**Bridging the gap between bioinformatics and the clinical and public health microbiology laboratory: An ISO-accredited genomics workflow for antimicrobial resistance**

Norelle L. Sherry[a,b], Kristy Horan[a], Susan A. Ballard[a], Anders Gonçalves da Silva[a], Claire L. Gorrie[c], Mark B. Schultz[a], Kerrie Stevens[a], Mary Valcanis[a], Michelle Sait[a], Timothy P. Stinear[b], Benjamin P. Howden[a,b*] and Torsten Seemann[a*]

[a] Microbiological Diagnostic Unit Public Health Laboratory (MDU-PHL), Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

[b] Department of Infectious Diseases, Austin Health, Heidelberg, Victoria, Australia

[c] Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Australia

*Co-senior authors

**Corresponding author:** Professor Benjamin P. Howden, Director, MDU Public Health Laboratory, Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, 792 Elizabeth Street, Melbourne, Victoria, Australia 3000. Email: bhowden@unimelb.edu.au. Phone: +61 3 8344 5701, fax +61 3 8344 7833.

**Running title:** Genomics workflow for antimicrobial resistance

**ABSTRACT**

Realising the promise of genomics to revolutionise routine AMR diagnosis and surveillance has been a long-standing challenge in clinical and public health microbiology. We have directly addressed this issue by creating and validating *abritAMR,* an ISO-accredited bioinformatics platform for genomics-based bacterial AMR gene detection. *abritAMR* utilises the NCBI *AMRFinderPlus* for detection of AMR genes and mutations, with additional features to classify AMR determinants into an antibiotic class. We validated *abritAMR* by comparing with multiplex PCR or gold-standard reference genomes, together representing 1500 different bacteria across 29 genera and covering 415 antibiotic resistance alleles. We also assessed inference of phenotypic resistance by comparing genomic predictions with agar dilution results for 864 *Salmonella* spp. Performance of *abritAMR* was excellent, detecting AMR genes with 99.9% accuracy (95% CI 99.9-99.9%), 97.9% sensitivity (95% CI 97.5-98.4%) and 100% specificity (100-100%). Phenotypic inference of resistance for *Salmonella* spp. was equally impressive, with 98.9% accuracy (98.7-99.1%). Validation data were submitted to the governing authority and ISO15189 accreditation was achieved. Implementation of *abritAMR* resulted in streamlined bioinformatics and reporting pathways, and it was readily updated and re-verified with database revisions or changes in reporting requirements. *abritAMR* is publicly and freely available to assist clinical and public health microbiology laboratories everywhere harness the power of AMR genomics in their professional practice.

**INTRODUCTION**

Antimicrobial resistance (AMR) is an increasingly well-recognised threat to global health[1-3]. A clear understanding of the genomic and mechanistic basis for AMR is required to inform clinicians and public health teams, from the level of individual patients through to population-level surveillance[4,5]. By providing additional, timely data on acquired AMR genes or gene mutations that confer resistance, genomic sequencing has the potential to significantly enhance AMR surveillance and inform patient treatment beyond conventional phenotypic susceptibility testing methods[6,7].

The use of genomics in detection and surveillance of bacterial AMR is lagging behind other applications of genomics, such as strain typing and phylogenetic analysis. Contributors to the lack of uptake include the fact that phenotypic testing can be performed more rapidly than genotypic testing for many common pathogens, and the correlation between genotype and phenotype can be variable due to incomplete knowledge of AMR mechanisms that impact function[4,8]. However, technological advances in whole genome sequencing (WGS) means the process is becoming more cost-effective and the turnaround time for sequencing a microbial genome is decreasing significantly[9].

A lack of international standards for genomic detection of AMR mechanisms means it is difficult to compare results between laboratories. To facilitate implementation, the development of standardised and extensive open-access AMR databases and the validation of bioinformatic analytical tools for the detection of AMR is crucial[4,10]. Another hurdle to the acceptance and implementation of AMR genomics is how the data can be meaningfully reported outside of research or reference laboratory settings[4]. If the implementation of WGS for AMR is going to be accepted for detection of AMR resistance, it is important to consider the way in which complex

3

genomic data is presented to clinicians, nurses, public health surveillance teams, and other stakeholders with varying understanding of genomics, and thus how to interpret findings[11].

Here we developed and validated a bioinformatic platform for detection of AMR mechanisms from whole genome sequencing data, suitable for clinical and public health microbiology reporting. We envisage that this pipeline could be adopted in public health and clinical settings to assist those involved in AMR surveillance and clinical applications.

## MATERIALS AND METHODS

**Setting and existing genomics workflow.** The Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL) is a state reference laboratory for bacterial pathogens, including carbapenemase-producing *Enterobacterales* (CPE), *Acinetobacter* spp., *Pseudomonas* spp., vancomycin-resistant enterococci (VRE) and enteric pathogens[12-14]. The laboratory has a strong emphasis on genomics, primarily for epidemiologic surveillance, with increasing applications for clinical purposes. In conjunction with the Department of Health Victoria, we have embarked upon a broad program to increase implementation of pathogen genomics for public health purposes, either enhancing or superseding current laboratory methods.

Our existing genomics workflow (incorporating sample receipt, nucleic acid extraction, library preparation, short-read sequencing (Illumina NextSeq or MiSeq), and quality control (QC) of reads including *de novo* genome assembly) has already been validated and accredited by the National Association of Testing Authorities Australia (NATA, analogous to Clinical Laboratory Improvement Amendments [CLIA] in USA)[15,16]. Briefly, single colonies from overnight pure bacterial sub-cultures were selected and placed in lysis buffer. DNA extraction was performed on the QIAsymphony using the DSP Virus/Pathogen Mini Kit, and library preparation performed using Nextera XT (Illumina Inc.) according to manufacturer's instructions. WGS was performed on NextSeq 500/550 or MiSeq platforms (Illumina Inc.), generating 150bp or 300bp paired-end reads respectively. Reads were trimmed using Trimmomatic[17], and assembled *de novo* using Shovill[18], based on SPAdes[19]. QC requirements for fastq reads to be included in subsequent analysis were (i) Q-score > 30, (ii) data with a minimum estimated average genome coverage of >40X, and (iii) estimated genome size within range for observed species.

**The *abritAMR* bioinformatics pipeline.** The aims for development of this bioinformatic pipeline were to detect AMR genes and mutations accurately and reliably from bacterial whole genome sequencing (WGS) data, which could be validated against PCR and other data sources, implemented in a public health or clinical microbiology laboratory, and successfully accredited by governing bodies. The *abritAMR* bioinformatic platform takes a genome assembly from short-read data (fasta file) as input (once it has met defined QC parameters), and includes five main components (Figure 1):

i.   *NCBI's* AMRFinderPlus *tool* – *abritAMR* implements this tool to identify AMR genes in genome sequences, using a combination of BLASTx (matching the protein sequences of AMR genes to the protein sequence of the query isolate) and Hidden Markov Models (HMMs)[20].

ii.  *NCBI's* AMRFinderPlus *database* – *abritAMR* uses this frequently updated database (https://github.com/ncbi/amr/wiki/AMRFinderPlus-database), which is a comprehensive and extensively curated database of AMR gene sequences. Current functionality includes mainly AMR genes ('core' database), with point mutations (species-specific) and virulence genes increasingly being included in the 'plus' database. In more recent iterations, AMR genes and point mutations include information about the antimicrobial class and subclass (or specific antimicrobials) that they confer resistance to.

iii. *Classification database* – While the *AMRFinderPlus* database includes some information about the antibiotic class and subclass affected for each AMR gene, these classifications are not always easily translatable for clinical and public health practice. For example, the beta-lactam subclass 'cephalosporin' includes AMR genes conferring resistance to first-generation cephalosporins (narrow-spectrum cephalosporinases, such as $bla_{OXA-1}$), or

6

third-generation cephalosporins (such as $bla_{CTX-M}$ ESBLs), which have very different implications for AMR surveillance and patient management. The local *abritAMR* classification database is based on the current version of the *AMRFinderPlus* database, with an added field ('Enhanced subclass') to translate the NCBI subclasses into more functional versions for our purposes (logic detailed in Supplementary Table S1, examples in Figure 2). This field is updated following each new database release (logic detailed in Supplementary Table S2).

iv.  *Species-specific reporting logic (AMR genes, all species)* – Currently, most AMR genes detected by this pipeline are not required to be reported for surveillance or clinical purposes; reporting data on *all* AMR genes found in an isolate runs the risk of overwhelming clients with unnecessary data and missing the most pertinent AMR genes detected. As such, we developed a reporting logic process to filter the AMR genes detected in each isolate into 'reportable' or 'non-reportable' categories, to mirror the usual reporting requirements diagnostic laboratories (Supplementary Figure S1). This logic takes into account the species when determining what is reportable, limiting the reporting of intrinsic AMR genes (such as $bla_{OXA-51}$ subtypes in *Acinetobacter baumannii*), and differentiating between AMR genes that are only reportable in certain species (e.g. ESBL genes reportable for national surveillance of *Salmonella spp.*), while always reporting significant AMR genes that are not limited by species (e.g. carbapenemase and *mcr* genes). Non-reportable genes are also made available to the reporting pathologists and senior scientists and recorded in the laboratory information management system (LIMS), enabling detailed review of all detected AMR genes, correlation with phenotype, and movement between reportable and non-reportable categories when required as part of any routine results review process before reporting.

v. *Inferred phenotype (AMR genes and mutations, validated species only)* – The pattern of AMR genes and mutations detected can be used to infer phenotype for a given isolate. In *abritAMR*, this is currently validated for *Salmonella* spp., and reported for epidemiologic purposes in our laboratory, replacing routine antimicrobial susceptibility testing (AST) of *Salmonella* spp. for public health surveillance (reporting logic detailed in Supplementary Figure S2, Inferred Antibiogram Report example shown in Supplementary Data File 2).

