



OPEN

Myanmar *Burkholderia pseudomallei* strains are genetically diverse and originate from Asia with phylogenetic evidence of reintroductions from neighbouring countries

Jessica R. Webb¹✉, Mo Mo Win³, Khwar Nyo Zin⁸, Kyi Kyi Nyein Win³, Thin Thin Wah³, Elizabeth A. Ashley^{4,6}, Frank Smithuis^{4,6}, Myo Maung Maung Swe⁴, Mark Mayo¹, Bart J. Currie^{1,2} & David A. B. Dance^{4,5,7}

Melioidosis was first identified in Myanmar in 1911 but for the last century it has remained largely unreported there. *Burkholderia pseudomallei* was first isolated from the environment of Myanmar in 2016, confirming continuing endemicity. Recent genomic studies showed that *B. pseudomallei* originated in Australia and spread to Asia, with phylogenetic evidence of repeated reintroduction of *B. pseudomallei* across countries bordered by the Mekong River and the Malay Peninsula. We present the first whole-genome sequences of *B. pseudomallei* isolates from Myanmar: nine clinical and seven environmental isolates. We used large-scale comparative genomics to assess the genetic diversity, phylogeography and potential origins of *B. pseudomallei* in Myanmar. Global phylogenetics demonstrated that Myanmar isolates group in two distantly related clades that reside in a more ancestral Asian clade with high amounts of genetic diversity. The diversity of *B. pseudomallei* from Myanmar and divergence within our global phylogeny suggest that the original introduction of *B. pseudomallei* to Myanmar was not a recent event. Our study provides new insights into global patterns of *B. pseudomallei* dissemination, most notably the dynamic nature of movement of *B. pseudomallei* within densely populated Southeast Asia. The role of anthropogenic influences in both ancient and more recent dissemination of *B. pseudomallei* to Myanmar and elsewhere in Southeast Asia and globally requires further study.

The opportunistic environmental biothreat agent *Burkholderia pseudomallei* causes melioidosis, a disease of diverse manifestations with mortality rates ranging from 9 to 40%^{1,2}. Melioidosis was first described in Rangoon (now Yangon: 21.9° N, 95.9° E), Myanmar (formerly Burma) in 1911 by British pathologist, Alfred Whitmore, in a 40-year-old morphine addict who died from a fever associated with multiple subcutaneous abscesses and pneumonia^{3,4}. Over the next 5 years, Whitmore's assistant Krishnaswamy saw over 200 cases of what was then known as 'morphia injectors' septicæmia' in Rangoon⁵. Since then, however, reports of melioidosis in Myanmar

¹Global and Tropical Health Division, Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia. ²Department of Infectious Diseases and Northern Territory Medical Program, Royal Darwin Hospital, Darwin, NT, Australia. ³Department of Medical Research, Yangon, Myanmar. ⁴Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Old Road Campus, Oxford, UK. ⁵Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Microbiology Laboratory, Mahosot Hospital, Vientiane, Laos. ⁶Myanmar-Oxford Clinical Research Unit, Yangon, Myanmar. ⁷Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK. ⁸Microbiology Laboratory, Yangon General Hospital, Yangon, Myanmar. ✉email: Jessica.webb@menzies.edu.au

have dwindled and become sporadic, with only a handful of cases reported in the 1940s^{6,7} before the disease was rediscovered more than half a century later, in 2000 in Mandalay⁸. The gradual re-emergence of melioidosis in Myanmar has recently been reviewed⁶.

The climate in Myanmar varies from north to south but is generally tropical, with rainfall varying from 5500 mm in the coastal regions to less than 1000 mm in the central ‘Dry Zone’⁹. Recent modelling predicted that the environment of much of Myanmar would be suitable for the persistence of *B. pseudomallei*¹⁰, although the organism was only first detected in environmental samples from Myanmar in 2016, more than a century after its initial description by Whitmore. The diagnosis of culture-confirmed human melioidosis cases at Yangon General Hospital in 2009 prompted soil sampling on two farms in nearby townships, Hmawbi and Thanlyin. *B. pseudomallei* was isolated from seven soil samples: five from Hmawbi and two from Thanlyin¹¹. Subsequently a systematic survey of the environment in Myanmar was initiated in 2018 in order to identify areas of high melioidosis risk (Smithuis, manuscript in preparation).

Although melioidosis was initially discovered in Myanmar, at the time of initiating this study there were no isolates of *B. pseudomallei* included in the international Multilocus Sequence Typing (MLST) database (<https://pubmlst.org/bpseudomallei/>) to inform understanding of the place of the country in the global phylogeography of the species. Prior studies have shown that *B. pseudomallei* originally evolved in the environment of Australia and subsequently spread to Southeast Asia and East Asia during the last glacial period (between 16 and 225 thousand years ago)¹². From Asia, *B. pseudomallei* subsequently spread to Madagascar and Africa and more recently from West Africa to the Americas (between 1650 and 1850), likely linked to the slave trade^{12–15}. This strong geographic signal of the *B. pseudomallei* genome is notably due to infection in humans and animals being acquired directly from an environmental source, with human–human and zoonotic transmission being exceedingly rare. Such epidemiology and large continental geographic barriers restrict gene flow and have resulted in distinct geographic populations of *B. pseudomallei*.

The isolation of *B. pseudomallei* from patients and the environment in Myanmar not only confirmed the continuing endemicity of melioidosis in the country but enabled phylogenetic comparisons of *B. pseudomallei* from Myanmar with those from neighbouring countries and the global *B. pseudomallei* dataset in order to infer the origins and investigate whether the recent re-emergence of melioidosis in Myanmar reflected persistence or recent re-introductions. This study included sixteen *B. pseudomallei* isolates from Myanmar (nine human and seven environmental). We present the first Myanmar *B. pseudomallei* whole-genome sequences (WGS) and, using large-scale comparative-genomics, we attempt to determine the origin of *B. pseudomallei* in Myanmar.

Methods

Study setting and Myanmar *B. pseudomallei* isolates. Myanmar has a total land area of 676,577 km², with a length of 2,090 km from north to south, has four important river systems, with the Ayeyarwaddy River the main waterway, and comprises 14 states and regions. It extends south towards the Malay Peninsula, is bordered by China to the north and northeast, Laos to the east, Thailand to the south and east, Bangladesh to the west and India to the northwest and is situated within the biogeographical region of Sundaland (Fig. 1A).

