# Taking the next-gen step: comprehensive antibiotic resistance detection from *Burkholderia pseudomallei* genomes

**Danielle E. Madden[1,2], Jessica R. Webb[3], Eike J. Steinig[4], Mark Mayo[3], Bart J. Currie[3,5], Erin P. Price[1,2,3,*], and Derek S. Sarovich[1,2,3,*]**

[1]GeneCology Research Centre, University of the Sunshine Coast, Sippy Downs, Queensland, Australia; [2]Sunshine Coast Health Institute, Sunshine Coast University Hospital, Birtinya, Queensland, Australia; [3]Global and Tropical Health Division, Menzies School of Health Research, Tiwi, Northern Territory, Australia; [4]Australian Institute of Tropical and Health Medicine, James Cook University, Townsville, Queensland, Australia; [5]Department of Infectious Diseases and Northern Territory Medical Program, Royal Darwin Hospital, Tiwi, Northern Territory, Australia

This preprint was compiled on August 20, 2019

**Antimicrobial resistance (AMR) is emerging as a major threat to human health worldwide. Whole-genome sequencing (WGS) holds great potential for rapidly and accurately detecting AMR from genomic data in the diagnostic laboratory setting. However, most work to date has focussed on identifying only horizontally-acquired AMR-conferring genes, with chromosomally-encoded AMR determinants remaining largely undetected. Here, we present an improved tool for Antibiotic Resistance Detection and Prediction (ARDaP) from WGS data. ARDaP was designed with three priorities: 1) to accurately identify a wide range of AMR genetic determinants (i.e. horizontally-acquired gene gain, single-nucleotide polymorphisms, insertions-deletions, copy-number variation, and functional gene loss); 2) to predict enigmatic AMR determinants based on novel mutants with moderate- or high-consequence impacts in known AMR-conferring genes, and 3) to detect minor AMR allelic determinants in mixed (e.g. metagenomic) sequence data. ARDaP performance was demonstrated in the melioidosis pathogen, *Burkholderia pseudomallei*, due to its exclusively chromosomally-encoded AMR determinants and inherently limited treatment options. Using a well-characterised collection of 1,063 clinical strains, ARDaP accurately detected all currently known AMR determinants in *B. pseudomallei* (~50 determinants), including stepwise AMR mutations. Additionally, ARDaP accurately predicted meropenem resistance in four previously uncharacterised *B. pseudomallei* isolates. In mixed strain data, ARDaP identified AMR determinants down to ~5% allelic frequency, enabling the early detection of emerging AMR. We demonstrate that ARDaP is an accurate tool for identifying and predicting all confirmed *B. pseudomallei* AMR determinants from WGS data, including from mixed strain data. Finally, our study illustrates that manual cataloguing and functional verification of putative AMR determinants in individual pathogens is essential for truly comprehensive AMR detection. ARDaP is open source and available at: github.com/dsarov/ARDaP**

Antibiotic | Antimicrobial resistance | Whole genome sequencing | Next-generation sequencing | *Burkholderia pseudomallei* | Bioinformatics | Database | Efflux Pumps

Antimicrobial resistance (AMR) is a major threat to human health worldwide and an increasing contributor to morbidity and mortality [1]. Antibiotic use and misuse has resulted in an alarming increase in the number of multi-drug resistant infections emerging worldwide [2], resulting in an urgent need to improve global AMR detection and surveillance [1, 3]. In addition to pathogen identification, AMR detection is one of the primary goals of diagnostic microbiology, with far-reaching consequences for both infection control and effective treatment [4].

Whole-genome sequencing (WGS) enables the prediction of AMR profiles from bacterial pathogens based on their genomic sequence [5], circumventing the need for multiple and often laborious diagnostic methods, and with the potential to identify all AMR determinants in a single genome or metagenome [6, 7]. Although existing bioinformatic tools have proven effective for the detection of resistance genes acquired from horizontal gene transfer events [8], such as the SCC*mec* element in *Staphylococcus aureus* [9], many bacterial pathogens also develop AMR via chromosomal mutations [10, 11]. For instance, missense or nonsense single-nucleotide polymorphism (SNP) mutations in $\beta$-lactamase-encoding genes, in-frame or frameshift insertion-deletions (indels) in efflux pump regulators [12–14], and loss of outer membrane porins that decrease permeability of the cell membrane [15] can all cause AMR. Recent advances in AMR prediction software have integrated the detection of some chromosomal mutations in their algorithms [16, 17]. For example, Mykrobe Predictor is effective for identifying AMR-conferring SNPs in *S. aureus* and *Mycobacterium tuberculosis* [17]. Nevertheless, other genetic variants, including indels, gene loss or truncation, inversion, and gene amplification via copy-number variations (CNVs) have received little attention, despite their important role in conferring AMR [18].

To address these deficits in existing software, we present ARDaP, a bioinformatic tool for Antibiotic Resistance Detec-

* EP and DS contributed equally to the study.

No conflicts of interest are declared by the authors.

Correspondence: dsarovich@usc.edu.au

tion and Prediction from WGS data, which has been designed to detect both horizontally-acquired AMR genes and chromosomal mutations conferring AMR. We demonstrate the applicability and performance of ARDaP using the Tier 1 Select Agent pathogen, *Burkholderia pseudomallei* as a model organism due to i) its high intrinsic resistance towards many antibiotic classes, which greatly limits treatment options; ii) its exclusively chromosomally-encoded AMR determinants; and iii) its high mortality rate (10-40%), even with antibiotic treatment [12, 19]. *B. pseudomallei* causes melioidosis, with modelling estimating that 165,000 cases occur globally each year, resulting in potentially ~89,000 deaths [20]. Melioidosis severity ranges from mild, self-limiting skin abscesses to pneumonia, neurological disease, and septic shock, the latter of which has a mortality rate up to 95% [21]. Fortunately, *B. pseudomallei* transmission between humans is exceedingly rare [22], with almost all cases acquired from contact with a contaminated soil or water source [21]. As an environmentally acquired pathogen, isolates collected prior to antibiotic exposure (i.e. 'primary' isolates from melioidosis patients) are almost universally susceptible to the drugs used for melioidosis treatment; ceftazidime (CAZ), amoxicillin-clavulanate (AMC), trimethoprim/sulfamethoxazole (SXT), doxycycline (DOX), meropenem (MEM) and imipenem (IPM) [23]. However, melioidosis requires prolonged antibiotic therapy of at least three months to prevent relapse, which can lead to resistance acquisition over the course of a *B. pseudomallei* infection [23]. Although acquired AMR in *B. pseudomallei* has conventionally been considered an uncommon event, it has now been reported for all clinically-relevant antibiotics [12, 24], and new AMR determinants towards these key antibiotics continue to be uncovered. The consequences of AMR development in *B. pseudomallei* are significant, being linked to treatment failure and higher mortality rates in melioidosis patients [12].

