

The use of next-generation sequencing in fighting antimicrobial resistant pathogens

Edited by

Ibrahim Bitar, Costas C. Papagiannitsis and Alberto Antonelli

Published in

Frontiers in Microbiology



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83251-691-1
DOI 10.3389/978-2-83251-691-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The use of next-generation sequencing in fighting antimicrobial resistant pathogens

Topic editors

Ibrahim Bitar — Charles University, Czechia

Costas C. Papagiannitsis — University of Thessaly, Greece

Alberto Antonelli — University of Florence, Italy

Citation

Bitar, I., Papagiannitsis, C. C., Antonelli, A., eds. (2023). *The use of next-generation sequencing in fighting antimicrobial resistant pathogens*.

Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-691-1

Table of contents

- 05 **A Practical Approach for Predicting Antimicrobial Phenotype Resistance in *Staphylococcus aureus* Through Machine Learning Analysis of Genome Data**
Shuyi Wang, Chunjiang Zhao, Yuyao Yin, Fengning Chen, Hongbin Chen and Hui Wang
- 14 **Ultrafast and Cost-Effective Pathogen Identification and Resistance Gene Detection in a Clinical Setting Using Nanopore Flongle Sequencing**
Ekaterina Avershina, Stephan A. Frye, Jawad Ali, Arne M. Taxt and Rafi Ahmad
- 24 **Optimization of Early Antimicrobial Strategies for Lung Transplant Recipients Based on Metagenomic Next-Generation Sequencing**
Xiao-qin Zhang, Yu Lei, Xiao-li Tan, Lu Guo, Xiao-bo Huang, Fu-xun Yang, Hua Yu, Xiao-shu Liu, Yi-ping Wang, Sen Lu and Ling-ai Pan
- 33 **Successful Intra- but Not Inter-species Recombination of *msr(D)* in *Neisseria subflava***
Tessa de Block, Natalia González, Saïd Abdellati, Jolein Gyonne Elise Laumen, Christophe Van Dijck, Irith De Baetselier, Dorien Van den Bossche, Sheeba S. Manoharan-Basil and Chris Kenyon
- 42 **Co-expression Mechanism Analysis of Different Tachyplesin I-Resistant Strains in *Pseudomonas aeruginosa* Based on Transcriptome Sequencing**
Jun Hong, Xinyang Li, Mengyao Jiang and Ruofei Hong
- 56 **Prediction of Antibiotic Susceptibility Profiles of *Vibrio cholerae* Isolates From Whole Genome Illumina and Nanopore Sequencing Data: CholerAegon**
Valeria Fuesslin, Sebastian Krautwurst, Akash Srivastava, Doris Winter, Britta Liedigk, Thorsten Thye, Silvia Herrera-León, Shirlee Wohl, Jürgen May, Julius N. Fobil, Daniel Eibach, Manja Marz and Kathrin Schuldt
- 67 **CARB-ES-19 Multicenter Study of Carbapenemase-Producing *Klebsiella pneumoniae* and *Escherichia coli* From All Spanish Provinces Reveals Interregional Spread of High-Risk Clones Such as ST307/OXA-48 and ST512/KPC-3**
Javier E. Cañada-García, Zaira Moure, Pedro J. Sola-Campoy, Mercedes Delgado-Valverde, María E. Cano, Desirée Gijón, Mónica González, Irene Gracia-Ahufinger, Nieves Larrosa, Xavier Mulet, Cristina Pitart, Alba Rivera, Germán Bou, Jorge Calvo, Rafael Cantón, Juan José González-López, Luis Martínez-Martínez, Ferran Navarro, Antonio Oliver, Zaira R. Palacios-Baena, Álvaro Pascual, Guillermo Ruiz-Carrascoso, Jordi Vila, Belén Aracil, María Pérez-Vázquez, Jesús Oteo-Iglesias and the GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 Study Group

- 80 **Early Transcriptional Response to Monensin in Sensitive and Resistant Strains of *Eimeria tenella***
Hongtao Zhang, Lei Zhang, Hongbin Si, Xianrong Liu, Xun Suo and Dandan Hu
- 89 **Next-generation sequencing and PCR technologies in monitoring the hospital microbiome and its drug resistance**
Carolina Cason, Maria D'Accolti, Irene Soffritti, Sante Mazzacane, Manola Comar and Elisabetta Caselli
- 99 **Automated antimicrobial susceptibility testing and antimicrobial resistance genotyping using Illumina and Oxford Nanopore Technologies sequencing data among *Enterobacteriaceae***
Rick Conzemius, Yehudit Bergman, Peter Májek, Stephan Beisken, Shawna Lewis, Emily B. Jacobs, Pranita D. Tamma and Patricia J. Simner
- 109 **Resistance and virulence features of hypermucoviscous *Klebsiella pneumoniae* from bloodstream infections: Results of a nationwide Italian surveillance study**
Fabio Arena, Giulia Menchinelli, Vincenzo Di Pilato, Riccardo Torelli, Alberto Antonelli, Lucia Henrici De Angelis, Marco Coppi, Maurizio Sanguinetti and Gian Maria Rossolini for the *Klebsiella pneumoniae* Hypermucoviscous Italian Network (KHIN) of the Associazione Microbiologi Clinici Italiani (AMCLI)
- 121 **Clinical utility of target amplicon sequencing test for rapid diagnosis of drug-resistant *Mycobacterium tuberculosis* from respiratory specimens**
Kenneth Siu-Sing Leung, Kingsley King-Gee Tam, Timothy Ting-Leung Ng, Hiu-Yin Lao, Raymond Chiu-Man Shek, Oliver Chiu Kit Ma, Shi-Hui Yu, Jing-Xian Chen, Qi Han, Gilman Kit-Hang Siu and Wing-Cheong Yam
- 131 **The red thread between methylation and mutation in bacterial antibiotic resistance: How third-generation sequencing can help to unravel this relationship**
Stella Papaleo, Alessandro Alvaro, Riccardo Nodari, Simona Panelli, Ibrahim Bitar and Francesco Comandatore
- 145 **Molecular epidemiology and antimicrobial susceptibility of *Pseudomonas* spp. and *Acinetobacter* spp. from clinical samples at Jimma medical center, Ethiopia**
Tsegaye Sewunet, Daniel Asrat, Yimtubezinash Woldeamanuel, Abraham Aseffa and Christian G. Giske
- 158 **A resistome survey across hundreds of freshwater bacterial communities reveals the impacts of veterinary and human antibiotics use**
Susanne A. Kraemer, Naila Barbosa da Costa, Anais Oliva, Yannick Huot and David A. Walsh
- 171 **Transcriptome profiling in response to Kanamycin B reveals its wider non-antibiotic cellular function in *Escherichia coli***
Yaowen Chang, Xuhui Zhang, Alastair I. H. Murchie and Dongrong Chen



A Practical Approach for Predicting Antimicrobial Phenotype Resistance in *Staphylococcus aureus* Through Machine Learning Analysis of Genome Data

Shuyi Wang^{1,2†}, Chunjiang Zhao^{2†}, Yuyao Yin², Fengning Chen^{1,2}, Hongbin Chen² and Hui Wang^{1,2*}

¹ Institute of Medical Technology, Peking University Health Science Center, Beijing, China, ² Department of Clinical Laboratory, Peking University People's Hospital, Beijing, China

OPEN ACCESS

Edited by:

Ibrahim Bitar,
Charles University, Czechia

Reviewed by:

Vittoria Mattioni Marchetti,
Charles University, Czechia
Balig Panossian,
Lebanese American University,
Lebanon

*Correspondence:

Hui Wang
whuibj@163.com;
wanghui@pkuph.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 22 December 2021

Accepted: 11 February 2022

Published: 02 March 2022

Citation:

Wang S, Zhao C, Yin Y, Chen F,
Chen H and Wang H (2022) A
Practical Approach for Predicting
Antimicrobial Phenotype Resistance
in *Staphylococcus aureus* Through
Machine Learning Analysis
of Genome Data.
Front. Microbiol. 13:841289.
doi: 10.3389/fmicb.2022.841289

With the reduction in sequencing price and acceleration of sequencing speed, it is particularly important to directly link the genotype and phenotype of bacteria. Here, we firstly predicted the minimum inhibitory concentrations of ten antimicrobial agents for *Staphylococcus aureus* using 466 isolates by directly extracting k-mer from whole genome sequencing data combined with three machine learning algorithms: random forest, support vector machine, and XGBoost. Considering one two-fold dilution, the essential agreement and the category agreement could reach >85% and >90% for most antimicrobial agents. For clindamycin, cefoxitin and trimethoprim-sulfamethoxazole, the essential agreement and the category agreement could reach >91% and >93%, providing important information for clinical treatment. The successful prediction of cefoxitin resistance showed that the model could identify methicillin-resistant *S. aureus*. The results suggest that small datasets available in large hospitals could bypass the existing basic research and known antimicrobial resistance genes and accurately predict the bacterial phenotype.

Keywords: *Staphylococcus aureus*, k-mer algorithm, antimicrobial resistance (AMR), machine learning, WGS

INTRODUCTION

Traditional microbial identification and antimicrobial susceptibility rely on microbial culture technology, which is a time-consuming process. Owing to the optimization of the cultivation technology for more than a hundred years, the cultivation time has been shortened to within 24–48 h, depending on the specific strain. However, traditional culture techniques cannot meet the ever-increasing demands for rapid diagnosis. Although mass spectrometry technology has now been widely used, and it is possible to quickly identify pathogenic bacteria after obtaining purely cultured isolates, antimicrobial susceptibility results are still not widely available.

In recent years, gene sequencing technology has developed rapidly. The price of whole genome sequencing has fallen below 26.3 dollars a isolate recently, and the time spent on data yield can be archived within 18 h. Metagenomics technology can identify the types of pathogens in samples,

including those which cannot be identified by traditional methods, within 8 h; thus expanding the boundaries of clinical microbiology and filling the gap between the genotype and phenotype of bacteria (Chen et al., 2020a,b).

After obtaining genome data, a relatively rough prediction of resistance can be achieved by studying the antimicrobial resistance genes in the genome. Nowadays, many studies have focused on antimicrobial resistance prediction; some studies only predicted whether isolates were resistant or susceptible, and did not predict the specific minimum inhibitory concentration (MIC) values (Brinda et al., 2020; Macesic et al., 2020; Avershina et al., 2021), while others predicted resistance and susceptibility based on the presence or absence of known resistance genes or single nucleotide polymorphisms. However, the updation of pre-existing basic research is quite slow (Satta et al., 2018; Khaleedi et al., 2020; Kim et al., 2020; Avershina et al., 2021). The k-mer algorithm can count the size of genomic repeats and the degree of genomic heterozygosity. Currently, training set data used by k-mer based research is relatively large, while longer k-mers are better at exhibiting specificity of genomic features. This poses serious challenges to data acquisition, storage, and processing, thereby limiting the promotion and application of this technology. According to Moore's Law (Moore, 1965), the speed of technological process double about every 2 years, revealing the speed of technological progress. It is foreseeable that these limitations will be overcome quickly in the future, and that follow-up developments will come rapidly, increasing the speed of the genome identification and analysis.

Methicillin-resistant *Staphylococcus aureus* (MRSA) is one of the most serious multi-antimicrobial-resistant threats (Tacconelli et al., 2018) and is the leading cause of many systemic infections (Lowy, 1998). In this study, we created a novel *Staphylococcus aureus* resistance prediction model that could accurately predict MICs of antimicrobial agents using a relatively small number of training isolates based on antimicrobial-resistant phenotypes and k-mer calculation, combined with machine learning algorithms.

MATERIALS AND METHODS

Isolates

We prospectively investigated the pathogen spectrum of bloodstream infections, hospital-acquired pneumonia, and intra-abdominal infections. We collected 466 *S. aureus* [249 MRSA and 217 methicillin-susceptible *Staphylococcus aureus* (MSSA)], belonging to 23 different sequence types (STs), from 14 states across China between 2005 and 2020.

Whole Genome Sequencing

The *S. aureus* isolates were sequenced using the Illumina NextSeq 500, NovaSeq, HiSeq Xten, and Illumina HiSeq platforms. Multilocus sequence typing (MLST) was performed according to the PubMLST scheme.¹ We obtained clean DNA sequences after using the fastp program² (Lee et al., 2018) to optimize raw FASTQ

data quality and clean the raw data obtained after whole genome sequencing. Staphylococcal cassette chromosome mec (SCCmec) was assigned using SCCmecFinder-1.2 (De Oliveira et al., 2020)³ (Supplementary Table 1).

Antibiotic Susceptibility Testing

Antibiotic susceptibility testing was performed on all the *S. aureus* isolates. The following antimicrobial agents were evaluated in this study: clindamycin (CLI), ceftiofur (FOX), oxacillin (OXA), levofloxacin (LVX), trimethoprim-sulfamethoxazole (SXT), vancomycin (VAN), linezolid (LNZ), erythromycin (ERY), daptomycin (DAP), and gentamicin (GEN). The MIC of commonly used antimicrobials was determined using the agar dilution method according to the protocol and the susceptibility spot of the Clinical and Laboratory Standards Institute guidelines (M100-S31,2021) and susceptible (S), intermediate (I) and resistant (R) categories were adjudicated.

Data Set Preparation

For all isolates, the KMC application (Deorowicz et al., 2015)⁴ was used to extract features from cleaned DNA sequences. We used pandas (version 1.0.5) to combine and resize the obtained k-mer data, which were then combined into the matrix, with the rows indicating strain names and columns indicating k-mer counts. This matrix was used as the input data for machine learning ($k = 11$). The larger the value of k , the better the features obtained from the genomic information. In this study, $k = 11$ was used because of the limitations of computing memory. We used LabelBinarizer (scikit-learn, version 0.23.1) to convert the MIC labels to one-hot codes, which facilitated subsequent training of the model.

A total of 466 isolates with their antimicrobial susceptibility data were randomly divided into training sets (372 isolates) and testing sets (94 isolates). For GEN, ERY, and DAP, the data set was extracted separately because of the lack of antimicrobial susceptibility data. Among the 466 isolates, 69 isolates had GEN antimicrobial susceptibility data, 431 isolates had DAP antimicrobial susceptibility data, 454 isolates had ERY antimicrobial susceptibility data, and 20% were selected as the testing set. In order to ensure that all MICs of all antimicrobial agents could be trained as a category in the training set, we used the StratifiedShuffleSplit software (scikit-learn, version 0.23.1) to obtain stratified randomized folds. The folds were made by preserving the percentage of samples for each MIC grade. MIC grades that had only one isolate were incorporated into a lower grade for optimal classification.

Machine Learning Analysis

The support vector machine (SVM) and random forest, used for training in Python 3.6, were based on scikit-learn (version 0.23.1) (Pedregosa et al., 2011). XGBoost is a machine learning algorithm based on the gradient boosting framework. We used the XGBoost (Chen and Guestrin, 2016) sklearn interface for training. We trained random forests with 600 trees. For SVM, we

¹<https://pubmlst.org/saureus/>

²<https://github.com/OpenGene/fastp>

³<https://cge.cbs.dtu.dk/services/SCCmecFinder-1.2/>

⁴<https://github.com/lh3/kmer-cnt>

tried three kernel functions: linear, poly, and rbf. For XGBoost, we tried binary: logistic learning task parameters. For all training, we selected the best result of the 10-fold cross-validation as the final result. All source code is available at https://github.com/ShuyiWang-pku/sau_micprediction.

Interpretation of Results

For the final classification, the results were evaluated according to the Clinical and Laboratory Standards Institute guidelines (M100-S31,2021). We assumed that for the categories obtained by the classifier, the classification results of one two-fold dilution, also termed as essential agreement (EA), were correct. According to this standard, the receiver operating characteristic (ROC) curves, area under curve (AUC) values, EA, category agreement (CA), sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), very major error (VME), and major error (ME) were calculated to evaluate our model. ROC curves and AUC values can judge the predictive performance of models. Taking label imbalance into account, we calculated metrics globally by considering each element of the label indicator matrix as a label. We aggregated outcomes across all classes, drew the ROC curves and calculated the AUC values (Fawcett, 2006). CA refers to the accuracy of the prediction model that only considers the classification of susceptible, intermediate, and resistant categories of antimicrobial agents. Sensitivity refers to the number of predicted resistant and intermediate results divided by the actual number of resistant and intermediate samples, and specificity refers to the number of predicted susceptible results divided by the actual number of susceptible samples. NPV refers to the number of true negatives (a “susceptible” prediction for which the actual event had a “susceptible” result), divided by the total number of isolates predicted to be “susceptible” (sum of true negatives and false negatives). PPV refers to the number of true positives (a “resistant” or “intermediate” prediction, for which the actual result was “resistant” or “intermediate”), divided by the total number of isolates predicted to be “resistant” or “intermediate.” VME refers to the number of isolates predicted to be “susceptible” while the actual result was “intermediate” or “resistant,” divided by the actual number of “resistant” and “intermediate” results. ME refers to the number of isolates predicted to be “resistant” or “intermediate,” while their actual result was “susceptible” divided by the actual number of “susceptible” results.

RESULTS

Basic Characteristics of the *Staphylococcus aureus* Isolates

We analyzed 466 isolates of *S. aureus*, including 249 MRSA and 217 MSSA, from 14 states in China. The DNA sequence of all isolates was obtained using whole genome sequencing. Among the 466 *S. aureus* isolates, 23 STs were identified, among which ST59, ST239, and ST398 were higher in number (Figure 1). The ST of 2 isolates could not be identified; however, this could not be attributed to the quality of sequencing, as evidenced by the high quality sequencing data shown in the **Supplementary Material**.

The sample collection details of all the isolates are presented in **Table 1**.

The MICs of the isolates are shown in **Table 2**, where light green refers to susceptible isolates defined according to clinical breakpoints, light orange refers to intermediate isolates, and colorless refers to resistant isolates. As seen from the results, VAN, LNZ, and DAP isolates were all susceptible. Therefore, we analyzed the classification performance of other models of antimicrobial agents and the factors that may affect the classification results.

Extraction of the Characteristics of the Isolates and Rapid Identification of the *Staphylococcus aureus* Isolates by k-mer

The workflow of this study is shown in **Figure 2**. After collecting and culturing the samples, we performed whole genome sequencing and antibiotic susceptibility testing for all isolates to obtain the genomic and antimicrobial susceptibility data, respectively. For our data, the k-mer algorithm was used to calculate the k-mer characteristics from the genome of each strain. The k-mer statistics ($k = 11$) of all isolates were combined into a matrix, which was used as the input data for machine learning. We used the following machine learning algorithms: SVM, random forest, and XGBoost (Chen and Guestrin, 2016) to train our data, and selected the best model to test the testing set, considered as the output results in this study. The testing time was consistently less than 1 min. The specific classification results for testing MIC and susceptible (S), intermediate (I), and resistant (R) categories are shown in **Supplementary Tables 2–21**, and the AUC, EA, CA, sensitivity, specificity, NPV, PPV, VME, and ME results are shown in **Figure 3** and **Table 3**. To evaluate the reliability of the model, the ROC curves of five models with the best results in cross-validation were obtained (**Supplementary Figures 1–10**), and the average and standard deviations of the cross-validation results for all metrics are shown in **Supplementary Tables 22, 23**.

The STs of the isolates in this study were different. In addition, we did not consider the evolution of the strain, unlike in a previous study (Brinda et al., 2020). Nonetheless, it was observed that the EAs and CAs of all antimicrobial agents were almost more than 80%. Except for OXA, LVX, and ERY, the EAs of all antimicrobial agents reached over 85%, while that for CLI and SXT reached over 92%. Except for OXA, the CAs for all antimicrobial agents reached over 90% (**Figure 3**). Furthermore, except for OXA, the sensitivity of all antimicrobial agents was over 90%, and the specificity of other antimicrobial agents was all over 85%. The sensitivity of LVX, SXT, and GEN reached 100%, and the specificity of CLI and SXT reached 97%. The VMEs of LVX, SXT, and GEN were as low as 0% in this study (**Table 3**).

Effect of Data Volume and Structure on Prediction Results

For ERY and GEN, which had poor classification results, the specific data showed that the antimicrobial susceptibility data of these two antimicrobial agents were too biased toward

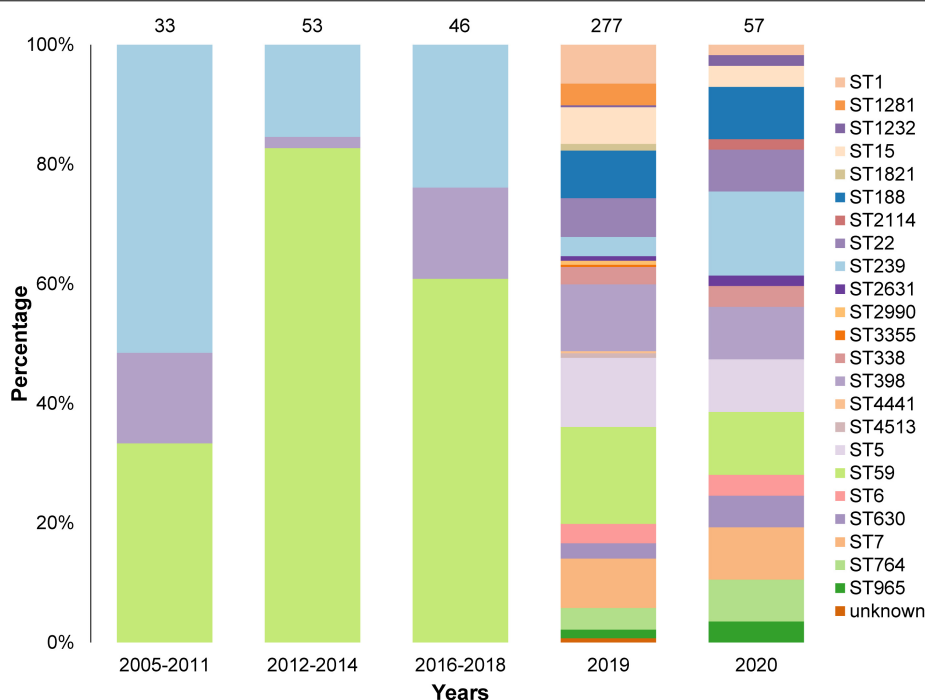


FIGURE 1 | Sequence Types (STs) and collection years of the 466 isolates in this study. A total of 23 STs were identified and the STs of 2 isolates were unknown. Most isolates were collected in 2019 and 2020.

the extreme value and the number of genomes in GEN was small, with only 69 isolates. When the MIC of GEN was 0.25 $\mu\text{g/mL}$, the number of isolates was 34, while the number of other MICs is relatively few, most of them are less than 5 (Table 1). When the MIC of ERY was 0.25 and 512 $\mu\text{g/mL}$, the number of isolates was 112 and 237, respectively, which was approximately 100–200 fold higher than that of some other MIC grades (Table 1).

TABLE 1 | States and collection years of the 466 isolates in this study.

Years States	2005–2011	2012–2014	2016–2018	2019	2020
Beijing	15	2	10	79	1
Chongqing	1	1	4	20	1
Guangdong	4	3	5	28	5
Hubei	2	-	4	19	-
Hunan	1	1	3	9	7
Jiangsu	1	2	3	22	-
Liaoning	1	1	-	-	-
Shaanxi	1	1	4	-	-
Shandong	1	38	1	20	6
Shanghai	4	1	2	5	6
Shanxi	-	-	3	-	25
Shenyang	-	2	2	-	-
Tianjin	-	-	2	26	-
Zhejiang	2	1	3	49	6

The 466 isolates were widely distributed across 14 states in China.

We speculated that the data structure and data volume limit the improvement in accuracy. Owing to the difference in the number of samples, the classifier did not fully learn MIC features with a small number of isolates, which led to a vague distinction between the features of different categories and reduced the accuracy of model classification and recognition.

In most hospitals, data related to isolates is usually unevenly distributed, as observed in this study (Table 1). In addition to the fact that some large national institutions can use large hospital networks to obtain the data and avoid the condition where there only 1 or 2 isolates in some MIC, the distribution of antimicrobial susceptibility in most hospitals is not uniform, and the extremes of distribution phenomena are serious. This study proved that even if the antimicrobial susceptibility data of isolates are not evenly distributed, the developed model could be used to perform a certain rapid antimicrobial susceptibility analysis in various hospitals with a small amount of uneven data, and is suitable for most general hospitals that have accumulated a certain number of isolates with the popularity of sequencing technology. In addition, with an increase in the number of isolates collected in the future, the accuracy of MIC and classification may be further improved.

Reasons for the Varied Prediction Results of Cefoxitin and Oxacillin

Methicillin, FOX, and OXA are β -lactam antimicrobials. In general, MRSA is resistant to OXA and FOX. This is because most MRSA isolates acquire the staphylococcal cassette chromosome *mec* (SCC*mec*) (Lakhundi and Zhang, 2018). However, in recent

TABLE 2 | Number of genomes with different minimum inhibitory concentration (MIC) to the 10 antimicrobial agents for the 466 *Staphylococcus aureus* isolates used in this study.

MICs ($\mu\text{g/mL}$) Antimicrobial agents	0.032	0.064	0.125	0.25	0.5	1	2	4	8	16	32	64	128	256	512	Total
Clindamycin	46	163	32	3	5	1			3	1	2	2	208			466
Cefoxitin							46	171	31	54	63	7	94			466
Oxacillin					217	40	39	39	15	13	5	5	93			466
Levofloxacin			51	207	45	26	5	9	33	90						466
Trimethoprim-Sulfamethoxazole	102	257	43	21	8	10	9	2	2	12						466
Vancomycin					61	396	9									466
Linezolid					22	278	166									466
Erythromycin		4	33	112	6			4	9	6	9	12	12	10	237	454
Daptomycin			13	175	221	22										431
Gentamicin			8	34	3	5	1		1	4	2	1	5	3	2	69

In this table, light green refers to susceptible isolates defined according to clinical breakpoints, light orange refers to intermediate isolates, and colorless refers to resistant isolates.

years, many studies have shown that even when carrying *SCCmec*, isolates are susceptible to OXA (termed as OS-MRSA). Previous studies have speculated that the development of OS-MRSA results from different *SCCmec* types, but there is no definitive conclusion (Boonsiri et al., 2020). Among the 466 isolates, 79 were resistant to FOX but susceptible to OXA. Among these 79 isolates, 56 were ST59, accounting for 70.89% of all OS-MRSA isolates (Figure 4). Among the ST59 isolates, OS-MRSA accounted for 46.67% of MRSA isolates (OS-MRSA/MRSA), which was almost half of all the ST59 MRSA isolates (Supplementary Table 24).

Our prediction model could directly capture the features of *SCCmec* from the FASTQ data and estimate the MICs of FOX, in which EA could reach 91.49% and CA could reach 93.62%. This result showed that the model could accurately and quickly distinguish whether the isolate was MRSA or MSSA. However, the prediction accuracy of OXA was relatively low, and CA was lower than that of EA because we considered one two-fold dilution for the calculation, while the MIC breakpoints of OXA were 2 and 4 $\mu\text{g/mL}$ (Supplementary Table 24). The reason for the high accuracy of FOX may be that there is basically no “OS-MRSA” in FOX; that is, *SCCmec* can be detected; however, the isolate is susceptible to FOX. When the k-mer learns the genes of antimicrobial resistance, it is easy to distinguish them, but the isolates of ST59, which have *SCCmec* genes, were susceptible or resistant to OXA with a 50:50 proportion and have a close genetic relationship. Therefore, even if the classifier learns *SCCmec* genes, there will not be any clear conclusion, and the isolates may become OS-MRSA in different ways. We believe that the limited number of OS-MRSA interfered with the results, and our prediction model did not fully understand the characteristics of OS-MRSA, which resulted in varied prediction accuracies of OXA and FOX.

Effect of Other Resistance Mechanism on Prediction Results

Previous studies have shown that ST59 is mostly susceptible to LVX and ST239 is mostly resistant to LVX (Li et al., 2018). Similar results were obtained in this study. Among the 133 ST59 isolates,

only 11 isolates were intermediate or resistant to LVX, and the rest were all susceptible to LVX. Of the 54 ST239 isolates, only two were susceptible to LVX, and the rest were all intermediate or resistant. In contrast, ST59 and ST239 accounted for the majority of the total data set, and the two STs were very far apart at the genome level, and hence the CA of LVX was high. The low EA of LVX may be due to the small differences within susceptible isolates and resistant isolates, which our prediction model did not recognize.

Timeliness of Prediction Using Genome Data

Antimicrobial resistance prediction was 6 h faster when using whole genome sequencing combined with k-mer detection and machine learning algorithms than the duration of routine clinical testing (Figure 5). The success of predicting MICs with pure bacterial culture with a limited amount of data, indicates that it is feasible to obtain susceptibility results directly from genomic data without requiring prior information.

DISCUSSION

In this study, the resistance characteristics of *S. aureus* to major antimicrobial agents were successfully predicted using k-mers and machine learning techniques. Among them, the CAs of all antimicrobial agents were generally high, and >90% pairs could provide important information for clinical treatment. Moreover, the EAs of most antimicrobial agents were higher than 85%.

Previous studies have predicted antimicrobial resistance phenotypes using antimicrobial-resistant genes or conserved genes in gene sequence of isolates (Kim et al., 2020; Macesic et al., 2020; Nguyen et al., 2020; Avershina et al., 2021), these methods of inferring antimicrobial resistance phenotypes from known resistance genes are highly accurate. At the same time, there are many public databases like PATRIC (Vanoeffelen et al., 2021), through which researchers can collect genomic data. The availability of such databases reduces the imbalance of data caused by regional differences by enabling the tracking of antimicrobial resistance evolution or prediction of antimicrobial

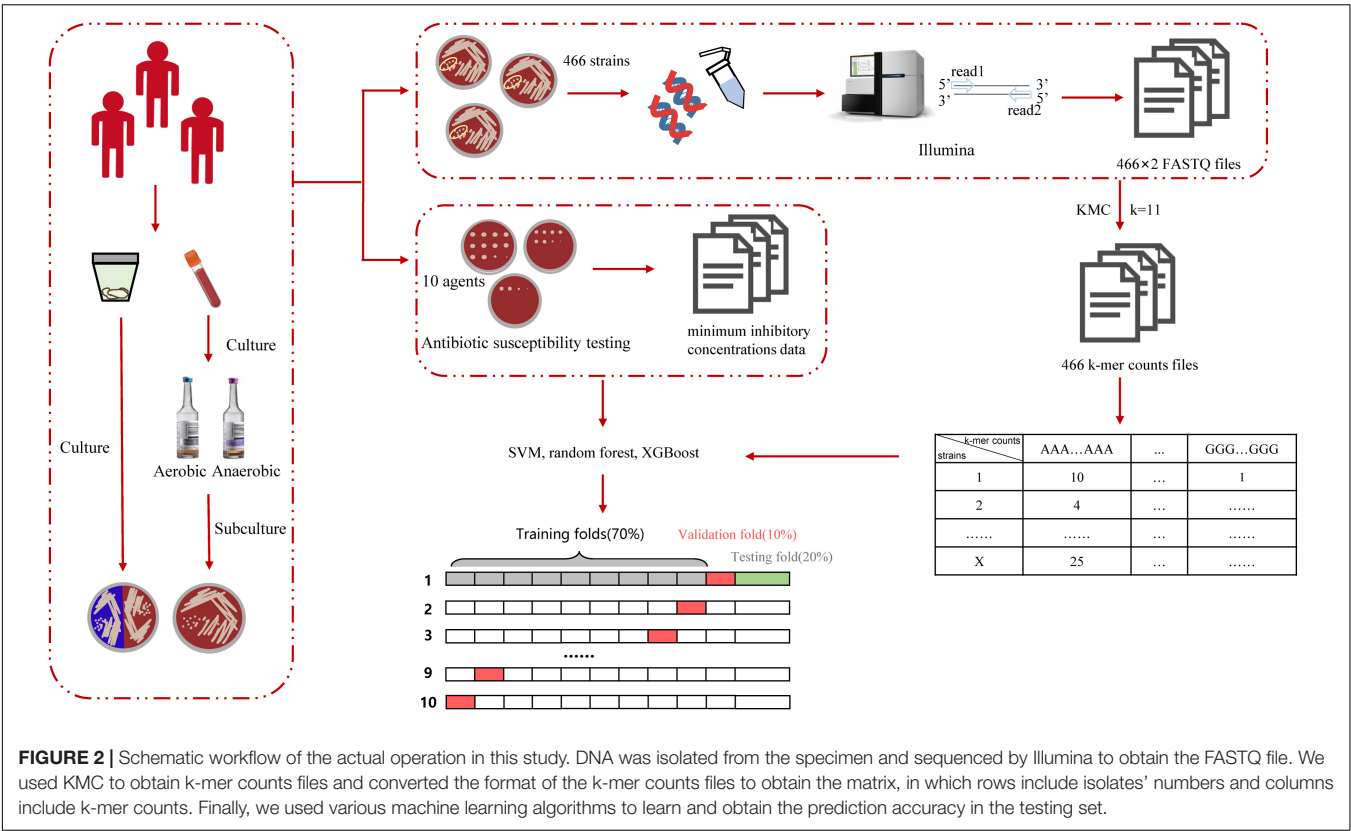


FIGURE 2 | Schematic workflow of the actual operation in this study. DNA was isolated from the specimen and sequenced by Illumina to obtain the FASTQ file. We used KMC to obtain k-mer counts files and converted the format of the k-mer counts files to obtain the matrix, in which rows include isolates' numbers and columns include k-mer counts. Finally, we used various machine learning algorithms to learn and obtain the prediction accuracy in the testing set.

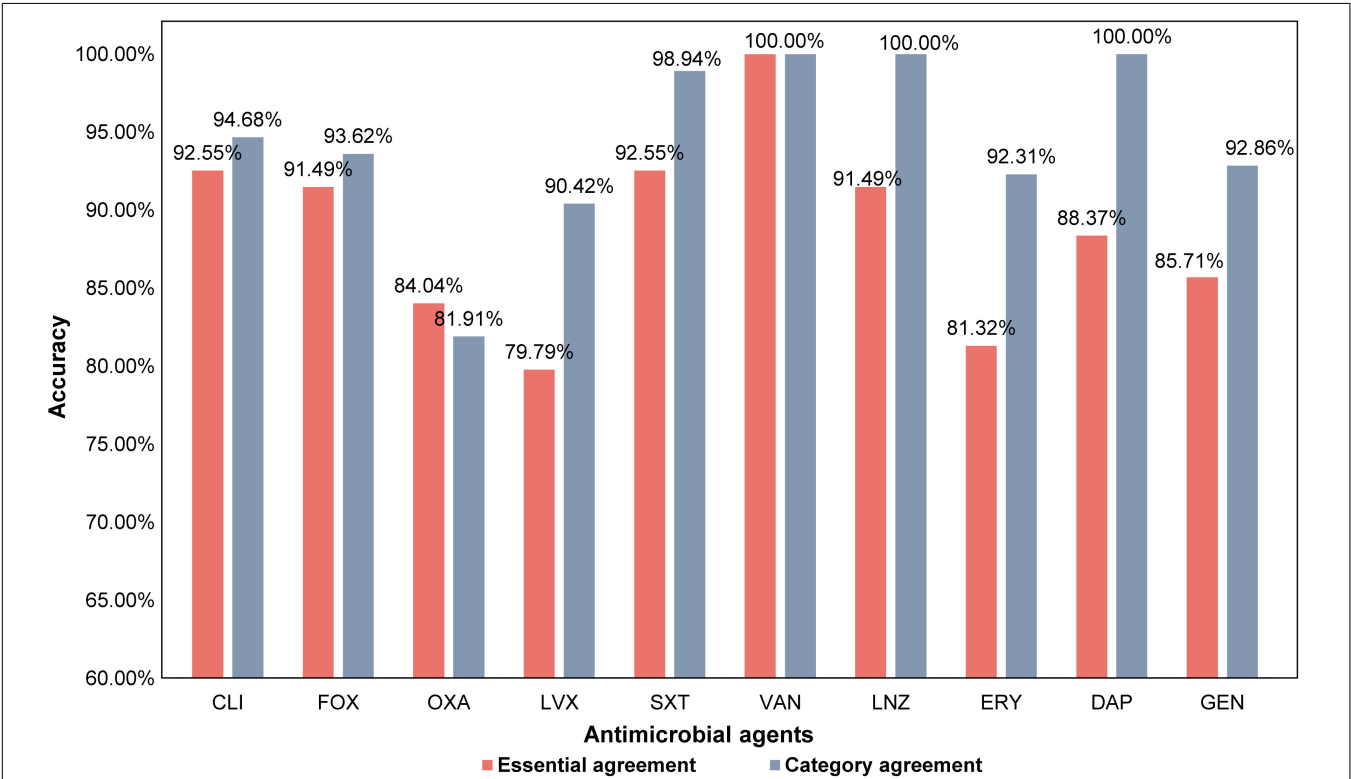
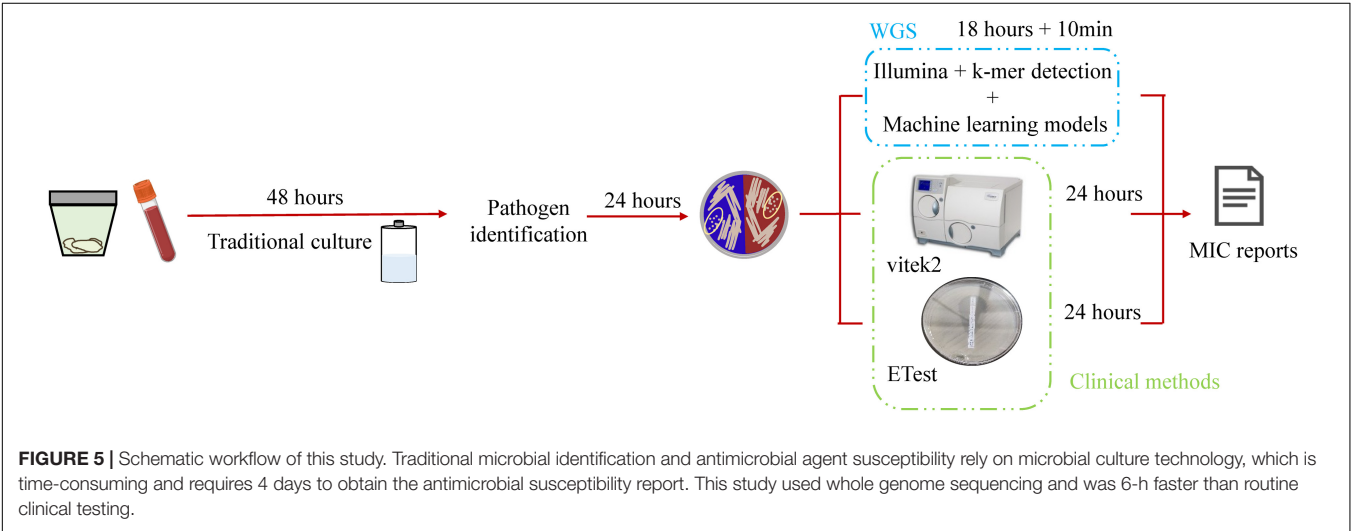
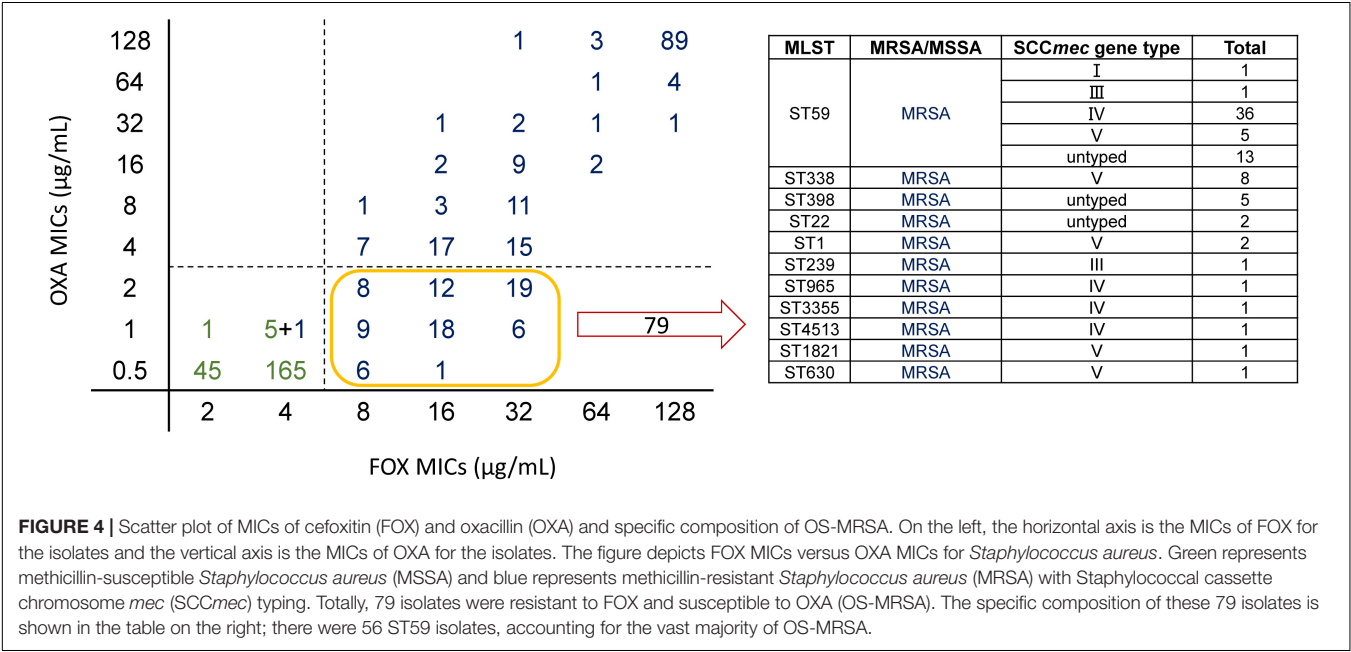


FIGURE 3 | The prediction accuracy within the one two-fold dilution, essential agreement (EA), and the category agreement (CA) of all antimicrobial agents.

TABLE 3 | The AUC (Area Under Curve), sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), very major error (VME), and major error (ME) of the prediction results of all antimicrobial agents evaluated in this study.

Antimicrobial agents	AUC (%)	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	VME (%)	ME (%)
Clindamycin	94.61	91.30	97.92	92.16	97.67	8.70	2.08
Cefoxitin	92.65	94.00	93.18	93.18	94.00	6.00	6.82
Oxacillin	94.47	86.96	80.28	95.00	58.82	13.04	20.59
Levofloxacin	88.99	100.00	88.00	100.00	67.86	0.00	12.00
Trimethoprim-Sulfamethoxazole	92.77	100.00	98.92	100.00	50.00	0.00	1.08
Vancomycin	96.13	-	100.00	100.00	-	-	0.00
Linezolid	82.02	-	100.00	100.00	-	-	0.00
Erythromycin	94.83	96.49	85.29	93.55	91.67	3.51	14.71
Daptomycin	82.24	-	100.00	100.00	-	-	0.00
Gentamicin	85.46	100.00	91.67	100.00	66.67	0.00	8.33



resistance phenotype through artificial intelligence. The method reported in this study is an independent prediction method without information on resistance genes that can predict the resistance of isolates with unknown resistance mechanisms. This is the first time that k-mer has been combined with machine learning to predict the MIC of *S. aureus* in China. Most previous studies predicted antimicrobial resistance through known resistance mechanisms, and the resistance phenotypes have not been obtained directly from FASTQ files (Gordon et al., 2014; Mason et al., 2018; Wang et al., 2021). The findings of this study are consistent with those of previous studies.

Staphylococcus aureus is a common pathogenic gram-positive coccus, while former research objects have been mainly gram-negative bacteria (Nguyen et al., 2019; Van Camp et al., 2020) or *Mycobacterium tuberculosis* (Satta et al., 2018). MRSA is more specific, and the resistance mechanism of MRSA corresponds to its resistant phenotype (Lee et al., 2018). However, the resistance mechanism of gram-negative bacteria is relatively complex; for example, *Enterobacteriaceae* can simultaneously develop resistance to many types of antimicrobials by producing different mechanisms, such as modifying enzymes, changing the action targets of antimicrobials, and reducing membrane permeability. Gram-negative bacteria often exhibit one resistance mechanism that causes resistance to multiple antimicrobial agents (De Oliveira et al., 2020). Therefore, in this study, for gram-positive bacteria, such as those with resistance mechanisms relatively consistent with the resistance phenotype, more accurate results can be obtained with a smaller number of training sets. For isolates with more complicated resistance mechanisms, such as gram-negative bacteria, more accurate prediction results need a larger number of training sets. The present study demonstrated the clinical value of this method in predicting the resistance phenotypes of gram-positive bacteria.

While a large number of training set isolates can greatly improve the accuracy of prediction results, it simultaneously limits the application scope of this method. For example, greater power is required for data computing and storage. As the k value increases, the size of the k-mer file increases by 4^k -fold, which poses a considerable challenge to data processing. In addition, the large amount of training set data in the previous studies was not easily accessible, and most hospitals hardly obtained such clean data. In contrast, the sequencing data for less than 500 isolates in this study were more accessible. Almost all big hospitals can collect the same number of isolates in 1–2 years, and a strain prediction model based on its own regional characteristics can be rapidly established, thus enhancing its application. All samples used in this study are from one country, and our methodology focused on large hospital data, which may not be applicable to all regions because the genotypes are clustered geographically

(Novembre et al., 2008). Collecting more samples from other countries or regions would solidify our approach.

In clinical practice, clinicians generally use antibiotics based on strain susceptibility results. The results from this study can help improve the accuracy of empirical treatment in the clinic, especially when there is no way to obtain antimicrobial susceptibility results quickly. In addition, extracting gene data characteristics using the k-mer alone can be linked with metagenomics. Analysis of the MIC directly after species determination can speed up pathogen diagnosis and antimicrobial susceptibility testing in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Peking University People's Hospital Institutional Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

HW conceived, designed, and supervised the study. SW, CZ, YY, FC, and HC collected and interpreted the data. SW and CZ conducted the analysis. SW, CZ, and HW drafted the manuscript. All the authors approved the final version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (81991533 and 81625014).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.841289/full#supplementary-material>

REFERENCES

- Avershina, E., Sharma, P., Taxt, A. M., Singh, H., Frye, S. A., Paul, K., et al. (2021). AMR-Diag: neural network based genotype-to-phenotype prediction of resistance towards beta-lactams in *Escherichia coli* and *Klebsiella pneumoniae*. *Comput. Struct. Biotechnol. J.* 19, 1896–1906. doi: 10.1016/j.csbj.2021.03.027
- Boonsiri, T., Watanabe, S., Tan, X. E., Thititanapakorn, K., Narimatsu, R., Sasaki, K., et al. (2020). Identification and characterization of mutations responsible for the beta-lactam resistance in oxacillin-susceptible *mecA*-positive *Staphylococcus aureus*. *Sci. Rep.* 10:16907. doi: 10.1038/s41598-020-73796-5
- Brinda, K., Callendrello, A., Ma, K. C., Macfadden, D. R., Charalampous, T., Lee, R. S., et al. (2020). Rapid inference of antibiotic resistance and susceptibility by

- genomic neighbour typing. *Nat. Microbiol.* 5, 455–464. doi: 10.1038/s41564-019-0656-6
- Chen, H., Fan, C., Gao, H., Yin, Y., Wang, X., Zhang, Y., et al. (2020a). Leishmaniasis diagnosis via metagenomic next-generation sequencing. *Front. Cell. Infect. Microbiol.* 10:528884. doi: 10.3389/fcimb.2020.528884
- Chen, H., Yin, Y., Gao, H., Guo, Y., Dong, Z., Wang, X., et al. (2020b). Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response. *Clin. Infect. Dis.* 71, S416–S426. doi: 10.1093/cid/ciaa1516
- Chen, T., and Guestrin, C. (2016). “XGBoost: a scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. (New York, NY: ACM). doi: 10.1145/2939672.2939785
- De Oliveira, D. M. P., Forde, B. M., Kidd, T. J., Harris, P. N. A., Schembri, M. A., Beatson, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* 33:e00181–19. doi: 10.1128/CMR.00181-19
- Deorowicz, S., Kokot, M., Grabowski, S., and Debudaj-Grabysz, A. (2015). KMC 2: fast and resource-frugal k-mer counting. *Bioinformatics* 31, 1569–1576. doi: 10.1093/bioinformatics/btv022
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010
- Gordon, N. C., Price, J. R., Cole, K., Everitt, R., Morgan, M., Finney, J., et al. (2014). Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing. *J. Clin. Microbiol.* 52, 1182–1191. doi: 10.1128/JCM.03117-13
- Khaledi, A., Weimann, A., Schniederjans, M., Asgari, E., Kuo, T. H., Oliver, A., et al. (2020). Predicting antimicrobial resistance in *Pseudomonas aeruginosa* with machine learning-enabled molecular diagnostics. *EMBO Mol. Med.* 12:e10264. doi: 10.15252/emmm.201910264
- Kim, J., Greenberg, D. E., Pifer, R., Jiang, S., Xiao, G., Shelburne, S. A., et al. (2020). VAMPr: VARIant mapping and prediction of antibiotic resistance via explainable features and machine learning. *PLoS Comput. Biol.* 16:e1007511. doi: 10.1371/journal.pcbi.1007511
- Lakhundi, S., and Zhang, K. (2018). Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. *Clin. Microbiol. Rev.* 31:e00020–18. doi: 10.1128/CMR.00020-18
- Lee, A. S., De Lencastre, H., Garau, J., Kluytmans, J., Malhotra-Kumar, S., Peschel, A., et al. (2018). Methicillin-resistant *Staphylococcus aureus*. *Nat. Rev. Dis. Primers* 4:18033. doi: 10.1038/nrdp.2018.33
- Li, S., Sun, S., Yang, C., Chen, H., Yin, Y., Li, H., et al. (2018). The changing pattern of population structure of *Staphylococcus aureus* from bacteremia in China from 2013 to 2016: ST239-030-MRSA replaced by ST59-t437. *Front. Microbiol.* 9:332. doi: 10.3389/fmicb.2018.00332
- Lowy, F. D. (1998). Medical progress: *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532. doi: 10.1056/NEJM199808203390806
- Macesic, N., Bear Don't Walk, O. J., Pe'er, I., Tatonetti, N. P., Peleg, A. Y., and Uhlemann, A. C. (2020). Predicting phenotypic polymyxin resistance in *Klebsiella pneumoniae* through machine learning analysis of genomic data. *mSystems* 5:e00656–19. doi: 10.1128/mSystems.00656-19
- Mason, A., Foster, D., Bradley, P., Golubchik, T., Doumith, M., Gordon, N. C., et al. (2018). Accuracy of different bioinformatics methods in detecting antibiotic resistance and virulence factors from *Staphylococcus aureus* whole-genome sequences. *J. Clin. Microbiol.* 56:e01815–17. doi: 10.1128/JCM.01815-17
- Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electron. Mag. Electron.* 38, 114–117.
- Nguyen, M., Long, S. W., Mcdermott, P. F., Olsen, R. J., Olson, R., Stevens, R. L., et al. (2019). Using machine learning to predict antimicrobial MICs and associated genomic features for nontyphoidal *Salmonella*. *J. Clin. Microbiol.* 57:e01260–18. doi: 10.1128/JCM.01260-18
- Nguyen, M., Olson, R., Shukla, M., Vanoeffelen, M., and Davis, J. J. (2020). Predicting antimicrobial resistance using conserved genes. *PLoS Comput. Biol.* 16:e1008319. doi: 10.1371/journal.pcbi.1008319
- Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R., Auton, A., et al. (2008). Genes mirror geography within Europe. *Nature* 456, 98–101. doi: 10.1038/nature07331
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Satta, G., Lipman, M., Smith, G. P., Arnold, C., Kon, O. M., and Mchugh, T. D. (2018). *Mycobacterium tuberculosis* and whole-genome sequencing: how close are we to unleashing its full potential? *Clin. Microbiol. Infect.* 24, 604–609. doi: 10.1016/j.cmi.2017.10.030
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* 18, 318–327. doi: 10.1016/S1473-3099(17)30753-3
- Van Camp, P. J., Haslam, D. B., and Porollo, A. (2020). Prediction of antimicrobial resistance in gram-negative bacteria from whole-genome sequencing data. *Front. Microbiol.* 11:1013. doi: 10.3389/fmicb.2020.01013
- Vanoeffelen, M., Nguyen, M., Aytan-Aktug, D., Brettin, T., Dietrich, E. M., Kenyon, R. W., et al. (2021). A genomic data resource for predicting antimicrobial resistance from laboratory-derived antimicrobial susceptibility phenotypes. *Brief. Bioinform.* 22:bbab313. doi: 10.1093/bib/bbab313
- Wang, W., Baker, M., Hu, Y., Xu, J., Yang, D., Maciel-Guerra, A., et al. (2021). Whole-genome sequencing and machine learning analysis of *Staphylococcus aureus* from multiple heterogeneous sources in china reveals common genetic traits of antimicrobial resistance. *mSystems* 6:e0118520. doi: 10.1128/mSystems.01185-20

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Zhao, Yin, Chen, Chen and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Ultrafast and Cost-Effective Pathogen Identification and Resistance Gene Detection in a Clinical Setting Using Nanopore Flongle Sequencing

Ekaterina Avershina¹, Stephan A. Frye², Jawad Ali¹, Arne M. Taxt² and Rafi Ahmad^{1,3*}

¹ Department of Biotechnology, Inland Norway University of Applied Sciences, Hamar, Norway, ² Division of Laboratory Medicine, Department of Microbiology, Oslo University Hospital, Oslo, Norway, ³ Faculty of Health Sciences, Institute of Clinical Medicine, UiT - The Arctic University of Norway, Tromsø, Norway

OPEN ACCESS

Edited by:

Costas C. Papagiannitsis,
University of Thessaly, Greece

Reviewed by:

Tamara Salloum,
Brigham and Women's Hospital
and Harvard Medical School,
United States
Jaroslav Hrabak,
Charles University, Czechia

*Correspondence:

Rafi Ahmad
rafi.ahmad@inn.no

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 25 November 2021

Accepted: 27 January 2022

Published: 17 March 2022

Citation:

Avershina E, Frye SA, Ali J,
Taxt AM and Ahmad R (2022)
Ultrafast and Cost-Effective Pathogen
Identification and Resistance Gene
Detection in a Clinical Setting Using
Nanopore Flongle Sequencing.
Front. Microbiol. 13:822402.
doi: 10.3389/fmicb.2022.822402

Rapid bacterial identification and antimicrobial resistance gene (ARG) detection are crucial for fast optimization of antibiotic treatment, especially for septic patients where each hour of delayed antibiotic prescription might have lethal consequences. This work investigates whether the Oxford Nanopore Technology's (ONT) Flongle sequencing platform is suitable for real-time sequencing directly from blood cultures to identify bacteria and detect resistance-encoding genes. For the analysis, we used pure bacterial cultures of four clinical isolates of *Escherichia coli* and *Klebsiella pneumoniae* and two blood samples spiked with either *E. coli* or *K. pneumoniae* that had been cultured overnight. We sequenced both the whole genome and plasmids isolated from these bacteria using two different sequencing kits. Generally, Flongle data allow rapid bacterial ID and resistome detection based on the first 1,000–3,000 generated sequences (10 min to 3 h from the sequencing start), albeit ARG variant identification did not always correspond to ONT MinION and Illumina sequencing-based data. Flongle data are sufficient for 99.9% genome coverage within at most 20,000 (clinical isolates) or 50,000 (positive blood cultures) sequences generated. The SQK-LSK110 Ligation kit resulted in higher genome coverage and more accurate bacterial identification than the SQK-RBK004 Rapid Barcode kit.

