Title: Using land runoff to survey the distribution and genetic diversity of *Burkholderia pseudomallei* in Vientiane, Laos

Running title: *B. pseudomallei* presence and diversity in Laos

Authors: Audrey Rachlin[a], Manophab Luangraj[b], Mirjam Kaestli[a,c], Sayaphet Rattanavong[b], Phonelavanh Phoumin[b], Jessica R. Webb[a], Mark Mayo[a], Bart J Currie[a,d], and David A. B. Dance[b,e,f#]

Author affiliations:

a. Global and Tropical Health Division, Menzies School of Health Research, Charles Darwin University, Darwin, Northern Territory, Australia
b. Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Microbiology Laboratory, Mahosot Hospital, Vientiane, Laos
c. Research Institute for the Environment and Livelihoods, Charles Darwin University, Darwin, Northern Territory 0811, Australia.
d. Department of Infectious Diseases and Northern Territory Medical Program, Royal Darwin Hospital, Darwin, Northern Territory, Australia
e. Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, Old Road Campus, University of Oxford, Oxford, UK
f. Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK

Address correspondence to David A. B. Dance
Email: david.d@tropmedres.ac

Abstract

Melioidosis is a disease of significant public health importance that is being increasingly recognized globally. The majority of cases arise through direct percutaneous exposure to its etiological agent, *Burkholderia pseudomallei*. In the Lao People's Democratic Republic (Laos), the presence and environmental distribution of *B. pseudomallei* are not well characterized, though recent epidemiological surveys of the bacterium have indicated that *B. pseudomallei* is widespread throughout the environment in the center and south of the country and that rivers can act as carriers and potential sentinels for the bacterium. The spatial and genetic distribution of *B. pseudomallei* within Vientiane Capital, from where the majority of cases diagnosed to date have originated, remains an important knowledge gap. We sampled surface runoff from drain catchment areas throughout urban Vientiane to determine the presence and local population structure of the bacterium. *B. pseudomallei* was detected in drainage areas throughout the capital, indicating it is widespread in the environment and that exposure rates in urban Vientiane are likely more frequent than previously thought. Whole-genome comparative analysis demonstrated that Lao *B. pseudomallei* isolates are highly genetically diverse, suggesting the bacterium is well-established and not a recent introduction. Despite the wide genome diversity, one environmental survey isolate was highly genetically related to a Lao melioidosis patient isolate collected 13 years prior to the study. Knowledge gained from this study will augment understanding of *B. pseudomallei* phylogeography in Asia and enhance public health awareness and future implementation of infection control measures within Laos.

Importance

The environmental bacterium *B. pseudomallei* is the etiological agent of melioidosis, a tropical disease with one model estimating a global annual incidence of 165,000 cases and 89,000 deaths. In the Lao People's Democratic Republic (Laos), the environmental distribution and population structure of *B. pseudomallei* remain relatively undefined, particularly in Vientiane Capital from where most diagnosed cases have originated. We used

surface runoff as a proxy for *B. pseudomallei* dispersal in the environment and performed whole-genome sequencing (WGS) to examine the local population structure. Our data confirmed that *B. pseudomallei* is widespread throughout Vientiane and that surface runoff might be useful for future environmental monitoring of the bacterium. *B. pseudomallei* isolates were also highly genetically diverse, suggesting the bacterium is well-established and endemic in Laos. These findings can be used to improve awareness of *B. pseudomallei* in the Lao environment and demonstrates the epidemiological and phylogeographical insights that can be gained from WGS.

84 **Introduction**

85 Melioidosis is a serious disease of humans and animals caused by the environmental Gram-
86 negative sapronotic bacterium *Burkholderia pseudomallei*. Infection results from inoculation,
87 inhalation or ingestion of *B. pseudomallei*, and is fatal in 10-40% of human cases (1, 2).
88 Melioidosis was first reported in a patient in the Lao People's Democratic Republic (Laos) in
89 1999 (3) and since that case, 1690 culture-positive Lao melioidosis patients have been
90 confirmed by the Microbiology Laboratory of Mahosot Hospital in Vientiane as part of an
91 ongoing prospective study (unpublished data)(4). While the infection has now been
92 established as being highly endemic in Laos, the true burden of melioidosis and the
93 environmental distribution of *B. pseudomallei* remain relatively undefined.
94
95 In soil, *B. pseudomallei* is recognized as being spatially heterogeneously distributed on
96 both broad and localized geographical scales (5). This restricted geographical distribution
97 has resulted in a robust global biogeographic structure, with genetic populations being highly
98 spatially clustered in the environment despite high levels of gene recombination and
99 sequence type (ST) diversity (6-8). Whole-genome sequencing (WGS) has facilitated the
100 examination of genetic populations of the bacterium on a fine-scale (6, 9-11) and has
101 revealed large-scale geographical partitioning between Australian and Southeast Asian
102 isolates as well as highly localized genetic spatial clustering (6, 8, 12, 13).
103
104 The spatial heterogeneity of isolates means that the use of random soil sampling to establish
105 the bacterial presence in a region can often be indeterminate and imprecise (5, 14). It has
106 been suggested that the identification of new environments endemic for melioidosis may be
107 effectively determined by analyzing catchment points along the water column, including
108 groundwater and surface runoff areas (14-16). Since stormwater is known to capture and
109 leach what is in the land, including particulates, contaminants and bacteria, it is thought that
110 it may provide a good indication of *B. pseudomallei* distribution within a catchment, as the
111 bacterium is able to disperse along the water table and via drainage lines (17, 18).

Moreover, direct sampling of the water column and surface water can also provide an indication of the associations between environmental physico-chemical factors within a catchment (14).

In Laos, recent surveys have shown that water may be a significant reservoir and transport vehicle for *B. pseudomallei* (14, 15, 19). In one recently published study, the bacterium was isolated in 57% of samples collected during the rainy season from the Mekong river and its tributaries in the center and south of the country (15), and it has also been detected in surface water and catchment areas in Salavan province in the south of the country (14, 19).

High levels of *B. pseudomallei* have also been isolated in groundwater and groundwater seeps in both Townsville (16) and Darwin, Australia (13) and groundwater isolates have been linked to clinical isolates using molecular typing (16). However, the extent to which groundwater and seasonal runoff are contaminated with *B. pseudomallei* and might contribute to melioidosis in Laos have not yet been assessed and there is limited knowledge about the sequence type (ST) distribution and genomic diversity of isolates. Moreover, very few environmental surveys of the bacterium have been undertaken in Vientiane Capital, where over 10% of the Lao population currently resides and where just over half (54.6%) of more than 1359 culture-confirmed Lao melioidosis patients between 1999 and 2017 reported living (4). As most of the culture-confirmed melioidosis patients from Mahosot Hospital reside in Vientiane, we surveyed surface runoff at drainage catchment areas across the city. Whole-genome sequencing (WGS) and large-scale comparative genomics were performed on cultured isolates to examine the phylogenetic relatedness and population structure of *B. pseudomallei* in Laos to improve knowledge of genotype diversity. This has important global relevance given the substantial numbers of undetected melioidosis cases and deaths predicted to occur annually throughout Laos and Southeast Asia (20).

