View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by OpenKnowledge@NAU

JCM
Journals.ASM.org

Whole-Genome Sequencing of *Burkholderia pseudomallei* Isolates from an Unusual Melioidosis Case Identifies a Polyclonal Infection with the Same Multilocus Sequence Type

Erin P. Price,[a,b] Derek S. Sarovich,[a,b] Linda Viberg,[a] Mark Mayo,[a] Mirjam Kaestli,[a] Apichai Tuanyok,[b*] Jeffrey T. Foster,[b*] Paul Keim,[b] Talima Pearson,[b] Bart J. Currie[a,c]

Global and Tropical Health Division, Menzies School of Health Research, Darwin, Northern Territory, Australia[a]; Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, Arizona, USA[b]; Department of Infectious Diseases and Northern Territory Medical Program, Royal Darwin Hospital, Darwin, Northern Territory, Australia[c]

Twelve *Burkholderia pseudomallei* isolates collected over a 32-month period from a patient with chronic melioidosis demonstrated identical multilocus sequence types (STs). However, whole-genome sequencing suggests a polyclonal infection. This study is the first to report a mixed infection with the same ST.

Burkholderia pseudomallei is a Gram-negative bacterium that causes the potentially fatal disease melioidosis (1). Melioidosis is the most common cause of community-acquired bacteremic pneumonia in the tropical Top End of the Northern Territory, Australia (2), with annual infection rates in recent years of up to 50 per 100,000 people (3). Multilocus sequence typing (MLST) is a widely adopted genotyping method for characterizing bacterial pathogens, including *B. pseudomallei* (4–6). MLST has been used to demonstrate that *B. pseudomallei* is among the most recombinogenic bacterial species studied to date, with a recombination-to-mutation ratio more than twice that of *Streptococcus pneumoniae* (7). Although useful for identifying strain relatedness, MLST is insensitive to genome-wide variation. To investigate genome-wide variation and better detect polyclonal infections, whole-genome sequencing is essential.

Case history. The Darwin Prospective Melioidosis Study (DPMS) has documented all known Top End melioidosis cases since October 1989 (2). The 103rd enrolled DPMS patient, P103, was a 49-year-old male who had a history of asthma and chronic lung disease and who had received intermittent therapy with oral prednisolone. Although P103 had positive *B. pseudomallei* serological titers since he was first tested in August 1990, serology itself is not an accurate diagnostic tool, particularly in regions where melioidosis is endemic; culture confirmation is required for a definitive melioidosis diagnosis (8). Given his positive serology, P103's sputum was repeatedly tested for *B. pseudomallei*, but it remained culture negative (9, 10) until September 1994, when *B. pseudomallei* isolate MSHR338 was obtained. Upon diagnosis, standard therapy comprising 2 weeks of intravenous ceftazidime followed by 3 months of oral doxycycline was commenced. Despite an improvement in symptoms, the sputum was again culture positive in December 1994. An additional round of intravenous ceftazidime followed by oral therapy with various combinations of doxycycline, trimethoprim-sulfamethoxazole, and chloramphenicol was administered. However, P103 still remained culture positive, with isolates retrieved from sputum or throat specimens 4, 7, 15, 21, and 30 months after obtaining MSHR338 (Table 1). All samples collected after March 1997 were *B. pseudomallei* negative. MIC testing of MSHRs 338 and 346 demonstrated that these isolates remained sensitive to ceftazidime, doxycycline, trimethoprim-sulfamethoxazole, and chloramphenicol. From the 888 melioidosis cases enrolled in the DPMS to August 2014, only 1 other case, P314, demonstrated long-term *B. pseudomallei* persistence. P314 has ongoing *B. pseudomallei* respiratory tract colonization that was first diagnosed in 2000 (11); in contrast, *B. pseudomallei* was eventually eradicated in P103 32 months after the initial diagnosis and the commencement of therapy.

MLST analysis. MLST was initially performed on seven isolates derived from six clinical specimens over a 32-month period (Table 1). DNA was extracted from single purified colonies as previously detailed (12). All seven isolates were sequence type 243 (ST-243), which has only been identified in isolates from this patient (http://bpseudomallei.mlst.net/).