Keywords: ONT sequencing, pathogen identification, antibiotic resistance gene (ARGs), clinical sample, Flongle

INTRODUCTION

The rising antimicrobial resistance of recent years poses a major threat to humanity and rapid diagnostic tools for bacterial infections are urgently needed (Avershina et al., 2021a). The current state-of-the-art in infection diagnostics are mostly based on biochemical analysis of cultured clinical samples (Vasala et al., 2020). The turnaround time is around 2–4 days for most samples, but a longer time is required for low bacterial load samples, such as bloodstream infection (BSI) [usually 1–100 colony-forming units (CFU)/mL], where identification and antibiotic susceptibility testing (AST) could take up to 5 days (Metzgar et al., 2016; Briggs et al., 2021). There are emerging

Abbreviations: AMR, antimicrobial resistance; ARG, antimicrobial resistance genes; ONT, Oxford Nanopore Technology; WGS, Whole-genome sequencing.

micro- and nanotechnologies for bacterial identification and AST, including both phenotypic (e.g., microfluidic-based bacterial culture) and molecular (e.g., multiplex PCR, hybridization probes, nanoparticles, synthetic biology, and mass spectrometry) methods (Li et al., 2017).

The very limited availability of rapid, easy-to-use and scalable methods to interpret whole-genome sequencing (WGS) data for clinical purposes make it challenging. Illumina sequencing is not time-efficient, but recent studies have shown that Oxford Nanopore Technology's (ONT) MinION could potentially be used for point of care sequencing and become a basis for WGS-based diagnostic strategies (Harstad et al., 2018; Text et al., 2020). Examples of its clinical usefulness include same-day diagnostic results for tuberculosis and blood cultures (Votintseva Antonina et al., 2017; Text et al., 2020). Also, bacterial pathogen identification in lower respiratory infection (sputum samples) within 6 h has been demonstrated (Charalampous et al., 2019).

The faster a resistance profile of an infectious agent is known, the faster the right treatment can be initiated (Trevas et al., 2021). We have previously demonstrated that MinION real-time sequencing can be successfully used for bacterial identification and antimicrobial resistance (AMR) detection in blood cultures within 4 h after the blood culture was flagged positive (Text et al., 2020). Flongle provides a less expensive sequencing setup than MinION (512 sequencing channels) but it has a lower total data output as it only has a maximum of 126 available sequencing channels, i.e., pores, with 60 guaranteed by the manufacturer (Grädel et al., 2019).

There is currently a very limited number of publications using Flongle flow cell sequencing available at PubMed. A search on August 16, 2021, with keywords “Flongle” and “nanopore Flongle,” resulted in a total of only 14 publications. With regard to its use in clinical microbiology, a Flongle-based assay has produced accurate and reproducible results for enterovirus identification and genotyping (Grädel et al., 2019). Also, recently Flongle and MinION have been used in profiling preterm microbiota and AMR in fecal samples (Leggett et al., 2020). However, only 1 out of 20 samples was sequenced with the Flongle flow cell (Leggett et al., 2020), which is too small of an amount to conclude on the clinical usefulness of Flongle.

This work assessed whether Flongle flow cells are as proficient for bacterial ID and antimicrobial resistance genes (ARGs) detection as MinION flow cells. We demonstrate the performance of Flongle in the sequencing of (a) cultured clinical isolates of *E. coli* and *K. pneumoniae*, (b) blood cultures spiked with clinical isolates using Flongle flow cells with 59–60 active pores, and (c) as low as 24–39 active pores. The Flongle data were validated using previously published MinION and Illumina MiSeq sequencing data for these isolates (Text et al., 2020; Avershina et al., 2021b; Khezri et al., 2021a,b). We demonstrate the sufficiency of Flongle for bacterial ID and ARG detection within the first obtained sequences both on pure isolates and direct blood cultures. We also demonstrate that whole-genome sequencing is preferable over plasmid isolation and sequencing. Flongle accuracy was, however, insufficient for correct CTX-M variant calling, but this task is relatively less critical for proper antibiotic treatment adjustment.

MATERIALS AND METHODS

Ethics Statement

In this study, we worked with bacterial isolates from the strain collection at the Department of Microbiology at Oslo University Hospital (OUH). Human blood was obtained from anonymous healthy donors via the blood bank at OUH. There was no intention to sequence human DNA and, therefore, QIAamp BiOstic Bacteremia DNA Kit for extraction of bacterial DNA was used. Moreover, any sequencing reads recognized as being generated from human DNA were omitted from further analysis and permanently discarded. Ethical approval was not deemed necessary. The clinical microbiology laboratory at OUH is approved for the described experimental work and carries out diagnostic work in the clinical routine.

Dataset Description

This study used two clinical isolates of *Escherichia coli* (*E. coli* 125 and *E. coli* A2-39) and two clinical isolates of *Klebsiella pneumoniae* (*K. pneumoniae* 225 and *K. pneumoniae* A2-37). The rationale for using these isolates is an extensive amount of sequencing data (both Illumina MiSeq and MinION) available from our previously published works (Text et al., 2020; Avershina et al., 2021b; Khezri et al., 2021b). These isolates were subject to whole genome sequencing and plasmid sequencing from the bacterial culture using Flongle flow cells. Sequencing was performed both using ONT Rapid Barcoding kit SQK-RBK004 (four isolates per flow cell) and Genomic DNA by Ligation SQK-LSK110 (one isolate per flow cell). In addition, we have spiked blood with the *E. coli* A2-39 and *K. pneumoniae* A2-37 isolates, cultured overnight, and performed WGS directly on the DNA isolated from the blood culture using flow cells with two times the difference in the number of active sequencing pores. The QIAamp BiOstic Bacteremia DNA Kit from Qiagen (Germany) was used for the extraction of bacterial DNA to be used for nanopore sequencing (Text et al., 2020). DNA extraction was performed according to the manufacturer's instructions. Human blood was obtained from healthy anonymous donors via the blood bank at OUH. The blood was used before the “best by” date, no later than 3 weeks after the blood was drawn. All experiments, as well as ARGs possessed by the isolates, are reported in **Supplementary Table 1**.

Sample Preparation

Genomic DNA was isolated from plate-grown bacterial cultures using the Midi format of the protocol for Gram-negative and some Gram-positive bacterial samples with Genomic-tip 100/G columns following the manufacturer's instructions (Qiagen, Germany). Plasmid DNA was isolated using the Qiagen QIAprep Spin Kit. Purified DNA was stored in Tris buffer at -20°C and quantified using the Qubit dsDNA HS assay kit (Thermo Fisher Scientific, United States). Sequencing samples were prepared for the run on Flongle flowcells following the manufacturer's instructions. Rapid Barcoding Sequencing (SQK-RBK004) and Genomic DNA by Ligation (SQK-LSK110) kits were used together with the solutions from the custom kit

FLP003. Sequencing was done on Flongle flow cells (R9.4.1) for at least 24 h. The availability of sufficient active pores was confirmed on arrival and directly before the sequencing run. The starting bias voltage was -180 mV with 1.5 h between channel scans. High accuracy base calling with Guppy version 4.3.4 run on GPU was performed. In the case of SQK-RBK004, reads were also demultiplexed at this step. The Qscore limit was 7 and no filtration for length was applied. The output was sent to *fastq* files with 1,000 reads per file. Reads recognized as generated from human DNA were omitted from further analysis and discarded.

Whole-Genome Sequencing Data Analysis

WGS reads from each *fastq* file, as well as combined reads from each whole sequencing run, were mapped to previously published Illumina-MinION hybrid assemblies (Taxt et al., 2020; Khezri et al., 2021b; (a) of the corresponding isolates using CLC Genomic Workbench v.20.0.4 (Qiagen, Germany). Each output file was searched with Blastn v 2.6.0 against NCBI RefSeq (version June 2021) and ResFinder (version June 2021) databases for bacterial ID and ARG, respectively. Plasmids were searched against the PLSDb database using a minimum identity of 90% as a cut-off (Galata et al., 2019).

WGS data with $> 100\times$ genome coverage was split into 12 subsets. Each subset was then filtered by length (1,000 bp) using Filtlong v.0.2.1 and assembled using Flye v2.8.3, Miniasm-0.3, or Raven v1.5.1, and resulting assemblies were reconciled and optimized using Trycycler v.0.5.0.

Plasmid Data Analysis

Reads generated from plasmid sequencing were mapped to reference hybrid assemblies (Taxt et al., 2020; Khezri et al., 2021b) using CLC Genomic Workbench, v.20.0.4 (Qiagen, Germany) and assembled using Unicycler v0.4.9 and Flye v2.8.3. Assembly graphs were visualized using Bandage v0.8.1.

Unless stated otherwise, all statistical analyses were performed in MATLAB R2020b (MathWorks Ltd., United States).

RESULTS

General Statistics of the Oxford Nanopore Technology's Flongle Runs

The number of active pores on the Flongle cells remained stable during storage in the cold room four of which lasted up to 8 months (Supplementary Table 2).

On average, $4,430 \pm 3,506$ [mean \pm standard deviation] and $151,435 \pm 95,325$ reads per sample were generated using SQK-RBK004 and SQK-LSK110 kits, respectively (Table 1). N50 values comprised $9,188 \pm 6,758$ bp for SQK-RBK004 and $11,867 \pm 5,081$ bp for SQK-LSK110, longest generated reads ranged from 24,745 to 124,329 bp for SQK-RBK004 and from 40,760 to 146,931 bp for SQK-LSK110. The total length generated for clinical isolates using the SQK-LSK110 kit was on average 126x higher compared to SQK-RBK004 ($p < 0.05$). When using SQK-RBK004, only one output file was generated after the whole sequencing was run

for one of the isolates, i.e., it required > 24 h to get data for that isolate. Additionally, 1.2% of the barcodes were misidentified during the demultiplexing step.

First Output Files of Whole-Genome Sequencing Are Sufficient for Bacterial Identification and Antimicrobial Resistance Gene Detection in Clinical Isolates

Taxonomic assignment of reads remained the same throughout the whole sequencing run (Figure 1), indicating that the first generated and outputted 1,000 sequences were as sufficient for bacterial identification as the data from the whole sequencing run. *K. pneumoniae* reads in A2-37 (experiments 1 and 5) were commonly misclassified as *K. quasipneumoniae*—the species that is closely related to *K. pneumoniae* (Long et al., 2017). In *E. coli* A2-39 (SQK-RBK004, experiment 1), only 189 sequences were generated, and nine percent of these reads were assigned to *K. pneumoniae*. Interestingly, there was no misidentification of *E. coli* to *K. pneumoniae* using SQK-LSK110 (experiment 4).

First hits to ARGs were found in the first or the second generated output file (5 min to 2.6 h after the sequencing start), and the whole resistome was detected at most within the first 3,000 sequences. However, assignment of the correct variant (SHV-187, CTX-M-2, CTX-M-14) was not possible using single Flongle reads (Supplementary Table 3).

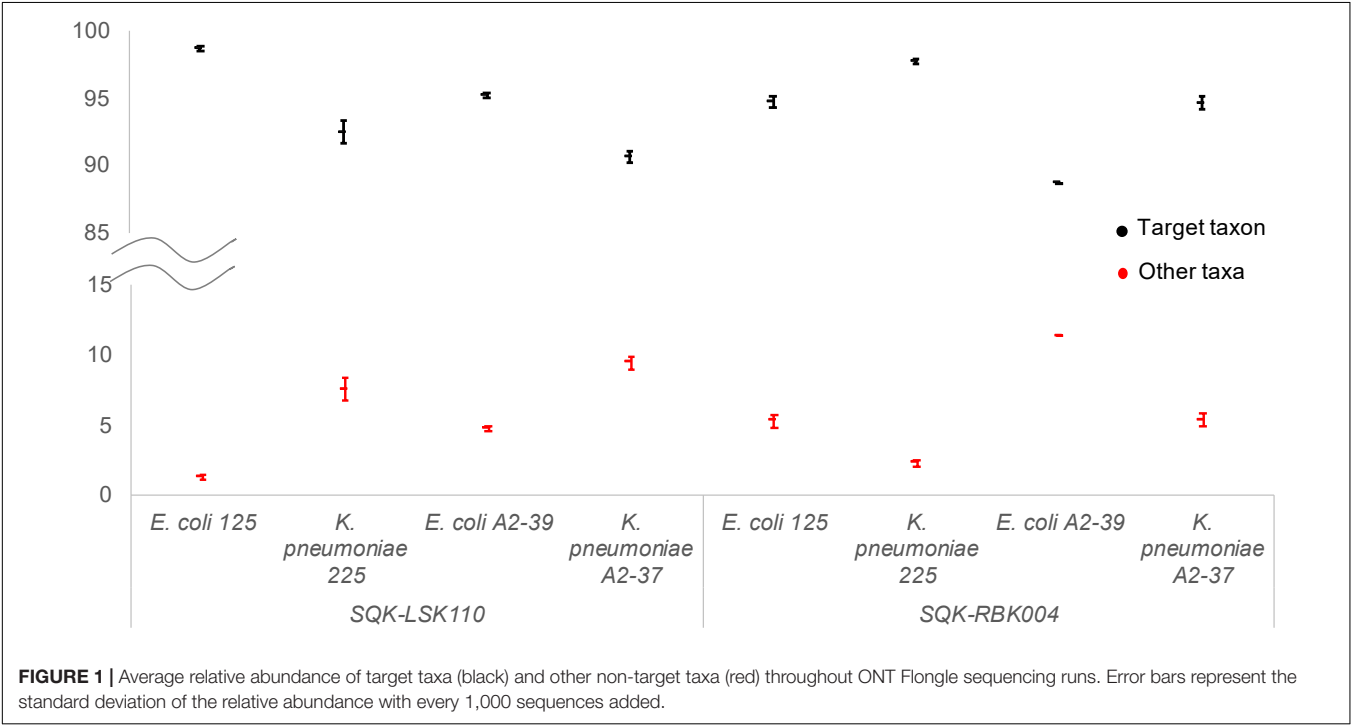
For *E. coli* 125, the target gene TEM-1B was not detected either when the isolate was sequenced using SQK-RBK004 [experiment 1; WGS data map to 75% of the reference hybrid assembly (Khezri et al., 2021b) with 3X coverage] or using Ligation kit [experiment 2, WGS data map to 99% of the reference hybrid assembly (Khezri et al., 2021b) with 132X coverage]. Interestingly, the TEM-1B gene in this isolate was also previously not detected by the MinION, although confirmed by Illumina sequencing (Khezri et al., 2021b). In the reference hybrid assembly, TEM-1B was located on a 10,909 bp plasmid contig (99.9% identity to *E. coli* O55:H7 strain RM12579 p12579_5; 5954 bp). This contig was supported only by the Illumina data (4,750 reads mapped, 27x coverage) and additionally contained Tn3 family transposase Tn2, Transposon Tn3 resolvase, and DNA relaxase MbeA (Supplementary Table 4). Neither Flongle nor previously published MinION generated reads (Khezri et al., 2021b) mapped to this contig. Coverage of the TEM-1B gene by Illumina data ranged from 7X to 25X. We then performed PCR to address the discrepancy between the platforms (Supplementary Text 1). The presence of TEM-1 was not supported by PCR (Supplementary Figure 1).

SQK-LSK110 Flongle Generated Data Is Sufficient for Whole-Genome Coverage of the Clinical Isolates Using at Most the First 20,000 Sequencing Reads Generated

Due to the large variation in time required for the generation of each 1,000 sequences [around 10–30 min per 1,000 sequences

TABLE 1 | General statistics of the ONT Flongle runs.

Exp Nr	ID	Sample type	ONT kit	Number of active pores	Number of reads	Total length (Mbp)	Longest read (bp)	N50 (bp)
1	<i>E. coli</i> 125	Clinical isolates	SQK-RBK004	69	7,840	25.6	89,995	7,956
	<i>K. pneumoniae</i> 225				9,493	76.6	124,329	21,552
	<i>E. coli</i> A2-39				879	1.9	58,672	11,868
	<i>K. pneumoniae</i> A2-37				6,744	35.1	136,325	14,239
2	<i>E. coli</i> 125	Plasmid preparation	SQK-LSK110	70	10,5368	683.5	122,281	13,232
3	<i>K. pneumoniae</i> 225			60	102,714	904.0	127,460	21,738
4	<i>E. coli</i> A2-39			67	313,395	861.6	124,559	9,871
5	<i>K. pneumoniae</i> A2-37			59	193,000	1082.9	146,931	14,684
6	<i>E. coli</i> 125	Spiked blood culture	SQK-RBK004	59	1,144	1.7	24,745	1,626
	<i>K. pneumoniae</i> 225				189	0.8	44,111	10,015
	<i>E. coli</i> A2-39				3,575	6.1	51,440	2,359
	<i>K. pneumoniae</i> A2-37				5,579	12.3	54,931	3,889
7	<i>E. coli</i> A2-39	Spiked blood culture	SQK-LSK110	62	293,814	1273.8	40,760	7,044
8	<i>K. pneumoniae</i> A2-37			59	127,537	756.8	41,029	7,961
9	<i>E. coli</i> A2-39			39	95,689	528.6	44,214	7,260
10	<i>K. pneumoniae</i> A2-37			24	93,892	583.5	39,975	8,339



when using one sample per sequencing run (SQK-LSK110) and up to 6 h when using four samples per run (SQK-RBK004)], we decided to report the data with regards to the number of sequences generated rather than an actual timeline.

With SQK-RBK004, only the *K. pneumoniae* 225 isolate had a sufficient amount of sequencing data to cover > 95% of its reference hybrid assembly at least once by the end of the run (experiment 1, **Figure 2**). For the other isolates, however, portions of genome sequenced ranged from < 10% (*E. coli* A2-39, experiment 4) to 80% (*K. pneumoniae* A2-37, experiment 5). When using the ligation kit, all isolates covered > 99%

reference hybrid assembly within at most the first 20,000 sequences. On average, these data were generated 4.5 h after the sequencing started.

The Flongle Based Assemblies of Clinical Isolates Are Not Suitable for Antimicrobial Resistance Gene Variant Calling

On average, Flye, Miniasm and Minipolish, and Raven assembled each isolate (experiments 2–5) into 6 ± 4 contigs. After clustering

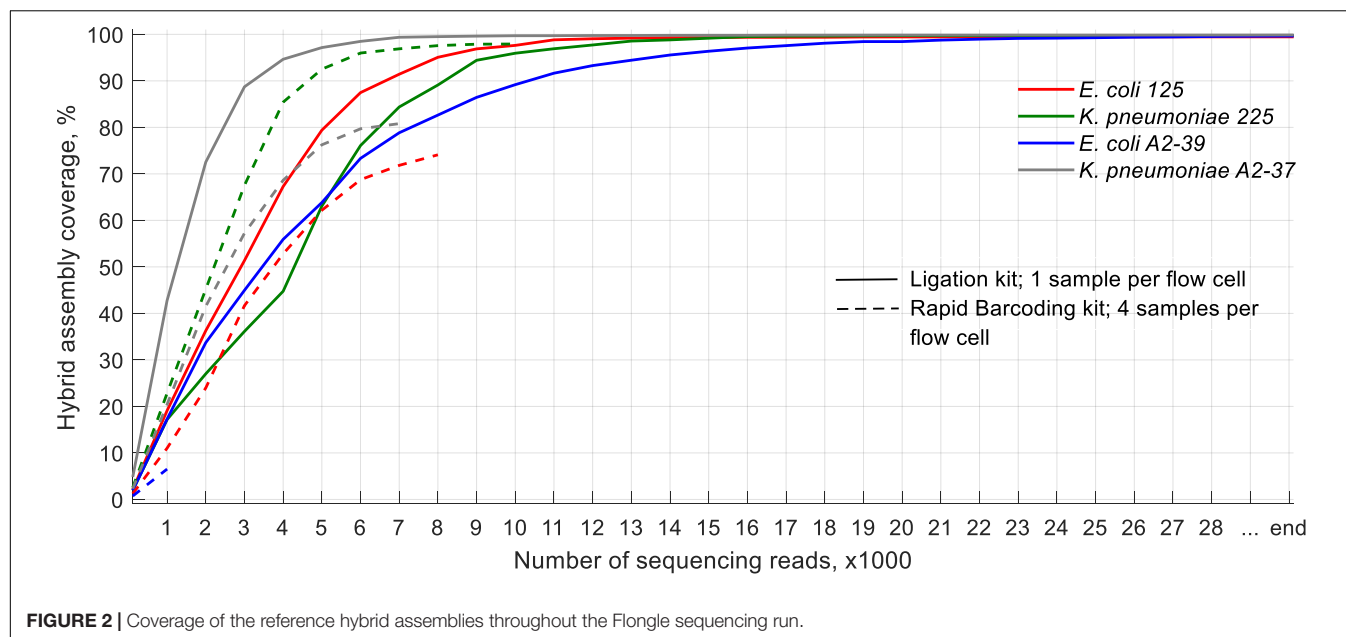


FIGURE 2 | Coverage of the reference hybrid assemblies throughout the Flongle sequencing run.

and assemblies' reconciliation, the number of contigs on average reduced to 4 ± 2 (Figure 3), and chromosomes were fully assembled. In the case of *K. pneumoniae* 225 (experiment 3) and *E. coli* A2-39 (experiment 4), target ARGs were correctly assigned to an expected variant, SHV-187, and CTX-M-2, respectively (Figure 3). However, unlike reference hybrid assembly (Khezri et al., 2021b), Flongle assembly suggests that SHV-187 was located on the *K. pneumoniae* 225 chromosome rather than on a plasmid. It is worth noting that using Tricycler assembly, an additional CTX-M-35 was detected on the same plasmid from *E. coli* A2-39 isolate upstreams of CTX-M-2. Twenty-four raw Flongle reads (8,816–71,122 bp long) had an additional CTX-M gene located approximately (\approx)4 kbp downstream from the first CTX-M gene. Both copies had 35–39x coverage by MinION reads, but none of the reads covered two copies simultaneously. Illumina reads supported only CTX-M-2 (73X), but not CTX-M-35.

Interestingly, one of the plasmids in *E. coli* A2-39 isolate (experiment 4), where 2.3% reads were assigned to *Salmonella enterica* (Figure 1), was closely related to *Salmonella enterica* subsp. *enterica* plasmid pSH1148_107 (Figure 3).

Flongle Sequencing of Plasmids From Clinical Isolates Recovered < 9 kbp Long Plasmids Only

On average, $2,622 \pm 2,433$ reads per isolate were generated using the SQK-RBK004 kit (experiment 6, Table 2). In the *K. pneumoniae* 225 isolate, a single output file with 189 reads was generated. The majority of reads mapped to chromosomal contigs of the reference hybrid assembly (Supplementary Table 5), and only 9.3% (*K. pneumoniae* 225) to 37.6% (*E. coli* A2-39) mapped to plasmid contigs. A minor part of generated reads could be assembled into circular or linear contigs of at most 16,324 bp (Table 2). Plasmid sequencing only

revealed < 9 kbp long plasmids from the isolates. Interestingly, for *K. pneumoniae* A2-37, Unicycler produced one circular plasmid of 4,447 bp closely related to *K. pneumoniae* plasmid pZZW20-4.4K, 4,436 bp, also detected using WGS Flongle sequencing of the clinical isolate (experiment 5). Conversely, Flye produced three contigs. Two of these contigs gave the closest PLSDb match to plasmids 4–10 times shorter, including a match to *K. pneumoniae* plasmid pZZW20-4.4K. The third contig did not produce any close match to a plasmid using PLSDb (Table 2).

The CTX-M gene was detected in a single read (14,495 bp) from *K. pneumoniae* A2-37 plasmid data (CTX-M, variant not assigned) (Table 2). After assembly, however, the CTX-M gene was not detected in any of the contigs. Based on the WGS data from the clinical isolate (experiment 5), the gene was located on the circular 84,438 bp plasmid closely related to *K. pneumoniae* pFAM22321 (Figure 3). Indeed, the 14,495 bp CTX-M containing read mapped to this plasmid contig with 87% coverage.

Direct Flongle Sequencing of Positive Blood Cultures Is as Efficient as That of the Pure Clinical Isolates for Bacterial ID and Antimicrobial Resistance Identification, but Not for Plasmid Detection

Same as with clinical isolates, the first file out was sufficient for bacterial ID as any other consequent file, albeit with a higher fluctuation in relative abundances (Supplementary Figure 2). Human reads comprised < 10% in all cases apart from *K. pneumoniae* A2-37 blood culture, where human data comprised 20% of the sequencing data. Like *K. pneumoniae* A2-37 clinical isolate sequencing, reads were also assigned to *K. quasipneumoniae*.

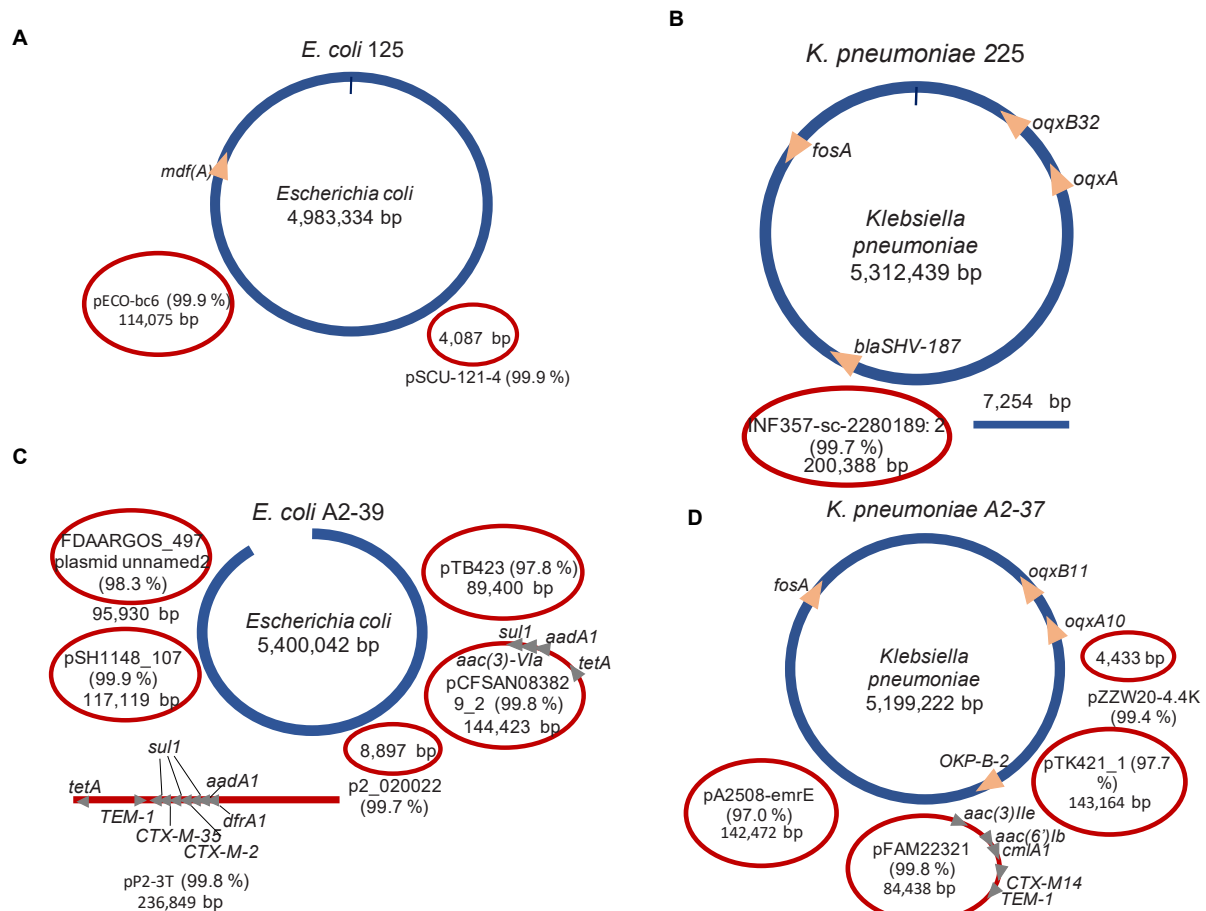


FIGURE 3 | Summary of the Tricycler assemblies' reconciliation of clinical isolates using SQK-LSK110 Flongle. Blue color depicts chromosomal contigs, red color—plasmid contigs. Circular contigs are represented as closed ellipses, linear—as lines, and open ellipses. For plasmid contigs, the closest search hit against the PLSDB plasmid database is given as a plasmid name. Identity% is provided in brackets. **(A)** *Escherichia coli* 125 (experiment 2); **(B)** *Klebsiella pneumoniae* 225 (experiment 3); **(C)** *Escherichia coli* A2-39 (experiment 4); **(D)** *Klebsiella pneumoniae* A2-37 (experiment 5).

Using a flow cell with ≈ 60 active pores required at most 75,000 reads to sequence $> 99\%$ of the genome at least once for the blood culture spiked with *K. pneumoniae* A2-37 (**Figure 4**). Using a flow cell with ≈ 30 active pores did not impair the sequencing performance. It took, at most, 56,000 reads to sequence $> 99\%$ of the genome, and in this case, *E. coli* A2-39 blood culture required more sequencing data. All ARGs were detected within the first 3,000–4,000 sequences apart from the sample spiked with *K. pneumoniae* A2-37 and sequenced using a flow cell with ≈ 60 active pores. In that case, the first indication of *fosA* and *oqxA* genes appeared only after 10,000 reads. Same as with sequencing of pure clinical isolates, ARG variant detection was not possible.










Comparison of plasmid contigs between assemblies from clinical isolates and blood cultures of these clinical isolates revealed 99.7–99.9% pairwise identity between the corresponding contigs (**Supplementary Table 6**). However, blood culture sequencing assemblies of *E. coli* A2-39 missed a plasmid closely related to *Salmonella enterica* subsp. *enterica* plasmid pSH1148_107 (NC_019123.1). Both *K. pneumoniae* A2-37 blood culture assemblies lacked a plasmid contig closely related to

K. pneumoniae plasmid pZZW20-4.4K, 4436 bp (NZ_CP058963). Same as with clinical isolate assembly, CTX-M-14 was found located on a plasmid contig in *K. pneumoniae* A2-37. In the case of *E. coli* A2-39, two copies of the CTX-M-35 gene were detected on plasmid contigs. Twelve raw Flongle reads contained both copies of the CTX-M gene, separated by ≈ 4 kbp region.

DISCUSSION

Fast identification of the infection is crucial for proper treatment. Previously we have established that real-time ONT MinION sequencing allows rapid identification of bacteria within 10 min after the sequencing run (Taxt et al., 2020). In this work, we demonstrate that a cheaper Flongle setup is also sufficient for fast detection and identification of bacteria, especially using one sample per flow cell. The first output files of 1,000 sequences produced enough data to identify target bacteria both when pure isolates and spiked blood cultures were sequenced, albeit the time needed to generate the first file varied extensively between flow

TABLE 2 | Summary of ONT Flongle sequencing of plasmids.

Plasmids from	Total number of reads generated	Longest raw read, bp	Assembly				Closest match at PLSDb					Target AMR	
			Assembler	Assembly length, bp	Assembly graph	Raw reads mapped to assembly	Plasmid ID	Plasmid length, bp	ID,%	AMR	Assembled in WGS	Raw reads	Assembly
<i>E. coli</i> 125	1,144	24,745	Unicycler	5,644		21	<i>E. coli</i> pSCU-121-4; CP054332.1	4,091	99.9	No	Yes	No	No
			Flye	4,083		20	<i>E. coli</i> pSCU-121-4; CP054332.1	4,091	99.8	No	Yes		No
<i>K. pneumoniae</i> 225	189	44,111	Unicycler	3,295		43	<i>K. pneumoniae</i> INF072-sc-2279995, plasmid: 4; LR890188.1	3,302	99.8	No	No	No	No
			Flye			Too little data, error in the assembly						No	—
<i>E. coli</i> A2-39	3,575	51,440	Unicycler	8,893		102	<i>E. coli</i> p2_020022; NZ_CP032881.1	8,899	99.7	No	Yes	No	No
			Flye	7,655		149	<i>E. coli</i> p2_020022; NZ_CP032881.1	8,899	98.5	No	Yes		No
				2,347			<i>E. coli</i> pSCU-172-7; NZ_CP054360.1	2,311	99.7	No	No		No
				1,033			no hits	—	—	—	—		No
<i>K. pneumoniae</i> A2-37	5,579	54,931	Unicycler	4,447		288	<i>K. pneumoniae</i> pZZW20-4.4K; NZ_CP058963.1	4,436	98.8	No	Yes	CTX-M, 1st file	No
<i>K. pneumoniae</i> A2-37	5,579	54,931	Flye	16,324		911	<i>Klebsiella pneumoniae</i> plasmid pZZW20-4.4K; NZ_CP058963.1	4,436	96.2	No	Yes	CTX-M, 1st file	No
				13,264			<i>Klebsiella grimontii</i> SS141 plasmid_3.2; NZ_CP044531.1	1,180	96.2	No	No		No
				4,519			—	—	—	—	—		No

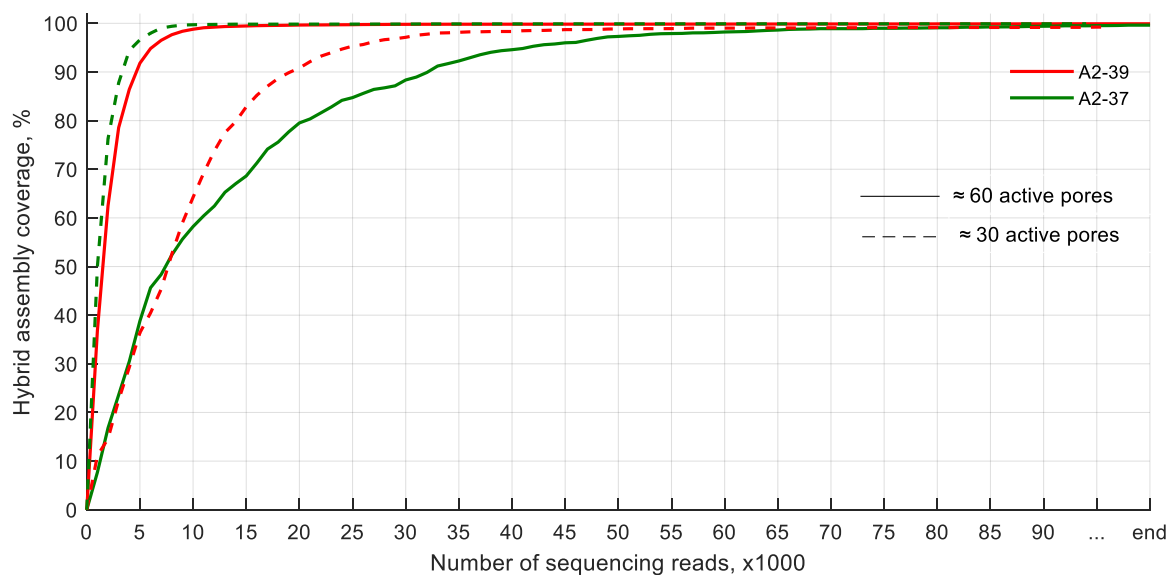


FIGURE 4 | Coverage of the genome by ONT Flongle sequencing of the spiked blood cultures.

cells and sequencing kits. When using the ligation kit, the first 1,000 sequences were available in 10–15 min after the sequencing started. Conversely, the rapid barcoding kit required up to 2 h to generate the first output file for each barcode. Additionally, we noticed a possible miscalling of used barcodes. A study conducted on MinION sequencing of three viruses suggested around 0.056% of the generated reads to be misassigned to a different barcode (Xu et al., 2018). In our work, Flongle sequencing has exhibited a higher rate of barcode misassignment, with 1.2% of generated reads assigned to barcodes that were not used in the run using standard Guppy base calling and demultiplexing provided by ONT. Therefore, we suggest that it is safer to sequence one sample per flow cell during clinical sample sequencing.

We observed high variation between throughput for each isolate when using rapid barcoding, with some genomes being only < 10% sequenced throughout the whole run. When one sample per flow cell was sequenced, conversely, the total throughput was > 100× genome coverage, and around 20,000–50,000 generated sequences covered the whole genome to be sequenced at least once.

The first 3,000 sequences were sufficient for detecting target ARGs both when sequencing clinical isolates and blood cultures spiked with these isolates in all instances. Same as with MinION, however, variant detection was not possible until the data were assembled (Taxt et al., 2020). Flongle generated sequencing data previously exhibited high concordance with Illumina data regarding gene fusion detection (Jeck et al., 2021). In our work, however, there were discrepancies between ONT flow cells and Illumina MiSeq generated data, especially regarding ARG content. Unlike MinION and MiSeq, Flongle sequencing data found the CTX-M-35 gene in addition to the CTX-M-2 gene in the *E. coli* A2-39 plasmid. These two copies were separated by a ≈4 kbp region which was supported by a total of 36 raw Flongle reads.

The other discrepancy between ONT (both MinION and Flongle) and Illumina MiSeq data with regards to ARGs is the TEM-1B penicillinase in *E. coli* 125. Phenotypically, *E. coli* 125 was susceptible to penicillin (Avershina et al., 2021b), which suggests that the TEM-1B gene in this isolate was not expressed. Due to numerous rounds of PCR, Illumina sequencing data are prone to chimera generation (Arroyo Mühr et al., 2020), and chimera removal is a common filtering step when using amplicon sequencing with Illumina (Prodan et al., 2020). However, it seems doubtful that a TEM-1B-identical sequence would be chimera-generated. Therefore, we believe that the absence of TEM-1B might have been caused by a loss of the TEM-1B containing plasmid by *E. coli* 125.

ARGs located on plasmids can be transferred across species and thus pose a higher threat of broad distribution than those located on a chromosome. Moreover, plasmids can be present in multiple copies in a bacterial cell, and the plasmid copy number can increase if environmental conditions favor the survival of plasmid carriers (f.ex. presence of antibiotic) (San Millan et al., 2015). It has been previously demonstrated that plasmids > 150 kbp in size are difficult to extract using column-based methods (Villa and Carattoli, 2020). WGS sequencing was recently recommended for plasmid reconstruction using ONT MinION over plasmid extraction and sequencing (Berbers et al., 2020). In our work, we also favor whole-genome sequencing to detect plasmid-borne ARGs and plasmid reconstruction with ONT Flongle.

Currently, the price per Flongle flow cell, which according to the manufacturer, can deliver up to 2.8 gigabytes of data, is \$90.¹ Given its fast turn-around time, and low cost, Flongle sequencing has a high potential for being used for infection diagnostics in a clinical setting. According to ONT,

¹<https://nanoporetech.com/>

the recommended Flongle flow cell storage time is 4 weeks.² However, we have stored the flow cells for as long as 8 months, and we observed that the number of active pores remained similar. This is a very relevant observation, which shows that the Flongle flow cells could still be used and provide sufficient data, thereby being more cost-effective. Besides, ONT is launching a ligation-based sequencing kit for multiplexing samples (\$600 for 6 reactions). Based on our positive results for the ligation-based sequencing kit, this would be an interesting approach for future studies.

CONCLUSION

In this work, we demonstrated that Flongle can be successfully used for the identification of bacteria and for the detection of antibiotic resistance genes both in clinical isolates and when using direct sequencing of blood spiked with these clinical isolates and cultured over-night. We suggest that currently, the optimal setup is to sequence one sample per flow cell, which provides an adequate amount of data to give the relevant information. Given that the task of whole-genome assembly is not a requirement in routine clinical diagnostics, therefore the sequencing run can be stopped after the first 20,000–50,000 sequences as with that amount of data, the whole genome was covered at least once. We also demonstrated that whole gDNA is preferable to plasmid isolation prior to sequencing since the latter poses the risk of missing ARGs. In conclusion, although the Flongle showed higher error in ARG variant calling than MinION, it still displays similar performance in real-time for pathogen ID and ARG detection. This, together with its cost-effectiveness, could lead to potential future use in clinical microbiology.

²<https://store.nanoporetech.com/flongle-flow-cell-pack.html>

REFERENCES

- Arroyo Mühr, L. S., Lagheden, C., Hassan, S. S., Kleppe, S. N., Hultin, E., and Dillner, J. (2020). De novo sequence assembly requires bioinformatic checking of chimeric sequences. *PLoS One* 15:e0237455. doi: 10.1371/journal.pone.0237455
- Avershina, E., Shapovalova, V., and Shipulin, G. (2021a). Fighting Antibiotic Resistance in Hospital-Acquired Infections: Current State and Emerging Technologies in Disease Prevention, Diagnostics and Therapy. *Front. Microbiol.* 12:707330. doi: 10.3389/fmicb.2021.707330
- Avershina, E., Sharma, P., Taxt, A. M., Singh, H., Frye, S. A., Paul, K., et al. (2021b). AMR-Diag: neural network based genotype-to-phenotype prediction of resistance towards β -lactams in *Escherichia coli* and *Klebsiella pneumoniae*. *Comp. Struct. Biotechnol. J.* 19, 1896–1906. doi: 10.1016/j.csbj.2021.03.027
- Berbers, B., Ceyssens, P.-J., Bogaerts, P., Vanneste, K., Roosens, N. H. C., Marchal, K., et al. (2020). Development of an NGS-Based Workflow for Improved Monitoring of Circulating Plasmids in Support of Risk Assessment of Antimicrobial Resistance Gene Dissemination. *Antibiotics* 9:503. doi: 10.3390/antibiotics9080503
- Briggs, N., Campbell, S., and Gupta, S. (2021). Advances in rapid diagnostics for bloodstream infections. *Diagn. Microbiol. Infect. Dis.* 99:115219. doi: 10.1016/j.diagmicrobio.2020.115219
- Charalampous, T., Kay, G. L., Richardson, H., Aydin, A., Baldan, R., Jeanes, C., et al. (2019). Nanopore metagenomics enables rapid clinical diagnosis of bacterial

DATA AVAILABILITY STATEMENT

The datasets used and analyzed during the current study are deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB49072.

AUTHOR CONTRIBUTIONS

EA, SF, AT, and RA designed the experiments. EA and RA wrote the manuscript. SF and JA performed experimental work in discussions with AT and RA. EA analyzed the data in discussions with RA. SF and AT edited the manuscript. All authors read and edited the manuscript.

FUNDING

This work was funded by the Regional Research Fund Inland Region (RFFINNL) of the Research Council of Norway (Grant No. 311664 for Rapid-Diag project to RA).

ACKNOWLEDGMENTS

We thank Ørjan Samuelsen, Universitetssykehus Nord-Norge, Tromsø, Norway, for providing the CTX-M positive bacterial strains *E. coli* A2-39 and *K. pneumoniae* A2-37.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.822402/full#supplementary-material>

lower respiratory infection. *Nat. Biotechnol.* 37, 783–792. doi: 10.1038/s41587-019-0156-5

Galata, V., Fehlmann, T., Backes, C., and Keller, A. (2019). PLSDb: a resource of complete bacterial plasmids. *Nucleic Acids Res.* 47, D195–D202. doi: 10.1093/nar/gky1050

Grädel, C., Terrazos Miani, M. A., Barbani, M. T., Leib, S. L., Suter-Riniker, F., and Ramette, A. (2019). Rapid and Cost-Efficient Enterovirus Genotyping from Clinical Samples Using Flongle Flow Cells. *Genes* 10:659. doi: 10.3390/genes10090659

Harstad, H., Ahmad, R., and Bredberg, A. (2018). Nanopore-based DNA sequencing in clinical microbiology: preliminary assessment of basic requirements. *bioRxiv* [Preprint]. doi: 10.1101/382580

Jeck, W. R., Iafrate, A. J., and Nardi, V. (2021). Nanopore Flongle Sequencing as a Rapid, Single-Specimen Clinical Test for Fusion Detection. *J. Mol. Diagn.* 23, 630–636. doi: 10.1016/j.jmoldx.2021.02.001

Khezri, A., Avershina, E., and Ahmad, R. (2021a). Hybrid Assembly Provides Improved Resolution of Plasmids, Antimicrobial Resistance Genes, and Virulence Factors in *Escherichia coli* and *Klebsiella pneumoniae* Clinical Isolates. *Microorganisms* 9:2560. doi: 10.3390/microorganisms9122560

Khezri, A., Avershina, E., and Ahmad, R. (2021b). Plasmid Identification and Plasmid-Mediated Antimicrobial Gene Detection in Norwegian Isolates. *Microorganisms* 9:52. doi: 10.3390/microorganisms9010052

- Leggett, R. M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T. C., Kujawska, M., et al. (2020). Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nat. Microbiol.* 5, 430–442. doi: 10.1038/s41564-019-0626-z
- Li, Y., Yang, X., and Zhao, W. (2017). Emerging Microtechnologies and Automated Systems for Rapid Bacterial Identification and Antibiotic Susceptibility Testing. *SLAS Technol.* 22, 585–608. doi: 10.1177/2472630317727519
- Long, S. W., Linson Sarah, E., Ojeda Saavedra, M., Cantu, C., Davis James, J., Brettin, T., et al. (2017). Whole-Genome Sequencing of Human Clinical *Klebsiella pneumoniae* Isolates Reveals Misidentification and Misunderstandings of *Klebsiella pneumoniae*, *Klebsiella variicola*, and *Klebsiella quasipneumoniae*. *mSphere* 2, e290–e217. doi: 10.1128/mSphereDirect.00290-17
- Metzgar, D., Frinder, M. W., Rothman, R. E., Peterson, S., Carroll, K. C., Zhang, S. X., et al. (2016). The IRIDICA BAC BSI Assay: Rapid, Sensitive and Culture-Independent Identification of Bacteria and Candida in Blood. *PLoS One* 11:e0158186. doi: 10.1371/journal.pone.0158186
- Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., and Levin, E. (2020). Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One* 15:e0227434. doi: 10.1371/journal.pone.0227434
- San Millan, A., Santos-Lopez, A., Ortega-Huedo, R., Bernabe-Balas, C., Kennedy Sean, P., and Gonzalez-Zorn, B. (2015). Small-Plasmid-Mediated Antibiotic Resistance Is Enhanced by Increases in Plasmid Copy Number and Bacterial Fitness. *Antimicrob. Agents Chemother.* 59, 3335–3341. doi: 10.1128/AAC.00235-15
- Taxt, A. M., Avershina, E., Frye, S. A., Naseer, U., and Ahmad, R. (2020). Rapid identification of pathogens, antibiotic resistance genes and plasmids in blood cultures by nanopore sequencing. *Sci. Rep.* 10:7622. doi: 10.1038/s41598-020-64616-x
- Trevas, D., Caliendo, A. M., Hanson, K., Levy, J., and Ginocchio, C. C. (2021). Diagnostic Tests Can Stem the Threat of Antimicrobial Resistance: Infectious Disease Professionals Can Help. *Clin. Infect. Dis.* 72, e893–e900. doi: 10.1093/cid/ciaa1527
- Vasala, A., Hytonen, V. P., and Laitinen, O. H. (2020). Modern Tools for Rapid Diagnostics of Antimicrobial Resistance. *Front. Cell Infect. Microbiol.* 10:308. doi: 10.3389/fcimb.2020.00308
- Villa, L., and Carattoli, A. (2020). “Plasmid Typing and Classification,” in *Horizontal Gene Transfer: Methods and Protocols*, ed. F. de la Cruz (New York, NY: Springer US), 309–321. doi: 10.1007/978-1-4939-9877-7_22
- Votintseva Antonina, A., Bradley, P., Pankhurst, L., del Ojo Elias, C., Loose, M., Nilgiriwala, K., et al. (2017). Same-Day Diagnostic and Surveillance Data for Tuberculosis via Whole-Genome Sequencing of Direct Respiratory Samples. *J. Clin. Microbiol.* 55, 1285–1298. doi: 10.1128/JCM.02483-16
- Xu, Y., Lewandowski, K., Lumley, S., Pullan, S., Vipond, R., Carroll, M., et al. (2018). Detection of Viral Pathogens With Multiplex Nanopore MinION Sequencing: Be Careful With Cross-Talk. *Front. Microbiol.* 9:2225. doi: 10.3389/fmicb.2018.02225

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Avershina, Frye, Ali, Taxt and Ahmad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Optimization of Early Antimicrobial Strategies for Lung Transplant Recipients Based on Metagenomic Next-Generation Sequencing

OPEN ACCESS

Edited by:

Costas C. Papagiannitsis,
University of Thessaly, Greece

Reviewed by:

Marc Finianos,
Charles University, Czechia
Adam Valcek,
Vrije University Brussel, Belgium

*Correspondence:

Yi-ping Wang
297285600@qq.com
Sen Lu
lu.sen.cool@163.com
Ling-ai Pan
panlingai2004@163.com

[†]These authors have contributed
equally to this work and share first
authorship

[‡]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 20 December 2021

Accepted: 18 January 2022

Published: 23 March 2022

Citation:

Zhang XQ, Lei Y, Tan XL, Guo L,
Huang XB, Yang FX, Yu H, Liu XS,
Wang YP, Lu S and Pan LA (2022)
Optimization of Early Antimicrobial
Strategies for Lung Transplant
Recipients Based on Metagenomic
Next-Generation Sequencing.
Front. Microbiol. 13:839698.
doi: 10.3389/fmicb.2022.839698

Xiao-qin Zhang^{1†}, Yu Lei^{1†}, Xiao-li Tan^{2†}, Lu Guo³, Xiao-bo Huang¹, Fu-xun Yang¹,
Hua Yu⁴, Xiao-shu Liu³, Yi-ping Wang^{1*†}, Sen Lu^{1*†} and Ling-ai Pan^{1*†}

¹ Department of Critical Care Medicine, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu, China, ² Genoxor Medical Science and Technology Inc., Taizhou, China, ³ Department of Pulmonary and Critical Care Medicine, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu, China, ⁴ Department of Microbiology Laboratory, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu, China

The management of perioperative antibiotic options after lung transplantation varies widely around the world, but there is a common trend to limit antibiotic use duration. Metagenomic next-generation sequencing (mNGS) has become a hot spot in clinical pathogen detection due to its precise, rapid, and wide detection spectrum of pathogens. Thus, we defined a new antibiotic regimen adjustment strategy in the very early stage (within 7 days) after lung transplantation mainly depending on mNGS reports combined with clinical conditions to reduce the use of antibiotics. To verify the clinical effect of the strategy, we carried out this research. Thirty patients who underwent lung transplantation were finally included, whose information including etiology, antibiotic adjustment, and the effect of our strategy was recorded. Lung transplant recipients in this study were prescribed with initial antibiotic regimen immediately after surgery; their antibiotic regimens were adjusted according to the strategy. According to our study, the entire effectiveness of the strategy was 90.0% (27/30). Besides, a total of 86 samples containing donor lung tissue, recipient lung tissue, and bronchoalveolar lavage fluid (BALF) were obtained in this study; they were all sent to mNGS test, while BALF was also sent to pathogen culture. Their results showed that the positive rate of BALF samples was higher (86.67%) than that of donor's lung tissue (20.0%) or recipient's lung tissue (13.33%) by mNGS test, indicating BALF samples are more valuable than other clinical samples from early postoperative period to guide the early adjustment of antibiotics after lung transplantation. It is effective for mNGS combined with traditional methods and clinical situations to optimize antibiotic regimens in lung transplantation recipients within 7 days after surgery.

Keywords: metagenomic next-generation sequencing, lung tissue, lung transplantation, bronchoalveolar lavage fluid, pathogen detection

INTRODUCTION

Infectious complications remain a significant cause of morbidity and mortality in lung transplant recipients (Burguete et al., 2013; Raskin et al., 2020), who are at increased risk of infection for multiple reasons, including continuous exposure of the allograft to environmental microorganisms and denervation of the lung resulting in impaired cough reflex, dysfunctional mucociliary clearance, impaired lymphatic drainage, and immunosuppression (Costa et al., 2017).

At present, broad-spectrum antibiotics are used to cover the infection after lung transplantation worldwide, and hospital-acquired bacteria are mostly targeted by perioperative antibiotic therapy even if no pathogen colonization occurs (Coiffard et al., 2020). However, this may cause problems of antibiotic abuse and resistance, and also potential liver and kidney injury. Therefore, precise anti-infective treatment during the perioperative period is particularly important.

A study conducted by Coiffard et al. (2020) analyzed worldwide clinical practices in perioperative antibiotic therapy for lung transplantation; they collected data from 99 hospitals in 24 countries, and the result showed that the duration of prophylaxis in this context was heterogeneous but mostly 7 days (33.3%) or shorter (26.3%), or until cultures of the donor and the recipients were reported as negative (12.1%). The antibiotic treatment was almost systematically adapted to the results of the donor samples (97.1%). After 4 days of empirical treatment, if the results of the bacteriological screening were negative, and there was no sign of infection, antibiotics were stopped in 52.5% of the centers. This suggests that the antibiotic regimen in the very early stage (within 7 days) after lung transplantation is very important.

Precise pathogen detection is a prerequisite for precise anti-infection treatment (Zhou et al., 2019). Traditional pathogen testing methods cannot meet clinical needs due to the disadvantages of a long detection cycle and low sensitivity. Metagenomic next-generation sequencing (mNGS) could yield higher sensitivity (Chen et al., 2020; Duan et al., 2021) for pathogen identification and is less affected by prior antibiotic exposure (Miao et al., 2018; Huang et al., 2020), thereby emerging as a promising technology for detecting infectious diseases (Miao et al., 2018; Li et al., 2020; Duan et al., 2021). Two previous studies (Liu et al., 2020; Lian et al., 2021) found that mNGS is effective in detecting pathogens after lung transplantation. mNGS combined with traditional methods and pathogen detection of different samples including donor lung tissue, recipient lung tissue, and BALF immediately after surgery can theoretically be more comprehensive obtaining the background information of the pathogen and making anti-infective treatment more targeted. At the same time, because mNGS is short time-consuming, the results can be obtained quickly, even if the infection occurs in the ultra-early postoperative period. Adjusting the antibiotics according to the pathogen detection results in advance will be more targeted; this approach is similar to preemptive therapy in anti-fungal therapy.

Therefore, we developed a new strategy for super early-stage (within 7 days) antibiotic optimization after lung transplantation:

donor lung tissue, recipient lung tissue, and BALF within 2 h after surgery were all sent to mNGS (BALF samples were also sent to pathogen culture). We gave the patients a basic antibiotic regimen (relatively narrow spectrum) soon after the operation and optimized the regimen according to their clinical indications, and pathogen detection results after pathogen detection reports were obtained.

The purpose of this study is to validate the clinical practical effect of the above antibiotic adjustment strategy; evaluate mNGS combined with traditional methods to detect donor lung tissue, receptor lung tissue, and BALF; and guide the adjustment of early antibacterial strategies, determine whether they can affect the early prognosis of lung transplantation recipients, and provide a clinical reference for anti-infection work in the super early stage post lung transplantation.

MATERIALS AND METHODS

Criteria for Donor Lungs Inclusion and Exclusion

Donor lungs inclusion criteria: (a) Age < 60 years old, smoking history < 20 packs/year. (b) No chest injury. (c) Continuous mechanical ventilation < 1 week. (d) PaO₂ > 300 mmHg (FiO₂ = 100%, PEEP = 5 cmH₂O). (e) X-ray or CT shows that the lung field is relatively clear. (f) There is no abscess secretion in the lung bronchus at all levels through bronchoscopy (Liu et al., 2020).