Materials and methods

Study sites and sample collection

Vientiane Capital is located along the southern edge of Vientiane Plain and is situated on the left bank of Mekong River. The topography is generally flat, with elevation varying no more than 164 to 175 metres above sea level (21). Forty drain sites were selected across five urban districts of Vientiane Capital encompassing an area of approximately 100km^2: Chanthabuly, Sisattanak, Xaysetha, Xaythany, and Sikhottabong districts (17.9°N, 102.6°S) (**Figures 1,2**). Sites were selected based on their accessibility, including proximity to the road and whether they were unfenced and uncovered. The environmental sources of drain water primarily consisted of surface land runoff from stormwater and irrigation and drains varied in their patterns of flow, shading and lining. Informed oral consent was obtained from landowners and written permission was obtained from the relevant authorities before commencement of sampling.

Samples were collected during the Lao rainy season in late June-July 2018. Eight samples consisting of both soil and water were collected at each of the 40 sites. Three water samples were collected inside each drain line using sterile one liter bottles, totalling 120 samples. Five soil samples were also simultaneously collected along adjacent drainage embankment areas (200 samples total), each spaced 10m apart (5).

On-site physico-chemical measurements were analyzed for each sample collected. For water samples this consisted of: nitrate (Horiba LAQUAtwin NO3-11), temperature, pH, electrical conductivity (EC), dissolved oxygen (DO), redox potential (ORP), water turbidity and total dissolved solids (TDS) using a portable multiparameter field probe (Hanna Instrument HI9829). Post-sampling *E. coli* and coliform counts were performed on all water samples at the Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU) laboratory in Vientiane Capital (3M Petrifilm). Additionally, drain lining (concrete lined or

unlined), degree of shading, and geographical coordinates (Garmin eTrex 30) were recorded at each survey location.

Sample processing and confirmation

Water samples

Water samples were processed in a Biosafety level 2 laboratory (BSL2) facility in the LOMWRU microbiology laboratory. 500ml of water was filtered in duplicate through 0.2 µm pore size, 47 mm diameter cellulose acetate filters (Merck & Co.) using an electrical pump. To detect *B. pseudomallei* on water filters and in sediment, we applied two independent methods: conventional culture techniques and PCR after an enrichment step.

Conventional culture

One filter was placed in 30ml Ashdown broth containing colistin (50mg/L) and incubated at 37°C. At 48 hours and seven days post-enrichment, 10µL and 100µL of broth was plated onto Ashdown's agar with gentamicin (8 mg/L) and incubated at 37°C aerobically for two days. All colonies resembling *B. pseudomallei* were sub-cultured onto Ashdown's agar. DNA from suspected colonies and from sweeps of bacterial growth on all plates was extracted using 10% Chelex-100 resin (22) and *B. pseudomallei* was confirmed in-house using the TTS1 real-time PCR assay targeting a signature 115-bp segment within the bacterial type three secretion system 1 (TTS1) gene (23).

Confirmed *B. pseudomallei* isolates were stored on Tryptone Soya Broth (TSB) agar slopes in 2ml screwcap tubes, incubated at 37°C for 48h and stored at room temperature until being shipped to Menzies School of Health Research (Menzies), Darwin, Australia.

Direct PCR following enrichment

For direct extraction, one water filter per sample was placed in 30ml Ashdown broth containing colistin as described above and shaken at 220rpm in a 37°C shaking incubator.

196 At 48 hours post-enrichment, the Ashdown broth was transferred to a sterile 50ml falcon
197 tube and spun at 3,400 g for 20 seconds. The supernatant was transferred to a clean 50ml
198 falcon tube, spun at 4,300 g for 45 minutes. The supernatant was discarded and the pellet
199 was transferred to a 2ml screwcap tube and stored at -20°C until being shipped to Menzies
200 for direct extraction and PCR confirmation. Direct extraction from pellets was done using the
201 Qiagen DNeasy PowerSoil DNA isolation kit (Qiagen Pty Ltd), and TTS1 qPCR confirmation
202 was performed at Menzies, Darwin.
203
204 **Soil sample processing**
205 Technical issues prohibited adequate on-site parallel processing of soil samples to enable
206 valid comparisons between soil and water at study sites and have consequently been
207 excluded from parallel statistical and geospatial analysis. Briefly, soil samples were stored in
208 the dark at room temperature at LOMWRU until Jan-Feb 2019. Samples were processed
209 using previously established methods and bacterial growth was sub-cultured onto TSB agar
210 slopes in 2ml screwcap tubes and shipped to Menzies, Darwin, Australia for culture
211 detection of *B. pseudomallei* (24-26). DNA was extracted using 10% Chelex-100 resin (22)
212 and *B. pseudomallei* was confirmed using the TTS1 qPCR assay (23).
213
214 **Mapping and statistical analysis**
215 Maps were created with ArcGIS (Version 10.4.1, ESRI Inc) using GPS coordinates
216 recorded at sampling sites. The spatial correlation across *B. pseudomallei* positive sites
217 was examined in ArcGIS implementing the Global indexes of spatial autocorrelation
218 (Moran's I) function with a fixed band distance. A positive Moran's Index and Z-score >
219 1.96 was considered significant at 95% confidence level ($p<0.05$) (27).
220
221 Statistical analyses were computed with Stata 14.0 (www.stata.com). A semiparametric
222 binomial generalized estimating equation (GEE) model with robust standard errors clustered
223 for site (40 sites) was used to analyse associations between the occurrence of *B.*

pseudomallei and different physico-chemical factors by estimating population averaged parameters which are robust to the unknown covariance structure within sites. An exchangeable intra-site correlation structure was estimated (ICC=0.04) and odds ratios (ORs) for *B. pseudomallei* occurrence were calculated. Results were considered significant if *P* values were less than 0.05. Multicollinearity between model predictors was assessed using the variance inflation factor (VIF); all VIF values were less than 1.2. Model residuals were checked and no patterns across predictors were found.