### *abritAMR* outputs

*abritAMR* outputs include a Detailed Report output, consisting of a table (comma separated values file) of AMR genes or mutations detected for each sample, listed by enhanced subclass (e.g. "Carbapenemase (MBL)", "Colistin"), or a Final AMR Gene Reports, a table of AMR genes detected for each sample binned into 'reportable' or 'not reportable' fields, when the species-specific reporting logic is applied (Figure 2). Additionally, when run on validated species (currently *Salmonella* spp.), *abritAMR* also produces an Inferred Antibiogram Report. All alleles listed in these outputs are either 'exact matches' (100% identity and 100% sequence coverage compared to the reference protein sequence) or 'close matches' (90-<100% identity and 90-<100% sequence coverage compared to the reference protein sequence, marked by an asterisk [*] to distinguish from exact matches), as defined by *AMRFinderPlus*. Partial matches (>90% identity, 50-<90% coverage compared to reference protein sequence) are listed separately, and must be examined further if deemed suitable for reporting. Where an internal stop codon (i.e. truncated gene) or HMM match are recorded by *AMRFinderPlus*, no result is reported by *abriTAMR*. Examples of *abritAMR* pipeline outputs are shown in Figure 2, demonstrating how the *AMRFinderPlus* output is modified by *abritAMR* (binned into enhanced subclass according to *abritAMR's* classification database, and separated into reportable and non-reportable categories by the reporting logic).

8

**Validation of the *abritAMR* pipeline.** To validate *abritAMR*, results from the pipeline were compared to results from PCR testing, Sanger sequencing and synthetic read sets as detailed below. For the purposes of validation, both 'exact' and 'close' matches were considered as 'detected'. Pre-specified sensitivity and specificity thresholds were defined for successful validation prior to analysis.

**Validation datasets**

*PCR.* This dataset included 1184 bacterial isolates (42 species), that had previously been tested by PCR, including a a carbapenemase and ESBL real-time multiplex PCR (*n*=1020 isolates, AusDiagnostics 16-well CRE panel, catalogue no. 21098, version 03; Sydney, Australia), *van* gene PCR (*n*=121, in-house assay for *vanA, vanB, vanC1* and *vanC2/3* genes[21]) and *mecA* PCR (*n*=43, in-house assay for *mecA*[22])(Supplementary Figures S3-S6).

*PCR and Sanger sequencing for allelic variants*. This dataset included 347 isolates (20 species) with carbapenemase resistance genes detected by a range of carbapenemase and ESBL PCR assays across six different carbapenemase resistance gene families (targets and primers detailed in Supplementary Table S3), with Sanger sequencing subsequently performed to identify the carbapenemase allelic variant (Supplementary Figures S7 & S8).

*Synthetic reads*. For the remaining AMR gene targets where PCR was not readily available to compare with *abritAMR*, we created synthetic short-read sequence data from complete, publicly available genomes from RefSeq or GenBank, and compared *abritAMR* results on synthetic short reads to *AMRFinderPlus* results from the complete genomes. To do this, we generated synthetic 150bp paired-end reads using the *art-illumina* tool[23] to fragment the complete genome sequences, incorporating error profile data from a NextSeq500 sequencer, at 40X to 150X average genome coverage (40X is the minimum coverage accepted for QC) (Figure 3). This dataset comprised 321 isolates (49 species) covering 415 unique AMR alleles from 43 resistance subclasses

(Supplementary Table S1 and Figures S9 & S10). *abritAMR* results from synthetic reads were compared to (native) *AMRFinderPlus* results from complete genome sequences. This allowed direct comparisons of presence or absence of AMR genes, therefore avoiding the problem of discrepancies in AMR gene nomenclature that may lead to false discordance if two different AMR gene databases were compared.

***Precision testing***. *abritAMR* results from a test panel of 13 organisms (12 genera, Supplementary Table S4) sequenced multiple times (both within and across sequencing runs) using different sequencing platforms in our laboratory (NextSeq and MiSeq) and a range of sequencing modes (low, mid and high throughput) and read lengths (75-300bp). Different combinations were compared to assess analytical precision (repeatability and reproducibility).

***Determination of limit of detection***. The limit of detection for molecular assays is normally the lowest amount of nucleic acid target that can be detected by the assay. This definition is not strictly applicable to whole genome sequencing, as the WGS assay is qualitative with a standardised DNA concentration being used in the sequencing reaction. Instead, the limit of detection in this context was calculated as the minimum average coverage across the genome required for accurate detection of gene targets or allele variants.  Synthetic paired-end reads (150bp) were generated at a range of sequencing coverages, from the minimum average coverage accepted for our routine QC (40 X) up to 150 X coverage.

***Determination of inferred phenotype (*Salmonella* spp.)***. We validated phenotypic inference (Susceptible/Intermediate/Resistant, S/I/R) against an existing dataset of 864 sequenced *Salmonella* spp. with antimicrobial susceptibility (AST) data generated by agar dilution from 2018-2019. For the fluoroquinolone drug class, the S/I/R phenotypes associated with combinations of AMR genes and mutations were analysed to determine the relative weighting of each AMR mechanism to infer a phenotype most reliably from *in silico* analysis.

10

**Discordant result resolution.** Discordant results were divided into two categories: firstly, PCR negative, WGS positive (false positive) - this may be due to the AMR gene detected by WGS not being included in the range of the PCR panel. If the gene was known to be included in the range of the PCR panel (as stated by the manufacturer), the isolate was retested by PCR and WGS to resolve this discrepancy. Secondly, PCR positive, WGS negative (false negative) – this may be due to an AMR gene being fragmented across two or more contigs, hence partial matches were assessed; if no partial matches were found, the sequence was interrogated using alternative tools; if this failed to resolve the discrepancy, the isolate was retested by PCR and WGS. Where possible, discrepancies between phenotypic and genotypic results were investigated through repeat phenotypic testing and/or repeat sequencing of the isolate.

**Re-verification processes.** In accordance with ISO standards, the *abritAMR* pipeline must be re-verified after each database or tool update. Database updates are reverified by confirming that the updated database performs to the same criteria as was defined in the original validation, using the synthetic dataset described above (*'abritAMR* test suite'). Updates to the *abritAMR* software may take the form of minor patches or major updates. Minor patches are changes that do not impact underlying structure or core logic of the pipeline, such as fixes for typographical errors or addition of functionality which does not impact the core logic of the tool, e.g. changes to log outputs. In these cases, a full reverification is deemed unnecessary and running of the *abritAMR* test suite is sufficient. However, other changes which may impact the core logic or structure of the outputs require a complete reverification as described for database updates. Any change in performance is assessed, the cause identified, and modifications made before the changes are implemented for reporting. All changes to *abritAMR* are tracked in GitHub and the versions implemented managed using *conda*.

**Data availability.** Code for the *abritAMR* pipeline is publicly available at https://github.com/MDU-PHL/abritAMR. Accession numbers for genomes used in generation of synthetic reads are supplied in Supplementary File 2 and uploaded to NCBI Sequence Read Archive (BioProjects PRJNA529744 (CPE), PRJNA842191 (non-CPE Gram negatives), PRJNA495857 (VRE), PRJNA319593 (*Salmonella*), and *Staphylococcus* spp. (PRJNA842192)).

**RESULTS**

The *abritAMR* pipeline performed very well against the four validation panels, with an overall accuracy of 99.9% (95%CI 99.9-99.9%), sensitivity 97.9% (97.5-98.4%), and specificity 100% (100-100%)(Table 1). Importantly, the *abritAMR* pipeline was reliable for the high-risk AMR gene classes that are notifiable as part of our national critical antimicrobial resistance surveillance system (CARAlert) in Australia[24], with 99.9% accuracy (95%CI 99.9-100%), 98.9% sensitivity (98.3-99.3%) and 100% specificity (100-100%) across these classes (carbapenemases, 16S ribosomal methyltransferases, mobile colistin resistance genes, ESBLs (including AmpCs), vancomycin resistance genes, and oxazolidinone and phenicol resistance (*optrA, cfr* and *poxtA* genes)).