Due to a clear need to improve the diagnosis and treatment of AMR *B. pseudomallei* infections, ARDaP was designed with two main aims: first, to accurately identify all currently known AMR genetic determinants in *B. pseudomallei*, including gene gain, SNPs, indels, CNVs, and gene loss or truncation, and second, to predict enigmatic AMR determinants in isolates with phenotypically-confirmed AMR. The predictive component of ARDaP reports novel high-consequence variants (i.e. nonsense mutations) in known AMR genes to identify candidate AMR determinants for further investigation. ARDaP was initially validated using a panel of 23 *B. pseudomallei* isolates with characterised AMR determinants and associated phenotypic data, with ARDaP being subsequently tested across a collection of 1,040 primary clinical *B. pseudomallei* isolates. We also tested ARDaP on synthetic strain mixtures at varying ratios to determine the lower limits of AMR detection from polyclonal sequence data, and a naturally-occurring AMR mixture in MSHR9021 [12] to verify the mixture-aware functionality of ARDaP.

## Results and Discussion

The alarming rise of AMR infections, which are associated with high economic burden and poor clinical outcomes, has highlighted the need for accurate, comprehensive, and rapid AMR diagnostics [25]. The efficiency, accuracy, and affordability of next-generation sequence technologies [16] has facilitated genomic- [6, 26] and transcriptomic- [25] based approaches for the personalised diagnosis and treatment of AMR infections [27], and has enabled coordinated, near-real-time AMR surveillance on a global scale [3]. Here, we developed a new bioinformatic tool, ARDaP, to enable the detection of both horizontally-acquired AMR genes via the Comprehensive Antibiotic Resistance Database (CARD) [3, 28], and chromosomally-encoded AMR determinants encoded by SNPs, indels, CNVs, and gene loss/truncation from WGS data. We chose the Tier 1 Select Agent and melioidosis pathogen, *B. pseudomallei*, as a model organism for several reasons: i) infection occurs via environmental acquisition only, meaning that AMR only arises in response to certain selective pressures (e.g. in its host during antibiotic treatment, or in the environment) and is not transferred horizontally among its human or animal hosts; ii) AMR in *B. pseudomallei* is exclusively chromosomally-encoded, rendering existing software insufficient for AMR detection; and iii) AMR development in *B. pseudomallei* results in treatment failure, higher mortality rates, and limited treatment options [12].

To assess the performance of ARDaP, this tool was validated using 23 previously characterised AMR *B. pseudomallei* strains with known antibiotic minimum inhibitory concentrations (MICs) (Table 1). These strains represent the spectrum of known AMR determinants in *B. pseudomallei* [12, 24, 29–40], with at least one strain being resistant towards one or more of the clinically-relevant antibiotics (Table 2). As the AMR profiles and AMR determinants for many of these strains have previously been characterised, we were able to accurately assess the ability of ARDaP to identify known AMR determinants, and to ignore genetic variants that do not cause AMR (Table 2). ARDaP correctly identified all AMR determinants (Table 1) with 100% specificity, with the exception of one false negative and three false positives. The false negative result was an 800kb inversion in strain 354e, which directly impacts *llpE*, a gene that resides between the efflux repressor gene, *bpeT*, and its associated efflux pump, encoded by *bpeEF-oprC* [40]. This inversion was not detected due to limitations in short-read Illumina data that led to this structural variant not easily being identified by Pindel, coupled with the occurrence of this inversion outside of an AMR determinant. One of the false-positive strains, MSHR5654, was predicted to be MEM resistant according to ARDaP, and two other strains, MSHR5666 and MSHR5669, were predicted to be DOX resistant, yet Etests showed sensitivity towards these respective antibiotics. MSHR5654, which was isolated from a cystic fibrosis patient with a chronic infection [30, 41], encodes a $BpeT_{Thr314fs}$ variant. Although this strain is considered MEM sensitive, when compared with wild-type strains, it exhibited an elevated MEM MIC (2 µg/mL) that was still below the resistance threshold (3 µg/mL). *bpeT* is a LysR-type transcriptional regulator that controls expression of the resistance-nodulation-division (RND) efflux pump, *bpeEF-oprC* [42]. Alterations in *bpeT* have previously been linked with MEM resistance in MSHR1300 (4 µg/mL) [12] and 354e (6 µg/mL) [12]. However, MSHR1300 also encodes an $AmrR_{K13fs}$ variant that likely causes the MEM resistance [12], and in 354e, the 800kb inversion that displaced *bpeT* by ~800kb may also affect other AMR-conferring genes. Our study supports prior work [29] suggesting that the contribution of the bpeEF-oprC efflux pump in conferring AMR in *B. pseudomallei* is currently poorly understood, with additional

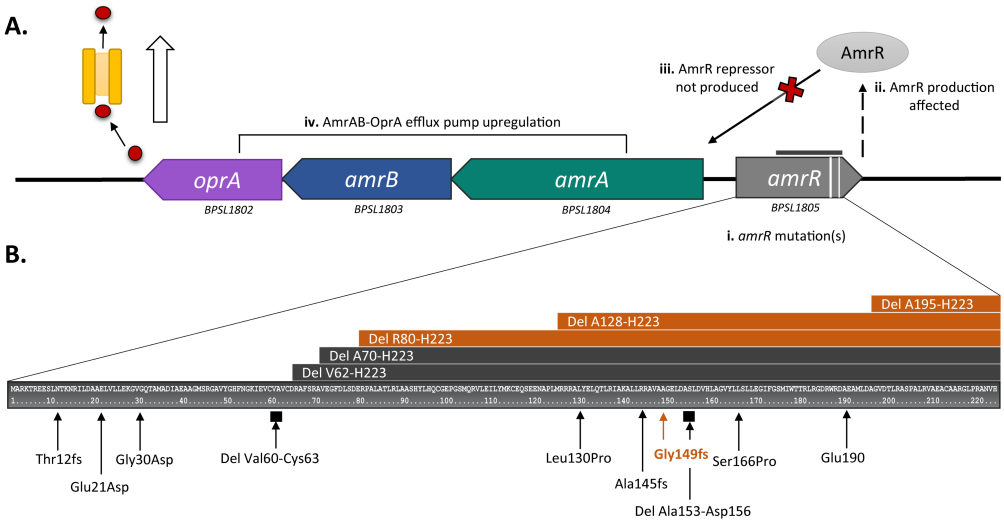


**Fig. 1.** Operon organisation of the *Burkholderia pseudomallei* AmrAB-OprA resistance-nodulation-division efflux pump and loss-of-function mutations in its TetR-type regulator, AmrR. **A.** Transcriptional organisation of the *amrR* (*BPSL1805*), *amrA* (*BPSL1804*), *amrB* (*BPSL1803*) and *oprA* (*BPSL1802*) operon, and summary of how (i) AmrR mutations cause (ii) loss-of-function of AmrR, which (iii) no longer represses expression of the resistance-nodulation-division AmrAB-OprA efflux pump, resulting in (iv) efflux pump over-expression and resistance to meropenem and aminoglycoside antibiotics. **B.** Distribution and annotation of AmrR mutations. Eleven previously observed AmrR mutations (in black) [12] have been augmented with four novel mutations identified in the current study (orange); $AmrR_{G149fs}$, $AmrR_{\Delta P81\text{-}H223}$, $AmrR_{\Delta A128\text{-}H223}$, and $AmrR_{\Delta A195\text{-}H223}$, all of which cause AmrR loss-of-function, resulting in efflux pump overexpression and antimicrobial resistance.

work needed to confirm the role of *bpeT* in MEM resistance. In line with these collective observations, we have modified ARDaP to flag *bpeT* variants as stepwise mutations towards MEM, rather than conferring MEM resistance in their own right (Table 1; Table 2).