Included in this study were sixteen *B. pseudomallei* isolates from Myanmar collected from 2017 to 2018 during ongoing enhanced surveillance for clinical cases of melioidosis and the national environmental survey. The collection comprised seven soil isolates from Kyaikmaraw township, Mon state (one isolate), Paukkhaung township, Bago region (two isolates), Shwe Taung township, Bago region (two isolates) and Magway township, Magway region (two isolates) (Fig. 1B), and nine isolates from melioidosis patients in Ayeyarwady region (two patients), Kayin state (one patient), Yangon region (four patients) and Rakhine state (two patients) (Fig. 1). All patients were believed to have acquired their infections in Myanmar with no history of travel. None of the soil isolates were epidemiologically linked to the melioidosis patients.

Sequencing, quality control and assembly. Culture of *B. pseudomallei* was performed using methods previously established^{17–19} and their identity was confirmed using a real-time PCR assay targeting a *B. pseudomallei*-specific 115-bp segment within the type three secretion system 1 (TTS1) gene²⁰. Genomic DNA was extracted from purified *B. pseudomallei* colonies as previously described²¹. WGS were performed on the Illumina NovaSeq platform, with the sequencing platform generating 150 bp paired-end reads for each genome. Genomic analysis included an additional 173 publicly available *B. pseudomallei* genomes and all genomes are available on the sequence read archive database (Table S1). Read quality was conducted using Trimmomatic v0.39²² 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 (contigs ≤ 90 and N50 $> 180,000$ bp).

MLST assignment. MLST was performed to assign STs to isolates by determining allele sequence in silico from bacterial WGS read data using the BIGSdb tool, which is accessible on the *B. pseudomallei* MLST website (<https://pubmlst.org/bpseudomallei/>).

Variant calling, phylogenetic analysis, recombination identification and pan-genome analysis. Single-nucleotide polymorphisms (SNP) and small insertions/deletion (Indels) variants were identified from WGS data using Genome Analysis Toolkit (GATK v4.1.0.0) (<https://github.com/broadinstitute/gatk>) wrapped in SPANdx v3.2²³ and the closed K96243 (accessions BX571965 and BX571966) genome was used as the reference for all phylogenetic analysis. The SNP and Indel variants identified by GATK were used for phylogenetic reconstruction using maximum parsimony (MP) in PAUP* 4.0.b5²⁴ or SNP variants for maximum likelihood (ML) in RAXML²⁵. Bootstrapping using 1000 replicates was carried out for MP and phylogenetic trees

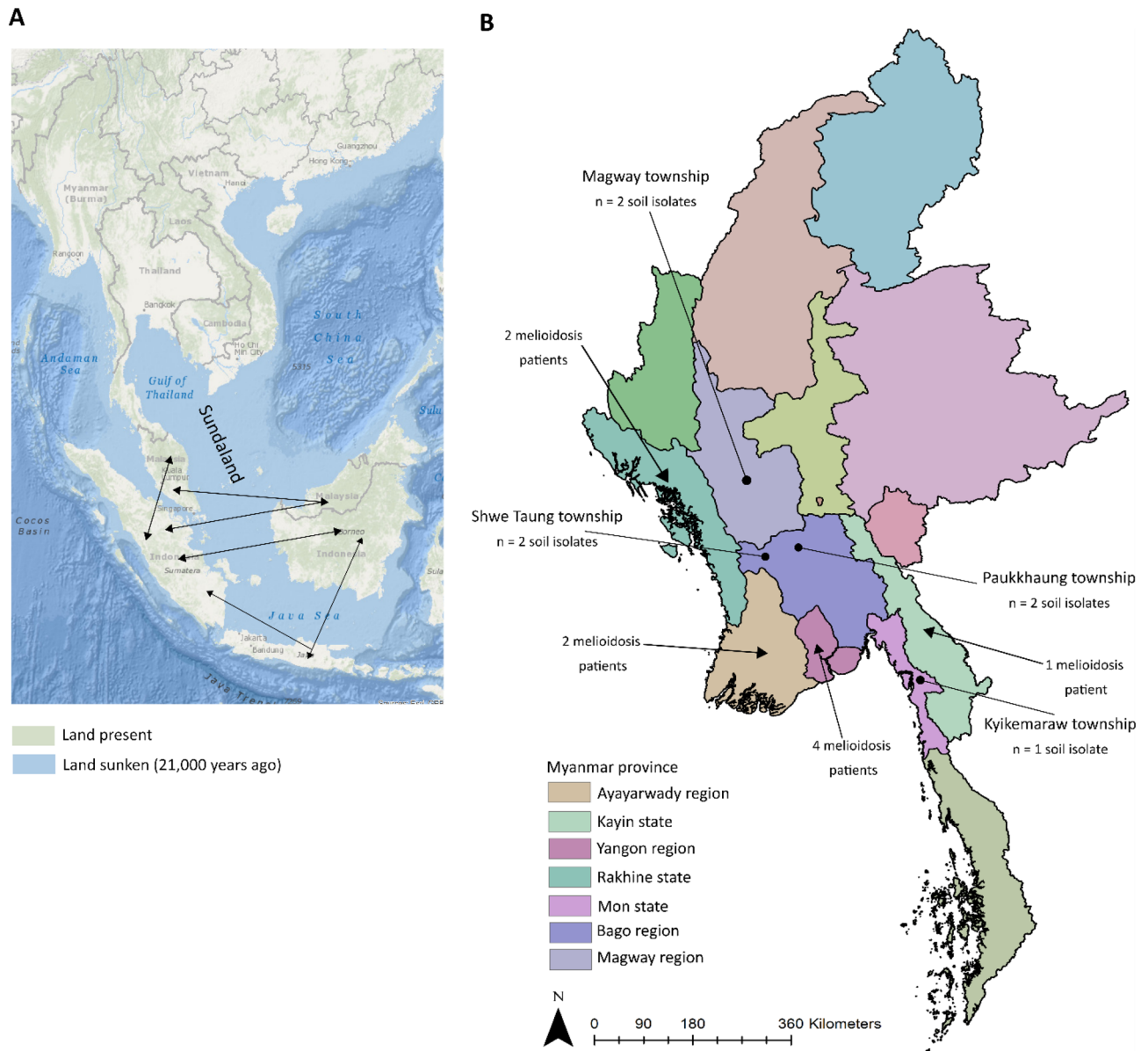


Figure 1. (A) Map of Sundaland region, demonstrating when a larger landmass was exposed in the region (including light blue and green shading) and after the Sunda shelf was inundated by rising sea levels (green shading only). Black arrows indicate potential dispersal routes for terrestrial organisms¹⁶. (B) Myanmar and its 14 regions, showing where *B. pseudomallei* was isolated from soil and regions where the nine melioidosis patients resided. Maps were created using ArcGIS V10.4.1 (<https://www.arcgis.com/index.html>).