**Validation results compared to PCR and Sanger sequencing**. The *abritAMR* pipeline was highly accurate compared to PCR (carbapenemase, ESBL, *van* and *mec* gene PCRs) with 1179/1184 (99.6%) resistance genes correctly detected, and compared well to Sanger sequencing (carbapenemase allele calling) with 355/356 (99.7%) alleles correctly identified by WGS. After discrepancy resolution (including repeat PCR and/or WGS, or examination of partial genes detected by *abritAMR*), five discrepancies between PCR and WGS results remained, including three potential false negatives (PCR positive, WGS negative) consisting of one CTX-M and two CMY genes not detected by *abritAMR;* at least one of these was due to the presence of

12

a contig break in the gene leading to smaller fragments not detected by *AMRFinderPlus*. Additionally, two potential false positives (PCR negative, WGS positive) were identified, one CMY-42 and one IMP-62, confirmed by repeat PCR and sequencing; both genes were reported to be within the inclusivity range of the assay as per the manufacturer's instructions, although this was validated by *in silico* PCR by the manufacturer (and observed in our dataset). Alternatively, these discrepancies may be due to plasmid dropout in culture (which is commonly observed with suspected CPE isolates such as these), as different colonies with and without the ESBL/AmpC gene may have been picked for PCR and WGS, potentially explaining these discrepancies. No discrepant results were detected for *mecA* and *van* gene detection, and only one allele was incorrectly assigned compared to Sanger sequencing (99.7% accuracy)(Table 1 and Figure 4). Overall performance of *abritAMR* against PCR yielded 99.6% accuracy (95%CI 99.0-99.9%), 99.6% sensitivity (99.0-99.9%) and 99.4% specificity (97.9-99.9%).

**Identification of AMR genes from synthetic reads.** The presence or absence of 415 AMR genes across 321 genomes (133215 alleles) was tested by running *abritAMR* on synthetic reads from complete reference genomes, and comparing to the (native) *AMRFinderPlus* results on the complete genome, considered the 'gold standard' (Figure 3). Overall accuracy of AMR gene detection by *abritAMR* was excellent, with 133127/133215 alleles called correctly, resulting in 99.9% accuracy (95% CI 99.9-99.9%), 97.5% sensitivity (96.9-98.0%), and 100% sensitivity (100-100%)(Figure 5). Note that any discrepancies here include differences in *abritAMR* performance compared to *AMRFinderPlus*, as well as differences between complete genomes and (synthetic) short-read data, which likely accounts for at least a proportion of the discrepant results.

The majority of discrepancies were false negatives, with the aminoglycoside AMR genes being most common (32/88, 36.4%), especially the *aac(6')-Ib* family, implicated in 18 false negatives, and specifically the *aac(6')-Ib-cr5* allele (11/18). Some of these were detected as partial genes at

13

the site of contig breaks, possibly related to slightly higher GC content (leading to lower sequence coverage) in these genes. The other major theme was difficulty resolving sequences with multiple alleles of the same gene family, which were often collapsed into a single gene detection by *abritAMR* or miscalled as a different allele. For example, this included a sequence with CTX-M-3, CTX-M-14 and CTX-M-65 identified by *AMRFinderPlus*, and called as CTX-M-3 and CTX-M-24 by *abritAMR*. Four out of the five 'false positive' detections were actually allele miscalls within the same gene family. Collapse of repeated regions or duplicate alleles is often a feature of short-read sequencing, hence the discrepancies here may be a feature of comparing (synthetic) short-read data to complete genomes, rather than a feature of *abritAMR*.

**Limit of detection and precision.** The limit of detection of the *abritAMR* pipeline was assessed to determine the minimum average sequencing depth for acceptable accuracy of AMR gene detection (as required for clinical microbiology validation and accreditation). Accuracy was found to be consistent (99.9%) across the 40X to 150X range, with 40X being the minimum coverage accepted by our accredited quality control (QC) pipeline. Repeatability and reproducibility (precision) were assessed (replicates within and across sequencing runs) and found to be 100% concordant.

**Validation of inferred antibiogram (*Salmonella* spp.).** Validation of inferred phenotype against phenotypic AST data demonstrated 98.9% accuracy (95% CI 98.7-99.1%), 98.9% sensitivity (98.4-99.3) and 98.9% specificity (98.7-99.1%) overall (Table 1, Supplementary Table S5 and Figure 6). Accuracy of phenotypic inference was ≥98% for 11/13 antimicrobials (85%), with lower accuracy identified for streptomycin (95.5%, 95% CI 93.7-96.9%) and ciprofloxacin (96.8%, 95% CI 95.4-97.8%), similar to previous findings using different bioinformatic methods[25].

14

A number of 'false positive' results were identified for streptomycin (resistant genotype [AMR genes or mutations detected], susceptible phenotype; *n*=30/716 (4.2%) isolates). The AMR genes detected in phenotypically susceptible isolates were also detected in non-susceptible isolates, although the non-susceptible isolates more often had >1 AMR gene (1 AMR gene, 22% phenotypically resistant; 2 or more AMR genes, 81.4% phenotypically resistant), suggesting that these AMR mechanisms had small but additive effects on phenotype. Evaluation of phenotype-genotype concordance for azithromycin identified five 'false negatives' (susceptible phenotype, no AMR mechanism detected) and two 'false positives' (AMR mechanism detected but phenotypically susceptible; neither isolate carried the dominant resistance mechanism for azithromycin in *Salmonella* spp. (*mph(A);* one carried *mef(B)*, an efflux pump with variable activity, and one carried *ere(A)*, an esterase with lower affinity for azithromycin[26]).

Similar to streptomycin, ciprofloxacin also had a number of phenotype-genotype mismatches, likely due to low-level resistance conferred by AMR mechanisms. Isolates with one AMR gene most often had an intermediate phenotype (81.3% intermediate, 12.4% resistant, 6.2% susceptible), whilst isolates with ≥2 AMR genes were all phenotypically resistant. Despite these discordances, the best correlation between genotype and phenotype (S/I/R) was determined according to the number of AMR mechanisms detected (any type). This was coded into the reporting logic: absence of AMR mechanisms, 'susceptible', one AMR mechanism, 'intermediate', two or more AMR mechanisms, 'resistant'. In this application, the use of an 'intermediate' category implies that MICs are likely to be borderline for these samples, i.e. may test susceptible or resistant on AST. Note that these results are only used for epidemiologic purposes (not for patient treatment), and hence over-calling resistance is more suitable for this purpose than non-detection of AMR mechanisms.

**Sample outputs and incorporation into clinical/public health microbiology report.** Two different outputs are used in the *abritAMR* pipeline: (i) Detailed Report output, where AMR genes or mutations are shown by enhanced subclass, as classified by the *abritAMR* database, and (ii) Final AMR Gene Report output (binned into reportable and non-reportable genes) after reporting logic is applied (Figure 2 and Supplementary File 2). An example of the output of the additional module incorporating mutational resistance and Inferred Antibiogram Report phenotype (currently validated for *Salmonella spp.*) is also shown in Supplementary File 2.

**Implementation processes.** After the validation process, implementation processes included modifying report outputs, integration with the existing LIMS, and documentation of the standard operating procedure (SOP). Subsequently, all staff involved in detection, reporting or interpretation of AMR results were trained in the use and interpretation of *abritAMR*, and clients were educated about the change (although only minimal differences were noticeable to clients, such as change in report formats) before full implementation into routine workflows. Implementation led to streamlined workflows, including rapid bioinformatic processing of large sequencing runs (AMR gene detection for a 96-sample run completed in <3 minutes with 256 CPUs), and less manual re-classification of AMR gene results by laboratory scientists (e.g. moving genes between reportable and non-reportable fields, removing intrinsic AMR genes from reportable fields).

**DISCUSSION**

The use of genomics in clinical and public health microbiology (CPHM) has increased substantially in the last decade, particularly in the fields of pathogen typing and outbreak investigations[5,7,27]. Detection of AMR from WGS data has somewhat lagged behind other applications of WGS, likely due its inherent complexity in comparison to simple and effective phenotypic AST[4]. This complexity is multi-faceted but includes the vast array of resistance mechanisms for testing (a single phenotype may be encoded by many different AMR mechanisms), and the limitations of phenotype-genotype correlation, particularly for less-common organisms and drug classes[28,29]. If not addressed systematically, these issues may render genomic AMR difficult to identify comprehensively across all pathogens seen in a CPHM laboratory, and difficult to communicate to clinicians and public health units[11,30].

Globally, the paucity of highly accurate, reproducible bioinformatic tools for detection of AMR mechanisms has been recognised as one of the main limiting factors to wider application of genomics in the CPHM setting[4,31]. Here, we have designed and validated a bioinformatic platform for genomic detection of AMR mechanisms across bacterial species, and implemented it to achieve an ISO-accredited genomic workflow for AMR. This was achieved by adapting an existing software tool and database (*AMRFinderPlus*), and adding a modified classification step plus reporting logic to produce tailored reports for a CPHM audience.

