

Whole-Genome Assemblies of 16 Burkholderia pseudomallei Isolates from Rivers in Laos

Nicole Liechti, Rosalie E Zimmermann, Jakob Zopfi, Matthew T Robinson, Alain Pierret, Olivier Ribolzi, Sayaphet Rattanavong, Viengmon Davong, Paul Newton, Matthias Wittwer, et al.

▶ **To cite this version:**

Nicole Liechti, Rosalie E Zimmermann, Jakob Zopfi, Matthew T Robinson, Alain Pierret, et al.. Whole-Genome Assemblies of 16 Burkholderia pseudomallei Isolates from Rivers in Laos. Microbiology Resource Announcements, 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.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Whole-Genome Assemblies of 16 *Burkholderia pseudomallei* Isolates from Rivers in Laos

Nicole Liechti,[a,b] Rosalie E. Zimmermann,[c,d,e] Jakob Zopfi,[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. Dance[c,f,g]

[a]Spiez Laboratory, Swiss Federal Office for Civil Protection, Spiez, Switzerland
[b]Interfaculty Bioinformatics Unit, Department of Biology, University of Bern, Bern, Switzerland
[c]Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Microbiology Laboratory, Mahosot Hospital, Vientiane, Laos
[d]Department of Environmental Sciences, Biogeochemistry, University of Basel, Basel, Switzerland
[e]Department of Medical Microbiology, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[f]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[g]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom
[h]iEES-Paris (IRD, Sorbonne Universités, UPMC Univ Paris 06, CNRS, INRA, UPEC, c/o Department of Agricultural Land Management (DALaM), Vientiane, Laos
[i]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.

Burkholderia 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, *Betaprotebacteria* 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

Citation Liechti N, Zimmermann RE, Zopfi J, Robinson MT, Pierret A, Ribolzi O, Rattanavong S, Davong V, Newton PN, Wittwer M, Dance DAB. 2021. Whole-genome assemblies of 16 *Burkholderia pseudomallei* isolates from rivers in Laos. Microbiol Resour Announc 10:e01226-20. https://doi.org/10.1128/MRA.01226-20.

Editor David Rasko, University of Maryland School of Medicine

Copyright © 2021 Liechti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Nicole Liechti, nicole.liechti@babs.admin.ch, or Rosalie E. Zimmermann, r.e.zimmermann@amsterdamumc.nl.

Received 28 October 2020
Accepted 3 January 2021
Published 28 January 2021

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

TABLE 1 Characteristics and accession numbers of the genomes of 16 *B. pseudomallei* isolates from rivers in Laos

Isolate[a]	River[a]	Latitude[a]	Longitude[a]	Raw read SRA no.	GenBank assembly accession no.	Sequence type	No. of contigs	Assembly length (Mbp)	GC content (%)	N_{50} (kbp)	No. of paired reads (million)	Genome coverage (×)
40S_S61	Mekong	15.11316	105.80506	SRR11097786	GCA_014713055.1	ST1787	186	7.15	68.19	119.6	4.8	203
43W_S62	Xe Bangnouan	16.00286	105.47903	SRR11097785	GCA_014713065.1	ST1801	205	7.17	68.14	107.1	4.7	183
43S_S63	Xe Bangnouan	16.00286	105.47903	SRR11097778	GCA_014713085.1	ST1787	176	7.15	68.19	128.2	4.4	195
44W_S64	Mekong	16.00421	105.42515	SRR11097777	GCA_014713025.1	ST1801	191	7.17	68.14	123.3	5.5	231
45W_S65	Xe Banghieng	16.09798	105.37699	SRR11097776	GCA_014713015.1	ST1815	184	7.15	68.15	126.5	4.6	194
46W_S66	Mekong	17.39898	104.80098	SRR11097774	GCA_014712945.1	ST1815	188	7.15	68.14	132.9	5.3	219
47W_S67	Xe Bangfai	17.07787	104.98503	SRR11097775	GCA_014712955.1	ST1801	209	7.17	68.15	119.7	5.3	221
49W_S68	Nam Kading	18.32559	104.00002	SRR11097773	GCA_014712965.1	ST1801	193	7.17	68.14	127.8	5.3	223
50W_S69	Nam Xan	18.39103	103.65572	SRR11097772	GCA_014712935.1	ST1852	192	7.24	68.1	138.5	5.0	208
51W_S70	Nam Gniep	18.41756	103.60212	SRR11097771	GCA_014712915.1	ST1852	222	7.25	68.26	110.8	4.9	201
52W_S71	Nam Mang	18.37017	103.19838	SRR11097784	GCA_014712835.1	ST54	168	7.04	68.09	134.4	5	213
53W_S72	Nam Ngum	18.17874	103.05594	SRR11097783	GCA_014712895.1	ST54	182	7.04	68.27	126.7	4.4	188
58W_S73	Nam Ngum	18.20194	102.58669	SRR11097782	GCA_014712875.1	ST1852	184	7.25	68.1	120.9	4.7	195
60W_S74	Nam Sang	18.22297	102.14228	SRR11097781	GCA_014712825.1	ST1852	186	7.24	68.11	123	4.6	189
64W_S75	Mekong	17.97309	102.50404	SRR11097780	GCA_014712815.1	ST1869	212	7.3	68.1	97.6	4.8	197
91W_S76	Nam Ngum	18.3555	102.57198	SRR11097779	GCA_014712775.1	ST1869	194	7.21	68.11	87.9	4.1	171