were visualised and annotated using Interactive Tree of Life (ITOL) (<https://itol.embl.de>). Regions of recombination were predicted using Gubbins v2.3.4 (Genealogies Unbiased by Recombination In Nucleotide Sequences, v.2.3.1) (<https://github.com/sanger-pathogens/gubbins>) with default settings. The High-quality isolate assemblies were annotated using Prokka (v1.13) and proteomes were used for pan-genomic analysis, which was calculated for the 16 Myanmar isolates using Roary (v3.12.0)²⁶. Detection of virulence genes (LPS A (*wbiI* to *apaH* in K96243 [GenBank ref: NC_006350]), LPS B (*BUC_3392* to *apaH* in *B. pseudomallei* 576 [GenBank ref: NZ_ACCE01000003]), LPS B2 (*BURP840_LPSb01* to *BURP840_LPSb21* in *B. pseudomallei* MSHR840 [GenBank ref: GU574442]), *bimA_{Bm}* (*BURPS668_A2118* in *B. pseudomallei* MSHR668 [GenBank ref: NZ_CP009545]) *bimA_{Bp}* (*BPSS1492* in *B. pseudomallei* K96243 [GenBank ref: NC_006351.1]) and *fhaB3* (*BPSS2053* in *B. pseudomallei* K96243 [GenBank ref: NC_006351.1])) was determined using the Large scale BLAST score ratio (LS-BSR) pipeline with standard settings and blastn or tblastn²⁷. We used a BSR threshold of >0.9 for presence of virulence genes and scores below the cut-off were deemed absent or variable.

Isolate	Aliases	Year	Myanmar province	Below surface (cm)	Sample ID	ST	FhaB3	LPS type	BimA type
Human									
MSHR12626	Y3	2017	Ayeyarwady region	N/A	N/A	1371	Negative	LPS B	BimA _{Bp}
MSHR12627	Y21	2017	Ayeyarwady region	N/A	N/A	56	Negative	LPS A	BimA _{Bp}
MSHR12628	Y22	2017	Kayin state	N/A	N/A	56	Negative	LPS A	BimA _{Bp}
MSHR12629	Y42	2018	Yangon region	N/A	N/A	90	Positive	LPS A	BimA _{Bp}
MSHR12630	M72	2018	Yangon region	N/A	N/A	56	Negative	LPS A	BimA _{Bp}
MSHR12632	M92	2018	Yangon region	N/A	N/A	90	Positive	LPS A	BimA _{Bp}
MSHR12833	MM73014	2018	Yangon region	N/A	N/A	1765*	Positive	LPS A	BimA _{Bp}
MSHR12633	Y43	2018	Rakhine state	N/A	N/A	1371	Negative	LPS B	BimA _{Bp}
MSHR12634	Y33	2018	Rakhine state	N/A	N/A	346	Positive	LPS A	BimA _{Bp}
Environmental									
MSHR12636	755	2018	Kyikemaraw township, Mon state	60	MON 20-005	1752*	Positive	LPS A	BimA _{Bp}
MSHR12637	2119	2018	Paukkhaung township, Bago region	90	BGO 02-009	1753*	Positive	LPS B	BimA _{Bp}
MSHR12640	2128	2018	Paukkhaung township, Bago region	90	BGO 03-008	1753*	Positive	LPS B	BimA _{Bp}
MSHR12644	2139	2018	Shwe Taung township, Bago region	90	BGO 04-009	1753*	Positive	LPS B	BimA _{Bp}
MSHR12645	2143	2018	Shwe Taung township, Bago region	30	BGO 05-003	1753*	Positive	LPS B	BimA _{Bp}
MSHR12647	2844	2018	Magway township, Magway region	60	MGY 16-004	1770*	Positive	LPS A	BimA _{Bp}
MSHR12837	2897	2018	Magway township, Magway region	60	MGY 21-007	1766*	Positive	LPS A	BimA _{Bp}

Table 1. Myanmar isolates included in this study. Symbol: *novel ST; N/A not applicable/not available.

Results

Myanmar *B. pseudomallei* isolates are genetically diverse. In silico MLST of the 16 Myanmar genomes identified nine STs [ST56 (n = 3), ST90 (n = 2), ST346 (n = 1), ST1371 (n = 2), ST1752 (n = 1), ST1753 (n = 4), ST1765 (n = 1), ST1766 (n = 1) and ST1770 (n = 1)] and of these nine STs, five STs were novel and so far only found in Myanmar (Table 1). ST1753 was the only ST identified from the environment in more than one location, being identified in two towns in the Bago region separated by ~ 50 km. MLST revealed that none of the STs of the clinical isolates matched the STs of the environmental isolates, with phylogenetics confirming that the environmental isolates recovered from Myanmar in this study were not closely related to the patient isolates, with > 18,000 SNPs-indels separating clinical isolates from environmental isolates (Fig. 2). This was not unexpected as the soil isolates were not specifically collected in locations linked to the individual cases of melioidosis.

On interrogating the *B. pseudomallei* PubMLST database, four of the nine STs were shared by neighbouring countries: ST56, Bangladesh (seven isolates), Cambodia (two isolates), Thailand (four isolates) and Vietnam (two isolates); ST90, United Kingdom with history of travel to Asia (eleven isolates), a USA isolate with unknown travel history, China (one isolate); ST1371, India (two isolates) and ST346, France (one isolate). The original source of the isolates linked to the United Kingdom, USA and France, which were presumably imported, is not available. Draft genomes from three ST56 genomes, one from Bangladesh and two from Vietnam were available for comparison to our Myanmar ST56s (see next section).

Comparative genomics revealed that the Myanmar isolates were genetically diverse with 60,204 core genome SNPs and indels identified among the 16 Myanmar genomes (Fig. 2). The SNPs and indels were used to construct a Myanmar-only phylogenetic tree, revealing two distinct clades: clade one (seven isolates) and clade two (nine isolates) which were separated by 4157 SNPs-indels (Fig. 2). There was some evidence of phylogenetic structure, with environmental isolates from the same province clustered together (Fig. 2). For example, the four ST1753 isolates from Bago region (Shwe Taung and Paukkhaung) clustered together on clade two and were genetically similar with two SNPs-indels being the largest genetic difference in the ST1753 cluster. There were instances where isolates from the same region clustered together but were genetically dissimilar and belonged to different STs, with two soil isolates from Magway region clustering together but separated by 18,464 SNPs-indels.