This platform relies heavily on the comprehensive, well-curated and frequently updated AMR database behind *AMRFinderPlus*, as well as the excellent software tool, which uses multiple search methods to best identify AMR genes and mutations (with results annotated by the type of 'match', to allow scientists and clinicians to understand the degree of confidence behind each call)[20]. Notably, outputs for other large AMR databases, such as CARD[32] and ResFinder[33], could

be modified to achieve similar tailored reports to *abritAMR*; our choice was based on the ease of integration into our existing workflows and reporting. *abritAMR's* speed allows for rapid detection of AMR genes in routine high-throughput workflows, with AMR gene detection completed on a 96-sample sequencing run within 3 minutes. The addition of mutations to *AMRFinderPlus* for an increasing number of species has been very useful in our early applications, enabling our public health laboratory to move to a fully-genomic workflow for *Salmonella* surveillance, as all samples were sequenced for typing and phylogenetic analysis (a sub-sample of isolates still undergo AST to ensure new AMR mechanisms are detected). The capacity to include AMR genes or mutations of local significance would be a welcome addition to *AMRFinderPlus*, further extending its utility.

When this work commenced, there were no classifications of AMR mechanisms into drug classes in the *AMRFinderPlus* database, hence we created our own classification database, which has also evolved in parallel with the great advances made by the *AMRFinderPlus* team. There are (now) a small number of essential differences in the drug class classifications that we feel are important to enhance its utility for CPHM. Key examples include separating carbapenem resistance into different groups based on their mechanisms; separating into 'carbapenemase', 'carbapenemase (MBL)' and 'carbapenemase (OXA-51 family)' enables reporting each group separately, as antibiotic choices differ with metallo-beta-lactamases (MBLs) compared to non-MBL carbapenemases, and OXA-51 family carbapenemases are weak and intrinsic to *Acinetobacter spp.* and routine reporting is not required (coded as part of reporting logic). This combination of tailored classification and reporting logic allows the vast and complex array of AMR mechanisms to be distilled into results and reports that can be understood by scientists, clinicians and public health teams alike, without a great deal of prior knowledge. Future development will focus on restructuring the database to include different levels (classes, subclasses) to take advantage of the higher resolution of classifications now included in *AMRFinderPlus*.

In a CPHM setting, it is critical to validate any new test or analytical process to ensure the veracity of results, and that the results (outputs or reports in this case) are fit-for-purpose[10]. However, formal test validation and accreditation procedures are based on wet-lab assays, and not always easily transferable to new methods such as WGS[34]. This may require some creative thinking about different ways to validate a new genomic test[35], as demonstrated here with the use of synthetic sequencing reads generated from complete reference genomes. Ideally, a broad range of publicly-available reference datasets with genotypic and phenotypic data would be made freely available to assist with validation and bench-marking for databases, tools and new pipelines such as this, greatly advancing the development of AMR genomics[4,36]. Initiatives such as the NCBI National Database of Antibiotic Resistant Organisms (NDARO)[37] and PATRIC[38] are promising, but currently limited in scale, and further global data sharing is required here to advance phenotype-genotype correlations.

As with all attempts to validate new WGS pipelines and workflows, our study has limitations. The absence of a 'gold-standard' dataset to compare results from our pipeline to means that we must compare to imperfect standards, such as existing testing methods with lower resolution (PCR) and use synthetic sequencing data to compare targets not covered by PCR in our laboratory. Until these issues are addressed globally, laboratories will have to persist with these challenging comparisons, and rely on new initiatives such as proficiency testing programs (PTPs) for WGS (with participation being a requirement for test accreditation in our setting) to start to standardise results across laboratories and countries. Whilst *abritAMR* was highly accurate overall, a small proportion of discrepant results were identified, of which the majority were false negative results. Most of these discrepancies are likely due to the comparison of synthetic short-read data to complete genomes, where contig breaks within a gene result in non-detection.

We envisage that the *abritAMR* pipeline will most likely be applied in CPHM settings, and hope that it may assist sequencing laboratories address the difficult question of how to best report these data to clinicians and public health teams with limited AMR knowledge. However, it may also have utility in other settings including research, particularly where complex AMR data need to be binned into functional classes to facilitate understanding when the user is less familiar with AMR. In our view, it is critical for medical microbiologists, scientists and bioinformaticians to continue to work together to navigate the challenges of communicating complex AMR data to clients, to advance the reach of genomic AMR and maximise the benefits of this potentially transformative technology.

## ACKNOWLEDGEMENTS

## REFERENCES

1.  O'Neill J. Review on antimicrobial resistance: Tackling a crisis for the health and wealth of nations. London, UK: UK Government; 2014.

2.  Centres for Disease Control and Prevention (CDC). Antibiotic resistance threats in the United States, 2019. Atlanta, GA: U.S. Department of Health & Human Services; 2019.

3.  World Health Organization. Global action plan on antimicrobial resistance. Geneva: WHO; 2015.

4.  Ellington MJ, Ekelund O, Aarestrup FM, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clin Microbiol Infect* 2017; **23**(1): 2-22.

5.  Motro Y, Moran-Gilad J. Next-generation sequencing applications in clinical bacteriology. *Biomol Detect Quantif* 2017; **14**: 1-6.

6.  Maugeri G, Lychko I, Sobral R, Roque ACA. Identification and antibiotic-susceptibility profiling of infectious bacterial agents: a review of current and future trends. *Biotechnology Journal* 2019; **14**(1): 1700750.

7.  Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect* 2018; **24**(4): 335-41.

8.  Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet* 2019; **20**(6): 356-70.

9.  Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ. Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. *J Microbiol Methods* 2017; **138**: 60-71.

10. Gargis AS, Kalman L, Lubin IM. Assuring the quality of next-generation sequencing in clinical microbiology and public health laboratories. *J Clin Microbiol* 2016; **54**(12): 2857-65.

11. Crisan A, McKee G, Munzner T, Gardy JL. Evidence-based design and evaluation of a whole genome sequencing clinical report for the reference microbiology laboratory. *PeerJ* 2018; **6**: e4218.

12. Lane CR, Brett J, Schultz M, et al. Search and Contain: Impact of an integrated genomic and epidemiological surveillance and response program for control of carbapenemase-producing Enterobacterales. *Clin Infect Dis* 2021; **73**(11): e3912-e20.

13.    Ingle DJ, Andersson P, Valcanis M, et al. Genomic epidemiology and antimicrobial resistance mechanisms of imported typhoid in Australia. *Antimicrobial Agents & Chemotherapy* 2021; **65**(12): e0120021-e.

14.    Ingle DJ, Andersson P, Valcanis M, et al. Prolonged outbreak of multidrug-resistant *Shigella sonnei* harboring *bla*CTX-M-27 in Victoria, Australia. *Antimicrobial Agents & Chemotherapy* 2020; **64**(12).

15.    National Association of Testing Authorities Australia (NATA). 2021. https://nata.com.au/.

16.    Centres for Disease Control and Prevention (CDC). Clinical Laboratory Improvement Amendments. 2022.

17.    Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**(15): 2114-20.

18.    Seemann T. Shovill: assemble bacterial isolate genomes from Illumina paired-end reads. GitHub; 2017.

19.    Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 2012; **19**(5): 455-77.

20.    Feldgarden M, Brover V, Gonzalez-Escalona N, et al. *AMRFinderPlus* and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific Reports* 2021; **11**(1).

21.    Dutka-Malen S, Evers S, Courvalin P. Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant enterococci by PCR. *Journal of clinical microbiology* 1995; **33**(1): 24-7.

22.    Louie L, Goodfellow J, Mathieu P, Glatt A, Louie M, Simor A. Rapid detection of methicillin-resistant staphylococci from blood culture bottles by using a multiplex PCR assay. *Journal of clinical microbiology* 2002; **40**(8): 2786-90.

23.    Huang W, Li L, Myers JR, Marth GT. ART: a next-generation sequencing read simulator. *Bioinformatics* 2011; **28**(4): 593-4.

24.    Australian Commission on Safety and Quality in Health Care. National Alert System for Critical Antimicrobial Resistances (CARAlert). 2021. https://www.safetyandquality.gov.au/our-work/antimicrobial-resistance/antimicrobial-use-and-resistance-australia-surveillance-system/national-alert-system-critical-antimicrobial-resistances-caralert.

25.    Sia CM, Baines SL, Valcanis M, et al. Genomic diversity of antimicrobial resistance in non-typhoidal *Salmonella* in Victoria, Australia. *Microbial Genomics* 2021; **7**(12): 000725.

26.     Gomes C, Martínez-Puchol S, Palma N, et al. Macrolide resistance mechanisms in Enterobacteriaceae: focus on azithromycin. *Critical reviews in microbiology* 2017; **43**(1): 1-30.

27.     Armstrong GL, Maccannell DR, Taylor J, et al. Pathogen genomics in public health. *New Engl J Med* 2019; **381**(26): 2569-80.

28.     Ruppe E, Cherkaoui A, Lazarevic V, Emonet S, Schrenzel J. Establishing genotype-to-phenotype relationships in bacteria causing hospital-acquired pneumonia: a prelude to the application of clinical metagenomics. *Antibiotics (Basel)* 2017; **6**(4).