Donor lungs exclusion criteria: (a) Age > 60 years old, smoking history > 20 packs/year. (b) Chest trauma and lung contusion. (c) Continuous mechanical ventilation > 1 week. (d) PaO₂ < 300 mmHg (FiO₂ = 100%, PEEP = 5 cmH₂O). (e) X-ray or CT shows that the lung field is infected. (f) There are purulent secretions at bronchoscopy in the donor lower airways. (g) The percentage of white blood cells, neutrophils, C-reactive protein, and procalcitonin increases gradually compared with the situation at the onset of the disease. (h) The donor's body temperature is higher than normal. (i) Blood culture is positive.

Data Acquisition and Samples Collection

Patients undergoing lung transplantation at Sichuan Provincial People's Hospital from October 2018 to July 2021 were included in this study, and baseline data (at the time of admission) and data during the hospitalization period were collected. Baseline data mainly include gender, age, date of admission, underlying disease information, basic lung pathogen colonization, and infection information.

After enrollment, we record information about lung transplantation including donor information, single or bilateral lung transplantation, and sampling protocols as follows: BALF samples from recipients were obtained within 2 h after surgery, which must be submitted to traditional pathogen culture and mNGS pathogen detection at the same time. The donor lung tissue samples from the basal segment of the lower lobe with a size of 0.5 cm × 0.5 cm were sent to mNGS test. Lung tissue samples of the recipients from the basal segment of the

lower lobe with a size of 0.5 cm × 0.5 cm were also sent to mNGS test.

Prognostic information on antimicrobial use of the enrolled patients and prognostic information on mechanical ventilation, ICU hospitalization, new-onset infection, and 30-day mortality were recorded after surgery.

mNGS Sequencing and Data Analysis

The samples were stored at 4°C and sent to Genoxor Medical Laboratory for mNGS detection within 24 h. The 1.5-ml microcentrifuge tube with the 0.6-ml (/g) sample, enzyme, and 1.0 g of 0.5-mm glass beads was attached to a horizontal platform on a vortex mixer and agitated vigorously at 2,800–3,200 rpm for 30 min. Then the 0.3-ml sample was separated into new 1.5-ml microcentrifuge tubes, and DNA of BALF samples was extracted using the TIANamp Micro DNA Kit (DP316, Tiangen Biotech) according to the manufacturer's instructions. DNA of lung tissue samples was extracted using the TIANamp Genomic DNA Kit (DP304, Tiangen Biotech) according to the manufacturer's instructions. Then, DNA libraries were constructed through DNA fragmentation, end-repair, adapter ligation, and PCR amplification. Agilent 2100 was used for quality control of the DNA libraries. Library concentration was measured by Qubit 2.0, and sequencing data was pre-quantified by q-PCR. Quality-qualified libraries were sequenced on the NextSeq™ 550DX platform in SE-75 sequencing type according to the manufacturer's instructions.

Raw data were split using bcl2fastq2 software, and the connector sequences and low-quality base sequences were removed using Trimmomatic software to obtain high-quality effective data. Sequences from the human genome were removed using the bowtie2 calibration software. Eventually, sequences that could not be mapped to the human genome were retained and the rest of the sequences were aligned to the microbial genome database, which was constructed using the sequences of bacteria, fungi, archaea, and viruses screened in the NCBI database, covering 16,834 microbes (7,982 bacteria, 7,811 viruses, and 124 parasites).

To enable a comparison between species within the same sample, the number of reads was homogenized. The numbers of reads were normalized with the genome length to calculate their RPK (reads per kilobase), and the species' relative abundance was further calculated based on the RPK.

Antibiotic Regimen Adjustment Strategy

An initial antimicrobial protocol was formulated based on the donor, recipient, and perioperative conditions. The principle of the initial protocol is: cover all the possible pathogens with minimal antimicrobial drugs.

After the traditional pathogen detection and mNGS reports were acquired, the adjustment of the antibiotic regimen was made according to the following conditions:

(I) Maintain the basic antibiotic regimen. (II) Change the basic antibiotic regimen based on the clinical characteristics and pathogen detection results of patients: (II-1) Antimicrobial de-escalation (ADE) treatment or simplified antibiotic regimen

because of no signs of infection; (II-2) There were no new-onset infections but we add some other antibiotics in advance or replace the original antibiotics according to the clinical characteristics and pathogen detection results; (II-3) There were new-onset infections and we change the antibiotics.

Effectiveness Evaluation of Antibiotic Adjustment Strategy

The effectiveness judgment of the antibiotic strategy was made according to the new-onset infections and other clinical indicators, such as MV time and ICU hospitalization. This effectiveness evaluation requires at least two transplant management experts. The effect of the strategy was determined as "Positive," "Negative," and "None," as indicated respectively:

"Positive": The patients have no new-onset infections within 1 week after surgery, or effectively controlled infection after adjusting antibiotic regimen by referring to the mNGS positive results.

"Negative": There were new-onset infections and aggravation, even if the antimicrobial regimen is optimized based on mNGS reports.

"None": The role of mNGS cannot be judged or the patient's infection status cannot be evaluated.

In detail, the judgment criteria for new-onset infections were based on the increase of body temperature after transplantation, continuously increased white blood cells, continuously increased procalcitonin (PCT), deteriorated sputum traits, hemodynamic instability, and CT scans by experienced lung transplant management experts.

Statistical Analysis of Data

Descriptive statistics were computed for the overall sample and stratified by the presence of positive pathogen detected by NGS or bacterial culture positive on BALF samples. All statistical analyses were performed using the EXCEL software. Mean ± standard deviation (SD) or median (interquartile range, IQR) was used for describing the continuous variables.

RESULTS

Clinical Characteristics of Patients in This Study

From October 2018 to July 2021, we completed 32 lung transplants in total, 30 cases of which were included in this study, including 26 men (86.67%) and 4 women (13.33%). The average age was 57.8 ± 1.03 years; the youngest recipient and the oldest recipient were 33 and 70 years old, respectively. Of the 30 lung transplantations, 15 were single-lung transplantations and 15 were bilateral lung transplantations. In this study, the primary disease of lung transplantation patients were interstitial lung disease (ILD) (50.0%), chronic obstructive pulmonary disease (COPD) (46.67%), and silicosis (3.33%). Other demographic details are listed in **Table 1** and clinical characteristics of individuals in **Table 2**.

TABLE 1 | Demographic and clinical characteristics of the 30 patients with lung transplantation.

Patient characteristics	Patients included (n = 30)
Male, n (%)	26 (86.67)
Age (years), mean (SD)	57.8 (1.03)
Underlying condition, n (%)	
ILD	15 (50.0)
COPD	14 (46.67)
Silicosis	1 (3.33)
Type of transplant, n (%)	
Single lung transplantation	15 (50.0)
Bilateral lung transplantation	15 (50.0)
Other parameters (days)	
MV time, median (interquartile range, IQR)	3 (IQR 13-1) (except 1 death case)
Length of stay in ICU, median (interquartile range, IQR)	8 (IQR 13-5) (except 6 death cases)

ILD, interstitial lung disease; COPD, chronic obstructive pulmonary disease; MV, mechanical ventilation; ICU, intensive care unit.

Bronchoalveolar Lavage Fluid Samples Have the Highest Positive Rate Compared With Donor and Recipient Tissues

The donor lung tissue, recipient lung tissue, and BALF samples were submitted for pathogen culture and mNGS according to the study flow chart (Figure 1) after lung transplantation. A total of 86 effective samples were obtained in this study; the BALF test showed the highest positive rate (86.67%, 26/30) by mNGS. Meanwhile, the positive rate of pathogen culture on the BALF was only 36.67% (11/30). In detail, pathogen cultures on nine BALF samples (30%, 9/30) grew out bacteria such as *Staphylococcus aureus*, *Staphylococcus haemolyticus*, *Burkholderia multivorans*, *Acinetobacter ursingii*, and *Klebsiella pneumoniae*; only one BALF sample (0.33%, 1/30) grew fungal (*Candida parapsilosis*). But there were 22 bacteria, 9 fungi, 3 viruses, and 2 mycoplasma detected by mNGS for the same BALF samples, and for the 26 donor and 30 recipient lung tissue samples, there were fewer pathogens tested in their mNGS reports (Figure 2), which may suggest that such samples are not well correlated with infection after lung transplantation. Thus, BALF may be the most appropriate sample for pathogen detection after lung transplantation.

The Pathogenic Spectrum of Bronchoalveolar Lavage Fluid Detected by mNGS in This Study

To clear all possible pathogens after lung transplantation, we used mNGS to make the pathogen spectrum of BALF. According to the mNGS results of BALF (Figure 3), 22 bacteria were detected in 30 samples, including Gram-negative bacteria (66.67%, 20/30), such as *Haemophilus*, *klebsiella*, *Pseudomonas*, *Acinetobacter*, *Burkholderia*, *Bacteroides*, *Stenotrophomonas*, and *Enterobacter*. Gram-positive bacteria (36.67%, 11/30) included *Staphylococcus*,

Enterococcus, *Tropheryma*, *corynebacterium*, and *Actinomyces*. Also, there were fungi (16.67%, 5/30) like *Candida*, *Penicillium*, *Moesziomyces*, and *Pneumocystis*. There were two pathogenic viruses detected in four BALF samples (13.33%, 4/30): *Human cytomegalovirus* and *Human herpesvirus 1*. Some rare pathogens were also discovered such as *Ureaplasma parvum* (0.33%, 1/30) and *Mycoplasma hominis* (0.33%, 1/30). This indicated that Gram-negative bacteria may be the first pathogen to be considered after lung transplantation.

Efficacy of Our Antibiotic Adjustment Strategies in This Study

Of the patients in this study, 50% (15/30) used β -lactamase inhibitor as the only antibiotic; 30.0% (9/30) of patients were administered with combined β -lactamase inhibitor and some other anti-fungal or anti-viral drugs, such as Caspofungin, voriconazole, cotrimoxazole, and ganciclovir, or other anti-microbial drugs such as moxifloxacin and vancomycin. The other 20.0% (6/30) of patients use carbapenems (amopenem or meropenem) that replaced β -lactamase inhibitors.

The basic antibiotic regimen was adjusted according to the pathogen detection reports from pathogen culture and mNGS and also the patient's clinical indications. After the strategy application, 40.0% (12/30) of patients maintained the initial medication regimen because it could cover the pathogen in the comprehensive clinical situation within 7 days after transplantation. The other 60.0% (18/30) of patients adjusted the antibiotic regimen, in which:

We reduced the antibiotic in one patient, whose basic antibiotic plan was piperacillin/tazobactam combined with linezolid. *Pseudomonas aeruginosa* and *Pseudomonas putida* were the main pathogens according to mNGS, so the linezolid was suspended. This patient had no new-onset infection within 7 days after prognosis, and the effect of the strategy was positive.

Eight patients had no new-onset infections within 7 days, but their antibiotic regimen was changed because lung transplant management experts predicted that there will occur infections according to the clinical situation and mNGS reports. Also, the effect of our strategy in the eight cases was all positive.

The remaining 9 patients had progressed infection status, so antibiotics regimens were adjusted for suspected pathogens mainly depending on reports of mNGS. In this study, most cases were supplemented with corresponding antibiotics for a possible fungal infection or viral infection. According to the effect evaluation results, the effect of the strategy in most cases (66.67%, 6/9) was positive; just one case (11.11%) was negative in which infection status cannot be controlled. The effect of the strategy in the last two patients cannot be evaluated because antibiotic adjustment did not refer to their mNGS reports.

According to the comprehensive judgment of professional clinical management experts of lung transplantation, using the antibiotic adjustment strategies after lung transplantation combined with mNGS and pathogen culture, 70.0% (21/30) of patients in this study did not get new-onset infections within 7 days after transplantation; in 6.67% (2/30) of patients, infection reduced; in 13.33% (4/30), infection was effectively controlled;

TABLE 2 | Characteristics of individual transplant recipients.

	Age (years)	Gender	Type of transplant	ECLS	Initial antibiotics	Antibiotic regimen adjusted	Antibiotics after initial protocol changed	Antibiotic regimen adjustment	New-onset infections	MV time (days)	Days of stay in ICU (days)	Effect of the antibiotic adjustment strategy
Patient 1	70	Male	Single	No ECLS	Cefoperazone and sulbactam	Yes	Meropenem	II-3	Yes	2	9	None
Patient 2	69	Male	Bilateral	No ECLS	Piperacillin tazobactam	Yes	Piperacillin tazobactam, caspofungin	II-2	No	2	13	Positives
Patient 3	67	Male	Single	ECMO	Piperacillin tazobactam	Yes	Piperacillin tazobactam, vancomycin	II-3	Yes	15	NA	Positives
Patient 4	57	Male	Single	No ECLS	Cefoperazone and sulbactam	Yes	Meropenem	II-3	Yes	2	7	Positives
Patient 5	49	Male	Bilateral	No ECLS	Cefoperazone and sulbactam	No	Cefoperazone and sulbactam	I	No	1	8	Positives
Patient 6	52	Male	Single	No ECLS	Cefoperazone and sulbactam	Yes	Meropenem	II-2	No	1	10	Positives
Patient 7	50	Female	Single	No ECLS	Cefoperazone and sulbactam	Yes	Cefoperazone and sulbactam, compound sulfamethoxazole	II-2	No	1	13	Positives
Patient 8	67	Male	Single	No ECLS	Cefoperazone and sulbactam	No	Cefoperazone and sulbactam	I	No	3	6	Positives
Patient 9	61	Male	Bilateral	VV-ECMO	Cefoperazone and sulbactam	Yes	Piperacillin tazobactam, caspofungin	II-2	No	10	NA	Positives
Patient 10	54	Male	Single	VV-ECMO	Piperacillin tazobactam, Ganciclovir	No	Piperacillin tazobactam, ganciclovir	I	No	1	11	Positives
Patient 11	33	Male	Bilateral	No ECLS	Piperacillin tazobactam	Yes	Piperacillin tazobactam, voriconazole	II-2	No	1	5	Positives
Patient 12	46	Female	Single	VV-ECMO	Cefoperazone and sulbactam	No	Cefoperazone and sulbactam	I	No	1	7	Positives
Patient 13	64	Male	Single	VV-ECMO	Imipenem, vancomycin, Ganciclovir	No	Imipenem, vancomycin, ganciclovir	I	No	3	NA	Positives
Patient 14	60	Male	Bilateral	VV-ECMO	Cefoperazone and sulbactam, vancomycin	No	Cefoperazone and sulbactam, vancomycin	I	No	12	16	Positives
Patient 15	66	Female	Bilateral	VV-ECMO	Piperacillin tazobactam, vancomycin, Caspofungin	No	Piperacillin tazobactam, vancomycin, Caspofungin	I	No	25	NA	Positives
Patient 16	66	Male	Single	VV-ECMO	Cefoperazone and sulbactam	Yes	Meropenem, daptomycin, caspofungin	II-3	Yes	28	41	None
Patient 17	53	Male	Bilateral	VV-ECMO	Piperacillin tazobactam, ganciclovir	No	Piperacillin tazobactam, ganciclovir	I	No	1	3	Positives
Patient 18	53	Male	Bilateral	VV-ECMO	Cefoperazone and sulbactam	Yes	Cefoperazone and sulbactam, caspofungin	II-2	No	3	5	Positives
Patient 19	49	Male	Bilateral	VV-ECMO	Imipenem, linezolid, caspofungin acetate for injection	No	Imipenem, linezolid, caspofungin acetate	I	No	3	5	Positives
Patient 20	65	Male	Bilateral	VV-ECMO	Clindamycin, aztreonam	No	Clindamycin, aztreonam	I	No	49	NA	Positives
Patient 21	59	Male	Bilateral	VV-ECMO	Piperacillin tazobactam, linezolid	Yes	Piperacillin tazobactam	II-1	No	3	8	Positives
Patient 22	67	Male	Single	VV-ECMO	Piperacillin tazobactam, caspofungin	Yes	Imipenem, Linezolid, caspofungin	II-2	No	6	9	Positives

(Continued)

TABLE 2 | (Continued)

	Age (years)	Gender	Type of transplant	ECLS	Initial antibiotics	Antibiotic regimen adjusted	Antibiotics after initial protocol changed	Antibiotic regimen adjustment	New-onset infections	MV time (days)	Days of stay in ICU (days)	Effect of the antibiotic adjustment strategy
Patient 23	66	Male	Single	VV-ECMO	Cefoperazone and sulbactam	Yes	Cefoperazone and sulbactam, caspofungin	II-2	No	1	4	Positives
Patient 24	58	Male	Bilateral	VV-ECMO	Cefoperazone and sulbactam	Yes	Cefoperazone and sulbactam, moxifloxacin	II-3	Yes	4	8	Positives
Patient 25	64	Female	Single	VV-ECMO	Imipenem, vancomycin	No	Imipenem, vancomycin	I	No	19	21	Positives
Patient 26	61	Male	Bilateral	VV-ECMO	Piperacillin tazobactam, linezolid, isoniazid, ethambutol	Yes	Imipenem, Linezolid, isoniazid, Ethambutol	II-3	Yes	2	5	Positives
Patient 27	57	Male	Bilateral	VV-ECMO	Tegacyclin, caspofungin, linezolid, cefoperazone and sulbactam	Yes	Tegacyclin, caspofungin, linezolid, Sulbactam	II-3	Yes	NA	NA	Positives
Patient 28	62	Male	Bilateral	VV-ECMO	Linezolid, meropenem	Yes	Linezolid, meropenem, caspofungin	II-3	Yes	14	14	Positives
Patient 29	53	Male	Single	VV-ECMO	Moxifloxacin, caspofungin, piperacillin tazobactam	No	Moxifloxacin, caspofungin, piperacillin tazobactam	I	No	2	5	Positives
Patient 30	37	Male	Single	No ECLS	Moxifloxacin, caspofungin, piperacillin tazobactam	Yes	Imipenem, polymyxin, sulbactam, voriconazole	II-3	Yes	28	28	Negative

ECLS, extracorporeal life support; VV ECMO, venous extracorporeal membrane oxygenation; ICU, intensive care unit; MV, mechanical ventilation; NA, not applicable.

in 3.33% (1/30), infection was aggravated and the strategies failed; and in 6.67% (2/30), the antibiotic regimen was adjusted but not referring to the mNGS reports. In conclusion, the entire effectiveness of this strategy was 90.0% (27/30). Therefore, our antibiotic adjustment strategy has the prospect of further application in the management of clinical lung transplantation.

DISCUSSION

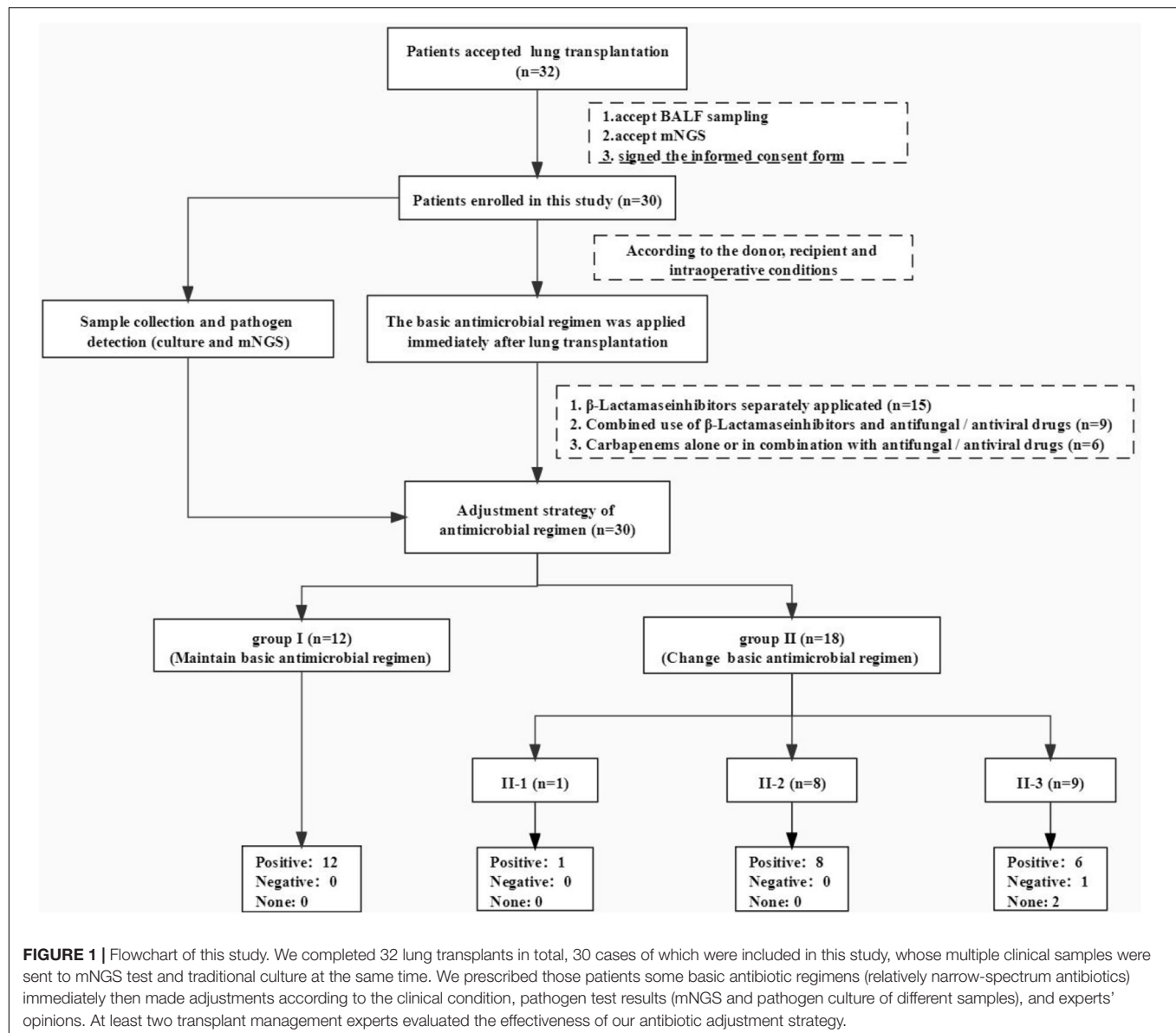
In this study, we made a new antibiotic regimen adjustment strategy in the very early stage (within 7 days) after lung transplantation mainly depending on pathogenic test results of different clinical samples and patients' clinical indications. We confirmed that the effectiveness of this strategy was 90.0% through 30 lung transplantation cases.

Half of the infectious episodes following lung transplantation occur in the first 30 days (Aguilar-Guisado et al., 2007), but it is not clear which period occurred infectious within 30 days. The data related to the infection in the very early stage (within 7 days) after lung transplantation is limited (Parada et al., 2010). In our study, of the 30 patients enrolled in this study, 10 (33.33%, 10/30) patients developed a pulmonary infection within 1 week after the operation. Among them, 7 patients (70%, 7/10) developed the infection within 72 h after the operation, indicating that early postoperative pulmonary infection (within 7 days) cannot be ignored.

A previous study (Liu et al., 2020) suggested that the colonized bacteria in different parts of the lung are inconsistent and there is no association between the colonized bacteria in donor lungs and the short-term outcome of lung transplantation patients. In our study, the positive rate by mNGS test of BALF samples was higher (86.67%) than that of donor lung tissue (20.0%) or recipient lung tissue (13.33%). It indicates BALF samples from the early postoperative period are more valuable than other clinical samples to guide the early adjustment of antibiotics after lung transplantation. Meanwhile, it is the first study about the value of pathogen detection of different clinical samples in the adjustment of the anti-infection scheme after lung transplantation.

The pathogens of pulmonary infection in the early stage after transplantation are complex (Nosotti et al., 2018; José et al., 2020). Our study suggested that the proportion of Gram-negative bacteria and Gram-positive bacteria detected by mNGS in the BALF of lung transplant recipients was 66.67 and 33.33%, respectively. The significance is that Gram-negative bacteria are still the pathogen that needs to be covered first when the pathogen is unknown. The positive rate of fungi (33.33%) also suggested that the prevention of fungal infection after lung transplantation is worthy of attention.

The postoperative antimicrobial strategies adopted by many centers are often for the sake of wide coverage of all possible pathogens (Coiffard et al., 2020). For example, the scheme (Ling et al., 2016) of "meropenem, vancomycin, caspofungin, and ganciclovir were used simultaneously" was adopted by



Guangzhou Institute of Respiratory Diseases in China. In our study, we developed a new strategy for super early-stage (within 7 days) antibiotic optimization after lung transplantation: donor lung tissue, recipient lung tissue, and BALF within 2 h after surgery were all sent to mNGS (BALF also sent to pathogen culture); we gave the patients a basic antibiotic regimen (relatively narrow spectrum) and optimized the regimen according to their clinical indications and pathogen detection results after a relevant pathogen detection report was obtained, of which the effective rate is 90% (27/30). It suggested that not only is our antibiotic optimization strategy effective, but also, it is not necessary to use so many antibiotics which may increase the chance of multi-drug-resistant bacteria, the economic burden of patients, and also the liver and kidney function injury of patients.

The average mechanical ventilation time of the patients enrolled in this study was 3 (IQR 13-1) days (excluding 1 death),

the average ICU hospitalization time was 8 (IQR 13-5) days (excluding 6 deaths), and the 30-day mortality was 80.0% (24/30). Liu's study on the application of mNGS in the management of lung transplantation in Wuxi People's Hospital showed that 17 patients were included in the study, and the average mechanical ventilation time was 73.3 (SD = 67.6) h, the average ICU hospitalization time was 112.6 (SD = 59.8) h, and the average hospital stay was 41.2 (SD = 46.9) days. Our strategy reduced the unnecessary postoperative antibiotics of wide coverage while it did not increase the incidence of infection, MV time, and ICU length of stay.

In this study, the strategy of adjusting antibiotics after lung transplantation based on traditional pathogen detection combined with mNGS is effective for the prevention and control of new-onset infections. However, due to the high difficulty of lung transplantation and less clinical practice in China,

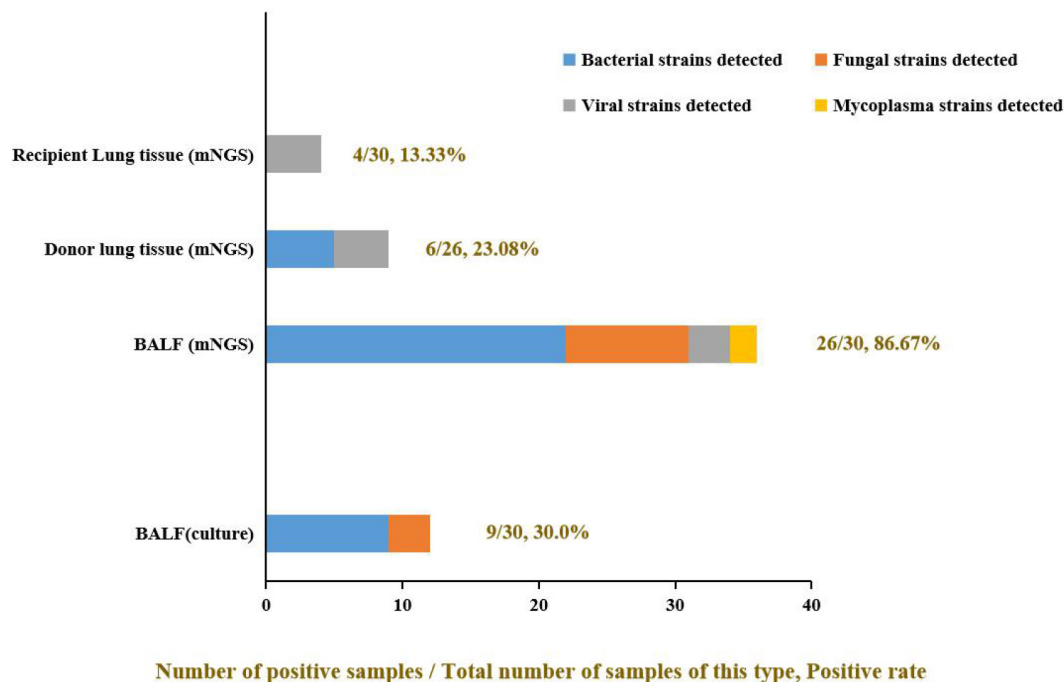


FIGURE 2 | The positive rate of different samples and the proportion of pathogens detected. The positive rate of mNGS detection was significantly higher than in the pathogen culture of the same BALF samples (mNGS vs. culture: 86.67% vs. 30.0%). Some viruses and mycoplasma were detected by mNGS but not by culture method. Among all three kinds of clinical samples, BALF samples showed the highest clinical pathogen diagnostic value: only virus detected in recipient lung tissue; only virus and mycoplasma detected in donor lung tissue, but most kinds of pathogens were detected in BALF and the bacteria made up the majority.

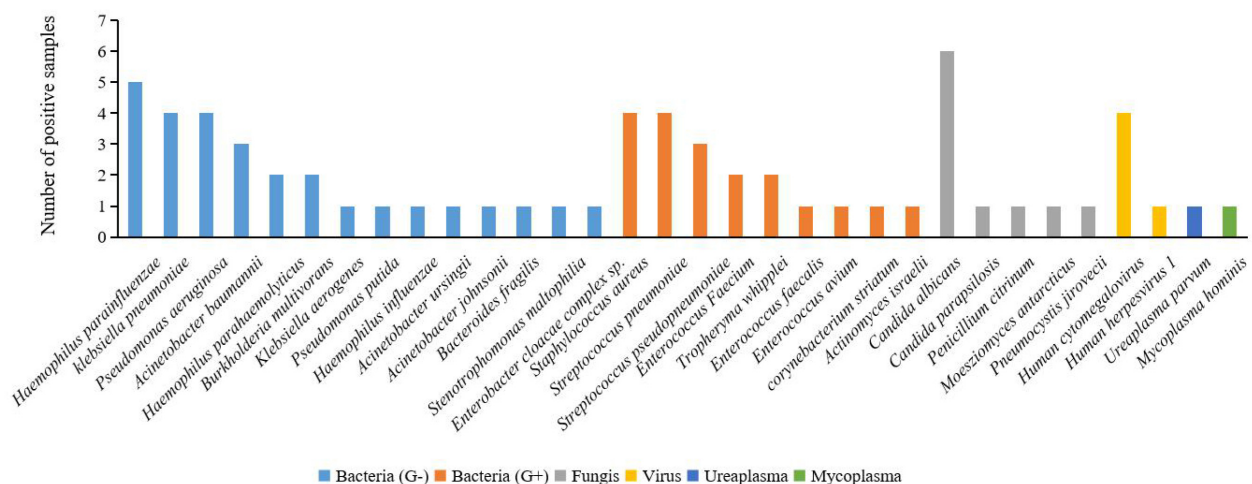


FIGURE 3 | Positive pathogen spectrum detected in BALF samples by mNGS in this study. Of all the bacteria detected by mNGS in BALF samples, G-bacteria (*Haemophilus parainfluenzae*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*) made up the majority. Several rare pathogens were also detected in the BALF samples, such as *Penicillium citrinum*, *Ureaplasma parvum*, and *Mycoplasma hominis*.

the number of cases in this study is small (30 cases totally). Therefore, there is no distinction between the types of primary diseases (such as pulmonary interstitial disease or COPD) and lung transplantation methods (such as single lung or bilateral lung transplantation). In addition, because of the particularity of lung transplantation, a specific control group cannot be set

in this study, but the effect of our strategy can be judged by comparing it with other studies. Finally, how this antibiotic optimization strategy works in different clinical situations and how it affects the long-term prognosis of lung transplantation patients should be verified by the clinical practice of larger population applications.

CONCLUSION

It is effective to apply mNGS combined with traditional pathogen detecting methods and clinical features to optimize antibiotic regimens in lung transplantation recipients within 7 days after surgery. For patients who accepted lung transplantation, BALF samples from the early postoperative period are more valuable than other clinical samples to guide the early adjustment of antibiotics.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Sichuan Provincial People's Hospital. The

patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XZ and SL designed and performed the study. XZ, XT, YL, HY, and LP collected and analyzed data. XT, XZ, and LP wrote the manuscript. SL, LP, and YW supervised the clinical research and revised the manuscript. All authors approved the final manuscript.

FUNDING

This work was supported by the Nursery Project of Sichuan Provincial People's Hospital (Grant No. 30805031031) and the Special Foundation for Young Medical Research of China International Medical Exchange Foundation (Grant No. Z-2018-35-1902).

REFERENCES

- Aguilar-Guisado, M., Givaldá, J., Ussetti, P., Ramos, A., Morales, P., Blanes, M., et al. (2007). Pneumonia after lung transplantation in the RESITRA Cohort: a multicenter prospective study. *Am. J. Transpl.* 7, 1989–1996. doi: 10.1111/j.1600-6143.2007.01882.x
- Burguete, S. R., Maselli, D. J., Fernandez, J. F., and Levine, S. M. (2013). Lung transplant infection. *Respirology* 18, 22–38. doi: 10.1111/j.1440-1843.2012.02196.x
- Chen, H., Yin, Y., Gao, H., Guo, Y., Dong, Z., Wang, X., et al. (2020). Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response. *Clin. Infect. Dis.* 71, S416–S426. doi: 10.1093/cid/ciaa1516
- Coiffard, B. A.-O., Prud'Homme, E., Hraiech, S., Cassir, N., Le Pavec, J., Kessler, R., et al. (2020). Worldwide clinical practices in perioperative antibiotic therapy for lung transplantation. *BMC Pulm. Med.* 20:109. doi: 10.1186/s12890-020-1151-9
- Costa, J., Benvenuto, L. J., and Sonett, J. R. (2017). Long-term outcomes and management of lung transplant recipients. *Best Pract. Res. Clin. Anaesthesiol.* 31, 285–297. doi: 10.1016/j.bpa.2017.05.006
- Duan, H., Li, X., Mei, A., Li, P., Liu, Y., Li, X., et al. (2021). The diagnostic value of metagenomic next-generation sequencing in infectious diseases. *BMC Infect. Dis.* 21:62.
- Huang, Z. D., Zhang, Z. J., Yang, B., Li, W. B., Zhang, C. J., Fang, X. Y., et al. (2020). Pathogenic detection by metagenomic next-generation sequencing in osteoarticular infections. *Front. Cell Infect. Microbiol.* 10:471. doi: 10.3389/fcimb.2020.00471
- José, R. J., Periseleris, J. N., and Brown, J. S. (2020). Opportunistic bacterial, viral and fungal infections of the lung. *Medicine* 48, 366–372. doi: 10.1016/j.mpmed.2020.03.006
- Li, Y., Sun, B., Tang, X., Liu, Y. L., He, H. Y., Li, X. Y., et al. (2020). Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in critically ill patients. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 369–374. doi: 10.1007/s10096-019-03734-5
- Lian, Q. Y., Chen, A., Zhang, J. H., Guan, W. J., Xu, X., Wei, B., et al. (2021). High-throughput next-generation sequencing for identifying pathogens during early-stage post-lung transplantation. *BMC Pulm. Med.* 21:348. doi: 10.1186/s12890-021-01723-z
- Ling, S., Xiaoqing, L., and Yimin, L. (2016). Perioperative management of lung transplantation. *Chin. J. Crit. Care Med.* 4, 205–210.
- Liu, D., Zhang, J., Wu, B., Liu, F., Ye, S., Wang, H., et al. (2020). Impact of donor lung colonized bacteria detected by next-generation sequencing on early post-transplant outcomes in lung transplant recipients. *BMC Infect. Dis.* 20:689. doi: 10.1186/s12879-020-05393-w
- Miao, Q., Ma, Y., Wang, Q., Pan, J., Zhang, Y., Jin, W., et al. (2018). Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin. Infect. Dis.* 67, S231–S240. doi: 10.1093/cid/ciy693
- Nosotti, M., Tarsia, P., and Morlacchi, L. C. (2018). Infections after lung transplantation. *J. Thorac. Dis.* 10, 3849–3868.
- Parada, M. T., Alba, A., and Sepúlveda, C. (2010). Early and late infections in lung transplantation patients. *Transpl. Proc.* 42, 333–335. doi: 10.1016/j.transproceed.2009.12.002
- Raskin, J., Vanstapel, A., Verbeken, E. K., Beeckmans, H., Vanaudenaerde, B. M., Verleden, S. A.-O., et al. (2020). Mortality after lung transplantation: a single-centre cohort analysis. *Transpl. Int.* 33, 130–141. doi: 10.1111/tri.13540
- Zhou, X., Su, L. X., Zhang, J. H., Liu, D. W., and Long, Y. (2019). Rules of anti-infection therapy for sepsis and septic shock. *Chin. Med. J.* 132, 589–596. doi: 10.1097/CM9.0000000000000101

Conflict of Interest: XT is employed by Genoxor Medical Science and Technology Inc., Taizhou, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Lei, Tan, Guo, Huang, Yang, Yu, Liu, Wang, Lu and Pan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Successful Intra- but Not Inter-species Recombination of *msr(D)* in *Neisseria subflava*

Tessa de Block^{1*}, Natalia González^{1†}, Saïd Abdellati¹, Jolein Gyonne Elise Laumen^{1,2}, Christophe Van Dijck^{1,2}, Irith De Baetselier¹, Dorien Van den Bossche¹, Sheeba S. Manoharan-Basil¹ and Chris Kenyon^{1,3}

¹Department of Clinical Sciences, Institute of Tropical Medicine, Antwerp, Belgium, ²Laboratory of Medical Microbiology, Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium, ³Department of Medicine, University of Cape Town, Cape Town, South Africa

OPEN ACCESS

Edited by:

Alberto Antonelli,
University of Florence,
Italy

Reviewed by:

Gianluca Morróni,
Marche Polytechnic University,
Italy

William Shafer,
Emory University,
United States

*Correspondence:

Tessa de Block
tdeblock@itg.be

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 15 January 2022

Accepted: 02 March 2022

Published: 30 March 2022

Citation:

de Block T, González N, Abdellati S,
Laumen JGE, Van Dijck C, De
Baetselier I, Van den Bossche D,
Manoharan-Basil SS and
Kenyon C (2022) Successful Intra-
but Not Inter-species Recombination
of *msr(D)* in *Neisseria subflava*.
Front. Microbiol. 13:855482.
doi: 10.3389/fmicb.2022.855482

Resistance acquisition *via* natural transformation is a common process in the *Neisseria* genus. Transformation has played an important role in the emergence of resistance to many antimicrobials in *Neisseria gonorrhoeae* and *Neisseria meningitidis*. In a previous study, we found that currently circulating isolates of *Neisseria subflava* had acquired an *msr(D)* gene that has been found to result in macrolide resistance in other bacteria but never found in *Neisseria* species before. To determine if this resistance mechanism is transferable among *Neisseria* species, we assessed if we could transform the *msr(D)* gene into other commensal and pathogenic *Neisseria* under low dose azithromycin pressure. Intraspecies recombination in commensal *N. subflava* was confirmed with PCR and resulted in high-level macrolide resistance. Whole-genome sequencing of these transformed strains identified the complete uptake of the *msr(D)* integration fragment. Sequence analysis showed that a large fragment of DNA (5 and 12 kb) was transferred through a single horizontal gene transfer event. Furthermore, uptake of the *msr(D)* gene had no apparent fitness cost. Interspecies transformation of *msr(D)* from *N. subflava* to *N. gonorrhoeae* was, however, not successful.

Keywords: horizontal gene transfer, *msr(D)*, transformation, *Neisseria subflava*, *Neisseria gonorrhoeae*, macrolide resistance

INTRODUCTION

Transformation is one of the genetic recombination methods *Neisseria gonorrhoeae* has used to acquire resistance to every class of antimicrobials used to treat it (Unemo and Shafer, 2014). *Via* this process, *Neisseria* species are able to take up environmental DNA and incorporate it into their chromosomes (Hamilton and Dillard, 2006; Rotman and Seifert, 2014). *Neisseria* species preferably take up DNA from closely related species, especially those that use the same DNA uptake sequence (DUS) for transformation (Duffin and Seifert, 2010). An important consequence of transformation is the transfer of resistance-associated DNA fragments from commensal *Neisseria* towards pathogenic *Neisseria* (Nakayama et al., 2016; Wadsworth et al., 2018). Commensal *Neisseria* are important members of a healthy oral microbiome and hence are present in all humans (Liu et al., 2015; Tedijanto et al., 2018). This high prevalence means they are more

likely to be exposed to antimicrobials used for any indication (bystander selection; Kenyon et al., 2021). As a result, commensal *Neisseria* are particularly at risk for developing antimicrobial resistance (AMR) to commonly used antimicrobials. Along these lines, recent studies have found alarmingly high minimum inhibitory concentrations (MIC) values for fluoroquinolones, macrolides and β -lactams in commensal *Neisseria* (Dong et al., 2020; Laumen et al., 2021b). Studies have confirmed that transformation of DNA from commensal *Neisseria* has played an important role in the genesis of resistance to a number of classes of antimicrobials in pathogenic *Neisseria*: macrolides (*mtrR*, *mtrCDE*, *rplD* and *rplY*; Wadsworth et al., 2018; Manoharan-Basil et al., 2021), β -Lactams (*penA*; Bowler et al., 1994; Ito et al., 2005), sulphonamides (*folP*) and fluoroquinolones (*gyrA*; Unemo and Shafer, 2014; Chen et al., 2020).

An additional pathway used by the pathogenic *Neisseria* to acquire AMR has been the uptake of whole genes from other species. Examples of these are the acquisition of the *tetM*, *ermB/C* and *bla_{TEM}* genes that confer resistance to tetracyclines, macrolides and β -Lactams, respectively (Roberts et al., 1999; Unemo and Shafer, 2014). In a previous study, we identified the recent acquisition of a new ribosomal protection protein (MsrD) in *N. subflava* as a novel resistance mechanism in *Neisseria* (de Block et al., 2021). The *msr(D)* gene is part of the antibiotic resistance ATP-binding cassette F (ABC-F) protein family. The four classes of Msr proteins (A, C, D and E) operate as ribosomal protection proteins by displacing macrolides and ketolides from the ribosome. Macrolide resistance conferring *msr* genes have been identified in various species of *Streptococcus*, *Staphylococcus*, *Enterococcus*, *Pseudomonas* and *Corynebacterium* (Sharkey et al., 2016; Dinis, 2017). Complementation studies in these species have clearly established that *msr(D)* has a powerful effect on macrolide MICs (Daly et al., 2004; Nunez-Samudio and Chesneau, 2013; Zhang et al., 2016; Iannelli et al., 2018; Fostier et al., 2020). In our previous study, we found that the *msr(D)* in *N. subflava* was likely derived from the macrolide efflux genetic assembly (MEGA) element in

Streptococcus pneumoniae, with whom it shared 100% sequence homology (de Block et al., 2021). As already described in other species, we found that the presence of the *msr(D)* gene in *N. subflava* was associated with higher azithromycin MICs (Iannelli et al., 2018; Fox et al., 2021).

In the current paper, we aimed to address four questions that emerged from the previous research: (1) Can the *msr(D)* gene be transformed into other strains of *N. subflava*? (2) If so, does this occur at the same insertion site? (3) Does uptake of *msr(D)* confer a fitness cost? (4) Can the *msr(D)* gene be transformed into *N. gonorrhoeae*?

MATERIALS AND METHODS

Intra- and Interspecies Transformation in Plates

The strains used in this experiment were all isolated from oropharyngeal swabs taken from men who have sex with men (MSM) attending our Sexually Transmitted Infections (STI) clinic in Antwerp, Belgium in 2019 (Laumen et al., 2021b). Nine *N. subflava* strains containing the *msr(D)* gene (azithromycin MIC ≥ 24 mg/L) were used as donor and two *N. subflava* and one *N. gonorrhoeae* strains without this gene were used as recipients (MIC < 1 mg/L; **Table 1**). Genomic DNA was extracted using the EpiCentre® kit. The DNA concentration (ng/ μ l) was determined using the NanoDrop® ND-1000 spectrophotometer (Thermo Scientific). One hundred μ l of three different donor pools (P1–P3), each containing a mix of three donor DNA extracts of *N. subflava* (150 ng/ μ l), were separately mixed with 100 μ l (4.0 McFarland) of the mid-log phase growth (6h) of three recipient strains: (i) *N. subflava* (ITM_Ns_9/1: azithromycin MIC 3), (ii) *N. subflava* (ITM_Ns_45/1: azithromycin MIC 6 mg/L and (iii) *N. gonorrhoeae* (ITM_Ng_38/1: azithromycin MIC 0.19 mg/L; **Table 2**). Azithromycin concentration of 1.5 \times MIC was added as a stress factor. Control experiments did

TABLE 1 | Characteristics of strains used in this study.

Isolate*	Species	Source of isolate	AZM MIC (mg/L)	<i>msr(D)</i>	Function in current experiment	Experiment
ITM_Ns_3/2	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 1 (P1)	Transformation in Plates (Table 2) Morbidostat
ITM_Ns_27/1	<i>N. subflava</i>	Laumen 2021	24	Present	Donor Pool 1 (P1)	Transformation in Plates (Table 2) Morbidostat MIC stability
ITM_Ns_36/1	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 1 (P1)	Transformation in Plates (Table 2) Morbidostat
ITM_Ns_9/2	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 2 (P2)	Transformation in Plates (Table 2)
ITM_Ns_27/2	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 2 (P2)	Transformation in Plates (Table 2)
ITM_Ns_29/1	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 2 (P2)	Transformation in Plates (Table 2)
ITM_Ns_36/2	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 3 (P3)	Transformation in Plates (Table 2)
ITM_Ns_41/1	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 3 (P3)	Transformation in Plates (Table 2) Template for PCR transformation
ITM_Ns_49/1	<i>N. subflava</i>	Laumen 2021	>256	Present	Donor Pool 3 (P3)	Transformation in Plates (Table 2)
ITM_Ns_9/1	<i>N. subflava</i>	Laumen 2021	3	Absent	Recipient	Transformation in Plates (Table 2) MIC stability Growth curve
ITM_Ns_45/1	<i>N. subflava</i>	Laumen 2021	6	Absent	Recipient	Transformation in Plates (Table 2)
ITM_Ng_38/1	<i>N. gonorrhoeae</i>	Laumen 2021	0.19	Absent	Recipient	Transformation in Plates
ITM_Ng_21.021	<i>N. gonorrhoeae</i>	Clinical sample	1	Absent	Recipient	Morbidostat
WHO-X	<i>N. gonorrhoeae</i>	Reference strain	0.004	Absent	Recipient	PCR transformation

*Ns: *N. subflava*; Ng: *N. gonorrhoeae*.

TABLE 2 | MIC values after intraspecies (*N. subflava*) transformation in plates.

		Intraspecies recombination in <i>N. subflava</i>					
		Recipient 1 (ITM_Ns_9/1)			Recipient 2 (ITM_Ns_45/1)		
	Sample		MIC AZM ¹	<i>msr(D)</i> qPCR	Sample	MIC AZM ¹	<i>msr(D)</i> qPCR
Transformation experiments	Donor DNA <i>N. subflava</i> P1	ITM_Ns_9/1_P1 ²	>256	Pos ³	ITM_Ns_45/1_P1 ²	>256	Pos
	Donor DNA <i>N. subflava</i> P2	ITM_Ns_9/1_P2	>256	Pos	ITM_Ns_45/1_P2	>256	Pos
	Donor DNA <i>N. subflava</i> P3	ITM_Ns_9/1_P3	>256	Pos	ITM_Ns_45/1_P3	>256	Pos
Control experiments	AB control	ITM_Ns_9/1_AZM	4	Neg	ITM_Ns_45/1_AZM	1	Neg
	DNA control	ITM_Ns_9/1_DNA	1.5	Neg	ITM_Ns_45/1_DNA	3	Neg
	Growth control	ITM_Ns_9/1	1.5	Neg	ITM_Ns_45/1	2	Neg

¹Post-transformation minimum inhibitory concentrations of azithromycin (MIC AZM) in mg/L.

²Transformed stains subjected to whole-genome sequencing.

³Pos: positive confirmation of *msr(D)* transformation by qPCR and Neg: negative confirmation.

not contain azithromycin and/or DNA. The reaction mixtures were plated on blood agar and incubated for 48 h. One colony from each blood agar culture was selected for azithromycin MIC determination E-test gradient strips (bioMérieux, France). All the experiments were conducted at 36°C and 6% CO₂.

Inter-species Transformation in Morbidostat

The transformation experiment was performed in a NGmorbidostat. The construction, optimisation and use of the NGmorbidostat have been described in detail elsewhere (Verhoeven et al., 2019; Laumen et al., 2021a). In brief, the NGmorbidostat is a bioreactor that measures bacterial growth *via* optical density measurements and is used to assess the evolution of antimicrobial resistance (AMR) over time within a constant temperature (35°C–36°C) and CO₂ range (5.5%–6%). In this experiment, we only used the incubator and turbidity measurement functions with the programme MATLAB, to record the growth rate of *N. gonorrhoeae* (The Math Works, Inc. MATLAB, version R2015b).

The experiment was conducted in four flasks with a total volume of 15 ml in each of gonococcal (GC) broth supplemented with (1%) IsoVitaleX, henceforth referred as GC medium. The conditions were as: (1) 1.5× MIC azithromycin + DNA from *msr(D)* containing *N. subflava*, (2) 1.5× MIC azithromycin, (3) DNA from *msr(D)* containing *N. subflava* and (4) GC medium (Supplementary Figure 1). To achieve this, firstly we added 200 µl of *N. gonorrhoeae* (ITM_Ng_21.021 with Azithromycin MIC 1 mg/L; Table 1) at 4.0–5.0 McFarland in all four flasks. After 6 h, the growth curve reached the mid-log phase and 100 µl of HLR-Azithromycin DNA from pool 1 of *N. subflava* (150 ng/µl) was added to flasks 1 and 3. At the same time point, azithromycin was added to a final concentration of 1.5 mg/L in flasks 1 and 2. After 24 h, 7.5 ml of the old medium was replaced by fresh medium and an additional 100 µl of HLR-Azithromycin DNA from *N. subflava* (150 ng/µl; flasks 1 and 3) and 1.5 mg/L of azithromycin of the ITM_Ng_21.021 was added (flasks 1 and 2; Table 1). This process was repeated daily for 7 days, after which the azithromycin concentration was increased to 3 mg/L for another 7 days.

Inter-species Transformation With *msr(D)*-DUS DNA Fragment

Msr(D) was PCR amplified from *N. subflava* isolate ITM_Ns_41/1 (Table 1) using primers containing a AT-DUS tag to facilitate inter-species transformation to *N. gonorrhoeae*, forward primer (5'-GAT GCC GTC TGA ACA AAT GAT AAC TGA GG-3') and reverse primer (5'-GAA TCA ATA CTG ACC AGC GAC-3'). This amplification was carried out as a touchdown PCR: the initial denaturation consisted of 5 min at 95°C, followed by amplification for 10 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 3 min. The next stage consisted of 35 cycles, lasting 5 more seconds at each cycle, at 94°C for 30 s, 60°C for 30 s and 72°C for 3 min. A final extension step was carried at 72°C for 7 min. The PCR fragment size was analysed on an agarose gel. The concentration of the amplicon was determined using the NanoDrop® ND-1000 spectrophotometer (Thermo Scientific). The PCR product was used for transformation using the 'Transformation in plates' methodology as described above with 100 µl (150 ng/µl) as DNA donor.

Confirmation of *msr(D)* Transformation With qPCR

Presence or absence of the *msr(D)* gene in transformant strains were confirmed using quantitative PCR (qPCR). The DNA of the recipient strains was extracted using the EpiCentre® kit. The primers used to amplify the internal region of the *msr(D)* (637–934) were as: Forward (5'-GCG GAG GAA AAG CGA AAA C-3') and Reverse (5'-ACA GAG CCT TAT CCC CAA ATAC-3'). The master mix was composed by 10× EHF PCR buffer (5 µl), 2 mM dNTPs (7 µl), 5 µM Primer Forward (3 µl), 5 µM Primer Reverse (3 µl), 3.5 U/µl EHF Taq Polymerase (0.5 µl), Rnase free water (21.5 µl) and DNA (10 µl). The qPCR protocol consisted of an initial denaturation stage at 95°C during 5 min followed by amplification for 45 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 3 min. This step was followed by the final stage consisting of a single cycle of 72°C for 7 min. The specificity of the amplicon was confirmed by conducting melting point analyses.

Assessment of Fitness Cost Based on MIC Stability

To test the stability of the transformed *N. subflava*, a single colony of HLR-azithromycin *N. subflava* strain (ITM_Ns_27/1; azithromycin MIC of 24 µg/ml) and a single colony of one transformant strain of *N. subflava* (ITM_Ns_9/1; azithromycin MIC of 256 µg/ml) were subcultured every 24 h in blood agar plates without additional azithromycin for 7 days, similar to the one described in O'Regan et al. (2010). The azithromycin MICs were tested daily on a single colony from each plate with E-tests (Table 3).

Evaluation of Fitness Cost in Transformants by Growth Curves Rate Variance

The NGmorbidityostat was used to compare the growth curves of *N. subflava* recipient and transformant strain. In a total volume of 15 ml of GC broth supplemented with 1% IsoVitalax (BD BBL™) for each experiment, 100 µl of a 4.0 McFarland suspension in PBS of *N. subflava* recipient (ITM_Ns_9/1) or *N. subflava* transformant (ITM_Ns_9/1transformant) strain was added in triplicate. The growth curves were assessed for 18 h, *via* measurement of optical density every 20 min. Difference in growth curves was assessed *via* analysis with R (R Core Team, 2019) package 'growthcurver' (Sprouffske and Wagner, 2016) with the data obtained from the NGmorbidityostat. R was also used to perform the t-test on the samples to confirm or deny or null hypothesis and to obtain the value of *p* (Supplementary Figure 2).

Whole-Genome Sequencing

For whole-genome sequencing (WGS) analysis, the following samples were chosen: (i) DNA recipients after transformation (ITM_Ns_9/1, ITM_Ns_45/1) and (ii) Transformation in morbidityostat ITM_Ng_21.021 (Time points 1, 7 and 14 for AZM+DNA and day 14 for the controls). Genomic DNA was extracted using the MasterPure Complete DNA and RNA Purification Kit (Epicentre, Madison, Wisconsin, United States) and suspended in nuclease-free water. Indexed paired-end libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA, United States) and sequenced on an Illumina MiSeq instrument (Illumina, San Diego, CA, United States). Data are available in GenBank: <https://www.ncbi.nlm.nih.gov/sra/PRJNA794044>. Processed Illumina reads were *de novo* assembled with Shovill (v1.0.4; <https://github.com/tseemann/shovill>) which uses SPAdes (v3.14.0) using the following parameters:

```
--trim --depth 150 --opts --isolate (Prjibelski et al., 2020). The quality of the contigs was verified with Quast (v5.0.2; Gurevich et al., 2013) followed by annotation using Prokka (v1.14.6; Seemann, 2014). WGS assemblies of the donor (ITM_Ns_3/2, ITM_Ns_27/1 and ITM_Ns_36/1) and recipient strains (ITM_Ns_9/1, ITM_Ns_45/1 and ITM_Ns_38/1) were available from a previous study by our group and included in the comparative analysis (de Block et al., 2021). BLAST Ring Image Generator (BRIG) was used for genome comparison (Alikhan et al., 2011). Mauve (Darling et al., 2004) was used to align contigs and MEGAX (Kumar et al., 2018) was used to align DNA fragments. Percent sequence identity of DNA fragments was calculated using Muscle (https://www.ebi.ac.uk/Tools/msa/muscle/, version 3.8.31).
```

RESULTS

Horizontal Gene Transfer of *msr(D)* From Commensal *Neisseria*

Intra-species Transformation of *msr(D)* on Agar Plates

After 48 h of exposure to each of the three pools of high-level resistance (HLR)-azithromycin DNA (donor) on agar plates, both *N. subflava* recipient strains (ITM_Ns_9/1 and ITM_Ns_45/1; Table 1) attained an azithromycin MIC >256 mg/L (*n*=6; Table 2). These isolates are henceforth referred to as transformants. There was no increase in azithromycin MIC in the control experiments. To confirm if the uptake of the *msr(D)* was successful in these six transformants, the presence of *msr(D)* was confirmed using qPCR (Table 2). One transformant strain of each recipient was used for WGS.