The two DOX false-positive strains, MSHR5666 and MSHR5669, were from sputum collected from a chronically-infected patient with cystic fibrosis, CF9 [30]. Both isolates encode a SAM-dependent methyltransferase variant, $BPSL3085_{A88fs}$ [44]. We also observed $BPSL3085_{A88fs}$ in an unrelated DOX-resistant strain, Bp1651 (Table 1), which was also cultured from sputum from a chronically-infected patient with cystic fibrosis [24]. BPSL3085 mutations have been shown to confer DOX resistance likely due to altered ribosomal methylation patterns [30, 31]. However, both MSHR5666 and MSHR5669 remained DOX-sensitive (1.5 µg/mL) despite other strains from CF9 encoding $BPSL3085_{A88fs}$ and being DOX-resistant (MSHR5665: MIC=6 µg/mL; MSHR5667: MIC=48 µg/mL; Table 1) [30]. The higher DOX MIC in MSHR5667 is attributable to an $AmrR_{L132P}$ mutation in combination with $BPSL3085_{A88fs}$ (Table 1). We postulate that MHSR5666 and MSHR5669 encode an unidentified compensatory mutation that reverts them to a DOX-sensitive phenotype, despite encoding the $BPSL3085_{A88fs}$ variant. Notably, all longitudinal CF9 isolates, including MSHR5666 and MSHR5669, encode *mutS* (*BPSL2252*) mutations, resulting in a hypermutator phenotype [30, 44]. Unfortunately, identifying the causative basis for this reversion is non-trivial due to the large number of mutations (range: 97-157) accrued by these hypermutator strains [30], with further work needed to identify the specific mutant/s responsible for this phenomenon. As with any software designed for AMR detection from WGS data [3, 16], ARDaP is only as comprehensive as its underlying AMR database. Fortunately, as new AMR determinants in *B. pseudomallei* are identified and verified, ARDaP can easily be updated to incorporate these new determinants.

Our study demonstrates that accurate prediction of novel chromosomal AMR determinants requires cataloguing of natural variation in the antibiotic-sensitive strain population to avoid false-positive AMR calls. The inclusion of 1,040 primary (i.e. predominantly pre-antibiotic treatment) *B. pseudomallei* genomes in our study enabled robust investigation of all putative AMR determinants described in the literature to date. The rationale for using a large primary isolate dataset was that AMR determinants would not be present in this strain cohort due to their antibiotic-sensitive nature towards the drugs used to treat melioidosis, except in cases where the patient had begun receiving antibiotic treatment prior to primary isolate retrieval. Of the 1,040 primary strains, ARDaP predicted AMR in 22 strains, with the remaining strains classed as antibiotic-sensitive. The majority (17/22; 77%) of AMR strains possessed a Ser72Phe mutation in the PenA $\beta$-lactamase (K96243 numbering: $PenA_{S78F}$; encoded by *BPSS0946*), which has previously been linked to AMC resistance [24, 35, 36]. To investigate further, we performed MIC testing on 10 of these strains, which revealed that all were sensitive towards AMC (MIC=1.5 µg/mL). These results confirm that $PenA_{S72F}$ does not by itself confer AMC resistance, with this variant present in the wild-type *B. pseudomallei* population at a rate of ~1.6%. Previous work has suggested that either over-expression of PenA or $PenA_{S72F}$ can lead to AMC resistance [33, 35]; however, the exact MICs differ between studies, and further complicating this issue, the contribution of PenA expression differences versus enzyme modifications is variable [33, 35, 45]. We therefore propose that AMC resistance is conferred by both $PenA_{S72F}$ and PenA up-regulation, the latter of which can be caused by mutations within the 5' untranslated region [33], *penA* CNVs [30], or as-yet-undiscovered in trans regulatory changes. To account for this possibility, we have included the $PenA_{S72F}$ variant as a putative stepwise AMR variant in the ARDaP database (Table 2), with an additional PenA mutation required to defini-

**BURKHOLDERIA PSEUDOMALLEI GENOME SEQUENCING REPORT**
NOT FOR DIAGNOSTIC USE

| | | | |
|---|---|---|---|
| Patient Name | JAMES SMITH | Barcode | BARCODE |
| Birth Date | 1/01/1990 | Patient ID | MSHR3683 |
| Location | DARWIN | Sample Type | BLOOD |
| Sample Source | BLOOD | Sample Date | 2019-06-27 |
| Sample ID | MSHR3683 | SequencedFrom | CULTURED ISOLATE |
| Reporting Lab | RDH | Report Date/Time | 2019-06-27 |
| Requested By | DR. REQUESTOR NAME | Requester Contact | CONTACT@GENOME.COM |

**Summary**

The specimen was positive for **Burkholderia pseudomallei**. It is predicted to be **resistant to Doxycycline and Gentamicin**.

**Drug Susceptibility**

Resistance is reported when a high-confidence resistance-conferring mutation is detected. **"No mutation detected" does not exclude the possibility of resistance**

☐ No drug resistance predicted
☐ Mono-resistance predicted
☑ Multi-drug resistance predicted
☐ Extensive drug resistance predicted

| Drug class | Interpretation | Drug | Resistance Gene (Amino Acid Mutation) |
|---|---|---|---|
| First-line | Susceptible | Meropenem | No resistance detected |
| | | Ceftazidime | No resistance detected |
| Second-line | Susceptible | Imipenem | No resistance detected |
| | | Trimethoprim/sulfamethoxazole | No resistance detected |
| | | Amoxicillin/Clavulanic acid | No resistance detected |
| | Resistant | Doxycycline | Multiple mechanisms |

**Fig. 2.** *B. pseudomallei* AMR clinical report produced by ARDaP. Adapted from [43]. The final step in the ARDaP pipeline is the production of a clinician-friendly report that summarises patient and sample details, confirms the given isolate is *B. pseudomallei* and denotes any predicted AMR with what mutation has been detected and what antibiotic/s (first- or second-line) have been affected.

tively call AMC resistance in a given strain. We anticipate that future studies will assist with elucidating the precise role of $PenA_{S72F}$ in AMC resistance, at which point the *B. pseudomallei* ARDaP database can be adjusted accordingly.

