<table>
<thead>
<tr>
<th>Title</th>
<th>Draft Genome Sequence of the Haloacid-Degrading Burkholderia caribensis Strain MBA4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Pan, Y; Kong, KF; Tsang, JSH</td>
</tr>
<tr>
<td>Citation</td>
<td>Genome Announcements, 2014, v. 2 n. 1, p. e00047-14</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2014</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/196403">http://hdl.handle.net/10722/196403</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Genome Announcements. Copyright © American Society for Microbiology.; Copyright © American Society for Microbiology, Genome Announcements, v. 2 n. 1, p. e00047-14, 2014; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Draft Genome Sequence of the Haloacid-Degrading \textit{Burkholderia caribensis} Strain MBA4

Yanling Pan, Ka Fai Kong, Jimmy S. H. Tsang
Molecular Microbiology Laboratory, School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China

\textit{Burkholderia caribensis} MBA4 was isolated from soil for its ability to utilize 2-haloacid. An inducible haloacid operon, encoding a dehalogenase and a permease, is mainly responsible for the biotransformation. Here, we report the draft genome sequence of this strain.

Haloacettes such as monochloroacetate (MCA) are toxic and mutagenic and can be produced accidentally during disinfection of water. \textit{Burkholderia caribensis} MBA4 is a Gram-negative bacterium that can utilize 2-haloacid as a growth substrate. This bacterium was characterized for its production of a dimeric hydrolytic dehalogenase (Deh4a) (1, 2) that removes the halogen from the carbon backbone. Here we describe the draft genome sequence of \textit{Burkholderia caribensis} MBA4.

Analysis of \textit{B. caribensis} MBA4 with pulsed-field gel electrophoresis showed that it has a genome size of more than 9 Mb with at least three replicons (data not shown). Whole-genomic sequencing was obtained with 454 GS FLX Titanium and Illumina HiSeq 2000. With low-quality short reads discarded, the 454 sequencing has 929,485 reads and 380,525,001 bp after trimming. Four sets of Illumina paired-end libraries with insert sizes of 100, 300, 500, and 2,000 bp were constructed and sequenced. After trimming and filtering, the four libraries have sizes of 61, 61, 69, and 39 bp, respectively. The overall coverage is 37,483,321, 36,788,695, 23,594,431, and 12,689,821 high-quality reads identified by Tandem Repeats Finder (6). Among the 9,022 protein-coding sequences, 76% were annotated as encoding known proteins while the remaining 24% encode hypothetical products. Among these RAST-annotated genes, 3,666 coding DNA sequences (CDS) were assigned to 27 subsystems. Analysis of the CDS with the KEGG Automatic Annotation Server (version 1.6a) (7) has specified 34 groups with 191 pathways.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AXDD00000000. The version described in this paper is version AXDD01000000.

ACKNOWLEDGMENTS
We thank S. Lok, A. Tong, N. Lin, J. Jiang, F. C. C. Leung, and the University Centre for Genomic Sciences for advice.

This work has been supported by grants from the University Small Project Funding 2010 and the General Research Fund (project number HKU 780511) of the Research Grants Council of the Hong Kong Special Administrative Region, China.

REFERENCES