Genomic analysis. Illumina GAIIx or HiSeq 2000 (Illumina, Inc., San Diego, CA) whole-genome sequencing was performed on 12 P103 *B. pseudomallei* isolates, including 5 isolates derived from three clinical specimens to more thoroughly sample within-host *B. pseudomallei* diversity (Table 1). Comparative genomic analyses were performed using SPANDx v2.3 (13), and maximum parsimony phylogenetic reconstruction of orthologous core single-nucleotide polymorphism (SNP) variants was carried out using PAUP* v4.0b (14). 454 GS FLX+ sequencing (454 Life Sciences, Branford, CT) was also performed on MSHR338. The 454 reads

Received 4 September 2014 Returned for modification 2 October 2014 Accepted 9 October 2014

Accepted manuscript posted online 22 October 2014

Citation Price EP, Sarovich DS, Viberg L, Mayo M, Kaestli M, Tuanyok A, Foster JT, Keim P, Pearson T, Currie BJ. 2015. Whole-genome sequencing of *Burkholderia pseudomallei* isolates from an unusual melioidosis case identifies a polyclonal infection with the same multilocus sequence type. J Clin Microbiol 53:282–286. doi:10.1128/JCM.02560-14.

Editor: D. J. Diekema

Address correspondence to Erin P. Price, erin.price@menzies.edu.au.

* Present address: Apichai Tuanyok, Department of Tropical Medicine, Medical Microbiology and Pharmacology, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii, USA; Jeffrey T. Foster, Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, New Hampshire, USA.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JCM.02560-14

Downloaded from http://jcm.asm.org/ on October 14, 2015 by NORTHERN ARIZONA UNIV

TABLE 1 Bacterial strains collected from P103

Strain	Collection date
MSHR338	September 1994
MSHR338-5[a] (MSHR338)	September 1994
MSHR346A[b]	January 1995
MSHR346B[b]	January 1995
MSHR376	April 1995
MSHR2844[a] (MSHR376)	April 1995
MSHR2845[a] (MSHR376)	April 1995
MSHR391	December 1995
MSHR443	June 1996
MSHR487	March 1997
MSHR2848[a] (MSHR487)	March 1997
MSHR2849[a] (MSHR487)	March 1997

[a] An additional subculture from the original stock (denoted in parentheses) to identify mixed genotypes.
[b] Morphological variants that were observed within the same clinical specimen.

were combined with Illumina data for *de novo* genome assembly, as previously described (11), and checked for errors with iterative correction of reference nucleotides (iCORN2) (15) using the Illumina reads. The final MSHR338 hybrid assembly (GenBank accession no. ATJY00000000.1) is in 59 high-quality contigs totaling 7,317,227 bp.

Phylogenetic reconstruction of the P103 genomes using default (i.e., permissive) SNP density filtering parameters in SPANDx (whereby only regions containing ≥3 SNPs within 10 bp are excluded [16]) demonstrated that MSHRs 346A, 346B, 376, 391, 443, 2845, and 2849 were identical. MSHRs 487, 2844, and 2848 each differed by one SNP (Fig. 1A). In contrast, 842 SNPs separated MSHR338 from the dominant genotype. No large deletions (>50 bp) were observed in any P103 strains compared with MSHR338, although 36 unique insertions or deletions (indels) were found in MSHR338 (Table 2). No indels were found among the other 11 strains. Given the relatively large genetic distance between MSHR338 and the subsequent P103 isolates, the potential for additional genotypes within the original MSHR338 culture stock was explored further. This stock was plated onto chocolate agar, and 10 individual colonies were screened using allele-specific real-time PCR (17) interrogation of an SNP at position 332 of the MSHR338 Seq0033 contig, which differentiates MSHR338 from other P103 isolates. All 10 colonies matched the dominant genotype. One colony (MSHR338–5) was fully sequenced (Table 1) and was identical to the dominant genotype, confirming the presence of this genotype in the original clinical specimen.

Given the ST-243 diversity in P103, the relative roles of recombination and mutation were investigated. A recombination filtering parameter (i.e., excluding regions with ≥3 SNPs within 300 bp) was first applied to the 12 P103 genomes. Using this filter, the majority of SNP differences (n = 778; 94%) in MSHR338 were removed (Fig. 1B). Closer examination of these SNPs in Integrative Genomics Viewer 2.3.34 (18) showed that the vast majority (99.9%) of SNPs in MSHR338 were colocated within a 1.3-Mbp region (Fig. 2). These results strongly suggest that one or more recent recombination events led to the ST-243 diversity in P103. Given that no other isolates with similar variants were found in P103, this recombination event probably occurred prior to infection, although the possibility of within-host recombination with an unrelated ST that was not sampled or that became extinct cannot be ruled out.