For the clinical isolates, isolates belonging to the same ST clustered together, but in most instances corresponding patients resided in different states/regions. For example, three ST56 clinical isolates (MSHR12630,

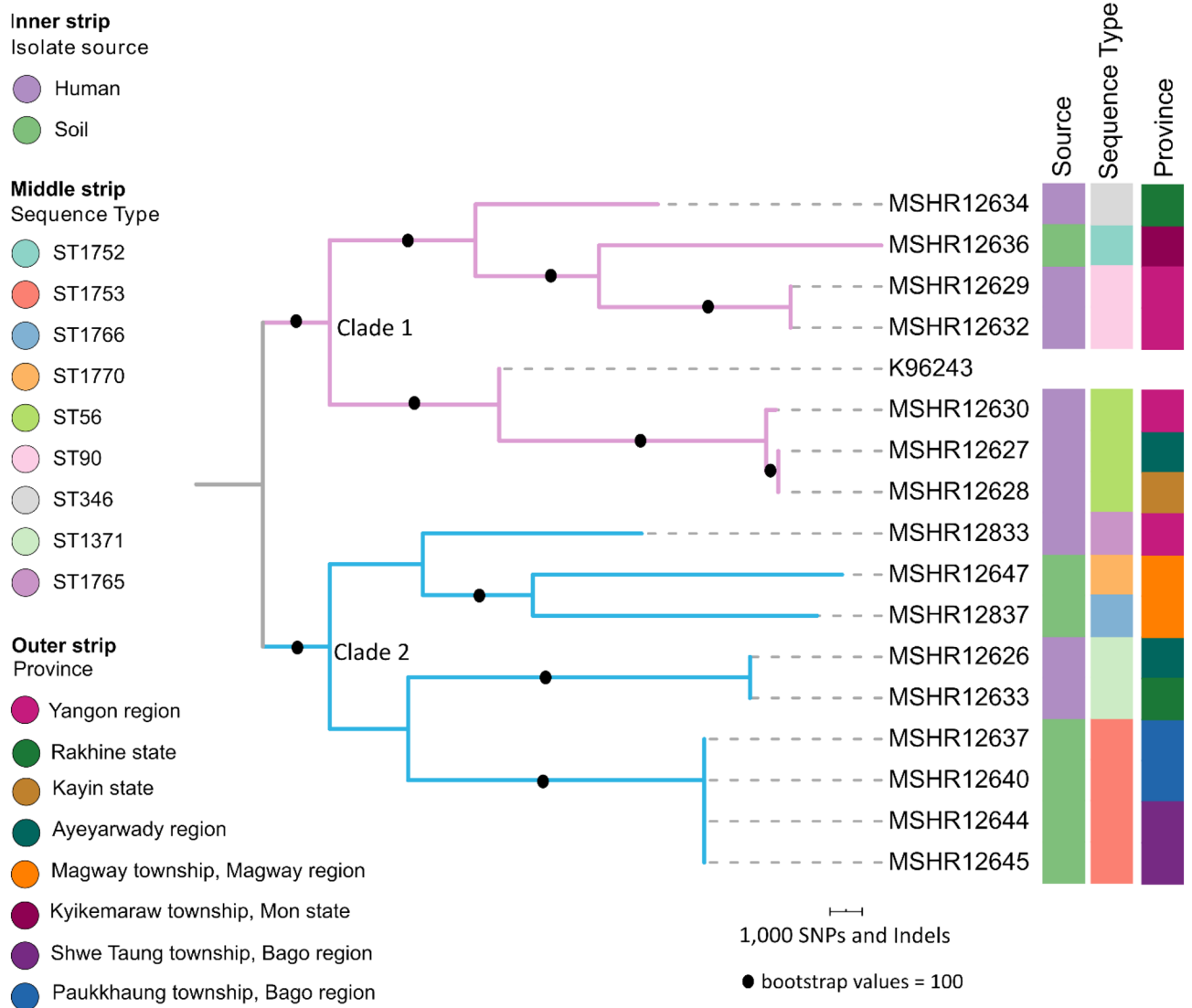


Figure 2. Midpoint rooted maximum parsimony phylogeny of *B. pseudomallei* (n = 16) from Myanmar, based on 60,204 core genome SNPs-indels.

MSHR12627 and MSHR12628) clustered together with the greatest genetic difference of six SNPs-indels, but patients resided in; Yangon region, Kayin state or Ayeyarwaddy region with the greatest distance difference of 291 km. In another example, two clinical isolates belonging to ST1371 differed by only four SNPs-indels but the corresponding patients resided many kilometres apart in Rakhine state and Ayeyarwaddy region. There was one instance where clinical isolates clustered together based on ST and geography: ST90 clinical isolates from two patients residing in Yangon region clustered together were separated by only five SNPs-indels.

We found considerable evidence of recombination among the 16 Myanmar isolates, (Figure S1B; recombination events are represented by the colour blocks) although removing regions associated with recombination did not significantly alter the phylogenetic structure of the isolates (Figure S1A).

Phylogenetic analysis reveals reintroductions into Myanmar from neighbours. Comparative analysis of the 16 Myanmar *B. pseudomallei* genomes with a global set of 173 genomes, identified 168,934 core genome SNPs-indels, which were used to construct a phylogenetic tree to identify the origin of *B. pseudomallei* from Myanmar. The Myanmar isolates (purple branches, Fig. 3) resided in two clades within the Asian clade. For the purpose of this study we defined clades containing Myanmar isolates as clade one and clade two (Fig. 3). Myanmar isolates located on clade one were the most ancient given their proximity to the ancient Australian isolates located at the tree base, and clade two was located further away from the tree base so these isolates were the least ancient. Clade one isolates had the longest branches out of the Myanmar isolates and so were the most diverse, with Myanmar isolates in clade two the least diverse with the shortest branches.