IPM was the first carbapenem antibiotic used in melioidosis studies [46]; however, in most guidelines, IPM has been replaced by meropenem due to its lower neurotoxicity [22]. Reported IPM resistance rates are exceedingly low [47], and recent documentation of MEM-resistant *B. pseudomallei* infections has resurrected the potential role of IPM as an attractive alternative for treating such infections due to no cross-resistance between IPM and MEM [12]. In 2017, Bugrysheva and coworkers reported a $PenA_{T147A}$ mutation (K96243 numbering: $PenA_{T153A}$) in Bp1651, which raised the IPM MIC to 8 µg/mL when expressed at high levels in this strain [24]. We subsequently refuted the role of this variant in IPM resistance by identifying sensitive IPM MICs in three $PenA_{T147A}$-encoding strains [12]. In the current study, we provide further evidence that this variant does not cause IPM resistance as most primary DPMS strains (789/1,040; 76%) encoded the $PenA_{T147A}$ variant. However, it remains possible that this variant confers AMR in a stepwise manner with other PenA mutations, particularly those leading to *penA* upregulation. Given that $PenA_{T147A}$ occurs at a very high rate in the wild-type *B. pseudomallei* population, and that it may lower the barrier for IPM resistance emergence, we have included this mutant as a stepwise variant in our ARDaP database to facilitate the putative detection of potential IPM resistance, especially if identified alongside *penA* upregulation (Table 2). It is important to note that additional work is needed to verify this putative pathway to IPM resistance, as the basis for IPM resistance in *B. pseudomallei* is currently speculative. In addition to the IPM AMR-associated PenA variants reported in Bp1651, a novel mutation in the PenA $\beta$-lactamase, $PenA_{D239G}$ (K96243 numbering: $PenA_{D245G}$; Bugrysheva et al. numbering: $PenA_{D240G}$), has been linked to CAZ resistance (>128 µg/mL) [24]. We interrogated our DPMS primary dataset for $PenA_{D239G}$ to determine its frequency in primary

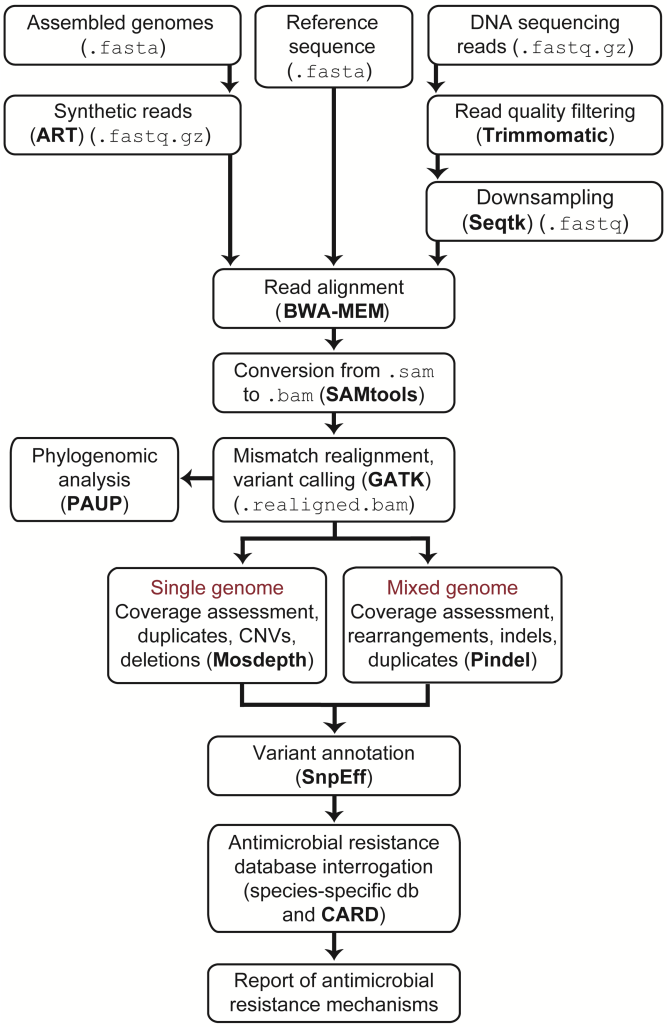


**Fig. 3.** Overview of the ARDaP pipeline. The user inputs assembled genome/s or raw sequencing reads, and a reference genome sequence. ARDaP then performs read alignment, read processing, mismatch realignment and variant identification. An optional phylogenetic analysis is also performed (if specified). Coverage assessment is undertaken on either single or mixed genomes (if specified), genetic variants are then annotated and AMR databases are interrogated. Finally, ARDaP produces a summary report of the detected and annotated AMR determinants for each strain (Figure 2).

isolates. None encoded $PenA_{D239G}$, consistent with the hypothesis that this variant is likely causal for CAZ resistance. CAZ resistance in *B. pseudomallei* has also been linked to various mutations in PenA ($PenA_{P167S}$ [36], $PenA_{C69Y}$ [34], *penA* 10x CNV [30], and *penA* -78 G→A [33]), and penicillin-binding protein 3 loss [32]. None of these CAZ resistance-conferring determinants were identified in our DPMS dataset, indicating CAZ sensitivity in all primary isolates.

ARDaP detected known AMR markers in two strains: $BPSL3085_{Ala88fs}$ in MSHR3683, and AmrR loss in MSHR1043 (Table 3). Subsequent review of clinical histories found that both patients had received antibiotic treatment prior to isolate retrieval, confirming that AMR emergence in these patients was very likely caused by antibiotic selective pressure. MSHR1043 represents an exception to the rule, however, due to the presence of a second mutation that overrides the predicted AMR phenotype. Ordinarily, the AmrR mutation in this strain should cause MEM resistance due to AmrAB-OprA de-repression [12]. However, we have previously shown that MSHR1043 exhibits unusual aminoglycoside sensitivity (MIC~1 µg/mL) due to an $AmrA_{L247fs}$ mutation, resulting in *AmrAB-OprA* loss-of-function and a MEM MIC of just 0.75 µg/mL (Table 3) [41]. AmrAB-OprA loss-of-function variants have also been described in Bp1651 ($AmrB_{A254fs}$), and the $AmrB_{T367R}$ variant is naturally present in many Malaysian Borneo strains from Sarawak [38]. Because of their unusual aminoglycoside sensitivity, all AmrA and AmrB mutated strains fail to grow on Ashdown's agar, which includes 4 µg/mL gentamicin (GEN) for *B. pseudomallei* selection [48]. Given their ability to override predicted AMR genotypes, we incorporated these antimicrobial-sensitive genotypes into ARDaP to better reflect the strain phenotype. These unusual variants may provide additional treatment options for melioidosis patients who otherwise could not be treated with aminoglycosides due to inherent resistance to this antibiotic class. AmrA or AmrB mutants are also rendered MEM-sensitive, meaning that infections with such strains are at a far lower risk of developing MEM resistance than wild-type strains. The inclusion of sensitivity-conferring variants is thus just as important as including AMR determinants, and should be a considered in all

AMR prediction databases.