Inter-species Transformation of *msr(D)* on Agar Plates

In the three experiments where *N. gonorrhoeae* was used as recipient, the azithromycin MIC did not increase following incubation with the three donor DNA pools. qPCRs confirmed that *msr(D)* was not taken up by *N. gonorrhoeae* in any of these experiments (ct value >30 or NA).

Transformation of *Neisseria gonorrhoeae* in the NGmorbidityostat

Differences were noted in the azithromycin MIC trajectories in the four flasks (Figure 1). The azithromycin MIC of the *N. gonorrhoeae* recipient increased by day 5 in the flask containing DNA + azithromycin (condition 1) However, the

TABLE 3 | Observation of the azithromycin MIC evolution in ITM_Ns_27/1 donor and ITM_Ns_9/1_P1 transformant strain (both *Neisseria subflava*) after serial subculturing in plates with absence of azithromycin as stress factor.

Isolate	Day 1 (AZM MIC) (mg/L)	Day 2 (AZM MIC) (mg/L)	Day 3 (AZM MIC) (mg/L)	Day 4 (AZM MIC) (mg/L)	Day 5 (AZM MIC) (mg/L)	Day 6 (AZM MIC) (mg/L)	Day 7 (AZM MIC) (mg/L)
ITM_Ns_27/1	24	16	16	32	16	12	16
ITM_Ns_9/1	>256	>256	>256	>256	>256	>256	>256

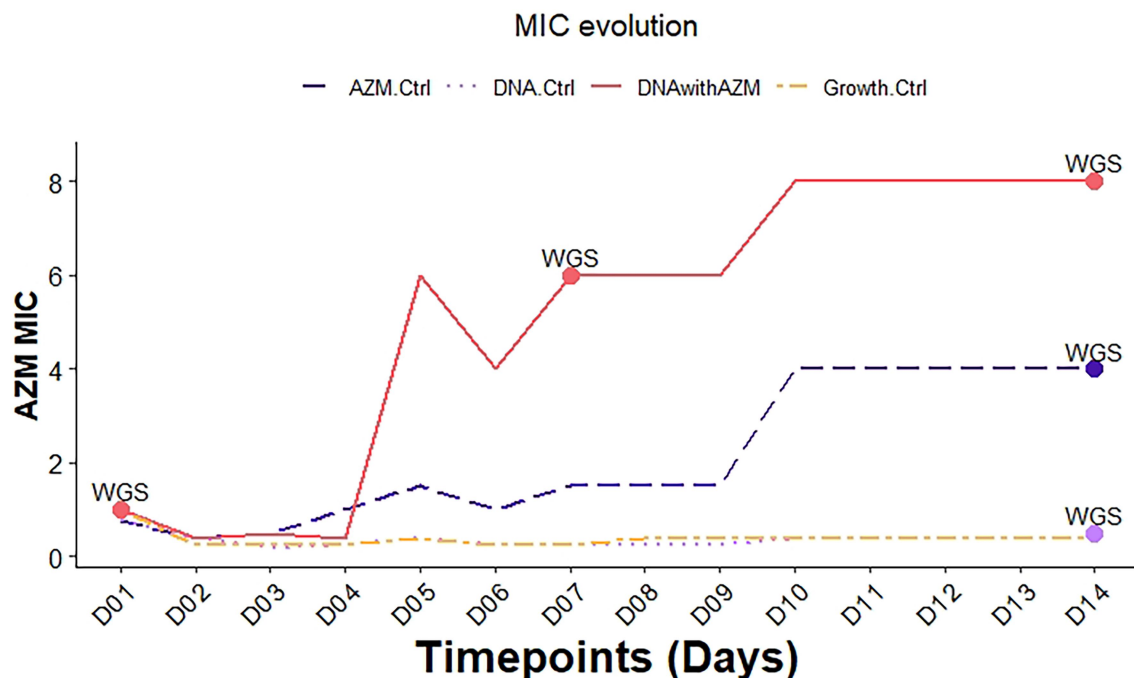


FIGURE 1 | Azithromycin MIC evolution of *Neisseria gonorrhoeae* in the morbidostat transformation experiment in different conditions. (AZM.Ctrl: azithromycin control (condition 2); DNA.Ctrl: DNA control (condition 3); DNAwithAZM: DNA with azithromycin (condition 1); Growth.Ctrl: Growth control (condition 4); WGS—whole-genome sequencing, D01—day 1, etc., MIC—minimal inhibitory concentrations). The time points when samples were subjected to WGS are indicated with a dot.

qPCR of *msr(D)* remained negative in all samples. WGS of samples from day 7 and day 14 revealed a well-known mutation previously linked to macrolide resistance in *N. gonorrhoeae*: C2611T (*Escherichia coli* numbering) in the 23S rRNA gene.

WGS of the *N. gonorrhoeae* recipient in the azithromycin control on day 14 (condition 2) revealed that the recipient acquired the recently described macrolide resistance-associated mutation (RAM) G70D in the 50S ribosomal protein L4 (*rplD*; Ma et al., 2020; Laumen et al., 2021a).

There was no increase in azithromycin MICs of the *N. gonorrhoeae* recipient in the DNA control (condition 3) and the growth control (condition 4).

Transformation of *Neisseria gonorrhoeae* With *msr(D)* PCR Product

There was no increase in the azithromycin MIC of the *N. gonorrhoeae* recipient strain after incubation on agar plates for 48h with a dsDNA fragment containing *msr(D)* and a DUS. qPCR analysis confirmed that the *msr(D)* gene was not acquired by the recipient (ct value >30).

Fitness Cost of Transformants

MIC Stability

There was no significant decline in the azithromycin MIC value in the transformant or donor strain during 7 days of subculturing (Table 3).

Growth Rate

There was no statistically significant difference obtained in the growth rate ratios between recipient (ITM_Ns_9/1, mean: 0.64) and transformant strain (ITM_Ns_9/1_P1, mean: 0.67; value of *p*: 0.3673; Supplementary Figure 2).

Whole-Genome Sequencing of *msr(D)* Transformants

Two PCR-confirmed *msr(D)*-transformant *N. subflava* strains (ITM_Ns_45/1_P1 and ITM_Ns_9/1_P1; Table 2) were subjected to WGS to identify the exact integration site of *msr(D)*. DNA sequences including the upstream (6,725bp) and downstream (9,134bp) region of *msr(D)* with a maximum total length of 17,803bp were extracted for further analyses. Donor (ITM_Ns_3/2, ITM_Ns_27/1 and ITM_Ns_36/1), recipient (ITM_Ns_45/1 and ITM_Ns_9/1) and transformant (ITM_Ns_9/1_P1 and ITM_Ns_45/1_P1) DNA sequences were aligned. This alignment revealed the acquisition of the *msr(D)* gene at the same site (GCATA-acquisition of *msr(D)*-ATTGA) in the chromosome in both recipients, 32bp downstream of a DUS sequence (Figure 2). Genome comparison of donor, recipient and transformant revealed that the transformants had acquired a new *msr(D)*-containing DNA fragment, which originated from the donor, and was not present in the recipient (Figure 3).

A more global alignment conducted in Mauve illustrates the chromosomal organisation around the acquired *msr(D)* in the transformant ITM_Ns_45/1_P1 compared to the recipient strain

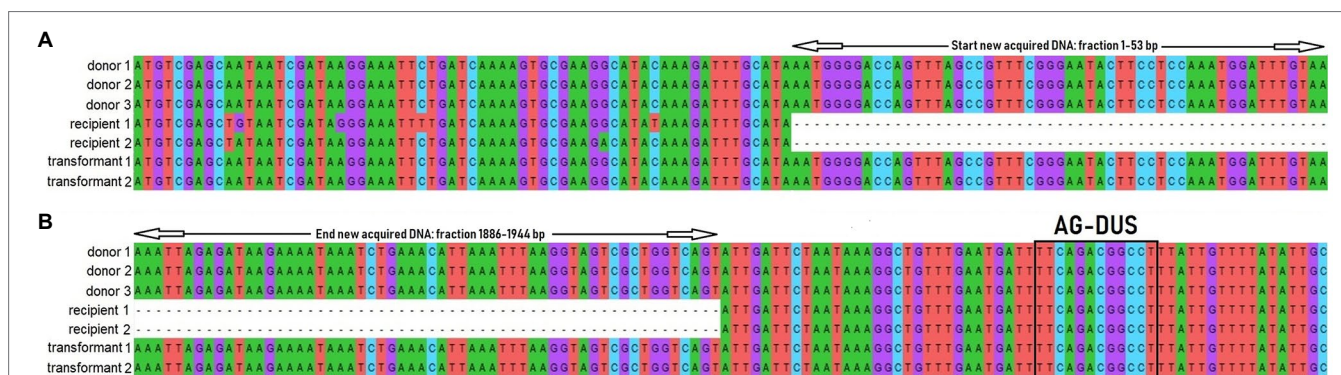


FIGURE 2 | Fragment of DNA sequence alignment of the start (A) and end (B) point (black triangle) of the integration of the new acquired DNA fragment containing *msr(D)*. AG-DUS 31 bp upstream *msr(D)* is indicated with a black box. Transformation of recipient strains (recipient 1; ITM_Ns_9/1, recipient 2; ITM_Ns_9/1) with donor DNA containing *msr(D)* (donor 1; ITM_Ns_3/2, donor 2; ITM_Ns_27/1 and donor 3; ITM_Ns_36/1) resulted in transformant 1 (ITM_Ns_9/1_P1) and 2 (ITM_Ns_45/1_P1).

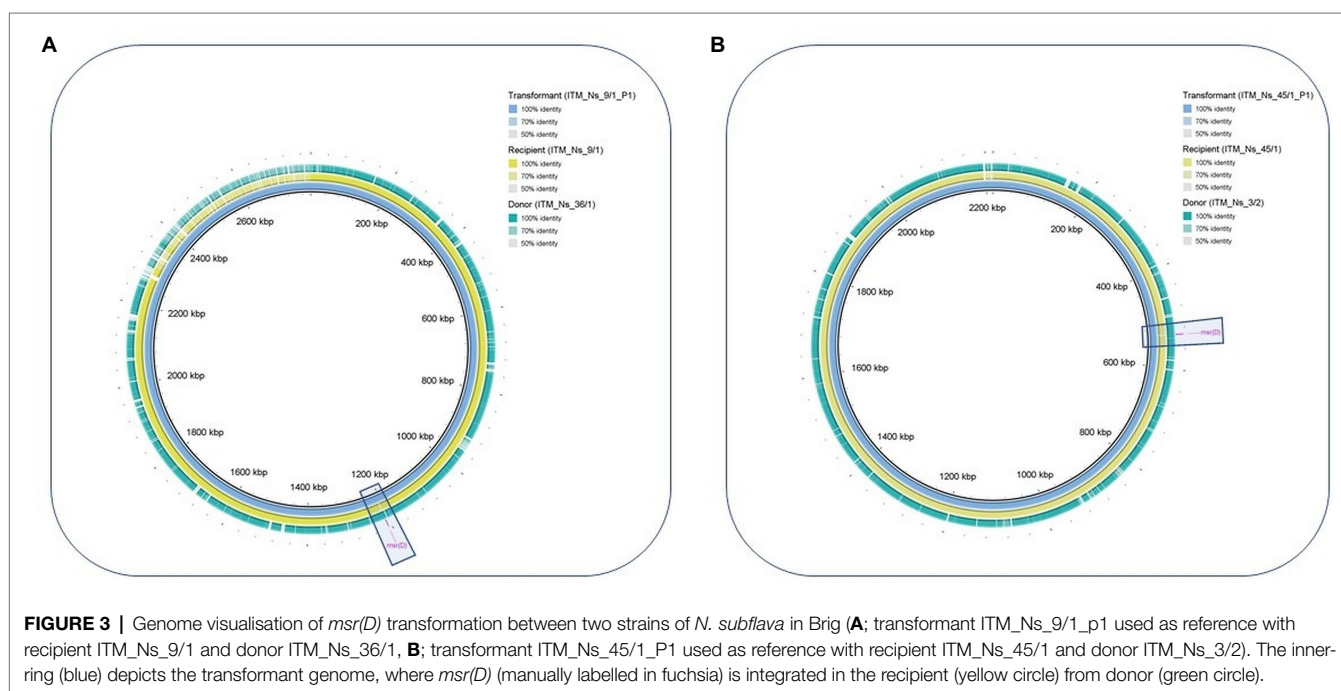


FIGURE 3 | Genome visualisation of *msr(D)* transformation between two strains of *N. subflava* in Brig (A; transformant ITM_Ns_9/1_P1 used as reference with recipient ITM_Ns_9/1 and donor ITM_Ns_36/1; B; transformant ITM_Ns_45/1_P1 used as reference with recipient ITM_Ns_45/1 and donor ITM_Ns_3/2). The inner-ring (blue) depicts the transformant genome, where *msr(D)* (manually labelled in fuchsia) is integrated in the recipient (yellow circle) from donor (green circle).

(Figure 4). SNP analysis revealed that transformant ITM_Ns_9/1_P1 had taken up a larger DNA fragment than transformant ITM_Ns_45/1_P1 (Figure 2; Table 4; Supplementary Table 1). The length between the first and last SNP of transformants compared to the recipient strains was 12,033bp and 5,113bp for transformants ITM_Ns_9/1_P1 and ITM_Ns_45/1_P1, respectively. The acquired DNA extended from upstream of *msr(D)* (7,234bp in ITM_Ns_9/1_P1; 2,883bp in ITM_Ns_45/1_P1) to downstream of *msr(D)* (3,335bp in ITM_Ns_9/1_P1, 766bp in ITM_Ns_45/1_P1; Supplementary Table 1).

The uptake-fragment of transformant ITM_Ns_9/1_P1 showed high similarity to ITM_Ns_36/1 (97.57% identical to ITM_Ns_3/2, 92.78% to ITM_Ns_27/1 and 99.96% to ITM_Ns_36/1) and the uptake-fragment of transformant ITM_Ns_45/1_P1

showed high similarity to ITM_Ns_3/2 (99.92% identical to ITM_Ns_3/2, 92.41% to donor ITM_Ns_27/1 and 98.43% to donor ITM_Ns_36/1; Table 4). These data suggest that for both transformants, the complete *msr(D)* containing fragment was taken up from a single (but different) donor in a single transformation event (ITM_Ns_3/2 as the donor for ITM_Ns_45/1 and ITM_Ns_36/1 for ITM_Ns_9/1_P1).

DISCUSSION

We studied the intra- and inter-species transformability of the resistance conferring *msr(D)* gene in *Neisseria* spp. We found that intraspecies transformation in commensals under

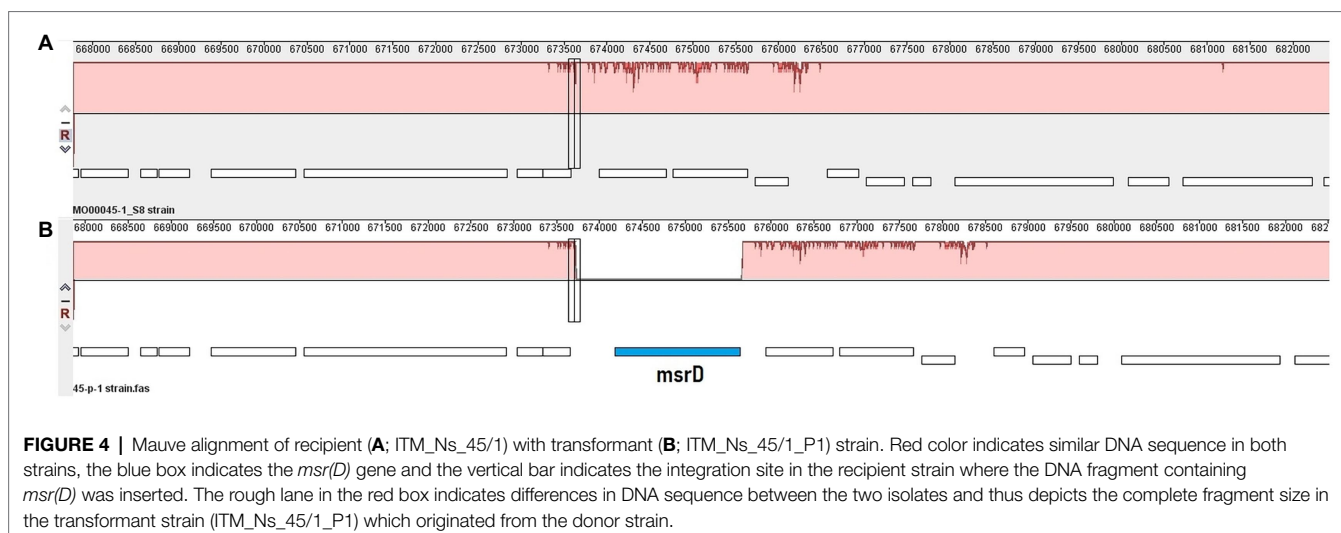


TABLE 4 | Characteristics of integrated DNA fraction in transformant ITM_Ns_9/1_P1 and ITM_Ns_45/1_P1.

Transformant strain	Integrated DNA fraction			% Identical to donor strain		
	Upstream <i>msr(D)</i>	Downstream <i>msr(D)</i>	Complete length	ITM_Ns_3/2	ITM_Ns_27/1	ITM_Ns_36/1
ITM_Ns_9/1_P1	7,234bp	3,335bp	12,033bp	97.57	92.78	99.96
ITM_Ns_45/1_P1	2,883bp	766bp	5,113bp	99.92	92.41	98.43

azithromycin pressure in *N. subflava* was very efficient. Azithromycin triggered the integration of *msr(D)* into strains of *N. subflava* with low level azithromycin resistance (3–6 mg/l). The *msr(D)* gene could be acquired without any apparent fitness cost and was universally associated with an elevation of azithromycin MICs to >256 mg/L. We did not conclusively establish that *msr(D)* is responsible for macrolide resistance in *N. subflava*. This was, however, not one of the study aims as this has been clearly established for a range of gram negative and positive bacterial species (Daly et al., 2004; Nunez-Samudio and Chesneau, 2013; Zhang et al., 2016; Iannelli et al., 2018; Fostier et al., 2020).

In a previous study, we found that nine out of 11 clinical *N. subflava* strains had the *msr(D)* gene integrated in the same place in the genome (de Block et al., 2021). The complete integrated DNA sequence originates from the MEGA element in *S. pneumoniae*. The integration in *N. subflava* was located 32bp downstream of a DUS sequence, suggesting that this DUS enhances the transformation efficiency. WGS of transformant *N. subflava* strains in the current study revealed that the chromosomal integration of the *msr(D)* gene was integrated into the same position in the genome as the donor strains. The complete fragment size in the recipients included up- and downstream regions of *msr(D)* with a total length of 5 and 12 kb, respectively. Thus, a DNA insert up to 12 kb can be transformed into the cell and integrated into the chromosome in a single event. Other studies have found similar sized transformation events in *Neisseria* spp. (Chen et al., 2020). A previous core genome MLST analysis revealed that the *msr(D)*

gene was present in different clusters of clinical isolates of *N. subflava* (de Block et al., 2021). This implies that horizontal gene transfers such as transformation either took place on more than one occasion, or that the *msr(D)* has been taken up and lost in sub-lineages. This suggests that single transformation events of *msr(D)* could also take place *in vivo*.

The MEGA element in *S. pneumoniae* contains both the *msr(D)* gene (which is responsible for displacing bound macrolides) and *mef(A)* which codes an efflux pump that pumps the displaced macrolide out of the cell. Together, these genes belong to the two-gene efflux transport system of the ATP-Binding Cassette (ABC) superfamily and are responsible for type M resistance to macrolides (Iannelli et al., 2018). In *N. subflava*, the *mef(A)* is truncated and likely non-functional (de Block et al., 2021). This suggests that another efflux pump may be used by *N. subflava* to expel the dissociated macrolide. This function is may be carried out by the mtrCDE efflux pump. Interestingly, the *N. subflava*s used in this study all contained the K823E *mtrD* mutant which is known to enhance the ability of the mtrCDE pump to export macrolides (Lyu et al., 2020).

Although it has been proven that interspecies recombination is successful between commensals and pathogenic *Neisseria in vitro*, we were unable to transform *msr(D)* into *N. gonorrhoeae* (Chen et al., 2020). There are a number of possible explanations for this finding. Firstly, the chromosomal organisation around *msr(D)* is very similar in the *N. subflava* donor and recipient strains but is divergent to *N. gonorrhoeae* strains (Supplementary Figures 3, 4). We have previously established

that the core genome of the strains used in this study varies considerably between *N. subflava* and *N. gonorrhoeae* (De Block et al., 2021). This could affect efficient chromosomal integration of *msr(D)* and explain why interspecies transformation between *N. subflava* and *N. gonorrhoeae* was not successful (Qvarnstrom and Swedberg, 2006). Secondly, the relative frequency of the 12-bp DUS sequences varies considerably between *N. subflava* and *N. gonorrhoeae*. The 5'-ATGCCGTCTGAA-3' DUS is more prevalent in *N. gonorrhoeae*, whereas the 5'-AGGCCGTCTGAA-3' DUS is more prevalent in *N. subflava* (Supplementary Figure 4; Berry et al., 2013). These differences in the relative frequency of DUS-subtypes have been shown to influence the probability of transformation (Duffin and Seifert, 2010). This provided the rationale for using dsDNA fragments containing *msr(D)* combined with the predominant *N. gonorrhoeae* DUS for the transformation experiments. However, this approach did not result in transformation. Thirdly, the differential DNA methylation pattern between species of *Neisseria* may result in the uptake of *msr(D)* containing DNA from *N. subflava* being toxic to *N. gonorrhoeae* but not *N. subflava* (Kim et al., 2020). Finally, the failure to transform *msr(D)* into *N. gonorrhoeae* may be due to limitations in our experimental approach. Although we used three different experimental approaches to transform *msr(D)* into *N. gonorrhoeae*, we did so in a limited number of strains. Furthermore, while we have previously been able to conduct successful transformation experiments with two of these strains of *N. gonorrhoeae* using the same experimental protocol, we did not include positive controls in the current experiments (Abdellati et al., 2019). These limitations mean that we cannot conclude that the *msr(D)* gene could not be transformed into *N. gonorrhoeae*. A further limitation of our study is the crude methods we used to measure the fitness cost associated with the acquisition of the *msr(D)* gene.

Another transformation pathway to evaluate in a future study is the transformability between different commensal strains. It may be possible to transform *msr(D)* from *N. subflava* to another commensal *Neisseria* species, such as *N. lactamica*, which is then able to transform the *msr(D)* in *N. gonorrhoeae*

or *N. meningitidis*. For example, *N. lactamica* is known to be an efficient AMR donor to *N. meningitidis* (Chen et al., 2020).

Our study showed that intraspecies transformation of *msr(D)* under azithromycin pressure is very efficient within *N. subflava*. We were unable to transform *msr(D)* into *N. gonorrhoeae*. The limitations noted above mean that we cannot exclude the possibility of this occurring in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/>, PRJNA794044.

AUTHOR CONTRIBUTIONS

SA conducted the wet laboratory experiments. TB and NG conducted the bioinformatic analysis and wrote the first draft. CK, SA and SM-B conceptualized the study. SA, CD, JL, SM-B, IB, DB and CK reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by SOFI 2021 grant—"Preventing the Emergence of untreatable STIs via radical Prevention, (PRESTIP).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.855482/full#supplementary-material>

REFERENCES

- Abdellati, S., Verhoeven, E., Irith, D. B., Crucitti, T., and Kenyon, C. (2019). *Transfer of High-Level Macrolide Resistance in Neisseria Gonorrhoeae*. London: BMJ Publishing Group Ltd. P668.
- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Berry, J. L., Cehovin, A., McDowell, M. A., Lea, S. M., and Pelicic, V. (2013). Functional analysis of the interdependence between DNA uptake sequence and its cognate ComP receptor during natural transformation in *Neisseria* species. *PLoS Genet.* 9:1004014. doi: 10.1371/journal.pgen.1004014
- Bowler, L. D., Zhang, Q.-Y., Riou, J.-Y., and Spratt, B. G. (1994). Interspecies recombination between the penA genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: natural events and laboratory simulation. *Journal of Bacteriology* 176, 333–337. doi: 10.1128/jb.176.2.333-337.1994
- Chen, M., Zhang, C., Zhang, X., and Chen, M. (2020). Meningococcal quinolone resistance originated from several commensal *Neisseria* species. *Antimicrob Agents Chemother* 64, e01494–19. doi: 10.1128/AAC.01494-19
- Daly, M. M., Doktor, S., Flamm, R., and Shortridge, D. (2004). Characterization and prevalence of MefA, MefE, and the associated *msr (D)* Gene in *Streptococcus pneumoniae* Clinical Isolates. *J. Clin. Microbiol.* 42, 3570–3574. doi: 10.1128/JCM.42.8.3570
- Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704.tion
- de Block, T., Laumen, J. G. E., Van Dijk, C., Abdellati, S., Baetselier, I. De, Manoharan-Basil, S. S., et al. (2021). Wgs of commensal *Neisseria* reveals acquisition of a new ribosomal protection protein (Msrd) as a possible explanation for high level azithromycin resistance in Belgium. *Pathogens* 10:384. doi:10.3390/pathogens10030384.
- Dinos, G. P. (2017). The macrolide antibiotic renaissance. *Br. J. Pharmacol.* 174, 2967–2983. doi: 10.1111/bph.13936
- Dong, H. V., Pham, L. Q., Nguyen, H. T., Nguyen, M. X. B., Nguyen, T. V., May, F., et al. (2020). Decreased cephalosporin susceptibility of Oropharyngeal *Neisseria* species in antibiotic-using men who have sex with men in Hanoi, Vietnam. *Clin. Infect. Dis.* 70, 1169–1175. doi: 10.1093/cid/ciz365
- Duffin, P. M., and Seifert, H. S. (2010). DNA uptake sequence-mediated enhancement of transformation in *Neisseria gonorrhoeae* is strain dependent. *J. Bacteriol.* 192, 4436–4444. doi: 10.1128/JB.00442-10

- Fostier, C. R., Singh, S., and Hunt, J. F. (2020). ABC-F translation factors: from antibiotic resistance to immune response. 1–32. doi: 10.1002/1873-3468.13984
- Fox, V., Santoro, F., Pozzi, G., and Iannelli, F. (2021). Predicted transmembrane proteins with homology to Mef (A) are not responsible for complementing *mef* (A) deletion in the *mef* (A)–*msr* (D) macrolide efflux system in *Streptococcus pneumoniae*. *BMC. Res. Notes* 14, 1–7. doi: 10.1186/s13104-021-05856-6
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hamilton, H. L., and Dillard, J. P. (2006). Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol. Microbiol.* 59, 376–385. doi: 10.1111/j.1365-2958.2005.04964.x
- Iannelli, F., Santoro, F., Santagati, M., Docquier, J. D., Lazzeri, E., Pastore, G., et al. (2018). Type M resistance to macrolides is due to a two-gene efflux transport system of the ATP-binding cassette (ABC) superfamily. *Front. Microbiol.* 9, 1–9. doi: 10.3389/fmicb.2018.01670
- Ito, M., Deguchi, T., Mizutani, K.-S., Yasuda, M., Yokoi, S., Ito, S.-I., et al. (2005). Emergence and spread of *Neisseria gonorrhoeae* clinical isolates harboring mosaic-like structure of penicillin-binding protein 2 in Central Japan. *Antimicrob. Agents Chemother.* 49, 137–143. doi: 10.1128/AAC.49.1.137-143.2005
- Kenyon, C., Laumen, J., and Manoharan-Basil, S. (2021). Choosing new therapies for gonorrhoea: we need to consider the impact on the pan-neisseria genome. A viewpoint. *Antibiotics* 10:515. doi: 10.3390/antibiotics10050515
- Kim, W. J., Higashi, D., Goytia, M., Rendón, M. A., Pilligua, M., Bronnimann, M., et al. (2020). Commensal *Neisseria* kill *Neisseria gonorrhoeae* through a DNA-dependent mechanism in mice. *Cell Host Microbe* 26, 228–239. doi: 10.1016/j.chom.2019.07.003
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Laumen, J. G. E., Manoharan-Basil, S. S., Verhoeven, E., Abdellati, S., De Baetselier, I., Crucitti, T., et al. (2021a). Molecular pathways to high-level azithromycin resistance in *Neisseria gonorrhoeae*. *J. Antimicrob. Chemother.* 76, 1752–1758. doi: 10.1093/jac/dkab084
- Laumen, J. G. E., Van Dijk, C., Abdellati, S., Manoharan-Basil, S. S., De Baetselier, I., Martiny, D., et al. (2021b). Markedly reduced azithromycin and ceftriaxone susceptibility in commensal neisseria species in clinical Samples from Belgian men who have sex with men. *Clin. Infect. Dis.* 72, 363–364. doi: 10.1093/cid/ciaa565
- Liu, G., Tang, C. M., and Exley, R. M. (2015). Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus. *Microbiology* 161, 1297–1312. doi: 10.1099/mic.0.000086
- Lyu, M., Moseng, M. A., Reimche, J. L., Holley, C. L., Dhulipala, V., Su, C. C., et al. (2020). Cryo-EM structures of a gonococcal multidrug efflux pump illuminate a mechanism of drug recognition and resistance. *MBio* 11, e00996–e00920. doi: 10.1128/mBio.00996-20
- Ma, K. C., Mortimer, T. D., Duckett, M. A., Hicks, A. L., Wheeler, N. E., Sánchez-Busó, L., et al. (2020). Increased power from conditional bacterial genome-wide association identifies macrolide resistance mutations in *Neisseria gonorrhoeae*. *Nat. Commun.* 11, 5374. doi: 10.1038/s41467-020-19250-6
- Manoharan-Basil, S. S., Laumen, J. G. E., Van Dijk, C., De Block, T., De Baetselier, I., and Kenyon, C. (2021). Evidence of horizontal gene transfer of 50S ribosomal genes *rplB*, *rplD*, and *rplY* in *Neisseria gonorrhoeae*. *Front. Microbiol.* 12, 1–17. doi: 10.3389/fmicb.2021.683901
- Nakayama, S. I., Shimuta, K., Furubayashi, K. I., Kawahata, T., Unemo, M., and Ohnishi, M. (2016). New ceftriaxone- and multidrug-resistant *Neisseria gonorrhoeae* strain with a novel mosaic *penA* gene isolated in Japan. *Antimicrob. Agents Chemother.* 60, 4339–4341. doi: 10.1128/AAC.00504-16
- Nunez-Samudio, V., and Chesneau, O. (2013). Functional interplay between the ATP binding cassette *Msr(D)* protein and the membrane facilitator superfamily *Mef(E)* transporter for macrolide resistance in *Escherichia coli*. *Res. Microbiol.* 164, 226–235. doi: 10.1016/j.resmic.2012.12.003
- O'Regan, E., Quinn, T., Frye, J. G., Pagès, J. M., Porwollik, S., Fedorka-Cray, P. J., et al. (2010). Fitness costs and stability of a high-level ciprofloxacin resistance phenotype in *Salmonella enterica* serotype enteritidis: reduced infectivity associated with decreased expression of *Salmonella* pathogenicity island 1 genes. *Antimicrobial agents and chemotherapy* 54, 367–374. doi: 10.1128/AAC.00801-09
- Prijbelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Curr. Protoc. Bioinform.* 70:e102. doi: 10.1002/cpbi.102
- Qvarnstrom, Y., and Swedberg, G. (2006). Variations in gene organization and DNA uptake signal sequence in the *folP* region between commensal and pathogenic *Neisseria* species. *BMC microbiology* 6:11. doi: 10.1186/1471-2180-6-11
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Roberts, M. C., Chung, W. O., Roe, D., Xia, M., Marquez, C., Borthagaray, G., et al. (1999). Erythromycin-resistant *Neisseria gonorrhoeae* and oral commensal *Neisseria* spp. carry known rRNA methylase genes. *Antimicrob. Agents Chemother.* 43, 1367–1372. doi: 10.1128/aac.43.6.1367
- Rotman, E., and Seifert, H. S. (2014). The genetics of *Neisseria* species. *Annu. Rev. Genet.* 48, 405–431. doi: 10.1146/annurev-genet-120213-092007
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sharkey, L. K. R., Edwards, T. A., and O'Neill, A. J. (2016). ABC-F proteins mediate antibiotic resistance through ribosomal protection. *MBio* 7, 1–10. doi: 10.1128/mBio.01975-15
- Sprouffske, K., and Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinform.* 17, 17–20. doi: 10.1186/s12859-016-1016-7
- Tedijanto, C., Olesen, S. W., Grad, Y. H., and Lipsitch, M. (2018). Estimating the proportion of bystander selection for antibiotic resistance among potentially pathogenic bacterial flora. *Proc. Natl. Acad. Sci. U. S. A.* 115, E11988–E11995. doi: 10.1073/pnas.1810840115
- Unemo, M., and Shafer, W. M. (2014). Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: past, evolution, and future. *Clin. Microbiol. Rev.* 27, 587–613. doi: 10.1128/CMR.00010-14
- Verhoeven, E., Abdellati, S., Nys, P., Laumen, J., De Baetselier, I., Crucitti, T., et al. (2019). Construction and optimization of a 'NG Morbidostat' - An automated continuous-culture device for studying the pathways towards antibiotic resistance in *Neisseria gonorrhoeae*. *F1000Research* 8:1343. doi: 10.12688/f1000research.18861.2
- Wadsworth, C. B., Arnold, B. J., Sater, M. R. A., and Grad, Y. H. (2018). Azithromycin resistance through interspecific acquisition of an epistasis-dependent efflux pump component and transcriptional regulator in *Neisseria gonorrhoeae*. *mBio* 9, e01419–e01418. doi: 10.1128/mBio.01419-18
- Zhang, Y., Tatsuno, I., Okada, R., Hata, N., Matsumoto, M., Isaka, M., et al. (2016). Predominant role of *msr(D)* over *mef(A)* in macrolide resistance in *Streptococcus pyogenes*. *Microbiology* 162, 46–52. doi: 10.1099/mic.0.000206

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 de Block, González, Abdellati, Laumen, Van Dijk, De Baetselier, Van den Bossche, Manoharan-Basil and Kenyon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Co-expression Mechanism Analysis of Different Tachyplesin I-Resistant Strains in *Pseudomonas aeruginosa* Based on Transcriptome Sequencing

Jun Hong^{1*}, Xinyang Li^{1†}, Mengyao Jiang¹ and Ruofei Hong²

¹ College of Life Science and Engineering, Henan University of Urban Construction, Pingdingshan, China, ² School of International Education, Henan University of Technology, Zhengzhou, China

OPEN ACCESS

Edited by:

Costas C. Papagiannitsis,
University of Thessaly, Greece

Reviewed by:

Vittoria Mattioni Marchetti,
Charles University, Czechia
Qi Jialong,
Chinese Academy of Medical
Sciences and Peking Union Medical
College, China
Marc Finianos,
Charles University, Czechia

*Correspondence:

Jun Hong
hongjun@hncj.edu.cn

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 08 February 2022

Accepted: 10 March 2022

Published: 07 April 2022

Citation:

Hong J, Li X, Jiang M and Hong R
(2022) Co-expression Mechanism
Analysis of Different Tachyplesin
I-Resistant Strains in *Pseudomonas*
aeruginosa Based on Transcriptome
Sequencing.
Front. Microbiol. 13:871290.
doi: 10.3389/fmicb.2022.871290

Tachyplesin I is a cationic antimicrobial peptide with 17 amino acids. The long-term continuous exposure to increased concentrations of tachyplesin I induced resistance in *Pseudomonas aeruginosa*. The global gene expression profiling of tachyplesin I-resistant *P. aeruginosa* strains PA-60 and PA-99 and the sensitive strain *P. aeruginosa* CGMCC1.2620 (PA1.2620) were conducted by transcriptome sequencing to analyze the common underlying mechanism of resistance to tachyplesin I in low- or high-resistance mutants. The co-expression patterns, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment, sRNA target genes, and single-nucleotide polymorphism (SNP) change were analyzed for the co-expressed genes in this study. A total of 661 differentially co-expressed genes under treatments of PA1.2620 vs. PA-99 and PA1.2620 vs. PA-60 (HL) were divided into 12 kinds of expression patterns. GO and KEGG pathway enrichment analyses indicated that the enrichment of co-expressed genes was mainly associated with oxidoreductase activity, mismatched DNA binding, mismatch repair, RNA degradation of GO terms, aminoacyl-tRNA biosynthesis, and aminobenzoate degradation pathways, and so forth. The co-expressed resistance-related genes were mainly involved in antibiotic efflux and antibiotic inactivation. Seven co-expressed genes had SNP changes. Some co-expressed sRNAs were involved in *P. aeruginosa* resistance to tachyplesin I by regulating target genes and pathways related to resistance. The common resistance mechanism of *P. aeruginosa* among different mutants to tachyplesin I was mainly associated with the expression alteration of several genes and sRNA-regulated target genes related to resistance; few genes had base mutations. The findings of this study might provide guidance for understanding the resistance mechanism of *P. aeruginosa* to tachyplesin I.

Keywords: co-expressed genes, *P. aeruginosa*, RNA-Seq, SNP, sRNA, tachyplesin I

INTRODUCTION

Pseudomonas aeruginosa (*P. aeruginosa*) is an important opportunistic pathogen of concern in the healthcare system due to increased rates of resistance to antibiotics or antimicrobial peptides (AMPs). Multidrug resistance in *P. aeruginosa* is increasingly becoming a threat to human health. The bacteria easily form biofilms, which significantly increases bacterial resistance to antibiotics

and innate host defense. A number of AMPs have been reported as potential anti-biofilm agents against multidrug-resistant bacteria (Chung and Khanum, 2017; Beaudoin et al., 2018).

During the co-evolution of hosts and bacterial pathogens, bacteria have developed an ability to sense and initiate an adaptive response to AMPs and thus resist their bactericidal activity. *P. aeruginosa* also could develop resistance to a number of AMPs, such as polymyxin (Han et al., 2018). A growing number of *P. aeruginosa* strains demonstrated resistance to AMPs due to mutations in two-component regulatory systems (e.g., PhoPQ, PmrAB, ParRS, and CprRS) (Barrow and Kwon, 2009; Fernández et al., 2012). Changes in drug efflux pumps are one of the main resistance mechanisms of antibiotics or AMPs of *P. aeruginosa*. Some bacteria can use efflux pumps or increase the expression of efflux pumps to mediate resistance against cationic antimicrobial peptides (CAMPs) (Andersson et al., 2016; Puja et al., 2020).

With the rapid development of high-throughput sequencing technology, the study of bacterial transcriptomics can help explore the genes differentially expressed before and after bacterial drug resistance and screen out the non-coding RNAs with regulatory effects. Small RNAs (sRNAs) are important components of post-transcriptional regulation. sRNAs in bacteria have evolved with diverse mechanisms to balance their target gene expression in response to changes in the environment (Dutta and Srivastava, 2018), such as bacterial stress responses, iron uptake, quorum sensing, virulence, and biofilm formation (Wagner and Romby, 2015). sRNAs, in most studied cases, can directly base pair with target mRNA to remodel its expression. As an example, two sRNAs (Sr0161 and ErsA) in *P. aeruginosa* interact with the mRNA encoding the major porin OprD responsible for the uptake of carbapenem antibiotics. The two sRNAs base pair with the 5' untranslated region (5'UTR) of oprD, leading to an increase in the bacterial resistance to meropenem. Another sRNA Sr0161 positively regulated PagL expression in the absence of Hfq, reducing its pro-inflammatory properties and leading to polymyxin resistance (Zhang et al., 2017). Revealing the post-transcriptional regulation of bacterial drug-resistant genes to screen the existing sRNA in bacteria and study the function of sRNA through transcriptome research is of great significance.

Single-nucleotide polymorphism (SNP) refers to DNA sequence polymorphism caused by nucleotide variations in the genome, including single-base conversion, transposition, and single-base insertion and deletion. In genomic DNA, any base may be mutated. Therefore, SNPs may exist in any sequence of a gene, either in a non-coding sequence or in a gene sequence.

Tachyplesin I, a CAMP with a typical cyclic antiparallel β -sheet structure, was originally isolated from hemocytes of marine horseshoe crabs in 1988 (Nakamura et al., 1988). With potent and broad-spectrum activities against both Gram-positive and Gram-negative bacteria, tachyplesin I has been found as a promising candidate for the development of anti-infection, anti-tumor, and anti-virus drugs (Xie et al., 2016; Kuzmin et al., 2018). Previous studies showed that tachyplesin I could kill bacteria by permeabilizing or damaging the bacterial membrane and acting on intracellular targets in bacteria such as inhibition of DNA,

RNA, and protein synthesis and enzyme inactivation (Hong et al., 2013; Kushibiki et al., 2014; Hong et al., 2015; Amiss et al., 2021), and could play a synergistic role. The cytotoxicity of tachyplesin I to mammalian cells was weak in a bacteriostatic dose, and PEGylated tachyplesin I could greatly reduce the cytotoxicity (Imura et al., 2007). Liu et al. (2018) indicated that tachyplesin might kill bacteria by targeting FabG, the conserved β -ketoacyl-acyl carrier protein reductase in unsaturated fatty acid biosynthesis, and provided one possible mechanism to explain how tachyplesin kills bacteria and causes cytotoxicity by targeting membranes.

Bacterial resistance to tachyplesin I has been induced under laboratory conditions, revealing the potential involvement of extracellular protease in mediating tachyplesin I-resistance in Gram-negative bacteria (Hong et al., 2016, 2018). Our results indicated that *P. aeruginosa* resistance to tachyplesin I was mainly related to the reduced entry of tachyplesin I into the bacterial cell and decreased outer membrane permeability (Hong et al., 2020). However, tachyplesin I-induced *P. aeruginosa* mutants were obtained from different generations. The common or specific resistance mechanism of *P. aeruginosa* among different mutants to tachyplesin I has remained elusive.

Therefore, we employed transcriptome sequencing to investigate the differentially co-expressed genes in genome-wide gene expression among tachyplesin I-induced PA-60 and PA-99 mutants and the original strain PA1.2620. Then, the co-expression patterns, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, annotation and analysis of Comprehensive Antibiotic Resistance Database (CARD), sRNA target genes, and SNP change in the co-expressed genes were analyzed in this study. The findings might enhance our understanding on the common resistance mechanism of *P. aeruginosa* to tachyplesin I.

MATERIALS AND METHODS

Strains, Media, and Growth Conditions

PA1.2620 was provided by China General Microbiological Culture Collection Center (CGMCC; Beijing, China). The minimal inhibitory concentration (MIC) value of tachyplesin I was 10 μ g/mL for PA1.2620. Tachyplesin I-resistant mutants (PA-99 and PA-60) were produced in our laboratory by exposure to increasing tachyplesin I concentrations as described previously (Hong et al., 2016, 2020). Namely, PA-60 and PA-99 (naming based on transfer generation) induction strains were developed after 60 and 99 serial transfers in PA1.2620 under long-term selection by exposure to increasing concentrations of tachyplesin. The MIC value of tachyplesin I was 30 and 150 μ g/mL for PA-60 and PA-99 mutants, respectively; the two strains had different levels of resistance to tachyplesin I. These strains were cultured in Mueller–Hinton broth (MHB) medium or nutrient agar plates at 30°C.

Tachyplesin I

Tachyplesin I (>95% purity) was synthesized by Gil Biochemical Co., Ltd. (Shanghai, China), and its sequence was as follows:

NH₂-K-W-C-F-R-V-C-Y-R-G-I-C-Y-R-R-C-R-CONH₂, including two disulfide bonds (C3–C16 and C7–C12) (10). It was dissolved in sterile water to yield 10 mg/mL stock solution, which was filter-sterilized before use. The peptide solution was freshly prepared on the day of the assay or stored at –20°C for a short period.

RNA Isolation, cDNA Library Construction, and Transcriptome Sequencing

PA1.2620 original strain and PA-60 and PA-99 mutants were grown overnight for 12 h in 3 mL of MHB medium at 30°C and shaken at 180 rpm. Then, the cultures were refreshed with medium to OD_{600nm} = 0.2 and grown to the mid-exponential phase (OD_{600nm} = 1.0) until the time of harvesting. The extraction and purification of the total RNA from the cells were carried out using a bacterial total RNA purification kit (Sangon Biotech Co., Ltd., Shanghai, China). The quality, quantity, and integrity of extraction were analyzed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, MA, United States), gel electrophoresis, and an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., CA, United States), respectively, to ensure the use of total RNA sequencing for transcriptome analysis. All experiments were performed in triplicate.

After the samples of total RNA passed the test, the library construction was carried out. The main procedures for library construction were the same as in the previous study (Hong et al., 2020). The RNA-Seq libraries of nine samples were prepared and subsequently sequenced using an Illumina HiSeq 2500 Platform (Illumina Inc., CA, United States) following the manufacturer's protocols. The construction of the libraries and the transcriptome sequencing were performed by the Biomarker Biotechnology Corporation (Beijing, China).

Analysis and Assembly of RNA-Seq Data

Generally, the raw reads contained tiny minority primer sequences, adapter sequences, sequencing connectors, and other potential contaminants. Prior to subsequent analysis, the clean reads were filtered from the raw reads by removing the reads with only adaptor and unknown nucleotides.

The clean reads obtained from RNA-Seq were mapped on the reference genome of *P. aeruginosa* PAO1 using Bowtie-2 software, and only mapped data were used for subsequent analysis (Langmead et al., 2009). Data analysis and base-calling were undertaken using Illumina sequencing software (Illumina Inc., CA, United States).

Gene Expression Analysis

Each UniGene cluster was compared via UniGene database using Bowtie-2 software to evaluate the expression level of all reads and normalized into Reads Per Kilobase of transcript per Million reads mapped (RPKM) values using the following formula: $RPKM = \frac{\text{Map the UniGene cluster}}{\text{Map all UniGene clusters (million)}} \times \frac{1}{\text{The length of UniGene}}$

cluster (kb)]. Also, differentially expressed genes (DEGs) with a false discovery rate (FDR < 0.01) and a fold change of ≥ 2 were selected for analysis according to the DESeq software (Wang et al., 2010).

Functional Annotation and Enrichment Analysis of Differentially Expressed Genes

The functional annotation of each UniGene cluster was searched against various protein databases and identified using annotation information of the given UniGene cluster that had the highest sequence similarity with the tested one. The Blast2GO program and Web Gene Ontology Annotation Plot (WEGO) software were also employed to obtain GO annotations and the distribution of gene functions for each UniGene cluster using a value of less than 10^{-5} (Ashburner et al., 2000). The assignment of each UniGene cluster to different pathways was carried out by searching in the KEGG database using KEGG Automatic Annotation Server (Kanehisa et al., 2004). DEGs were annotated in the CARD using Resistance Gene Identifier (RGI) to find out the antibiotic resistance genes (Jia et al., 2017).

Analysis of Differentially Expressed Small RNA and Regulatory Target Genes

The reads mapped on the reference genome sequence were checked to determine whether that region could form a hairpin structure and was located on the arm. The prediction result was obtained after screening using the miRDeep2 software. The differentially expressed sRNAs in different treatment groups were screened out. The differential sRNA target gene data were processed to identify the DEGs, and GO function and KEGG pathway enrichment analyses were performed.

Single-Nucleotide Polymorphism Analysis of Differentially Expressed Genes and Small RNA Target Genes

Based on the results of reads mapped on reference genome sequences of each sample using the Bowtie software, the Genome Analysis ToolKit (GATK) software was used to identify a single-base mismatch between the sequenced sample and reference genome and search for potential SNPs in gene regions. Comparison of the sample sequence with the reference transcriptome sequence and statistical comparison were performed on the SNPs in each sample of the control and treatment groups, and the statistical results of their types and numbers were analyzed. The location information of genes was extracted from the gff files of *P. aeruginosa* using the GFF information filtering tool based on the gene ID of the Biomarker Cloud platform. The number and types of SNPs generated in DEGs were obtained by comparing the location information of SNPs with the corresponding location information of genes. The obtained data were analyzed mainly through the Cloud platform of Biomarker.

RESULTS

Sequence Analysis and Assembly

The transcriptome sequencing of nine cDNA libraries was performed using total RNA extracted from PA-99 and PA-60 strains and the original strain PA 1.2620. After filtering the raw data and passing quality control, the clean data of each sample were no less than 2.11 Gb, the GC content of clean data was more than 62.57%, and the Q30 base reached more than 95.42% (**Supplementary Table 1**). These data were selected as high-quality reads for further analysis.

The evaluation of transcriptome sequencing data of nine samples, including those mapped to the reference genome, randomness testing of mRNA fragmentation, saturation testing of data, and distribution of reads on the reference genome were performed. The results revealed that the transcriptome data obtained were ideal and reliable in terms of quality. The statistical reads of clean data mapped to reference genome mapping statistical data are presented in **Supplementary Table 2**.

The transcriptome sequencing results revealed a correlation between one biological replicate (PA1.2620¹) and the other two replicates with low correlations (0.4689, 0.4735) of the original strain PA1.2620, whereas R^2 between the other two samples was 0.9931 (**Supplementary Table 3**). In addition, among the three biological replicates of PA-60 strain having high correlations, one of the biological replicates (PA-60¹) was far different from the overall dispersion of other samples. Considering the accuracy of the results, we deleted one biological replicate of PA1.2620 with poor correlation and one biological replicate of PA-60 that was different from the other two samples (**Supplementary Table 3** and **Supplementary Figure 1**).

Gene Expression Analysis

We selected tachyplesin I-induced 99 transfer strain (PA-99 mutant), 60 transfer strain (PA-60 mutant), and the original strain PA1.2620 to perform expression profiling analysis so as to investigate the changes in transcription levels before and after PA1.2620 resistance to tachyplesin I (**Figures 1, 2**).

As illustrated in **Figure 1A**, the PA-60 mutant had 1419 DEGs, including 490 (34.53%) upregulated genes and 929 (65.47%) downregulated genes, compared with the original strain PA1.2620; the log₂FC values of expression were mainly distributed between -2.5 and 2.5. Besides, the expression of 37 genes was downregulated with more than sixfold changes. The top three downregulated genes were *Novel_815*, *Novel_838*, and *Novel_43*. According to Basic Local Alignment Search Tool (BLAST) analysis, the homologous protein of *Novel_815* gene (log₂FC = -4.793) was ferredoxin. The *Novel_838* (log₂FC = -4.275) was permease protein MsbA. The *Novel_43* gene had no homologous proteins. The most significantly upregulated gene was *fruA*, and its homologous protein was the phosphotransferase system (PTS) fructose-specific IIBC component. Whether PA-60 low resistance to tachyplesin I was associated with the changes in highly expressed genes needs to be further explored.

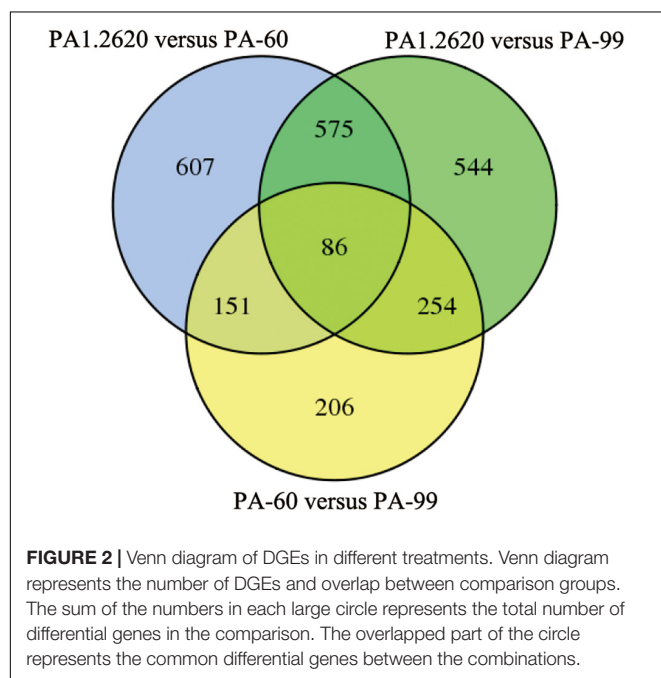
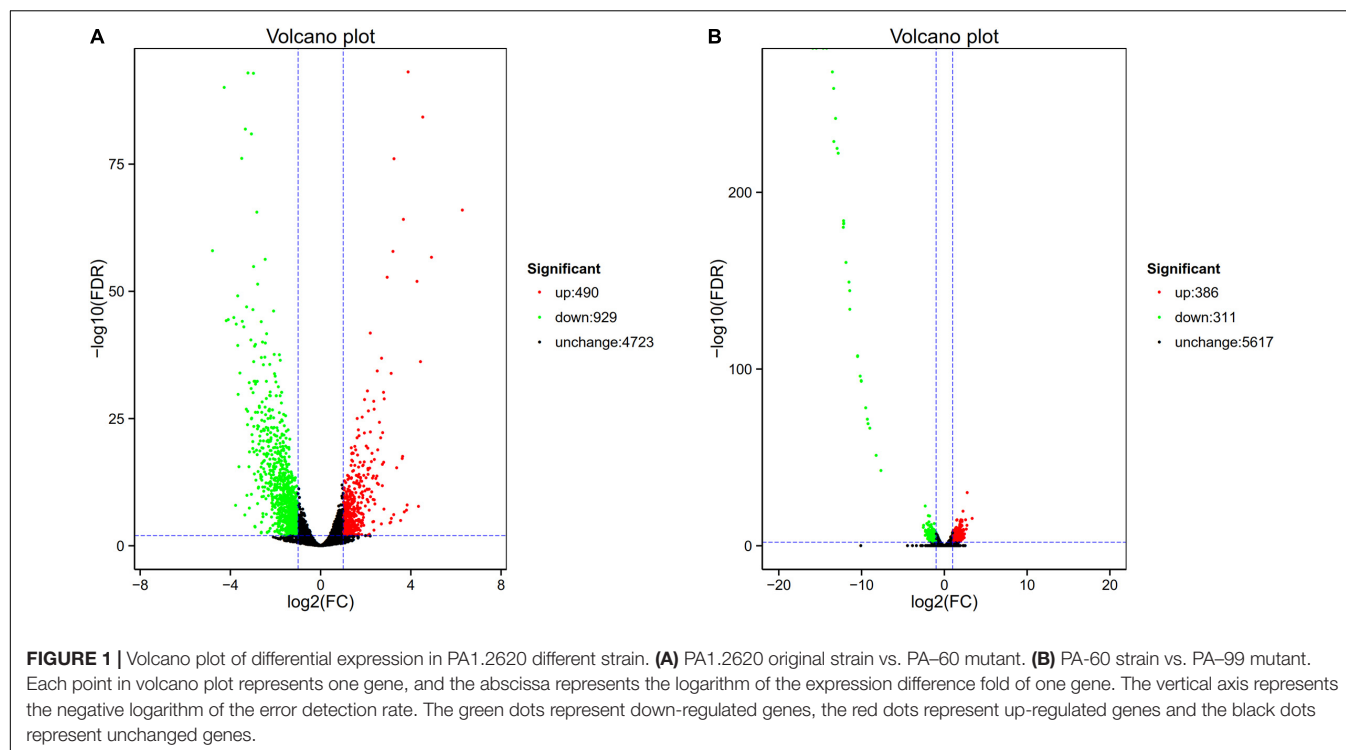
The PA-99 mutant had 1459 DEGs, including 672 (46.06%) upregulated genes and 787 (53.94%) downregulated genes, compared with the original strain PA1.2620; the log₂FC values of expression levels were mainly distributed between -5 and 5.