Phylogeography demonstrated that the Myanmar isolates shared nodes and clustered with isolates from neighbouring countries (Fig. 3). Clade one (thirteen Myanmar isolates); four Myanmar isolates shared a node with four Thai isolates (moderate bootstrap value 70%) and nine Myanmar isolates shared a node with a cluster

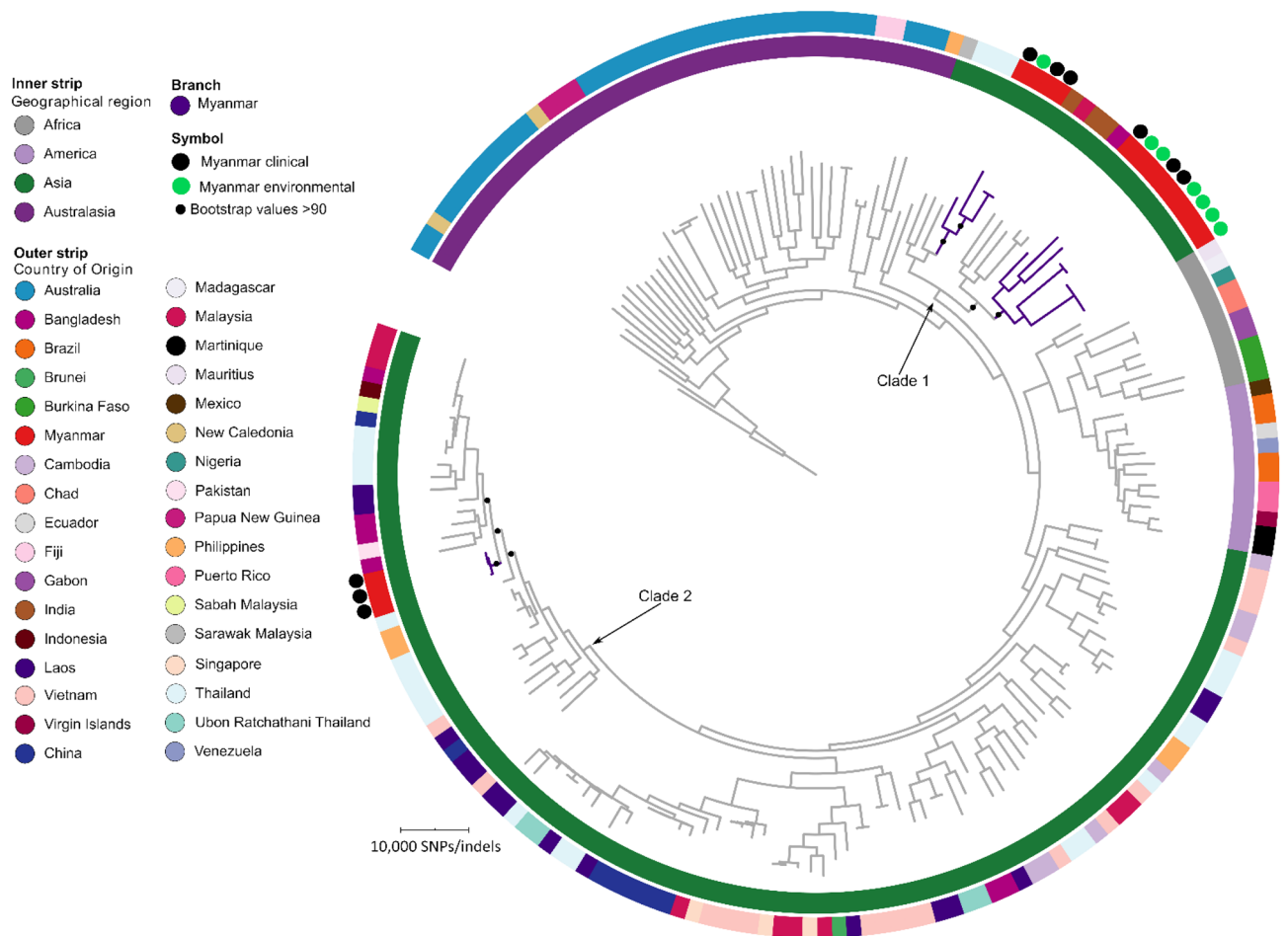


Figure 3. Maximum parsimony phylogeny of *B. pseudomallei* from Myanmar ($n=16$) with a global set of genomes ($n=173$) based on 167,781 core genome SNPs-indels. Rooted with MSHR0668, the most ancestral *B. pseudomallei* strain, as identified in a large phylogenetic study²⁹.

that contained isolates from India (three isolates), Bangladesh (one isolate) and Malaysia (one isolate) (bootstrap value 92%). Clade two (three ST56 Myanmar isolates): three Myanmar isolates shared a node with a cluster that contained isolates from Bangladesh (four isolates), Pakistan (one isolate: the true origin of this isolate is uncertain²⁸), Laos (two isolates), Thailand (four isolates), China (one isolate), Malaysia (four isolates) and Indonesia (one isolate) (well supported bootstrap value 97%), and descended from a Thai isolate. The closest isolate to a Myanmar isolate was a Thai isolate in clade two which differed by 4,317 SNPs-indels from a Myanmar isolate and the node was well supported (bootstrap value 97%). An ML global phylogeny supported the clustering observed here as did an Asian-only phylogeny (data not shown).

The ST56 isolates from Vietnam, Myanmar and Bangladesh were located in the Asian clade, but were distantly related to one another. The Myanmar clade containing the ST56s was separated from the ST56 from Bangladesh by 5,688 SNPs-indels (SID2889) and the ST56s from Vietnam by 12,102 SNPs-indels (MSHR6971) or 12,103 SNPs-indels (MSHR6972). Global phylogenetics demonstrated that the ST56 isolates from Vietnam were located nearest to the tree base, followed by ST56 isolates from Myanmar, with the ST56 isolate from Bangladesh the furthest away from the tree base.

Gene content of *B. pseudomallei* isolates from Myanmar. The virulence genes screened for in the Myanmar isolates varied. For BimA: all 16 Myanmar isolates contained BimA, *bimA_{Bp}*, for *fhaB3*: 11 (69%) isolates contained *fhaB3* and five (31%) lacked the gene, for LPS type: 10 (63%) isolates contained LPS A and the remaining six (37%) isolates contained LPS B. LPS A and LPS B were associated with different regions, with all environmental LPS B isolates from Bago region and LPS A from Magway region or Mon state (Table 1).

For the Myanmar pan-genome, which was calculated for the 16 isolates using Roary, the accessory genome contained heterogeneity indicating gene diversity among the Myanmar isolates, but the core genome (genes occurring in >99% of isolates) was stable (Figure S2). We identified a total of 9246 predicted coding sequences (CDS) with 5249 and 3997 assigned to the core and accessory genome. The number of core genes was comparable to that reported previously, but the number of accessory genes was greater than previously reported, (based on 37 *B. pseudomallei* genomes³⁰) highlighting the high genetic diversity of the Myanmar isolates.