Whole-genome sequencing of *B. pseudomallei* isolates

Twenty-five Vientiane study isolates were included in the comparative genomic analysis. One *B. pseudomallei* water isolate from each culture-positive drain site (n=8) was initially selected for WGS. All isolates were chosen at random. To examine genetic variation between sites, additional isolates cultured from all drain samples were screened by BOX-PCR and visualized via gel electrophoresis using methods previously described by Menzies School of Health Research (28). One isolate was selected from every culture-positive sample within a site (min number of isolates per site, n=1; max number of isolates examined per site, n=4) and screened against the single site isolate already selected for WGS. All sample isolates within a site that had a different banding patterns to the primary WGS isolate were also sent for sequencing (n=13). *B. pseudomallei* isolates recovered from four positive soil samples at two sites were also included in genomic analysis to increase diversity and phylogenetic resolution. Genomic DNA was extracted using the Qiagen DNeasy blood and tissue kit (Qiagen, Chadstone, Victoria, Australia) as previously described (28). Isolates were sequenced at Australian Genome Research Facility Ltd. (Melbourne, Australia) using the Illumina NovaSeq 6000 platform (Illumina, Inc., San Diego, CA). Genomic analysis included an additional 15 publicly available Lao and 159 global *B. pseudomallei* genomes and all genomes are available on the sequence read archive database **(Supplementary Data- Table S1)**. Read quality was conducted using Trimmomatic v0.39 (29) and FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and then assembled using the

MGAP pipeline (https://github.com/dsarov/MGAP-Microbial-Genome-Assembler-Pipeline) generating high quality draft assemblies. Multi-locus sequence typing (MLST) assignment of Lao *B. pseudomallei* environmental soil and water isolates (n=25) was assigned from WGS data *in silico* using the Bacterial Isolates Genome Sequence database (BIGSdb) tool accessible on the *B. pseudomallei* MLST website (http://pubmlst.org/bpseudomallei/) (30).

Orthologous core biallelic single-nucleotide polymorphisms (SNPs) and insertion and deletion events (Indels) were identified from WGS data using Genome Analysis Toolkit (GATK) in SPANDx v3.2 50 and the closed Thai K96243 genome was used as the reference for all phylogenetic analysis (31, 32). Maximum-parsimony (MP) trees were constructed from core orthologous SNPs and indels using PAUP (v4.0a165) (33) with 1000 bootstrap replicates. Trees were visualised in FigTree (v1.4.3) (http://tree.bio.ed.ac.uk/software/figtree/) and manipulated using Interactive Tree of Life (iTOL v4) (https://itol.embl.de) (34).

Results

Detection of *B. pseudomallei* at drain sites

Results of culture and direct detection of *B. pseudomallei* in water samples are shown in **Table 1**. *B. pseudomallei* was detected in water collected at 62.5% (25/40) of sites by either standard culture and/or direct PCR extraction techniques. At only two water-positive drain sites (7.7%, 2/26) did all three water samples test positive for the bacterium. *B. pseudomallei* was detected more frequently using molecular detection techniques than by conventional culture for water samples **(Table 1).**

Positive sites were scattered throughout the city, though some clustering was observed around the That Luang Marsh area in Xaysetha and Sisattanak districts. Global Moran's I and corresponding Z-score also suggested that there was positive spatial autocorrelation

between *B. pseudomallei* positive sites (I=0.31, Z=2.21, p < 0.05). Fewer positive sites were observed in west and northwest areas of the city. **(Figure 2)**.

Of the 1690 patients identified in the prospective Lao melioidosis study between October 1999 and September 2020, ten had incomplete residence data. Of the remaining 1680 patients, 55.5% (933/1680) reported having a residential address in Vientiane Capital, of whom 63.5% (592/933) resided within the five study districts. Cases were most frequently reported from Xaythany (44.8%, 265/592) and Sikhottabong (24.3%, 144/592) districts, with no obvious overrepresentation in the That Luang Marsh area, where spatial clustering of positive environmental samples was found. In view of the inherent biases in the selection of both sampling sites and patient residence, however, formal analysis of associations was not attempted.

Physico-chemical parameters

Characteristics of physico-chemical water parameters from sites (turbidity, temperature, total dissolved solids (TDS), nitrate, acidity (pH), salinity (EC), dissolved oxygen (DO), redox potential (ORP), coliform and *E. coli* counts, drain type, district where located) are shown in the Supplementary Data (**Table S2**). Conductivity differed considerably between samples (49-908 µS/cm) as did turbidity (1.1-851 FNU) and TDS (41-377 ppm). Nitrate content was also variable (8-28 mg/L) as were *E. coli* and coliform counts (both 0->250 CFU/ml) and redox potential (-150.8-192.1 mV). Temperature ranged between 26.3C° and 33.2C°. In contrast, pH only varied by approximately two units (6.36-8.45) and DO fluctuated between 0-4.4mg/L.

Physico-chemical associations with *B. pseudomallei* occurrence in drain water

For the water samples, there was a positive association between the presence of *B. pseudomallei* in water with turbidity, total dissolved solids (TDS), unlined drain sites, as well as slightly cooler temperature (univariable GEE models, p < 0.05 for all) **(Table 2,**

Supplementary Data- Figure S1). *B. pseudomallei* was also less likely to be isolated from Sikhottabong (p=0.049) and Chanthabuly districts (p=0.017) compared to Xanthany district (Table 2). There was no association observed between *B. pseudomallei* and additional variables measured as part of the study **(Supplementary Data- Figure S2)**.

A multivariable GEE model showed that *B. pseudomallei* was negatively associated with drains that were lined with cement rather than those that were sediment-lined and was less likely to be detected in Sikhottabong than in Sisattanak, Chanthabuly, Xaythany, and Xaysetha districts. Water samples higher in turbidity and total dissolved solids were also positively associated with the detection of *B. pseudomallei*. However, after accounting for district, turbidity and TDS, water temperature was no longer significantly associated with the presence of *B. pseudomallei* in the multivariable GEE model **(Table 2, Figure 3)**.

Population structure of *B. pseudomallei* in Laos

From the 25 *B. pseudomallei* Vientiane soil and water isolates selected for WGS we identified ten distinct MLST genotypes **(Table 3)**. ST-507 was the most frequently observed molecular type (n=11), followed by ST-376 (n=3). ST-1792 was the only novel ST type identified. There were two STs (ST-70 and ST-654) that were identified in soil only, while seven STs were isolated in water but not soil. Only one ST, ST-507, was isolated from both sample types.

All nine non-novel STs had been recorded in at least one nearby Asian country. This included Thailand, Cambodia, Vietnam, and China as well as Bangladesh, Malaysia, Singapore, and Indonesia. Three of the nine environmental STs, ST-70, ST-376 and ST-507, had also been identified in Lao melioidosis patients previously (https://pubmlst.org/bpseudomallei/).

A phylogenetic tree was constructed using the 25 study isolates as well as an additional 15 publicly available clinical and environmental Lao genomes, comprising 24 individual STs **(Figure 4)**. Concurrent with the high degree of ST diversity, comparative genomics demonstrated that the Lao isolates were highly genetically diverse, with 56,532 orthologous core genome SNPs and indels detected amongst the 40 isolates. Two distinct clades separated by 4,460 SNPs/indels were also identified, comprising 13 and 11 ST types respectively. Isolates did not group by whether they were clinical or environmental and while most isolates grouped by ST, two ST-507 isolates (MSHR12347 and MSHR12414) did not cluster with the other ST-507 genomes. Both isolates, which were recovered from the same soil sample and differed from one another by only three SNPs/indels, were separated from the other nine ST-507 isolates by more than 8,355 SNPs/indel variants. This distance is consistent with previously reported occurrences of *B. pseudomallei* MLST homoplasy (35, 36) and likely represents homoplasy occurring within Laos. Additionally, phylogenetic analysis demonstrated that one publicly available ST-507 Lao melioidosis patient isolate, MM70, was closely related to a survey water isolates. The water sample isolate (MSHR12012), which was recovered from Saphanthong Tai Village in eastern Vientiane City, differed from the clinical patient genome by only 66 SNP/indel variants **(Figure 4)**. This is despite MM70 having been isolated in 2005 from a Lao melioidosis patient with a residential address in Bolikhamxai Province, approximately 150km northeast of Vientiane City.