AmrR, the regulator responsible for controlling the expression of the RND efflux pump AmrAB-OprA, is a mutational hotspot in *B. pseudomallei*, with multiple frameshift mutations and deletions previously found to be associated with MEM resistance [12] (Figure 1). In the remaining three strains, ARDaP predicted three novel AmrR mutations that all resulted in AmrR loss-of-function, either via aberrant or truncated AmrR (Figure 1), leading to AmrAB-OprA de-repression and subsequent MEM and aminoglycoside resistance. Consistent with this prediction, these novel variants ($AmrR_{\Delta P81\text{-}H223}$ in MSHR1058; $AmrR_{A128fs}$ in MSHR8777; $AmrR_{G149fs}$ in MSHR1174) were phenotypically confirmed by MEM MIC testing (range: 4-12 µg/mL; Table 3). Evaluation of patient records revealed that all had received MEM treatment prior to isolate retrieval. In the case of MSHR8777, the patient was treated with MEM for six days prior to isolate retrieval. Similarly, MSHR1058 was obtained from a relapsed melioidosis case, with this patient having previously received MEM injected directly into an infected melioidosis lung cavity. MSHR1174 was also collected following MEM treatment, during a period of prolonged blood culture positivity (approx. 6 weeks), which was correlated with MEM resistance development [12]. These findings provide further confirmation of a link between MEM treatment and potential treatment failure due to AmrR mutability [12], and demonstrate the value of ARDaP for identifying novel AMR determinants in *B. pseudomallei*.

Our ARDaP algorithm is mixture-aware, an important feature for detecting emerging AMR determinants in both single strains (e.g. emerging AMR determinants from non-purified colonies) and in mixed sequence data (e.g. culture sweeps or metagenomic data). To verify the performance of the mixture function in ARDaP, Illumina reads from both sensitive and resistant *B. pseudomallei* strains with known AMR status (Table 4) were mixed at ratios ranging from 5:95 to 95:5, respectively. In addition, we tested ARDaP performance on a previously detected AmrR mixture from strain MSHR9021, which was obtained from a DPMS patient six days after MEM treatment and which had evolved two AmrR mutations, $AmrR_{S166P}$ and $AmrR_{A145fs}$, at ~66% and 33% allele frequency, respectively (Table 1). For MSHR9021, we detected the $AmrR_{S166P}$ and $AmrR_{A145fs}$ variants at allele frequencies of 63% and 31%, respectively, which closely reflects their previously estimated proportions [12]. For the synthetic mixture dataset, ARDaP identified three AMR determinants down to the lowest tested ratio of 5% minor allele frequency: a *penA* 10x CNV from MSHR8441, and a *penA* 30x CNV and $PenA_{C69Y}$ in MSHR5654 (Table 4). The other determinants were identified at minor allele frequencies of 10% ($Ptr1_{R21fs}$, $Ptr1_{A22_G23ins_R\text{-}R\text{-}A}$, $BpeT_{T314fs}$, and $GyrA_{Y77S}$) 15% ($BPSL3085_{S130L}$), and 50% ($AmrR_{\Delta V62\text{-}H223}$) (Table 4). The high sensitivity of the *penA* CNVs and $PenA_{C69Y}$ can be explained by the high copy numbers of this gene in MSHR8441 and MSHR5654. The $PenA_{C69Y}$ missense variant likely has a sensitivity closer to 10-15% when present as a single copy, as seen with the $GyrA_{Y77S}$ and $BPSL3085_{S130L}$ mixtures (Table 5). The gene truncations had the lowest sensitivity in the mixed dataset, with allelic frequency requiring a higher threshold than other variant types due to the challenge of discriminating gene loss from normal sequence coverage variation, and the inherent limitations of short-read data. Overall, ARDaP confidently identified all mixtures, albeit with varying sensitivities. Further validation on specific variant mixtures is recommended when new mixtures are identified to determine their sensitivity. Deeper sequencing (e.g. 100-500x) should enable more robust mixture detection at lower allele frequencies.

The final output from ARDaP is the generation of an easy-to-interpret report that summarises the AMR determinants and associated antibiotic phenotype/s, if applicable, for each genome under investigation. This clinician-friendly report, which is based on the format developed by Crisan and colleagues [43], summarises AMR findings for both first- and second-line antibiotic treatments (Figure 2), and has been designed to prioritise the clinical workflow. AMR results are ordered hierarchically to impart relevant information at a quick glance, enabling end users such as clinicians to make informed decisions about patient treatment based on the reported AMR profile without requiring a detailed understanding of the underlying AMR determinants and their mechanisms of action. This ARDaP report also lists stepwise AMR determinants, which are critical to report as they can inform early treatment shifts that minimise the risk of AMR emergence. This easy-to-interpret report represents a major improvement over current software for AMR annotation, which requires an intimate understanding of AMR determinants and associated AMR mechanisms to disambiguate technically-detailed outputs. The simple and accurate AMR report produced by ARDaP represents an important step towards the incorporation of WGS as a routine tool for guiding best-practice AMR stewardship and personalised treatment regimens in the clinical diagnostic setting.

## Conclusion

WGS is becoming a routine and essential tool for the rapid detection of AMR determinants as it provides the best diagnostic method to mitigate the predicted devastating impact of AMR on global health in the coming decades. Whilst existing software can readily detect horizontally-acquired AMR genes, these tools currently have limited capacity to detect AMR determinants caused by chromosomal mutation, leading to substantial underreporting of AMR in many bacterial pathogens. To overcome this major limitation in AMR research, we developed ARDaP, which detects AMR caused by both horizontal gene acquisition and chromosomal mutation events. As the mechanisms by which pathogens develop AMR are diverse and complex, we tailored ARDaP for the detection of >50 *B. pseudomallei* known and novel AMR determinants, although this tool has been designed for implementation with any pathogen of interest. ARDaP also incorporates a mixture-aware feature that enables the detection of emerging AMR determinants in genomic data, which can inform early treatment shifts that will ultimately improve antibiotic stewardship efforts and patient survival. Our work demonstrates that novel AMR determinants can be reliably predicted from WGS data; however, large collections of susceptible strains and functional verification (e.g. gene knockouts, heterologous expression, or RNA sequencing) are still required to validate novel AMR determinants. Finally, we assert that software designed for AMR detection from WGS data requires well-curated pathogen-specific databases for the most accurate, comprehensive, and relevant AMR detection. Of utmost importance in ongoing research efforts is to uncover

the myriad ways that bacteria evolve to evade antibiotics.