To the best of our knowledge, this study is the first to report a polyclonal infection with the same ST, and it demonstrates that MLST can be insensitive for detecting polyclonality. Two previous studies of mixed *B. pseudomallei* infections using various genotyping methods estimated a *B. pseudomallei* polyclonal infection rate of between 1.5 and 28% (19, 20); however, the true rate remains unknown. Within-host evolution in P103 is unlikely given the relatively low diversity in other reported melioidosis

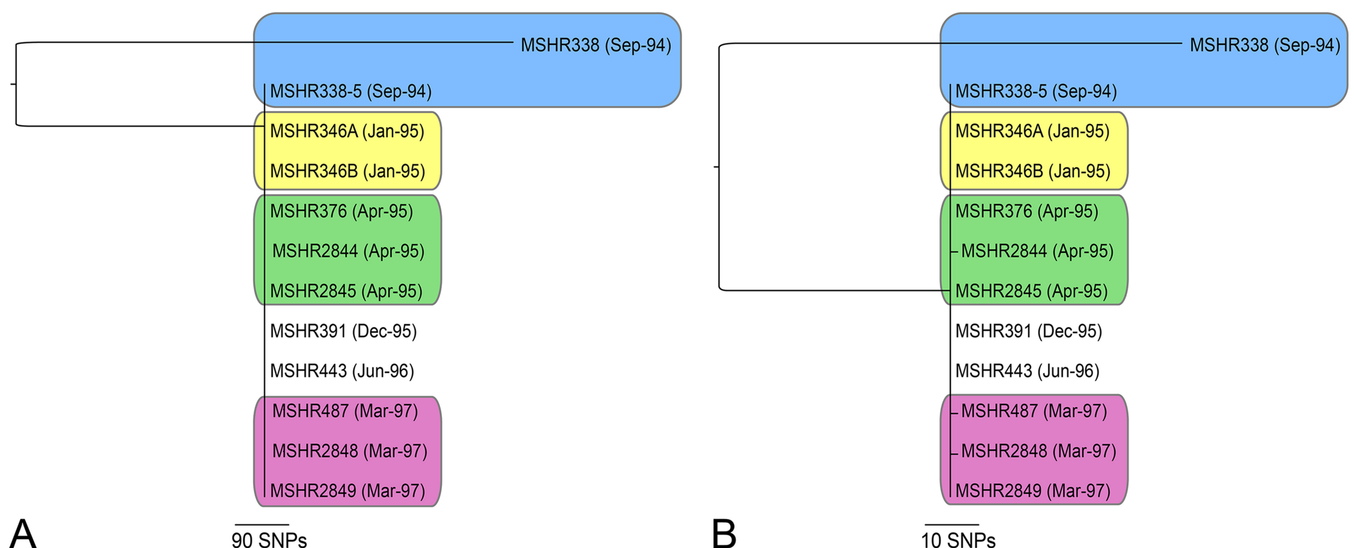

FIG 1 Maximum parsimony phylogenetic analyses of P103 *Burkholderia pseudomallei* isolates over a 32-month infection. All isolates are ST-243; however, MSHR338 is an outlier according to whole-genome sequencing. Colored shading indicates isolates that were derived from the same clinical specimen. (A) Default single-nucleotide polymorphism (SNP) density filtering in SPANDx (excluding regions with ≥3 SNPs per 10 bp). Single SNPs in MSHR487, MSHR2844, and MSHR2489 are present but not visible due to scale. (B) Moderate recombination SNP filtering in SPANDx (excluding regions with ≥3 SNPs per 300 bp). Only SNPs on the MSHR338 branch were removed with the recombinogenic filter. SNPs in MSHR487, MSHR2844, and MSHR2489 are visible. The consistency index for both trees is 1.0.