Comparative analysis demonstrates *B. pseudomallei* is diverse and well-established in Laos

Whole-genome comparison of the 40 Lao *B. pseudomallei* genomes with an additional 159 global isolates identified 168,934 core SNPs and indel variants. Lao isolates (green branches, **Figure 5**) clustered in multiple distinct groups within the Asian clade, with some appearing to have arisen earlier based on their proximity to the more ancestral Australian strains. These strains (ST-52, ST-491 and ST-535) were more genetically diverse and had

longer branches than strains residing at the end of the global phylogeny. ST-507 isolates appeared to be the least diverse and most recently evolved strain we detected from Laos.

The 40 Lao isolates also shared nodes and clustered with strains from multiple nearby Asian countries, indicating distinct recent common ancestors. These included isolates from Thailand, China, Vietnam, Cambodia, Singapore, Bangladesh, Malaysia, and Indonesia. **(Figure 5)**. ST-507 strains grouped closely and shared nodes with Thai and Chinese isolates, in some instances differing by fewer than 1,100 SNPs and indels. One Lao isolate (MSHR12071; ST- 46) was separated by strains from Bangladesh and Malaysia by 233 and 327 SNPs/indels, respectively.

Discussion

Given the spatial heterogeneity of *B. pseudomallei* distribution in soil, unknown regions endemic for melioidosis may be effectively identified through the analysis of integrated catchment points along a water column (5, 14). We investigated the presence and genetic diversity of *B. pseudomallei* in urban Vientiane, Laos by assessing surface runoff and drainage catchment points throughout the city center. *B. pseudomallei* was detected at the majority of sites surveyed across all districts, indicating that it is well-established there and that surface runoff, particularly during periods of increased rainfall, might be useful for future environmental monitoring of the bacterium. Whole-genome comparison of isolates also demonstrated that Lao *B. pseudomallei* are highly genetically diverse, suggesting that introduction to Laos has not been a recent occurrence and that the bacterium has long been endemic there.

During periods of heavy rainfall and increased surface discharge, *B. pseudomallei* is likely washed out of the soil and channelled into drainage areas along with other eroded particulate matter. Consequently, turbidity and increased suspended solids are thought to be important correlates of the presence of *B. pseudomallei* in water, since bacteria tend to

attach to soil and sediment particles rather than exist in their free-state (17). This association has been observed previously with fecal indicator bacteria after heavy rainfall events (14, 15, 17). Accordingly, we identified a positive association between *B. pseudomallei* and turbid, particle-rich water, as has been observed previously with *B. pseudomallei* isolated from rural domestic water supplies in Northern Australia (38) and in rivers and tributaries throughout southern Laos (14, 15). Additionally, presence of the bacterium was also found to be associated with sediment-laden unlined drains, suggesting sediment might act as an additional reservoir for the pathogen and also supports a link between bank erosion and *B. pseudomallei* particle-bound transport (39).This finding also supports the lack of correlation we detected between *B. pseudomallei*-positive water samples and the presence of fecal coliforms. Abundant enteric microorganisms have been shown to outcompete *B. pseudomallei* in the environment previously and fewer coliforms in a sample may enable the growth of *B. pseudomallei* (17). The lack of correlation identified may reflect variations in the origins of increased turbidity such as soil runoff rather than fecal contamination.

Land use can play an integral role in the transfer of bacteria through soils to downstream aquatic systems and catchment areas (14, 40, 41). As Vientiane Capital continues to develop and expand, changes in land use may ultimately lead to increased soil erosion and runoff. This could potentially affect the distribution and dispersal of the bacterium there, particularly during periods of heavy rainfall (14, 18). Thus, the potential for increased rates of *B. pseudomallei* transmission and its propagation to uncontaminated areas should be considered as the city continues to grow.

Moreover, though we detected *B. pseudomallei* in all districts surveyed as part of the investigation, there was evidence for spatial clustering of the bacterium across the city. Despite the small geographical area surveyed, *B. pseudomallei* was detected at a lower rate in west and northwest areas of Vientiane, again indicative of the heterogeneous

nature of the bacterium in the environment. Additionally, some clustering of positive sites was also observed around That Luang Marsh, located on the eastern edge of the city. The marsh, which is the largest wetland area in Vientiane Municipality, has been designed to collect and treat runoff and drainage water from Vientiane and surrounding areas and also provides local irrigation to farmers (42, 43). The increased degree of runoff could indicate why we observed some clustering of positive sites in this region and why *B. pseudomallei* was detected at a lower rate in the western areas of the city. However, bias caused by non-random sampling due to accessibility and site approval from local authorities should not be discounted as possible study limitations.

For water samples, we applied two separate detection methods including direct TTS1 qPCR on DNA extracted post-enrichment. This has been demonstrated to be a more sensitive technique for the detection of *B*. *pseudomallei* in the environment than standard bacterial culture. Despite this, six of the 120 water samples were detected by culture-only methods and were negative by qPCR post-direct DNA extraction, confirming that no single detection technique is 100% sensitive (44-46). In contrast, soil samples were processed using less-sensitive culture methods than those that are usually recommended and confirmation of *B. pseudomallei* was only performed on small quantities of shipped bacterial cultures due to constraints of project time and budget. While water samples were also filtered and processed promptly after collection, soils were stored for several months before being cultured and shipped back to Darwin, potentially decreasing the viable bacterial count to below the limit of detection. Tropical soils have been demonstrated to be the natural environmental reservoir for *B. pseudomallei* detected in rivers and groundwater, with the bacterium leached out of the soil along with eroded particulate matter during periods of heavy rainfall (14, 41). Consequently, it is likely that the pathogen was present in many of the soils collected in water-positive survey sites but was not detected by our collection and processing methods. Future comparisons between the roles

and links of water and soil will require more intensive soil sampling and on-site processing and analysis.

