EBioMedicine 63 (2021) 103152

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Research paper

Taking the next-gen step: Comprehensive antimicrobial resistance detection from *Burkholderia pseudomallei*

Danielle E. Madden, BSc (Hons)[1,2], Jessica R. Webb, PhD[3], Eike J. Steinig, BSc (Hons)[4], Bart J. Currie, FRACP[3,5], Erin P. Price, PhD[1,2,3,^], Derek S. Sarovich, PhD[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, Charles Darwin University, 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*

ARTICLE INFO

Article History:
Received 8 September 2020
Revised 18 October 2020
Accepted 16 November 2020
Available online xxx

Keywords:
ARDaP
Antimicrobial resistance
Comparative genomics
Next-generation sequencing
Melioidosis, Database

ABSTRACT

Background: Antimicrobial resistance (AMR) poses a major threat to human health. Whole-genome sequencing holds great potential for AMR identification; however, there remain major gaps in accurately and comprehensively detecting AMR across the spectrum of AMR-conferring determinants and pathogens.
Methods: Using 16 wild-type *Burkholderia pseudomallei* and 25 with acquired AMR, we first assessed the performance of existing AMR software (ARIBA, CARD, ResFinder, and AMRFinderPlus) for detecting clinically relevant AMR in this pathogen. *B. pseudomallei* was chosen due to limited treatment options, high fatality rate, and AMR caused exclusively by chromosomal mutation (i.e. single-nucleotide polymorphisms [SNPs], insertions-deletions [indels], copy-number variations [CNVs], inversions, and functional gene loss). Due to poor performance with existing tools, we developed ARDaP (Antimicrobial Resistance Detection and Prediction) to identify the spectrum of AMR-conferring determinants in *B. pseudomallei*.
Findings: CARD, ResFinder, and AMRFinderPlus failed to identify any clinically-relevant AMR in *B. pseudomallei*; ARIBA identified AMR encoded by SNPs and indels that were manually added to its database. However, none of these tools identified CNV, inversion, or gene loss determinants, and ARIBA could not differentiate AMR determinants from natural genetic variation. In contrast, ARDaP accurately detected all SNP, indel, CNV, inversion, and gene loss AMR determinants described in *B. pseudomallei* ($n \approx 50$). Additionally, ARDaP accurately predicted three previously undescribed determinants. In mixed strain data, ARDaP identified AMR to as low as ~5% allelic frequency.
Interpretation: Existing AMR software packages are inadequate for chromosomal AMR detection due to an inability to detect resistance conferred by CNVs, inversions, and functional gene loss. ARDaP overcomes these major shortcomings. Further, ARDaP enables AMR prediction from mixed sequence data down to 5% allelic frequency, and can differentiate natural genetic variation from AMR determinants. ARDaP databases can be constructed for any microbial species of interest for comprehensive AMR detection.
Funding: National Health and Medical Research Council (BJC, EPP, DSS); Australian Government (DEM, ES); Advance Queensland (EPP, DSS).

1. Introduction

Antimicrobial resistance (AMR) poses a major threat to human health worldwide and is an increasing contributor to morbidity and mortality. Antibiotic use and misuse have resulted in an alarming increase in multidrug-resistant infections worldwide, provoking an urgent need to improve global AMR detection and surveillance. Alongside pathogen identification, AMR detection is one of the primary goals of diagnostic microbiology, with far-reaching consequences for both infection control and effective treatment [1].

Whole-genome sequencing (WGS) permits comprehensive AMR detection and prediction from bacterial genomes by identifying all AMR determinants in a single genome or metagenome [2], circumventing the need for multiple and often laborious diagnostic methods. Existing bioinformatic tools such as ARG-ANNOT [3], Antibiotic Resistance Identification By Assembly (ARIBA) [4], Comprehensive

E-mail address: dsarovich@usc.edu.au (D.S. Sarovich).
^ Authors contributed equally
# Corresponding author.

https://doi.org/10.1016/j.ebiom.2020.103152

Research in context

Evidence before this study

If unchecked, antimicrobial resistance (AMR) is predicted to have a devastating impact on global health in the coming decades. Next-generation sequencing (NGS) is an essential tool for combatting AMR, providing a comprehensive and accurate diagnostic tool for AMR detection and unveiling the molecular basis underpinning the evolution of AMR in many dangerous multidrug-resistant pathogens. Whilst currently available AMR software readily detects horizontally-acquired AMR genes and some chromosomally-encoded variants, no existing tool can detect AMR determinants caused by the spectrum of chromosomal mutations, leading to considerable underreporting of AMR in many microbes.

Added value of this study

To overcome current software limitations, we were prompted to develop ARDaP. Using NGS or genome assembly data as input, ARDaP can detect and predict AMR caused by gene acquisition, point mutations, insertions-deletions, gene copy-number variation, inversions, and gene loss or truncation. We tailored ARDaP for AMR determinant detection in the formidable melioidosis pathogen, *Burkholderia pseudomallei*, which has limited treatment options due to intrinsic multidrug resistance and poor or no AMR detection support with existing AMR software. ARDaP also incorporates a mixture-aware feature that enables the detection of emerging AMR determinants, thereby informing early treatment shifts and improving antibiotic stewardship efforts and patient survival. Although we demonstrate its application in *B. pseudomallei*, ARDaP databases can be developed to identify AMR in any microbe of interest.

Implications of all the available evidence

Using ARDaP, both known and novel AMR determinants can be accurately identified from NGS data, and non-AMR-conferring variants can be ignored, representing important advances over existing AMR detection software. Inclusion of antimicrobial-susceptible strains, an important yet often-overlooked component of AMR database development and validation, is critical for accommodating natural genetic variation and mitigating high false-positive rates. Functional verification of novel AMR determinants (e.g. phenotypic testing, gene knockouts, heterologous expression, or RNA sequencing), remains a limiting factor in our understanding of AMR. Our study highlights the essential need for well-curated and meticulous pathogen-specific databases for the most accurate, comprehensive, and clinically relevant AMR detection. Ongoing efforts are needed to continue uncovering the myriad ways that microorganisms evolve to evade antimicrobial agents.

Antibiotic Resistance Database (CARD) [5], ResFinder [6], AMRFinder [7], and MEGARes [2] can readily detect AMR genes acquired from horizontal gene transfer events. Many bacterial pathogens also develop AMR via chromosomal mutations, including missense single-nucleotide polymorphism (SNP) mutations in β-lactamase-encoding genes, SNPs or insertion-deletions (indels) in efflux pump regulators [8–10], gene amplification via copy-number variations (CNVs) [11], inversions [9], and functional gene loss [8]. Recent improvements in AMR identification software mean that chromosomal mutations, particularly SNPs, are now identifiable. For example, ARIBA can identify AMR-conferring SNPs and indels in multiple species [4]. Nevertheless, other types of genetic variants – gene loss or truncation, inversions, and CNVs – remain poorly identified using existing tools, despite their crucial role in conferring AMR [12].

The Tier 1 Select Agent bacterium, *Burkholderia pseudomallei*, causes the often-fatal tropical disease melioidosis. Melioidosis severity ranges from mild, self-limiting skin abscesses to pneumonia, neurological disease, and septic shock. *B. pseudomallei* is naturally resistant to many antibiotics, including aminoglycosides, penicillins, macrolides, and polymyxins [13,14]. Fortunately, human-to-human *B. pseudomallei* transmission is rare; almost all infections are acquired from the environment. As such, isolates collected prior to antibiotic treatment are almost universally susceptible to the following clinically-relevant antibiotics: ceftazidime (CAZ), amoxicillin-clavulanate (AMC), co-trimoxazole (SXT), doxycycline (DOX), meropenem (MEM) and imipenem (IPM) [15]. To prevent melioidosis relapse, treatment involves prolonged (3-6 month) antibiotic therapy, which increases AMR risk and treatment failure [8]. AMR in *B. pseudomallei* has been reported for all clinically-relevant antibiotics [8], with novel AMR determinants towards these key antibiotics continuing to be uncovered.

Here, we tested 47 characterised *B. pseudomallei* genomes with known antibiotic phenotype profiles and associated AMR determinants, and three MEM-resistant (MEMr) strains with previously unidentified AMR determinants, against existing tools (ARIBA, CARD and AMRFinderPlus) to determine their AMR detection efficacy. Among the characterised strains, 25 were phenotypically-confirmed as resistant towards at least one clinically relevant antibiotic, 16 were sensitive, and the remainder encoded unusual sensitivity towards aminoglycosides and macrolides, or stepwise AMR variants. Following testing against the current AMR tools, we developed a new tool, Antibiotic Resistance Detection and Prediction (ARDaP), to permit comprehensive AMR detection from microbial genomes. ARDaP was designed to meet four main aims: first, to accurately identify AMR determinants caused by a spectrum of mutational mechanisms (i.e. gene gain, SNPs, indels, CNVs, inversions, and functional gene loss); second, to predict enigmatic AMR determinants in isolates with phenotypically-confirmed AMR, third, to detect minor AMR allelic determinants in mixed (e.g. metagenomic) sequence data; and finally, to provide a user-friendly report that summarises the AMR determinants (if any) and associated AMR phenotypes, stepwise variants, unusual antimicrobial sensitivity determinants, and genetic variants associated with natural variation that do not confer AMR. Although we illustrate its utility in *B. pseudomallei*, ARDaP is amenable to AMR identification across all microbial species.