Discussion

The scarceness of genotype and phylogenetic overlap between Australian and Asian *B. pseudomallei* is largely due to ancient oceanic biogeographical barriers, including Wallace's Line, restricting gene flow between populations³¹. The same cannot be said about *B. pseudomallei* isolates originating within Asia. Southeast Asian countries bordered by the Mekong River and the Malay Peninsula have shared genotypes and phylogenetic overlap¹²; this is likely due to the lack of geographical barriers, both as exist currently, but even more extensively during the last ice age within an expanded Sundaland (Fig. 1A). Here we present the first genomic analysis of *B. pseudomallei* from Myanmar, where melioidosis was first identified in 1911. Given that Myanmar shares geographical borders with its Southeast Asian neighbours, and its proximity to the Mekong River, we hypothesised, (1) high genetic diversity of *B. pseudomallei* from Myanmar, (2) genotype overlap between Myanmar and its neighbours and (3) multiple introduction events of *B. pseudomallei* into Myanmar.

Our MLST and phylogenetic analysis demonstrated that *B. pseudomallei* from Myanmar are genetically diverse and heterogeneous, with nine STs identified amongst only sixteen isolates. This finding is consistent with the high levels of genetic and ST diversity in *B. pseudomallei* from neighbouring countries including India³², China³³ and Thailand³⁴ and in the ancient *B. pseudomallei* population from the Northern Territory of Australia^{35,36}. Five of the Myanmar STs had not been previously identified, with four being environmental isolates. Furthermore, based on phylogenetics at a local level, environmental isolates clustered by geography, suggesting geographical restriction of *B. pseudomallei* within Myanmar, as demonstrated in Australia^{36–38}. However, more environmental sampling in Myanmar is necessary to elucidate whether genotype-geographical links exist and on what spatial scale. Four STs had previously been reported and are shared by neighbouring countries: ST56 ST90, ST346 and ST1371, and additional ST56 genomes were publicly available. Our phylogenetic analysis demonstrated that, despite the ST56 genomes being of Asian origin, the Myanmar ST56s are genetically distinct from those from Bangladesh and Vietnam (separated by 5,688 to 12,103 SNPs-indels) and did not all share a common ancestor, clustering in three phylogenetic locations. This represents another instance of ST homoplasy, which with increasing whole genome analyses is being recognised increasingly and which highlights the limitations of MLST, with unrecognised ST homoplasy suggesting incorrect associations between *B. pseudomallei* genotypes that confounds analysis of melioidosis epidemiology^{39–42}. The genetic distance of ST56 genomes from the Australasian node suggests that ST56 may have spread from Vietnam to Myanmar and then from Myanmar on towards Bangladesh, but additional ST56 isolates are needed to confirm this hypothesis.

It is now clear that *B. pseudomallei* arrived into Myanmar from elsewhere in Asia, with all Myanmar isolates nested within the Asian clade in our global phylogeny. Based on our global phylogenetic tree, the Myanmar isolates clustered within two distinct clades (representing two distinct populations) rather than in a single clade and so a recent common ancestor is not shared among all Myanmar isolates. This suggests that *B. pseudomallei* was introduced into Myanmar on at least two independent occasions. Based on our global phylogeny, it seems likely that *B. pseudomallei* initially arrived into Myanmar firstly from neighbours including India, Bangladesh and Malaysia and the node was well supported on our phylogeny (node shared with Myanmar isolates: strong bootstrap support 92%). With subsequent reintroductions from Thailand, Laos, Malaysia, India, Bangladesh, Pakistan, and China and the node on the phylogeny was well supported (node shared with Myanmar isolates: strong bootstrap support of 97%). In support of this our ST data showed overlap between isolates from Myanmar with isolates from Bangladesh, India, Thailand or China and so it was no surprise that Myanmar isolates grouped with isolates from those countries on the whole-genome level. Likewise, previous phylogenomic studies of *B. pseudomallei* have described repeated introduction of *B. pseudomallei* between countries bordered by the Mekong River and the Malay Peninsula¹², with the Mekong River representing a hotspot for *B. pseudomallei* evolution and transmission throughout Southeast Asia¹².

The observed connectivity, mode and time period of *B. pseudomallei* transmission into and within Myanmar is likely explained by severe weather events, especially flooding, geographical proximity to neighbours, human and animal migration, cultural links, prehistoric and current trading networks associated with the Mekong River. Our global phylogeny demonstrated that the Myanmar isolates located in clade 1 nearest to the tree base represent possibly the earliest introduction of *B. pseudomallei* into Myanmar, which could have been facilitated by the land bridges of Sundaland present during the last ice age (Fig. 1A). A similar theory has been postulated for *B. pseudomallei* populations in Papua New Guinea, the Sahul land bridge between Australia and New Guinea during the last ice age⁴³. It is also worthy of note, however that genetic analysis of humans from Myanmar and Southwestern China (~6,000 individuals) demonstrated that these individuals share haplogroups (sharing a common ancestor)⁴⁴, with individuals from China migrating to Burma likely along river valleys that connect Myanmar and China (e.g. Ayeyarwady River) around 25,000 years ago. This migration provides another potential for anthropogenically driven introduction of *B. pseudomallei* into Myanmar and fits with the period when *B. pseudomallei* was thought to have spread from Australia to Asia (between 16 and 225 thousand years ago¹²). Additionally, animal migration may have played a role in the introduction of *B. pseudomallei* to Myanmar, for example birds, which can both asymptotically carry *B. pseudomallei* and develop melioidosis^{45,46}. Myanmar has over 1,000 native bird species of which ~200 are migratory, so it is plausible that *B. pseudomallei* has been introduced into Myanmar by returning migratory birds. Whatever routes were responsible for the initial introduction of *B. pseudomallei* into Myanmar, the underlying genetic diversity demonstrated by our global phylogeny and extensive recombination supports the notion that it was not a recent event, with enough time elapsed since its introduction for divergence to occur. Future phylogenetic and molecular clock investigations using a larger number of Myanmar and Asian isolates will provide further clues as to the origin of *B. pseudomallei* in Myanmar.