Although *B. pseudomallei* has one of the most highly recombinogenic genomes of any bacteria, certain features of its biology mean reliable inferences about geographic origin and population structure can still be made, particularly when high-resolution WGS data is used (37). In 2009 Pearson and colleagues were the first to hypothesize an Australian origin for *B. pseudomallei*. Combining WGS, Bayesian inference and molecular clock estimates, they predicted that *B. pseudomallei* moved into Southeast Asia during the last glacial period (16-225 thousand years ago), when the Sahul and Sunda land masses were in close proximity due to low sea levels (6). Studies across larger more diverse sets of data have supported this hypothesis and it has recently been shown that there have been several successive *B. pseudomallei* re-introductory events within Southeast Asia. This was particularly evident amongst countries bordering the Mekong River and Malay Peninsula, where there was a high degree of genetic relatedness and shared ancestry amongst isolates (8). Given Laos has geographical borders with five Southeast Asian countries and the Mekong River runs along its western boundary, the extent of ST diversity and genetic relatedness we observed amongst *B. pseudomallei* isolates from Laos and those from neighboring countries was unsurprising. Collectively, the diversity and divergence of isolates within Laos suggests that the original introduction of *B. pseudomallei* did not happen recently and the disease has long been endemic there.

Moreover, Lao genomes clustered in different clades within our global phylogeny. This could indicate that *B. pseudomallei* was introduced to Laos on multiple separate occasions, with isolates having distinct recent common ancestors. These repeated introductory events and the subsequent dispersal of *B. pseudomallei* within Laos are likely multifaceted. Severe weather and flooding during the monsoonal season have probably played an important role, as has its close proximity to neighboring countries, which would enable transmission by both

humans and animals. More environmental sampling and sequencing of isolates throughout Laos will be necessary to elucidate the timeframe in which these introductory events may have occurred and further explore the phylogeographic relatedness with other Southeast Asian isolates.

Our results also revealed that *B. pseudomallei* isolates from Laos are highly genetically diverse, with 24 STs and 56,532 orthologous core SNPs and indels identified amongst the 40 sequenced isolates. Likewise, we identified ten individual STs amongst the 25 environmental survey isolates sent for WGS within the small 100km^2 study radius. Despite the overall degree of genetic diversity, whole-genome comparison identified two highly related ST-507 isolates: a Lao clinical isolate from 2005 and a Vientiane survey water isolate from 2018 collected from Saphanthong Tai Village. Despite being isolated 13 years apart, the genomes were separated by only 66 core SNPs and indels. Additionally, the patient's residential address was approximately 150km northeast of the source of the environmental isolate. Genetic populations of *B. pseudomallei* have been demonstrated to spatially cluster in the environment on a highly localized scale despite frequent opportunities to spread within the water table, via agricultural and migratory animals, or in transported soil (7, 47-49). Genetic clustering has also been shown to match the spatial distribution of clinical cases previously (13). This might indicate the patient did not acquire their infection at their residential address but closer to the location where the survey isolate was collected in Vientiane, although more detailed clinical epidemiological data would be required to determine this. Alternatively, this finding could suggest that ST-507 is comparatively widespread throughout central Laos and that there is limited intra-ST-507 diversity. This is supported by results from our global phylogeny, which demonstrated that ST-507 is a more recently evolved and less genetically diverse *B. pseudomallei* sequence type than many others. Additional sampling and WGS of clinical and environmental isolates from Laos are needed to further examine this, since relatively few Lao isolates have had MLST or WGS completed.

Previous environmental surveys undertaken in Laos have shown that *B. pseudomallei* is widespread throughout central and southern regions of the country, with high concentrations of the bacterium identified in rice paddies in rural Vientiane Province (4, 50-52). Results from this study indicate that *B. pseudomallei* is also widespread in the environment throughout urban Vientiane, where more than 10% of the Lao population currently resides and where over half of the individuals diagnosed with melioidosis at Mahosot Hospital lived (4). While the true distribution and epidemiology of melioidosis is still not well characterized in Laos, our results indicate that infection from contact with the environment is a significant risk in urban Vientiane, with drains and surface runoff being potential sources in addition to more conventional sources such as agricultural land. The rate of *B. pseudomallei* detection in drain water across the study area also corresponded with the high proportion of Lao melioidosis patients reporting residential addresses within the five urban districts, comprising a third (35.2%, 592/1680) of all cases confirmed at Mahosot Hospital. Despite this, patient addresses did not appear to cluster around That Luang Marsh like the positive environmental sites. This finding may reflect the non-random aspects of environmental sampling in addition to population density, underlying risk factors, access to healthcare in these areas, and the fact that not all patients will have been infected at their residential address.

Author statements

Acknowledgements

We thank the Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU) and Mahosot Hospital Microbiology staff for their advice and support in undertaking this investigation, particularly Anisone Chanthongthip and Sengkham Simanivong. Additionally, we thank the microbiology staff at Menzies School of Health and Research for their assistance with culture and processing of samples, especially Vanessa Rigas. We are grateful to Professor Paul Newton, Director of LOMWRU, and Dr. Manivanh Vongsouvath,

529 Head of the Mahosot Hospital Microbiology Laboratory, and the staff at the Ministry of
530 Health, the University of Health Sciences, and the Directors of Mahosot Hospital, for their
531 support of the this study and the work of LOMWRU.

532

533 **Authors and contributors**

534 The study was conceptualized by AR, MM, MK, DABD, and BJC. Funding was acquired by
535 BJC and DABD. Data collection was undertaken by AR, ML, and SR. Formal analysis was
536 done by AR, MK, and JRW. AR wrote the original draft, and MK, BJC, MM, JRW and DABD
537 revised and edited the manuscript. All authors saw and approved the final version.

538

539 **Conflicts of interest**

540 None of the authors have any conflicts of interest to declare.

541

542 **Funding information**

543 This study was supported by grants from the Australian National Health and Medical
544 Research Council: grant numbers 1046812, 1098337 and 1131932 (The HOT NORTH
545 initiative) and AR is supported by a Charles Darwin University International PhD scholarship.
546 DABD was supported by the Wellcome Trust (grant number 106698/Z/14/Z).

547

548 **Ethical approval**

549 This study was approved by the Lao National Ethics Committee for Health Research
550 (2018.32.MC) and the Menzies School of Health Research (HREC 02/38). Sample site
551 approval was obtained by local Lao landowners before survey commencement.

552

553 **Data availability:**

554 Raw sequence data from this study are available in the Short Read Archive in Bioproject
555 PRJNA659606 (http://www.ncbi.nlm.nih.gov/bioproject/ PRJNA659606), biosample
556 accessions SAMN15327852-15327876.

557

558 **References**

559 1. Cheng AC, Currie BJ. 2005. Melioidosis: epidemiology, pathophysiology, and
560 management. Clin Microbiol Rev 18:383-416.

561 2. Limmathurotsakul D, Wongratanacheewin S, Teerawattanasook N, Wongsuvan G,
562 Chaisuksant S, Chetchotisakd P, Chaowagul W, Day NP, Peacock SJ. 2010.
563 Increasing incidence of human melioidosis in Northeast Thailand. Am J Trop Med
564 Hyg 82:1113-7.

565 3. Phetsouvanh R, Phongmany S, Newton P, Mayxay M, Ramsay A, Wuthiekanun V,
566 White NJ. 2001. Melioidosis and Pandora's box in the Lao People's Democratic
567 Republic. Clin Infect Dis 32:653-654.