2. Methods

Ethics. Ethics approval was obtained from the Human Research Ethics Committee of the Northern Territory Department of Health and Menzies School of Health Research (HREC 02/38, "Clinical and Epidemiological Features of Melioidosis"). Written informed consent was provided by study participants.

Isolates. Forty-seven *B. pseudomallei* strains were included in this study, including 25 with elevated MICs towards one or more clinically-relevant antibiotics (Table 1) and genotypically-confirmed AMR determinants. These isolates were selected as they represent the spectrum of known AMR determinants in *B. pseudomallei* (Table S1). Strains encoding unusual aminoglycoside- and macrolide-sensitivity, and stepwise mutations that lower the barrier to AMR development, were also examined (Table 1). A further 16 strains sensitive to all clinically-relevant antibiotics were included to test software efficacy (Table 2). Finally, three previously uncharacterised clinical strains exhibiting MEMr (MSHR1058 MIC=12 μg/mL; MSHR1174 MIC=6 μg/mL; MSHR8777 MIC=4 μg/mL; Table 3) were included to test the predictive capacity of ARDaP.

Table 1
Antimicrobial resistance (AMR) determinants in 25 *Burkholderia pseudomallei* strains with verified AMR phenotypes, plus strains conferring unusual antimicrobial susceptibility and stepwise AMR variants[§].

Patient ID	Isolate	ST	Genome accession	Antibiotic MIC (μg/mL)[¥]	Stepwise variant[†]	AMR determinant	Reference/s
Thai patient	316c	17	SRR2975745	CAZr (64)	—	$PenA_{P173S}$	[22,23]
Thai patient	354e	78	AHJD00000000.1	CAZi (6)	—	*penA* -78G>A	[9]
				SXTi (3)	BpeT structural variant[‡]	$Ptr1_{R21fs}$	
Australian patient	Bp1651	880	SRR2102060	CAZr (≥128)	*penA* -78G>A	$PenA_{D245G}$	[38]
				Unusual sensitivity to aminoglycosides (GEN, KAN) and macrolides (AZM)	—	$AmrB_{A254fs}$	
				DOXr (16)	—	$BPSL3085_{A88fs}$	
				AMCr (64/32)	*penA* -78G>A	$PenA_{S78F}$	
Pre-DPMS 89	MSHR0052	722	SRR5818275	MEMr (8)	—	$AmrR_{E190*}$	[8]
				DOXr (48)	$AmrR_{E190*}$	$BPSL3085_{V211M}$	
Australian patient	MSHR0292	236	SRR4254580	DOXr (16)	$AmrR_{S174P}$	$BPSL3085_{V40A}$	[29]
P215	MSHR0663	36	SRR2887062	SXTr (≥32)	$BpeT_{H278Y}$	$Ptr1_{R21fs}$	[8]; This study
	MSHR0937	36	SRR2886988	AMCi (12/6), MEMr (6)	—	$BpeR_{D176A}$	
				SXTi (3)	$BpeR_{D176A}$	$MetF_{Q142*}$	
P179	MSHR0678	114	SRR6075118	MEMr (3)	—	$AmrR_{E21D}$	[8]
	MSHR0800	114	SRR6075115	MEMr (6), DOXi (8)	—	$BpeR_{L85fs}$	
P337	MSHR1226	333	SRR9598635	CAZr (≥256)	—	$PenA_{C75Y}$	[22]
	MSHR1300	333	SRR6075114	CAZr (≥256)	*penA* -78G>A	$PenA_{C75Y}$	[22]
				MEMr (4)	—	$AmrR_{K13fs}$	[8]
				SXTi (3)	—	$AmrR_{K13fs}$	This study
P595	MSHR3683	144	SRR11678542	DOXi (12)	—	$BPSL3085_{A88fs}$	This study
P608	MSHR4083	36	SRR2887030	SXTr (24)	$AmrR_{\Delta A153-D156}$	Dut_{G91A}	[8]
				MEMr (6)	—	$AmrR_{\Delta A153-D156}$	
CF6	MSHR5654	1040	SRR3404570	CAZr (≥256)	PenA 30x CNV	$PenA_{C75Y}$	[11]
				CIPr (≥32)	—	$GyrA_{Y77S}$	
				SXTr (≥32)	$BpeT_{T314fs}$	$Ptr1_{R21fs}$	
CF9	MSHR5665	252	SRR3404582	SXTr (≥32)	$Ptr1_{W116R}$; $MetF_{Q28P}$	Dut_{V77A}	[11,25]
				DOXi (6)	—	$BPSL3085_{A88fs}$	
	MSHR5666	252	SRR3404597	DOXs (1)[#]	—	$BPSL3085_{A88fs}$	
				SXTr (≥32)	$Ptr1_{W116R}$; $MetF_{Q28P}$	Dut_{V77A}	
	MSHR5667	252	SRR3404598	SXTr (≥32)	$MetF_{N162T;}$ $AmrR_{L132P}$	Dut_{N99S}	
				MEMr (4)	—	$AmrR_{L132P}$	
				DOXr (48)	$AmrR_{L132P}$	$BPSL3085_{A88fs}$	
	MSHR5669	252	SRR3404599	DOXi (6)	—	$BPSL3085_{A88fs}$	
				SXTr (≥32)	$Ptr1_{W116R}$; $MetF_{Q28P}$	Dut_{V77A}	
P726	MSHR6755	975	SRR6075122	MEMr (3)	—	$AmrR_{\Delta V60-C63}$	[8]
P797	MSHR7587	437	SRR6075129	MEMr (4)	—	$AmrR_{G30D}$	[8]
	MSHR7929	437	SRR6075126	SXTr (4)	$AmrR_{G30D}$	$BPSL2263_{\Delta D110-G116}$	
				MEMr (4)	—	$AmrR_{G30D}$	
CF11	MSHR8441	46	SRR3382162	CAZi (12)	—	PenA 10x CNV	[8,11]
				SXTr (≥32)	$AmrR_{\Delta V62-H223}$[c]	$Ptr1_{A22-G23ins\_R-R-A}$	
				Decreased DOX susceptibility (4)	$AmrR_{\Delta V62-H223}$[c]	$BPSL3085_{S130L}$	
				Decreased MEM susceptibility (2)	—	$AmrR_{\Delta V62-H223}$[c]	
	MSHR8442	46	SRR3404603	SXTr (≥32)	$AmrR_{\Delta V62-H223}$[c]	$Ptr1_{A22-G23ins\_R-R-A}$	
				DOXi (8)	$AmrR_{\Delta V62-H223}$[c]	$BPSL3085_{S130L}$	
				MEMr (3)	—	$AmrR_{\Delta V62-H223}$[c]	
Non-DPMS QP09	MSHR8481	1378	SRR6075123	MEMr (6)	—	$AmrR_{\Delta A70-H223}$[d]	[8]
P989	MSHR9021[a]	132	SRR6075127	MEMr (3)		$AmrR_{S166P}$; $AmrR_{A145fs}$[b]	[8]
Strains with unusual antibiotic sensitivity							
P314	MSHR1043[§]	131	SRR6380769	Unusual GEN sensitivity	—	$AmrA_{L247fs}$	[19]
Malaysian patient	MSHR5089[§]	881	SRR2975737	Unusual sensitivity to aminoglycosides (GEN, KAN) and macrolides (AZM)	—	$AmrB_{T368R}$	[18]
Strains with stepwise variants							
P179	MSHR0535	114	SRR6075120	DOXs (4)[§]	$BPSL3085_{R104fs}$	—	[8]
Thai patient	354e	78	AHJD00000000.1	AMCs (4/2)[§]	*penA* -78G>A	—	[9]

Abbreviations: AZM, azithromycin; CAZ, ceftazidime; CIP, ciprofloxacin; CNV, copy-number variation; DOX, doxycycline; GEN, gentamicin; KAN, kanamycin; MEM, meropenem; MIC, minimum inhibitory concentration; SXT, co-trimoxazole

[¥] i, intermediate; r, resistant; s, sensitive

[†] Where applicable

[‡] Encodes an 800kb inversion that affects the 3' end of *bpeT*, the transcriptional regulator of the BpeEF-OprC efflux pump[9]

[#] False positive; isolate remained sensitive despite encoding a known AMR determinant. Cause of reversion unknown

[§] Strains MSHR1043[19] and MSHR5089[18] encode unusual antimicrobial susceptibility, and are susceptible to all five clinically-relevant antibiotics (i.e. AMC, CAZ, DOX, MEM, SXT), and strains MSHR0535 and 354e encode stepwise variants that elevate the DOX and AMC MICs, respectively, but do not exceed the established resistance cut-off for these antibiotics

[a] MSHR9021 was intentionally sequenced from a potentially mixed population to capture population diversity. Two AmrR mutants, $AmrR_{S166P}$ and $AmrR_{A145fs}$, were present at ratios of ~66% and 33%, respectively[8]. In the non-mixture mode, only the dominant variant, $AmrR_{S166P}$, is detected.