In addition to our phylogeographic analysis, we screened the Myanmar *B. pseudomallei* assemblies against variably present virulence genes. *Burkholderia* intracellular motility factor A (BimA) is responsible for actin-based motility of *B. pseudomallei*⁴⁷ and all Myanmar isolates contained the more common *bimA*_{Bp}, which is

typically identified across Asia and has been associated with pneumonia⁴⁸. *bimA_{Bm}* has only been identified in Australian and Indian isolates and is associated with neurological melioidosis^{39,48–50}. Filamentous hemagglutinin (*fha*) is a surface protein that functions as an adhesin⁵¹. Although *B. pseudomallei* contains three *fha* genes that vary genetically, it is the presence of *fhaB3* variant that has been associated with positive blood cultures and the absence of *fhaB3* with localised skin abscess⁴⁸. *fhaB3* was variably present among the Myanmar isolates, absent from five and present in eleven, and similar variability has been noted in isolates from India, Australia and Thailand^{48,49}. Lastly, we investigated the LPS types. LPS of *B. pseudomallei* is required for serum resistance, with three types being described: LPS A, LPS B and LPS B2. LPS A was dominant amongst Myanmar isolates but LPS B was also detected and prevalence of LPS B in Myanmar was higher than in Thailand⁵². Interestingly, LPS B is the most common LPS type in India, a close neighbour of Myanmar⁴⁹.

In conclusion we have used WGS and large-scale comparative-genomics to describe for the first time the genetic diversity and phylogeography of *B. pseudomallei* in Myanmar, the country where melioidosis was originally recognised. We have demonstrated that *B. pseudomallei* from Myanmar are genetically diverse, supporting the long-term endemicity of *B. pseudomallei* in Myanmar. Phylogenomic analysis revealed that Myanmar *B. pseudomallei* cluster in the Asian clade linking to isolates from neighbouring countries, including those bordered by the Mekong River. It is thus likely that *B. pseudomallei* has been introduced into Myanmar from these close neighbours on more than one occasion. Further studies are required to elucidate the role of anthropogenic factors in both the past and evolving regional and global dispersal of *B. pseudomallei*.

Received: 23 July 2020; Accepted: 11 September 2020

Published online: 01 October 2020

References

- Currie, B. J., Ward, L. & Cheng, A. C. The epidemiology and clinical spectrum of melioidosis: 540 cases from the 20 year Darwin prospective study. *PLoS Negl. Trop. Dis.* **4**, e900 (2010).
- Wiersinga, W. J., Currie, B. J. & Peacock, S. J. Melioidosis. *N. Engl. J. Med.* **367**, 1035–1044 (2012).
- Whitmore, A. An account of a glanders-like disease occurring in Rangoon. *J. Hyg. (Lond.)*. **13**(1–34), 1 (1913).
- Whitmore, A. & Krishnaswami, C. S. A Hitherto undescribed infective disease in Rangoon. *Ind. Med. Gaz.* **47**, 262–267 (1912).
- Krishnaswami, C. S. Morphia injector's septicaemia. *Indian Med. Gaz.* **52**, 296–299 (1917).
- Win, M.M., Ashley, E.A., Zin, K.N., et al. Melioidosis in Myanmar. *Trop. Med. Infect. Dis.* **2018**, 3.
- Cox, C. D. & Arbogast, J. L. Melioidosis. *Am. J. Clin. Pathol.* **15**, 567–570 (1945).
- Aung, M. K. Case report: Melioidosis—A hidden disease in Myanmar. *Myanmar J. Curr. Med. Pract.* **5**, 57–59 (2000).
- Climate WW. World Weather & Climate. Accessed 13/07/2020.
- Limmathurtsakul, D., Golding, N., Dance, D.A., et al. Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. *Nat. Microbiol.* **2016**, 1.
- Win, M. M. et al. A study of *Burkholderia pseudomallei* in the environment of farms in Thanlyin and Hmawbi Townships, Myanmar. *Am. J. Trop. Med. Hyg.* **100**, 1082–1084 (2019).
- Chewapreecha, C. et al. Global and regional dissemination and evolution of *Burkholderia pseudomallei*. *Nat. Microbiol.* **2**, 16263 (2017).
- Sarovich, D.S., Garin, B., De Smet, B., et al. Phylogenomic analysis reveals an Asian origin for African *Burkholderia pseudomallei* and further supports melioidosis endemicity in Africa. *mSphere*. **2016**, 1.
- Hall, C. M. et al. *Burkholderia pseudomallei*, the causative agent of melioidosis, is rare but ecologically established and widely dispersed in the environment in Puerto Rico. *PLoS Negl. Trop. Dis.* **13**, e0007727 (2019).
- Cossaboom, C. M. et al. Melioidosis in a resident of Texas with no recent travel history, United States. *Emerg Infect Dis.* **26**, 1295–1299 (2020).
- Husson, L., Boucher, F. C., Sarr, A.-C., Sepulchre, P. & Cahyarini, S. Y. Evidence of Sundaland's subsidence requires revisiting its biogeography. *J. Biogeogr.* **47**, 843–853 (2020).
- Currie, B. J. et al. Use of whole-genome sequencing to link *Burkholderia pseudomallei* from air sampling to mediastinal melioidosis, Australia. *Emerg. Infect. Dis.* **21**, 2052–2054 (2015).
- Limmathurtsakul, D. et al. Systematic review and consensus guidelines for environmental sampling of *Burkholderia pseudomallei*. *PLoS Negl. Trop. Dis.* **7**, e21105 (2013).
- Mayo, M. et al. *Burkholderia pseudomallei* in unchlorinated domestic bore water, Tropical Northern Australia. *Emerg. Infect. Dis.* **17**, 1283–1285 (2011).
- Novak, R. T. et al. Development and evaluation of a real-time PCR assay targeting the type III secretion system of *Burkholderia pseudomallei*. *J. Clin. Microbiol.* **44**, 85–90 (2006).
- Currie, B. J. et al. Using BOX-PCR to exclude a clonal outbreak of melioidosis. *BMC Infect. Dis.* **7**, 68 (2007).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Sarovich, D. S. & Price, E. P. SPANDx: A genomics pipeline for comparative analysis of large haploid whole genome re-sequencing datasets. *BMC Res Notes*. **7**, 618 (2014).
- Swofford DL. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods) 4.0.b5*. (Sunderland, Sinauer Associates, 2001).
- Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
- Page, A. J. et al. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
- Sahl, J. W., Caporaso, J. G., Rasko, D. A. & Keim, P. The large-scale blast score ratio (LS-BSR) pipeline: A method to rapidly compare genetic content between bacterial genomes. *PeerJ*. **2**, e332 (2014).
- Mukhopadhyay, C., Shaw, T., Varghese, G.M., Dance, D.A.B. Melioidosis in South Asia (India, Nepal, Pakistan, Bhutan and Afghanistan). *Trop. Med. Infect. Dis.* **2018**, 3.
- Price, E. P., Currie, B. J. & Sarovich, D. S. Genomic insights into the melioidosis pathogen, *Burkholderia pseudomallei*. *Curr. Trop. Med. Rep.* **4**, 95–102 (2017).
- Spring-Pearson, S. M. et al. Pangenome analysis of *Burkholderia pseudomallei*: Genome evolution preserves gene order despite high recombination rates. *PLoS ONE* **10**, e0140274 (2015).
- Pearson, T. et al. Phylogeographic reconstruction of a bacterial species with high levels of lateral gene transfer. *BMC Biol.* **7**, 78 (2009).