Downloaded from http://jcm.asm.org/ on October 14, 2015 by NORTHERN ARIZONA UNIV

TABLE 2 Indels between MSHR338 and other P103 strains

Location[a]	Protein	MSHR338, others[b]	Effect	Amino acid change
M218_20600: Seq0023, 74246	Methyltransferase	GGCC, G	Upstream	
M218_20665: Seq0023, 99625	Polyketide synthase	A, ACCGAAGCGACCGAAGTGG	Insertion	T3810 → TEATEVA
M218_20685: Seq0024, 19320	Polyketide synthase	ACGG, A	Deletion	DG1463 → D
Seq0025, 264		CTT, C	Intergenic	
Seq0026, 155		G, GC	Intergenic	
Seq0026, 19908		C, CG	Intergenic	
Seq0026, 23627		AAGTGC, A	Intergenic	
Seq0026, 23632		CGAAGTGG, C	Intergenic	
Seq0026, 23634		AAG, A	Intergenic	
Seq0026, 24939		G, GGCC	Intergenic	
M218_20910: Seq0026, 40186	2-Keto-4-pentenoate hydratase	CCG, C	Frameshift	P110: 254 → 257 amino acids
Seq0026, 40687		CCG, C	Intergenic	
Seq0026, 40707		GCAT, G	Intergenic	
Seq0026, 46450		CGCGCG, C	Intergenic	
Seq0026, 48033		G, GT	Intergenic	
M218_20940: Seq0026, 49266	Hypothetical protein	CGGCGCGCGGGCGCGCG, C	Deletion and frameshift	PARPRA140: 188 → 204 amino acids
M218_20945: Seq0026, 50476	Membrane protein	TGGCGC, T	Upstream	
M218_21440: Seq0026, 166557	Serine transporter	GTC, G	Frameshift	
M218_21485: Seq0026, 181817	Nonribosomal peptide synthase	GTCGACA, G	Deletion	DVD576 → D
M218_21485: Seq0026, 181833	Nonribosomal peptide synthase	ACATCG, A	Frameshift	VDV571: 1354 → 1352 amino acids
M218_21530: Seq0026, 198738	Entericidin	C, CG	Upstream	
Seq0026, 204399		AG, A	Intergenic	
Seq0026, 204633		TTG, T	Intergenic	
M218_21560: Seq0026, 205821	Transposase	CGTCA, C	Upstream	
Seq0027, 10298		G, GT	Intergenic	
M218_21660: Seq0027, 22185	3-Demethyl-ubiquinone-9 3-methyltransferase	G, GA	Upstream	
M218_22155: Seq0027, 148292	Aminopeptidase	C, CCCCGTGTGT	Insertion	G282 → GTHG
Seq0028, 89723		CG, C	Intergenic	
Seq0028, 90812		T, TG	Intergenic	
Seq0028, 90857		GC, G	Intergenic	
M218_22960: Seq0028, 93133	Membrane protein	TGC, T	Frameshift	AQ212: 481 → 538 amino acids
Seq0030, 121		AGGCGC, A	Intergenic	
M218_23840: Seq0031, 84459	Hypothetical protein	G, GCCGCATCACCCGCATCAC	Insertion	A822 → GDAGDAA
M218_23840: Seq0031, 85558	Hypothetical protein	C, CCGGCGTTGT	Insertion	P455 → PTTP
M218_23840: Seq0031, 87016	Hypothetical protein	GCC, G	Upstream	
M218_23880: Seq0031, 94561	Hypothetical protein	T, TA	Upstream	

[a] Nucleotide position in the MSHR338 genome (GenBank accession no. ATJY00000000).
[b] All other strains from P103 besides MSHR338.

Downloaded from http://jcm.asm.org/ on October 14, 2015 by NORTHERN ARIZONA UNIV

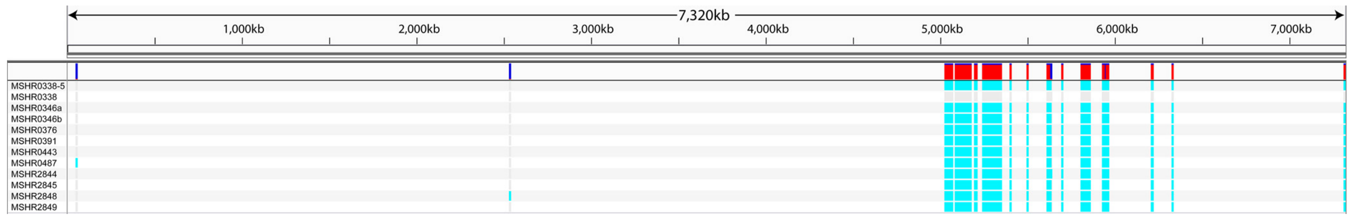

FIG 2 Contribution of recombination to ST-243 diversity in P103. Comparison of MSHR338 relative to other P103 isolates shows that 99.9% of SNPs found in MSHR338 are located within a single ~1.3-Mbp region on chromosome 2, indicating that one large (or several small) recombination events likely led to the recent divergence of MSHR338 from other P103 isolates. Two single-nucleotide polymorphism (SNP) mutations, one each in MSHR0487 and MSHR2848, are also visible. MSHR338 contigs were reordered relative to MSHR305 (23) using progressiveMauve (24) prior to visualization in Integrative Genomics Viewer (18).