[b] Frameshift indel shortens protein length from 223 to 183 residues

[c] Frameshift indel increases protein length from 223 to 285 residues

[d] Frameshift indel shortens protein length from 223 to 117 residues

Table 2
Burkholderia pseudomallei strains phenotypically confirmed to be sensitive towards the five clinically-relevant antibiotics, and associated genome data.

Patient ID	Isolate	ST	Genome accession	Antibiotic MIC (μg/mL)*					Reference/s
				AMC	CAZ	DOX	MEM	SXT	
Australian patient	MSHR0293	236	SRR4254579	2/1	1	1	0.5	0.4	[29]
P179	MSHR0492	114	SRR6075119	1.5/0.75	1.5	1	1.5	1	[8]
	MSHR0934	114	SRR6075116	2/1	2	1	0.75	1	
P337	MSHR1141	333	SRR2975732	1.5/0.75	1.5	1	0.75	0.75	[22]
P608	MSHR3763	36	SRR2887021	4/2	2	0.75	0.75	3	[8]
P726	MSHR5864	975	SRR6075121	3/1.5	1.5	1	0.75	1.5	[8]
P797	MSHR6522	437	SRR6075128	2/1	1.5	1	0.5	1.5	[8]
CF1	MSHR0913	279	SRR3404575	2/1	3	1	1	0.5	[11]
	MSHR1053	279	SRR3404578	1.5/0.75	2	1	0.5	1	
Malaysian patient	MSHR5093	881	SRR2975738	6/3	4	1.5	1	3	[18]
Malaysian patient	MSHR5104	881	SRR2975740	2/1	0.75	1	0.19	0.75	[18]
CF6	MSHR5651	1040	SRR3381886	1.5/0.75	1.5	0.38	0.5	0.75	[11,25]
CF9	MSHR5662	252	SRR3381885	2	2	2	0.75	0.5	[11,25]
	MSHR5670[#]	252	SRR3404600	1.5/0.75	2	2	0.75	0.5	
CF10	MSHR8438	442	SRR3382015	2/1	3	1	1	0.25	[11,25]
	MSHR8440	442	SRR3404601	2/1	3	1	0.75	0.38	

Abbreviations: AMC, amoxicillin-clavulanate; CAZ, ceftazidime; DOX, doxycycline; MEM, meropenem; ST, multilocus sequence type; SXT, co-trimoxazole
*According to Etesting
[#] False positive; isolate remained sensitive despite encoding known AMR determinants towards DOX ($BPSL3085_{A88fs}$) and SXT ($Ptr1_{W116R}$; $MetF_{Q28P}$; Dut_{V77A}). Cause of reversion unknown

Table 3
ARDaP prediction in three meropenem-resistant *Burkholderia pseudomallei* isolates with previously unknown antimicrobial resistance (AMR) determinants.

Strain ID	ARDaP-predicted AMR determinant	MEM MIC (μg/mL)	Reference/s
MSHR1058	$AmrR_{\Delta P81\text{-}H223}$[a]	12 (MEM)	[8]; This study
MSHR1174	$AmrR_{G149fs}$[b]	6 (MEM)	This study
MSHR8777	$AmrR_{\Delta A128\text{-}H223}$[c]	4 (MEM)	This study

[a] Previously undescribed 3' *amrR* deletion shortens protein length from 223 to 130 residues.
[b] Previously undescribed 11bp insertion shortens protein length from 223 to 178 residues.
[c] Previously undescribed 3' *amrR* insertion increases protein length from 223 to 621 residues.

Culturing, WGS, and genome assembly. *B. pseudomallei* culture, DNA extraction, and WGS were performed as described elsewhere [11]. Genomic data for MSHR1058, MSHR1174, and MSHR8777 were uploaded to the NCBI Sequence Read Archive database (BioProject PRJNA641249). Accession numbers for all other genomic data are listed in Tables 1 and 2. For genomes lacking a publicly-available assembly, MGAP v1.1 (https://github.com/dsarov/MGAP—Microbial-Genome-Assembler-Pipeline) was used, with archetypal strain K96243 (RefSeq accessions NC_006350.1 and NC_006351.1) provided as the scaffolding reference.

Minimum Inhibitory Concentrations (MICs). MICs were determined using Etests (bioMérieux, Murarrie, Australia). Sensitive, intermediate and resistant cut-offs were based on the Clinical and Laboratory Standards Institute (CLSI) M100-S17 guidelines for *B. pseudomallei* ($\leq$8/4, 16/8, and $\geq$32/16 μg/mL for AMC; $\leq$8, 16, $\geq$32 μg/mL for CAZ; $\leq$4, 8, $\geq$16 μg/mL for DOX and IPM; $\leq$2/38, nil, $\geq$4/76 μg/mL for SXT). CLSI guidelines do not list MEM for *B. pseudomallei*; however, based on prior work [16,17], and recent proposed EUCAST breakpoints for *B. pseudomallei*, we categorised MEMr as $\geq$3 μg/mL. Likewise, the CLSI guidelines do not list gentamicin (GEN) MIC values for *B. pseudomallei* due to almost ubiquitous resistance ($>$16 μg/mL) towards this antibiotic; however, there are notable exceptions [18,19]. We chose a GEN-sensitive cut-off of $\leq$4 μg/mL, which also reflects those strains unable to grow on Ashdown's agar, a selective medium for *B. pseudomallei* isolation that contains 4 μg/mL GEN. For the four false-positive strains, Etests were performed on a minimum of two occasions by different operators to ensure MIC robustness.

AMR software parameters. The default RGI v5.1.0 database parameters of CARD v3.0.9 (https://card.mcmaster.ca/analyze/rgi; accessed 25Jun20), ARIBA v2.14.5 (https://github.com/sanger-pathogens/ariba), ResFinder v4.1 (https://cge.cbs.dtu.dk/services/ResFinder/), and AMRFinderPlus v3.8.28 (https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/) were examined for performance across the *B. pseudomallei* genomes.

ARDaP AMR database construction. ARDaP is available at: https://github.com/dsarov/ARDaP. All reported *B. pseudomallei* AMR determinants, including stepwise AMR mutations and unusual antimicrobial susceptibility mutations (Table 1; Table S1), were annotated relative to K96243. The AMR determinants (as of version 1.7) are summarised in an SQLite database (Table S1; most up-to-date version available at: https://github.com/dsarov/ARDaP/tree/master/Databases/Burkholderia_pseudomallei_k96243). Briefly, CAZ resistance (CAZr) is caused by altered PenA β-lactamase substrate specificity [20–23], *penA* upregulation [9,22,24,25] (including CNVs [11]), or loss of penicillin-binding protein 3 [26]; AMC resistance (AMCr) is caused by *penA* upregulation [19,22]; MEMr is caused by AmrAB-OprA, BpeAB-OprB, or BpeEF-OprC resistance-nodulation-division (RND) multidrug efflux pump regulator loss-of-function [8]; SXT intermediate (SXTi) or full resistance (SXTr) is caused by cumulative mutations in core metabolism pathways coupled with AmrAB-OprA, BpeAB-OprB, or BpeEF-OprC RND efflux pump regulator loss-of-function [8,11,27,28]; and DOX intermediate (DOXi) or full resistance (DOXr) is caused by loss-of-function mutations within the SAM-dependent methyltransferase gene, *BPSL3085*, often in combination with AmrAB-OprA, BpeAB-OprB, or BpeEF-OprC regulator loss-of-function [29]. Our *B. pseudomallei* ARDaP database also includes AmrA and AmrB mutants that are associated with unusual aminoglycoside and macrolide susceptibility [18,19]. To avoid poor-quality WGS data or incorrect species assignments, the database also includes two conserved genetic targets (Table S1) found only in this bacterium; strains lacking these loci are flagged for further user assessment.