# Methods

## Ethics

This study was approved by the Human Research Ethics Committee of the Northern Territory Department of Health and Menzies School of Health Research (HREC 02/38). Culture conditions, DNA isolation, and WGS. Culture conditions, DNA extraction, and WGS were performed as outlined previously [49].

## Minimum Inhibitory Concentration (MIC) Determination

MICs were determined using Etests (bioMérieux, Murarrie, Australia) with sensitive, intermediate and resistant cut-offs based on the Clinical and Laboratory Standards Institute (CLSI) M100-S17 guidelines for *B. pseudomallei* (i.e. ≤8/4, 16/8, and ≥32/16 µg/mL for AMC; ≤8, 16, ≥32 µg/mL for CAZ; ≤4, 8, ≥16 µg/mL for DOX and IPM; ≤2/38, nil, ≥4/76 µg/mL for SXT). The CLSI guidelines do not list MEM MIC values for *B. pseudomallei*; however, based on previous testing of 234 primary *B. pseudomallei* isolates [50], and establishment of an epidemiological cut-off value [51], we categorised MEM resistance as MIC 3 µg/mL. This value is identical to recent proposed EUCAST breakpoints for *B. pseudomallei* [52]. Likewise, the CLSI guidelines do not list GEN MIC values for *B. pseudomallei* due to almost universal resistance (>16 µg/mL) towards this antibiotic; however, there are notable exceptions [38, 41]. We chose a cut-off of 4 µg/mL to denote GEN-sensitive *B. pseudomallei*, which also reflects those strains that are unable to grow on Ashdown's agar, a selective medium developed for *B. pseudomallei* isolation [48].

## Bacterial isolate collection

We examined a 30-year collection of 1,040 primary (i.e. initial and predominantly pre-antibiotic treatment) clinical *B. pseudomallei* isolates for this study. Isolates were collected as part of the Darwin Prospective Melioidosis Study (DPMS) [53], which commenced in October 1989 and which has catalogued all known cases of melioidosis occurring in the tropical "Top End" of the Northern Territory, Australia, over the past ~30 years, with at least one *B. pseudomallei* isolate stored in ~95% of cases [50]. In each case, associated epidemiological and antibiotic treatment metadata were also collected.

## ARDaP algorithm

The ARDaP workflow is shown in Figure 3. ARDaP requires WGS data (clonal genomes or metagenomes in paired-end Illumina v1.8+ FASTQ format, or assembled genomes in FASTA format) and an annotated reference genome (FASTA) as input. For assembled genomes, ARDaP first converts the assemblies to synthetic Illumina v1.8+ reads using ART (version Mount Rainier 2016-06-05) [54]. For genomes in FASTQ format, ARDaP performs read quality filtering using Trimmomatic v0.39 [55] followed by random down-sampling of large genomes to user-defined coverage (default=50x) using Seqtk (https://github.com/lh3/seqtk) to facilitate more rapid analysis. ARDaP maps reads against an annotated reference using BWA-MEM v0.7.17-r1188 [56], followed by SAMTools v1.9 [57] for alignment processing and BAM creation, Genome Analysis Toolkit (GATK v4.1.0.0) [58] for local realignment around poorly mapped regions and for SNP and indel calling. Mosdepth (v0.2.3) [59] and Pindel (v0.2.5b9) [60] are used for coverage assessment and structural variant identification, respectively. Genome-wide variants (SNPs, small indels [<50bp], CNVs, gene gain, and gene loss or truncation) are identified with this comparative genomics approach. Variants are annotated with SnpEff [61]. ARDaP interrogates two databases: first, the CARD [3, 28] (v3.0.1) is screened to identify horizontally-acquired genes with removal of false-positive hits (i.e. genes present in all strains), and second, a species-specific database of known AMR determinants is compared against the annotated variants identified in the comparative genomic steps. ARDaP databases are created in SQLite and can be manually curated as additional AMR determinants are identified. Finally, ARDaP explores the annotated variant output to identify novel high-consequence mutations (i.e. those resulting in a frameshift or nonsense mutation) in known AMR genes that may cause AMR. These putative mutants can then be further explored with phenotypic AMR testing. The entire ARDaP pipeline has been implemented in Nextflow to ensure portability, scalability and compatability.

## Mixture detection

ARDaP incorporates extensive minor allelic variant analysis to enable variant identification (down to 5%) from mixed genomes or metagenomes, enabling the detection of emerging AMR determinants that have not yet become fixed in the bacterial population. Minor-variant SNPs and indels are identified using the ploidy-aware HaplotypeCaller tool in GATK v4.1, and deletions, CNVs and structural mutations are identified with the ploidy-aware function of Pindel. *B. pseudomallei* strains with known AMR status were mixed at ratios of 5% increments ranging from 5:95 to 95:5, respectively, to a total depth of 55-60x coverage. Two mixtures were created: MSHR6555 (sensitive to all standard antibiotics used to treat melioidosis) and MSHR5654 (MEM-, SXT- and CAZ-resistant), and MSHR6555 and MSHR8441 (MEM-, SXT-, CAZ- and DOX-resistant). MSHR5654 and MSHR8441 were chosen as they represent the spectrum of AMR towards the clinically relevant antibiotics used in melioidosis treatment.

## AMR database construction

ARDaP interrogates two databases to identify AMR determinants. The first is the public CARD database, and the second is a custom user-created database that lists all known chromosomally-encoded AMR determinants in the target organism. To develop the latter database, *B. pseudomallei* AMR determinants, including stepwise mutations, were identified from published literature, and associated AMR determinants were annotated relative to the archetypal K96243 *B. pseudomallei* genome [62]. The 50+ AMR determinants (as of version 1.4) identified from this search are summarised in an SQLite database (Table 2). Briefly, CAZ resistance is caused by altered substrate specificity of the *penA* $\beta$-lactamase [33–36, 63], *penA* upregulation [40, 63, 64] or duplication [30], or loss of penicillin-binding protein 3 [32]; AMC resistance is caused by *penA* upregulation [33]; MEM resistance is caused by regulator loss of the clinically relevant RND multidrug efflux pump systems in *B. pseudomallei* [12]; SXT resistance is caused by cumulative mutations occurring in core metabolism pathways coupled with loss of efflux pump regulation [12, 29, 30]; and DOX resistance is caused by loss-of-function mutations within

BPSL3085, a SAM-dependent methyltransferase, often but not always in combination with loss-of-function mutations affecting RND efflux pump regulators [31]. In addition to AMR determinants, the *B. pseudomallei* ARDaP database also includes AmrA or AmrB mutants that are associated with rare GEN susceptibility [38, 41]. To avoid poor-quality WGS data or incorrect species assignment affecting AMR outputs, the custom AMR database also includes two conserved genetic targets found only in *B. pseudomallei* [65, 66]. Their inclusion enables ARDaP to flag any strain lacking these internal control loci for further user assessment.