568 4. Dance DAB, Luangraj M, Rattanavong S, Sithivong N, Vongnalaysane O,
569 Vongsouvath M, Newton PN. 2018. Melioidosis in the Lao People's Democratic
570 Republic. Trop Med Infect Dis 3.

571 5. Limmathurotsakul D, Wuthiekanun V, Chantratita N, Wongsuvan G, Amornchai P,
572 Day NPJ, Peacock SJ. 2010. *Burkholderia pseudomallei* Is Spatially Distributed in
573 Soil in Northeast Thailand. PLOS Negl Trop Dis 4:e694.

574 6. Pearson T, Giffard P, Beckstrom-Sternberg S, Auerbach R, Hornstra H, Tuanyok A,
575 Price EP, Glass MB, Leadem B, Beckstrom-Sternberg JS, Allan GJ, Foster JT,
576 Wagner DM, Okinaka RT, Sim SH, Pearson O, Wu Z, Chang J, Kaul R, Hoffmaster
577 AR, Brettin TS, Robison RA, Mayo M, Gee JE, Tan P, Currie BJ, Keim P. 2009.
578 Phylogeographic reconstruction of a bacterial species with high levels of lateral gene
579 transfer. BMC Biol 7:78.

580 7. McRobb E, Kaestli M, Price EP, Sarovich DS, Mayo M, Warner J, Spratt BG, Currie
581 BJ. 2014. Distribution of *Burkholderia pseudomallei* in northern Australia, a land of
582 diversity. Appl Environ Microbiol 80:3463-8.

583 8. Chewapreecha C, Holden MT, Vehkala M, Valimaki N, Yang Z, Harris SR, Mather
584 AE, Tuanyok A, De Smet B, Le Hello S, Bizet C, Mayo M, Wuthiekanun V,

Limmathurotsakul D, Phetsouvanh R, Spratt BG, Corander J, Keim P, Dougan G, Dance DA, Currie BJ, Parkhill J, Peacock SJ. 2017. Global and regional dissemination and evolution of *Burkholderia pseudomallei*. Nat Microbiol 2:16263.

9. 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 53:282-6.

10. Sarovich DS, Garin B, De Smet B, Kaestli M, Mayo M, Vandamme P, Jacobs J, Lompo P, Tahita MC, Tinto H, Djaomalaza I, Currie BJ, Price EP. 2016. Phylogenomic Analysis Reveals an Asian Origin for African *Burkholderia pseudomallei* and Further Supports Melioidosis Endemicity in Africa. mSphere 1.

11. De Smet B, Sarovich DS, Price EP, Mayo M, Theobald V, Kham C, Heng S, Thong P, Holden MTG, Parkhill J, Peacock SJ, Spratt BG, Jacobs JA, Vandamme P, Currie BJ. 2015. Whole-genome sequencing confirms that *Burkholderia pseudomallei* multilocus sequence types common to both Cambodia and Australia are due to homoplasy. J Clin Microbiol 53:323-326.

12. Dale J, Price EP, Hornstra H, Busch JD, Mayo M, Godoy D, Wuthiekanun V, Baker A, Foster JT, Wagner DM, Tuanyok A, Warner J, Spratt BG, Peacock SJ, Currie BJ, Keim P, Pearson T. 2011. Epidemiological tracking and population assignment of the non-clonal bacterium, *Burkholderia pseudomallei*. PLoS Negl Trop Dis 5:e1381.

13. 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. Nat Sci Rep 10:5443.

14. Ribolzi O, Rochelle-Newall E, Dittrich S, Auda Y, Newton PN, Rattanavong S, Knappik M, Soulileuth B, Sengtaheuanghoung O, Dance DAB, Pierret A. 2016. Land use and soil type determine the presence of the pathogen *Burkholderia pseudomallei* in tropical rivers. Environ Sci Poll Res 23:7828-7839.

15. 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. Nat Sci Rep 8:8674.

16. Baker A, Tahani D, Gardiner C, Bristow KL, Greenhill AR, Warner J. 2011. Groundwater Seeps Facilitate Exposure to *Burkholderia pseudomallei*. Appl Environ Microbiol 77:7243-7246.

17. Rochelle-Newall E, Nguyen TMH, Le TPQ, Sengtaheuanghoung O, Ribolzi O. 2015. A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions. Front Microbiol 6:308-308.

18. Chuah CJ, Tan EKH, Sermswan RW, Ziegler AD. 2017. Hydrological connectivity and *Burkholderia pseudomallei* prevalence in wetland environments: investigating rice-farming community's risk of exposure to melioidosis in North-East Thailand. Environ Monit Assess 189:287.

19. Vongphayloth K, Rattanavong S, Moore CE, Phetsouvanh R, Wuthiekanun V, Sengdouangphachanh A, Phouminh P, Newton PN, Buisson Y. 2012. *Burkholderia pseudomallei* detection in surface water in southern Laos using Moore's swabs. Am J Trop Med Hyg 86:872-877.

20. Limmathurotsakul D, Golding N, Dance DA, Messina JP, Pigott DM, Moyes CL, Rolim DB, Bertherat E, Day NP, Peacock SJ, Hay SI. 2016. Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. Nat Microbiol 1:15008.

21. Rafiqui PS, Gentile M. 2009. Vientiane. Cities 26:38-48.

22. de Lamballerie X, Zandotti C, Vignoli C, Bollet C, de Micco P. 1992. A one-step microbial DNA extraction method using “Chelex 100” suitable for gene amplification. Res Microbiol 143:785-790.

23. Novak RT, Glass MB, Gee JE, Gal D, Mayo MJ, Currie BJ, Wilkins PP. 2006. Development and evaluation of a real-time PCR assay targeting the type III secretion system of *Burkholderia pseudomallei*. J Clin Microbiol 44:85-90.