ARDaP algorithm. To achieve high-quality variant calls, ARDaP incorporates several tools into its workflow (full list available at: https://github.com/dsarov/ARDaP). In addition to an organism-specific SQLite database (https://github.com/dsarov/ARDaP/tree/master/

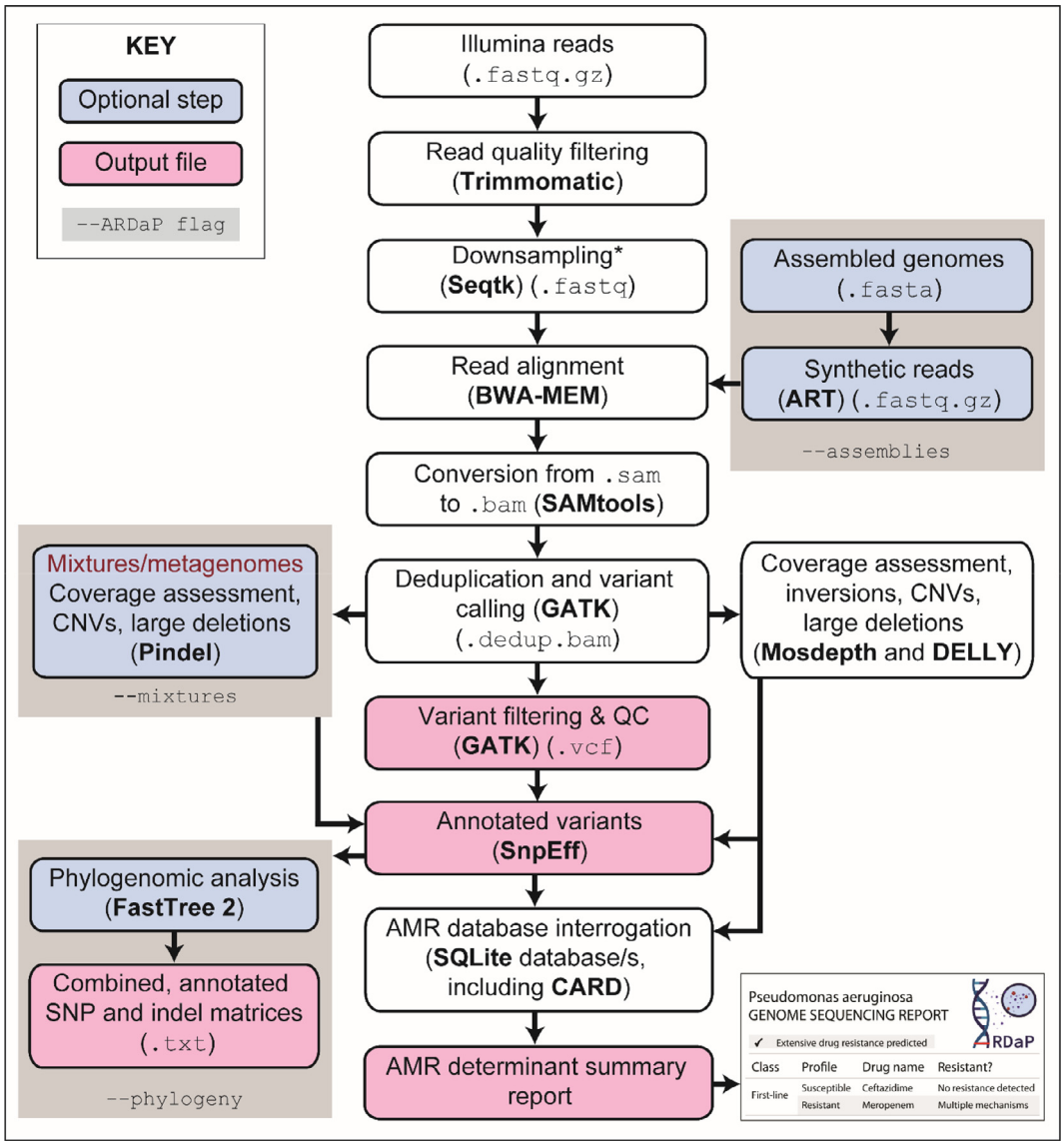

Fig. 1. ARDaP pipeline. The user inputs assembled genome/s or raw sequencing reads. ARDaP then performs read alignment, read processing, deduplication, and variant identification. An optional phylogenetic analysis is also performed (if specified). Coverage assessment is undertaken on either single or mixed genomes (if specified); genetic variants are then annotated and antimicrobial resistance database/s interrogated. Finally, ARDaP produces a summary report of antimicrobial resistance determinants for each strain (Figure 2). *Downsampling is carried out by default but can be turned off using the –size 0 flag in ARDaP.

Databases), ARDaP requires WGS data, either genomes or metagenomes in paired-end Illumina v1.8+ FASTQ format, or assembled genomes in FASTA format, as input (Fig. 1). For genomes in FASTA format, ARDaP first converts to synthetic Illumina v1.8+ reads using ART (version Mount Rainier 2016-06-05).[30] For genomes in FASTQ format, ARDaP performs quality filtering using Trimmomatic v0.39 followed by optional random down-sampling to a user-defined coverage (default=50x) using Seqtk (https://github.com/lh3/seqtk) to permit more rapid analysis. ARDaP then performs comparative genomic analysis to identify AMR determinants by mapping reads against an annotated reference using BWA-MEM (v0.7.17-r1188),[31] followed by SAMTools (v1.9)[32] for alignment processing and BAM creation, Genome Analysis Toolkit (GATK v4.1.0.0)[33] for SNP and indel identification, Mosdepth (v0.2.3)[34] for coverage assessment, Pindel (v0.2.5b9)[35] for CNV detection, and DELLY (v0.8.3)[36] for inversion identification. High-quality genetic variants (SNPs, indels [<50bp], CNVs, gene gain, inversions, and gene loss or truncation) are then annotated with SnpEff (v4.3.1t).[37] ARDaP next interrogates two databases: i) a customisable CARD[5] database is screened to identify horizontally-acquired AMR genes and to ignore conserved genes that do not confer AMR, and ii) a bespoke AMR determinant database (in this study, a *B. pseudomallei* database) containing species-specific AMR determinants. ARDaP databases are created in SQLite and can be readily updated as additional AMR determinants are identified. This database also accommodates stepwise mutations and AMR conferred by ≥2 mutations. Finally, ARDaP can predict AMR by identifying novel high-consequence mutations (i.e. those resulting in a frameshift or nonsense mutations, loss of coverage, or inversion) in known AMR genes. These putative mutants can then be flagged for further investigation with phenotypic AMR testing. ARDaP outputs are presented in a comprehensive, human-readable report (Fig. 3).

Mixture detection. ARDaP incorporates a minor allelic variant analysis function to permit variant identification from mixed genomes/metagenomes, enabling the detection of emerging AMR determinants (down to 5% abundance). Minor-variant SNPs and indels are identified using the ploidy-aware HaplotypeCaller tool in GATK v4.1; deletions and CNVs are identified with the ploidy-aware function of Pindel. *B. pseudomallei* strains with known AMR status

Table 4
ARDaP detection limits for AMR determinants in synthetic mixtures.

AMR determinant	K96243 gene ID	Mutation type	AMR minor allele detection (%)
Mixture #1: MSHR0913 (sensitive strain) and MSHR5654 (resistant to CAZ, CIP, and SXT)			
$BpeT_{T314fs}$	*BPSS0290*	Gene loss/truncation	10
$Ptr1_{R21fs}$	*BPSS0039*	Gene loss/truncation	10
penA 30x	*BPSS0946*	Copy-number variation (30x)	5
$PenA_{C75Y}$	*BPSS0946*	Missense SNP	5
$GyrA_{Y77S}$	*BPSL2521*	Missense SNP	10
Mixture #2: MSHR0913 (sensitive strain) and MSHR8441 (resistant to SXT; intermediate resistance to CAZ; decreased susceptibility to MEM and DOX)			
$AmrR_{\Delta V62\text{-}H223}$	*BPSL1805*	Gene loss/truncation	50[†]
$BPSL3085_{S130L}$	*BPSL3085*	Missense SNP	15
$Ptr1_{A22\text{-}G23ins\_R\text{-}R\text{-}A}$	*BPSS0039*	In-frame insertion	10
penA 10x	*BPSS0946*	Copy-number variation (10x duplication)	5

AMR, antimicrobial resistance; CAZ, ceftazidime; CIP, ciprofloxacin; DOX, doxycycline; MEM, meropenem; ND, not detected; SNP, single-nucleotide polymorphism; SXT, co-trimoxazole
[†] Represents the lowest allele frequency that was consistently detected in our mixed strain dataset.

were mixed at ratios of 5% increments ranging from 5:95 to 95:5, to 55-60x total depth. Two mixtures were created: MSHR0913 (sensitive to all clinically-relevant antibiotics) and MSHR5654 (SXTr, CAZr, ciprofloxacin-resistant), and MSHR0913 and MSHR8441 (SXTr; intermediate resistance to CAZ; decreased susceptibility to MEM and DOX). MSHR5654 and MSHR8441 were chosen as they represent a wide spectrum of clinically-relevant AMR and mutation types (Table 4).