## References


1. On Antimicrobial Resistance, R. *Antimicrobial resistance: tackling a crisis for the health and wealth of nations* Report (2014).
2. Smith, R. D. & Coast, J. Antimicrobial resistance: a global response. *Bull World Health Organ* **80,** 126–33. ISSN: 0042-9686 (Print) 0042-9686 (2002).
3. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* **45,** D566–d573. ISSN: 0305-1048 (2017).
4. Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A. & Crook, D. W. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* **13,** 601–612 (2012).
5. Boolchandani, M., D'Souza, A. W. & Dantas, G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet.* ISSN: 1471-0056. doi:10.1038/s41576-019-0108-4. https://www.nature.com/articles/s41576-019-0108-4.pdf (2019).
6. Lakin, S. M. *et al.* MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res* **45,** D574–d580. ISSN: 0305-1048 (2017).
7. Li, B. *et al.* Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME Journal* **9,** 2490–2502 (2015).
8. Zankari, E. *et al.* Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother* **68,** 771–7. ISSN: 0305-7453 (2013).
9. Hiramatsu, K. *et al.* Genomic basis for methicillin resistance in Staphylococcus aureus. *Infect Chemother* **45,** 117–36. ISSN: 2093-2340 (Print) 1598-8112 (2013).
10. Martinez, J. L. General principles of antibiotic resistance in bacteria. *Drug Discov Today Technol* **11,** 33–39 (2014).
11. Fernandez, L. & Hancock, R. E. Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance. *Clin Microbiol Rev* **25,** 661–81. ISSN: 0893-8512 (2012).
12. Sarovich, D. S. *et al.* Raising the stakes: Loss of efflux-pump regulation decreases meropenem susceptibility in Burkholderia pseudomallei. *Clin Infect Dis.* ISSN: 1058-4838. doi:10.1093/cid/ciy069 (2018).
13. Lister, P. D. Chromosomally-encoded resistance mechanisms of Pseudomonas aeruginosa: therapeutic implications. *Am J Pharmacogenomics* **2,** 235–43. ISSN: 1175-2203 (Print) 1175-2203 (2002).
14. Tristram, S., Jacobs, M. R. & Appelbaum, P. C. Antimicrobial resistance in Haemophilus influenzae. *Clin Microbiol Rev* **20,** 368–389 (2007).
15. Richardot, C. *et al.* Carbapenem resistance in cystic fibrosis strains of Pseudomonas aeruginosa as a result of amino acid substitutions in porin OprD. *Int J Antimicrob Agents* **45,** 529–32. ISSN: 0924-8579 (2015).
16. Hunt, M. *et al.* ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genom* **3,** e000131. ISSN: 2057-5858 (2017).
17. Bradley, P. *et al.* Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. *Nat Commun* **6,** 10063. ISSN: 2041-1723 (2015).
18. Nicoloff, H., Hjort, K., Levin, B. R. & Andersson, D. I. The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification. *Nat Microbiol* **4,** 504–514. ISSN: 2058-5276 (2019).
19. Schweizer, H. P. Mechanisms of antibiotic resistance in Burkholderia pseudomallei: implications for treatment of melioidosis. *Future Microbiol* **7,** 1389–99. ISSN: 1746-0913 (2012).
20. Limmathurotsakul, D. *et al.* Predicted global distribution of Burkholderia pseudomallei and burden of melioidosis. *Nat Microbiol* **1,** 15008. ISSN: 2058-5276 (2016).
21. Wiersinga, W. J. *et al.* Melioidosis. *Nat Rev Dis Primers* **4.** doi:10.1038/nrdp.2017.107 (2018).
22. Currie, B. J. Melioidosis: evolving concepts in epidemiology, pathogenesis, and treatment. *Semin Respir Crit Care Med* **36,** 111–25. ISSN: 1098-9048 (Electronic) 1069-3424 (Linking) (2015).
23. Limmathurotsakul, D. & Peacock, S. J. Melioidosis: A clinical overview. *Br Med Bull* **99,** 125–139 (2011).
24. Bugrysheva, J. V. *et al.* Antibiotic resistance markers in Burkholderia pseudomallei strain Bp1651 identified by genome sequence analysis. *Antimicrob Agents Chemother* **61.** doi:10.1128/AAC.00010-17 (2017).
25. Khaledi, A. *et al.* Fighting antimicrobial resistance in Pseudomonas aeruginosa with machine learning-enabled molecular diagnostics. *bioRxiv,* 643676 (2019).
26. Gupta, S. K. *et al.* ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* **58,** 212–20. ISSN: 0066-4804 (2014).
27. Lim, Y. W. *et al.* Mechanistic model of Rothia mucilaginosa adaptation toward persistence in the CF lung, based on a genome reconstructed from metagenomic data. *PLoS ONE* **8.** doi:10.1371/journal.pone.0064285 (2013).
28. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57,** 3348–57. ISSN: 1098-6596 (Electronic) 0066-4804 (Linking) (2013).
29. Podnecky, N. L. *et al.* Mechanisms of resistance to folate pathway inhibitors in Burkholderia pseudomallei: deviation from the norm. *MBio* **8.** doi:10.1128/mBio.01357-17. https://mbio.asm.org/content/mbio/8/5/e01357-17.full.pdf (2017).
30. Viberg, L. T. *et al.* Within-host evolution of Burkholderia pseudomallei during chronic infection of seven Australasian cystic fibrosis patients. *MBio* **8.** doi:10.1128/mBio.00356-17. http://mbio.asm.org/content/8/2/e00356-17.full.pdf (2017).
31. Webb, J. R., Price, E. P., Currie, B. J. & Sarovich, D. S. Loss of methyltransferase function and increased efflux activity leads to doxycycline resistance in Burkholderia pseudomallei. *Antimicrob Agents Chemother* **61.** ISSN: 0066-4804. doi:10.1128/aac.00268-17. http://aac.asm.org/content/61/6/e00268-17.full.pdf (2017).
32. Chantratita, N. *et al.* Antimicrobial resistance to ceftazidime involving loss of penicillin-binding protein 3 in Burkholderia pseudomallei. *Proc Natl Acad Sci U S A* **108,** 17165–70. ISSN: 0027-8424 (2011).
33. Sarovich, D. S. *et al.* Characterization of ceftazidime resistance mechanisms in clinical isolates of Burkholderia pseudomallei from Australia. *PLoS One* **7,** e30789. ISSN: 1932-6203 (2012).
34. Sam, I. C., See, K. H. & Puthucheary, S. D. Variations in ceftazidime and amoxicillin-clavulanate susceptibilities within a clonal infection of Burkholderia pseudomallei. *J Clin Microbiol* **47,** 1556–8. ISSN: 0095-1137 (2009).
35. Rholl, D. A. *et al.* Molecular investigations of PenA-mediated ?-lactam resistance in Burkholderia pseudomallei. *Front Microbiol* **2,** 139–139. ISSN: 1664-302X (2011).
36. Tribuddharat, C., Moore, R. A., Baker, P. & Woods, D. E. Burkholderia pseudomallei class A ?-lactamase mutations that confer selective resistance against ceftazidime or clavulanic acid inhibition. *Antimicrob Agents Chemother* **47,** 2082–7. ISSN: 0066-4804 (Print) 0066-4804 (2003).
37. Price, E. P. *et al.* Within-Host Evolution of Burkholderia pseudomallei over a Twelve Year Chronic Carriage Infection. *mBio* **4,** e00388–13 (2013).
38. Podin, Y. *et al.* Burkholderia pseudomallei isolates from Sarawak, Malaysian Borneo, are predominantly susceptible to aminoglycosides and macrolides. *Antimicrob Agents Chemother* **58,** 162–6. ISSN: 0066-4804 (2014).
39. Trunck, L. A. *et al.* Molecular basis of rare aminoglycoside susceptibility and pathogenesis of Burkholderia pseudomallei clinical isolates from Thailand. *PLoS Negl Trop Dis* **3,** e519. ISSN: 1935-2735 (Electronic) 1935-2727 (Linking) (2009).
40. Hayden, H. S. *et al.* Evolution of Burkholderia pseudomallei in recurrent melioidosis. *PLoS One* **7,** e36507. ISSN: 1932-6203 (2012).
41. Price, E. P. *et al.* Within-host evolution of Burkholderia pseudomallei over a twelve-year chronic carriage infection. *MBio* **4,** e00388–13 (2013).
42. Rhodes, K. A. *et al.* Molecular determinants of Burkholderia pseudomallei BpeEF-OprC efflux pump expression. *Microbiology* **164,** 1156–1167. ISSN: 1350-0872 (2018).
43. Crisan, A., McKee, G., Munzner, T. & Gardy, J. L. Evidence-based design and evaluation of a whole genome sequencing clinical report for the reference microbiology laboratory. *PeerJ* **6,** e4218. ISSN: 2167-8359 (Print) 2167-8359 (2018).
44. Price, E. P. *et al.* Transcriptomic analysis of longitudinal Burkholderia pseudomallei infecting the cystic fibrosis lung. *Microb Genom* **4.** ISSN: 2057-5858 (Electronic) 2057-5858 (Linking). doi:10.1099/mgen.0.000194. https://www.ncbi.nlm.nih.gov/pubmed/29989529 (2018).
45. Godfrey, A. J., Wong, S., Dance, D. A., Chaowagul, W. & Bryan, L. E. Pseudomonas pseudomallei resistance to b-lactam antibiotics due to alterations in the chromosomally encoded b-lactamase. *Antimicrob Agents Chemother* **35,** 1635–40. ISSN: 0066-4804 (Print) 0066-4804 (Linking) (1991).
46. Simpson, A. J. *et al.* Comparison of imipenem and ceftazidime as therapy for severe melioidosis. *Clin Infect Dis* **29,** 381–7. ISSN: 1058-4838 (Print) 1058-4838 (Linking) (1999).
47. Wuthiekanun, V. *et al.* Survey of antimicrobial resistance in clinical Burkholderia pseudomallei isolates over two decades in Northeast Thailand. *Antimicrob Agents Chemother* **55,** 5388–91. ISSN: 1098-6596 (Electronic) 0066-4804 (Linking) (2011).
48. Ashdown, L. R. An improved screening technique for isolation of Pseudomonas pseudomallei from clinical specimens. *Pathology* **11,** 293–7. ISSN: 0031-3025 (Print) 0031-3025 (Linking) (1979).
49. 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,** 954–63. ISSN: 0099-2240 (2016).

