Whole-Genome Assemblies of 16 Burkholderia pseudomallei Isolates from Rivers in Laos
Nicole Liechti, Rosalie Zimmermann, Jakob Zopfi, Matthew Robinson, Alain Pierret, Olivier Ribolzi, Sayaphet Rattanavong, Viengmon Davong, Paul Newton, Matthias Wittwer, et al.

To cite this version:
Nicole Liechti, Rosalie Zimmermann, Jakob Zopfi, Matthew Robinson, Alain Pierret, et al.. Whole-Genome Assemblies of 16 Burkholderia pseudomallei Isolates from Rivers in Laos. Microbiology Resource Announcements, American Society for Microbiology, 2021, 10 (4), pp.e01226-20. 10.1128/mra.01226-20. hal-03127000

HAL Id: hal-03127000
https://hal.sorbonne-universite.fr/hal-03127000
Submitted on 1 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Whole-Genome Assemblies of 16 *Burkholderia pseudomallei* Isolates from Rivers in Laos

Nicole Liechti,a,b Rosalie E. Zimmermann,c,d,e Jakob Zopf,d Matthew T. Robinson,c,f Alain Pierret,h Olivier Ribolzi,i Sayaphet Rattanavong,c Viengmon Davong,c Paul N. Newton,c,f Matthias Wittwer,a David A. B. Dancec,f,g

Spiez Laboratory, Swiss Federal Office for Civil Protection, Spiez, Switzerland
Interfaculty Bioinformatics Unit, Department of Biology, University of Bern, Bern, Switzerland
Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Microbiology Laboratory, Mahosot Hospital, Vientiane, Laos
Department of Environmental Sciences, Biogeochemistry, University of Basel, Basel, Switzerland
Department of Medical Microbiology, Amsterdam University Medical Centers, Amsterdam, The Netherlands
Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom
IEES-Paris (IRD, Sorbonne Universités, UPMC Univ Paris 06, CNRS, INRA, UPEC, c/o Department of Agricultural Land Management (DALaM), Vientiane, Laos
Géosciences Environnement Toulouse (GET), Université de Toulouse, IRD, CNRS, UPS, Toulouse, France

Nicole Liechti and Rosalie E. Zimmermann contributed equally to this work. Author order was determined by the focus of analyses (bioinformatics).

**ABSTRACT** We report 16 *Burkholderia pseudomallei* genomes, including 5 new multilocus sequence types, isolated from rivers in Laos. The environmental bacterium *B. pseudomallei* causes melioidosis, a serious infectious disease in tropical and subtropical regions. The isolates are geographically clustered in one clade from around Vientiane, Laos, and one clade from further south.

*B. pseudomallei* causes the human infectious disease melioidosis and is found in tropical and subtropical soils and freshwater (1). Survival and replication in various ecological niches and within hosts is possibly enabled by the large and highly variable accessory genome of *B. pseudomallei* (2, 3). Genome descriptions of *B. pseudomallei* isolates contribute to research on links between environment-associated and disease-associated genes of *B. pseudomallei* and their functions (3, 4).