3. Role of funders

This study was funded by the National Health and Medical Research Council (awards 1046812, 1098337, and 1131932 [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). The funders had no role in study design; in the collection, analysis, or interpretation of data; in the writing of this report; or in the decision to submit for publication. The corresponding author had full access to all the data in the study and has final responsibility for the decision to submit for publication.

4. Results

Performance of existing AMR tools in *B. pseudomallei*. The validated dataset of 47 *B. pseudomallei* isolates was used to assess the performance and capacity of existing AMR tools to identify AMR determinants in AMR but not antimicrobial-sensitive strains. According to CARD and AMRFinderPlus, all 47 genomes were found to harbour AMR determinants; however, these determinants corresponded with conserved genes in all *B. pseudomallei* (Table S2). In addition, CARD, ResFinder, and AMRFinderPlus failed to identify any clinically relevant AMR determinants in the 25 AMR strains. ARIBA outperformed CARD and AMRFinderPlus due to its ability to include missense – although not nonsense – SNPs in its database construction, and to identify SNPs and indels in its report outputs. However, ARIBA cannot identify CNVs or inversions, and it requires considerable user expertise and assessment time to determine the validity of variant outputs and to distinguish real AMR determinants from natural variation. Due to these limitations, we did not pursue this tool further.

ARDaP development and performance in *B. pseudomallei*. Given the shortcomings of existing AMR software, ARDaP was designed to both identify known AMR determinants and to ignore non-causal genetic variants (Table 2). When tested against the 47 validated isolates, ARDaP correctly identified all *B. pseudomallei* AMR determinants (Table 1) and yielded no false negatives; however, four false positives (MSHR5654, MSHR5666, MSHR5669, and MSHR5670), all of which were isolated from chronic cystic fibrosis (CF) infections, were identified. The first of these, MSHR5654 (from CF6) [11], was predicted to be MEMr due to the presence of $BpeT_{Thr314fs}$. The remaining false-positive strains (all from CF9) [11], encode a SAM-dependent methyltransferase truncation ($BPSL3085_{A88fs}$) [25] and were predicted by ARDaP to be DOXr. We also observed $BPSL3085_{A88fs}$ in an unrelated DOXr chronic CF strain, Bp1651 [38] (Table 1). BPSL3085 mutations confer DOXr likely by altering ribosomal methylation patterns [11,29]. However, all three strains remained DOX-sensitive (1.5 μg/mL) despite other CF9 strains encoding $BPSL3085_{A88fs}$ and being DOXr (MSHR5665: MIC=6 μg/mL; MSHR5667: MIC=48 μg/mL; Table 1) [11]. The much higher DOX MIC in MSHR5667 is attributable to a second mutation ($AmrR_{L132P}$; Table 1).

Importance of including natural genetic variation in AMR databases. Accurate prediction of novel AMR determinants requires thorough cataloguing of both confirmed AMR-causing mutations and natural variation in AMR-encoding genes to avoid false positives. To illustrate this point, a PenA β-lactamase missense mutation (K96243 numbering: $PenA_{S78F}$; encoded by *BPSS0946*) has previously been linked to AMCr [20,23,38]. However, we found that $PenA_{S78F}$ alone is unlikely to cause AMCr due to its presence in genetically diverse AMC-sensitive strains (MIC=3-4 μg/mL in strains MSHR0291, MSHR0668, MSHR0848, MSHR0911, MSHR1711, MSHR2212, MSHR3902, MSHR4797, MSHR8392, and MSHR9887). Instead, AMCr is likely conferred by both $PenA_{S78F}$ and *penA* upregulation, the latter of which can be caused either by mutations within the 5' untranslated region [22], or by *penA* CNVs [11]. We therefore included $PenA_{S78F}$ as a putative stepwise mutation in the *B. pseudomallei* ARDaP database (Table 2), with an additional *penA* upregulation mutation required to confer the AMCr phenotype. In another example, we observed that both AMR and antimicrobial-sensitive strains can possess 3'-truncated *amrR* (Table 2). Multiple frameshift mutations and deletions in *amrR* are associated with MEMr[8] due to loss-of-repressor function (Fig. 3). However, certain 3' region mutations (i.e. those affecting residues ~210-223) do not cause MEMr (Table 2). To accommodate this natural genetic variation, we coded ARDaP to ignore these non-causal 3' variants, thereby greatly reducing false-positive MEMr rates.

AMR predictive capacity of ARDaP. We tested ARDaP's predictive capacity to identify the causative mutation/s in three clinical MEMr strains (MSHR1058, MSHR1174, and MSHR8777; MEM MIC range: 4–12 μg/mL; Table 3) with no previously reported AMR determinants. All patients had received MEM treatment prior to isolate retrieval. ARDaP identified novel *amrR* mutations in each strain, all of which resulted in AmrR loss-of-function ($AmrR_{\Delta P81\text{-}H223}$ in MSHR1058; $AmrR_{G149fs}$ in MSHR1174; $AmrR_{\Delta A128\text{-}H223}$ in MSHR8777; Fig. 3).

Reversions and unusual antimicrobial susceptibility. Aminoglycoside- and macrolide-class antibiotics are typically not included in melioidosis treatment regimens due to near-ubiquitous intrinsic resistance; indeed, GEN resistance is commonly used for *B. pseudomallei* selection [18]. However, rare cases of sensitivity have been documented, such as in ST-881 and ST-997 strains from Sarawak, Malaysian Borneo, which naturally encode $AmrB_{T368R}$, resulting in AmrAB-OprA loss-of-function and unusual aminoglycoside and macrolide susceptibility [18]. An AmrAB-OprA loss-of-function variant has also been described in Bp1651 ($AmrB_{A254fs}$) [38]. Although ARDaP detected *amrR* loss in MSHR1043, the co-presence of *amrA* loss ($AmrA_{L247fs}$) resulted in reversion of MEMr to a wild-type MIC (0.75 μg/mL). This reversion also causes unusual gentamicin (MIC=1 μg/mL) [19] and presumably kanamycin and azithromycin sensitivity. Given their confounding potential, we incorporated these reversions into ARDaP to more accurately reflect the true strain phenotype.

ARDaP performance on mixed sequence data. To assess the performance of the mixture function in ARDaP, Illumina reads from antimicrobial-sensitive and AMR *B. pseudomallei* strains (Table 4) were mixed at ratios ranging from 5:95 to 95:5. ARDaP identified three AMR determinants down to the lowest tested ratio of 5% minor allele frequency: a *penA* 10x CNV from MSHR8441, a *penA* 30x CNV in MSHR5654, and $PenA_{C69Y}$ (K96243 numbering: $PenA_{C75Y}$) in MSHR5654 (Table 4). The other determinants were identified by ARDaP when present at minor allele frequencies of 10% ($Ptr1_{R21fs}$, $Ptr1_{A22\_G23ins\_R-R-A}$, $BpeT_{T314fs}$, and $GyrA_{Y77S}$), 15% ($BPSL3085_{S130L}$), and 50% ($AmrR_{\Delta V62\text{-}H223}$) (Table 4). The high sensitivity of $PenA_{C75Y}$ detection can be explained by the multicopy nature of this gene in MSHR5654; this missense variant likely has a sensitivity closer to 10−15% when present as a single copy, as observed with missense variants Gyr_{Y77S} and $BPSL3085_{S130L}$ (Table 4). Gene truncations (e.g. $AmrR_{\Delta V62\text{-}H223}$) had the lowest sensitivity.

Next, ARDaP was tested on a previously detected AmrR mixture from strain MSHR9021, which encodes $AmrR_{S166P}$ and $AmrR_{A145fs}$ variants at ~66% and ~33% allele frequencies, respectively [8] (Table 1). ARDaP detected $AmrR_{S166P}$ and $AmrR_{A145fs}$ at allele frequencies of 63% and 31%, respectively, thus closely reflecting their known proportions.

ARDaP reports. ARDaP generates an easy-to-interpret report that summarises the AMR determinants and associated antibiotic phenotype/s for each genome (Fig. 2). This report summarises AMR findings for first-line, second-line, and tertiary antibiotics, along with instances of unusual antibiotic susceptibility, and has been designed to prioritise a clinical workflow. In addition, the ARDaP report lists stepwise AMR determinants, thereby informing early treatment shifts aimed at mitigating the risk of AMR emergence and fixation.