The study further analyzed the difference in gene expression between different mutants. The PA-99 mutant had 697 DEGs, including 386 upregulated genes and 311 downregulated genes, compared with the induced strain PA-60 (**Figure 1B**). A total of 29 genes had downregulated expression with more than tenfold changes. The top five genes were *Novel_392*, *PA2754a*, *PA2765*, *PA2764*, and *PA2758*, all with downregulated expression. According to BLAST analysis, the *Novel_392* and *PA2754a* had no homologous proteins. The homologous protein of *PA2765* was peroxidase. The homologous proteins of *PA2764* and *PA2758* belonged to the alpha/beta hydrolase superfamily and LysR family of transcriptional regulators.

Common or specific DEGs in different strains were shown in the Venn diagram (**Figure 2**). A total of 661 co-expressed genes DEGs were found in the treatments of PA1.2620 vs. PA-60 and PA1.2620 vs. PA-99 (HL). A total of 237 co-expressed genes were under the treatments of PA1.2620 vs. PA-60 and PA-60 vs. PA-99, while 340 co-expressed genes were under treatments of PA1.2620 vs. PA-99 and PA-60 vs. PA-99. Moreover, 86 common DEGs were found among the 3 treatment groups. The results showed DEGs and co-expressed genes in different strains of induced transfers. However, 785 specific DEGs were identified in the PA1.2620 vs. PA-60 treatment group and 798 specific DEGs in the PA1.2620 vs. PA-99 treatment group. The result indicated that different tachyplesin-induced transfers influenced the gene expression levels of DEGs of *P. aeruginosa* mutants compared with the original strain. In other words, the expression levels of DEGs changed to varying degrees with the increase in *P. aeruginosa* resistance to tachyplesin I. We further analyzed the role of the co-expressed genes in the resistant mechanism of *P. aeruginosa* to tachyplesin I.

Analysis of Co-expression Patterns of HL Treatments

The 661 common DEGs in treatments of HL were analyzed to explore the co-expression patterns among PA1.2620 different mutants. Using the K-means clustering algorithm in BMKCloud Platform, 661 co-expressed genes were grouped into 12 clusters (subcluster_01–12) with the increasing resistance to tachyplesin or the increasing tachyplesin-induced transfers. The gene expression levels of six subclusters mainly had an increasing trend, while others had decreased trends. The part results (subcluster 01, 05, 08, and 09) are shown in **Figure 3**. These gene expression levels of subcluster 06 (75 genes), subcluster 07 (7 genes), subcluster 09 (148 genes), and subcluster 12 (44 genes) primarily followed similar patterns of expression with the increasing resistance to tachyplesin, which decreased from T04 (PA1.2620) to T09 and reached a steady state in T09–T11 (PA-99). These gene expression levels of subcluster 01 (115 genes) and subcluster 11 (63 genes) had primarily similar patterns with the increasing resistance to tachyplesin, which obviously decreased from T04 to T07–T08, and slightly increased from T08



to T09, then decreased in T10. In contrast, the gene expression levels of subcluster 03 with 7 genes and subcluster 05 with 20 genes showed similar patterns, which both obviously increased from T04 to T07–T08 (PA-60) and decreased in T09. Then reached a steady state in T10–T11. The gene expression levels of subcluster 08 with 100 genes increased from T04 to T05 to T07, reached a steady state in T07–T09, and then slightly

fluctuated in T10–T11. The co-expression patterns indicated that different tachyplesin-induced transfers influenced the gene expression level in *P. aeruginosa*.

Functional Classification and Enrichment Analysis of the Co-expressed Genes of HL

A further functional classification and enrichment analysis of 661 common DEGs between HL was performed using GO and KEGG databases. The GO annotation classification indicated that common DEGs were enriched in cellular components, molecular functions, and biological processes; it was subdivided into 46 GO_classify2 functional terms. As shown in **Figure 4A**, the top 20 GO terms were significantly enriched (P -value < 0.05) in the biological process and molecular function. Detailed information about the enriched GO terms was provided in **Supplementary Table 4**. Of these enriched GO terms, nine terms were related to molecular function. The top four enriched GO terms included “oxidoreductase activity, acting on NAD(P)H, NAD(P) as acceptor,” “oxidoreductase activity, acting on other nitrogenous compounds as donors,” “mismatched DNA binding,” and “FMN reductase activity.” Further, 11 terms related to biological process mainly included “mismatch repair,” “intracellular sequestering of iron ion,” “SOS response,” “quorum sensing,” and so forth. The GO enrichment indicated that tachyplesin significantly affected oxidoreductase activity and mismatched DNA binding in *P. aeruginosa*.

A total of 186 DEGs were assigned to 71 KEGG pathways. The KEGG pathway analysis revealed 20 over-represented pathways (P -value < 0.5), as shown in **Figure 4B** and

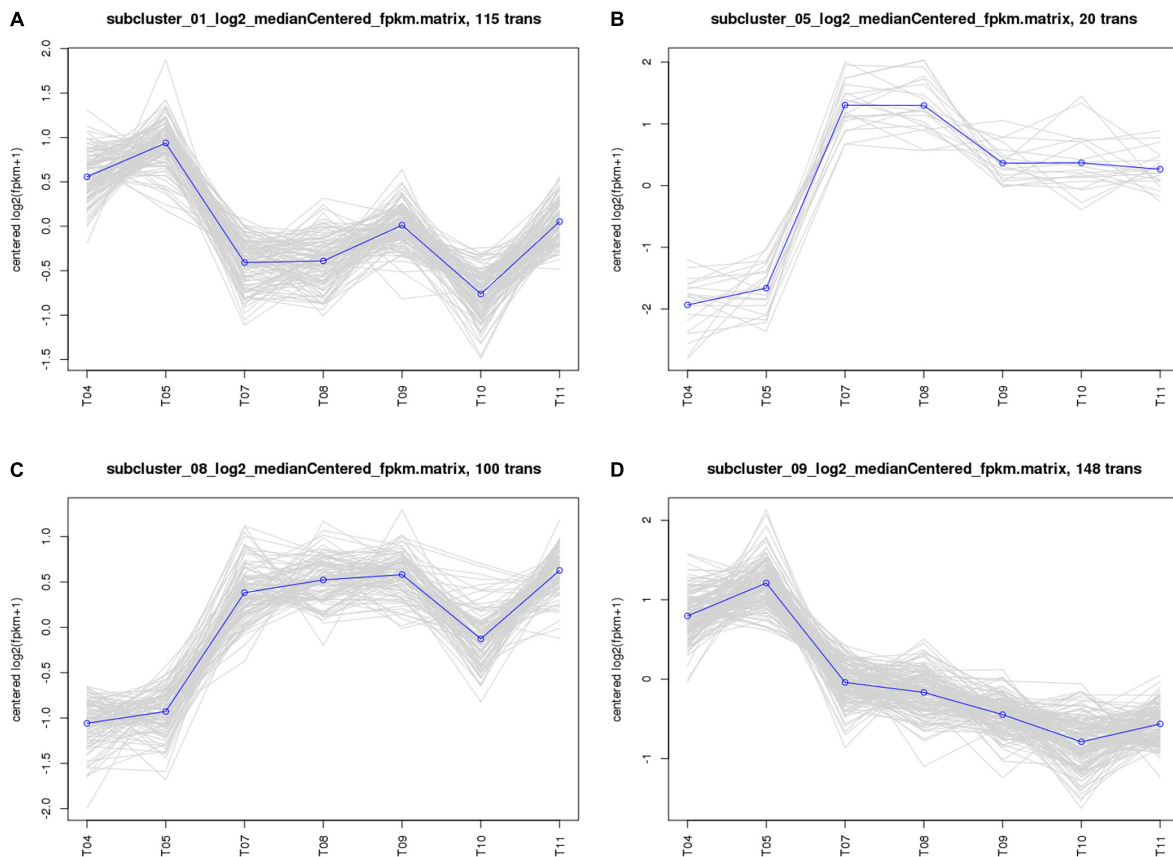


FIGURE 3 | Analysis of co-expressed patterns of HL treatments. **(A)** Subcluster_01. **(B)** Subcluster_05. **(C)** Subcluster_08. **(D)** Subcluster_09. The abscissa represents different sample, the vertical axis represents centered log₂ (fpm + 1). Four clusters were identified based on expression levels in seven different samples. T04–T05 are the two biological duplicates of the original strain PA1.2620; T07–T08 are the two biological duplicates of the PA-60 strain; T09–T11 are the three biological duplicates of PA-99 strain.

Supplementary Table 5. The top three enriched pathways (corrected P -value < 0.5) were RNA degradation (ko03018), aminoacyl-tRNA biosynthesis (ko00970), and aminobenzoate degradation (ko00627). Eight DEGs were involved in RNA degradation pathways, 12 DEGs in aminoacyl-tRNA biosynthesis pathways, and 7 DEGs in aminobenzoate degradation pathways. In addition, the two-component system (ko02020, 18 DEGs), lipopolysaccharide biosynthesis (ko00540, 2 DEGs), and amino sugar and nucleotide sugar metabolism (ko00520, 3 DEGs) pathways were involved in several known CAMP resistance DEGs. The beta-lactam resistance (ko01501, 3 DEGs), flagellar assembly (ko02040, 2 DEGs), bacterial secretion system (ko03070, 8 DEGs), and bacterial chemotaxis (ko02030, 10 DEGs) were also related to resistance. The enriched pathway showed that *P. aeruginosa* resistance to tachyplesin I resulted in some alterations in several pathways related to resistance.

Annotation and Analysis of Co-expressed Resistance Genes of HL

First, DEGs were annotated in CARD using RGI to select the known antibiotic resistance genes. Under RGI criteria with the perfect, strict, complete genes only which is the most

commonly used, seven co-expressed genes in **HL** treatments were mainly involved in antibiotic efflux for resistance–modulation–cell division (RND) antibiotic efflux pump (six genes, including *armR*, *nalC*, *PA1435*, *PA3677*, *mexC*, and *mexE*) and antibiotic inactivation (*PA5514* coding for OXA betalactamase), as shown **Figure 5A**. The result showed that *P. aeruginosa* resistance to tachyplesin I mainly involved the overexpression of the RND efflux pump, which was independent of the induction time.

Under RGI criteria with the perfect, strict and loose, and complete genes only which can be used for reference to predict drug resistance genes, 74 co-expressed potential resistance genes had high homology with 46 known antibiotic resistance genes in treating **HL** (**Figure 5A**). The common resistance mechanisms of tachyplesin in *P. aeruginosa* mainly involved antibiotic efflux pump (44 genes), antibiotic inactivation (11 genes), antibiotic target alteration (17 genes), and antibiotic target protection (2 genes). The mechanism of the antibiotic efflux pump involved these genes of RND, major facilitator superfamily (MFS), ATP-binding cassette (ABC), and multidrug and toxic compound extrusion (MATE) transporter families. Their expression levels were most upregulated with $\log_2\text{FC} \leq 2$ fold changes in the RND efflux pump genes. The mechanism of antibiotic target alteration

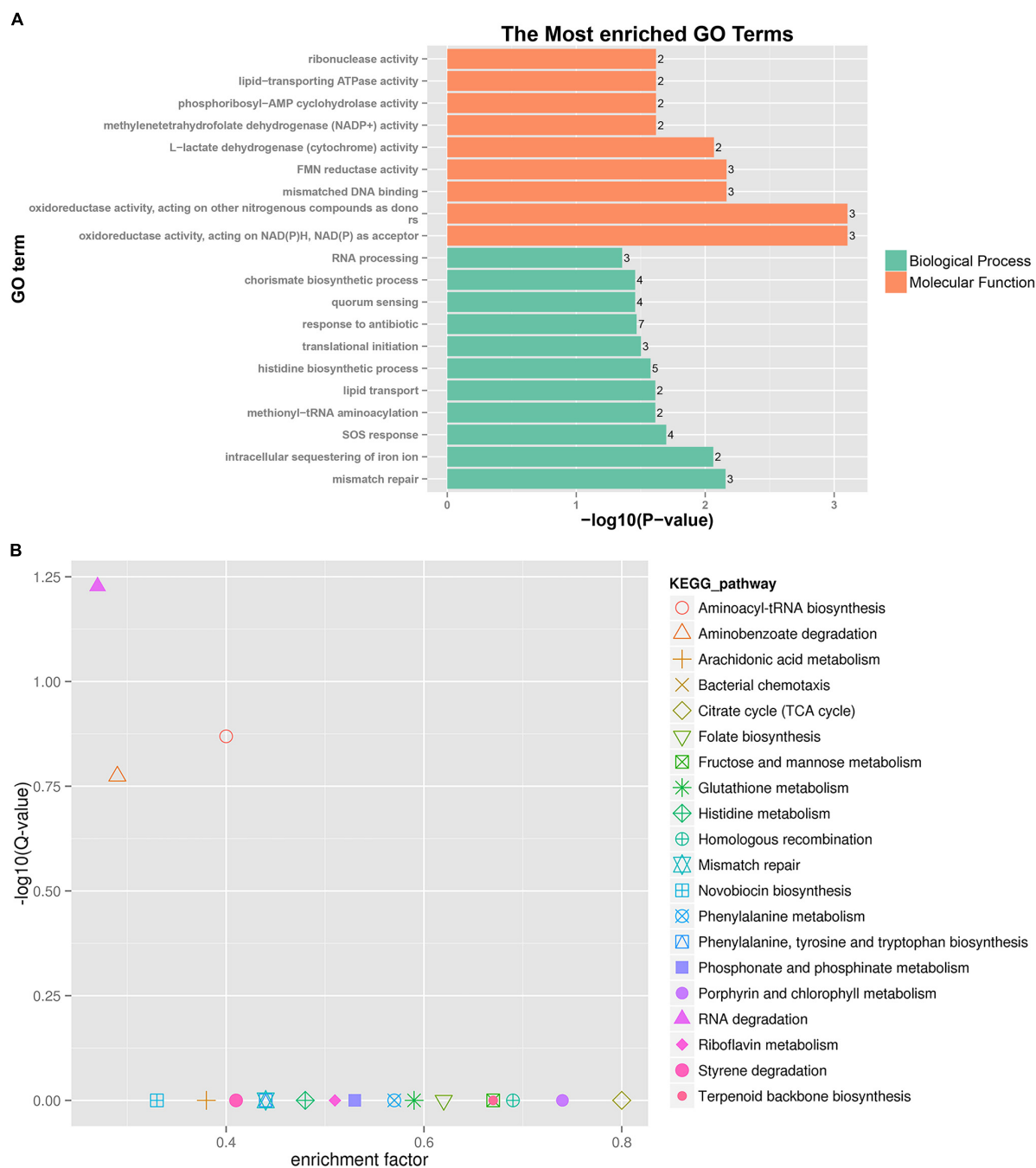
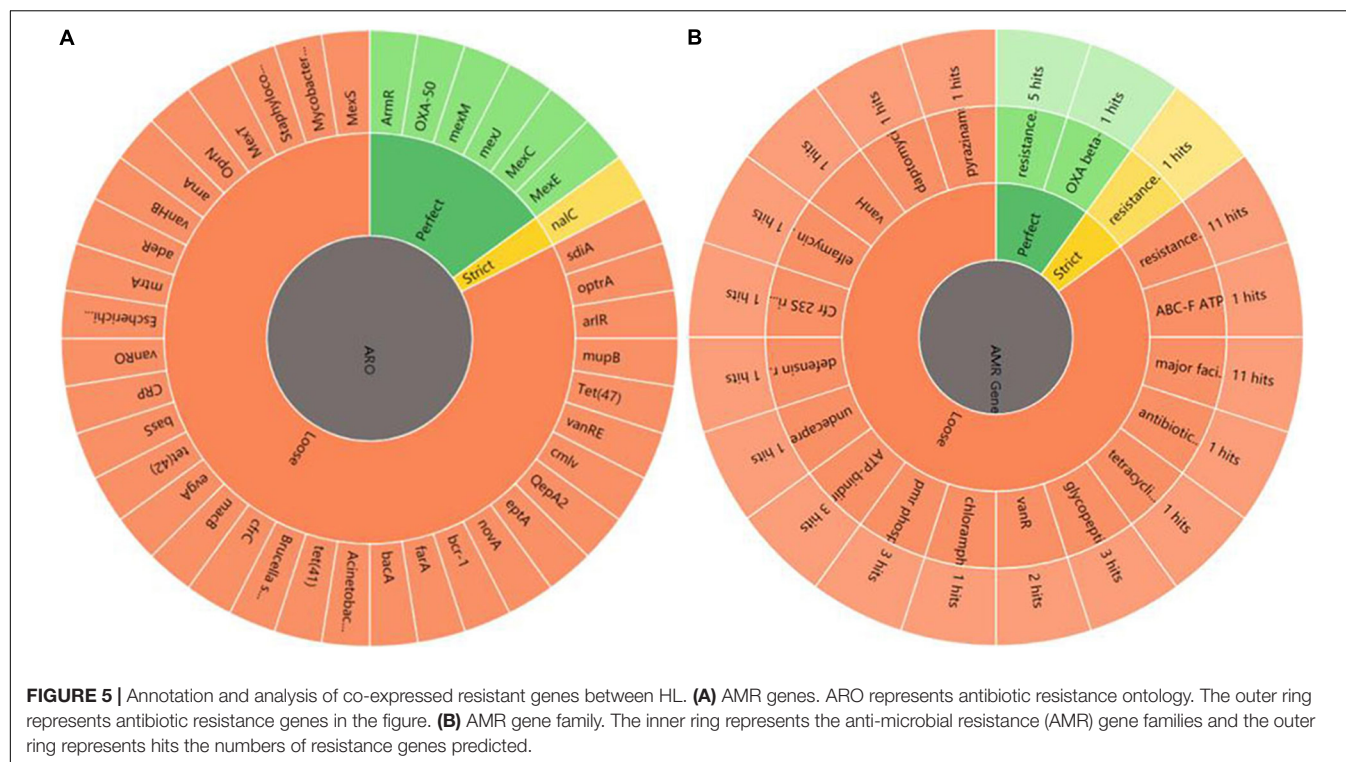


FIGURE 4 | GO and KEGG pathway enrichment analysis of DEGs between HL treatments. **(A)** GO enrichment. The bigger the value of $-\log_{10}(p\text{-value})$, the more enriched. $p\text{-value} < 0.05$ is the significant difference, and $p\text{-value} < 0.01$ is the most significant difference. The ordinate is "GO term." **(B)** KEGG pathway enrichment. Enrichment factor = Amount of all genes/Amount of DEGs enriched in the pathway in the background gene set. The smaller the enrichment factor, the more significant. The same below.

was mainly involved in 11 anti-microbial resistance (AMR) gene families. The mechanism of antibiotic inactivation was mainly involved in beta-lactamase, fosfomycin thiol transferase, and chloramphenicol phosphotransferase of AMR gene families.

The top three members of AMR gene families were RND, MFS, and ABC antibiotic efflux pump, detected using homologous

alignment in RGI analysis. AMR gene families mainly included 23 potential genes with high homology to 17 members of the RND efflux pump family, 16 potential genes with high homology to the 11 members of the MFS efflux pump family, 4 potential genes (*PA4223*, *Novel_141*, *rbsA*, and *PA3376*) with high homology to the 3 members of the ABC efflux pump family, and 3



potential genes (*fnt*, PA2480, and PA4517) with high homology to the 3 members of the pmr phosphoethanolamine transferase family. Three genes (gene ID: *Novel_737*, *Novel_736*, and *Novel_612*) were involved in the antibiotic resistance isoleucyl-tRNA synthetase (*ileS*) family (Figure 5B). The MFS efflux pump family was involved in more DEGs in this study. However, the function of MFS genes in *P. aeruginosa* resistance to tachyplesin was unclear and needed further investigation.

Single-Nucleotide Polymorphism Analysis

The data statistical analysis of SNP loci numbers, proportion and types of transition, and transversion and heterozygosity of SNPs on the samples showed that the frequency of SNPs was relatively consistent in each sample relative to the control PA1.2620 genome, in which the frequency of transition types was much higher than that of transversion types. Among the transition types, the most conversions were from A to G, followed by C to T. With the increasing tachyplesin resistance from PA-60 to PA-99 strains, the transition frequency tended to decrease while the transversion frequency tended to increase (Table 1). Also, differences were found in the frequency of occurrence between different groups of biological duplicates in the same sample.

Single-Nucleotide Polymorphism Analysis of Co-expressed Genes Between HL

Using the PA1.2620 strain as the control group, the PA-60 and PA-99 strains were compared to find the common SNP

TABLE 1 | Statistical table of mutation sites and mutation category of SNP.

SNP category Name	PA-60 sample		PA-99 sample		
	PA-60 ²	PA-60 ³	PA-99 ¹	PA-99 ²	PA-99 ³
Transition number	153	363	105	196	91
A/G	81	208	56	107	46
C/T	72	155	49	89	45
Transversion number	31	32	33	29	34
A/C	7	10	7	5	7
A/T	7	5	9	7	6
C/G	11	10	9	11	14
G/T	6	7	8	6	7
T/C, G	1	0	1	1	1
Total SNP number	185	395	139	226	126
Transition%	82.70	91.90	75.54	86.73	72.22
Transversion%	16.76	8.10	23.74	12.83	26.98
Heterozygosity%	0.54	0.00	0.72	0.44	0.79

PA-60² and PA-60³ are two biological duplicates of the PA-60 strain, while PA-99¹, PA-99², and PA-99³ are three biological duplicates of the PA-99 strain.

changes of DEGs. For these SNP genes, the base mutations that occurred only in one biological repeat were filtered out. A total of 22 DEGs (12 upregulated and 10 downregulated) had SNP changes between PA1.2620 and PA-99 strains, while 14 DEGs had SNP changes between PA1.2620 and PA-60 strains. Seven co-expressed genes had SNP changes, with four upregulated and three downregulated, as shown in Table 2.

Among the seven genes with SNP change, two were novel genes. SNPs mainly occurred in log2FC values of gene expression from onefold to threefold change. However, only two base

TABLE 2 | SNP change of the co-expressed genes of HL.

Gene #ID or name	Protein_name	PA1.2620 vs. PA-60 (log2FC)	PA1.2620 vs. PA-99 (log2FC)	Regulated	Start	End	SNP number	Snps
<i>pvdL</i>	Peptide synthase	1.10009	1.15621	Up	2,707,666	2,720,694	2	2712019C > A; 2715424T > C
<i>Novel_244</i>	-	-1.10470	-1.71631	Down	1,720,641	1,732,256	1	1732254T > C
<i>PA1938</i>	Hypothetical protein	-1.03418	-1.95428	Down	2,118,926	2,119,747	1	2119153T > C
<i>Novel_745</i>	-	1.86917	2.58303	Up	5,130,955	5,135,726	1	5130955T > C
<i>PA0193</i>	Hypothetical protein	1.39140	1.186966	Up	221,585	222,487	1	222358C > T
<i>exoY</i>	Adenylate cyclase	-1.08212	-1.75024	Down	2,410,344	2,411,480	1	2411150G > T
<i>vreR</i>	Sigma factor regulator VreR	2.20828	1.68468	Up	735,487	736,446	1	735564C > T

mutations occurred in the *pvdL* gene. According to BLAST analysis, *Novel_745*, *PA1938* and *PA0193* genes were described as hypothetical proteins, while a homology of *Novel_244* gene was 2-oxoglutarate dehydrogenase.

Analysis of Co-expressed Small RNA of HL

A total of 17 differentially expressed novel sRNAs were found under treatments of PA1.2620 vs. PA-60, and 11 differentially expressed sRNAs under treatments of PA1.2620 vs. PA-99, as detected by RNA sequencing. Six novel sRNAs with downregulated expression were co-expressed under treatments of HL, as shown in Table 3. Considering different sRNAs with the same target genes, 91 upregulated differentially expressed target genes were found after filtering out the duplicates.

Functional Enrichment Analysis of Co-expressed Small RNA Target Genes

A further functional classification and enrichment analysis of six co-expressed sRNA target genes (91) for HL was performed using GO and KEGG databases. As shown in Figure 6A, the top 20 GO terms were significantly enriched (P -value < 0.05) in the biological process and molecular function. Detailed information about the enriched GO terms is provided in Supplementary Table 6A. Of these enriched 20 GO terms, 12 terms were related to the molecular function and 8 terms were related to the biological process. The top four enriched GO terms (P -value < 0.01) included “carbon-nitrogen ligase activity, with glutamine as amido-N-donor (two genes),” “porphyrin-containing compound biosynthetic process (two genes),” “spermidine biosynthetic process (two genes),” and “respiratory electron transport chain (two genes).”

A total of 91 differentially co-expressed sRNA target genes were assigned to 29 KEGG pathways. The KEGG pathway analysis revealed the top 20 pathways (P -value < 0.5), as shown in Figure 6B and Supplementary Table 6B. The top four enriched pathways included aminobenzoate degradation (ko00627, three target genes), phenylalanine metabolism (ko00360, three target genes), styrene degradation (ko00643, two target genes), and arginine and proline metabolism (ko00330, four target genes). Besides the first 20 pathways, the two-component system (ko02020, three target genes), aminoacyl-tRNA biosynthesis (ko00970, one target gene), ABC transporters (one target gene), and glutathione, nitrogen, and sulfur metabolism were also detected. The enriched pathway showed that some sRNA played a role in *P. aeruginosa* resistance to tachyplesin I by regulating the pathway changes of target genes, such as aminobenzoate degradation, amino acid metabolism, bacterial chemotaxis, two-component system, and so on.

Single-Nucleotide Polymorphism Change of the Co-expressed Small RNA Target Genes

The SNPs of six co-expressed sRNA of HL showed no change. The target genes of 91 differentially co-expressed sRNAs were analyzed for changes in SNP. A total of three target genes

TABLE 3 | Co-expressed differentially expressed sRNA of HL.

sRNA (ID)	Gene length (nt)	PA1.2620 vs. PA-60 (log2FC)	PA1.2620 vs. PA-99 (log2FC)	Regulated
Novel_33	188	-2.904338527	-3.101050802	Down
Novel_377	378	-1.584854106	-2.065876992	Down
Novel_583	78	-3.066999834	-3.461144761	Down
Novel_584	116	-1.973275535	-2.305895532	Down
Novel_6	94	-1.43082481	-1.576755833	Down
Novel_853	117	-1.028188198	-1.507421364	Down

(upregulated) showed SNP changes after removing SNPs that occurred only in one biological replicate gene (Table 4). The expression levels of two sRNA target genes (*vreR* and *Novel_745*) were different in PA1.2620 vs. PA-60 and PA1.2620 vs. PA-99. Few SNP changes were found in sRNA target genes, and most of them occurred in target genes with low differential expression. Only the *pvdL* gene had two site mutations, while the other two genes had one site mutation. Also, the *pvdL* encoding peptide synthase gene was mainly involved in the synthesis of *P. aeruginosa* iron carrier and the biological synthesis of lipids; it participated in the molecular function of catalytic activity. How sRNA participates in the function of these target genes remains to be further explored.

DISCUSSION

By analyzing DEGs in drug-resistant strains, this study might help researchers select from several DEGs closely related to bacterial resistance genes and participating in metabolic pathways, thus providing new ideas for revealing the mechanism of bacterial drug resistance. sRNA is a class of newly discovered gene expression regulators that control cellular physiological functions in response to various environmental changes by pairing with target mRNA or target protein, especially in the development of bacterial drug resistance. These differentially expressed sRNAs mainly act on corresponding target mRNAs and then play their regulatory role. At the genomic level, SNP can be used to analyze the association between complex traits and reveal the genetic mechanism of related traits. Drug resistance caused by gene mutations is a process of long-term accumulation. Different from transcriptome, the genome is relatively stable and not susceptible to mutation. After special induction treatment, directional mutations can be generated to cope with selective pressure in the environment. These mutations may help adapt to the new environment after environmental changes in the strain and may have nothing to do with specific mutations in genes related to drug resistance and virulence. Therefore, based on transcriptome sequencing, the analysis of DEGs, sRNA target genes, and SNP changes among tachyplesin I-resistant strains and the original strain can be used to predict the common cause of *P. aeruginosa* resistance to tachyplesin at the transcriptome and genome levels, to provide certain theoretical support for the generation of *P. aeruginosa* resistance to tachyplesin.

This study attempted to analyze further the common underlying mechanism between different mutants resistant to

tachyplesin I using mRNA expression profiling. From the transcriptome data, we found that several common genes exhibited significantly altered expression. To validate the results of transcriptome sequencing, 10 DEGs from co-expressed genes of HL were randomly selected for RT-qPCR. All the selected DEGs showed concordant expression patterns between the RNA-Seq and the results of RT-qPCR. The part results of 10 DEGs from PA1.2620 vs. PA-99 for RT-qPCR had been published in previous study (Hong et al., 2020). A total of 661 differentially co-expressed genes were divided into 12 kinds of expression patterns, which might be beneficial for understanding the resistance mechanism of *P. aeruginosa* to tachyplesin I.

Aminoacyl-tRNA biosynthesis was a significantly enriched pathway in treating HL with 12 DEGs (1 upregulated and 11 downregulated). The result showed that *P. aeruginosa* resistance to tachyplesin involved the expression of some genes of aminoacyl-tRNA synthesis pathways, but no gene had SNP. The function of aminoacyl-tRNA synthesis is to precisely match amino acids with tRNAs containing the corresponding anticodon. Besides having essential roles in protein synthesis and Gram-positive peptidoglycan cross-linking or membrane phospholipid modification in some bacteria, aminoacyl-tRNAs are also involved in pathways directly implicated in antibiotic biosynthesis and resistance (Shepherd and Ibba, 2013). For example, MprF proteins mediated drug resistance to control the permeability of the cell wall to cationic antimicrobials by catalyzing the aminoacylation of inner membrane lipids (Ernst and Peschel, 2011). Aminoacylation of membrane lipids adjusts the net negative charge of the membrane bilayer in a manner dependent on the aminoacyl-tRNA substrate.

P. aeruginosa has Ala-phospholipid phosphatidyl-glycerols (PG), which can aminoacylate phosphatidylglycerol with Ala and thus neutralize the overall charge of the membrane to enhance the resistance to a multitude of negatively charged β -lactam antibiotics. Ala-PG also has a significant influence on the permeability and fluidity of the lipid bilayer. The *PA0920* gene, as the *mprF* ortholog, was responsible for alanyl-PG synthesis (Klein et al., 2009), and the tRNA-dependent enzyme was located in the inner membrane. ORF PA0919 coded for an alanyl-PG hydrolase that was anchored to the periplasmic surface of the inner membrane (Arendt et al., 2013). PA0919 was termed alanyl-PG hydrolase. In our study, *PA0919* was upregulated in PA-60 (Log2FC = 2.28) and PA-99 (Log2FC = 3.749); the upstream gene *PA0920* was also upregulated in PA-60 (Log2FC = 2.74) and PA-99 (Log2FC = 4.20). The expression

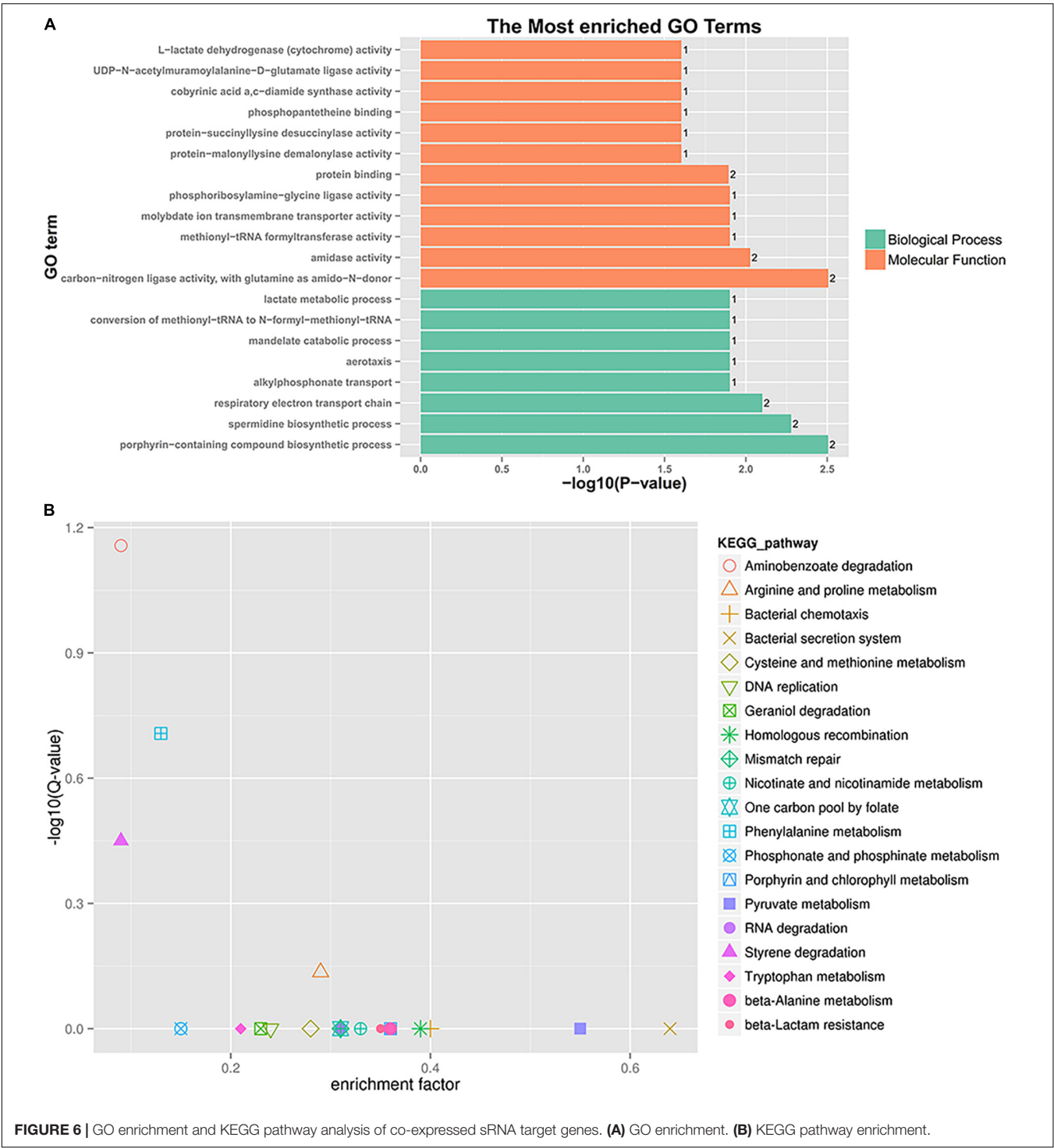


TABLE 4 | SNP change of the co-expressed sRNA target genes of HL.

Gene ID or name	PA1.2620 vs. PA-60 (log2FC)	PA1.2620 vs. PA-99 (log2FC)	Regulated	Start	End	SNP number	Snp
<i>pvdL</i>	1.100094764	1.156212	UP	2,707,666	2,720,694	2	2712019C > A; 2715424T > C
<i>vreR</i>	2.208278525	1.684682	Up	735,487	736,446	1	735564C > T
<i>Novel_745</i>	1.86916584	2.583033	Up	5,130,955	5,135,726	1	5130955T > C

trends of two genes showed that the *PA0920* gene positively regulated *PA0919* gene expression, and the expression level gradually increased with the increased levels of drug resistance in mutant strains. Our previous results indicated that the PA-99 strain resistant to tachyplesin I was mainly related to the reduced entry of tachyplesin I into the bacterial cell due to the overexpression of the efflux pump, in addition to a decrease in outer membrane permeability (Hong et al., 2020). We inferred that different *P. aeruginosa* resistance to tachyplesin might be related to a decrease in the overall net negative charge of the membrane or a decrease in the permeability due to lipid modification.

One of the mechanisms of bacterial resistance to AMPs is the reduced entry of the AMPs into the bacterial cell. The transcriptional regulation of HL indicated that several genes encoded by the outer membrane efflux protein or multidrug resistance protein were all upregulated, three genes (ID *Novel_156*, *Novel_157*, *Novel_443*) encoded by outer membrane porin were downregulated, and only one *opdB* gene was upregulated. The results showed that the common mechanism of *P. aeruginosa* resistance to tachyplesin was also related to overexpression of genes encoding the outer membrane porins and multidrug resistance protein.

Drug efflux through efflux pumps is one of the main resistance mechanisms of *P. aeruginosa* against antibiotics or AMPs (Puja et al., 2020; Zahedi bialvaei et al., 2021). Some bacteria could increase the expression of efflux pumps to mediate resistance against CAMPs (Andersson et al., 2016). The efflux pumps of the RND family could be directly related to the efflux of antibiotics in *P. aeruginosa*. MFS and ABC families were also two important superfamilies of membrane transporters, playing a significant role in substance exchange, energy metabolism, and drug resistance. In the present study, the co-expressed DEGs of HL involved in high-homology resistance genes were found in ABC, MFS, RND, and MATE transporter efflux pump families. Most genes were encoded by RND and MFS efflux pumps, whereas only one gene related to the MATE efflux pump was found, indicating that the RND and MFS antibiotic efflux pumps had an important role in *P. aeruginosa* resistance to tachyplesin, especially RND efflux pumps.

A large number of studies also found that drug efflux pumps could further enhance drug resistance by promoting biofilm formation. The drug resistance mechanism of biofilm formation played a pivotal role in *P. aeruginosa*. The biofilm formation pathway in *P. aeruginosa* mainly involved quorum sensing, bacterial secretion system, and flagellar assembly. We found that several co-expressed genes in treating HL and co-expressed sRNA target genes were assigned to flagellar assembly and bacterial secretion system pathway. In a previous study, PA-99 and PA-60 mutants formed biofilms easily compared with the original strain PA1.2620 with increased resistance to tachyplesin; the contents of extracellular polysaccharides were also higher than that in the original strain (data not shown). When the content of exopolysaccharide reached a certain level, it promoted biofilm formation. The results further indicated that the common mechanism of *P. aeruginosa* resistance to

tachyplesin I could be involved in expressing of biofilm-related genes. However, whether the resistance mechanism of efflux pumps was related to biofilm formation needs to be further investigated.

A two-component regulatory system plays a substantial role in the pathogenicity, virulence, biofilm formation, and drug resistance in *P. aeruginosa* (Bhagirath et al., 2019). In the present study, we detected 18 co-expressed DEGs for a two-component regulatory system under the treatment of HL and 3 co-expressed sRNA target genes, including some regulated resistance genes, such as response regulator *phoP* (gene ID: *Novel_194*) and alkaline phosphatase encoded by *phoA* (gene ID: *Novel_498*, 499). The result showed that a two-component regulatory system might also play a significant role in *P. aeruginosa* resistance to tachyplesin, and sRNA might also be involved in regulating the expression of the two-component system. A previous study demonstrated the importance of regulatory genes, such as *phoP*, *phoQ* and *pmrA*-*pmrB*, as well as the addition of aminoarabinose to lipid A (Barrow and Kwon, 2009).

In summary, mutants resistant to tachyplesin I obtained at different induction times had 661 co-expressed DEGs under the treatment of HL, which were divided into 12 kinds of expression patterns. Some co-expressed genes were regulated by sRNA. A few SNPs with relatively low expressed genes had base mutations, which were mainly related to iron metabolism, adenylate cyclase, and encoding of putative proteins. Several known CAMP resistance pathways and antibiotic resistance mechanisms might be involved in the common potential mechanism of *P. aeruginosa* resistance to tachyplesin I. The findings might enhance our understanding of the common resistance mechanism of different *P. aeruginosa* strains to tachyplesin I. The specific resistance mechanism of *P. aeruginosa* among different mutants to tachyplesin I needs further exploration.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JH and XL conducted the data analysis and wrote the manuscript. JH conceived the idea, collected the data, and aided with the writing. MJ and RH contributed to refine and reorganized the data used in the manuscript and figure preparation. All authors have read and approved the manuscript.

FUNDING

This study was financially supported by the National Natural Science Foundation of China (Grant No. 31540060) and the Natural Science Foundation of Henan Province (Grant No. 202300410035).

ACKNOWLEDGMENTS

We would like to thank Biomarker Biotechnology Corporation (Beijing, China) for their technical assistance via transcriptome sequencing.

REFERENCES

- Amiss, A. S., von Pein, J. B., Webb, J. R., Condon, N. D., Harvey, P. J., Phan, M. D., et al. (2021). Modified horseshoe crab peptides target and kill bacteria inside host cells. *Cell. Mol. Life. Sci.* 79:38. doi: 10.1007/s00018-021-04041-z
- Andersson, D. I., Hughes, D., and Kubicek-Sutherland, J. Z. (2016). Mechanisms and consequences of bacterial resistance to antimicrobial peptides. *Drug. Resist. Updat.* 26, 43–57. doi: 10.1016/j.drup.2016.04.002
- Arendt, W., Groenewold, M. K., Hebecker, S., Dickschat, J. S., and Moser, J. (2013). Identification and characterization of a periplasmic aminoacyl-phosphatidylglycerol hydrolase responsible for *Pseudomonas aeruginosa* lipid homeostasis. *J. Biol. Chem.* 288, 24717–24730. doi: 10.1074/jbc.M113.482935
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Barrow, K., and Kwon, D. H. (2009). Alterations in two-component regulatory systems of phoPQ and pmrAB are associated with polymyxin B resistance in clinical isolates of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 53, 5150–5154. doi: 10.1128/AAC.00893-09
- Beaudoin, T., Stone, T. A., Glibowicka, M., Adams, C., Yau, Y., Ahmadi, S., et al. (2018). Activity of a novel antimicrobial peptide against *Pseudomonas aeruginosa* biofilms. *Sci. Rep.* 8:14728.
- Bhagirath, A. Y., Li, Y., Patidar, R., Yerex, K., Ma, X., Kumar, A., et al. (2019). Two component regulatory systems and antibiotic resistance in gram-negative pathogens. *Int. J. Mol. Sci.* 20:1781. doi: 10.3390/ijms20071781
- Chung, P. Y., and Khanum, R. (2017). Antimicrobial peptides as potential antibiofilm agents against multidrug-resistant bacteria. *J. Microbiol. Immunol. Infect.* 50, 405–410. doi: 10.1016/j.jmii.2016.12.005
- Dutta, T., and Srivastava, S. (2018). Small RNA-mediated regulation in bacteria: a growing palette of diverse mechanisms. *Gene* 656, 60–72. doi: 10.1016/j.gene.2018.02.068
- Ernst, C. M., and Peschel, A. (2011). Broad-spectrum antimicrobial peptide resistance by MprF-mediated aminoacylation and flipping of phospholipids. *Mol. Microbiol.* 80, 290–299. doi: 10.1111/j.1365-2958.2011.07576.x
- Fernández, L., Jenssen, H., Bains, M., Wiegand, I., Gooderham, W. J., and Hancock, R. E. (2012). The two-component system CprRS senses cationic peptides and triggers adaptive resistance in *Pseudomonas aeruginosa* independently of ParRS. *Antimicrob. Agents Chemother.* 56, 6212–6222. doi: 10.1128/AAC.01530-12
- Han, M. L., Zhu, Y., Creek, D. J., Lin, Y. W., Anderson, D., Shen, H. H., et al. (2018). Alterations of metabolic and lipid profiles in polymyxin-resistant *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 62:e02656-17. doi: 10.1128/AAC.02656-17
- Hong, J., Guan, W. T., Jin, G., Zhao, H., Jiang, X., and Dai, J. G. (2015). Mechanism of tachyplesin I injury to bacterial membranes and intracellular enzymes, determined by laser confocal scanning microscopy and flow cytometry. *Microbiol. Res.* 170, 69–77. doi: 10.1016/j.micres.2014.08.012
- Hong, J., Hu, J. Y., and Ke, F. (2016). Experimental induction of bacterial resistance to the antimicrobial peptide tachyplesin I and investigation of the resistance mechanisms. *Antimicrob. Agents Chemother.* 60, 6067–6075. doi: 10.1128/AAC.00640-16
- Hong, J., Hu, J. Y., Liu, K., Wang, C., Zhou, C., Yan, K. M., et al. (2018). Characteristics and resistance of tachyplesin-I resistance in *Pseudomonas aeruginosa*. *Acta Microbiol. Sin.* 58, 1–12.
- Hong, J., Hu, J. Y., Qu, E. J., Shu, J., Liu, R. F., Zhou, Z., et al. (2013). Mechanisms of tachyplesin I against *Escherichia coli*. *Microbiol. China* 40, 1018–1026.
- Hong, J., Jiang, H. H., Hu, J. Y., Wang, L. Z., and Liu, R. F. (2020). Transcriptome analysis reveals the resistance mechanism of *Pseudomonas aeruginosa* to tachyplesin I. *Infect. Drug Resist.* 13, 155–169. doi: 10.2147/IDR.S226687
- Imura, Y., Nishida, M., Ogawa, Y., Takakura, Y., and Matsuzaki, K. (2007). Action mechanism of tachyplesin I and effects of PEGylation. *Biochim. Biophys. Acta* 1768, 1160–1169. doi: 10.1016/j.bbammem.2007.01.005
- Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 45, D566–D573. doi: 10.1093/nar/gkw1004
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280. doi: 10.1093/nar/gkh063
- Klein, S., Lorenzo, C., Hoffmann, S., Walther, J. M., Storbeck, S., Piekarski, T., et al. (2009). Adaptation of *Pseudomonas aeruginosa* to various conditions includes tRNA-dependent formation of alanylphosphatidylglycerol. *Mol. Microbiol.* 71, 551–565. doi: 10.1111/j.1365-2958.2008.06562.x
- Kushibiki, T., Kamiya, M., Aizawa, T., Kumaki, Y., Kikukawa, T., Mizuguchi, M., et al. (2014). Interaction between tachyplesin I, an antimicrobial peptide derived from horseshoe crab, and lipopolysaccharide. *Biochim. Biophys. Acta* 1844, 527–534. doi: 10.1016/j.bbapap.2013.12.017
- Kuzmin, D. V., Emel'yanova, A. A., Kalashnikova, M. B., Pantelev, P. V., and Ovchinnikova, T. V. (2018). In vitro study of antitumor effect of antimicrobial peptide tachyplesin I in combination with cisplatin. *Bull. Exp. Biol. Med.* 165, 220–224. doi: 10.1007/s10517-018-4134-6
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Liu, C., Qi, J., Shan, B., and Ma, Y. (2018). Tachyplesin causes membrane instability that kills multidrug-resistant bacteria by inhibiting the 3-Ketoacyl carrier protein reductase fabG. *Front. Microbiol.* 9:825. doi: 10.3389/fmicb.2018.00825
- Nakamura, T., Furunaka, H., Miyata, T., Tokunaga, F., Muta, T., Iwanaga, S., et al. (1988). Tachyplesin, a class of antimicrobial peptide from the hemocytes of the horseshoe crab (*Tachyplesus tridentatus*). Isolation and chemical structure. *J. Biol. Chem.* 263, 16709–16713. doi: 10.1016/s0021-9258(18)37448-9
- Puja, H., Bolard, A., Noguès, A., Plésiat, P., and Jeannot, K. (2020). The efflux pump MexXY/OprM contributes to the tolerance and acquired resistance of *Pseudomonas aeruginosa* to colistin. *Antimicrob. Agents Chemother.* 64:e02033-19. doi: 10.1128/AAC.02033-19
- Shepherd, J., and Ibba, M. (2013). Direction of aminoacylated transfer RNAs into antibiotic synthesis and peptidoglycan-mediated antibiotic resistance. *FEBS. Lett.* 587, 2895–2904. doi: 10.1016/j.febslet.2013.07.036
- Wagner, E., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* 90, 133–208. doi: 10.1016/bs.adgen.2015.05.001
- Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612
- Xie, H., Wei, J., and Qin, Q. (2016). Antiviral function of tachyplesin I against iridovirus and nodavirus. *Fish. Shellfish Immun.* 58, 96–102. doi: 10.1016/j.fsi.2016.09.015

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.871290/full#supplementary-material>

- Zahedi bialvaei, A., Rahbar, M., Hamidi-Farahani, R., Asgari, A., Esmailkhani, A., Mardani Dashti, Y., et al. (2021). Expression of RND efflux pumps mediated antibiotic resistance in *Pseudomonas aeruginosa* clinical strains. *Microb. Pathog.* 153:104789. doi: 10.1016/j.micpath.2021.104789
- Zhang, Y. F., Han, K., Chandler, C. E., Tjaden, B., Ernst, R. K., and Lory, S. (2017). Probing the sRNA regulatory landscape of *P. aeruginosa*: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility. *Mol. Microbiol.* 106, 919–937. doi: 10.1111/mmi.13857

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hong, Li, Jiang and Hong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Prediction of Antibiotic Susceptibility Profiles of *Vibrio cholerae* Isolates From Whole Genome Illumina and Nanopore Sequencing Data: CholerAegon

Valeria Fuesslin^{1†}, Sebastian Krautwurst^{2†}, Akash Srivastava², Doris Winter^{1,3}, Britta Liedigk¹, Thorsten Thye¹, Silvia Herrera-León⁴, Shirlee Wohl⁵, Jürgen May^{1,3,6}, Julius N. Fobil⁷, Daniel Eibach¹, Manja Marz^{2†} and Kathrin Schuldt^{1*†}

OPEN ACCESS

Edited by:

Ibrahim Bitar,
Charles University, Czechia

Reviewed by:

Adam Valcek,
Vrije University Brussel, Belgium
Marc Finianos,
Charles University, Czechia

*Correspondence:

Kathrin Schuldt
schuldt@bnim.de

[†]These authors have contributed
equally to this work and share first
authorship

[‡]These authors have contributed
equally to this work and share last
authorship

Specialty section:

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 31 March 2022

Accepted: 02 May 2022

Published: 22 June 2022

Citation:

Fuesslin V, Krautwurst S, Srivastava A,
Winter D, Liedigk B, Thye T,
Herrera-León S, Wohl S, May J,
Fobil JN, Eibach D, Marz M and
Schuldt K (2022) Prediction of
Antibiotic Susceptibility Profiles of
Vibrio cholerae Isolates From Whole
Genome Illumina and Nanopore
Sequencing Data: CholerAegon.
Front. Microbiol. 13:909692.
doi: 10.3389/fmicb.2022.909692

¹ Infectious Disease Epidemiology Department, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany, ² RNA Bioinformatics and High-Throughput Analysis, Friedrich Schiller University Jena, Jena, Germany, ³ German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems, Hamburg, Germany, ⁴ National Center of Microbiology, Institute of Health Carlos III, Madrid, Spain, ⁵ Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States, ⁶ Tropical Medicine II, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany, ⁷ Department of Biological, Environmental and Occupational Health Sciences, School of Public Health, University of Ghana, Accra, Ghana

During the last decades, antimicrobial resistance (AMR) has become a global public health concern. Nowadays multi-drug resistance is commonly observed in strains of *Vibrio cholerae*, the etiological agent of cholera. In order to limit the spread of pathogenic drug-resistant bacteria and to maintain treatment options the analysis of clinical samples and their AMR profiles are essential. Particularly, in low-resource settings a timely analysis of AMR profiles is often impaired due to lengthy culturing procedures for antibiotic susceptibility testing or lack of laboratory capacity. In this study, we explore the applicability of whole genome sequencing for the prediction of AMR profiles of *V. cholerae*. We developed the pipeline CholerAegon for the *in silico* prediction of AMR profiles of 82 *V. cholerae* genomes assembled from long and short sequencing reads. By correlating the predicted profiles with results from phenotypic antibiotic susceptibility testing we show that the prediction can replace *in vitro* susceptibility testing for five of seven antibiotics. Because of the relatively low costs, possibility for real-time data analyses, and portability, the Oxford Nanopore Technologies MinION sequencing platform—especially in light of an upcoming less error-prone technology for the platform—appears to be well suited for pathogen genomic analyses such as the one described here. Together with CholerAegon, it can leverage pathogen genomics to improve disease surveillance and to control further spread of antimicrobial resistance.

Keywords: antimicrobial resistance, AMR, *Vibrio cholerae*, nanopore, MinION, Illumina

1. INTRODUCTION

The vast number of pathogenic bacterial strains that are resistant to antimicrobial agents today has become a difficult public health issue (World Health Organization, 2014). Early identification and epidemiological surveillance of resistance in bacteria are key in order to guide the appropriate use of antimicrobials and to prevent further spread of antimicrobial resistance (AMR).

Commonly, antimicrobial susceptibility testing (AST) is performed phenotypically *in vitro* by a microbiological procedure based on the growth of a single bacterial isolate in the presence or absence of an anti-microbiological agent. Especially for clinical settings phenotypic AST represents an important method which can deliver robust and reproducible results. However, the phenotypic AST is limited in the sense that it only allows for assessing resistance to known antibiotics and due to the time needed for the growth of the microorganism results are often not provided timely. Depending on the microorganism in question it may take up to several days after the isolation of a clinical isolate until an AMR profile can be provided (Didelot et al., 2012; van Belkum et al., 2020). Moreover, *in vitro* AST is difficult to scale-up and requires laboratory facilities with suitable equipment, trained microbiologists, as well as the availability of the antibiotics in question. In low- and middle-income countries (LMICs), where AMR rates appear to be higher than in high income countries (HICs) (World Health Organization, 2021), these requirements are often not given and the lack of laboratory capacity hinders the necessary assessment of AMR in bacteria causing life-threatening infections.

AMR in bacteria is genetically encoded and can be intrinsically mediated or acquired by *de novo* mutations or the intake of new resistance genes from other organisms. Hence, an alternative approach to analyze the presence of AMR in bacteria is whole genome analysis, allowing to quickly identify the presence of resistance genes and resistance-conferring mutations (Didelot et al., 2012; Boolchandani et al., 2019). For instance, whole genome sequencing methods have been successfully implemented to accelerate diagnostics and prediction of drug resistance profiles of *Mycobacterium tuberculosis* (Consortium et al., 2018; Doyle et al., 2018). Multiple methods have been developed to detect AMR genes directly from sequencing read from isolates or from metagenomic reads, as well as from assembled genomes (for a detailed review see Boolchandani et al., 2019). Most tools utilize alignment and/or HMM strategies to find genes with individual curated AMR gene reference databases, e.g., Resistance Gene Identifier (CARD) (Alcock et al., 2019), Resfinder (Bortolaia et al., 2020), ARG-ANNOT (Gupta et al., 2014), or the NCBI database used for AMRfinderPlus (Feldgarden et al., 2021). ABRicate (Seemann, 2016) integrates all these databases with a BLAST alignment method, but is restricted to homolog models.