[a] Data from reference 5.

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

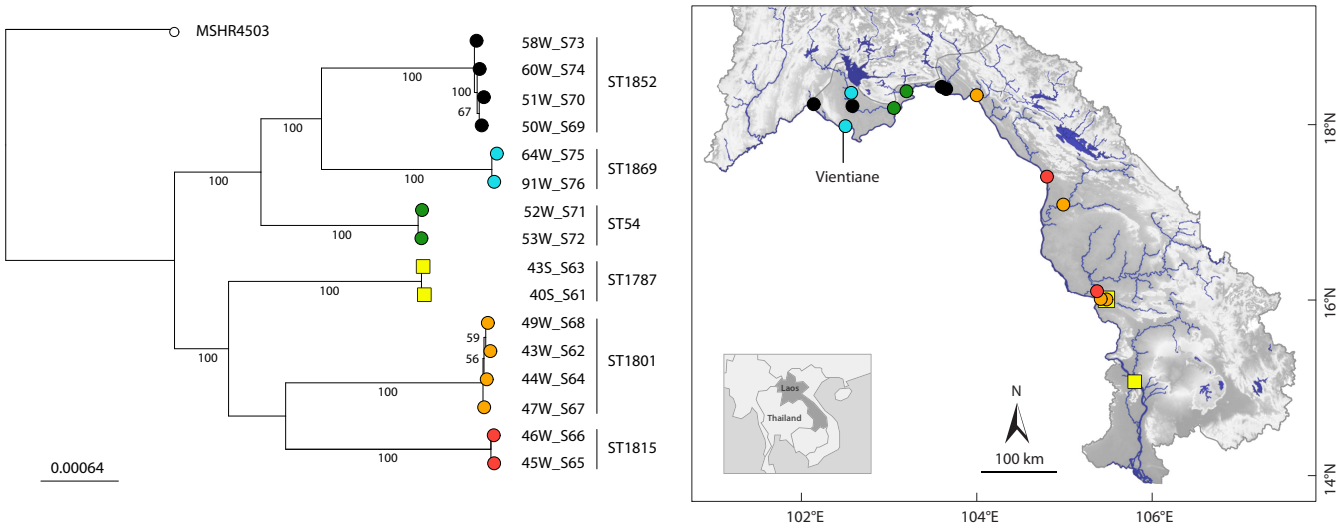


FIG 1 Maximum likelihood phylogeny and geographic locations of 16 environmental *B. pseudomallei* isolates. Clusters of isolates are displayed in different colors and correspond to sequence types (ST); circles represent water isolates, and squares represent sediment isolates. *B. pseudomallei* strain MSHR4503 from northern Australia (15) was used as the reference. The tree scale indicates changes per nucleotide; bootstrap values were calculated from 1,000 bootstrap replicates and are reported as percentages. The phylogenetic tree was visualized using Microreact (17), and the map was adapted from reference 5.