29.     Mahfouz N, Ferreira I, Beisken S, Von Haeseler A, Posch AE. Large-scale assessment of antimicrobial resistance marker databases for genetic phenotype prediction: a systematic review. *Journal of Antimicrobial Chemotherapy* 2020; **75**(11): 3099-108.

30.     Rossen JWA, Friedrich AW, Moran-Gilad J. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clinical Microbiology and Infection* 2018; **24**(4): 355-60.

31.     World Health Organization. GLASS whole-genome sequencing for surveillance of antimicrobial resistance. Geneva: WHO; 2020.

32.     Mcarthur AG, Waglechner N, Nizam F, et al. The Comprehensive Antibiotic Resistance Database. *Antimicrobial Agents & Chemotherapy* 2013; **57**(7): 3348-57.

33.     Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012; **67**(11): 2640-4.

34.     Kozyreva VK, Truong C-L, Greninger AL, Crandall J, Mukhopadhyay R, Chaturvedi V. Validation and implementation of clinical laboratory improvements act-compliant whole-genome sequencing in the public health microbiology laboratory. *Journal of clinical microbiology* 2017; **55**(8): 2502-20.

35.     Angers-Loustau A, Petrillo M, Bengtsson-Palme J, et al. The challenges of designing a benchmark strategy for bioinformatics pipelines in the identification of antimicrobial resistance determinants using next generation sequencing technologies. *F1000Research* 2018; **7**.

36.     Bogaerts B, Winand R, Fu Q, et al. Validation of a bioinformatics workflow for routine analysis of whole-genome sequencing data and related challenges for pathogen typing in a European national reference center: *Neisseria meningitidis* as a proof-of-concept. *Front Microbiol* 2019; **10**: 362.

37. National Center for Biotechnology Information. National Database of Antibiotic Resistant Organisms (NDARO). 2022. https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/ (accessed 2022-04-21 2022).

38. Davis JJ, Wattam AR, Aziz RK, et al. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic acids research* 2020; **48**(D1): D606-D12.

**Table 1. Performance characteristics of *abritAMR* bioinformatic pipeline for detection of acquired AMR genes and inferred phenotype from WGS data**

| Validation panel | Accuracy (%, 95% CI) | Sensitivity (%, 95% CI) | Specificity (%, 95% CI) | PPV (%, 95% CI) | NPV (%, 95% CI) |
|---|---|---|---|---|---|
| **Detection of acquired AMR genes** | | | | | |
| **PCR (*n*=1184 isolates)** | 99.6 (99.0-99.9) | 99.6 (99.0-99.9) | 99.4 (97.9-99.9) | 99.8 (99.1-100) | 99.1 (97.5-99.8) |
| **Synthetic reads (*n*=321 isolates, 133215 alleles[a,b])** | 99.9 (99.9-99.9) | 97.5 (96.9-98.0) | 100 (100-100) | 99.8 (99.6-100) | 99.9 (99.9-99.9) |
| **Overall performance (PCR & synthetic data)** | 99.9 (99.9-99.9) | 97.9 (97.5-98.4) | 100 (100-100) | 99.8 (99.6-99.9) | 99.9 (99.9-99.9) |
| **Critical AMR subclasses[c] (all validation sets)** | 99.9 (99.9-100) | 98.9 (98.3-99.3) | 100 (100-100) | 99.8 (99.5-100) | 100 (99.9-100) |
| **Sanger sequencing for allelic variants (*n*=356 alleles)** | Accuracy 99.7% (355/356 alleles correctly identified) | | | | |
| **Inferred phenotype concordance with AST (all antimicrobials)[d]** | 98.9 (98.7-99.1) | 98.9 (98.4-99.3) | 98.9 (98.7-99.1) | 96.1 (95.2-96.8) | 99.7 (99.6-99.8) |
| **Limit of detection** | Accuracy 99.9% at all coverage levels tested (40X – 150X) | | | | |
| **Precision (*n*=13 isolates)** | 100% repeatability and reproducibility (within- and between-run precision) | | | | |

'Detected' refers to AMR gene detected in reference dataset (PCR/allelic variant/synthetic reads) and also in *abritAMR* results from WGS.

'Not detected' refers to AMR genes detected in reference dataset, but not WGS.

CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; WGS, whole genome sequencing; AMR, antimicrobial resistance.

[a]Presence or absence of each of 415 AMR gene alleles was assessed across the 321 genomes, resulting in 133215 alleles for comparison across the dataset.

[b]Calculated by drug class (enhanced subclass from classification database). Range of sensitivity varied from 70-100% for each subclass (see Supplementary Table S1 for subclasses, and Figure 5 for performance across subclasses).

cCritical AMR subclasses – defined as classes containing AMR genes nationally-reportable to the

CARAlert program. Includes all carbapenemases, ESBL, ESBL (AmpC type), ribosomal

methyltransferases, colistin, oxazolidinone & phenicol resistance, vancomycin.

d Inferred phenotype validation for *Salmonella* spp., compared to agar dilution (CLSI methods and 2020

breakpoints). Detailed performance metrics available in Supplementary Table S4 and Figure 6.

**FIGURE LEGENDS**

**Figure 1. Overview of *abritAMR* pipeline**

Assembled short-read bacterial genomic sequence data are inputted in to the *abritAMR* pipeline to identify acquired AMR genes and mutations. The pipeline implements the *AMRFinderPlus* tool, identifying AMR genes (BLASTx search), and optionally identifies mutations associated with AMR for specified species where mutational data are available. Identified AMR genes +/- mutations are then binned into functional AMR classes according to the classification database (Detailed Report output, Figure 2 and Supplementary File 2). An additional step then applies reporting logic tailored to local requirements to produce per-isolate reports with acquired AMR genes (excluding common intrinsic AMR genes from being reported in specific species)(Final AMR Gene Report), and phenotypic inference for specified species, where validated (Inferred Antibiogram Report), Figure 2 and Supplementary File 2). AMR, antimicrobial resistance.

**Figure 2. Examples of *abritAMR* pipeline outputs**

This figure demonstrates *abritAMR* features and outputs across four different species (genome sequence A, *Escherichia coli;* genome sequence B, *Klebsiella pneumoniae;* genome sequence C, *Acinetobacter baumannii*; genome D, *Salmonella enterica).* The horizontal lanes represent the different output stages: top lane, raw *AMRFinderPlus* output; middle lane, *abritAMR* Detailed Report output (aligned to and binned by Enhanced Subclass from *abritAMR's* classification database); bottom lane, *abritAMR's* Final AMR Gene Report (genomes A-C) after application of tailored reporting logic to meet local requirements, and Inferred Antibiogram Report (genome D), currently validated for *Salmonella spp*. Key features include: (i) simplification of mechanism or drug class bins, (ii) identification and separation of high-priority AMR groups (such as notifiable AMR mechanisms) from lower-priority groups, (iii) identification of clinically-relevant AMR

mechanisms within a drug class (e.g. separation of ESBLs and AmpCs from genes encoding first-generation cephalosporin resistance; separation of metallo-beta-lactamase (MBL) carbapenemases and other carbapenemases due to differences in patient treatment), and (iv) application of tailored reporting logic to separate reportable and non-reportable genes according to local requirements, thus de-cluttering the report for clinicians and public health teams. Note, only 'Exact' matches (100% sequence identity and coverage) and 'Close' matches (90-<100% identity and 90-<100% sequence coverage) from the *AMRFinderPlus* tool are reported by *abritAMR.*

**Figure 3. Validation of *abritAMR* outputs compared to PCR and synthetic read data**

A. Validation compared to PCR data – Assembled short-read sequence data from bacterial isolates are run through the *abritAMR* pipeline and compared to multiplex real-time PCR results for the same AMR genes.

B. Synthetic data – where no validation dataset or PCR assay was available for comparison, synthetic read data were generated from publicly-available closed reference genomes by fragmentation with the *art-illumina* tool[23] (using the error profile from local sequencing platforms) to mimic library preparation from bacterial DNA. Synthetic reads then underwent the same analytical processes as for usual usage (genome assembly and input into *abritAMR* pipeline). These results were then compared to *AMRFinderPlus* results on the complete bacterial reference genomes, minimizing the risk of discordant results due to disparities between AMR databases.

AMR, antimicrobial resistance; PCR, polymerase chain reaction.

**Figure 4. Performance of *abritAMR* pipeline to detect AMR mechanisms compared to PCR**

Each panel details the identification of AMR mechanisms by *abritAMR* compared to the 'gold standard' multiplex PCR assays used in our laboratory. True positive, detected by both PCR and *abritAMR*; true negative, not detected by either PCR or *abritAMR*; false positive, detected by *abritAMR* but not multiplex PCR, *and* within the known range of the PCR assay; false negative, detected by multiplex PCR but not by *abritAMR*. Panel A, *mec* genes compared to multiplex PCR (*mecA/mecC*, no *mecC* detected by either method). Panel B, *van* genes compared to multiplex PCR (*vanA/B/C*). Panel C, detection of genes within carbapenemase and ESBL gene families compared to multiplex PCR panel (AusDiagnostics CRE panel), asterisks represent groups where discrepancies were identified, and expanded out in Panel D to show the specific gene discrepancies between the two methods.

