Whole-genome sequences of five burkholderia pseudomallei isolates from australian cystic fibrosis patients

Viberg, Linda T.; Price, Erin P.; Kidd, Timothy J.; Bell, Scott C.; Currie, Bart J.; Sarovich, Derek

Published in:
Genome Announcements (genomeA)

DOI:
10.1128/genomeA.00254-15

Published: 16/04/2015

Citation for published version (APA):
Whole-Genome Sequences of Five *Burkholderia pseudomallei* Isolates from Australian Cystic Fibrosis Patients

Linda T. Viberg, Erin P. Price, Timothy J. Kidd, Scott C. Bell, Bart J. Currie, Derek S. Sarovich

We report here five improved high-quality draft genomes of *Burkholderia pseudomallei* isolated from Australian cystic fibrosis (CF) patients. This pathogen is rarely seen in CF patients. These genomes will be used to better understand chronic carriage of *B. pseudomallei* in the CF lung and the within-host evolution of longitudinal isolates from these patients.

*B. pseudomallei* is the causative agent of melioidosis, a potentially lethal disease with multiple clinical presentations, of which pneumonia is most common (1–3). Infection with *B. pseudomallei* is acquired following accidental percutaneous inoculation with contaminated soil or water, or by inhalation or ingestion. Most reported cases of melioidosis are from the regions of northern Australia and Southeast Asia that are highly endemic for *B. pseudomallei*, with sporadic reports from other endemic regions, including the Middle East, Africa, South and Central America, and the Caribbean (4, 5).

Cystic fibrosis (CF) patients typically develop infections caused by a range of opportunistic pathogens, such as *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Haemophilus influenzae*, which may be eradicated in the early phases, but with increasing age, these can develop into a chronic infection (6). The factors that promote the development of infection are complex but include the abnormal composition of the airway lining fluid as a consequence of abnormal expression of the CF transmembrane regulator protein (CFTR). *B. pseudomallei* has been identified in the CF lung in a small proportion of CF patients living or traveling to regions endemic for the pathogen (7–12).

*B. pseudomallei* is known to cause chronic infections that can be difficult to treat and, in certain instances, persist for years (13, 14). Although many conventional CF-associated pathogens have been studied in detail, little is known about how *B. pseudomallei* behaves within the CF lung. Reports have demonstrated that, like most *B. pseudomallei* infections, clinical symptoms can present as either acute or chronic disease (9). Here, we present improved high-quality draft genome sequences (15) of the initial *B. pseudomallei* strains isolated from five Australian CF patients with chronic *B. pseudomallei* infection.

DNA was extracted from purified culture, as previously reported (16). The DNA samples were subjected to whole-genome sequencing (WGS) from a paired-end Nextera library and with a ~300-bp insertion size using the Illumina HiSeq 2000 platform (Illumina, Inc., San Diego, CA) at Macrogen, Inc. (Geumcheon-gu, Seoul, Republic of Korea). The genomes of the isolates (excluding MSHR8441) were also sequenced using the 454 Genome Sequencer FLX + instrument (454 Life Sciences, Branford, CT, USA). The sequence reads were quality-filtered and subsequently assembled via hybrid assembly using MIRA (17), followed by PAGIT (18), SSPACE version 2.0 (19), and GapFiller version 1.10 (20) polishing. MSHR8441 was assembled using Velvet version 1.2.10 (21) instead of MIRA but included the same quality improvement steps listed above. The contigs were reordered against *B. pseudomallei* MSHR1153 (GenBank accession numbers CP009271 and CP009272 [22]) or K96243 (GenBank accession numbers NC_006350 and NC_006351 [23]) using Mauve (24). The contiguous contigs were manually checked for synteny among reference genomes using BLAST and were stitched, if possible. A summary of the final genome assembly statistics is provided in Table 1.

These genomes will provide useful reference strains for use in

### TABLE 1 Statistics for the 5 *Burkholderia pseudomallei* draft genome sequences

<table>
<thead>
<tr>
<th>Strain</th>
<th>Alternate ID</th>
<th>Accession no.</th>
<th>Genome size (bp)</th>
<th>No. of contigs</th>
<th>N\textsubscript{500} (bp)</th>
<th>G+C content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCMRI_BP07</td>
<td>MSHR5651</td>
<td>JYBG000000000</td>
<td>7,767,989</td>
<td>139</td>
<td>227,078</td>
<td>67.6</td>
</tr>
<tr>
<td>QCMRI_BP13</td>
<td>MSHR8436</td>
<td>JYBH000000000</td>
<td>7,356,204</td>
<td>98</td>
<td>241,256</td>
<td>68.0</td>
</tr>
<tr>
<td>QCMRI_BP18</td>
<td>MSHR5662</td>
<td>JYBI000000000</td>
<td>7,391,892</td>
<td>83</td>
<td>755,997</td>
<td>68.0</td>
</tr>
<tr>
<td>QCMRI_BP28</td>
<td>MSHR8438</td>
<td>JYBJ000000000</td>
<td>7,565,815</td>
<td>161</td>
<td>398,485</td>
<td>67.7</td>
</tr>
<tr>
<td>QCMRI_BP32</td>
<td>MSHR8441</td>
<td>JYBK000000000</td>
<td>7,108,439</td>
<td>39</td>
<td>371,315</td>
<td>68.2</td>
</tr>
</tbody>
</table>
analyses of longitudinal isolates from the same patient and provide novel insights into the within-host evolution and adaptation of B. pseudomallei in the CF lung. More broadly, these genomes will be useful for identifying parallel evolutionary mechanisms with other Gram-negative pathogens affecting the CF lung.

Nucleotide sequence accession numbers. The genome accession numbers for the assemblies deposited in DDBJ/ENA/GenBank are listed in Table 1.

ACKNOWLEDGMENTS

We thank Mark Mayo and Vanessa Theobald for laboratory assistance, Roche Diagnostics Australia Pty Ltd, for kindly providing the 454 sequence data used in our study, and the Rebecca L. Cooper Medical Research Foundation for provision of the NanoDrop 2000 spectrophotometer.

This project was funded by the Australian National Health and Medical Research Council via award 1046812.

REFERENCES