5. Discussion

This study describes the development and first-described implementation of the new AMR tool, ARDaP, for truly comprehensive AMR determinant identification from NGS and genome assembly data, including from mixed (e.g. metagenomic) data. Using the melioidosis pathogen, *B. pseudomallei*, as a model organism, we demonstrate that ARDaP provides several key advantages over existing AMR software. From 47 well-validated isolate genomes, we found that CARD, ResFinder, and AMRFinderPlus failed to identify any AMR determinants in *B. pseudomallei*, and for CARD and AMRFinderPlus, all 47 isolates were found to harbour AMR determinants, despite 16 being antimicrobial-sensitive. Two reasons underpin this shortcoming of existing AMR software: first, *B. pseudomallei* exclusively acquires AMR through chromosomal mutation, thereby limiting the value of tools that are heavily biased towards gene gain identification; and second, these AMR tools are unable to identify AMR variants conferred by indels, CNVs, inversions, or gene loss/truncation. Although ARIBA can detect indels, in our hands, this tool provided comparable information to variant report outputs generated by comparative genomic pipelines; the user requires extensive domain-specific knowledge to accurately identify and interpret outputs, particularly when trying to differentiate AMR-conferring variants from naturally-occurring genetic variation. Other shortcomings of ARIBA include cumbersome and labour-intensive input file requirements, restrictions on database construction (e.g. reference genomes with indels and nonsense mutations cannot be included), and an inability to identify CNVs. In contrast, ARDaP uses standardised variant annotation, can differentiate natural gene variation from known and putative AMR determinants, can detect CNVs and inversions, and provides a user-friendly output that does not require domain-specific knowledge for accurate interpretation.

Assessment of ARDaP's performance across the 47 characterised isolates demonstrated that this software accurately identified all AMR determinants (including stepwise variants) in all strains, except for four false positives. The first of these, MSHR5654, was predicted to be MEMr due to a BpeT truncation;[11] however, Etesting showed MEM sensitivity (2 μg/mL) in this strain, just below the MEMr threshold (≥3 μg/mL) [11,25]. Although alterations in BpeT have been putatively linked with MEMr in MSHR1300 (4 μg/mL) [8] and 354e (6 μg/mL),[9] the role of BpeT mutations in conferring MEMr is contentious[27]. In support of this notion, MSHR1300 also encodes $AmrR_{K13fs}$, a TetR-family *cis*-acting repressor of the AmrAB-OprA RND efflux pump, which likely causes MEMr in its own right[8], and in 354e, the ~800kb inversion likely also affects other AMR-conferring genes besides *bpeT*. As such, the *B. pseudomallei* ARDaP database was updated to flag *bpeT* variants as stepwise mutations rather than solely conferring MEMr (Table 1; Table 2), thereby correcting the original false-positive call for MSHR5654. This issue highlights the complexity of unravelling AMR determinants, and in this case, the need for additional work to determine a role, if any, for *bpeT* mutations in conferring MEMr.

The remaining three false-positive strains, all of which are longitudinal isolates retrieved from a single chronic CF airway infection (patient CF9) [11], were predicted by ARDaP to be DOXi or DOXr due to the presence of a $BPSL3085_{A88fs}$ variant in these strains [25]. Indeed, other CF9 isolates that harbour the $BPSL3085_{A88fs}$ variant exhibit DOXi (MSHR5665; MIC=6) or DOXr (MSHR5667; MIC=48) phenotypes (Table 1). This variant is also found in unrelated strains MSHR3683 (DOXi) and Bp1651 (DOXr). Taken together, there is strong evidence that $BPSL3085_{A88fs}$ confers DOXi or DOXr in *B. pseudomallei*. We thus postulate that MSHR5666, MSHR5669, and MSHR5670 encode an unidentified mutation that reverts them to a DOX-sensitive phenotype. Notably, all longitudinal CF9 isolates, including MSHR5666 and MSHR5669, encode *mutS* mutations, resulting in a hypermutator phenotype [11,25]. Therefore, identifying the causal basis for this reversion is non-trivial due to the large number of mutations (range: 112−157) accrued by these hypermutator strains [11]. In addition, MSHR5670 was predicted to be SXTr due to $Ptr1_{W116R}$, $MetF_{Q28P}$, and Dut_{V77A} variants, yet exhibited SXT sensitivity (Table 2). The cause of SXTr reversion in this strain is also currently unknown and requires further exploration. Our results show that chronic infections, particularly those in which hypermutated strains have emerged, represent the most challenging scenario from which to accurately predict AMR phenotypes. We therefore recommend that chronically infecting strains be subjected to conventional phenotypic testing to confirm AMR profiles predicted from NGS data.

In most melioidosis treatment guidelines, IPM has been replaced by MEM due to neurotoxicity concerns [39]. However, the recent discovery of MEMr *B. pseudomallei* has resurrected IPM as a treatment option due to a lack of cross-resistance between these carbapenems [8] and exceedingly low rates of reported IPM resistance (IPMr) [40]. The one study reporting an IPMr (MIC=8 μg/mL) *B. pseudomallei* strain, Bp1651, attributed this phenotype to a $PenA_{T147A}$ mutation

BURKHOLDERIA PSEUDOMALLEI

GENOME SEQUENCING REPORT

NOT FOR DIAGNOSTIC USE

ARDaP

Patient Name	James Smith	Barcode	BARCODE
Birth Date	1/01/1990	Patient ID	MSHR5654
Location	Darwin	Sample Type	Blood
Sample Source	Blood	Sample Date	2020-07-30
Sample ID	MSHR5654	Sequenced From	Cultured isolate
Reporting Lab	RDH	Report Date/Time	2020-07-30
Requested By	Dr. Requestor Name	Requestor Contact	req_contact@genome.com

Summary
The specimen was interogated with the *Burkholderia pseudomallei* database
It is predicted to be resistant to Ceftazidime, Trimethoprim/sulfamethoxazole, Ciprofloxacin.

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 determinant
First-line	Sensitive	Meropenem	No resistance detected
	Resistant	Ceftazidime	Multiple determinants
Second-line	Sensitive	Imipenem	No resistance detected
	Resistant	Trimethoprim/sulfamethoxazole	Multiple determinants
	Sensitive	Amoxicillin/Clavulanic acid	No resistance detected
	Sensitive	Doxycycline	No resistance detected

Extended/non-clinical drug susceptibility

Drug class	Interpretation	Drug	Resistance determinant
Tertiary	Resistant	Ciprofloxacin	gyrA BPSL2521 Tyr77Ser

Intrinsic drug resistance/unusual drug susceptibility

Drug class	Interpretation	Drug	Mechanism of sensitivity
Intrinsic	Resistant	Gentamicin	No sensitivity detected

Antimicrobial determinant details
gyrA|BPSL2521|Tyr77Ser|CIPr|Confers CIP MIC⩾32 µg/mL
penA|BPSS0946|Cys75Tyr|CAZr|Confers CAZ MIC⩾256 µg/mL; also known as PenA Cys69Tyr
BPSS0039|ptr1|Arg21fs|SXTr|<PRECURSOR MUTATION>. Confers SXT MIC⩾32 µg/mL in combination with BpeT Thr314fs, BpeT His274Tyr, or BpeT loss. Confers intermediate resistance by itself
BPSS0290|bpeT|Thr314fs|SXTr|<PRECURSOR MUTATION>. Confers SXT MIC⩾32 µg/mL in combination with Ptr1 Arg20fs
penA|BPSS0946|upregulation|CAZr|0|10x copy number confers MIC=12 ug/uL; 2x copy number does not increase MIC (Viberg et al., 2017 MBio 8(2):00356-17)
BPSS0039|ptr1|loss|SXTr|1|Confers SXT resistance in combination with efflux pump upregulation. Confers intermediate resistance by itself
BPSS0290|bpeT|loss|SXTr|0|<PRECURSOR MUTATION>. Isolates with this mutation are more likely to develop SXT resistance

Natural variation that does not confer antimicrobial resistance
amrR|BPSL1805|Val222fs|None|Known variation in amrR that does not cause antibiotic resistance

Fig. 2. Example antimicrobial resistance (AMR) summary report produced by ARDaP for strain MSHR5654. The final step in the ARDaP pipeline is the production of a user-friendly report that summarises patient and sample details, confirms that the given isolate was interrogated against the correct database (in this case, *Burkholderia pseudomallei*), identifies any predicted AMR (including what mutation/s has/have been detected) and what antibiotic/s have been affected, identifies unusual antimicrobial sensitivity and natural variation that does not confer AMR, and identifies stepwise mutations that lower the barrier to AMR development.

(K96243 numbering: $PenA_{T153A}$) combined with *penA* upregulation due to a promoter mutation [38]. We subsequently refuted the role of the $PenA_{T147A}$ variant alone in conferring IPMr by identifying three genetically unrelated $PenA_{T153A}$-encoding strains that were IPM-sensitive [8]. Further, this variant is dominant (>50%) in publicly available *B. pseudomallei* genomes, none of which have been reported as IPMr. Given that $PenA_{T147A}$ occurs at a very high rate in the wild-type *B. pseudomallei* population, and none have been shown to exhibit IPMr, this mutant has not been included in our ARDaP database. However, this variant can readily be added as a stepwise AMR determinant should further evidence come to light about its role in conferring AMR.