(14). 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. pseudomallei* 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 bootstraps (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 Bounthaphany Bounxouei, 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

1. Wiersinga WJ, Virk HS, Torres AG, Currie BJ, Peacock SJ, Dance DAB, Limmathurotsakul D. 2018. Melioidosis. Nat Rev Dis Primers 4:17107. https://doi.org/10.1038/nrdp.2017.107.
2. Holden MTG, Titball RW, Peacock SJ, Cerdeño-Tárraga AM, Atkins T, Crossman LC, Pitt T, Churcher C, Mungall K, Bentley SD, Sebaihia M, Thomson NR, Bason N, Beacham IR, Brooks K, Brown KA, Brown NF, Challis GL, Cherevach I, Chillingworth T, Cronin A, Crossett B, Davis P, DeShazer D, Feltwell T, Fraser A, Hance Z, Hauser H, Holroyd S, Jagels K, Keith KE, Maddison M, Moule S, Price C, Quail MA, Rabbinowitsch E, Rutherford K, Sanders M, Simmonds M, Songsivilai S, Stevens K, Tumapa S, Vesaratchavest M, Whitehead S, Yeats C, Barrell BG, Oyston PCF, Parkhill J. 2004. Genomic plasticity of the causative agent of melioidosis, Burkholderia

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

pseudomallei. Proc Natl Acad Sci U S A 101:14240–14245. https://doi.org/10.1073/pnas.0403302101.

3. Chewapreecha C, Mather AE, Harris SR, Hunt M, Holden MTG, Chaichana C, Wuthiekanun V, Dougan G, Day NPJ, Limmathurotsakul D, Parkhill J, Peacock SJ. 2019. Genetic variation associated with infection and the environment in the accidental pathogen Burkholderia pseudomallei. Commun Biol 2:428. https://doi.org/10.1038/s42003-019-0678-x.
4. Rachlin A, Mayo M, Webb JR, Kleinecke M, Rigas V, Harrington G, Currie BJ, Kaestli M. 2020. Whole-genome sequencing of Burkholderia pseudomallei from an urban melioidosis hot spot reveals a fine-scale population structure and localised spatial clustering in the environment. Sci Rep 10:5443. https://doi.org/10.1038/s41598-020-62300-8.
5. Zimmermann RE, Ribolzi O, Pierret A, Rattanavong S, Robinson MT, Newton PN, Davong V, Auda Y, Zopfi J, Dance DAB. 2018. Rivers as carriers and potential sentinels for Burkholderia pseudomallei in Laos. Sci Rep 8:8674. https://doi.org/10.1038/s41598-018-26684-y.
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
8. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
9. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.
10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
11. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
12. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.
13. Seemann T. mlst. https://github.com/tseemann/mlst.
14. Vesaratchavest M, Tumapa S, Day NPJ, Wuthiekanun V, Chierakul W, Holden MTG, White NJ, Currie BJ, Spratt BG, Feil EJ, Peacock SJ. 2006. Nonrandom distribution of Burkholderia pseudomallei clones in relation to geographical location and virulence. J Clin Microbiol 44:2553–2557. https://doi.org/10.1128/JCM.00629-06.
15. Johnson SL, Baker AL, Chain PS, Currie BJ, Daligault HE, Davenport KW, Davis CB, Inglis TJ, Kaestli M, Koren S, Mayo M, Merritt AJ, Price EP, Sarovich DS, Warner J, Rosovitz MJ. 2015. Whole-genome sequences of 80 environmental and clinical isolates of Burkholderia pseudomallei. Genome Announc 3:e01282-14. https://doi.org/10.1128/genomeA.01282-14.
16. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
17. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, Feil EJ, Holden MTG, Yeats CA, Grundmann H, Spratt BG, Aanensen DM. 2016. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genom 2:e01282-14. https://doi.org/10.1099/mgen.0.000093.

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