32. Kamthan, A., Shaw, T., Mukhopadhyay, C. & Kumar, S. Molecular analysis of clinical *Burkholderia pseudomallei* isolates from southwestern coastal region of India, using multi-locus sequence typing. *PLoS Negl. Trop. Dis.* **12**, e0006915 (2018).
33. Fang, Y. *et al.* Multilocus sequencing-based evolutionary analysis of 52 strains of *Burkholderia pseudomallei* in Hainan, China. *Epidemiol. Infect.* <https://doi.org/10.1017/S0950268818002741>:1-6 (2018).
34. U'Ren, J. M. *et al.* Fine-scale genetic diversity among *Burkholderia pseudomallei* soil isolates in northeast Thailand. *Appl. Environ. Microbiol.* **73**, 6678–6681 (2007).
35. Price, E. P. *et al.* Unprecedented melioidosis cases in Northern Australia caused by an Asian *Burkholderia pseudomallei* strain identified by using large-scale comparative genomics. *Appl. Environ. Microbiol.* **82**, 954–963 (2016).
36. Rachlin, A. *et al.* 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 (2020).
37. McRobb, E. *et al.* Distribution of *Burkholderia pseudomallei* in northern Australia, a land of diversity. *Appl. Environ. Microbiol.* **80**, 3463–3468 (2014).
38. Chapple, S. N. *et al.* *Burkholderia pseudomallei* genotype distribution in the Northern Territory, Australia. *Am. J. Trop. Med. Hyg.* **94**, 68–72 (2016).
39. Webb, J. R. *et al.* Tracing the environmental footprint of the *Burkholderia pseudomallei* lipopolysaccharide genotypes in the tropical “Top End” of the Northern Territory, Australia. *PLoS Negl. Trop. Dis.* **13**, e0007369 (2019).
40. Morosini, M. I. *et al.* Melioidosis in traveler from Africa to Spain. *Emerg. Infect. Dis.* **19**, 1656–1659 (2013).
41. Garin, B. *et al.* Autochthonous melioidosis in humans, Madagascar, 2012 and 2013. *Emerg. Infect. Dis.* **20**, 1739–1741 (2014).
42. Aziz, A., Sarovich, D.S., Harris, T.M., *et al.* Suspected cases of intracontinental *Burkholderia pseudomallei* sequence type homoplasy resolved using whole-genome sequencing. *Microb. Genom.* **2017**, 3.
43. Baker, A. L. *et al.* *Burkholderia pseudomallei* distribution in Australasia is linked to paleogeographic and anthropogenic history. *PLoS ONE* **13**, e0206845 (2018).
44. Li, Y. C. *et al.* Ancient inland human dispersals from Myanmar into interior East Asia since the Late Pleistocene. *Sci. Rep.* **5**, 9473 (2015).
45. Choy, J. L., Mayo, M., Janmaat, A. & Currie, B. J. Animal melioidosis in Australia. *Acta Trop.* **74**, 153–158 (2000).
46. Hampton, V. *et al.* Melioidosis in birds and *Burkholderia pseudomallei* dispersal, Australia. *Emerg. Infect. Dis.* **17**, 1310–1312 (2011).
47. Sithidit, C. *et al.* Prevalence and sequence diversity of a factor required for actin-based motility in natural populations of *Burkholderia* species. *J. Clin. Microbiol.* **46**, 2418–2422 (2008).
48. Sarovich, D. S. *et al.* Variable virulence factors in *Burkholderia pseudomallei* (melioidosis) associated with human disease. *PLoS ONE* **9**, e91682 (2014).
49. Shaw, T., Tellapragada, C., Kamath, A., Kalwaje Eshwara, V. & Mukhopadhyay, C. Implications of environmental and pathogen-specific determinants on clinical presentations and disease outcome in melioidosis patients. *PLoS Negl. Trop. Dis.* **13**, e0007312 (2019).
50. Mukhopadhyay, C. *et al.* Molecular characterization of clinical *Burkholderia pseudomallei* isolates from India. *Am. J. Trop. Med. Hyg.* **85**, 121–123 (2011).
51. Sim, S. H. *et al.* The core and accessory genomes of *Burkholderia pseudomallei*: implications for human melioidosis. *PLoS Pathog.* **4**, e1000178 (2008).
52. Tuanyok, A. *et al.* The genetic and molecular basis of O-antigenic diversity in *Burkholderia pseudomallei* lipopolysaccharide. *PLoS Negl. Trop. Dis.* **6**, e1453 (2012).

Acknowledgements

We thank Vanessa Rigas from Menzies School of Health Research for assistance with laboratory work. We are grateful to Clare Ling for help with confirmation of the identity of some of the isolates. Thanks also to U Kyaw Soe, Saw Kan Sint Tun, Josh Cohen, staff from Medical Action Myanmar for their work on the environmental surveys. We would like to thank Direk Limmathuratsakul and Premjit Amornchai from Mahidol Oxford Tropical Medicine Research Unit who provided a practical training workshop for isolation and identification of *B. pseudomallei* in soil samples at the Department of Medical Research, Yangon, Myanmar.

Author contributions

The study was conceptualized by D.A.B.D., M.M.W., J.R.W and B.J.C. Funding was acquired by D.A.B.D. and M.M.W. Data collection was undertaken by D.A.B.D., M.M.W., K.N.Z., K.K.N.W., T.T.W., E.A.A., F.S and M.M.M.S. Formal analysis was done by J.R.W. J.R.W wrote the original draft, and D.A.B.D., M.M.W., B.J.C., E.A.A., F.S., M.M and M.M.M.S. revised and edited the manuscript.

Funding

The research was funded under the Australian National Health and Medical Research Council Grant numbers 1046812, 1098337, and 1131932 (The HOT NORTH initiative). DABD was supported by Wellcome funding of the Lao-Oxford-Mahosot Hospital- Wellcome Trust Research Unit (Grant number 106698/Z/14/Z). The Myanmar Oxford Clinical Research Unit is part of the MORU Tropical Health Network, funded by Wellcome Trust of Great Britain.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-73545-8>.

Correspondence and requests for materials should be addressed to J.R.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020