Abbreviations: AMR, antimicrobial resistance; PCR, polymerase chain reaction; CP, carbapenemase; ESBL, extended-spectrum beta-lactamase.

**Figure 5. Performance of *abritAMR* pipeline compared to synthetic read data**

This figure shows the presence or absence of 415 AMR genes across 321 genomes identified by *abritAMR* (performed on assembled synthetic read data) compared to *AMRFinderPlus* on complete genome sequences ('gold standard'). True positive, detected by *AMRFinderPlus* and *abritAMR*; true negative, not detected by either *AMRFinderPlus* or *abritAMR;* false positive, detected by *abritAMR* but not by *AMRFinderPlus*; false negative, detected by *AMRFinderPlus* but not by *abritAMR.* Panel A, *abritAMR* results by enhanced subclasses (further grouped to simplify visualisation); asterisks represent classes with any discrepant result (false positive or false negative), and are examined in more detail in Panel B. Panel B shows a detailed view of genes

with discrepant results within each subclass. A full list of AMR subclasses included in the validation set can be found in Supplementary Table S1.

Abbreviations: EBL, extended-spectrum beta-lactamase; Rmt, ribosomal methyltransferase; Agly, aminoglycoside; Macro., Linco. & Strepto, macrolides, lincosamides and streptogramins (combined class); Carb., carbapenemase; Pen., penicillin resistance (*S. aureus*); Other BL, other beta-lactamase; Other discrep. Agly, other discrepant aminoglycoside subclass; Ami./Quin., amikacin/quinolone subclass; Quin., quinolones; Mac., Lin. & Strep., macrolides, lincosamides and streptogramins; Sulfon., sulfonamides; Rif., rifampicin; Tet., tetracyclines.

**Figure 6. Performance of inferred phenotype from *abritAMR* compared to antimicrobial susceptibility testing (AST)**

Classification of genotype (AMR mechanism detection) compared to the 'gold standard' phenotypic AST for each isolate and antimicrobial. True positive, genotypic and phenotypic resistance; true negative, no AMR mechanisms detected in phenotypically susceptible isolate; false positive, AMR mechanism identified in phenotypically susceptible isolate; false negative, no AMR mechanisms detected in phenotypically resistant isolate. For ciprofloxacin, 'true positive' defined as concordant intermediate or resistant results (phenotype and genotype). AST, antimicrobial susceptibility testing; AMR, antimicrobial resistance; Trim-sulfa, trimethoprim-sulfamethoxazole.
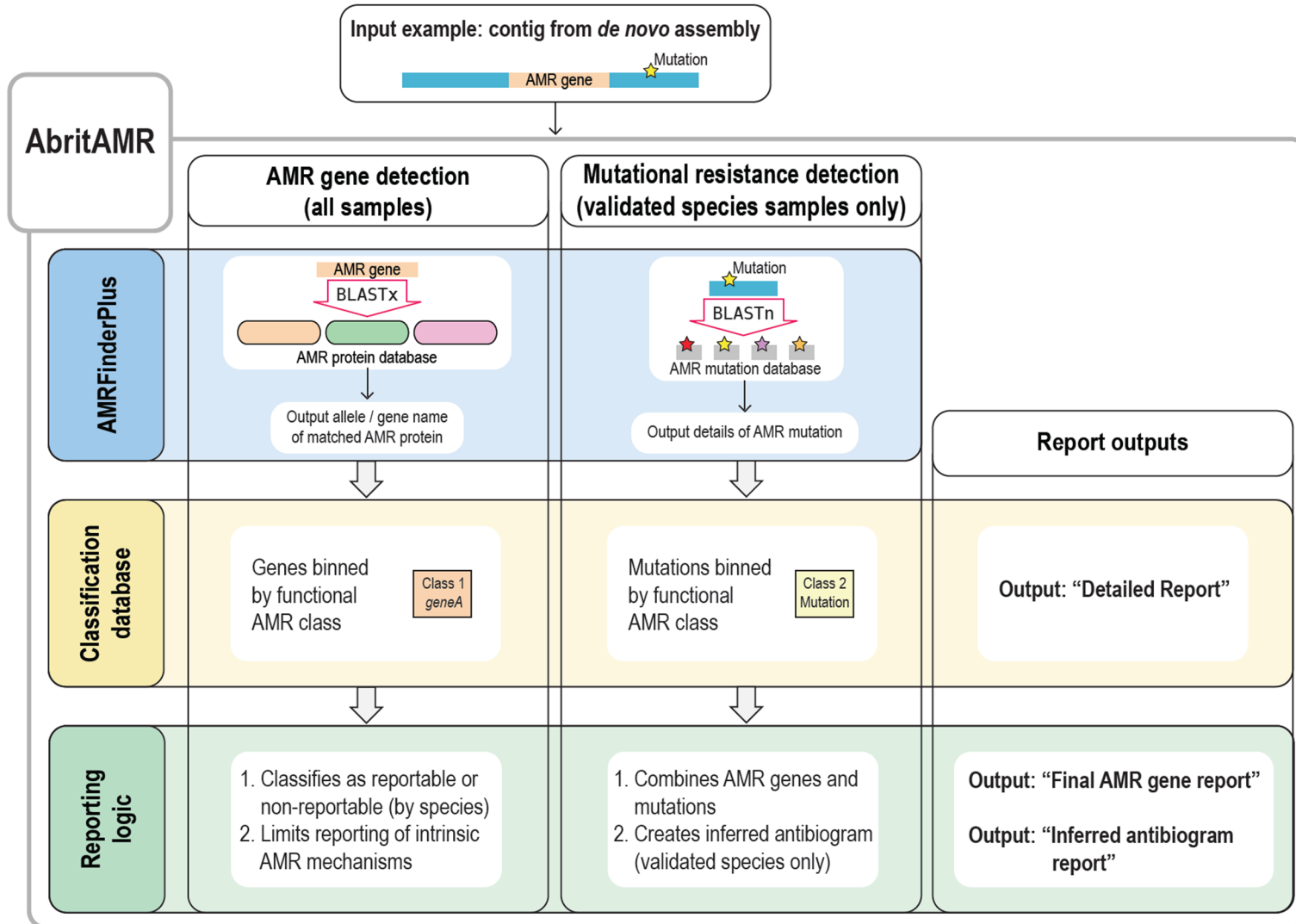
**Figure 1. Overview of *abritAMR* pipeline**

# Figure 2. Examples of *abritAMR* pipeline outputs



**AMR GENE DETECTION MODE**

**Genome sequence A: *Escherichia coli***

*AMR FinderPlus Output*

| | |
|---|---|
| *bla*NDM-1 | Subclass B1 metallo-beta-lactamase NDM-1 |
| *bla*OXA-181 | OXA-48 family carbapenem-hydrolyzing class D beta-lactamase OXA-181 |
| *rmtA* | RmtA family 16S rRNA (guanine(1405)-N(7))-methyltransferase |
| *bla*EC-8 | Cephalosporin-hydrolyzing class C beta-lactamase EC-8 |

**Genome sequence B: *Klebsiella pneumoniae***

| | |
|---|---|
| *bla*KPC-2 | Carbapenem-hydrolyzing class A beta-lactamase KPC-2 |
| *bla*SHV-12 | Class A extended-spectrum beta-lactamase SHV-12 |
| *bla*TEM-1 | Class A broad-spectrum beta-lactamase TEM-1 |
| *aac(6')-Ib-cr* | Fluoroquinolone-acetylating aminoglycoside 6'-N-acetyltransferase AAC(6')-Ib-cr3 |

**Genome sequence C: *Acinetobacter baumannii***

| | |
|---|---|
| *bla*OXA-23 | Carbapenem-hydrolyzing class D beta-lactamase OXA-23 |
| *bla*OXA-66 | OXA-51 family carbapenem-hydrolyzing class D beta-lactamase OXA-66 |
| *mcr*-1.1 | Phosphoethanolamine--lipid A transferase MCR-1.1 |
| *armA* | ArmA family 16S rRNA (guanine(1405)-N(7))-methyltransferase |

**INFERRED ANTIBIOGRAM MODE**

**Genome sequence D: *Salmonella enterica***

| | |
|---|---|
| *bla*CTX-M-15 | Class A extended-spectrum beta-lactamase CTX-M-15 |
| *gyrA*_A67P | DNA gyrase subunit A GyrA |
| *mph(A)* | Mph(A) family macrolide 2'-phosphotransferase |
| *dfrA12* | Trimethoprim-resistant dihydrofolate reductase DfrA12 |