We sequenced the genomes of 16 *B. pseudomallei* isolates from 14 filtered water samples and two sediment samples from rivers in Laos, cultured and confirmed as previously described (5). After storage at −80°C and pure culture on nutrient agar in air at 37°C for 24 h, genomic DNA was extracted using the Qiagen DNeasy blood and tissue kit and submitted to Microsynth AG (Balgach, Switzerland) for Nextera XT library preparation and sequencing using an Illumina NextSeq 500 instrument (paired-end [PE], 150-bp reads). Reads were quality trimmed using Trimmomatic 0.36 (slidingwindow:4:8, minlen:127) (6) and assembled using SPAdes 3.11.1 (-careful, -mismatch-correction, -k 21, 33, 55, 77, 99, 127 bp) (7). Pilon 1.22 (8) was applied to improve the quality of the draft assemblies. Scaffolds of <200 bp or with low coverage were removed. Finally, contaminants were removed manually using a BLAST search against the NCBI nucleotide database. The quality and completeness of the *de novo*-assembled genomes were accessed using BUSCO 3.0.1 (lineage, *Betaproteobacteria* odb9) (9), and basic assembly statistics were compared using QUAST 4.6.3 (10). The genomes were annotated automatically using the NCBI Prokaryotic Annotation Pipeline 4.11 (11). Default settings were used for all software unless otherwise specified. A summary of the assembly results is provided in Table 1. The 16 isolates were found to belong to 6 different sequence types using the multilocus sequence typing pipeline (12, 13), 5 of which were new. Sequence type 54 (ST54) (two isolates) was previously described and is common in neighboring Thailand.
<table>
<thead>
<tr>
<th>Isolate</th>
<th>River</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Raw read SRA no.</th>
<th>GenBank assembly accession no.</th>
<th>Sequence type</th>
<th>No. of contigs</th>
<th>Assembly length (Mbp)</th>
<th>GC content (%)</th>
<th>( N_{50} ) (kbp)</th>
<th>No. of paired reads (million)</th>
<th>Genome coverage (( \times ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>40S_S61</td>
<td>Mekong</td>
<td>15.11316</td>
<td>105.80506</td>
<td>SRR11097786</td>
<td>GCA_014713055.1</td>
<td>ST1787</td>
<td>186</td>
<td>7.15</td>
<td>68.19</td>
<td>119.6</td>
<td>4.8</td>
<td>203</td>
</tr>
<tr>
<td>43W_S62</td>
<td>Xe Bangnouan</td>
<td>16.00286</td>
<td>105.47903</td>
<td>SRR11097785</td>
<td>GCA_014713065.1</td>
<td>ST1801</td>
<td>205</td>
<td>7.17</td>
<td>68.14</td>
<td>107.1</td>
<td>4.7</td>
<td>183</td>
</tr>
<tr>
<td>43S_S63</td>
<td>Xe Bangnouan</td>
<td>16.00286</td>
<td>105.47903</td>
<td>SRR11097778</td>
<td>GCA_014713085.1</td>
<td>ST1787</td>
<td>176</td>
<td>7.15</td>
<td>68.19</td>
<td>128.2</td>
<td>4.4</td>
<td>195</td>
</tr>
<tr>
<td>44W_S64</td>
<td>Mekong</td>
<td>16.00421</td>
<td>105.42515</td>
<td>SRR11097777</td>
<td>GCA_014713025.1</td>
<td>ST1801</td>
<td>191</td>
<td>7.17</td>
<td>68.14</td>
<td>123.3</td>
<td>5.5</td>
<td>231</td>
</tr>
<tr>
<td>45W_S65</td>
<td>Xe Banghieng</td>
<td>16.09798</td>
<td>105.37699</td>
<td>SRR11097776</td>
<td>GCA_014713015.1</td>
<td>ST1815</td>
<td>184</td>
<td>7.15</td>
<td>68.15</td>
<td>126.5</td>
<td>4.6</td>
<td>194</td>
</tr>
<tr>
<td>46W_S66</td>
<td>Mekong</td>
<td>17.39898</td>
<td>104.80098</td>
<td>SRR11097774</td>
<td>GCA_014712945.1</td>
<td>ST1815</td>
<td>188</td>
<td>7.15</td>
<td>68.14</td>
<td>132.9</td>
<td>5.3</td>
<td>219</td>
</tr>
<tr>
<td>47W_S67</td>
<td>Xe Bangfai</td>
<td>17.07787</td>
<td>104.98503</td>
<td>SRR11097775</td>
<td>GCA_014712955.1</td>
<td>ST1801</td>
<td>209</td>
<td>7.17</td>
<td>68.15</td>
<td>119.7</td>
<td>5.3</td>
<td>221</td>
</tr>
<tr>
<td>49W_S68</td>
<td>Nam Kading</td>
<td>18.32559</td>
<td>104.00002</td>
<td>SRR11097773</td>
<td>GCA_014712965.1</td>
<td>ST1801</td>
<td>193</td>
<td>7.17</td>
<td>68.14</td>
<td>127.8</td>
<td>5.3</td>
<td>223</td>
</tr>
<tr>
<td>50W_S69</td>
<td>Nam Xan</td>
<td>18.39103</td>
<td>103.65572</td>
<td>SRR11097772</td>
<td>GCA_014712935.1</td>
<td>ST1852</td>
<td>192</td>
<td>7.24</td>
<td>68.1</td>
<td>138.5</td>
<td>5.0</td>
<td>208</td>
</tr>
<tr>
<td>51W_S70</td>
<td>Nam Ngiep</td>
<td>18.41756</td>
<td>103.60212</td>
<td>SRR11097771</td>
<td>GCA_014712915.1</td>
<td>ST1852</td>
<td>222</td>
<td>7.25</td>
<td>68.26</td>
<td>110.8</td>
<td>4.9</td>
<td>201</td>
</tr>
<tr>
<td>52W_S71</td>
<td>Nam Mang</td>
<td>18.37017</td>
<td>103.19838</td>
<td>SRR11097784</td>
<td>GCA_014712835.1</td>
<td>ST54</td>
<td>168</td>
<td>7.04</td>
<td>68.09</td>
<td>134.4</td>
<td>5</td>
<td>213</td>
</tr>
<tr>
<td>53W_S72</td>
<td>Nam Ngum</td>
<td>18.17874</td>
<td>103.05994</td>
<td>SRR11097783</td>
<td>GCA_014712895.1</td>
<td>ST54</td>
<td>182</td>
<td>7.04</td>
<td>68.27</td>
<td>126.7</td>
<td>4.4</td>
<td>188</td>
</tr>
<tr>
<td>58W_S73</td>
<td>Nam Ngum</td>
<td>18.20194</td>
<td>102.58694</td>
<td>SRR11097782</td>
<td>GCA_014712875.1</td>
<td>ST1852</td>
<td>184</td>
<td>7.25</td>
<td>68.1</td>
<td>120.9</td>
<td>4.7</td>
<td>195</td>
</tr>
<tr>
<td>60W_S74</td>
<td>Nam Sang</td>
<td>18.22297</td>
<td>102.14228</td>
<td>SRR11097781</td>
<td>GCA_014712825.1</td>
<td>ST1852</td>
<td>186</td>
<td>7.24</td>
<td>68.11</td>
<td>123</td>
<td>4.6</td>
<td>189</td>
</tr>
<tr>
<td>64W_S75</td>
<td>Mekong</td>
<td>17.97309</td>
<td>102.50404</td>
<td>SRR11097780</td>
<td>GCA_014712815.1</td>
<td>ST1869</td>
<td>212</td>
<td>7.3</td>
<td>68.1</td>
<td>97.6</td>
<td>4.8</td>
<td>197</td>
</tr>
<tr>
<td>91W_S76</td>
<td>Nam Ngum</td>
<td>18.3555</td>
<td>102.57198</td>
<td>SRR11097779</td>
<td>GCA_014712775.1</td>
<td>ST1869</td>
<td>194</td>
<td>7.21</td>
<td>68.11</td>
<td>87.9</td>
<td>4.1</td>
<td>171</td>
</tr>
</tbody>
</table>