Downloaded from http://jcm.asm.org/ on October 14, 2015 by NORTHERN ARIZONA UNIV

cases (20–22), including a 12-year chronic-carriage case (11). We therefore conclude that the genotypes recovered from P103 resulted from inoculation with two or more genotypes. The MSHR338 genotype was only detected in the first culture specimen from this patient and at low frequency (<10%) and might have been missed with less intensive sampling. Increased adoption of whole-genome sequencing, in combination with greater sampling efforts, will unveil further instances of polyclonality in clinical specimens. We speculate that polyclonal infection may have led to the difficulties in pathogen eradication observed in this case. However, a greater understanding of pathogen evolution and adaptation within the human host is needed to better inform the precise clinical ramifications of polyclonal infections.

Nucleotide sequence accession number. The sequence determined in this study has been deposited in GenBank under accession no. ATJY00000000.1.

ACKNOWLEDGMENTS

We thank our colleagues in the microbiology laboratory at Royal Darwin Hospital for expertise in identifying *B. pseudomallei* isolates. We also thank Leisha Richardson (Menzies School of Health Research), James Cook, and Kevin Drees (Northern Arizona University) for laboratory and genomics assistance.

This study was funded by the Australian National Health and Medical Research Council via project grant award 1046812 and by the U.S. Department of Homeland Security S&T CB Division Bioforensics R&D Program (HSHQDC-10-C-00139).

MSHR338 454 sequencing data were kindly provided by Henry S. Gibbons at the U.S. Army Edgewood Chemical Biological Center.

We declare no conflicts of interest.

REFERENCES

1. **Wiersinga WJ, Currie BJ, Peacock SJ.** 2012. Melioidosis. N Engl J Med **367:**1035–1044. http://dx.doi.org/10.1056/NEJMra1204699.
2. **Currie BJ, Fisher DA, Howard DM, Burrow JN, Lo D, Selva-Nayagam S, Anstey NM, Huffam SE, Snelling PL, Marks PJ, Stephens DP, Lum GD, Jacups SP, Krause VL.** 2000. Endemic melioidosis in tropical northern Australia: a 10-year prospective study and review of the literature. Clin Infect Dis **31:**981–986. http://dx.doi.org/10.1086/318116.
3. **Parameswaran U, Baird RW, Ward LM, Currie BJ.** 2012. Melioidosis at Royal Darwin Hospital in the big 2009–2010 wet season: comparison with the preceding 20 years. Med J Aust **196:**345–348. http://dx.doi.org/10.5694/mja11.11170.
4. **Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A **95:**3140–3145. http://dx.doi.org/10.1073/pnas.95.6.3140.
5. **Vesaratchavest M, Tumapa S, Day NP, Wuthiekanun V, Chierakul W, Holden MT, 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. http://dx.doi.org/10.1128/JCM.00629-06.
6. **Godoy D, Randle G, Simpson AJ, Aanensen DM, Pitt TL, Kinoshita R, Spratt BG.** 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. J Clin Microbiol **41:**2068–2079. http://dx.doi.org/10.1128/JCM.41.5.2068-2079.2003.
7. **Pearson T, Giffard P, Beckstrom-Sternberg S, Auerbach R, Hornstra H, Tuanyok A, Price EP, Glass MB, Leadem B, Beckstrom-Sternberg JS, Allan GJ, Foster JT, Wagner DM, Okinaka RT, Sim SH, Pearson O, Wu Z, Chang J, Kaul R, Hoffmaster AR, Brettin TS, Robison RA, Mayo M, Gee JE, Tan P, Currie BJ, Keim P.** 2009. Phylogeographic reconstruction of a bacterial species with high levels of lateral gene transfer. BMC Biol **7:**78. http://dx.doi.org/10.1186/1741-7007-7-78.
8. **Dance DAB.** 2002. Melioidosis. Curr Opin Infect Dis **15:**127–132. http://dx.doi.org/10.1097/00001432-200204000-00005.
9. **Cheng AC, Wuthiekanun V, Limmathurosakul D, Wongsuvan G, Day NP, Peacock SJ.** 2006. Role of selective and nonselective media for isolation of *Burkholderia pseudomallei* from throat swabs of patients with melioidosis. J Clin Microbiol **44:**2316. http://dx.doi.org/10.1128/JCM.00231-06.
10. **Ashdown LR.** 1979. An improved screening technique for isolation of *Pseudomonas pseudomallei* from clinical specimens. Pathology **11:**293–297. http://dx.doi.org/10.3109/00313027909061954.
11. **Price EP, Sarovich DS, Mayo M, Tuanyok A, Drees KP, Kaestli M, Beckstrom-Sternberg SM, Babic-Sternberg JS, Kidd TJ, Bell SC, Keim P, Pearson T, Currie BJ.** 2013. Within-host evolution of *Burkholderia pseudomallei* over a twelve-year chronic carriage infection. mBio **4**(4)**:**e00388–13. http://dx.doi.org/10.1128/mBio.00388-13.
12. **Currie BJ, Gal D, Mayo M, Ward L, Godoy D, Spratt BG, LiPuma JJ.** 2007. Using BOX-PCR to exclude a clonal outbreak of melioidosis. BMC Infect Dis **7:**68. http://dx.doi.org/10.1186/1471-2334-7-68.
13. **Sarovich DS, Price EP.** 2014. SPANDx: a genomics pipeline for comparative analysis of large haploid whole genome re-sequencing datasets. BMC Res Notes **7:**618. http://dx.doi.org/10.1186/1756-0500-7-618.
14. **Swofford DL.** 2002. PAUP*: phylogenetic analysis using parsimony (and other methods), version 4.0 beta. Sinauer Associates, Sunderland, MA.
15. **Otto TD, Sanders M, Berriman M, Newbold C.** 2010. Iterative correction of reference nucleotides (iCORN) using second generation sequencing technology. Bioinformatics **26:**1704–1707. http://dx.doi.org/10.1093/bioinformatics/btq269.
16. **DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ.** 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet **43:**491–498. http://dx.doi.org/10.1038/ng.806.
17. **Germer S, Holland MJ, Higuchi R.** 2000. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. Genome Res **10:**258–266. http://dx.doi.org/10.1101/gr.10.2.258.
18. **Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP.** 2011. Integrative genomics viewer. Nat Biotechnol **29:**24–26. http://dx.doi.org/10.1038/nbt.1754.
19. **Pitt TL, Trakulsomboon S, Dance DA.** 2007. Recurrent melioidosis: possible role of infection with multiple strains of *Burkholderia pseu-*