**CLASSIFICATION DATABASE**

*abritAMR Detailed Report Output*

| | |
|---|---|
| *bla*NDM-1 | Carbapenemase (MBL) |
| *bla*OXA-181 | Carbapenemase |
| *rmtA* | Aminoglycosides (ribosomal methyltransferase) |
| *bla*EC-8 | ESBL (AmpC type) |

| | |
|---|---|
| *bla*KPC-2 | Carbapenemase |
| *bla*SHV-12 | ESBL |
| *bla*TEM-1 | Beta-lactam resistance (not ESBL or carbapenemase) |
| *aac(6')-Ib-cr* | Amikacin/Quinolone resistance |

| | |
|---|---|
| *bla*OXA-23 | Carbapenemase |
| *bla*OXA-66 | Carbapenemase (OXA-51 family) |
| *mcr*-1.1 | Colistin |
| *armA* | Aminoglycoside resistance (ribosomal methyltransferase) |

| | |
|---|---|
| *bla*CTX-M-15 | ESBL |
| *gyrA*_A67P | Quinolone |
| *mph(A)* | Macrolide |
| *dfrA12* | Trimethoprim |

**REPORTING LOGIC**

*abritAMR Final AMR Gene Report*

| | |
|---|---|
| *bla*NDM-1 | Carbapenemase (MBL) |
| *bla*OXA-181 | Carbapenemase |
| *rmtA* | Aminoglycoside resistance (ribosomal methyltransferase) |
| *bla*EC-8 | Not routinely reported (Intrinsic AmpC in *E. coli*) |

| | |
|---|---|
| *bla*KPC-2 | Carbapenemase |
| *bla*SHV-12 | ESBL |
| *bla*TEM-1 *aac(6')-Ib-cr* | Not routinely reported (not notifiable, low clinical significance) |

| | |
|---|---|
| *bla*OXA-23 | Carbapenemase |
| *mcr*-1.1 | Colistin |
| *armA* *bla*OXA-66 | Not routinely reported (Intrinsic in *A. baumannii*) |

*abritAMR Inferred Antibiogram Report*

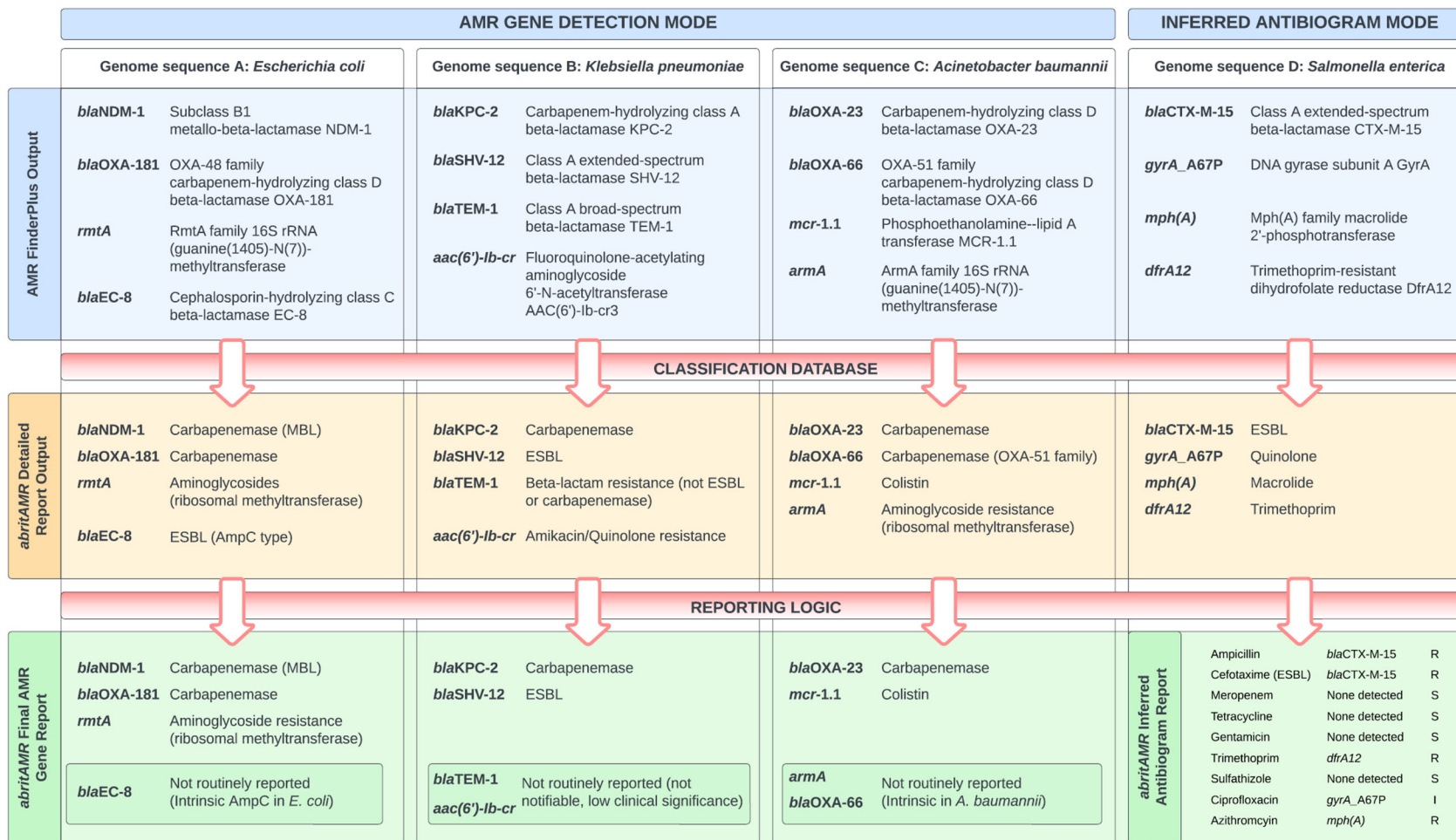| | | |
|---|---|---|
| Ampicillin | *bla*CTX-M-15 | R |
| Cefotaxime (ESBL) | *bla*CTX-M-15 | R |
| Meropenem | None detected | S |
| Tetracycline | None detected | S |
| Gentamicin | None detected | S |
| Trimethoprim | *dfrA12* | R |
| Sulfathizole | None detected | S |
| Ciprofloxacin | *gyrA*_A67P | I |
| Azithromcyin | *mph(A)* | R |

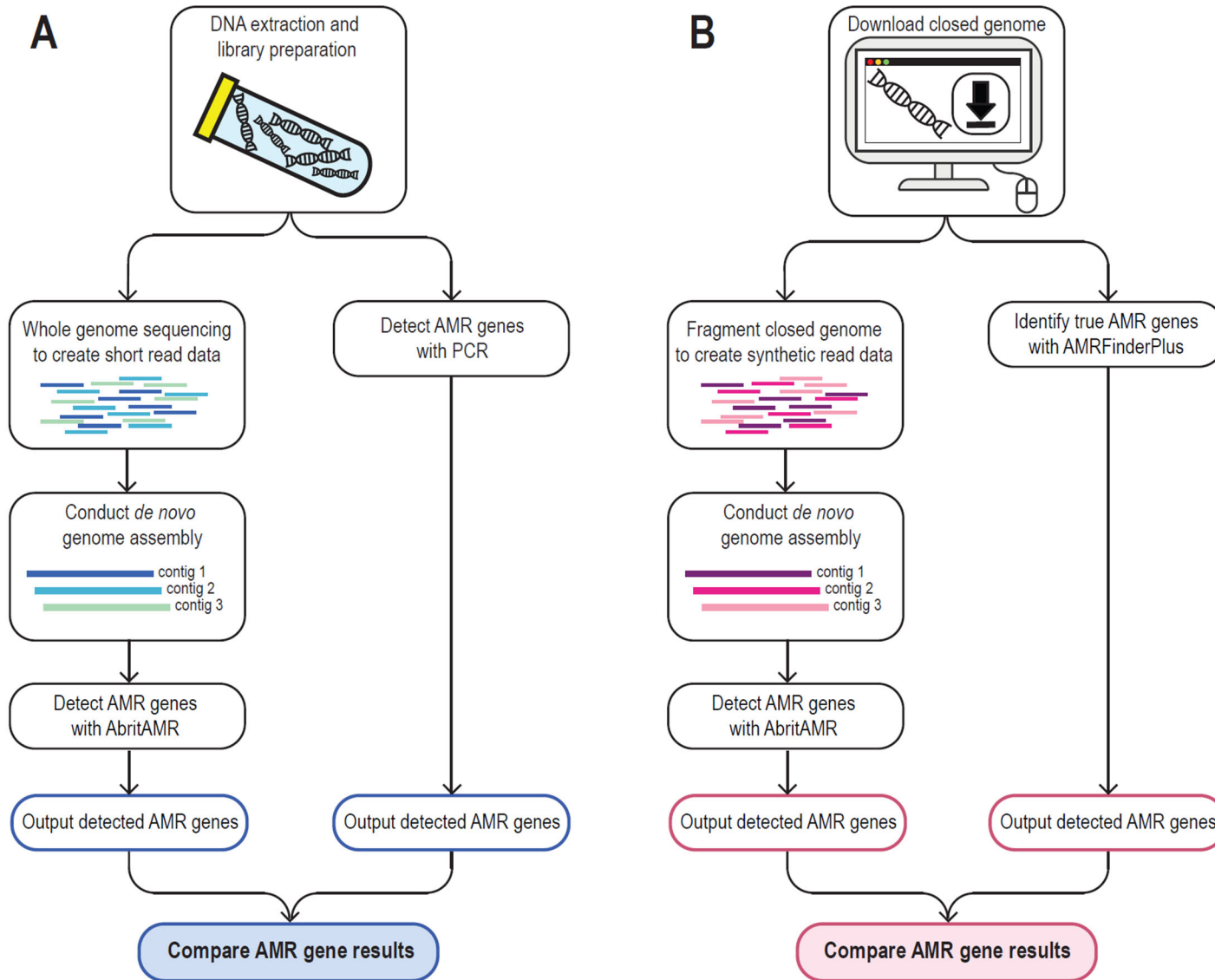**Figure 3. Validation of *abritAMR* outputs compared to PCR and synthetic read data**