\( ^a \) Data from reference 5.
To unravel the phylogenetic relationship of the isolates, we first constructed a core single nucleotide polymorphism genome alignment using Snippy 4.4.3 with *B. pseudo-mallei* MSHR4503 (15) as the reference. Then, we built a maximum likelihood tree using RAxML 8.2.11 (16) with a general time-reversible nucleotide substitution model including 1,000 bootstrap replicates (Fig. 1). The six main branches of the tree correspond to the sequence types and are geographically clustered in two different clades. One clade includes isolates from or around Vientiane, Laos (city and province), whereas the other consists of isolates from further south. The sediment isolate from Xe Bangnouan, Laos, is more closely related to the sediment isolate from the Mekong River than to the corresponding water isolate (Fig. 1). However, with relatively few samples taken at one point in time from rivers with large catchment areas, the interpretation of these clusters remains speculative. It is hoped that sequencing more isolates of *B. pseudomallei* from Laos will improve our understanding of the phylogeography of the organism within the country and enable comparisons to be made between clinical and environmental isolates.

**Data availability.** Illumina raw reads and genome assemblies were deposited at the NCBI and DDBJ/ENA/GenBank, respectively. The accession numbers are listed in Table 1. The isolates are linked to the respective sequence types on the PubMLST database.

**ACKNOWLEDGMENTS**

We are very grateful to the director and staff of the Microbiology Laboratory, Mahosot Hospital, to Bunthaphany Bouxouxiei, past director of Mahosot Hospital, to Bounnack Saysanasongkham, past director of the Department of Health Care, Ministry of Health, and to H. E. Bounkong Syhavong, Minister of Health, Lao PDR.

The project was funded by the U.S. Defense Threat Reduction Agency Cooperative Biological Engagement Program (contract HDTRA-16-C-0017).

**REFERENCES**


---

Volume 10 Issue 4 e01226-20 mra.asm.org

Downloaded from http://mra.asm.org on February 1, 2021 by guest