Both, traditional short read sequencing (e.g., Illumina) and long read sequencing (e.g., Oxford Nanopore Technologies - ONT, Pacific Biosciences) are commonly used for the genome assembly of bacterial isolates (De Maio et al., 2019). Long reads enable higher contiguity and can resolve long repeat regions, while short reads have lower error rates. Hybrid approaches harness the advantages of both types, leading to complete genomes with few assembly errors (De Maio et al., 2019). The portability of the MinION device by ONT allows in-field sequencing and rapid result turnaround, suited for applications using direct patient samples (Maestri et al., 2019; Leggett et al., 2020).

Cholera is an acute diarrheal disease and nowadays considered a neglected disease mainly affecting populations in low-income countries, where contaminated water and poor sanitation are the main drivers for the bacterial spread (Lippi et al., 2016). With approximately 2.9 million cases and an estimated 95,000 cholera-associated deaths each year the vast majority of cholera cases are observed on the African continent (Clemens et al., 2017). The disease's causing agent is the Gram-negative bacterium *Vibrio cholerae*, which belongs to the family of *Vibrionaceae*, mostly found in aquatic environments (Momba and El-Liethy, 2017). *V. cholerae* comprises more than 200 O serogroups, of which serogroups O1 and O139 strains are pathogenic for the human host (Shimada et al., 1994; Momba and El-Liethy, 2017). The O1 serogroup is subdivided into two biotypes, the classical and the El Tor biotype, the latter being responsible for the current pandemic (Ramamurthy et al., 2019).

The standard treatment for acute watery diarrhoea caused by *V. cholerae* is oral rehydration therapy (ORT) (Nalin and Cash, 2018). Antibiotic therapy is recommended for severely ill cholera patients, who have severe dehydration and continue to pass a large volume of stool during rehydration treatment, pregnant women, and patients with comorbidities, such as an infection with the human immunodeficiency virus (HIV) or severe malnutrition. The administration of antibiotic regimen helps to shorten the length of disease and to limit the number of infectious bacteria in the stool (Ramamurthy et al., 2019; Kaunitz, 2020). In recent years, however, treatment failures due to the recurrent emergence of antimicrobial resistant *V. cholerae* have been observed frequently (Clemens et al., 2017). The occurrence of *V. cholerae* strains that are not susceptible to antibiotics is especially challenging since the bacteria are able to pass on resistance-mediating genetic sequences as part of a highly mobile elements conferring antimicrobial resistance to other bacterial species *via* horizontal gene transfer (Kitaoka et al., 2011; Verma et al., 2019; Das et al., 2020).

In this study, our overarching aim is to assess the applicability of whole genome sequencing with short and long reads for the prediction of clinically relevant AMR in *V. cholerae*. We subjected 82 *V. cholerae* isolates to whole genome sequencing and established a bioinformatic analysis workflow for the identification of genetic resistance loci in *V. cholerae*. Subsequently, we correlated the results from the whole genome analysis to results from phenotypic *in vitro* characterization by AST for the same isolates. Furthermore, we estimate time and costs needed to apply this approach for the prediction of AMR profiles from bacterial genomes.

2. MATERIALS AND METHODS

2.1. *V. cholerae* Isolates

Patient samples were collected during cholera outbreaks in Ghana during the years 2011, 2012, and 2014 by the Ghanaian National Public Health & Reference Laboratory (NPHRL) as described by Eibach et al. (2016). In this study, 82 of the previously characterized *V. cholerae* isolates were subjected to whole genome sequencing (2011 *n* = 16, 2012 *n* = 8, 2014 *n* = 58).

2.2. Phenotypic Antibiotic Susceptibility Testing

The characterization of the phenotypic AST of the here sequenced isolates has been described previously (Eibach et al., 2016). Briefly, AST was performed for 80 isolates by the Kirby-Bauer disk diffusion method for the following seven antimicrobial drugs: ampicillin, chloramphenicol, gentamicin, nalidixic acid, sulfamethoxazole/trimethoprim, tetracycline and ciprofloxacin. Classifications of isolates into susceptible (S), susceptible dose-dependent (SD), intermediate (I), and resistant (R) phenotypes were done according to the 2015 Clinical and Laboratory Standards Institute (CLSI) guidelines (<http://www.clsi.org>) and when no interpretive criteria were available for *V. cholerae*, breakpoints for *Enterobacteriaceae* based on the 2015 European Committee on Antimicrobial Susceptibility Testing (EUCAST, <http://www.eucast.org>) were applied.

2.3. Culture and DNA Isolation

Vibrio cholerae isolates were stored in skim milk at -20°C. Culturing was performed on Columbia Agar with sheep blood. Plated bacteria were incubated at 26,5°C for 20–24 h and were then stored at 4°C for a maximum of 7 days. The LGC MasterPure Complete DNA and RNA Purification Kit (Lucigen Corporation, Cat. No. MC85200) was used for DNA extraction from the 82 *V. cholerae* isolates. Quality control of the isolated DNA was done by fluorometric quantification using the Quant-iTTM broad-range dsDNA assay kit (Thermo Fisher Scientific) and gel electrophoresis.

2.4. Illumina Library Preparation and Sequencing

Illumina library preparation was performed using the Nextera XT Library Prep Kit and the IDT for Illumina Nextera DNA Unique Dual Indexes Set A (both Illumina Inc.). Sequencing was performed on Illumina NextSeq with the NextSeq 500/550 Mid Output Reagent Cartridge v2.5 (Illumina Inc). Quality assessments of short reads were performed using fastqc (Andrews, 2010). The number of reads etc. can be found in the results and **Supplementary Material**.

2.5. Oxford Nanopore Technologies Library Preparation and Sequencing

An input mass of 1 µg of DNA was used for each sample for library preparation, following the instructions of ONT 1D Native barcoding genomic DNA protocol from May 2018. The following procedures complemented the ONT protocol: DNA repair was carried out using the NEBNext Companion Module (New England Biolabs). This step was followed by the barcode ligation using NEB Blunt/TA Ligase (New England Biolabs) and EXP-NBD104 native barcode kit (ONT). After barcoding, 10–12 DNA samples were pooled resulting in a total of 1 µg DNA. For adapter ligation, the NEBNext Quick Ligation Module (New England Biolabs) and SQK-LSK109 ligation sequencing kit (ONT) were used. Throughout the protocol, AMPure XP beads (Beckman Coulter) were used to purify the samples. After

DNA repair and barcode ligation the Elution Buffer (Promega Corporation) was selected for DNA Elution. Following adapter ligation, the DNA was eluted with elution buffer of SQK-LSK109 ligation sequencing kit (ONT). MinION sequencing was performed on the SpotON Flow Cell with the Flow Cell Priming Kit (Flow Cell Type R9.4.1, ONT). On average, ~250 ng of library DNA were loaded onto the flow cell for one run.

2.6. Basecalling of ONT Data

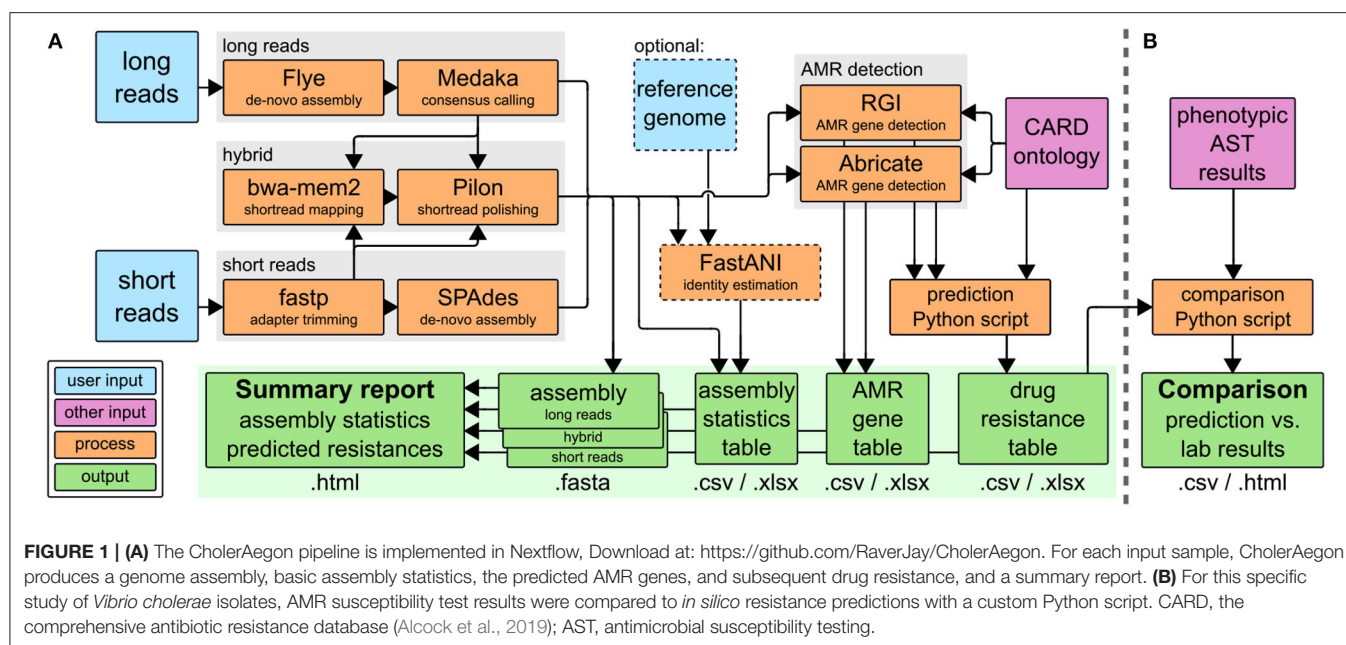
Conversion from raw data to nucleotide sequences (“basecalling”) of the raw nanopore data was performed with Guppy (available from Oxford Nanopore community portal) v5.0.16 using the “superior accuracy” configuration (dna_r9.4.1_450bps_sup.cfg). Guppy was run on HPC cluster nodes equipped with double NVIDIA Tesla GPUs (P100 16 GB or V100 16 GB) with the following parameter settings: enabled barcode demultiplexing for the utilized barcoding kit (-barcode_kits EXP-NDB104), usage of both GPUs (-device “cuda:0 cuda:1”), enabled barcode trimming (-trim_barcodes), disabled filtering of reads by median accuracy (-disable_qscore_filtering), and disabled sending of telemetry data (-disable_pings). The resulting fastq files were merged into a single fastq for each barcode. Quality control for the long reads was performed with pycoQC (Leger and Leonardi, 2019) v2.5.0.3.

2.7. CholerAegon Pipeline

We implemented the main analysis steps for short and long read data in an automated pipeline called CholerAegon. The pipeline is implemented in Nextflow (Tommaso et al., 2017) and Python3, and carries out short read assembly, long read assembly, and/or hybrid assembly, followed by assembly polishing, and finally assessment of resistance gene presence. Based on the found genes, CholerAegon predicts the antibiotic resistances of the sequenced isolates based on the CARD ontology (Alcock et al., 2019). We applied CholerAegon to 82 sequenced isolates. The individual steps of the pipeline are described in the next sections.

2.8. Genome Assembly

A *short read only* assembly was performed with Spades (Bankevich et al., 2012) v3.15.2 with default parameters except for the number of concurrent computation threads (-t 12). *Long read only* assembly was performed by running Flye (Kolmogorov et al., 2019) v2.9 with parameters set to treat input data as uncorrected nanopore reads (-nano-raw), and use multiple concurrent threads (-threads). Draft assemblies were polished with Medaka (ONT, 2018) v1.4.4 with the model (r941_min_sup_g507) which corresponds to the Guppy model used during basecalling. *Hybrid assembly* was performed by applying Pilon (Walker et al., 2014) v1.23 to the long read assembly, which uses the short reads to polish the long read draft assembly. Parameters were set to allocate 8 GB of memory (-Xmx8g) and use multiple threads (-threads 12). This approach harnesses the benefits of both data types: long reads lead to long contiguous sequences (and commonly complete genomes for bacterial samples like *V. cholerae*), and the very low error rate of the short reads can correct localized assembly



errors. Average nucleotide identity (ANI) to the reference sequence (*V. cholerae* O1 biovar El Tor str. N16961, accession NC_002505.1) was determined with FastANI (Jain, 2018) v1.32 using default parameters.

2.9. Assessment of Resistance Loci

For each bacterial isolate and each assembly type, we performed the assessment of resistance gene presence with Resistance Gene Identifier (Alcock et al., 2019) (RGI) and Abricate (Seemann, 2016). RGI uses the full CARD database, which includes protein homolog models as well as sequence variant models. Abricate can only use protein homolog models and was used with the respective sequences for resistance genes in the CARD (Alcock et al., 2019) database. CholerAegon uses the RGI results in order to include available protein variant models. For each detected resistance gene, CholerAegon reviews Abricate results to yield hits with higher reference gene coverage. Finally, CholerAegon will filter out false positives by applying coverage and identity cutoffs (minimum 80 and 80%, respectively) on the remaining hits.

2.10. Comparison With Phenotypic Characterization

The correlation of the predicted antibiotic resistance from the whole genome sequencing data with the results from the phenotypic AST was performed with custom Python scripts. Specifically, we predicted an isolate as resistant to an antibiotic, if any of the detected genes have a `confers_resistance_to_antibiotic` relation to that antibiotic in the CARD ontology. Additionally, if a gene has a `confers_resistance_to_drug_class` related to a class of antibiotics, we considered the isolate resistant to all members of that drug class. Resistance against multi-component drugs are only assumed if resistance were predicted against all

components. For each drug in the AST panel, we then compared the phenotypic AST result to the *in silico* predicted resistance profile.

We consider a prediction correct if for the “susceptible” (S) phenotype, no resistance has been predicted (no relevant gene or gene variant present), or if for the “intermediate” (I), “susceptible dose-dependent” (SD) and “resistant” (R) statuses, the resistance was predicted (relevant gene or gene variant found). All other combinations are considered false predictions.

3. RESULTS AND DISCUSSION

3.1. CholerAegon—A Fully Automated Nextflow Pipeline for Antibiotic Resistance Prediction

The CholerAegon pipeline, refer to **Figure 1A** assembles reads, annotates antibiotic resistance genes based on homology models and predicts the resistance of any sequenced bacterial isolate. CholerAegon can process short reads such as Illumina reads, for which after an initial adapter trimming an assembly is created. In case long reads (created with e.g., ONT MinION) serve as input, a *de novo* assembly with a subsequent polishing step is given. In case long and shorts read are given only the hybrid assembly is used for further analysis.¹ We evaluated our long-reads-first hybrid approach (Flye+Pilon) against Unicycler (Wick et al., 2017) v0.5.0, a short-reads-first hybrid assembler, and found that our approach outperformed Unicycler in terms of runtime and ANI (refer to **Supplementary Figure 2**). A basic assembly statistics table is given, which contains the total length of assemblies, number of contigs and N50 measure. In

¹The parameter `-do_all_assemblies` allows proceeding with all three assemblies.

case a reference genome is given by `-genome_reference`, additionally, the ANI of the assembly is reported.

Furthermore, CholerAegon detects the AMR genes by combining the results of two tools: Abricate with RGI. While Abricate detects homologs more accurate than RGI, the latter is able to detect genetic variants of AMR genes. We chose these tools for their capability to detect more genes than other tools, such as AMRFinderPlus or Resfinder (see **Supplementary Table 5**). The default minimum thresholds for the identity and coverage of detected genes are set to 80% and 80%, respectively (adopted from Abricate), but can be customized with the parameters `-min_identity_percent` and `-min_coverage_percent`.

As a result, for each bacterial isolate, an antibiotic drug resistance profile is provided. Previously, two web-based tools have been developed that detect *V. cholerae* genotypes from sequencing data, which also allow to screen for the presence of AMR-conferring genes: CholeraeFinder 1.0 and VicPred (Siriphap et al., 2017; Lee et al., 2021). CholeraeFinder 1.0 uses Resfinder and VicPred RGI for AMR prediction. Due to the integration of the two best performing tools in our case, CholerAegon provides the most comprehensive AMR profiles (refer to **Supplementary Table 5**).

In case the user has performed wet-lab experimental AST, refer to **Figure 1B**, we give a comparison of predicted drug resistance (**Figure 1A**) with AST results.

Finally, the summary report (*.html output) collects all results as described above, complemented by the total run time.

The CholerAegon pipeline has been applied and in detail analyzed in this study for *V. cholerae*. However, individual processes of the pipeline can readily be improved or replaced. CholerAegon is configured to run all processes in virtual environments to ensure complete reproducibility. Docker containers are used for each program, which guarantees the exact same behavior independent of the hardware architecture used to run the pipeline. CholerAegon utilizes Bioconda's (Grüning et al., 2018) pre-built containers hosted on `quay.io` for most programs. Updating any process is easily accomplished by switching to a new container version. This modular design also allows for ease of expansion or adaptation for other bacterial pathogens.

3.2. *Vibrio cholerae* Assemblies Significantly Improved by Nanopore Sequencing

We obtained about 7 Mio. reads of length 73 nt (throughput 522 Mb and 111 X) per bacterial isolate with Illumina sequencing and on average 94,000 reads of about 5.8 kb in length (920 Mb throughput and 209 X) with ONT MinION sequencing. The length of long reads and subsequently the throughput varied by about one order of magnitude between the different bacterial isolates, refer to **Supplementary Figure 1**.

The advantage of the Illumina sequencing technology of a lower sequencing error rate is compromised by an average assembly length of about 4,042 Mb, whereas the long read sequencing strategy resulted in the full assembly length of

4,107 Mb, similar to a combined hybrid assembly strategy, refer to **Table 1** and **Supplementary Figure 1**. Assemblies based on Nanopore data yielded the two chromosomes present in the *V. cholerae* genome *V. cholerae* (~3 Mb and ~1 Mb in length; **Table 1**) in complete contiguity for all isolates (erroneously fused in two strains). Illumina assemblies were highly fragmented but achieved a higher genome identity to the reference sequence. The hybrid assemblies harness the advantages of both sequencing approaches yielding few assembly errors at high sequence contiguity.

The average accuracy of the assembled genomes was 99.942% for ONT MinION, 99.978% for Illumina, and 99.974% for the hybrid assembly. Therefore, when being interested in the absence or presence of genes, their genomic orientation, larger gene duplications and more importantly plasmids reconstructions, as desired for predicting AMR based on gene presence, we recommend using ONT MinION sequencing (Lemon et al., 2017; Golparian et al., 2018). Because plasmids are rare in pandemic *V. cholerae* strains (Jaskólska et al., 2022) we did not apply a particular library preparation method for recovery of plasmids in this study. However, when applying this method to other bacterial species, where small plasmids play a major role in conferring AMR, the ONT rapid library preparation method may be more suitable (Wick et al., 2021).

Contrary, when focusing on specific antibiotic agents, for which resistance phenotypes are associated with the

TABLE 1 | General statistics of whole genome sequencing data.

	Illumina	ONT	Hybrid
# reads	7,055,057	93,778	–
Read length	74	5,768	–
Throughput	521.8Mb	920.0 Mb	–
Coverage	111 X	209 X	–
[assembly]	4,042,192	4,107,176	4,106,391
# contigs	275	2	2
% ANI	99.97815	99.94205	99.97385
AMR (predicted from 82 <i>V. cholerae</i> genomes)			
APH(3")-Ib	74	74	74
APH(6)-Id	74	74	74
CRP	82	82	82
VC varG	79	79	79
almG	82	82	82
catB9	79	79	79
dfrA1	79	79	79
floR	74	74	74
rsmA	82	82	82
sul2	74	74	74
EC parE	82	35	82

Displayed are average numbers over the 82 isolates. Details for each isolate can be viewed in **Supplementary Table 1**. Gene identifiers are the exact names used by CARD. # reads, average number of reads; [assembly], average assembly length in nt; # contigs, median number of contigs; % ANI, average nucleotide identity; VC varG, *Vibrio cholerae* varG; EC parE, *Escherichia coli* parE conferring resistance to fluoroquinolones.

presence of single nucleotide polymorphisms (SNPs), e.g., the fluoroquinolone Ciprofloxacin, we recommend Illumina sequencing (Dahl et al., 2021; Hodges et al., 2021). Currently, the ONT MinION sequencing produces too many errors for a reliable SNP analysis, however, with the upcoming R10 pore model of ONT MinION, we expect both sequencing methods to be equal in performance.

In summary, for the prediction of AMR, we conclude, that ONT MinION sequencing is similarly suitable for broad AMR studies as Illumina sequencing. However, ONT MinION sequencing brings the advantages of in-field studies, and in case a clinical sample is applied directly, can provide a faster result (refer below) and financial handling (refer below).

3.3. Antibiotic Resistance Genes Present in Illumina and Nanopore Sequencing

In total, ABricate and RGI detected ten different resistance-conferring genes against different antibiotics for the 82 *V. cholerae* isolates: *APH(3'')-Ib*, and *APH(6)-Id* confer resistance to aminoglycosides; *rsmA* against fluoroquinolones, diaminopyrimidines and phenicols; *CRP* against macrolides, fluoroquinolones, and penams; *dfrA1* against diaminopyrimidines; *floR* and *catB9* against phenicols; *Vibrio cholerae* *varG* against carbapenems; *almG* against polymyxins; and *sul2* against sulfonamides. There was no difference detected in Illumina and ONT MinION reads regarding the detection of AMR genes, refer to **Table 1**.

Several point mutations present in the gene *parE* in *V. cholerae* confer resistance against fluoroquinolones (Zhou et al., 2013) (*EC parE* in **Table 1**). Only RGI is able to detect this type of genetic variation. Accordingly, this resistance marker could be reliably detected in the accurate Illumina read based and in the hybrid assembly, whereas ONT reads this genetic variant was detected only in 35 of the 82 isolates analyzed.

3.4. Sequencing Can Replace AST for Five of Seven Antibiotics Tested

Independent of the chosen sequencing method, phenotypic AST can be replaced for most but not all antibacterial agents tested in this study. In the following, we present results from the hybrid assembly method for 80 isolates, for which both, sequencing as well as phenotypic AST data, were generated (refer to Section 2.2).

3.4.1. Trimethoprim/Sulfamethoxazole (SxT)

In *V. cholerae*, the genes *dfrA1* and *sul2* are encoding trimethoprim and sulfamethoxazole resistance, respectively (Das et al., 2020). AST resulted in resistance to SxT for 77 isolates, and computationally, we were able to predict resistance correctly for 72 isolates, for which both genes were identified. In four of the five remaining phenotypically resistant isolates *dfrA1* was detected and *sul2* was missing, which should theoretically result in susceptibility of the bacterial isolates to SxT; and for one isolate, although resistant in the AST, none of the two genes was detected in the assembled genome. Three isolates were susceptible to SxT and CholerAegon correctly did not identify both genes, refer to **Table 2** and for details

TABLE 2 | Numbers in boxes indicate the number of isolates, for which a gene (responsible for this phenotype) has been found.

Genes	SxT	NA	CIP	CN	TE	C	AMP
	<i>dfrA1</i> + <i>sul2</i>	<i>EC parE</i>	<i>EC parE</i>	<i>var</i>	<i>var</i>	<i>catB9, floR</i>	<i>var</i>
S	0/3	17/17	18/18	0/80	0/80	77/80	0/20
SD	–	–	62/62	–	–	–	–
I	–	–	–	–	–	–	0/60
R	72/77	63/63	–	–	–	–	–
Correct	75	63	62	80	80	3	20
False	5	17	18	0	0	77	60
Seq?	YES	YES	YES	YES	YES	NO	NO

Details can be found in the **Supplementary Material**. S, susceptible; SD, susceptible dose-dependent; R, resistant; I, intermediate; according to CLSI and EUCAST guidelines 2015 (see Section 2.2). Seq?, can sequencing replace AST testing?; AMP, ampicillin; C, chloramphenicol; CIP, ciprofloxacin; CN, Gentamicin A; TE, tetracycline; SxT, trimethoprim-sulfamethoxazole; NA, nalidixic acid; EC *parE*, *E. coli parE* conferring resistance to fluoroquinolones; var, various genes.

Supplementary Table 4. Despite five discrepancies found we come to the conclusion that sequencing can replace AST testing for SxT, emphasizing that, in practice, the detection of either *dfrA1* or *sul2* could be considered sufficient to computationally classify a bacterial isolate as resistant. In these cases, in order to ensure treatment success, SxT would not be recommended for treatment.

3.4.2. Nalidixic Acid (NA), Ciprofloxacin (CIP)

In *V. cholerae* quinolone resistance is mediated by point mutations in the genes encoding Topoisomerase IV and DNA Gyrase (topoisomerase II) enzymes, *parCE* and *gyrAB*, respectively (De, 2021). An isolate exhibits resistance to NA or CIP when SNPs are present in these genes, protecting the bacterial enzyme from inhibition by the antibiotic agent (Aldred et al., 2014). For NA, AST resulted in 63 resistant and 17 susceptible isolates. With respect to CIP, the majority of isolates (62) appeared to be AB dose dependent susceptible, and 18 isolates were susceptible isolates. For all 80 isolates the D476N mutation in the *EC parE* gene has been found, providing evidence that the sole presence of this variant appears not sufficient to confer resistance to NA and CIP in approximately one quarter of our *V. cholerae* isolates. As a result, if AST was replaced by sequencing, for 63 and 62 of the 80 isolates a treatment with NA and CIP, respectively, would not be recommended, which is in accordance with the AST results. In the case of the susceptible 17 and 18 isolates, however, the NA/CIP treatment would not be recommended, although it would have been possible to manage the infection with these ABs.

3.4.3. Gentamicin A (CN), Tetracycline (TE)

Differing resistance mechanisms to both aminoglycosides, CN and TE, have been described and the respective genes in *V. cholerae* are located on self-transmissible plasmids (Garneau-Tsodikova and Labby, 2016; Das et al., 2020). In our study, the phenotypic AST found in 80/80 and 80/80 isolates to be susceptible to CN and TE, respectively. In accordance with these

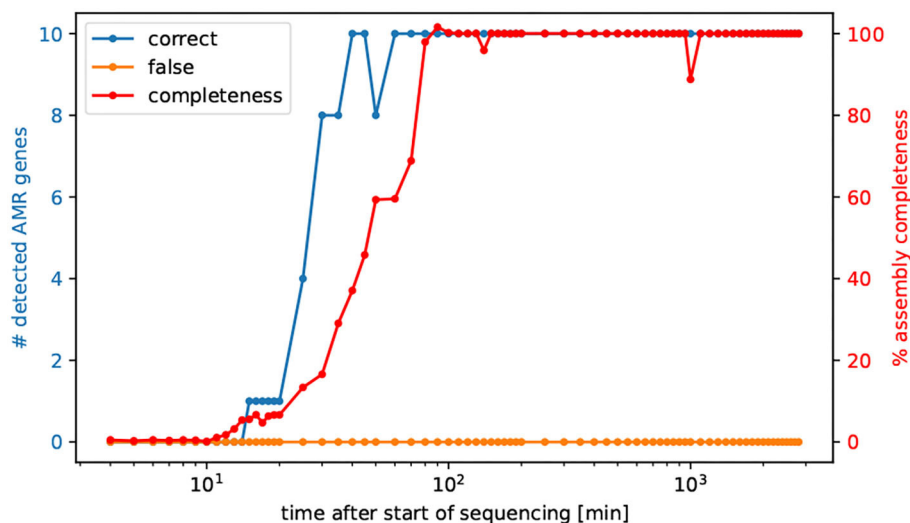


FIGURE 2 | Number of detected AMR genes by the elapsed time of sequencing. We selected cumulative temporal subsets of the nanopore read data of isolate Iso02507. We used CholerAegon for assembly and subsequent AMR detection on these subsets. With 10 multiplexed isolates on this flowcell, all resistance genes from homolog models are detectable in the assembly after 60 min.

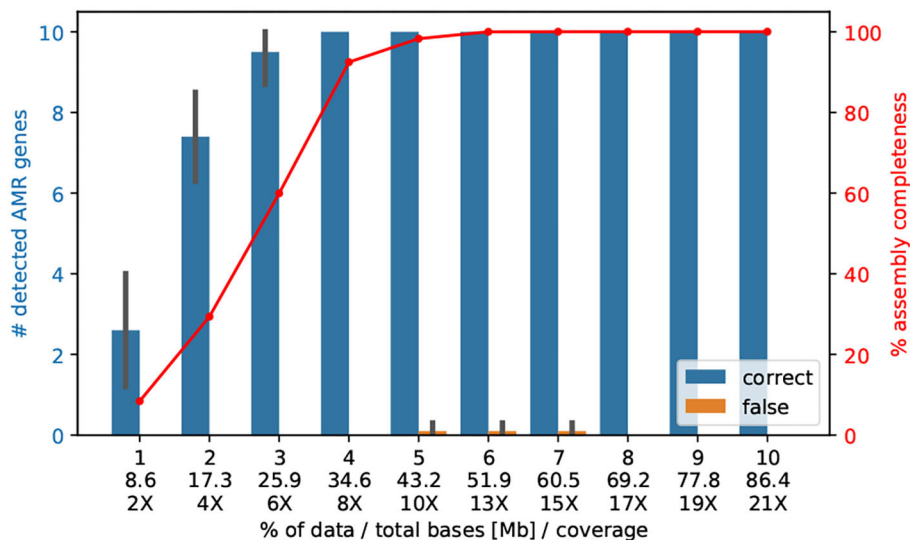


FIGURE 3 | The number of AMR genes found with a randomly selected lower amount of Nanopore data. Here, we subsampled the Nanopore data ($n = 10$ replicates) of isolate Iso02507 at various percentages and performed assembly and AMR detection with CholerAegon for all subsets. About 4% (35 Mb) of the total Nanopore data (864 Mb) is enough to achieve the same detection result as using the full data for protein homolog models of AMR genes. This demonstrates that a throughput of ~35 Mb (8X coverage of the 4.1 Mb genome) is sufficient for these models.

in vitro findings, for both antibiotics no resistance-mediating gene was found when applying CholerAegon. We conclude that the prediction of the profile for these two ABs by sequencing is possible.

3.4.4. Chloramphenicol (C)

Similar to the aminoglycoside antibiotics, a number of genes have been associated with resistance to C in *V. cholerae*, among them *catB9* and *florB*, encoding a C acetyl transferase (CAT) and a C efflux pump (Alcock et al., 2019; Das et al., 2020), respectively.

AST has shown that all 80 isolates had a susceptible phenotype to C. Of these, CholerAegon identified three as susceptible, too. However, at least one of the two resistance genes *catB9* or *florB* was detected in 77 isolates, which—based on the genome analyses only—would have led to a resistant isolate. Although a medical treatment based on the sequencing results would not cause any harm to the patient (due to avoiding treatment with C), we would not recommend replacing AST with sequencing for C, because in almost all cases this treatment would have helped the patient.

3.4.5. Ampicillin (AMP)

With regard to AMP, several genes are known to be responsible for AMR (Das et al., 2020). In the AST, 20 isolates have been found to be susceptible and 60 isolates have been found with an intermediate phenotype. Whereas for the 20 susceptible isolates the analysis of the genomes correctly did not reveal any resistance-associated gene, also no genetic resistance marker was found for the 60 intermediate phenotypes. In the case of AMP the AST intermediate phenotype could not be identified by the genome analysis. Possibly, the intermediate phenotypes maybe explained by resistance mechanisms that are independent of resistance genes, such as protein-regulatory mechanisms, which have been described in *V. cholerae* O1 strains resistant to ampicillin (Nguyen et al., 2009).

As described above, for NA, CIP, and C we observed discrepancies between the predicted *in silico* and *in vitro* AMR profiles. In these cases, *V. cholerae* isolates were found susceptible to AST, but resistance genes were found to be present. It is important to mention, that for phenotypic AST the classification into susceptible, resistant, dose-dependent and intermediate phenotype is solely defined by the breakpoints used. As a consequence, when applying different or updated breakpoints, as described in a recent study, correlation results may slightly differ (Opintan et al., 2021).

3.5. *Vibrio cholerae* Resistance Genes Visible After 40 min of Nanopore Sequencing

To successfully eliminate a pathogenic bacterial infection the time between diagnosis and initiation of treatment with an effective antibiotic drug is crucial. Hence, tracking the sequenced reads in minutes, we asked after how many minutes all AMR genes are detectable for the prediction of AMR, and after how much time of sequencing the entire genome is possible to assemble. Interestingly after already 60 min all AMR genes of *V. cholerae* are detectable in the same quality as the fully sequenced and assembled genomes, refer to **Figure 2**. More precisely, after already 40 min all AMR genes could be identified, however not in full length, refer to **Supplementary Table 3**.

For this study, we used bacterial isolates that had already undergone the selective culturing process and were used subsequently for DNA extraction and library preparation. In our case, the time needed for culturing was 1 day, and DNA extraction and library preparation took ~16h. In order to effectively apply this method directly sequencing a clinical sample from a patient with acute watery diarrhea could speed up the process. But, because of the plethora of bacterial species present in a stool sample, this approach will likely require a higher sequencing depth.

3.6. Ninety-Six Bacteria of About 4 Mb Can Be Sequenced for About 30 € on ONT MinION

We sequenced 10–12 genomes of *V. cholerae* (~4.1 Mb) on one ONT MinION Flowcell and received an average coverage of 209 X, refer to **Table 1**. We raised the question if with a lower

TABLE 3 | Sequencing costs in Euro (€) for Illumina and ONT per bacterial isolate (*) with an average genome length of 4.1 Mb for different sequencing strategies and resulting coverages.

Reagent	Illumina			ONT	
	96 samples	1 sample	10 samples	96 samples	96 samples (**)
Ultrasound lysing	0	0	0	0	0
Phenol/ Chloroform	288	3	30	288	288
Qubit	96	1	10	96	96
Flowcell	1,102	545	545	545	545
Library	3,703	109	109	109	109
Barcoding	631	0	24	109	–
Enzyme	–	38	267	2,158	1,518
Qubit	–	3	34	288	100
Ampure	58	2	19	165	305
Total	5,878	701	1,058	3,758	2,961
Total/ Isolate*	61	701	105.8	39.15	30.84
Coverage/ Isolate*	49.54 X	2,090 X	209 X	21.77 X	21.77 X

(*) Primer-ligated barcoding: 96 primers are ligated to 96 isolates and used as barcodes. Costs refer to current prices in Germany as of the time of publication.

coverage we could also (i) assemble the entire genome correctly; and (ii) find the important AMR genes. In fact, already 4% of the reads (8.36 X) are sufficient enough to detect all 10 AMR genes², and only 5% of the reads (43.2 Mb, 10X coverage) are necessary to assemble the entire genome correctly into the two chromosomes, refer to **Figure 3**.

By not aiming for the optimal cutoff of sequencing depth, but instead of allowing double coverage, we could bring 96 *V. cholerae* onto one ONT MinION Flowcell, resulting in costs of <40€ per sample with a comfortable coverage of 21 X to reconstruct the *V. cholera* genome, refer to **Table 3**. If we use primer-ligated barcoding, we reduce the costs of the barcoding procedure and can achieve a price of 31 per bacterial isolate with a genome size of 4.1 Mb.

The cost calculation for Illumina sequencing (refer to Material and Methods section 2.4) was based on a maximum output of 19.5 Gb, whereas for ONT (refer to Material and Methods section 2.5) 8.6 Gb data has been generated. During the preparation of the manuscript, the new chemistry of the SQK-LSK110 ligation sequencing kit in conjugation with the R9.4.1 flowcell has improved the output capacity considerably (average output now: 18 Gb). Therefore, we can expect for 40€ even a coverage of 40 X, or lower the sequencing costs further per sample.

4. CONCLUSION

Here, we present CholerAegon, a bioinformatic pipeline for the prediction of bacterial AMR profiles from whole genome

²Ten AMR genes are defined as (1) APH(3'')-Ib; (2) APH(6)-Id; (3) CRP; (4) *V. cholerae* varG; (5) almG; (6) catB9; (7) dfrA1; (8) floR; (9) sul2; (10) rsmA.

sequencing data. CholerAegon integrates long and short sequencing reads to produce genome assemblies and predicts an AMR profile based on a combined approach including the advantages of the two tools RGI and Abricate. By applying CholerAegon to *V. cholerae* sequences, we show that the prediction of antibiotic resistance profiles of *V. cholerae* works efficiently and with a comparable outcome for the Illumina and the ONT MinION sequencing platforms. When interested in specific SNPs as a resistance marker, we would recommend the less error-prone technology of Illumina. For the detection of larger genomic variations, which represent most resistance makers found in bacteria, the ONT MinION technology would preferably be used, because genome assemblies are significantly improved when using long reads.

Subsequently, the predicted AMR profiles based on sequencing data were correlated with phenotypic AST results for seven antibiotic agents for the same *V. cholerae* isolates. The correlation revealed, that in the case of *V. cholerae*, *in vitro* AST can be replaced by whole genome data analysis for the following antibiotics: Trimethoprim/Sulfamethoxazole, Nalidixic acid, Ciprofloxacin, Chloramphenicol and Tetracycline but not for Chloramphenicol and Ampicillin.

With special emphasis on the applicability of this approach in areas where cholera outbreaks occur frequently, we analyzed time and costs of the sequencing for the portable ONT MinION platform. After 40 min of sequencing the coverage was sufficient in order to detect all resistance genes present in the *V. cholerae* genome. Additionally, by applying a custom-made barcoding strategy, we could reduce the sequencing costs to ~30€ to produce one 4.1 Mb-sized genome with sufficient coverage of 21X on the ONT MinION platform. In order to use the herein described approach as a standardized diagnostic tool in the future it will be necessary to establish specific criteria for ABs and organisms in question (e.g., minimum coverages) to ensure quality of sequencing data and reproducible results. Furthermore, the availability of curated AMR databases, in particular, species-centric databases with continuous integration of new resistance genes and mechanisms are key for this method.

The application will be particularly useful for strengthening AMR surveillance, at the individual, but also at the community

level where it may be used to produce resistome profiles from metagenomes. Hence, genome sequencing, together with CholerAegon can be implemented widely, as part of a preparedness strategy to ensure that countries are ready to rapidly detect and respond to AMR in *V. cholerae* and other bacteria to reduce the spread of AMR.

DATA AVAILABILITY STATEMENT

Whole genome sequences have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB51675 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB51675>). CholerAegon is freely available at <https://github.com/RaverJay/CholerAegon>. Supplementary code used for writing this manuscript is available at <https://github.com/RaverJay/CholerAegon/tree/main/paper>.

AUTHOR CONTRIBUTIONS

VF, SK, SH-L, SW, DE, JM, MM, and KS conceived and designed the experiments. VF, SK, AS, DW, BL, and SH-L performed the experiments. SK and MM developed CholerAegon. SK, VF, TT, MM, and KS analyzed the data. VF, SK, MM, and KS wrote the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Dr. Daniel Cadar and Heike Baum from the NGS core facility of the Bernhard Nocht Institute for Tropical Medicine for technical support. We thank the Carl-Zeiss-Stiftung (FKZ 0563-2.8/738/2), TWMMG DigLeben (5575/10-9), and DFG iDIV (FZT 118, 202548816) for financial support. Figures were finalized with Inkscape v1.0.2.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.909692/full#supplementary-material>

REFERENCES

- Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2019). CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935
- Aldred, K. J., Kerns, R. J., and Osheroff, N. (2014). Mechanism of quinolone action and resistance. *Biochemistry* 53, 1565–1574. doi: 10.1021/bi5000564
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Boolchandani, M., D'Souza, A. W., and Dantas, G. (2019). Sequencing-based methods and resources to study antimicrobial resistance. *Nat. Rev. Genet.* 20, 356–370. doi: 10.1038/s41576-019-0108-4
- Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345
- Clemens, J. D., Nair, G. B., Ahmed, T., Qadri, F., and Holmgren, J. (2017). Cholera. *Lancet* 390, 1539–1549. doi: 10.1016/S0140-6736(17)30559-7
- Consortium, C., Allix-Béguec, C., Arandjelovic, I., Bi, L., Beckert, P., Bonnet, M., Bradley, P., et al. (2018). Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. *N. Engl. J. Med.* 379, 1403–1415. doi: 10.1056/NEJMoa1800474
- Dahl, L. G., Joensen, K. G., Østerlund, M. T., Kiil, K., and Nielsen, E. M. (2021). Prediction of antimicrobial resistance in clinical *Campylobacter jejuni* isolates

- from whole-genome sequencing data. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 673–682. doi: 10.1007/s10096-020-04043-y
- Das, B., Verma, J., Kumar, P., Ghosh, A., and Ramamurthy, T. (2020). Antibiotic resistance in *Vibrio cholerae*: understanding the ecology of resistance genes and mechanisms. *Vaccine* 38, A83–A92. doi: 10.1016/j.vaccine.2019.06.031
- De Maio, N., Shaw, L. P., Hubbard, A., George, S., Sanderson, N. D., Swann, J., et al. (2019). Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microbiol. Genom.* 5. doi: 10.1099/mgen.0.000294
- De, R. (2021). Mobile genetic elements of *Vibrio cholerae* and the evolution of its antimicrobial resistance. *Front. Trop. Dis.* 7:691604. doi: 10.3389/ftd.2021.691604
- Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., and Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* 13, 601–612. doi: 10.1038/nrg3226
- Donatella, L., Eduardo, G., Saverio, C., Michel, D., and Didier, R. (2016). Cholera. *Microbiol. Spect.* 4, 4.4.06. doi: 10.1128/microbiolspec.PoH-0012-2015
- Doyle, R. M., Burgess, C., Williams, R., Gorton, R., Booth, H., Brown, J., et al. (2018). Direct whole-genome sequencing of sputum accurately identifies drug-resistant mycobacterium tuberculosis faster than MGIT culture sequencing. *J. Clin. Microbiol.* 56:e00666-18. doi: 10.1128/JCM.00666-18
- Eibach, D., Herrera-León, S., Gil, H., Hogan, B., Ehlkes, L., Adjabeng, M., et al. (2016). Molecular epidemiology and antibiotic susceptibility of *Vibrio cholerae* associated with a large cholera outbreak in Ghana in 2014. *PLoS Negl. Trop. Dis.* 10:e0004751. doi: 10.1371/journal.pntd.0004751
- Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., et al. (2021). AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* 11, 1–9. doi: 10.1038/s41598-021-91456-0
- Garneau-Tsodikova, S., and Labby, K. J. (2016). Mechanisms of resistance to aminoglycoside antibiotics: overview and perspectives. *Medchemcomm* 7, 11–27. doi: 10.1039/C5MD00344J
- Golparian, D., Doná, V., Sánchez-Busó, L., Foerster, S., Harris, S., Endimiani, A., et al. (2018). Antimicrobial resistance prediction and phylogenetic analysis of *Neisseria gonorrhoeae* isolates using the oxford nanopore minion sequencer. *Sci. Rep.* 8, 1–12. doi: 10.1038/s41598-018-35750-4
- Grüning, B., Dale, R., Sjödin, A., Chapman, B. A., Rowe, J., Tomkins-Tinch, C. H., et al. (2018). Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods* 15, 475–476. doi: 10.1038/s41592-018-0046-7
- Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., et al. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* 58, 212–220. doi: 10.1128/AAC.01310-13
- Hodges, L. M., Taboada, E. N., Koziol, A., Mutschall, S., Blais, B. W., Inglis, G. D., et al. (2021). Systematic evaluation of whole-genome sequencing based prediction of antimicrobial resistance in *Campylobacter jejuni* and *C. coli*. *Front. Microbiol.* 12, 776967–776967. doi: 10.3389/fmicb.2021.776967
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9, 5114. doi: 10.1038/s41467-018-07641-9
- Jaskólska, M., Adams, D. W., and Blokesch, M. (2022). Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature* 604, 323–329. doi: 10.1038/s41586-022-04546-y
- Kaunitz, J. D. (2020). Oral defense: how oral rehydration solutions revolutionized the treatment of toxigenic diarrhea. *Digest. Dis. Sci.* 65, 345–348. doi: 10.1007/s10620-019-06023-5
- Kitaoka, M., Miyata, S. T., Unterwiesing, D., and Pukatzki, S. (2011). Antibiotic resistance mechanisms of *Vibrio cholerae*. *J. Med. Microbiol.* 60, 397–407. doi: 10.1099/jmm.0.023051-0
- Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi: 10.1038/s41587-019-0072-8
- Lee, I., Ha, S.-M., Baek, M.-G., Kim, D. W., Yi, H., and Chun, J. (2021). Vicpred: A *Vibrio cholerae* genotype prediction tool. *Front. Microbiol.* 2021:691895. doi: 10.3389/fmicb.2021.691895
- Leger, A., and Leonardi, T. (2019). pycoQC, interactive quality control for oxford nanopore sequencing. *J. Open Source Softw.* 4:1236. doi: 10.21105/joss.01236
- Leggett, R. M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T. C., Kujawska, M., et al. (2020). Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nat. Microbiol.* 5, 430–442. doi: 10.1038/s41564-019-0626-z
- Lemon, J. K., Khil, P. P., Frank, K. M., and Dekker, J. P. (2017). Rapid nanopore sequencing of plasmids and resistance gene detection in clinical isolates. *J. Clin. Microbiol.* 55, 3530–3543. doi: 10.1128/JCM.01069-17
- Maestri, S., Cosentino, E., Paterno, M., Freitag, H., Garces, J. M., Marcolungo, L., et al. (2019). A rapid and accurate minion-based workflow for tracking species biodiversity in the field. *Genes* 10:468. doi: 10.3390/genes10060468
- Momba, M., and Azab El-Liethy, M. (2017). “*Vibrio cholerae* and Cholera biotypes,” in *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project)*, eds J. B. Rose and B. Jiménez-Cisneros; Part 3: Specific Excreted Pathogens: Environmental and Epidemiology Aspects - Section 2: Bacteria, eds A. Pruden, N. Ashbolt, and J. Miller (East Lansing, MI: Michigan State University; UNESCO). doi: 10.14321/waterpathogens.28
- Nalin, D. R., and Cash, R. A. (2018). 50 years of oral rehydration therapy: the solution is still simple. *Lancet* 392, 536–538. doi: 10.1016/S0140-6736(18)31488-0
- Nguyen, D. T., Ngo, T. C., Tran, H. H., Nguyen, T. P. L., Nguyen, B. M., Morita, K., and Ehara, M. (2009). Two different mechanisms of ampicillin resistance operating in strains of *Vibrio cholerae* o1 independent of resistance genes. *FEMS Microbiol. Lett.* 298, 37–43. doi: 10.1111/j.1574-6968.2009.01693.x
- Opintan, J. A., Will, R. C., Kuma, G. K., Osei, M., Akumwena, A., Boateng, G., et al. (2021). Phylogenetic and antimicrobial drug resistance analysis of *Vibrio cholerae* o1 isolates from Ghana. *Microb. Genomics* 7:668. doi: 10.1099/mgen.0.000668
- Oxford Nanopore Technologies (2018) *Medaka*. Available online at: <https://github.com/nanoporetech/medaka>
- Ramamurthy, T., Mutreja, A., Weill, F.-X., Das, B., Ghosh, A., and Nair, G. B. (2019). Revisiting the global epidemiology of cholera in conjunction with the genomics of *Vibrio cholerae*. *Front. Publ. Health* 7:203. doi: 10.3389/fpubh.2019.00203
- Seemann, T. (2016). Abricate, *Github*. Available online at: <https://github.com/tseemann/abricate>
- Shimada, T., Arakawa, E., Itoh, K., Okitsu, T., Matsushima, A., Asai, Y., et al. (1994). Extended serotyping scheme for *Vibrio cholerae*. *Curr. Microbiol.* 28, 175–178. doi: 10.1007/BF01571061
- Siriphap, A., Leekitcharoenphon, P., Kaas, R. S., Theethakaew, C., Aarestrup, F. M., Suthikul, O., et al. (2017). Characterization and genetic variation of *Vibrio cholerae* isolated from clinical and environmental sources in Thailand. *PLoS ONE* 12:e0169324. doi: 10.1371/journal.pone.0169324
- Tommaso, P. D., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* 35, 316–319. doi: 10.1038/nbt.3820
- van Belkum, A., Burnham, C.-A. D., Rossen, J. W., Mallard, F., Rochas, O., and Dunne, W. M. (2020). Innovative and rapid antimicrobial susceptibility testing systems. *Nat. Rev. Microbiol.* 18, 299–311. doi: 10.1038/s41579-020-0327-x
- Verma, J., Bag, S., Saha, B., Kumar, P., Ghosh, T. S., Dayal, M., et al. (2019). Genomic plasticity associated with antimicrobial resistance in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 6226–6231. doi: 10.1073/pnas.1900141116
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9:e112963. doi: 10.1371/journal.pone.0112963
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Wick, R. R., Judd, L. M., Wyres, K. L., and Holt, K. E. (2021). Recovery of small plasmid sequences via oxford nanopore sequencing. *Microb. Genomics* 7:000631. doi: 10.1099/mgen.0.000631
- World Health Organization (2014). *Antimicrobial Resistance: Global Report on Surveillance*. World Health Organization.
- World Health Organization (2021). *Global Antimicrobial Resistance and Use Surveillance System (Glass) Report: 2021*. World Health Organization.

Zhou, Y., Yu, L., Li, J., Zhang, L., Tong, Y., and Kan, B. (2013). Accumulation of mutations in DNA gyrase and topoisomerase iv genes contributes to fluoroquinolone resistance in *Vibrio cholerae* o139 strains. *Int. J. Antimicrob. Agents* 42, 72–75. doi: 10.1016/j.ijantimicag.2013.03.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Fuesslin, Krautwurst, Srivastava, Winter, Liedigk, Thye, Herrera-León, Wohl, May, Fobil, Eibach, Marz and Schuldt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



CARB-ES-19 Multicenter Study of Carbapenemase-Producing *Klebsiella pneumoniae* and *Escherichia coli* From All Spanish Provinces Reveals Interregional Spread of High-Risk Clones Such as ST307/OXA-48 and ST512/KPC-3

Javier E. Cañada-García^{1†}, Zaira Moure^{1†}, Pedro J. Sola-Campoy¹, Mercedes Delgado-Valverde^{2,3}, María E. Cano⁴, Desirée Gijón^{3,5}, Mónica González^{3,6}, Irene Gracia-Ahufinger^{3,7}, Nieves Larrosa^{3,8}, Xavier Mulet^{3,9}, Cristina Pitart¹⁰, Alba Rivera¹¹, Germán Bou^{3,6}, Jorge Calvo^{3,4}, Rafael Cantón^{3,5}, Juan José González-López^{3,8}, Luis Martínez-Martínez^{3,7,12}, Ferran Navarro¹¹, Antonio Oliver^{3,9}, Zaira R. Palacios-Baena^{2,3,13}, Álvaro Pascual^{2,3,13}, Guillermo Ruiz-Carrascoso¹⁴, Jordi Vila^{3,10}, Belén Aracil^{1,3}, María Pérez-Vázquez^{1,3†}, Jesús Oteo-Iglesias^{1,3*†} and the GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 Study Group[§]

OPEN ACCESS

Edited by:

Costas C. Papagiannitsis,
University of Thessaly, Greece

Reviewed by:

Marc Finianos,
Charles University, Czechia
Tamara Salloum,
Brigham and Women's Hospital and
Harvard Medical School,
United States

*Correspondence:

Jesús Oteo-Iglesias
jesus.oteo@isciii.es

[†]These authors have contributed
equally to this work

[‡]These authors have contributed
equally to this work and share senior
authorship

[§]Members of the Group are listed in
Acknowledgments

Specialty section:

This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 12 April 2022

Accepted: 23 May 2022

Published: 30 June 2022

¹ Laboratorio de Referencia e Investigación en Resistencia a Antibióticos e Infecciones Relacionadas con la Asistencia Sanitaria, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Madrid, Spain, ² Unidad de Enfermedades Infecciosas y Microbiología, Hospital Universitario Virgen Macarena, Instituto de Biomedicina de Sevilla (Hospital Universitario Virgen Macarena/CSIC/Universidad de Sevilla), Seville, Spain, ³ CIBER de Enfermedades Infecciosas (CIBERINFEC), REIPI, Instituto de Salud Carlos III, Madrid, Spain, ⁴ Servicio de Microbiología, Hospital Universitario Marqués de Valdecilla, IDIVAL, Santander, Spain, ⁵ Servicio de Microbiología, Hospital Universitario Ramón y Cajal, Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS), Madrid, Spain, ⁶ Servicio Microbiología, Hospital Universitario A Coruña, Instituto Investigación Biomédica A Coruña (INIBIC), A Coruña, Spain, ⁷ Microbiology Unit, Reina Sofía University Hospital, Maimonides Biomedical Research Institute of Córdoba (IMIBIC), Córdoba, Spain, ⁸ Departament de Genètica i Microbiologia, Servei de Microbiologia, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain, ⁹ Servicio de Microbiología, Hospital Universitario Son Espases, Instituto de investigación sanitaria Illes Balears (IdiSBa), Palma de Mallorca, Spain, ¹⁰ Servicio de Microbiología, Hospital Clínic de Barcelona, ISGlobal Barcelona Institute for Global Health, Barcelona, Spain, ¹¹ Microbiology Department, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona (UAB), Sant Pau Biomedical Research Institute (IIB Sant Pau), Barcelona, Spain, ¹² Department of Agricultural Chemistry, Soil Science and Microbiology, University of Córdoba, Córdoba, Spain, ¹³ Departamento de Microbiología, Universidad de Sevilla, Seville, Spain, ¹⁴ Servicio de Microbiología Clínica, Hospital Universitario La Paz (IdiPAZ), Madrid, Spain

Objectives: CARB-ES-19 is a comprehensive, multicenter, nationwide study integrating whole-genome sequencing (WGS) in the surveillance of carbapenemase-producing *K. pneumoniae* (CP-Kpn) and *E. coli* (CP-Eco) to determine their incidence, geographical distribution, phylogeny, and resistance mechanisms in Spain.

Methods: In total, 71 hospitals, representing all 50 Spanish provinces, collected the first 10 isolates per hospital (February to May 2019); CPE isolates were first identified according to EUCAST (meropenem MIC > 0.12 mg/L with immunochromatography, colorimetric tests, carbapenem inactivation, or carbapenem hydrolysis with MALDI-TOF). Prevalence and incidence were calculated according to population denominators.

Antibiotic susceptibility testing was performed using the microdilution method (EUCAST). All 403 isolates collected were sequenced for high-resolution single-nucleotide polymorphism (SNP) typing, core genome multilocus sequence typing (cgMLST), and resistome analysis.

Results: In total, 377 (93.5%) CP-Kpn and 26 (6.5%) CP-Eco isolates were collected from 62 (87.3%) hospitals in 46 (92%) provinces. CP-Kpn was more prevalent in the blood (5.8%, 50/853) than in the urine (1.4%, 201/14,464). The cumulative incidence for both CP-Kpn and CP-Eco was 0.05 per 100 admitted patients. The main carbapenemase genes identified in CP-Kpn were *bla*_{OXA-48} (263/377), *bla*_{KPC-3} (62/377), *bla*_{VIM-1} (28/377), and *bla*_{NDM-1} (12/377). All isolates were susceptible to at least two antibiotics. Interregional dissemination of eight high-risk CP-Kpn clones was detected, mainly ST307/OXA-48 (16.4%), ST11/OXA-48 (16.4%), and ST512-ST258/KPC (13.8%). ST512/KPC and ST15/OXA-48 were the most frequent bacteremia-causative clones. The average number of acquired resistance genes was higher in CP-Kpn (7.9) than in CP-Eco (5.5).