domallei. J Clin Microbiol **45:**680–681. http://dx.doi.org/10.1128/JCM.01974-06.

20. **Limmathurotsakul D, Wuthiekanun V, Chantratita N, Wongsuvan G, Thanwisai A, Biaklang M, Tumapa S, Lee S, Day NP, Peacock SJ.** 2007. Simultaneous infection with more than one strain of *Burkholderia pseudomallei* is uncommon in human melioidosis. J Clin Microbiol **45:**3830–3832. http://dx.doi.org/10.1128/JCM.01297-07.
21. **Hayden HS, Lim R, Brittnacher MJ, Sims EH, Ramage ER, Fong C, Wu Z, Crist E, Chang J, Zhou Y, Radey M, Rohmer L, Haugen E, Gillett W, Wuthiekanun V, Peacock SJ, Kaul R, Miller SI, Manoil C, Jacobs MA.** 2012. Evolution of *Burkholderia pseudomallei* in recurrent melioidosis. PLoS One **7:**e36507. http://dx.doi.org/10.1371/journal.pone.0036507.
22. **Limmathurotsakul D, Holden MT, Coupland P, Price EP, Chantratita N, Wuthiekanun V, Amornchai P, Parkhill J, Peacock SJ.** 2014. Microevolution of *Burkholderia pseudomallei* during an acute infection. J Clin Microbiol **52:**3418–3421. http://dx.doi.org/10.1128/JCM.01219-14.
23. **Stone JK, Johnson SL, Bruce DC, Detter JC, Mayo M, Currie BJ, Gelhaus HC, Keim P, Tuanyok A.** 2013. Complete genome sequence of the encephalomyelitic *Burkholderia pseudomallei* strain MSHR305. Genome Announc **1**(4):e00656–13. http://dx.doi.org/10.1128/genomeA.00656-13.
24. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.

Downloaded from http://jcm.asm.org/ on October 14, 2015 by NORTHERN ARIZONA UNIV