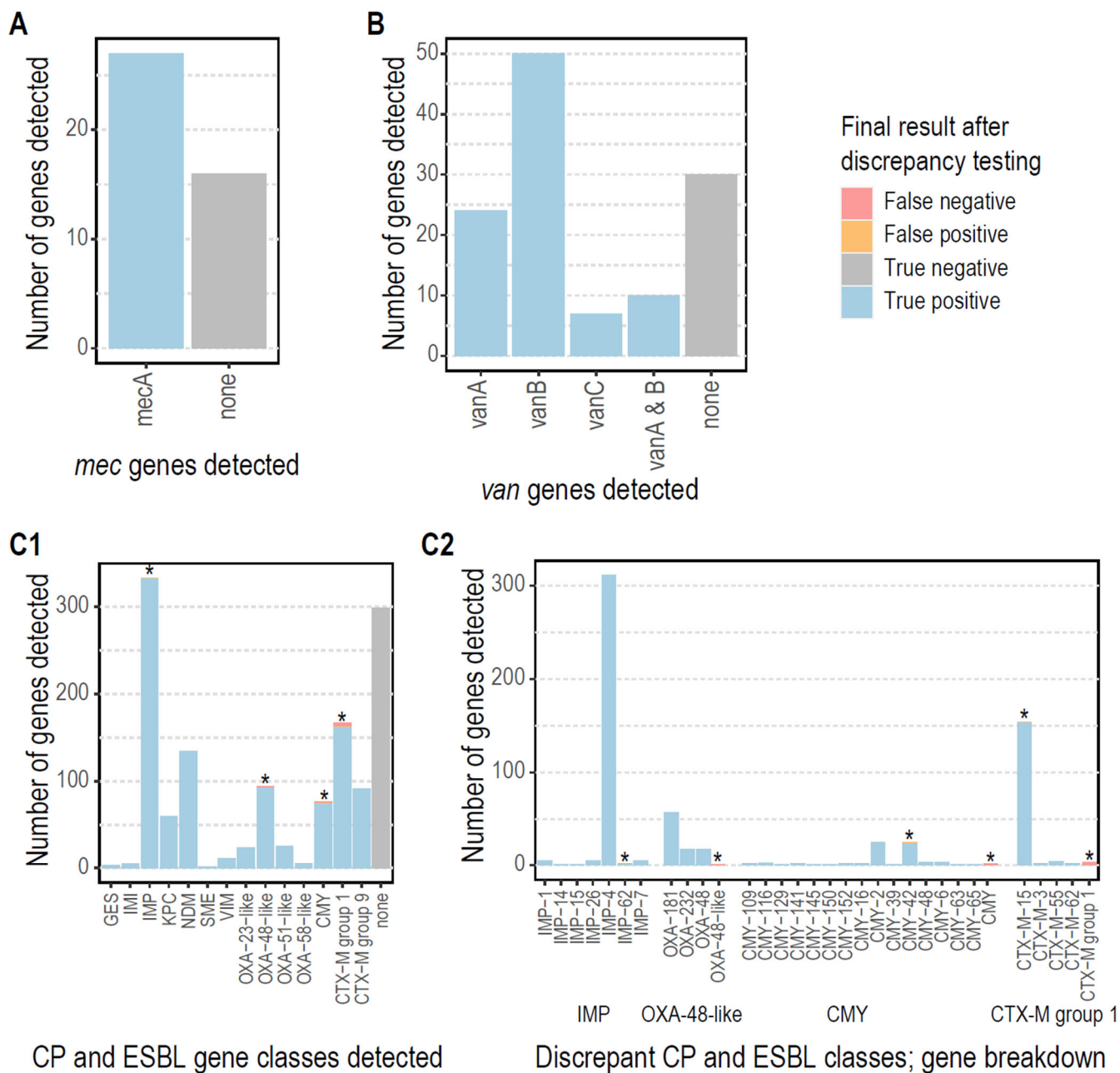**Figure 4. Performance of *abritAMR* pipeline to detect AMR mechanisms compared to PCR**

**Figure 5. Performance of *abritAMR* pipeline compared to synthetic read data**
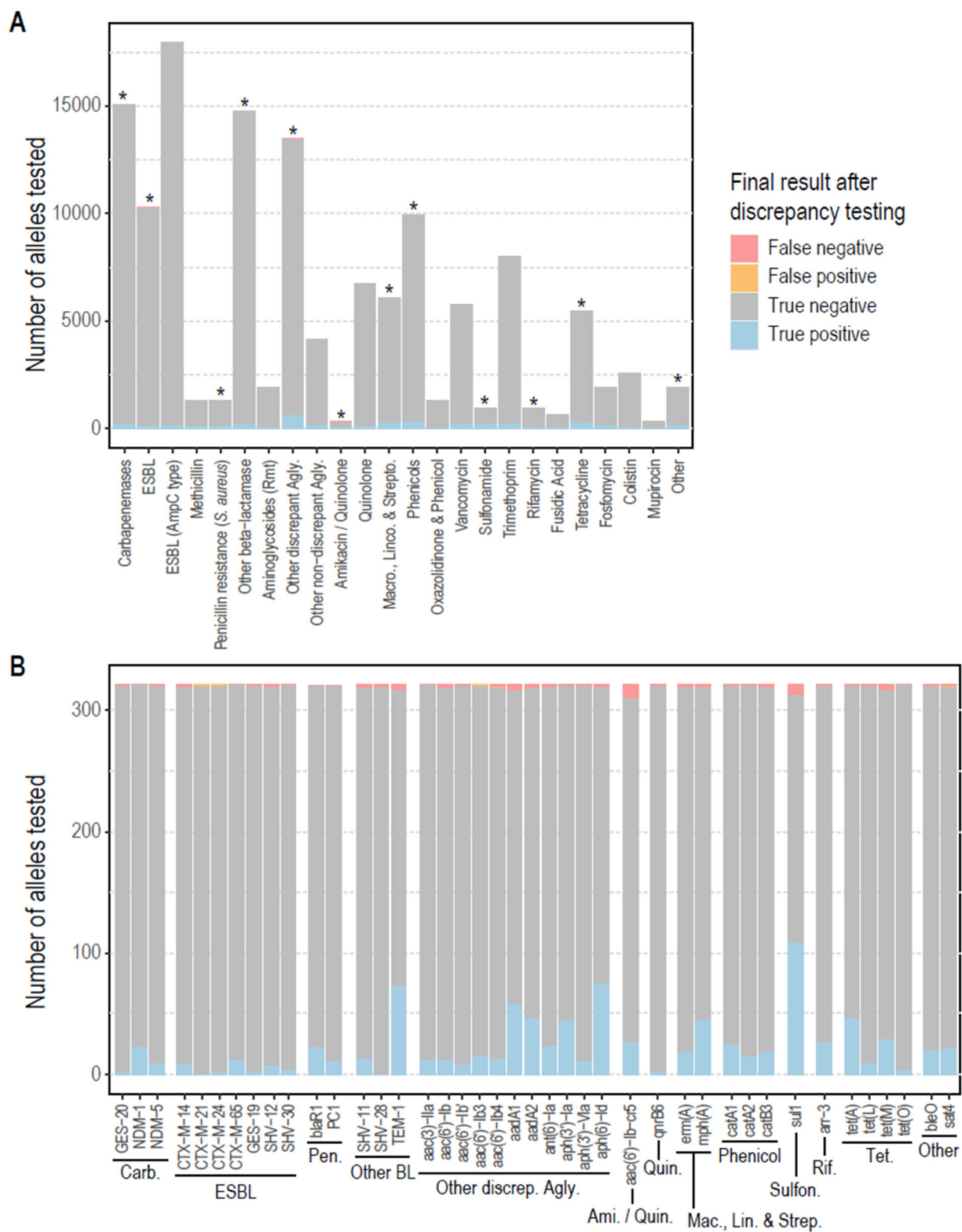
**Figure 6. Performance of inferred phenotype from *abritAMR* compared to antimicrobial susceptibility testing (AST)**