Conclusion: This study serves as a first step toward WGS integration in the surveillance of carbapenemase-producing Enterobacterales in Spain. We detected important epidemiological changes, including increased CP-Kpn and CP-Eco prevalence and incidence compared to previous studies, wide interregional dissemination, and increased dissemination of high-risk clones, such as ST307/OXA-48 and ST512/KPC-3.

Keywords: CARB-ES-19 study, carbapenemases, whole genome sequencing, *Klebsiella pneumoniae*, high-risk clones

INTRODUCTION

The rapid spread of carbapenemase-producing Enterobacterales (CPE) is a threat to individual and public health worldwide; infections caused by CPE significantly increase morbidity and mortality (Barrasa-Villar et al., 2017). The World Health Organization has included CPE as a critical priority issue (Tacconelli et al., 2018). The carbapenemases most frequently found in Enterobacterales are KPC, OXA-48, VIM, IMP, and NDM, although prevalence rates differ according to the geographical area considered (Grundmann et al., 2017).

Previous studies performed in Spain revealed the rapid evolution of CPEs from isolated cases in 2009 (Miró et al., 2013) to interregional dissemination in 2013 (Oteo et al., 2015). Recent European studies have indicated epidemiological changes in CPE infections (Haller et al., 2019; Ludden et al., 2020; Oteo-Iglesias et al., 2020; Di Pilato et al., 2021). According to the first structured survey on the occurrence of carbapenemase-producing *K. pneumoniae* and *E. coli* in European hospitals (Grundmann et al., 2017), Spain had the fourth highest incidence in Europe (0.04 cases per 100 patients) after Italy, Greece, and Montenegro.

However, recent multicenter studies with adequate geographical representation are scarce, since regions with low prevalence are often underrepresented.

Accurate data collection at the national level is required for the successful implementation of CPE control measures.

The CARB-ES-19 project utilized the national antibiotic-resistance surveillance framework and was designed to provide continuity with previous national, multicenter studies (Miró et al., 2013; Oteo et al., 2015). This large-scale, nationwide, structured survey integrated whole-genome sequencing (WGS) analyses of two CPE species with high clinical and epidemiological impacts, *Klebsiella pneumoniae* (CP-Kpn) and *Escherichia coli* (CP-Eco), to (i) determine the prevalence and incidence of these microorganisms, (ii) describe their inter-regional distribution and molecular epidemiology, and (iii) describe their resistance mechanisms and susceptibility profiles.

MATERIALS AND METHODS

Study Design and Isolates

CARB-ES-19 is a prospective, multicenter study designed to identify clinical cases associated with CP-Kpn and CP-Eco. In total, 71 hospitals, representing all 50 Spanish provinces, collected the first 10 non-duplicate consecutive isolates of carbapenem non-susceptible CP-Eco or CP-Kpn isolated from clinical samples from individual consecutive patients between February and May 2019. The geographical distribution of participating hospitals is available in the free and interactive online access tool Microreact.¹

Not all participating hospitals had 10 CP-Eco or CP-Kpn isolates during the study period. Isolates from rectal exudates for

¹<https://microreact.org/project/vXbq3eF5WXY5qZe9HBUJL-carb-es-2019>

the detection of EPC carriers were not included. All provinces (NUTS-3 regions in Spain) were represented by at least one hospital; seven of the provinces with the largest population were represented by more than one hospital (range 2–6). A similar study design was used by the European Centre for Disease Prevention and Control (ECDC) in the European survey of carbapenemase-producing Enterobacteriaceae (EuSCAPE; Grundmann et al., 2017).

Initial assays were performed at each participating hospital using standard microbiological methods. CPE isolates were identified according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) established meropenem cutoff value for CPE (meropenem MIC > 0.12 mg/L; European Society of Clinical Microbiology and Infectious Diseases [EUCAST], 2017). Confirmation of carbapenemase production was verified using at least one EUCAST-recommended method (European Society of Clinical Microbiology and Infectious Diseases [EUCAST], 2017), such as immunochromatography, biochemical (colorimetric) tests, carbapenem inactivation, or detection of carbapenem hydrolysis with MALDI-TOF.

In total, 10 hospitals and the Spanish National Centre of Microbiology (CNM) acted as node centers, performing molecular confirmation of standard carbapenemase genes using PCR (Supplementary Table 1). Confirmed CPE isolates were submitted to the Antibiotic Reference Laboratory of the CNM for WGS. Prevalence was estimated as the proportion of CP-Kpn and CP-Eco isolates relative to the total collected clinical *K. pneumoniae* and *E. coli* isolates, respectively. Overall cumulative incidence and incidence density estimates are reported as the number of admitted patients diagnosed with either CP-Kpn or CP-Eco per 100 admitted patients and per 1,000 patient-days, respectively. The denominators for the incidence and prevalence estimates were adjusted to the date on which the last isolate included in the study was collected in each hospital.

Drug Susceptibility Testing

Antibiotic susceptibility testing was performed using the broth microdilution susceptibility method (DKYMGN Sensititre™ panels, Thermo Fisher Scientific, United States) (International Organization for Standardization [ISO], 2006). Antibiotic gradient strips were used to study susceptibility to meropenem/vaborbactam and cefepime (bioMérieux, Marcy-l'Étoile, France) and to imipenem/relebactam, plazomicin, and cefiderocol (Liofilchem, Roseto degli Abruzzi, Italy) in Mueller Hinton agar (bioMérieux, Marcy-l'Étoile, France). EUCAST v12.0 clinical breakpoints and guidelines for Enterobacterales were used for interpretation. An FDA-approved susceptibility breakpoint of ≤ 2 mg/L was used for plazomicin.

Genomic Library Preparation and Sequence Analysis

Genomic DNA paired-end libraries were generated using the Nextera XT DNA Sample Preparation Kit (Illumina Inc., San Diego, CA, United States). These libraries were sequenced using the Illumina HiSeq 500 next-generation sequencer with

2 × 150 bp paired-end reads (Illumina Inc.) Raw sequence data were submitted to the European Nucleotide Archive (PRJEB50822). The quality of short reads was assessed using FASTQC, and assembly into contigs was performed with Unicycler 0.4.8 (Wick et al., 2017). The quality of the assembly was assessed with QUAST.² Prokka v1.14-beta (Seemann, 2014) was used for automatic *de novo* assembly annotation.

Phylogenetic Analysis

Assembly contigs were used as input for Roary version 3.13.0 (Page et al., 2015). An alignment of 2,645 core genes (present in > 99% of isolates), comprising 2,415,034 bp, was generated for *K. pneumoniae*. Variable positions were extracted (85,696 single-nucleotide polymorphisms [SNPs]), and a maximum-likelihood phylogenetic tree of SNPs was constructed using RAXML version 7.0.4 (Stamatakis, 2006) with a general time-reversible model and gamma correction for site variation. The phylogenetic tree and associated metadata were visualized using Microreact and iTOL.³

Sequence types (STs) were calculated according to multilocus sequence typing (MLST) schemes of the Institut Pasteur and the University of Warwick for *K. pneumoniae* and *E. coli*, respectively, using Ariba version 2.6.2 (Hunt et al., 2017). A simple diversity index (SDI; Gastmeier et al., 2006) was applied to analyze population diversity. Core genome MLST (cgMLST) was performed, consisting of 2,538 *K. pneumoniae* targets provided by SeqSphere+ version 3.5.0 (Ridom, Münster, Germany).

Analysis of Antimicrobial Resistance, Virulence Genes, and Capsular Locus

Antibiotic resistance genes were analyzed by SRST2 (Inouye et al., 2014) using the ARG-ANNOT database (Gupta et al., 2014) and ResFinder (CGE server⁴), with ID thresholds of 100% for β -lactamase variants and 98% for other genes. The K-locus and virulence genes were characterized using Kleborate.⁵ The presence of *ybt*, *clb*, and *iuc* was used to assign a virulence score, as described by Lam et al. (2021).

Characterization of Plasmids Carrying Carbapenemase Genes

The plasmids carrying the carbapenemase genes in six representative *K. pneumoniae* isolates (ST307/KPC-3, ST512/KPC-3, ST512/KPC-23, ST147/NDM-1, ST307/OXA-48, and ST11/VIM-1) were reconstructed by the in-house script (PlasmidID⁶). PlasmidID was used to (i) map reads over a plasmid curated database to find those with the higher coverage and *de novo* assembly of these reads and (ii) make local alignments to localize resistance and replicative genes (Pérez-Vázquez et al., 2019a).

²<http://quast.sourceforge.net/>

³<https://itol.embl.de/>

⁴<http://www.genomicepidemiology.org/>

⁵<https://github.com/katholt/Kleborate>

⁶<https://github.com/BU-ISCIH/plasmidID>

Ethics Statement

This study was authorized by the Spanish Agency for Medicines and Health Products (code JOI-AVI-2019-01). The Research Ethics Committee of the University Hospitals Virgen Macarena and Virgen del Rocio (Sevilla, Spain) approved this study.

RESULTS

Bacterial Isolates, Prevalence, Incidence, and Carbapenemase Types

Of 403 CPE cases, 377 (93.5%) were identified as CP-Kpn and 26 (6.5%) as CP-Eco. Patients were mainly men (211, 52.4%) older than 65 years (280, 69.5%). CPE isolates were collected from urine (215, 53.3%), wounds and abscesses (61, 15.1%), blood (52, 12.9%), respiratory samples (47, 11.7%), and other locations (28, 6.9%).

At least one case was identified in 62 (87.3%) of participating hospitals and in 46 (92%) of 50 Spanish provinces. Participating hospitals isolated a total of 15,100 and 70,760 isolates of *K. pneumoniae* and *E. coli*, respectively. The average CP-Kpn prevalence was 2.5% (377/15100; interprovincial range: 0–17.3%), with 16 (22.5%) hospitals reporting prevalences greater than 5%. The average prevalence of CP-Eco was 0.04% (26/70,760; interprovincial range: 0–0.5%; **Supplementary Table 2**). The prevalence distribution by province is detailed in **Figure 1** and **Supplementary Table 2**. CP isolates were more prevalent in blood (CP-Kpn: 5.8%, 50/853; CP-Eco: 0.06%, 2/3,353) than in urine (CP-Kpn: 1.4%, 201/14,464; CP-Eco: 0.02%, 14/56,848).

Overall, cumulative incidence and incidence density estimates were 0.05 per 100 admitted patients (interprovincial range: 0–0.34) and 0.08 per 1,000 patient-days (interprovincial range: 0–0.58), respectively (**Supplementary Table 2**).

The carbapenemase genes detected in CP-Kpn were *bla*_{OXA-48} (263 isolates, 69.8%), *bla*_{KPC-3} (62, 16.4%), *bla*_{VIM-1} (28, 7.4%), *bla*_{NDM-1} (12, 3.2%), *bla*_{KPC-23} (7, 1.9%), *bla*_{KPC-2} (5, 1.3%), *bla*_{OXA-245} (3, 0.8%), with *bla*_{IMP-8}, *bla*_{GES-2}, *bla*_{NDM-3}, *bla*_{NDM-23}, and *bla*_{OXA-505} in one isolate each (0.3%). In CP-Eco, *bla*_{OXA-48} (19, 73.1%) was predominant, followed by *bla*_{VIM-1} (5, 19.2%), *bla*_{KPC-3} (2, 7.7%), and *bla*_{NDM-5} (1, 3.8%). Eight CP-Kpn (two OXA-48 + VIM-1, two OXA-48 + KPC, and one each of OXA-48 + NDM-1, VIM-1 + NDM-1, NDM-1 + GES-2, and VIM-1 + NDM-3) and one CP-Eco (OXA-48 + VIM-1) isolates co-produced two carbapenemase types.

The Canary Islands and northern Spain had more OXA-48-producing CP-Kpn, whereas southern Spain had more KPC-producing CP-Kpn (**Figure 1**, **Supplementary Table 2**).

Antibiotic Susceptibility Testing

The antibiotic susceptibilities of CP-Kpn isolates are listed in **Tables 1** and **2**. The overall rates of susceptibility to carbapenems were 54.6% for imipenem, 52.3% for meropenem, and 3.4% for ertapenem (**Table 1A**), with all isolates non-susceptible to at least one carbapenem.

The most active antibiotics *in vitro* were cefiderocol (93.9% susceptibility), plazomicin (93.4%), colistin (90.5%), meropenem/vaborbactam (89.4%), ceftazidime/avibactam (84.1%), and imipenem/relebactam (78%; **Table 1A**). However, these numbers varied significantly depending on the carbapenemase type (**Table 1B**).

In general, CP-Eco isolates were more susceptible to antibiotics (**Supplementary Table 3**) than CP-Kpn isolates, with the greatest differences observed for tigecycline (84.6% in CP-Eco vs. 29.2% in CP-Kpn).

High-Risk Clones of Carbapenemase-Producing *Klebsiella pneumoniae*

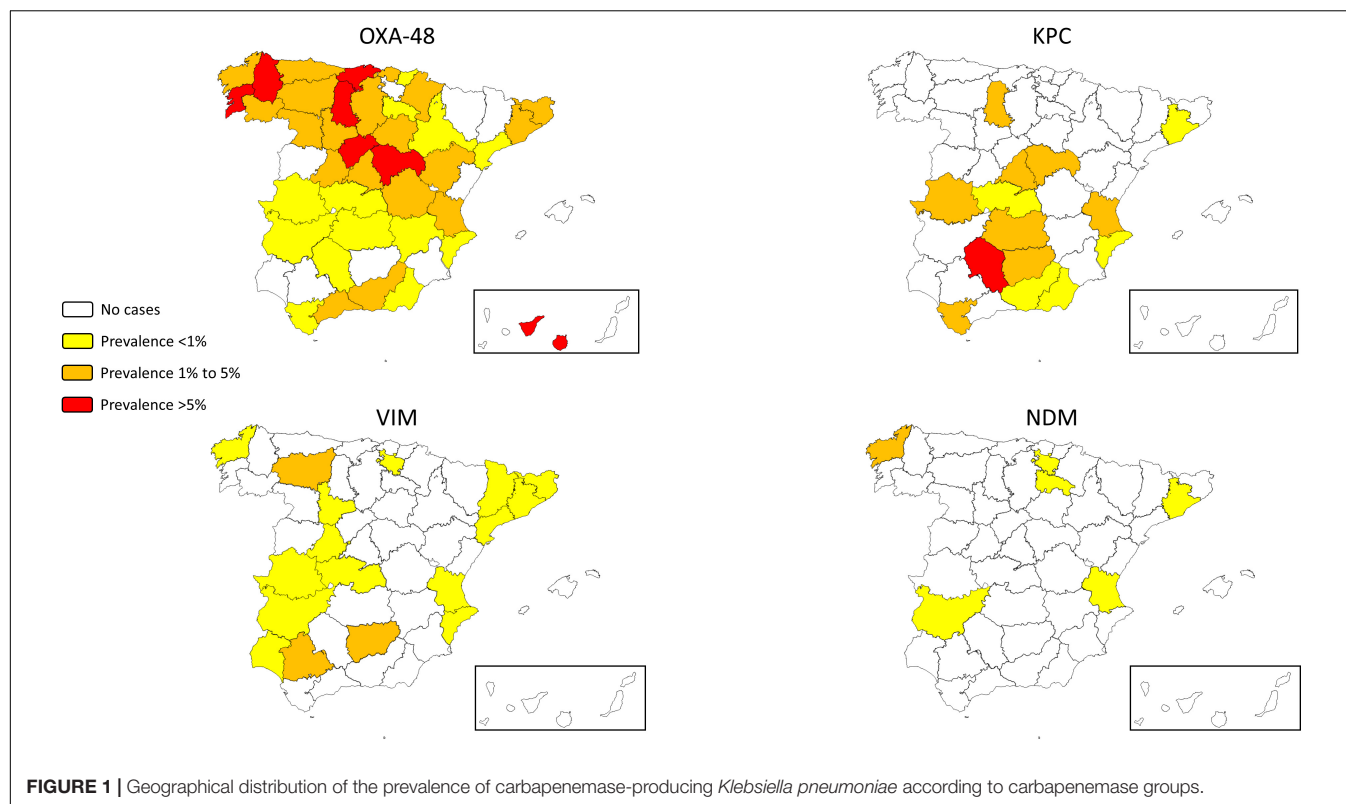
CP-Kpn isolates were grouped into 48 STs (SDI: 12.7; mean: 7.8 isolates per ST; range: 1–82). A new ST was detected in an OXA-48-producing *K. pneumoniae* strain. The most prevalent STs (≥ 10 isolates) were ST307 (82, 21.7%), ST11 (68, 18%), ST258/512 complex (4/48, 13.8%), ST15 (48, 12.7%), ST147 (36, 9.5%), and ST392 (15, 4%), accounting for 79.7% of all CP-Kpn isolates. ST307 and ST11 expressed four different carbapenemase types (OXA-48, KPC, VIM-1, and NDM-1), but ST258/512 only expressed KPC. Among CP-Eco isolates, 21 different STs were identified (SDI: 80.8; mean: 1.2 isolates per ST; range: 1–4), with only ST131 (4 isolates, 15.4%) expressing more than two carbapenemases.

Considering ST/carbapenemase combinations, eight high-risk CP-Kpn clones (≥ 7 isolates) were detected, with ST307/OXA-48 (16.4%), ST11/OXA-48 (16.4%), ST512-ST258/KPC (13.8%), and ST15/OXA-48-like (11.4%) as the most common combinations (**Table 2**). These high-risk clones were detected in at least four hospitals and two autonomous communities, suggesting interregional dissemination (**Table 2**). ST512-258/KPC and ST15/OXA-48-like were the most frequent bacteremia-producing clones, responsible for 24 and 16% of CP-Kpn-induced bacteremia, respectively (**Table 2**).

Phylogenetic Analysis of Carbapenemase-Producing *Klebsiella pneumoniae*

A total of 92,608 high-quality SNPs, identified by referencing the sequence of *K. pneumoniae* strain NTUH-K2044, were used to construct a maximum-likelihood phylogenetic tree (**Figure 2**). The median pairwise distance between isolates was 9,678 SNPs (range: 0–13,254 SNPs). CP-Kpn isolates were grouped into six major clusters (clusters 1–6; **Figure 2**). The main characteristics of these six clusters are detailed in **Supplementary Table 4**. The average difference between isolates from different clusters was 10,580 SNPs.

Genome assemblies of all CP-Kpn isolates, analyzed using the gene-by-gene approach and allelic distance from cgMLST, are reflected in a minimum-spanning tree (**Figure 3**). The average allelic distances in pairwise comparisons of isolates were 45 alleles in cluster 1 (range: 0–235 alleles), 68 alleles in cluster 2 (range: 0–49), 64 alleles in cluster 3 (range: 0–139), 21 alleles in cluster 4



(range: 0–43), 30 alleles in cluster 5 (range: 0–106), and 39 alleles in cluster 6 (range: 0–64).

Applying a threshold of both 5 alleles and 10 SNPs, 12 groups with more than four closely related isolates were detected, with ST512 (4), ST11 (4), and ST307 (3) commonly identified (Figure 3).

Identification and Distribution of Resistance and Virulence Genes

An average of 7.9 acquired resistance genes (ARGs) were detected in CP-Kpn isolates (range: 1–20 ARGs), and 5.5 ARGs were detected in CP-Eco isolates (range: 1–18 ARGs). The most frequent extended-spectrum β -lactamase (ESBL) gene was *bla*_{CTX-M-15} (247 overall, 61.3%), detected in 240 CP-Kpn (63.7%) and 7 CP-Eco (26.9%). Other ESBL genes identified were *bla*_{CTX-M-9} (7, 1.7%), *bla*_{SHV-12} (4, 1%), *bla*_{CTX-M-3} (2, 0.5%), *bla*_{CTX-M-88} (2, 0.5%), and *bla*_{CTX-M-65} (1, 0.25%).

The predominant aminoglycoside resistance genes encoded N-acetyltransferases *aac*(6′)-*Ib-cr* (189, 46.9%), *aac*(3)-*Ila* (141, 35%), and *aac*(6′)-*Ib* (70, 17.4%). Acquired 16S rRNA methyltransferase *rmtF*1 was detected in 19 CP-Kpn (5%) isolates, all belonging to ST147 and encoding *bla*_{OXA-48} (14) or *bla*_{NDM-1} (5).

ARGs encoding resistance to chloramphenicol, sulfonamides, trimethoprim, and tetracyclines were detected in 261 [64.7%; mainly *catB3* (46.2%)], 280 [69.5%; *sul1* (36.2%) and *sul2* (40.9%)], 252 [62%; mainly *dfrA14* (42.7%)], and 106 [26.3%, mainly *tetA* (23.1%)] isolates, respectively. Plasmid-mediated

quinolone resistance *qnr*-like determinants were detected in 202 (50.1%) isolates, with *qnrB1*-like (37.7%) as the most frequent.

Seven isolates contained colistin ARGs; *mcr-9* in four CP-Kpn and two CP-Eco and *mcr-1* in one CP-Eco. However, all six isolates harboring *mcr-9* were susceptible to colistin, as previously described (Macetic et al., 2021). There were 36 colistin-resistant CP-Kpn isolates, all of them showing amino acid substitution in both proteins of the PhoQ/PhoP regulatory system: D1509G and H406Y in PhoQ protein and R114A and L26Q in PhoP (Elias et al., 2021). Additional MgrB mutations were detected in four isolates: stop codon in L4 (3 ST512/KPC-3 isolates) and W47C (1 ST307/OXA-48 isolates).

The most prevalent ARGs associated with high-risk CP-Kpn clones are detailed in Table 2. *bla*_{SHV-28} allele was detected in 96.3% of ST307 but only in 21.1% of other clones. Additionally, the association of the *rmtF* and *arr* genes with ST147 was detected in 52.8 and 66.7%, respectively, compared with 0 and 6.2% in other clones.

CP-Kpn isolates belonged to 46 capsular polysaccharide loci (K-loci); 24 K-loci contained more than one isolate, and six included 78.3% of isolates, namely, KL102 (80), KL24 (75), KL107 (56), KL64 (39), KL112 (29), and KL27 (17). High correlations between K-loci and STs were observed: all KL102 isolates were ST307, all KL112 isolates were ST15, 92% of KL24 isolates were ST11, 87% of KL64 isolates were ST147, and 94.5% of KL107 isolates were ST512-258.

The yersiniabactin-encoding locus (*ybt*) was detected in 43% of CP-Kpn isolates, particularly ST11 (40.2%), ST147 (15.9%), and ST15 (14.6%). Nine different *ybt* lineages

were identified, with *ybt10*, associated with ICEKp4 and ST11, being the most frequent (58.5%; **Supplementary Figure 1**). Colibactin (*clb*) and aerobactin (*iuc*) loci were

detected in two isolates each (**Supplementary Figure 1**). One ST147/NDM-1 isolate contained the *rmrA2* gene associated with a hypermuroid/hypervirulent phenotype.

TABLE 1 | Antibiotic susceptibility of 377 carbapenemase-producing *Klebsiella pneumoniae* isolates as determined by the microdilution method and antibiotic gradient strips (antibiotics with *) according to EUCAST clinical breakpoints: (A) General results. (B) Results according to carbapenemase groups.

A					
Antibiotics	S (%)	R (%)	MIC₅₀^a	MIC₉₀^a	Range^a
Cefiderocol*	93.9	6.1	0.12	1	≤0.015–16
Plazomicin*	93.4	6.1	1	2	0.25 to >256
Colistin	90.5	9.5	1	2	0.5 to >8
Meropenem/vaborbactam*	89.4	10.6	0.5	16	≤0.015 to >128
Ceftazidime/avibactam	84.1	15.9	2	>16	≤0.5 to >16
Imipenem/relebactam*	77.4	22	1	8	0.12 to >32
Amikacin	72.9	27.1	8	>32	≤4 to >32
Imipenem	54.6	36.9	2	>16	≤0.5 to >16
Meropenem	52.3	28.4	2	>16	0.25 to >16
Gentamicin	45.6	54.4	4	>8	≤0.5 to >8
Trimethoprim/sulfamethoxazole	26.8	71.1	>8	>8	≤1 to >8
Tobramycin	25.7	74.3	>8	>8	≤1 to >8
Aztreonam	14.9	83.3	>32	>32	≤0.5 to >32
Cefepime*	11.9	84.1	32	>256	0.12 to >256
Ceftazidime	10.6	85.9	>16	>16	≤0.5 to >16
Ceftolozane/tazobactam	9.8	90.2	>32	>32	≤0.5 to >32
Cefotaxime	6.1	90.5	>8	>8	≤0.5 to >8
Ciprofloxacin	5.6	94.2	>2	>2	≤0.06 to >2
Ertapenem	3.4	96.6	>2	>2	0.25 to >2
B					
Antibiotic	Susceptibility (%)				
	OXA-48-group-producing isolates (n = 262)	KPC-group-producing isolates (n = 72)	MBL-group-producing isolates (n = 37)		
Cefiderocol*	95.8	86.1	94.6		
Plazomicin*	93.5	98.6	86.5		
Colistin	92.4	81.9	91.9		
Meropenem/vaborbactam*	89.3	100	73		
Ceftazidime/avibactam	95.4	90.3	0		
Imipenem/relebactam*	75.2	100	56.7		
Amikacin	86.6	33.3	56.8		
Imipenem	68.3	13.9	43.2		
Meropenem	64.9	13.9	45.9		
Gentamicin	47.3	48.6	32.4		
Trimethoprim/sulfamethoxazole	34.7	6.9	13.5		
Tobramycin	32.8	9.7	8.1		
Aztreonam	17.6	0	27		
Cefepime*	16.8	0	0		
Ceftazidime	15.3	0	0		
Ceftolozane/tazobactam	14.1	0	0		
Cefotaxime	8.4	0	0		
Ciprofloxacin	7.6	1.4	0		
Ertapenem	2.3	0	16.2		

S, susceptible, standard dosing regimen; R, resistant; MIC, minimum inhibitory concentrations.

^aExpressed in mg/L.

Isolates with two carbapenemases of different groups are excluded.

TABLE 2 | High-risk carbapenemase-producing *Klebsiella pneumoniae* clones defined according to combinations of sequence type/carbapenemase.

High-risk clones (n)	Hospitals (n)	Geographical distribution*	Evolution trends (%) 2013→2019**	Representation in bacteremia (%)	Carbapenemase genes (%)	Other prevalent acquired resistance genes (%)
ST307/OXA-48 (62)	23	Andalucía, Castilla La Mancha, Castilla y León, Extremadura, Cataluña, Galicia, Canarias, País Vasco, Comunidad Valenciana, Madrid	1.4→16.4	14	<i>bla</i> _{OXA-48} (100)	<i>bla</i> _{SHV-28} (95.1), <i>bla</i> _{CTX-M-15} (88.5), <i>aac</i> (6')-Ib-cr (78.7), <i>qnrB1</i> (78.7), <i>ant</i> (3')-Ia (75.4), <i>sul2</i> (73.8), <i>bla</i> _{OXA-1} (73.8), <i>aph</i> (3'')-Ib (73.8), <i>aph</i> (6)-Id (73.8), <i>catB3</i> (73.8%), <i>dfrA14</i> (73.8), <i>aac</i> (3)-IIa (72.1), <i>bla</i> _{TEM-1b} (70.5).
ST11/OXA-48 (62)	22	Andalucía, Castilla La Mancha, Castilla y León, Cantabria, Cataluña, La Rioja, Canarias, País Vasco, Madrid	24.1→16.4	12	<i>bla</i> _{OXA-48} (100),	<i>bla</i> _{CTX-M-15} (96.7), <i>bla</i> _{SHV-182} (95.1), <i>bla</i> _{OXA-1} (60.7), <i>aac</i> (6')-Ib-cr (63.9), <i>catB3</i> (63.9), <i>aac</i> (3)-IIa (55.7), <i>qnrB1</i> -like (57.4).
ST512-ST258/KPC (52)	10	Andalucía, Castilla La Mancha, Comunidad Valenciana, Cataluña, Madrid	0→13.8	24	<i>bla</i> _{KPC-3} (86.5) <i>bla</i> _{KPC-23} (13.5)	<i>bla</i> _{SHV-182} (97.9), <i>aph</i> (3'')-Ia (95.8), <i>bla</i> _{OXA-9} (95.8), <i>aac</i> (6')-Ib (93.8), <i>ant</i> (3')-Ia (93.8), <i>sul1</i> (93.8), <i>dfrA12</i> (93.3), <i>catA1</i> (91.7), <i>dfrA12</i> (91.7), <i>mphA</i> (91.7), <i>bla</i> _{TEM-187} (64.6).
ST15/OXA-48 like (43)	16	Andalucía, Castilla La Mancha, Madrid, Aragón, Cataluña, Canarias, País Vasco, Galicia	8.2→11.4	16	<i>bla</i> _{OXA-48} (93) <i>bla</i> _{OXA-245} (7)	<i>bla</i> _{SHV-28} (100), <i>ant</i> (3')-Ia (83.7), <i>bla</i> _{OXA-1} (74.4%), <i>bla</i> _{CTX-M-15} (72.1), <i>aac</i> (6')-Ib-cr (76.7), <i>catB3</i> (74.4), <i>dfrA14</i> (74.4), <i>aph</i> (3'')-Ib (60.5), <i>aph</i> (6)-Id (58.1), <i>sul2</i> (58.1), <i>bla</i> _{TEM-1b} (53.5).
ST147/OXA-48 (22)	10	Andalucía, Cataluña, Galicia, País Vasco, Asturias, Navarra	2.5→5.8	0	<i>bla</i> _{OXA-48} (95.5) <i>bla</i> _{OXA-505} (4.5)	<i>ant</i> (3')-Ia (90.9), <i>bla</i> _{CTX-M-15} (86.4), <i>bla</i> _{SHV-67} (86.4), <i>aac</i> (6')-Ib (72.7), <i>AAR-2</i> (72.7), <i>rmtF</i> (63.6), <i>qnrB1</i> (47.6), <i>dfrA14</i> (68.2), <i>mphA</i> (63.6), <i>dfrA12</i> (59.1), <i>sul1</i> (59.1).
ST307/KPC-3 (15)	4	Madrid, Extremadura, Castilla La Mancha	0→4	4	<i>bla</i> _{KPC-3} (100)	<i>bla</i> _{CTX-M-15} (100), <i>bla</i> _{OXA-1} (100), <i>bla</i> _{OXA-9} (100), <i>aac</i> (6')-Ib-cr (100), <i>aac</i> (3)-IIa (100), <i>ant</i> (3')-Ia (100), <i>aph</i> (3'')-Ib (100), <i>aph</i> (6)-Id (100), <i>catB3</i> (100), <i>dfrA14</i> (100), <i>bla</i> _{SHV-67} (100), <i>sul2</i> (100), <i>qnrB1</i> -like (71.4).
ST392/OXA-48 (14)	11	Andalucía, Castilla La Mancha, Madrid, Cataluña, Comunidad Valenciana, Galicia	0→3.7	4	<i>bla</i> _{OXA-48} (100).	<i>bla</i> _{SHV-67} (100), <i>aph</i> (3'')-Ib (92.9), <i>aph</i> (6)-Ib (92.9), <i>sul2</i> (86.7), <i>bla</i> _{TEM1-b} (85.7), <i>bla</i> _{CTX-M-15} (78.6), <i>dfrA14</i> (50), <i>qnrB1</i> (40), <i>aac</i> (6')-Ib-cr (57.1), <i>bla</i> _{OXA-1} (50), <i>catB3</i> (50).
ST147/NDM-1 (7)	4	Cataluña, Galicia	0→1.9	2	<i>bla</i> _{NDM-1} (100)	<i>aadA1</i> (100), <i>ARR</i> (100), <i>bla</i> _{CTX-M-1} -group (85.7), <i>bla</i> _{SHV-67} (100), <i>sul1</i> (100), <i>aac</i> (6')-Ib (71.4), <i>dfrA12</i> (71.4), <i>dfrA14</i> (71.4), <i>qnrB1</i> (71.4), <i>rmtF1</i> (71.4), <i>aph</i> (3'')-Ib (57.1), <i>aph</i> (6)-Id (57.1).

*Autonomous Communities in which the high-risk clones were detected.

**Evolution trends detected based on the results of a previous Spanish study conducted in 2013 (Oteo et al., 2015) relative to the study in 2019.

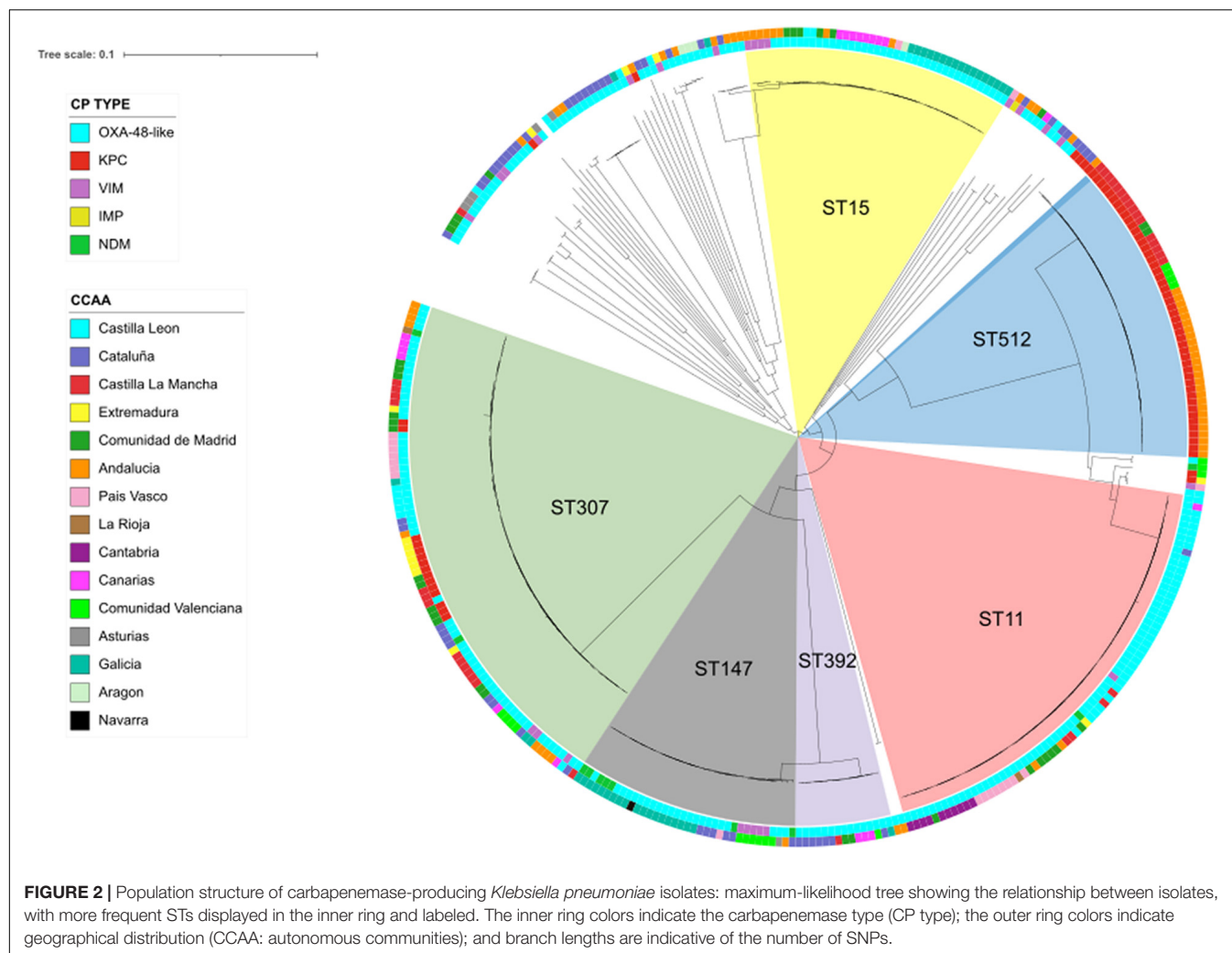
Characterization of Plasmids Harboring Carbapenemases Genes

The plasmidID mapping tool reconstructed three IncFIB plasmids carrying *bla*_{KPC} genes, all of them highly similar to NZ_CP027056 of ~90,000 bp (average identity > 95 and 99.9% coverage). In these three plasmids, *bla*_{KPC} genes were carried in the transposon Tn4401b, which was modified by the insertion of Tn5403, similar to that previously described (Rada et al., 2020).

In the ST147/NDM-1 isolate, *bla*_{NDM-1} was carried by an IncFIB plasmid of ~105,000 bp (highly similar to NZ_CM008884, average identity > 95% and 99.82% coverage). The largest genetic environment constructed for this gene

was 5,943 bp, with the sequence *groL-groS-nagA-trpF-ble-bla*_{NDM-1}-(trun)ISAb, which is similar to the previously described pNDM-11_IncFIB_KPN_Spain (Pérez-Vázquez et al., 2019a).

The plasmids detected carrying *bla*_{OXA-48} (ST307/OXA-48 isolate) and *bla*_{VIM-1} (ST11/VIM-1 isolate) were IncL. The first one was highly similar to NZ_CP023251 (~63,000 bp, average identity > 95% and 100% coverage) and the other to NZ_CP023419 (~70,000 bp, average identity > 95% and 100% coverage). The *bla*_{OXA-48} gene was located in a Tn1999 in which *lysR* and *bla*_{OXA-48} were flanked by two copies of *IS1999* (Mairi et al., 2018). The *bla*_{VIM-1} gene was located in the class

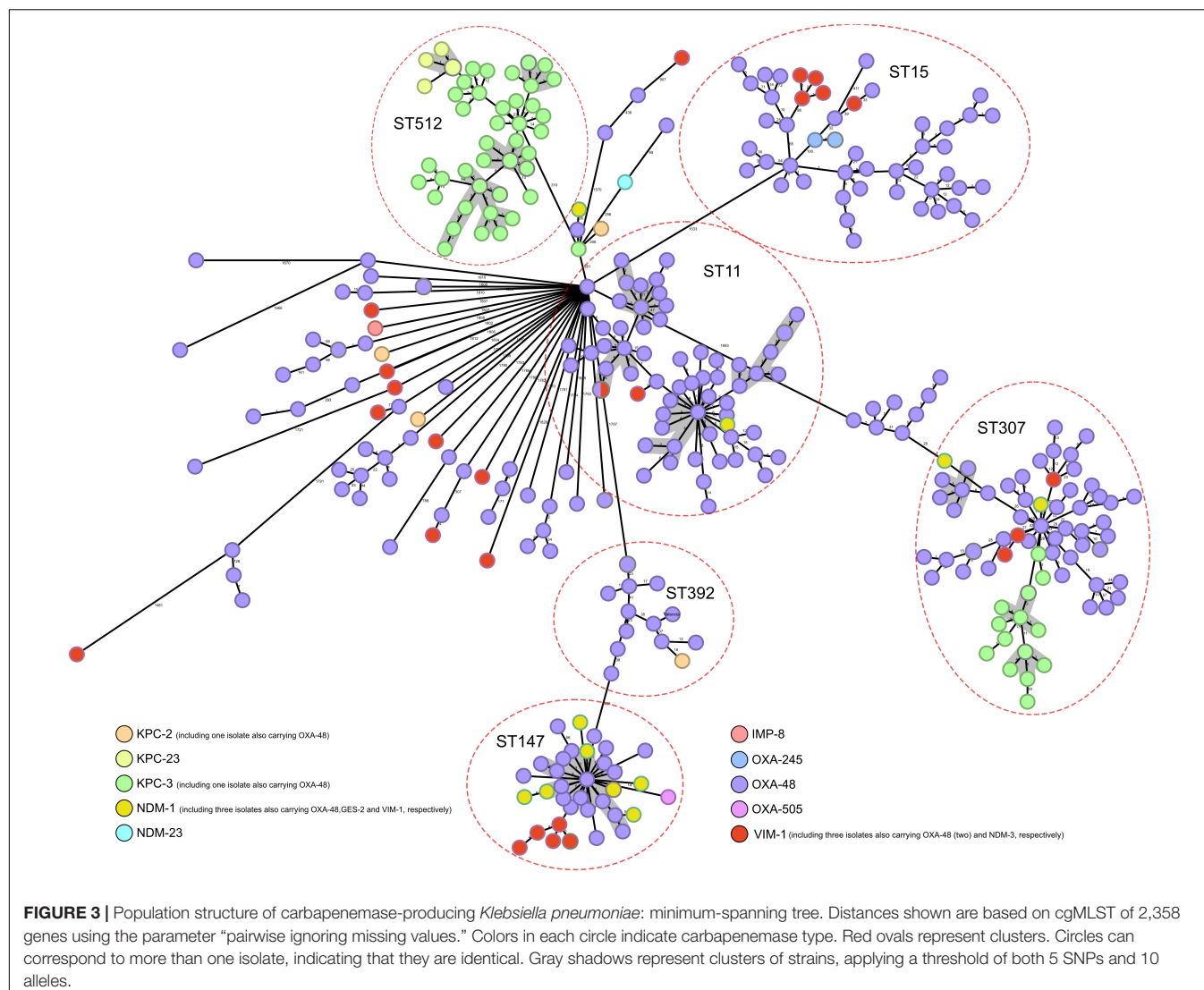


1 integron In624, showing the sequence *Int1-blav_{VIM-1}-aacA4-dfrB1-aadA1-catB2-qacEdelta1-sul1* (Villa et al., 2014).

DISCUSSION

We conducted a comprehensive analysis using WGS in CP-Kpn and CP-Eco isolates in Spain, revealing an overall cumulative incidence of 0.05 per 100 admitted patients and the wide interregional dissemination of high-risk clones. The study's strengths include representation of all Spanish provinces, application of population denominators for prevalence and incidence calculations, and the possibility of establishing evolutionary trends compared with previous studies using a similar methodology conducted by the same research team (Miró et al., 2013; Oteo et al., 2015; Grundmann et al., 2017). Although the use of only 10 strains per hospital might be a limitation, this approach aligns with ECDC strategies (Grundmann et al., 2017), avoiding the overrepresentation of geographical regions/clones observed in other studies.

The overall incidence of CP-Kpn and CP-Eco increased by 25%, from 0.04 cases per 100 patients in 2014 (Grundmann et al., 2017), with 13 provinces reporting incidences at least 2.5-fold higher in this study than the general one found in 2014. Compared with the 2.5% CP-Kpn prevalence reported here, the prevalences detected in 2009 and 2013 in Spain were significantly lower, at 0.2% (Miró et al., 2013) and 1.7% (Oteo et al., 2015), respectively. Increased prevalence was accompanied by a wide geographical spread, with strains detected in 92% of the 50 Spanish provinces and the presence of seven “high-risk clones” in at least three provinces. In a recent Italian study (Di Pilato et al., 2021), 80% of 30 participating hospitals reported CP-Kpn cases. The large difference in CP-Kpn prevalences between provinces detected in this study (four provinces had prevalences > 10%, **Supplementary Table 2**) highlights the importance of designing a study with representation from all geographical regions and with a design that minimizes the possible overrepresentation of a specific clone or a region due to the existence of an outbreak. In this sense, our study provides a realistic and representative view of the Spanish situation, minimizing possible biases.



Our results confirm previous reports of the preponderance of OXA-48 carbapenemase in Spain (detected in 82% of Spanish provinces; Oteo et al., 2015; Grundmann et al., 2017; Vázquez-Ucha et al., 2021). The *bla*_{OXA-505} gene (OXA-48 family, GenBank number: NG_049783) was detected in a carbapenem non-susceptible isolate with a positive colorimetric lacking other genes encoding carbapenemases, so it was included in the study despite the absence of published evidence on its carbapenemase activity.

New to this study was the observation of significant epidemiological evolution among CP-Kpn strains, from the predominance of the OXA-48-producing ST11 and ST405 (Oteo et al., 2015) to the emergence of ST307/OXA-48 and the dispersion of ST512/KPC (detected in seven provinces in Southern Spain). Compared with a multicenter study conducted in 2013 (Oteo et al., 2015), our study shows a significant increase in all high-risk clones, except ST11/OXA-48, which decreased in frequency from 24.1 to 16.4%, and ST15, which remained stable (Table 2). The frequency of ST307 increased from 1.4%

in 2013 (Oteo et al., 2015) to 16.4%, making it the predominant clone with a wide geographical distribution. Other recent studies also have identified that ST307 is an emerging clone worldwide (Wyres et al., 2019; Di Pilato et al., 2021), and the hyperepidemic clonal complex ST258/ST512 is widely predominant in Italy and Greece (David et al., 2019; Di Pilato et al., 2021). This epidemiological shift is associated with the increased CP-Kpn population diversity, from an SDI of 10.6 in 2013 (Oteo et al., 2015) to 12.7 in this study. Regarding the NDM-producers, the main high-risk *K. pneumoniae* clone ST147/NDM-1 detected in this study (Table 2) was the cause of one of the great NDM-1-producing *K. pneumoniae* outbreaks previously described in Spain (Pérez-Vázquez et al., 2019b).

CPE bacteremia is associated with high mortality (Tamma et al., 2017), mainly associated with delays in adequate treatment (Gutiérrez-Gutiérrez et al., 2017). This study showed a significantly higher prevalence of CP-Kpn in bacteremia than in the total infections considered altogether (more than double) or urinary tract infections (more than fourfold), with

ST512/KPC and ST15/OXA-48 causing 40% of CP-Kpn-induced bacteremia. Accurate and timely diagnosis could be critical in providing effective care.

Regarding the level of resistance to carbapenem antibiotics as the main target of carbapenemases, it should be noted that the susceptibility profile of meropenem and/or imipenem susceptibility with ertapenem resistance was frequently detected in this study. This profile was mainly due to the high prevalence of OXA-48 isolates, although it can also be found in VIM-1 producers (Oteo et al., 2015; Vázquez-Ucha et al., 2021).

Vaborbactam and relebactam do not inhibit metallo- β -lactamases; however, differences in susceptibility were observed between imipenem/imipenem-relebactam and meropenem/meropenem-vaborbactam in this collection (Table 1B). These discrepancies were due to the different cutoff points established by EUCAST in the case of meropenem-vaborbactam and to the five isolates that had MICs of 4 mg/L and 2 mg/L for imipenem and imipenem/relebactam, respectively.

New antibiotics (cefiderocol, plazomicin, meropenem/vaborbactam, and imipenem/relebactam) have significantly improved the treatment options for CPE infections (Doi, 2019). In our study, all CP-Kpn isolates showed > 75% susceptibility to these antibiotics. In accordance with previous studies (Doi, 2019), plazomicin was not active against ST147 isolates carrying *rmtF1* (OXA-48 or NDM-1 producers), and meropenem/vaborbactam and imipenem/relebactam activities were higher in KPC-producers compared with OXA-48 and metallo- β -lactamases producers. As previously described (Yamano, 2019), cefiderocol showed good activity *in vitro* against CP-Kpn, irrespective of the carbapenemase types. The seven KPC-23-producing isolates were resistant (CMI > 16 mg/L) to ceftazidime/avibactam. This carbapenemase was previously associated with a decrease in ceftazidime/avibactam susceptibility (Galani et al., 2019). Overall susceptibility to colistin and meropenem decreased from 95.5 to 81.3%, respectively, in 2013 (Oteo et al., 2015) to 90.5% and 52.3% in this study, due primarily to the rise in KPC and NDM. The rate of colistin resistance was lower than previously reported, associated with KPC-producers (Di Pilato et al., 2021) but was consistent with the rate reported in a recent Spanish study (Vázquez-Ucha et al., 2021).

The increasing worldwide dispersion of carbapenemases is due to a mixed spread: i) clonal, with the existence of high-risk clones predominant in the carbapenemase-producing *K. pneumoniae* population, and ii) polyclonal, with the spread of conjugative epidemic plasmids capable of carrying the different carbapenemases genes (Villa et al., 2014; Mairi et al., 2018; Pérez-Vázquez et al., 2019a; Kraftova et al., 2021).

Our study elucidates the epidemiology of CP-Kpn and CP-Eco in Spain using WGS in 403 clinical isolates, representing a first step toward the integration of WGS in CPE surveillance in Spain, compliant with the high-priority recommendations of the ECDC (ECDC, 2019). Our findings will aid in the development of the Network of Laboratories for the Surveillance of Resistant Microorganisms (RedLabRA; Cañada-García et al., 2021). The data generated by this study may serve as a reliable benchmark for CP-Kpn status in Spain in the year before the COVID-19 pandemic, facilitating implementation of control measures.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Spanish Agency for Medicines and Health Products (code JOI-AVI-2019-01) and the Research Ethics Committee of the University Hospitals Virgen Macarena and Virgen del Rocío (Sevilla, Spain). Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 STUDY GROUP

Members of the GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 Collaborating Group are as follows:

Mariela Martínez Ramírez (H. Universitario de Guadalajara, Guadalajara); Pilar Zamarrón (Hospital Virgen de la Salud, Toledo); Miriam Albert Hernández (Hospital Virgen de la Concha Complejo Asistencial de Zamora, Zamora); M. Pilar Ortega Lafont (Complejo Asistencial Universitario de Burgos, Burgos); Emilia Cercenado (H. General Universitario Gregorio Marañón, CIBERES Madrid); Cristobal del Rosario and Jose Luis Perez Arellano (Hospital Universitario Insular Materno Infantil de Gran Canaria, Las Palmas); María Lecuona (Hospital Universitario de Canarias, La Laguna, Sta. Cruz Tenerife); Luis López-Urrutia Lorente (H. Universitario Río Hortega, Valladolid); José Leiva and José Luis del Pozo (Clínica Universidad de Navarra, Navarra); Salvador Giner and Juan Frasset (H. Universitario La Fe de Valencia, Valencia); Lidia García Agudo and Soledad Illescas (H General Universitario de Ciudad Real, Ciudad Real); Pedro de la Iglesia (Hospital de Cabueñes, Asturias); Rosario Sánchez Benito (Hospital San Pedro de Alcántara, Cáceres); Eugenio Garduño (Hospital Universitario Badajoz, Badajoz); M^a Isabel Fernández Natal and Marta Arias (Complejo Asistencial Universitario de León, León); Marta Lamata Subero (Fundación Hospital de Calahorra Megalab, La Rioja); Mar Olga Pérez Moreno (H. Verge de la Cinta de Tortosa, Tarragona); Ana Isabel López-Calleja (H. Universitario Miguel Servet, Zaragoza); Luis Torres Sopena (H. San Jorge de Huesca, Huesca); José Manuel Azcona (H. San Pedro, La Rioja); Alba Belles and Mercè García González (H. Universitario Arnau de Vilanova, Lleida); Miriam Valverde Troya and Begoña Palop (Hospital Regional Universitario de Málaga, Málaga); Fernando García Garrote (Hospital Universitario Lucus Augusti, Lugo); Jose Luis Barrios Andrés and Leyre López Soria (Hospital Universitario Cruces, Vizcaya); Adelina Gimeno (H. General Universitario de Alicante, Alicante);

Susana Sabater (Hospital General de Castellón, Castellón); Ester Clapés Sanchez (Hospital Dr. Josep Trueta de Girona, Girona); Jennifer Villa (H. Universitario 12 de Octubre, Madrid); Nuria Iglesias Nuñez and Rafael Sánchez Arroyo (Hospital Nuestra Señora de Sonsoles Ávila); Inmaculada García García (Complejo Hospitalario de Salamanca, Salamanca); Susana Hernando (Real Hospital General de Segovia, Segovia); Cristina Seral and Javier Castillo (H. Clínico Universitario Lozano Blesa, Zaragoza); Eva Riquelme Bravo, Caridad Sainz de Baranda, and Oscar Esparcia Rodríguez (C. Hospitalario Universitario de Albacete, Albacete); Jorge Gaitán, María Huertas (Hospital General La Mancha Centro Alcazar, Ciudad Real); M.^a José Rodríguez Escudero (Hospital General Virgen de La Luz, Cuenca); Carmen Aldea and Nerea Sanchez (Complejo Asistencial de Soria, Soria); Antonio Casabella Pernas (Hospital Universitari Parc Taulí, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Sabadell, Barcelona); M.^a Dolores Quesada (H. Germans Trias i Pujol, Badalona Barcelona); Maria Pilar Chocarro and Francisco Javier Ramos (Hospital Obispo Polanco, Teruel); Carmina Martí Sala (H. General de Granollers, Barcelona); Laura Mora, Encarnación Clavijo (H. Virgen de la Victoria, Málaga); Natalia Chueca and Federico García (H. Clínico San Cecilio, Granada); José Gutierrez Fernández (Hospital Virgen de las Nieves, Granada); Juan Manuel Sánchez Hospital de Jerez (Jerez de la Frontera, Cádiz); Fátima Galán Sánchez (H. Puerta del Mar, Cádiz); Carmen Liébana and Carolina Roldán (Hospital Universitario de Jaén, Jaén); M.^a Isabel Cabeza (Hospital de Poniente, Almería); José María Saavedra (Hospital Juan Ramón Jiménez Huelva); M.^a Teresa Cabezas Fernández (Complejo Hospitalario de Torrecárdenas, Almería); Lucía Martínez Lamas and Sonia Rey Cao (Complejo Hospitalario Universitario de Vigo, Pontevedra); M.^a Isabel Paz Vidal (Complejo Hospitalario de Orense, Orense); Raquel Elisa Rodríguez Tarazona (Hospital Santos Reyes, Aranda de Duero, Burgos); Amparo Coira Nieto and M.^a Luisa Pérez del Molino Bernal (Hospital Clínico Universitario de Santiago de Compostela A Coruña); María Gomáriz Díaz (H. Universitario de Donostia, Guipuzcoa); Matxalen Vidal-García and Jose Luis Díaz de Tuesta (Hospital Universitario de Basurto, IIS Biocruces, Bizkaia); Moises García Bravo and Almudena Tinajas (H. General Río Carrión, Palencia); Andrés Canut Blasco and M.^a Luz Albina Cordón Rodríguez (H. Universitario de Álava, Álava); Nieves Gonzalo Jiménez (H. General Universitario de Elche, Alicante); Genoveva Yagüe Guirao (Hospital Virgen de la Arrixaca de Murcia, Murcia); Fe Tubau Quintano (Hospital Universitario de Bellvitge, Barcelona); Carmen Aspiroz (H. Royo Villanova Zaragoza); Nuria Prim (H. del Mar, Barcelona); Jesús Rodríguez-Baño (H. Virgen Macarena, CIBERINFEC, Seville).

REFERENCES

Barrasa-Villar, J. I., Aibar-Remón, C., Prieto-Andrés, P., Mareca-Doñate, R., and Moliner-Lahoz, J. (2017). Impact on morbidity, mortality, and length of stay of hospital-acquired infections by resistant microorganisms. *Clin. Infect. Dis.* 65, 644–652. doi: 10.1093/cid/cix411

AUTHOR CONTRIBUTIONS

JO-I and MP-V conceived and designed the study. JC-G, ZM, ÁP, ZP-B, AO, JV, RC, GB, JG-L, FN, LM-M, GR-C, and JO-I coordinated the study. JC-G, PS-C, MD-V, MC, DG, MG, IG-A, NL, XM, CP, AR, BA, and GEMARA-SEIMC/REIPI CARB-ES-19 Study Group performed the experiments. ZM, JC, MP-V, and JO-I wrote the manuscript. All authors have read, edited, and approved the final manuscript.

