords: Microbacterium aurum, aromatic amino acids, quinic acid, shikimate pathway

Microbacterium aurum strain KACC 15219T (= IFO 15204T = DSM 8600T) was isolated from corn steep liquor. A polyphasic taxonomic study revealed that M. aurum KACC 15219T could utilize 16 types of sole carbon substrates (Yokota et al., 1993). This strain has the potential to utilize quinic acid as a sole carbon source. Quinic acid is the substrate used to synthesize aromatic amino acids via the shikimate pathway (Guo et al., 2014). These aromatic amino acids are useful in food and pharmaceutical industries (Koma et al., 2012). Although many DNA sequences from the genus Microbacterium are available, genome sequencing of M. aurum KACC 15219T has not been conducted. We obtained the whole genome sequence and reported the complete genome of the strain.

Genomic DNA was extracted using QIAmp DNA Mini kit (Qiagen) through the manufacturer’s instructions. A single molecule real-time sequencing platform from PacBio RS II (Pacific Biosciences) was used to obtain the whole genome sequence (Eid et al., 2009; Nzila et al., 2018). The sub-reads from the raw sequencing reads following adapter-removal were used for de novo assembly using HGAP version 2 (Chin et al., 2013) based on 85,020 quality reads with a mean length of 12,521 bp. It produced a circular chromosome having 3,424,892 bp with 192.19-fold average coverage and was annotated automatically by using the RAST server (Aziz et al., 2008) and PGAAP from NCBI (Angiuoli et al., 2008). Genome annotation was performed by the prediction of protein-coding

*For correspondence. E-mail: jhshin@knu.ac.kr; Tel.: +82-53-950-5716; Fax: +82-53-953-7233
Complete genome sequence of *M. aurum* strain KACC 15219<sup>T</sup>

Genome features of *Microbacterium aurum* strain KACC15219<sup>T</sup> are summarized in Table 1 and Fig. 1. The genome was composed of a circular chromosome having the size of 3,424,892 bp with a GC content of 69.9%. In total, 3,302 genes, including 154 pseudogenes, were predicted from the genome sequence, and 52 RNA genes (3 rRNAs, 46 tRNAs) were identified. Approximately 72.00% of the total genes encoded proteins with known functions and 886 genes were annotated as hypothetical protein. Among the total genes, 2,012 genes were assigned to COGs. Genome analysis revealed that this strain

<table>
<thead>
<tr>
<th>Genomic features</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome size (bp)</td>
<td>3,424,892</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>69.9</td>
</tr>
<tr>
<td>Total genes</td>
<td>3,302</td>
</tr>
<tr>
<td>Protein coding genes</td>
<td>3,096</td>
</tr>
<tr>
<td>rRNAs (5S, 16S, 23S)</td>
<td>1, 1, 1</td>
</tr>
<tr>
<td>tRNAs</td>
<td>46</td>
</tr>
<tr>
<td>Pseudogenes</td>
<td>154</td>
</tr>
</tbody>
</table>

Table 1. Genome features of *Microbacterium aurum* strain KACC15219<sup>T</sup>

Fig. 1. Circular representation and subsystem category distribution of the genome of *Microbacterium aurum* KACC15219<sup>T</sup>. Circles are numbered from 1 (the outermost circle) to 6 (the innermost circle). The outer four circles show the forward and reverse strand of COG categories and coding sequence (CDS). The fifth circle represents the GC content (black). The sixth circle demonstrates the GC skew curve (green, positive GC skew; violet, negative GC skew). The genome position scaled in kb from base 1 is shown on the inner circle.
possessed putative enzymes for central carbohydrate metabolism to assimilate these carbon sources through different metabolic pathways (Justice et al., 2014). Putative enzymes responsible for the utilization of diverse sole carbons were found in the genome. All key enzymes in the Embden-Meyerhof-Parnas pathway and TCA cycle were present in *M. aurum* KACC 15219T. The presence of dihydrolipoamide acyltransferase (APZ34619.1), pyruvate dehydrogenase (acetyl-transferring) E1 component subunit alpha (APZ34621.1), and dihydrolipoamide acyltransferase (APZ34487.1) in the pyruvate metabolism pathway suggests that pyruvate is converted to acetyl-CoA. Moreover, the presence of type II 3-dehydroquinate dehydratase (APZ34549.1), shikimate dehydrogenase (APZ34542.1), shikimate kinase (APZ34544.1), 3-phosphoshikimate dehydratase (APZ33558.1) indicated that *M. aurum* KACC 15219T may utilize quinic acid to synthesize three aromatic amino acids via the shikimate pathway (Guo et al., 2014). This strain can use various carbon sources including quinic acid and may provide insight to improve our understanding of this organism and allow more efficient application of the strain to industry.

**Availability of the sequence data and strain**

The complete genome sequence of *M. aurum* strain KACC 15219T has been deposited in EMBL/GenBank under the accession number CP018762.1. The genome project for this strain is listed in the JGI GOLD under project Gp0191354.

**References**


**Acknowledgements**

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ013383)” Rural Development Administration, Republic of Korea.

이 연구에서는 *Microbacterium aurum* KACC 15219T (=IFO 15204T = DSM 8600T)의 완전한 유전체 서열이 해독되었다. 하나의 원형 염색체는 3.42 Mbp였으며 G+C 함량이 69.9%였다. 해당 염색체 염기서열을 주석화한 결과, 총 3,096개의 유전자 서열이 발견되었다. 16중 이상의 탄소원을 분해하는 것으로 알려진 *M. aurum* KACC 15219T에는 방향족 아미노산 합성 가스의 quinic acid를 비롯한 다양한 탄소원의 이용과 관련된 유전자가 존재하였다. *M. aurum* KACC 15219T의 유전체 정보는 이 미생물에 대한 이해를 높이고 산업적인 이용을 위한 기반이 될 것이다.