ARDaP has not just been designed to detect known AMR determinants; its databases can also be configured to ignore natural variation, and to predict novel AMR variants from known AMR loci, both of which are essential facets of accurate AMR prediction from WGS data. For example, our initial analyses identified several *amrR* mutants that were predicted to confer MEMr. However, Etesting of

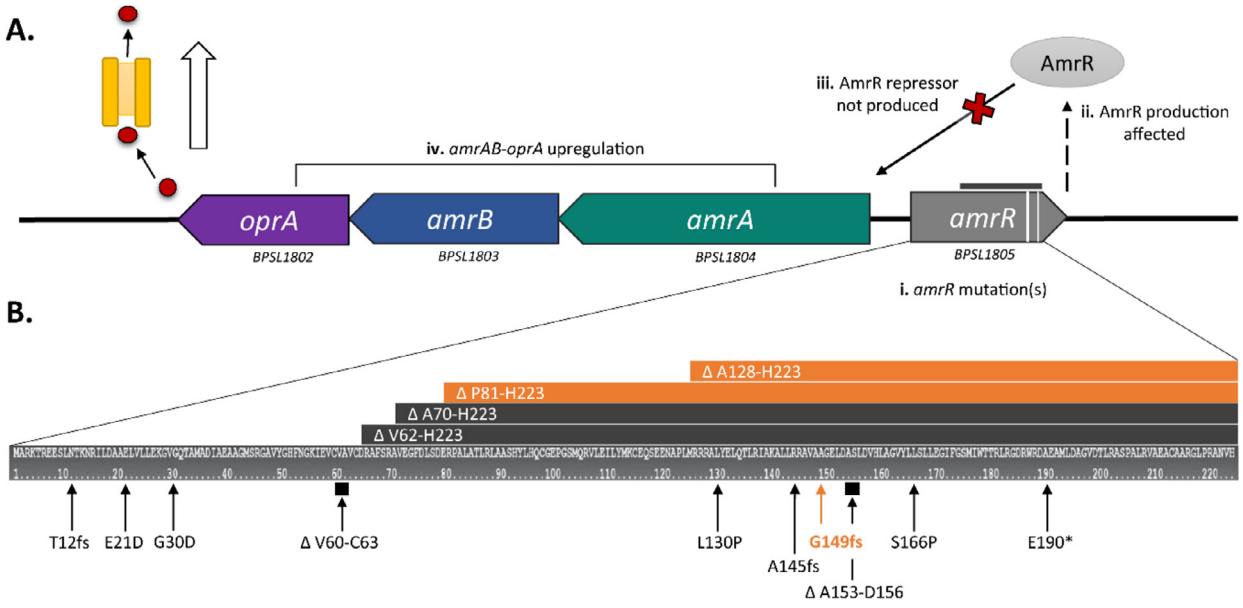

Fig. 3. Operon organisation of the *Burkholderia pseudomallei* AmrAB-OprA resistance-nodulation-division efflux pump and loss-of-function mutations in its TetR-type regulator, AmrR. A. Transcriptional organisation of the *amrR* (*BPSL1805*), *amrA* (*BPSL1804*), *amrB* (*BPSL1803*) and *oprA* (*BPSL1802*) 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 overexpression and resistance to meropenem and aminoglycoside antibiotics. B. Distribution and annotation of *amrR* mutations. Eleven previously observed *amrR* mutations (in black) [8] have been augmented with three novel mutations identified in the current study (orange); $AmrR_{G149fs}$, $AmrR_{\Delta P81\text{-}H223}$, and $AmrR_{\Delta A128\text{-}H223}$, all of which cause *amrR* loss-of-function, resulting in efflux pump overexpression and meropenem resistance.

these strains showed that most strains were MEM-sensitive. Closer inspection of the *amrR* gene found considerable variability spanning residues 210 to 223 in these strains, indicating that these 3' mutations do not impact the regulator or repressor activity of the AmrAB-OprA RND efflux pump. By ignoring this highly mutable portion of the *amrR* gene, we dramatically reduced the number of false positive AMR determinants identified by ARDaP.

To predict AMR determinants, ARDaP will flag known AMR genes encoding novel high-consequence (i.e. nonsense, frameshift, or gene loss) mutations for further user assessment. This prediction capacity of ARDaP was tested in three previously genetically uncharacterised MEMr strains: MSHR1058, MSHR1174, and MSHR8777, each of which was isolated from clinical infections where MEMr emerged during MEM therapy. In each case, ARDaP identified novel, high-consequence mutations affecting *amrR*, the local regulator of the AmrAB-OprA RND efflux pump (Fig. 3; Table 3). This result provides further confirmation of the link between MEM administration and potential treatment failure due to AmrR mutability [8], and demonstrates the value of ARDaP for predicting novel AMR determinants.

Genetic variants conferring unusual antimicrobial susceptibility, including those brought about by reversions, represent an important yet commonly overlooked aspect of AMR detection and prediction software. Most *B. pseudomallei* strains are naturally resistant to aminoglycosides and macrolides, meaning that these antibiotic classes are almost universally excluded from melioidosis treatment regimens due to inherent AMR towards these antibiotic classes; however, there are notable exceptions. For example, certain *B. pseudomallei* clones from Malaysian Borneo are naturally susceptible to gentamicin, kanamycin, and azithromycin due to AmrAB-OprA loss-of-function [18], and this phenotype can also arise *in vivo* due to within-host evolution. Importantly, such strains can conceivably be effectively treated with aminoglycoside and macrolide antibiotics, which are not typically considered for melioidosis treatment due to assumed inherent resistance. Strains encoding AmrAB-OprA loss-of-function variants (e.g. $AmrB_{T368R}$, $AmrB_{A254fs}$, $AmrA_{L247fs}$) are also at far lower risk of developing MEMr than wild-type strains due to the abrogation of deleterious *amrR* mutations that would otherwise cause MEMr. The identification of strains encoding *amrAB-oprA* loss-of-function mutations would thus strongly support long-term MEM use due to a far lower risk of MEMr development in such cases. These findings highlight the value of including sensitivity-conferring variants in AMR databases by increasing the antibiotic arsenal in naturally multidrug-resistant pathogens where treatment options are limited.

The ARDaP algorithm is mixture-aware, an important feature for detecting emerging AMR determinants in mixed strain data (e.g. non-purified colonies, culture sweeps, total clinical specimens). Using mixtures of AMR and antimicrobial-sensitive strains at varying ratios, we defined the limits of mixture detection in ARDaP for common AMR variants in *B. pseudomallei*. Overall, ARDaP confidently identified AMR determinants in the tested mixtures, albeit with varying sensitivities. CNVs were most readily detected by ARDaP, with 10x and 30x CNVs able to be distinguished at the lowest tested allele frequency of 5%. AMR-conferring SNPs and indels were robustly detected at minor allele frequencies of 10-15% (Table 4). Gene truncations were the least sensitive AMR variant type to detect from mixtures, with the one truncation examined in this study ($AmrR_{\Delta V62\text{-}H223}$) only detectable when present at $\geq$50% allele frequency. A possible explanation for the much lower sensitivity of gene truncation variant detection in mixed data is the challenge of discriminating gene loss from Illumina depth coverage variation, coupled with inherent limitations in short-read data mapping. Further validation of specific variant mixtures is recommended when new mixtures are identified to determine their sensitivity. In addition, deeper sequencing (e.g. 100–500x) should enable more robust mixture detection at lower allele frequencies.

The easy-to-interpret AMR summary report generated by ARDaP (Fig. 2) represents a major improvement over current AMR software such as AMRFinderPlus, ARIBA, CARD, and ResFinder, which require an intimate understanding of AMR determinants to correctly interpret outputs and to ignore naturally occurring genetic variation. The AMR report produced by ARDaP represents a crucial 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, and will help to accelerate the translation of NGS-to-bedside diagnostics.