50. Crowe, A., McMahon, N., Currie, B. J. & Baird, R. W. Current antimicrobial susceptibility of first-episode melioidosis Burkholderia pseudomallei isolates from the Northern Territory, Australia. *Int J Antimicrob Agents* **44,** 160–2. ISSN: 0924-8579 (2014).

51. Maloney, S., Engler, C. & Norton, R. Epidemiological cut-off value of clinical isolates of Burkholderia pseudomallei from northern Queensland to meropenem, ceftazidime, trimethoprim/sulfamethoxazole and doxycycline by the microbroth dilution method. *J Glob Antimicrob Resist* **10,** 291–294. ISSN: 2213-7165 (2017).

52. Kahlmeter, G. & Commitee, E. S. *2019. European Committee on Antimicrobial Susceptibility Testing proposed breakpoints for Burkholderia pseudomallei. [last accessed 24Jun19]* Report ().

53. 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. ISSN: 1935-2727 (2010).

54. Huang, W., Li, L., Myers, J. R. & Marth, G. T. ART: a next-generation sequencing read simulator. *Bioinformatics* **28,** 593–594. ISSN: 1367-4811 1367-4803 (2012).

55. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–20. ISSN: 1367-4803 (2014).

56. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997* (2013).

57. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–9. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking) (2009).

58. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20,** 1297–303. ISSN: 1549-5469 (Electronic) 1088-9051 (Linking) (2010).

59. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics* **34,** 867–868. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking) (2018).

60. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25,** 2865–71. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking) (2009).

61. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6,** 80–92. ISSN: 1933-6942 (Electronic) 1933-6934 (Linking) (2012).

62. Holden, M. T. *et al.* Genomic plasticity of the causative agent of melioidosis, Burkholderia pseudomallei. *Proc Natl Acad Sci U S A* **101,** 14240–5. ISSN: 0027-8424 (Print) 0027-8424 (2004).

63. Sarovich, D. S. *et al.* Development of ceftazidime resistance in an acute Burkholderia pseudomallei infection. *Infect Drug Resist* **5,** 129–132. ISSN: 1178-6973 (2012).

64. Chirakul, S. *et al.* Transcriptional and post-transcriptional regulation of PenA ?-lactamase in acquired Burkholderia pseudomallei ?-lactam resistance. *Sci Rep* **8,** 10652. ISSN: 2045-2322 (2018).

65. Price, E. P. *et al.* Development and validation of Burkholderia pseudomallei-specific real-time PCR assays for clinical, environmental or forensic detection applications. *PLoS One* **7,** e37723. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking) (2012).

66. 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. ISSN: 0095-1137 (Print) 0095-1137 (Linking) (2006).

## Funding

This study was funded by the National Health and Medical Research Council (awards 1046812, 1098337, and 1131932 from the HOT NORTH initiative). DEM was supported by an Australian Government Research Training Scholarship. ES was supported by an International Postgraduate Research Scholarship from James Cook University. EPP and DSS were supported by Advance Queensland fellowships (awards AQIRF0362018 and AQRF13016-17RD2, respectively).

**ACKNOWLEDGMENTS.** We thank Associate Professor Rob Baird and the microbiology staff at Royal Darwin Hospital for their support and expertise in identifying and characterising *B. pseudomallei* isolates.