641 24. Mayo M, Kaesti M, Harrington G, Cheng AC, Ward L, Karp D, Jolly P, Godoy D,
642 Spratt BG, Currie BJ. 2011. *Burkholderia pseudomallei* in unchlorinated domestic
643 bore water, Tropical Northern Australia. Emerg Infect Dis 17:1283-1285.
644 25. Currie BJ, Price EP, Mayo M, Kaestli M, Theobald V, Harrington I, Harrington G,
645 Sarovich DS. 2015. Use of Whole-Genome Sequencing to Link *Burkholderia*
646 *pseudomallei* from Air Sampling to Mediastinal Melioidosis, Australia. Emerg Infect
647 Dis 21:2052-4.
648 26. Limmathurotsakul D, Dance DAB, Wuthiekanun V, Kaestli M, Mayo M, Warner J,
649 Wagner DM, Tuanyok A, Wertheim H, Yoke Cheng T, Mukhopadhyay C,
650 Puthucheary S, Day NPJ, Steinmetz I, Currie BJ, Peacock SJ. 2013. Systematic
651 review and consensus guidelines for environmental sampling of *Burkholderia*
652 *pseudomallei*. PLoS Negl Trop Dis 7:e2105-e2105.
653 27. Waller LA, Gotway CA. 2004. Applied spatial statistics for public health data, vol 368.
654 John Wiley & Sons.
655 28. Currie BJ, Gal D, Mayo M, Ward L, Godoy D, Spratt BG, LiPuma JJ. 2007. Using
656 BOX-PCR to exclude a clonal outbreak of melioidosis. BMC Infect Dis 7:68.
657 29. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina
658 sequence data. Bioinformatics 30:2114-20.
659 30. Jolley KA, Maiden MCJ. 2010. BIGSdb: Scalable analysis of bacterial genome
660 variation at the population level. BMC Bioinformatics 11:595.
661 31. Sarovich DS, Price EP. 2014. SPANDx: a genomics pipeline for comparative
662 analysis of large haploid whole genome re-sequencing datasets. BMC Res Notes
663 7:618.
664 32. Holden MT, Titball RW, Peacock SJ, Cerdeno-Tarraga AM, Atkins T, Crossman LC,
665 Pitt T, Churcher C, Mungall K, Bentley SD, Sebaihia M, Thomson NR, Bason N,
666 Beacham IR, Brooks K, Brown KA, Brown NF, Challis GL, Cherevach I, Chillingworth
667 T, Cronin A, Crossett B, Davis P, DeShazer D, Feltwell T, Fraser A, Hance Z, Hauser
668 H, Holroyd S, Jagels K, Keith KE, Maddison M, Moule S, Price C, Quail MA,

669 Rabbinowitsch E, Rutherford K, Sanders M, Simmonds M, Songsivilai S, Stevens K,
670 Tumapa S, Vesaratchavest M, Whitehead S, Yeats C, Barrell BG, Oyston PC,
671 Parkhill J. 2004. Genomic plasticity of the causative agent of melioidosis,
672 *Burkholderia pseudomallei*. Proc Natl Acad Sci U S A 101:14240-5.
673 33. Wilgenbusch JC, Swofford D. 2003. Inferring Evolutionary Trees with PAUP*. Current
674 Protocols in Bioinformatics 00:6.4.1-6.4.28.
675 34. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display
676 and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-W245.
677 35. Aziz A, Sarovich DS, Harris TM, Kaestli M, McRobb E, Mayo M, Currie BJ, Price EP.
678 2017. Suspected cases of intracontinental *Burkholderia pseudomallei* sequence type
679 homoplasy resolved using whole-genome sequencing. Microb Genom 3.
680 36. Rachlin A, Kleinecke M, Kaestli M, Mayo M, Webb JR, Rigas V, Shilton C, Benedict
681 S, Dyrting K, Currie BJ. 2019. A cluster of melioidosis infections in hatchling
682 saltwater crocodiles (*Crocodylus porosus*) resolved using genome-wide comparison
683 of a common north Australian strain of *Burkholderia pseudomallei*. Microb Genom 5.
684 37. Price EP, Currie BJ, Sarovich DS. 2017. Genomic Insights Into the Melioidosis
685 Pathogen, *Burkholderia pseudomallei*. Current Trop Med Rep 4:95-102.
686 38. Draper AD, Mayo M, Harrington G, Karp D, Yinfoo D, Ward L, Haslem A, Currie BJ,
687 Kaestli M. 2010. Association of the melioidosis agent *Burkholderia pseudomallei* with
688 water parameters in rural water supplies in Northern Australia. Appl Environ Microbiol
689 76:5305-7.
690 39. Rochelle-Newall E, Nguyen TMH, Le TPQ, Sengtaheuanghoung O, Ribolzi O. 2015.
691 A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge
692 gaps and future directions. Front Microbiol 6.
693 40. Causse J, Billen G, Garnier J, Henri-des-Tureaux T, Olasa X, Thammahacksa C,
694 Latsachak KO, Soulileuth B, Sengtaheuanghoung O, Rochelle-Newall E. 2015. Field
695 and modelling studies of *Escherichia coli* loads in tropical streams of montane agro-
696 ecosystems. Journal of Hydro-Environment Research 9:496-507.

41. Deiner K, Fronhofer EA, Mächler E, Walser J-C, Altermatt F. 2016. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nat Comm 7:12544.

42. Gerrard P. 2004. Integrating wetland ecosystem values into urban planning: the case of that Luang Marsh. Vientiane, Lao PDR, WWF Lao PDR and IUCN–The World Conservation Regional Environmental Economics Programme Asia, Colombo.

43. Kyophilavong P. 2008. Impact of irrigation on aquatic wetland resources: a case study of That Luang Marsh, Lao PDR. EEPSEA, IDRC Regional Office for Southeast and East Asia, Singapore, SG.

44. Knappik M, Dance DAB, Rattanavong S, Pierret A, Ribolzi O, Davong V, Silisouk J, Vongsouvath M, Newton PN, Dittrich S. 2015. Evaluation of Molecular Methods To Improve the Detection *Burkholderia pseudomallei* Soil and Water Samples from Laos. Appl Environl Microbiol 81:3722-3727.

45. Kaestli M, Mayo M, Harrington G, Watt F, Hill J, Gal D, Currie BJ. 2007. Sensitive and specific molecular detection of *Burkholderia pseudomallei*, the causative agent of melioidosis, in the soil of tropical northern Australia. Appl Environ Microbiol 73:6891-7.

46. Dance D, Knappik M, Dittrich S, Davong V, Silisouk J, Vongsouvath M, Rattanavong S, Pierret A, Newton P, Amornchai P, Wuthiekanun V, Langla S, Limmathurotsakul D. 2018. Evaluation of consensus method for the culture of *Burkholderia pseudomallei* in soil samples from Laos. Wellcome Open Res 3:132.

47. Chapple SNJ, Price EP, Sarovich DS, McRobb E, Mayo M, Kaestli M, Spratt BG, Currie BJ. 2016. *Burkholderia pseudomallei* genotype distribution in the Northern Territory, Australia. Am J Trop Med Hyg 94:68-72.

48. Hampton V, Kaestli M, Mayo M, Choy JL, Harrington G, Richardson L, Benedict S, Noske R, Garnett ST, Godoy D. 2011. Melioidosis in birds and *Burkholderia pseudomallei* dispersal, Australia. Emerg Infect Dis 17:1310-1311.