Caveats and Limitations. We acknowledge that there are several limitations to our study. First, we have, to date, only developed one pathogen-specific AMR database for ARDaP; additional databases need to be populated for other microbes of interest, the curation of which is time-consuming and laborious. To begin addressing this task, we are currently developing ARDaP-compatible AMR databases for *Haemophilus influenzae* and *Pseudomonas aeruginosa*. Second, *B. pseudomallei* is hyperendemic in many resource-poor tropical regions, where access to NGS platforms and bioinformatics expertise is limited or non-existent. Therefore, ARDaP is unlikely to guide public health interventions in these regions until NGS capacity is better developed and funded, meaning that a large proportion of AMR infections in melioidosis-endemic regions will remain undetected. Despite this shortcoming, ARDaP provides a major advance towards the routine use of NNNGS for rapid and accurate acquired AMR detection in *B. pseudomallei* in well-resourced settings, and will be essential for informing treatment shifts and improving patient outcomes. Third, the lack of *B. pseudomallei* human-to-human and zoonotic transmission limits the use of AMR prediction in *B. pseudomallei* to individual cases rather than for larger epidemiological studies (e.g. outbreak tracking or global AMR dispersal). Finally, our study only included 25 AMR strains, the majority of which have been identified from our isolate collection. Whilst modest, these strains represent all publicly available, global, nonredundant AMR *B. pseudomallei* strains. Dual-use concerns in Select Agent pathogens such as *B. pseudomallei* mean that it is not possible to induce AMR in the laboratory setting, which has hampered the identification of novel AMR determinants as AMR identification is only possible from infected hosts. More work is needed to identify AMR strains and their associated determinants in *B. pseudomallei*, particularly from melioidosis hotspots in Asia, Africa, and Central and South America.

Acknowledgements

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, Rhys White (University of Queensland) for helpful discussions about software functionality, and Vanessa Rigas, Glenda Harrington, and Mark Mayo (Menzies School of Health Research) for isolate inventory support.

Contributors

DSS conceived of the study; EPP and DSS designed the study; JRW, EPP, and DSS generated laboratory data and performed laboratory analyses; EJS and DSS wrote the software; DEM, EPP, and DSS performed data analysis, literature searches, figure generation, software testing, and feature development; BJC provided clinical data and isolates; DEM, EPP, and DSS wrote the manuscript; and BJC, EPP, and DSS obtained funding for the study. All authors approved of the final manuscript.

Declaration of Competing Interests

The authors have no financial or non-financial competing interests.

Funding sources

This study was supported by the National Health and Medical Research Council (awards 1046812, 1098337, and 1131932 [the HOT NORTH initiative]), Advance Queensland (awards AQIRF0362018 and AQRF13016-17RD2), the Australian Government, and James Cook University. The funders had no role in study design, data acquisition, analysis, interpretation, writing or submission of the manuscript.

Data sharing statement

All genome sequence data examined in this study are publicly available on the NCBI GenBank or Sequence Read Archive databases (Table 1). The ARDaP code is freely available and accessible at: https://github.com/dsarov/ARDaP

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2020.103152.

References

[1] Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet 2012;13(9):601–12.
[2] Lakin SM, Dean C, Noyes NR, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res 2017;45(D1):D574–80.
[3] Gupta SK, Padmanabhan BR, Diene SM, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58(1):212–20.
[4] Hunt M, Mather AE, Sanchez-Buso L, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb Genom 2017;3(10):e000131.
[5] Jia B, Raphenya AR, Alcock B, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 2017;45(D1):D566–73.
[6] Bortolaia V, Kaas RS, Ruppe E, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 2020;75(12):3491–500.
[7] Feldgarden M, Brover V, Haft DH, et al. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 2019;63(11): e00483–19.
[8] Sarovich DS, Webb JR, Pitman MC, et al. Raising the stakes: loss of efflux-pump regulation decreases meropenem susceptibility in *Burkholderia pseudomallei*. Clin Infect Dis 2018;67(2):243–50.
[9] Hayden HS, Lim R, Brittnacher MJ, et al. Evolution of *Burkholderia pseudomallei* in recurrent melioidosis. PLoS One 2012;7(5):e36507.
[10] Tristram S, Jacobs MR, Appelbaum PC. Antimicrobial resistance in *Haemophilus influenzae*. Clin Microbiol Rev 2007;20(2):368–89.
[11] Viberg LT, Sarovich DS, Kidd TJ, et al. Within-host evolution of *Burkholderia pseudomallei* during chronic infection of seven Australasian cystic fibrosis patients. MBio 2017;8(2):e00356–17.
[12] Nicoloff H, Hjort K, Levin BR, Andersson DI. The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification. Nat Microbiol 2019;4(3):504–14.
[13] Schweizer HP. Mechanisms of antibiotic resistance in *Burkholderia pseudomallei*: implications for treatment of melioidosis. Future Microbiol 2012;7 (12):1389–99.
[14] Dance D. Treatment and prophylaxis of melioidosis. Int J Antimicrob Agents 2014;43(4):310–8.
[15] Limmathurotsakul D, Peacock SJ. Melioidosis: a clinical overview. Br Med Bull 2011;99(1):125–39.
[16] Crowe A, McMahon N, Currie BJ, Baird RW. Current antimicrobial susceptibility of first-episode melioidosis *Burkholderia pseudomallei* isolates from the Northern Territory, Australia. Int J Antimicrob Agents 2014;44(2):160–2.
[17] 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 2017;10:291–4.
[18] Podin Y, Sarovich DS, Price EP, et al. *Burkholderia pseudomallei* isolates from Sarawak, Malaysian Borneo, are predominantly susceptible to aminoglycosides and macrolides. Antimicrob Agents Chemother 2014;58(1):162–6.
[19] Price EP, Sarovich DS, Mayo M, et al. Within-host evolution of *Burkholderia pseudomallei* over a twelve-year chronic carriage infection. MBio 2013;4(4): e00388-13.
[20] Rholl DA, Papp-Wallace KM, Tomaras AP, Vasil ML, Bonomo RA, Schweizer HP. Molecular investigations of PenA-mediated β-lactam resistance in *Burkholderia pseudomallei*. Front Microbiol 2011;2:139.
[21] Sam IC, See KH, Puthucheary SD. Variations in ceftazidime and amoxicillin-clavulanate susceptibilities within a clonal infection of *Burkholderia pseudomallei*. J Clin Microbiol 2009;47(5):1556–8.
[22] Sarovich DS, Price EP, Limmathurotsakul D, et al. Development of ceftazidime resistance in an acute *Burkholderia pseudomallei* infection. Infect Drug Resist 2012;5:129–32.
[23] Tribuddharat C, Moore RA, Baker P, Woods DE. *Burkholderia pseudomallei* class A β-lactamase mutations that confer selective resistance against

ceftazidime or clavulanic acid inhibition. Antimicrob Agents Chemother 2003;47(7):2082–7.
[24] Chirakul S, Norris MH, Pagdepanichkit S, et al. Transcriptional and post-transcriptional regulation of PenA β-lactamase in acquired *Burkholderia pseudomallei* β-lactam resistance. Sci Rep 2018;8(1):10652.
[25] Price EP, Viberg LT, Kidd TJ, Bell SC, Currie BJ, Sarovich DS. Transcriptomic analysis of longitudinal *Burkholderia pseudomallei* infecting the cystic fibrosis lung. Microb Genom 2018;4(8):e000194.
[26] Chantratita N, Rholl DA, Sim B, et al. Antimicrobial resistance to ceftazidime involving loss of penicillin-binding protein 3 in *Burkholderia pseudomallei*. Proc Natl Acad Sci U S A 2011;108(41):17165–70.
[27] Podnecky NL, Rhodes KA, Mima T, et al. Mechanisms of resistance to folate pathway inhibitors in *Burkholderia pseudomallei*: deviation from the norm. MBio 2017;8(5) e01357-17.
[28] Rhodes KA, Somprasong N, Podnecky NL, Mima T, Chirakul S, Schweizer HP. Molecular determinants of *Burkholderia pseudomallei* BpeEF-OprC efflux pump expression. Microbiology 2018;164(9):1156–67.
[29] Webb JR, Price EP, Currie BJ, Sarovich DS. Loss of methyltransferase function and increased efflux activity leads to doxycycline resistance in *Burkholderia pseudomallei*. Antimicrob Agents Chemother 2017;61(6):e00268–17.
[30] Huang W, Li L, Myers JR, Marth GT. ART: a next-generation sequencing read simulator. Bioinformatics 2012;28(4):593–4.
[31] Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 2013 **1303.3997**.
[32] Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25(16):2078–9.
[33] McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20(9):1297–303.
[34] Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics 2018;34(5):867–8.
[35] Ye K, Schulz MH, 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 2009;25(21):2865–71.
[36] Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 2012;28(18):i333–i9.
[37] Cingolani P, Platts A, Wang le L, 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) 2012;6(2):80–92.
[38] Bugrysheva JV, Sue D, Gee JE, et al. Antibiotic resistance markers in *Burkholderia pseudomallei* strain Bp1651 identified by genome sequence analysis. Antimicrob Agents Chemother 2017;61(6).
[39] Currie BJ. Melioidosis: evolving concepts in epidemiology, pathogenesis, and treatment. Semin Respir Crit Care Med 2015;36(1):111–25.
[40] Wuthiekanun V, Amornchai P, Saiprom N, et al. Survey of antimicrobial resistance in clinical *Burkholderia pseudomallei* isolates over two decades in Northeast Thailand. Antimicrob Agents Chemother 2011;55(11):5388–91.