724 49. Baker A, Tahani D, Gardiner C, Bristow KL, Greenhill AR, Warner J. 2011.
725 Groundwater seeps facilitate exposure to *Burkholderia pseudomallei*. Appl Environ
726 Microbiol 77:7243-7246.
727 50. Rattanavong S, Wuthiekanun V, Langla S, Amornchai P, Sirisouk J, Phetsouvanh R,
728 Moore CE, Peacock SJ, Buisson Y, Newton PN. 2011. Randomized Soil Survey of
729 the Distribution of *Burkholderia pseudomallei* in Rice Fields in Laos. Appl Environ
730 Microbiol 77:532-536.
731 51. Wuthiekanun V, Mayxay M, Chierakul W, Phetsouvanh R, Cheng AC, White NJ, Day
732 NPJ, Peacock SJ. 2005. Detection of *Burkholderia pseudomallei* in Soil within the
733 Lao People's Democratic Republic. J Clin Microbiol 43:923-924.
734 52. Manivanh L, Pierret A, Rattanavong S, Kounnavongsa O, Buisson Y, Elliott I, Maeght
735 JL, Xayyathip K, Silisouk J, Vongsouvath M, Phetsouvanh R, Newton PN, Lacombe
736 G, Ribolzi O, Rochelle-Newall E, Dance DAB. 2017. *Burkholderia pseudomallei* in a
737 lowland rice paddy: seasonal changes and influence of soil depth and physico-
738 chemical properties. Nat Sci Rep 7:3031.
739
740
741
742
743
744
745
746
747
748
749
750
751

Tables

Table 1- *B. pseudomallei* positive sites and samples based on different detection techniques.

Method of detection	Sites	Water samples
Culture positive	22.5% (7/40)	11.7% (14/120)
Direct DNA extraction PCR positive	45.0% (18/40)	21.7% (26/120)
Direct DNA extraction PCR negative, culture positive	10% (4/40)	5% (6/120)
Both methods positive	12.5% (5/40)	6.7% (8/120)
Total	**62.5% (25/40)**	**33.3% (40/120)**

Table 2- Multivariable GEE analysis of water parameters associated with the presence of *B. pseudomallei* in Vientiane drains. Asterisks denote level of significance: * P<0.05 and ** P<0.01.

Variable	Median value		Multivariable GEE Model OR (95% CI) P value
	Positive	Negative	
Turbidity (FNU) (log transformed in GEE)	79.2 (13.1-851)	32.7 (1.1-461)	2.42 (1.31-4.5) 0.005**
TDS (ppm)	192 (103-377)	152 (41-304)	1.01 (1.01-1.02) 0.006**
	Number of samples		
Drain lining	**Positive**	**Negative**	
Cement-lined	2.5% (3/120)	20% (24/120)	0.14 (0.02-0.88) 0.036*
Unlined	30.8% (37/120)	46.7% (56/120)	Reference level
District			
Sisattanak District	15% (18/120)	15% (18/120)	0.9 (0.32-2.61) 0.87
Xaysetha District	7.5% (9/120)	12.5% (15/120)	0.43 (0.15-1.26) 0.13
Xaythany District	5.8% (7/120)	6.7% (8/120)	Reference level
Sikhottabong District	1.7% (2/120)	15.8% (19/120)	0.09 (0.009-0.72) 0.02*
Chanthabuly District	3.4% (4/120)	16.7% (20/120)	0.01 (0.08-1.57) 0.17

763 **Table 3-** Vientiane *B. pseudomallei* study isolates included in whole-genome

764 comparative analysis. *Asterisk denotes novel ST.

Isolate ID	Sample Type	Vientiane Village/District	ST Type
MSHR12071	Water	Phonphapao, Sisattanak	46
MSHR11998	Water	Phonphapao, Sisattanak	52
MSHR12336	Soil	Phonthan Neua, Xaysetha	70
MSHR12338	Soil	Phonthan Neua, Xaysetha	70
MSHR12048	Water	Phonphapao, Sisattanak	203
MSHR12122	Water	Tonphanthong, Sisattanak	203
MSHR11848	Water	Saphanthong, Sisattanak	368
MSHR12054	Water	Saphanthong, Sisattanak	368
MSHR11846	Water	Tonphanthong, Sisattanak	376
MSHR12046	Water	Phonphapao, Sisattanak	376
MSHR12097	Water	Tonphanthong, Sisattanak	376
MSHR11836	Water	Phonphapao, Sisattanak	507
MSHR11855	Water	Saphanthong Tai, Sisattanak	507
MSHR11859	Water	Saphanthong Tai, Sisattanak	507
MSHR11966	Water	Nongchan, Chanthabuly	507
MSHR12012	Water	Saphanthong Tai, Sisattanak	507
MSHR12020	Water	Sengsavanh, Sikhottabong	507
MSHR12061	Water	Saphanthong Tai, Sisattanak	507
MSHR12077	Water	Sengsavanh, Sikhottabong	507
MSHR12103	Water	Nongchan, Chanthabuly	507
MSHR12347	Soil	Sybounheung, Xathany	507
MSHR12414	Soil	Sybounheung, Xathany	507
MSHR12000	Water	Phonphapao, Sisattanak	535
MSHR12369	Soil	Phonthan Neua, Xaysetha	654
MSHR12059	Water	Sapangmor, Sisattanak	*1792

765

766

767

768

769

770

771

772

773

775 **Figures**

776

777 **Figure 1-** Map of the 18 provinces making up Lao PDR with Vientiane Capital
778 highlighted in red. Smaller insert shows Vientiane Capital comprised of its nine
779 districts. Numbers indicate the four urban and one semi urban districts where survey
780 sites were located (Chanthabuly:1, Sikhottabong:2, Xaysetha:3, Sisattanak:4,
781 Xathany:5).

782

783 **Figure 2**- Sampling site locations across urban Vientiane Capital. Sites where *B.*
784 *pseudomallei* was detected by culture and/or direct detection in water are denoted
785 by red triangles. Negative sites (green triangles) are those where *B. pseudomallei*
786 was not identified in water by either detection method. The base map is from
787 ArcGIS/Esri (sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO,
788 USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri
789 Japan, METI, Esri China [Hong Kong], OpenStreetMap contributors, and the GIS
790 User Community).

791

792 **Figure 3-** Generalized estimated equation (GEE) margins plots of adjusted
793 predicted probabilities of *B. pseudomallei* occurrence with all other variables in the
794 model held constant. Predicted probabilities as a function of increased water TDS
795 (A.) in each of the districts surveyed and (B.) turbidity, or probabilities in unlined or
796 cement-lined drains based on (C.) TDS or (D.) turbidity of water sample. Bars
797 denote 95% confidence intervals (CIs).

798

799 **Figure 4**- Midpoint rooted maximum parsimony phylogeny of 40 Lao *B. pseudomallei*
800 isolates based on 56,532 core genome SNPs and indels. Black circles denote
801 bootstrap values <80.
802
803 **Figure 5-** Maximum parsimony phylogeny of *B. pseudomallei* from Laos (n=40, green
804 branches) with a global set of genomes (n=159) based on 168,934 core genome SNPs
805 and indels. The closed Thai isolate K96243 was used as the reference strain and the
806 tree was rooted at MSHR0668, the most ancestral *B. pseudomallei* strain identified in
807 a previous study (37). Black circles denote bootstrap values <80.