FUNDING

This research was supported by grants from the Instituto de Salud Carlos III (numbers PI18CIII/00030 and PI21CIII/00039). It was also supported by Plan Nacional de I + D + i 2013–2016, Instituto de Salud Carlos III, Subdirección General de Redes y Centros de Investigación Cooperativa, Ministerio de Economía y Competitividad, Spanish Network for Research in Infectious Diseases (grants RD16CIII/0004/0002, RD16/0016/0001, RD16/0016/0003, RD16/0016/0004, RD16/0016/0006, RD16/0016/0007, RD16/0016/0008, RD16/0016/0010, and RD16/0016/0011). Cofinanced by the European Development Regional Fund “A way to achieve Europe,” Operative Program Intelligent Growth 2014–2020. CIBER – Consorcio Centro de Investigación Biomédica en Red (CB21/13/00095, CB21/13/00012, CB21/13/00049, CB21/13/00054, CB21/13/00055, CB21/13/00068, CB21/13/00081, CB21/13/00084, and CB21/13/00099) (CIBERINFEC) and Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación and Unión Europea – NextGenerationEU also supported this work.

ACKNOWLEDGMENTS

We thank all participating hospitals and the Genomics Unit of the Centro Nacional de Microbiología for support with DNA sequencing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.918362/full#supplementary-material>

Supplementary Figure 1 | Distribution of yersiniabactin lineages and K-types in high-risk carbapenemase-producing *Klebsiella pneumoniae* clones.

Cañada-García, J. E., Pérez-Vázquez, M., and Oteo-Iglesias, J. (2021). RedLabRA; a spanish network of microbiology laboratories for the surveillance of antibiotic resistant microorganisms. *Rev. Esp. Quimioter.* 34, 12–14. doi: 10.37201/req/s01.03.2021

David, S., Reuter, S., Harris, S. R., Glasner, C., Feltwell, T., Argimon, S., et al. (2019). Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven

- by nosocomial spread. *Nat. Microbiol.* 4, 1919–1929. doi: 10.1038/s41564-019-0492-8
- Di Pilato, V., Errico, G., Monaco, M., Giani, T., Del Grosso, M., Antonelli, A., et al. (2021). The changing epidemiology of carbapenemase-producing *Klebsiella pneumoniae* in Italy: toward polyclonal evolution with emergence of high-risk lineages. *J. Antimicrob. Chemother.* 76, 355–361. doi: 10.1093/jac/dkaa431
- Doi, Y. (2019). Treatment Options for Carbapenem-resistant Gram-negative Bacterial Infections. *Clin. Infect. Dis.* 69, S565–S575. doi: 10.1093/cid/ciz830
- ECDC (2019). *ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations – 2019–2021*. Solna: ECDC.
- Elias, R., Duarte, A., and Perdigão, J. (2021). A Molecular perspective on colistin and *Klebsiella pneumoniae*: mode of action, resistance genetics, and phenotypic susceptibility. *Diagnostics* 11:1165. doi: 10.3390/diagnostics11071165
- European Society of Clinical Microbiology and Infectious Diseases [EUCAST] (2017). *EUCAST guidelines for detection of resistance mechanisms and specific resistances of clinical and/or epidemiological importance, version 2.0*. 1–43. Available online at: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Resistance_mechanisms/EUCAST_detection_of_resistance_mechanisms_170711.pdf [accessed April 7, 2022].
- Galani, I., Antoniadou, A., Karaiskos, I., Kontopoulou, K., Giamarellou, H., and Souli, M. (2019). Genomic characterization of a KPC-23-producing *Klebsiella pneumoniae* ST258 clinical isolate resistant to ceftazidime-avibactam. *Clin. Microbiol. Infect.* 25, e5–e76. doi: 10.1016/j.cmi.2019.03.011
- Gastmeier, P., Schwab, F., Bärwolff, S., Rüden, H., and Grundmann, H. (2006). Correlation between the genetic diversity of nosocomial pathogens and their survival time in intensive care units. *J. Hosp. Infect.* 62, 181–186. doi: 10.1016/j.jhin.2005.08.010
- Grundmann, H., Glasner, C., Albiger, B., Aanensen, D. M., Tomlinson, C. T., Andrasevici, A. T., et al. (2017). Occurrence of carbapenemase-producing *Klebsiella pneumoniae* and *Escherichia coli* in the European survey of carbapenemase-producing *Enterobacteriaceae* (EuSCAPE): a prospective, multinational study. *Lancet Infect. Dis.* 17, 153–163. doi: 10.1016/S1473-3099(16)30257-2
- Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., et al. (2014). ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents Chemother.* 58, 212–220. doi: 10.1128/AAC.01310-13
- Gutiérrez-Gutiérrez, B., Salamanca, E., de Cueto, M., Hsueh, P. R., Viale, P., Paño-Pardo, J. R., et al. (2017). Effect of appropriate combination therapy on mortality of patients with bloodstream infections due to carbapenemase-producing *Enterobacteriaceae* (INCREMENT): a retrospective cohort study. *Lancet Infect. Dis.* 17, 726–734. doi: 10.1016/S1473-3099(17)30228-1
- Haller, S., Kramer, R., Becker, K., Bohnert, J. A., Eckmanns, T., Hans, J. B., et al. (2019). Extensively drug-resistant *Klebsiella pneumoniae* ST307 outbreak, north-eastern Germany. *Eurosurveillance* 24, 1–6. doi: 10.2807/1560-7917.ES.2019.24.50.1900734
- Hunt, M., Mather, A. E., Sánchez-Busó, L., Page, A. J., Parkhill, J., Keane, J. A., et al. (2017). ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb. Genomics* 3:131. doi: 10.1099/mgen.0.000131
- Inouye, M., Dashnow, H., Raven, L. A., Schultz, M. B., Pope, B. J., Tomita, T., et al. (2014). SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 6, 1–16. doi: 10.1186/s13073-014-0090-6
- International Organization for Standardization [ISO] (2006). *Part 1*. Available online at: <https://www.iso.org/standard/41630.html> [accessed April 7, 2022].
- Kraftova, L., Finianos, M., Studentova, V., Chudejova, K., Jakubu, V., Zemlickova, H., et al. (2021). Evidence of an epidemic spread of KPC-producing *Enterobacteriales* in Czech hospitals. *Sci. Rep.* 11:285. doi: 10.1038/s41598-021-95285-z
- Lam, M. M. C., Wick, R. R., Watts, S. C., Cerdeira, L. T., Wyres, K. L., and Holt, K. E. (2021). A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat. Commun.* 12:4483. doi: 10.1038/s41467-021-24448-3
- Ludden, C., Lötsch, F., Alm, E., Kumar, N., Johansson, K., Albiger, B., et al. (2020). Cross-border spread of *bla*_{NDM-1} and *bla*_{OXA-48} positive *Klebsiella pneumoniae*: a European collaborative analysis of whole genome sequencing and epidemiological data, 2014 to 2019. *Eurosurveillance* 25:627. doi: 10.2807/1560-7917.ES.2020.25.20.2000627
- Macesic, N., Blakeway, L. V., Stewart, J. D., Hawkey, J., Wyres, K. L., Judd, L. M., et al. (2021). Silent spread of mobile colistin resistance gene *mcr-9.1* on IncHI2 'superplasmids' in clinical carbapenem-resistant *Enterobacteriales*. *Clin. Microbiol. Infect.* 27, e7–e18. doi: 10.1016/j.cmi.2021.04.020
- Mairi, A., Pantel, A., Sotto, A., Lavigne, J. P., and Touati, A. (2018). OXA-48-like carbapenemases producing *Enterobacteriaceae* in different niches. *Eur. J. Clin. Microbiol. Infect. Dis.* 37, 587–604. doi: 10.1007/s10096-017-3112-7
- Miró, E., Agüero, J., Larrosa, M. N., Fernández, A., Conejo, M. C., Bou, G., et al. (2013). Prevalence and molecular epidemiology of acquired AmpC β -lactamases and carbapenemases in *Enterobacteriaceae* isolates from 35 hospitals in Spain. *Eur. J. Clin. Microbiol. Infect. Dis.* 32, 253–259. doi: 10.1007/s10096-012-1737-0
- Oteo, J., Ortega, A., Bartolomé, R., Bou, G., Conejo, C., Fernández-Martínez, M., et al. (2015). Prospective multicenter study of carbapenemase-producing *Enterobacteriaceae* from 83 hospitals in Spain reveals high in vitro susceptibility to colistin and meropenem. *Antimicrob. Agents Chemother.* 59, 3406–3412. doi: 10.1128/AAC.00086-15
- Oteo-Iglesias, J., Pérez-Vázquez, M., Campoy, P. S., Moure, Z., Romero, I. S., Benito, R. S., et al. (2020). Emergence of blood infections caused by carbapenemase-producing *Klebsiella pneumoniae* ST307 in Spain. *J. Antimicrob. Chemother.* 75, 3402–3405. doi: 10.1093/jac/dkaa301
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Pérez-Vázquez, M., Oteo-Iglesias, J., Sola-Campoy, P. J., CarrizoManzoni, H., Bautista, V., Lara, N., et al. (2019a). Characterization of carbapenemase-producing *Klebsiella oxytoca* in Spain, 2016–2017. *Antimicrob. Agents Chemother.* 63, e2518–e2529. doi: 10.1128/AAC.02529-18
- Pérez-Vázquez, M., Sola Campoy, P. J., Ortega, A., Bautista, V., Monzón, S., Ruiz-Carrascoso, G., et al. (2019b). Emergence of NDM-producing *Klebsiella pneumoniae* and *Escherichia coli* in Spain: Phylogeny, resistome, virulence and plasmids encoding *bla*_{NDM-like} genes as determined by WGS. *J. Antimicrob. Chemother.* 74, 3489–3496. doi: 10.1093/jac/dkz366
- Rada, A. M., De La Cadena, E., Agudelo, C., Orozco, N., Pallares, C., et al. (2020). Dynamics of *bla*_{KPC-2} dissemination from non-CG258 *Klebsiella pneumoniae* to other *Enterobacteriales* via IncN plasmids in an area of high endemicity. *Antimicrob. Agents Chemother.* 64, e1720–e1743. doi: 10.1128/AAC.01743-20
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Stamatakis, A. (2006). RAXML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* 18, 318–327. doi: 10.1016/S1473-3099(17)30753-3
- Tamma, P. D., Goodman, K. E., Harris, A. D., Tekle, T., Roberts, A., Taiwo, A., et al. (2017). Comparing the outcomes of patients with carbapenemase-producing and non-carbapenemase-producing carbapenem-resistant *enterobacteriaceae* bacteremia. *Clin. Infect. Dis.* 64, 257–264. doi: 10.1093/cid/ciw741
- Vázquez-Ucha, J. C., Seoane-Estévez, A., Rodiño-Janeiro, B. K., González-Bardanca, M., Conde-Pérez, K., Martínez-Gutián, M., et al. (2021). Activity of imipenem/relebactam against a Spanish nationwide collection of carbapenemase-producing *Enterobacteriales*. *J. Antimicrob. Chemother.* 76, 1498–1510. doi: 10.1093/jac/dkab043
- Villa, J., Viedma, E., Brañas, P., Orellana, M. A., Otero, J. R., and Chaves, F. (2014). Multiclonal spread of VIM-1-producing *Enterobacter cloacae* isolates associated

- with In624 and In488 integrons located in an IncHI2 plasmid. *Int. J. Antimicrob. Agents* 43, 451–455. doi: 10.1016/j.ijantimicag.2014.02.006
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:5595. doi: 10.1371/journal.pcbi.1005595
- Wyres, K. L., Hawkey, J., Hetland, M. A. K., Fostervold, A., Wick, R. R., Judd, L. M., et al. (2019). Emergence and rapid global dissemination of CTX-M-15-associated *Klebsiella pneumoniae* strain ST307. *J. Antimicrob. Chemother.* 74, 577–581. doi: 10.1093/jac/dky492
- Yamano, Y. (2019). In Vitro Activity of Cefiderocol Against a Broad Range of Clinically Important Gram-negative Bacteria. *Clin. Infect. Dis.* 69, S544–S551. doi: 10.1093/cid/ciz827

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Citation: Cañada-García JE, Moure Z, Sola-Campoy PJ, Delgado-Valverde M, Cano ME, Gijón D, González M, Gracia-Ahufinger I, Larrosa N, Mulet X, Pitart C, Rivera A, Bou G, Calvo J, Cantón R, González-López JJ, Martínez-Martínez L, Navarro F, Oliver A, Palacios-Baena ZR, Pascual Á, Ruiz-Carrascoso G, Vila J, Aracil B, Pérez-Vázquez M, Oteo-Iglesias J and the GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 Study Group (2022) CARB-ES-19 Multicenter Study of Carbapenemase-Producing *Klebsiella pneumoniae* and *Escherichia coli* From All Spanish Provinces Reveals Interregional Spread of High-Risk Clones Such as ST307/OXA-48 and ST512/KPC-3. *Front. Microbiol.* 13:918362. doi: 10.3389/fmicb.2022.918362

Copyright © 2022 Cañada-García, Moure, Sola-Campoy, Delgado-Valverde, Cano, Gijón, González, Gracia-Ahufinger, Larrosa, Mulet, Pitart, Rivera, Bou, Calvo, Cantón, González-López, Martínez-Martínez, Navarro, Oliver, Palacios-Baena, Pascual, Ruiz-Carrascoso, Vila, Aracil, Pérez-Vázquez, Oteo-Iglesias and the GEMARA/GEIRAS-SEIMC/REIPI CARB-ES-19 Study Group. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Early Transcriptional Response to Monensin in Sensitive and Resistant Strains of *Eimeria tenella*

Hongtao Zhang¹, Lei Zhang¹, Hongbin Si¹, Xianyong Liu², Xun Suo² and Dandan Hu^{1*}

¹ College of Animal Science and Technology, Guangxi University, Nanning, China, ² Key Laboratory of Animal Epidemiology and Zoonosis of Ministry of Agriculture, National Animal Protozoa Laboratory, College of Veterinary Medicine, China Agricultural University, Beijing, China

OPEN ACCESS

Edited by:

Ibrahim Bitar,
Charles University, Czechia

Reviewed by:

Damer Blake,
Royal Veterinary College (RVC),
United Kingdom
Vittoria Mattioni Marchetti,
Charles University, Czechia

*Correspondence:

Dandan Hu
hudandan@gxu.edu.cn

Specialty section:

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

Received: 02 May 2022

Accepted: 07 June 2022

Published: 04 July 2022

Citation:

Zhang H, Zhang L, Si H, Liu X, Suo X
and Hu D (2022) Early Transcriptional
Response to Monensin in Sensitive
and Resistant Strains of *Eimeria*
tenella. *Front. Microbiol.* 13:934153.
doi: 10.3389/fmicb.2022.934153

Eimeria parasites are the causative agents of coccidiosis, a common parasitic disease in poultry and livestock that causes significant economic losses to the animal husbandry industry. Ionophore coccidiostats, such as monensin and salinomycin, are widely used for prophylaxis of coccidiosis in poultry. Unfortunately, widespread drug resistance has compromised their efficacy. As a result, there is an increasing need to understand the targets and resistance mechanisms to anticoccidials. However, how *Eimeria* parasite genes respond to ionophores remains unclear. In this study, resistance to monensin was induced in *E. tenella* through serial generations of selection. Both sensitive and resistant *E. tenella* sporozoites were treated with 5 μ g/ml monensin for 0, 2, and 4 h, respectively. Gene transcription profiles were then compared by high-throughput sequencing. The results showed that protein translation-related genes were significantly downregulated after drug induction. A total of 1,848 DEGs were detected in the sensitive strain after 2 h of exposure, whereas only 31 were detected in the resistant strain. Among these DEGs in the sensitive strain, genes associated with protein degradation were significantly upregulated, supporting the autophagy-like parasite killing theory. Then, 4 h of exposure resulted in additional 626 and 621 DEGs for sensitive and resistant strains, respectively. This result implies that the gene transcription in sensitive strain is more susceptible to monensin treatment. Our results provide gene expression landscapes of *E. tenella* following monensin treatment. These data will contribute to a better understanding of the mechanism of drug resistance to polyether ionophores in coccidia.

Keywords: *Eimeria tenella*, monensin, drug resistance, RNA-seq, *in vitro*

INTRODUCTION

The intestinal disease coccidiosis, caused by protozoan parasites of the *Eimeria* species, is one of the most important diseases in poultry and livestock industry (Dubey, 2019). According to the latest estimates, chicken coccidiosis costs more than £10 billion annually (Blake et al., 2020). To control coccidiosis in poultry, anticoccidial drugs (e.g., monensin, salinomycin, diclazuril, etc.) are extensively used for chemoprophylaxis (Kadykalo et al., 2018), which has led to an on-board drug resistance of the parasite (Peek and Landman, 2011). Monensin, a highly effective polyether

ionophore coccidiostat, has been used commercially for more than 50 years to control coccidiosis in chickens (Chapman et al., 2010). Despite the fact that monensin-resistant *Eimeria* parasites were widely reported soon after their introduction (Jeffers, 1974, 1978; Augustine et al., 1987; Djemai et al., 2016), monensin is still extensively used in the poultry industry. Thus, there is an urgent need to clarify the mode of action and the mechanism of drug resistance in *Eimeria*.

Previous studies on the mode of action of monensin in *E. tenella* have shown that the activity of Na^+/K^+ ATPase increases after monensin treatment, leading to lactate accumulation and ATP depletion (Smith and Galloway, 1983). Consequently, a net influx of Na^+ into the parasite was suggested, which would lead to osmotic swelling and eventually to parasite bursts (Smith et al., 1980; Smith and Galloway, 1983). Besides, studies have also shown that the outer membrane of sporozoite is structurally altered after exposure to monensin (Wang et al., 2006; Del Cacho et al., 2007), and that autophagy may be involved in monensin-mediated killing of *E. tenella* (Qi et al., 2020). The current series of studies by Arrizabalaga's group has been highly successful in dissecting the mode of action and mechanism of resistance of monensin in *Toxoplasma gondii*. They reported that disruption of a mitochondrial MutS DNA repair enzyme, TgMSH-1, confers monensin resistance to the parasite (Garrison and Arrizabalaga, 2009). They also found that monensin causes TgMSH-1-dependent late-S-phase cell cycle arrest (Lavine and Arrizabalaga, 2011) and autophagy-like death in *T. gondii* (Lavine and Arrizabalaga, 2012). The autophagy can be indicated by the translocation of ATG8 to the autophagosomes, and the monensin-exposed parasites can be rescued by autophagy inhibitor 3-methyladenine (Lavine and Arrizabalaga, 2012). Monensin treatment also resulted in mitochondrial alterations in *T. gondii*, including reduced mitochondrial membrane potential and morphological changes (Lavine and Arrizabalaga, 2012; Charvat and Arrizabalaga, 2016). These deleterious effects could be mitigated by antioxidants or overexpression of antioxidant proteins, suggesting an oxidative stress induced by monensin in *T. gondii* (Charvat and Arrizabalaga, 2016).

An in-depth study on parasite response to monensin treatment is essential for clarifying the mode of action. Lavine and Arrizabalaga (Lavine and Arrizabalaga, 2011) highlighted that canonical histones were significantly upregulated in *T. gondii* after 24 h of exposure to monensin, suggesting that monensin alters the cell cycle of the parasite. More recently, Zhai et al. (2020) showed that protein biosynthesis-related pathways were significantly downregulated in *T. gondii* after monensin treatment. Quantitative proteomic analysis was also used to understand the differences in gene expression between *E. tenella*-sensitive and -resistant strains (Thabet et al., 2017). However, the different genetic backgrounds of the parasites may introduce noise into the results. Herein, we present early transcriptional profiles of *E. tenella*-resistant and its sensitive parental strain in response to monensin treatment, to better understand the mode of action and molecular mechanisms of drug resistance phenomenon.

MATERIALS AND METHODS

Ethical Statement

The use of animals in this study was approved by the Administration Committee of Laboratory Animals in Guangxi University and was performed in accordance with the Institutional Animal Care and Use Committee guidelines (Approval No: Gxu-2021-013).

Animals and Parasites

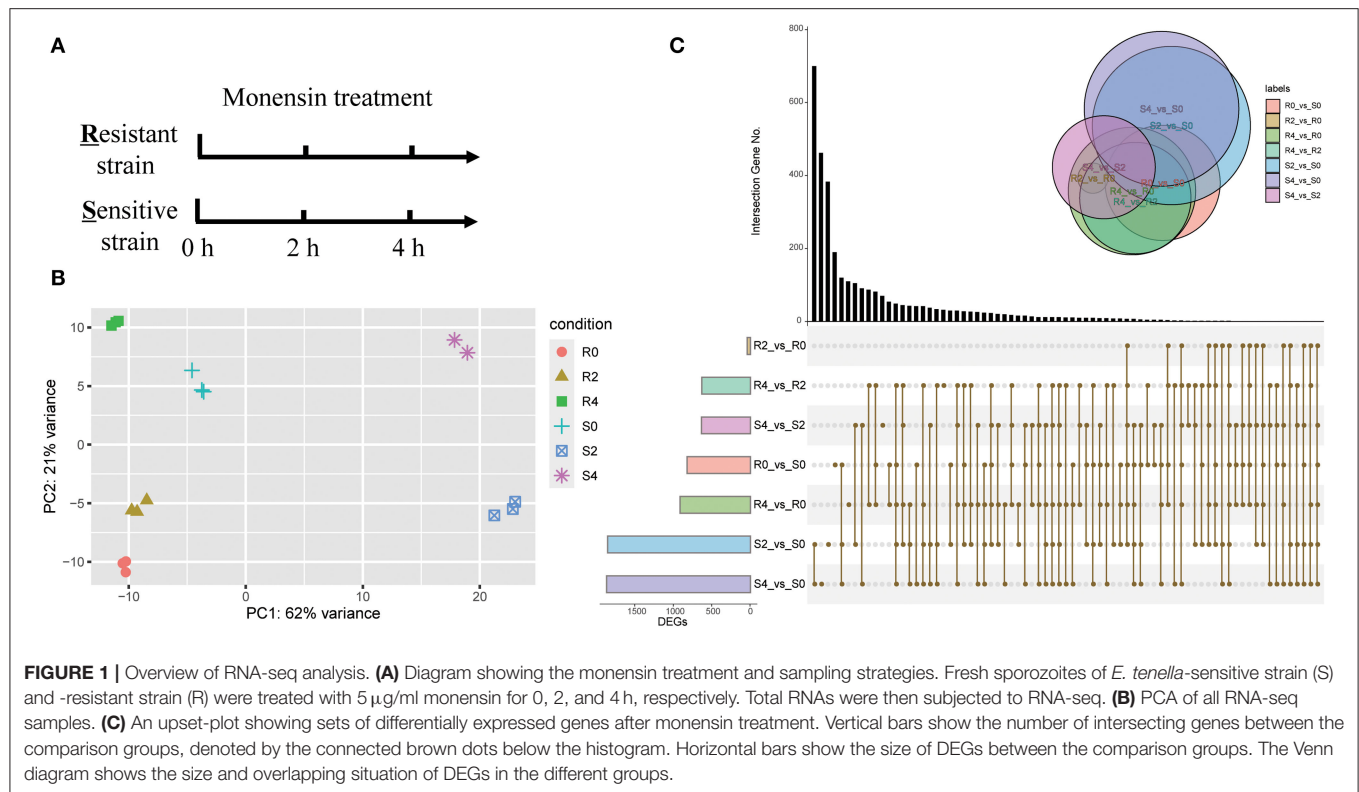
The 2–4-week-old San Huang broilers purchased from a local company were used for passaging. All birds were fed a coccidia-free diet and water *ad libitum*. The monensin-resistant Houghton strain (ETH-R) was generated by 16 generations of serial passage under gradient monensin treatment (from 50 to 250 ppm) using the *E. tenella* Houghton (ETH-S) strain as the parental strain. ETH-S was sensitive to monensin treatment at 100 ppm and the resistant strain ETH-R was susceptible to treatment at 250 ppm. Detailed procedures for resistant strain generation and genetic trajectory alteration will be published soon.

Sporozoites Purification and Monensin Treatment

ETH-S and ETH-R unsporulated oocysts were harvested after 9 days post-infection of six cages of chicken independently and sporulated in 2.5% K_2CrO_4 for 48 h, the sporozoites of both ETH-S and ETH-R were purified by Percoll density gradient method within 1 week as described previously (Dulski and Turner, 1988). Briefly, sporocysts were recovered from sporulated oocysts by a 50% Percoll density gradient after glass-dead grinding. The sporocysts were then excysted in excystation buffer at 42°C for 60 min. The sporozoites were purified by another gradient with 55% Percoll after excystation. The viability of sporozoites was tested before *in vitro* culture by trypan blue staining, and only sporozoites with >95% viability were used. For the monensin treatment, about 1.5×10^7 fresh sporozoites of ETH-S and ETH-R strains were cultured in DMEM containing 5 µg/ml monensin for 0, 2, and 4 h, respectively. They were named S0, S2, S4, R0, R2, and R4. The number represents the time of exposure and the letters “S” and “R” stand for ETH-S and ETH-R strains, respectively. Finally, all samples were washed two times with ice-cold PBS and immediately stored at −80°C in TRIzol (Invitrogen, Beijing, China). Each treatment consisted of two or three biological replicates.

RNA Extraction, Library Preparation, and RNA-Seq

Total RNAs were isolated using TRIzol reagent and genomic DNA was removed by DNase I (Tiangen Biotech Co., Ltd, Beijing, China). The purity, concentration, and integrity of RNAs were tested using NanoPhotometer® (IMPLEN, CA, USA), Qubit® RNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies, CA, USA), and RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA), respectively. Only qualified samples were used for library preparation. Sequencing libraries were generated using the Truseq™ RNA Sample Prep Kit (Illumina, CA, USA) according to the manufacturer's



recommendations. Sequencing was performed using the Illumina HiSeq TM platform to generate 150 bp paired-end reads. The original sequencing data could be found in the Sequence Read Archive database under the accession number PRJNA832043.

Bioinformatics

Paired-end clean reads were aligned to our newly generated reference genome of *E. tenella* H strain (a chromosome level genome to be published soon) using Hisat2 version 2.2.1 (Kim et al., 2019). The output SAM files were transformed into BAM files, which were then sorted and indexed. The sorted BAM files were then used for read count *via* htseq-count version 0.13.5 (Anders et al., 2015). Differentially expressed genes (DEGs) between groups were calculated by the R package DESeq2 (Love et al., 2014), and functional enrichment analysis (GO and KEGG) was performed using ClusterProfiler (v 4.0.5) (Wu et al., 2021). Gene expression with a fold change >2 or <-2 and an adjusted $p < 0.01$ was considered to be significantly differentially expressed. Transcripts per million (TPM) were calculated for each gene and used for clustered heatmap drawing.

qPCR

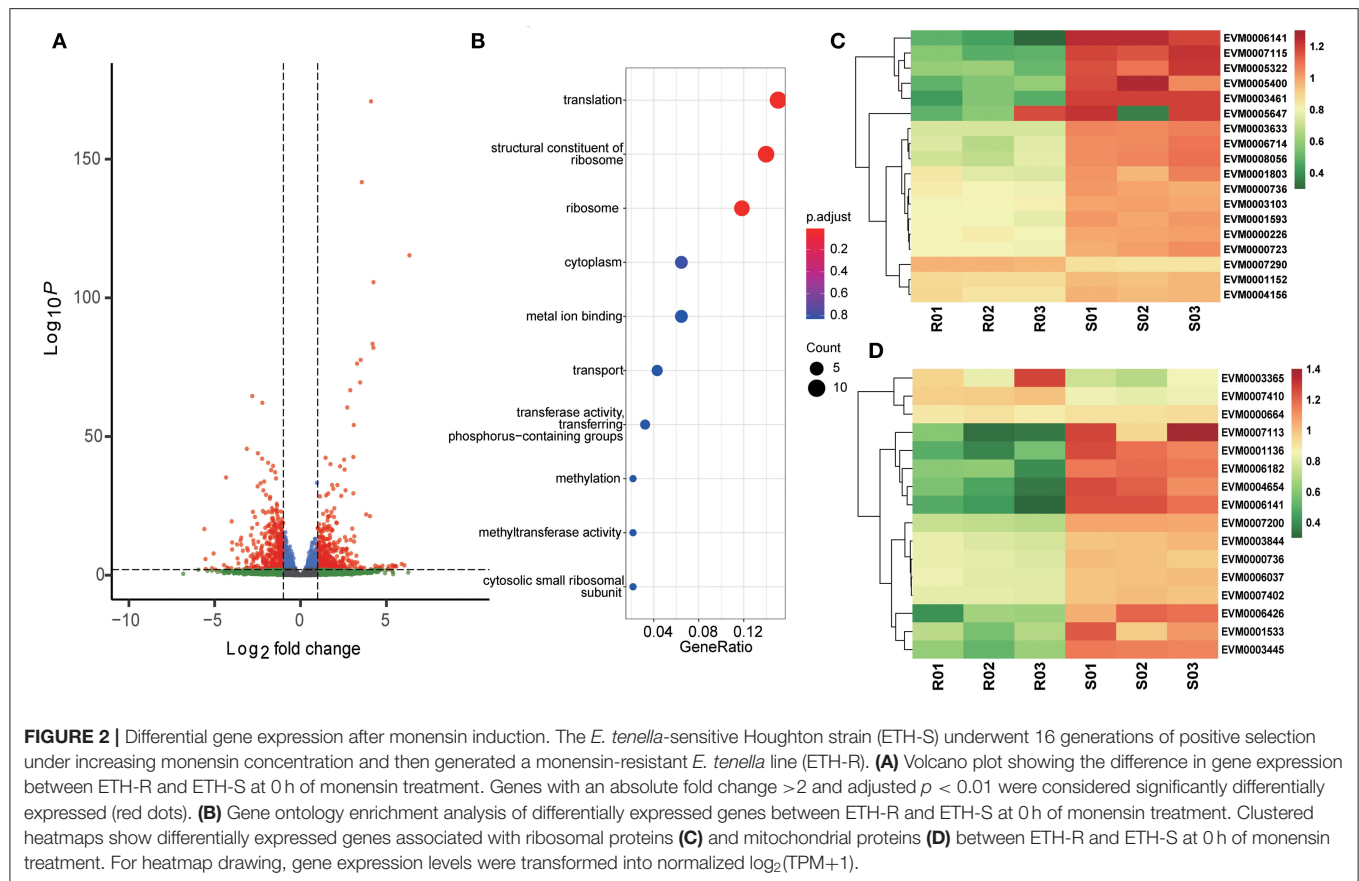
To validate the RNA-seq data, we selected four DEGs for qPCR experiments. The cDNA samples were synthesized from DNase-treated RNAs employed in the RNA-Seq using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, Beijing). PCRs were performed on Roche LightCycler® 480 system using TransScript® II Green One-Step qRT-PCR SuperMix (TransGen Biotech, Beijing). For each sample,

reactions were performed in three replicates. The primers are listed in the **Supplementary Table 1**. The expression of each gene was normalized to the reference gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as reported previously (Hu et al., 2018). Unpaired two-tailed Student's *t*-tests were used for statistical analysis using GraphPad Prism® version 9.0.0 (GraphPad Software Inc., USA).

RESULTS

Overview of the Sequencing Data and Gene Expression Patterns

As monensin primarily targets the sporozoite stage of *Eimeria*, we treated sporozoites of both resistant and sensitive strain with monensin *in vitro*. To gain insight into the early-stage gene expression landscape, sporozoites were treated for 0, 2, and 4 h and subsequently subjected to RNA-seq analysis (**Figure 1A**). A total of 410 million clean paired-end reads were generated and properly mapped to the reference genome (86.93–91.70% uniquely mapped), and samples could be grouped accordingly based on their background and monensin exposure (**Figure 1B**). To test the background of ETH-S and ETH-R, both heterozygous and homozygous SNPs were called out for all samples. It showed a relatively low level of mutants (SNP numbers ranging from 898 to 1,711, data not shown), suggesting that they have the same genetic background as the reference *E. tenella* Houghton strain. This study focuses on the transcriptome changes in the parasite in response to monensin, and the genetic variation and



evolutionary trajectory during drug-resistant induction will be published elsewhere.

To evaluate the gene expression patterns after monensin exposure, DEGs of different groups were determined (**Figure 1C**, **Supplementary Dataset 1**). The comparison between S4 and S0, as well as S2 and S0, had the highest number of DEGs in all comparisons (1,862 and 1,848, respectively), and they also had the most specific DEGs (462 and 383, respectively). On the contrary, the comparison between R2 and R0 had the minimal number (31 DEGs, **Figure 1C**, **Supplementary Dataset 2**). These results revealed considerable changes in gene expression profile after 2 h of treatment in ETH-S, but the response only started in ETH-R. We also noted a lower number of DEGs between S4 and S2 (626 DEGs, **Supplementary Dataset 3**) than between S2 vs. S0 (**Figure 1C**, **Supplementary Dataset 4**), suggesting limited gene alteration upon extended exposure in the sensitive strain.

Transcriptional Changes After Drug-Resistant Induction

The ETH-S and ETH-R sporozoite transcriptomes were directly compared to detect the differences in gene expression after generations of selection. A total of 813 genes were significantly differentially expressed, including 424 upregulated and 389 downregulated (**Figure 2A**, **Supplementary Dataset 5**). Among the downregulated genes, those responsible for ribosomes (GO:0005840, adjusted $p = 0.0681$), translation (GO:0006412,

adjusted $p = 0.0681$), and structural constituents of ribosomes (GO:0003735, adjusted $p = 0.0681$) were statistically significant (**Figure 2B**). These three GO terms shared the majority of their DEGs and were all associated with ribosomal proteins (**Figure 2C**), suggesting the downregulation of protein translation in the monensin-resistant strain. We also found that many proteins annotated as mitochondrial proteins were also downregulated in the resistant strain, including ribosomal proteins, tRNA synthetases, tRNA methyltransferase, cytochrome c, and cytochrome c oxidase subunit Vb (**Figure 2D**).

Gene Alteration in *E. tenella* Strains in Response to Monensin Treatment

After 2 h of monensin exposure, ETH-S and ETH-R parasites exhibited completely different scenarios. In the sensitive strain, a total of 1,848 DEGs were found after 2 h of treatment, of which 1,261 were upregulated and 587 were downregulated (**Figure 3A**). Among the upregulated genes, ubiquitin-dependent protein catabolic process (GO:0006511, $p = 0.004841$)-related DEGs were significantly enriched (**Figure 3B**), and more than 10 proteasome subunit proteins were also upregulated. These results suggest that protein degradation processes are activated in sensitive parasites after 2 h of treatment. Besides, many surface antigens and antioxidant proteins, such as thioredoxin and iron/manganese superoxide dismutase, were also upregulated

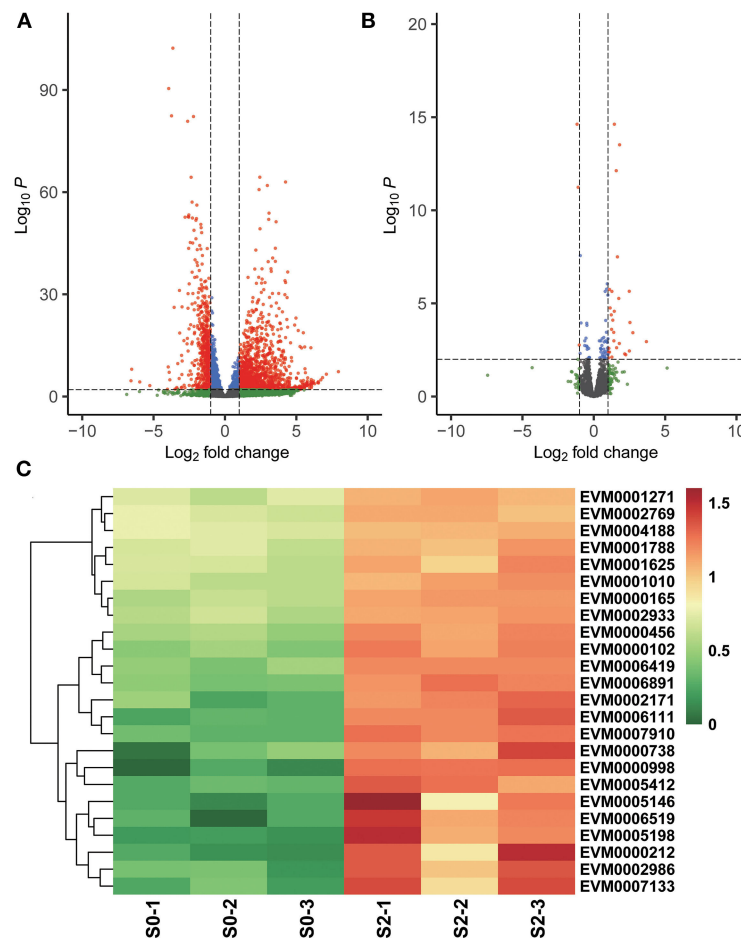


FIGURE 3 | Differential gene expression after 2 h of monensin treatment. Volcano plots showing differentially expressed genes in the monensin-sensitive ETH-S strain (A) and the monensin-resistant ETH-R strain (B) after 2 h of monensin treatment. Genes with an absolute fold change >2 and adjusted $p < 0.01$ were considered significantly differentially expressed (red dots). (C) Heatmaps showing differentially expressed genes associated with ubiquitin and proteasome in the monensin-sensitive ETH-S strain after 2 h of monensin treatment. For heatmap drawing, gene expression levels were transformed into normalized log₂(TPM+1).

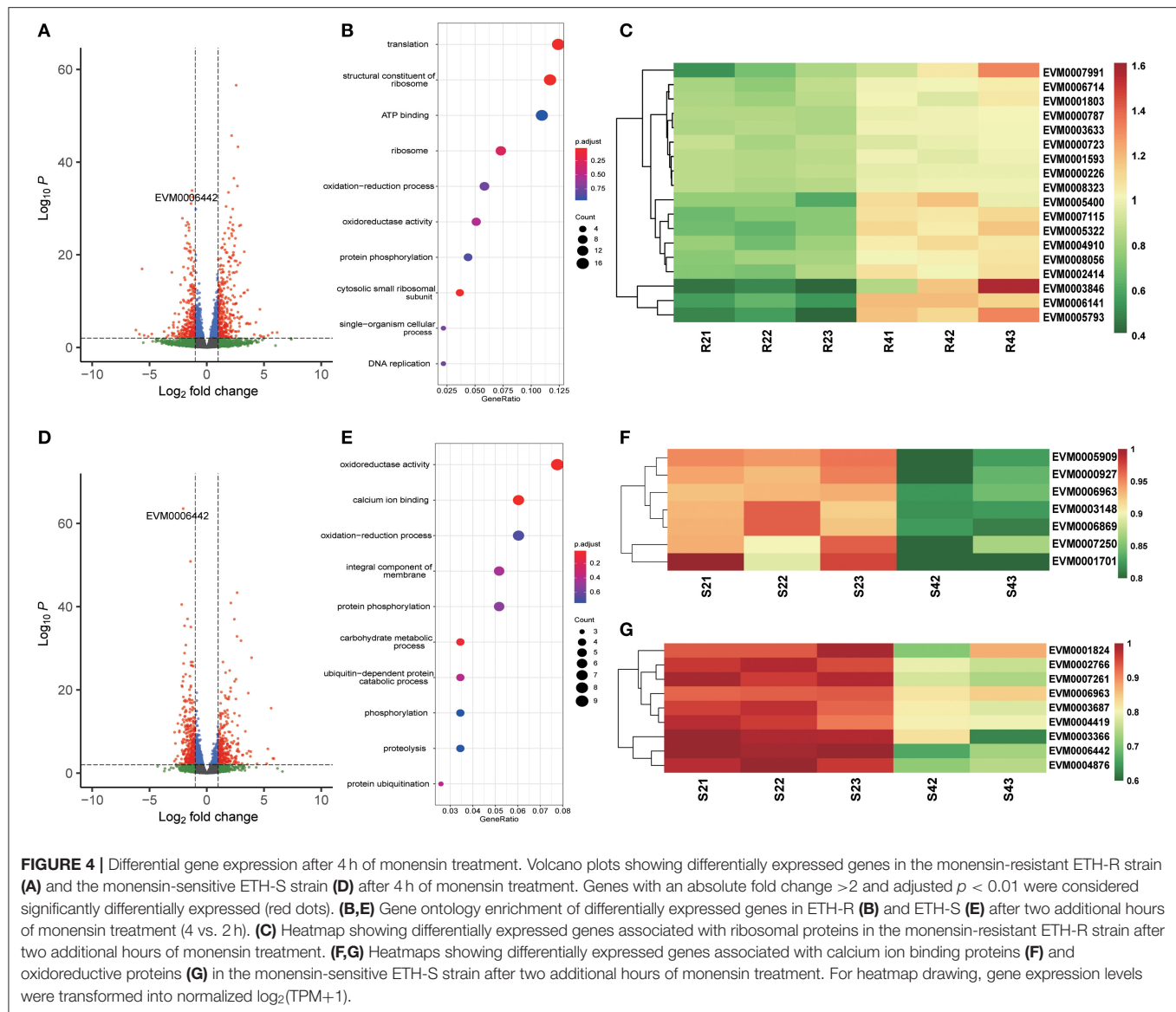
(Supplementary Dataset 4). Regarding the resistant strain, only 31 DEGs were found after 2 h of monensin exposure, of which 28 were upregulated and 3 were downregulated. An ApiAP2 transcription factor EVM0000268 was 5.79-folds higher after drug treatment, which may account for the early gene regulation.

After 4 h of monensin treatment, the monensin-resistant strain ETH-R showed more DEGs (621), of which 330 were upregulated and 291 were downregulated (Figure 4A, Supplementary Dataset 6). Translation-related genes (GO:0006412, adjusted p -value = 0.0000549) and structural constituents of ribosomes (GO:0003735, adjusted p = 0.0000549) were significantly enriched in the upregulated DEGs (Figures 4B,C). Compared to the 2-h treatment group, the sensitive ETH-S showed a total of 626 DEGs after 4 h exposure (Figure 4D), of which 292 were upregulated and 343 were downregulated. It can be noticed that the majority of gene expression alterations in the sensitive strain occurred in the first 2 h of treatment. Among the DEGs downregulated in ETH-S, genes involving in calcium ion binding (GO:0005509, adjusted

p = 0.0269) and oxidoreductase activity (GO:0016491, adjusted p = 0.0347) were significantly enriched (Figures 4E–G). Among the DEGs, we noticed that a FAD-dependent monooxygenase (EVM0006442) was significantly upregulated in ETH-R and significantly downregulated after monensin treatment in both ETH-S and ETH-R. Interestingly, EVM0006442 maintained higher transcription level in ETH-R at the same time point of treatment (Figure 5, Supplementary Figure 1).

DISCUSSION

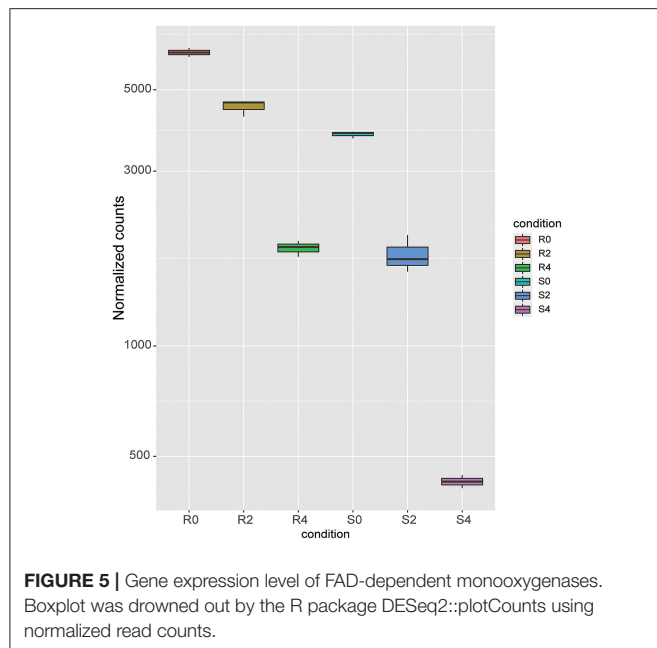
Ionophores are anticoccidial molecules commonly used in poultry industry, but resistance cases are frequently reported. To understand the molecular mechanism of drug resistance and the mode of action of ionophores, we induced a drug-resistant strain of *E. tenella* and compared the transcriptomes of the resistant strain and its sensitive ancestral strain in response to monensin treatment *in vitro*. We found that protein translation-related genes were significantly downregulated in the



monensin-resistant strain and that gene expression was more susceptible in the ancestral strain in response to monensin treatment. The *E. tenella*-sensitive strain responded very rapidly to monensin treatment, with a total of 1,848 DEGs found after 2 h of exposure, nearly 60 times more than the resistant strain. Even after 2 more h of treatment, only 621 DEGs were found in the monensin-resistant strain. This result implies that the parasite genes of the resistant strain are slightly affected by the monensin treatment. Our data give an early gene expression landscape of *E. tenella* strains in response to monensin treatment and provide important information for further studies on ionophore resistance in coccidia.

Autophagy has been reported as a potential mechanism for monensin-induced parasite killing in *Toxoplasma* and *Eimeria*. Charvat and Arrizabalaga (2016) found that monensin treatment caused oxidative stress to the parasites, resulting in reduced mitochondrial membrane potential and morphological

changes. The treatment also led to ATG8 re-localization from the cytoplasm to punctate autophagosomal structure parasite autophagy, suggesting that monensin induces autophagy in *T. gondii* (Lavine and Arrizabalaga, 2012). Monensin-induced autophagy and mitochondrial alterations (but not cell cycle arrest) can be blocked by the autophagy inhibitor 3-methyladenine (Lavine and Arrizabalaga, 2012). *In vitro* experiments on the sporozoites of *E. tenella* showed an elevated ratio of lipidated EtATG8 (form II) using western blotting with an anti-rEtATG8 antibody and also a concentrated localization of EtATG8 to the *E. tenella*-sensitive strain rather than to the resistant strain (Qi et al., 2020). In this study, a batch of proteasome subunits and ubiquitin proteins were significantly upregulated in the sensitive strain after 2 h of monensin exposure (Figure 3C), suggesting the activation of the proteasomal protein degradation process and monensin-mediated parasite killing. We also observed that several antioxidant genes, such as



thioredoxin, were upregulated in the sensitive strain after 2 h of monensin exposure, possibly suggesting that the parasite is under oxidative stress caused by monensin. These results support the previous theory that monensin-mediated oxidative stress leads to mitochondrial alterations and parasite autophagy.

After 16 generations of induction, ETH-R showed the downregulated genes in the ribosome, translation, and structural constituents of the ribosome (Figures 2B,C). These pathways were all crucial for parasite protein translation, demonstrating the downregulation of protein synthesis activities in the drug-resistant strain. Besides, this may result in reduced fitness in this field, considering that parasites require a large number of proteins to complete their lifecycle, especially during the rapid amplification schizogony stage. However, these genes were then upregulated after 4 h of treatment in the *E. tenella*-resistant strain. Different results were observed in the *T. gondii* RH strain, cytoskeletal and transmembrane proteins were differently expressed in the monensin-resistant strain determined by proteomics (Thabet et al., 2018), and the other study showed that protein biosynthesis-related pathways (spliceosome, ribosome, and protein processing in the endoplasmic reticulum) were downregulated in *T. gondii* RH strain in response to 24 h of monensin treatment (Zhai et al., 2020). These results suggest that monensin treatment affects parasite protein synthesis and may have different effects on different parasite stages (e.g., sporozoites and tachyzoites).

Cytochrome P450 (CYP450) is a well-known drug target and is also responsible for the metabolic resistance of several drugs (Zhang et al., 2017; Weedall et al., 2019; Tchouakui et al., 2021). Both FAD-dependent monooxygenases and CYP450 are crucial microsomal proteins involved in the metabolism of non-nutritive foreign compounds (known as xenobiotics), and their main function is to add molecular oxygen to xenobiotics, making them

soluble to ensure rapid excretion (Eswaramoorthy et al., 2006; Heine et al., 2018). In the malaria vector *Anopheles stephensi* (Vivekanandhan et al., 2021) and melon/cotton aphid *Aphis gossypii* (Chen et al., 2020), overexpression of the microsomal protein CYP450 contributes to the resistance of insecticides. In this study, we found that a FAD-dependent monooxygenase was significantly upregulated in the resistant strain and always maintained higher transcription level in the resistant strain at the same time point of treatment. We hypothesize that the higher expression level of FAD-dependent monooxygenase may increase the microsomal excretion capacity, which in turn promotes the metabolism of monensin in the parasite and makes it resistant to the drug. However, this should be validated by further experiments.

In general, we present here the comparative transcriptional profiles for sensitive and resistant strain of *E. tenella* in response to monensin treatment *in vitro*. This knowledge contributes to the understanding of the mode of action and mechanism of drug resistance to polyether ionophores in coccidia.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

ETHICS STATEMENT

The animal study was reviewed and approved by the Administration Committee of Laboratory Animals in Guangxi University.

AUTHOR CONTRIBUTIONS

HZ: investigation, writing the original draft, and visualization. LZ: validation and visualization. HS: writing the original draft. XL: writing, reviewing, and editing the manuscript. XS: conceptualization and writing, reviewing, and editing the manuscript. DH: software, funding acquisition, and supervision. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Natural Science Foundation of Guangxi Zhuang Autonomous Region (Grant No. 2021GXNSFBA220057), the National Natural Science Foundation of China (Grant No. 32102694), and the Specific Research Project of Guangxi for Research Base and Talents (Grant No. AD21075028).

ACKNOWLEDGMENTS

We would like to show our appreciation to Dr. Xingju Song for her technical help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.934153/full#supplementary-material>

Supplementary Figure 1 | qPCR validation of differentially expressed genes. The expression of each gene was normalized to the EtGAPDH, and the unpaired two-tailed Student's t-tests were used for statistical analysis.

Supplementary Table 1 | Primers used in qPCR experiments.

Supplementary Dataset 1 | *Eimeria tenella* gene expression under monensin treatment for different times. Gene transcription level is represented by transcripts per million (TPM).

Supplementary Dataset 2 | DEG analysis for the monensin-resistant ETH-R strain after 2h of monensin treatment.

Supplementary Dataset 3 | DEG analysis for the monensin-sensitive ETH-S strain after 4h of monensin treatment.

Supplementary Dataset 4 | DEG analysis for the monensin-sensitive ETH-S strain after 2h of monensin treatment.

Supplementary Dataset 5 | DEG analysis between *E. tenella* monensin-resistant strain and its parental strain.

Supplementary Dataset 6 | DEG analysis for the monensin-resistant ETH-R strain after 4h of monensin treatment.

REFERENCES

- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Augustine, P., Smith II, C., Danforth, H., and Ruff, M. (1987). Effect of ionophorous anticoccidials on invasion and development of *Eimeria*: comparison of sensitive and resistant isolates and correlation with drug uptake. *Poult. Sci.* 66, 960–965.
- Blake, D. P., Knox, J., Dehaeck, B., Huntington, B., Rathinam, T., Ravipati, V., et al. (2020). Re-calculating the cost of coccidiosis in chickens. *Vet. Res.* 51:115. doi: 10.1186/s13567-020-00837-2
- Chapman, H., Jeffers, T., and Williams, R. (2010). Forty years of monensin for the control of coccidiosis in poultry. *Poult. Sci.* 89, 1788–1801. doi: 10.3382/ps.2010-00931
- Charvat, R. A., and Arrizabalaga, G. (2016). Oxidative stress generated during monensin treatment contributes to altered *Toxoplasma gondii* mitochondrial function. *Sci. Rep.* 6:22997. doi: 10.1038/srep22997
- Chen, A., Zhang, H., Shan, T., Shi, X., and Gao, X. (2020). The overexpression of three cytochrome P450 genes CYP6CY14, CYP6CY22 and CYP6UN1 contributed to metabolic resistance to dinotefuran in melon/cotton aphid, *Aphis gossypii* Glover. *Pestic. Biochem. Physiol.* 167:104601. doi: 10.1016/j.pestbp.2020.104601
- Del Cacho, E., Gallego, M., Sánchez-Acedo, C., and Lillehoj, H. S. (2007). Expression of flotillin-1 on *Eimeria tenella* sporozoites and its role in host cell invasion. *J. Parasitol.* 93, 328–332. doi: 10.1645/GE-992R.1
- Djemai, S., Mekroud, A., and Jenkins, M. C. (2016). Evaluation of ionophore sensitivity of *Eimeria acervulina* and *Eimeria maxima* isolated from the Algerian to Jijel province poultry farms. *Vet. Parasitol.* 224, 77–81. doi: 10.1016/j.vetpar.2016.04.040
- Dubey, J. P. (2019). *Coccidiosis in Livestock, Poultry, Companion Animals, and Humans*. Boca Raton, FL: CRC Press, Taylor and Francis Group.
- Dulski, P., and Turner, M. (1988). The purification of sporocysts and sporozoites from *Eimeria tenella* oocysts using Percoll density gradients. *Avian Dis.* 32, 235–239.
- Eswaramoorthy, S., Bonanno, J. B., Burley, S. K., and Swaminathan, S. (2006). Mechanism of action of a flavin-containing monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9832–9837. doi: 10.1073/pnas.0602398103
- Garrison, E. M., and Arrizabalaga, G. (2009). Disruption of a mitochondrial MutS DNA repair enzyme homologue confers drug resistance in the parasite *Toxoplasma gondii*. *Mol. Microbiol.* 72, 425–441. doi: 10.1111/j.1365-2958.2009.06655.x
- Heine, T., van Berkel, W. J. H., Gassner, G., van Pée, K. H., and Tischler, D. (2018). Two-component FAD-dependent monooxygenases: current knowledge and biotechnological opportunities. *Biology* 7, 42. doi: 10.3390/biology7030042
- Hu, D., Wang, C., Wang, S., Tang, X., Duan, C., Zhang, S., et al. (2018). Comparative transcriptome analysis of *Eimeria maxima* (Apicomplexa: Eimeriidae) suggests DNA replication activities correlating with its fecundity. *BMC Genomics* 19, 699. doi: 10.1186/s12864-018-5090-2
- Jeffers, T. (1974). *Eimeria acervulina* and *E. maxima*: incidence and anticoccidial drug resistance of isolants in major broiler-producing areas. *Avian Dis.* 18, 331–342.
- Jeffers, T. (1978). *Eimeria tenella*: sensitivity of recent field isolants to monensin. *Avian Dis.* 22, 157–161.
- Kadykalo, S., Roberts, T., Thompson, M., Wilson, J., Lang, M., and Espeisse, O. (2018). The value of anticoccidials for sustainable global poultry production. *Int. J. Antimicrob. Agents* 51, 304–310. doi: 10.1016/j.ijantimicag.2017.09.004
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Lavine, M. D., and Arrizabalaga, G. (2011). The antibiotic monensin causes cell cycle disruption of *Toxoplasma gondii* mediated through the DNA repair enzyme TgMSH-1. *Antimicrob. Agents. Chemother.* 55, 745–755. doi: 10.1128/AAC.01092-10
- Lavine, M. D., and Arrizabalaga, G. (2012). Analysis of monensin sensitivity in *Toxoplasma gondii* reveals autophagy as a mechanism for drug induced death. *PLoS ONE* 7, e42107. doi: 10.1371/journal.pone.0042107
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8
- Peek, H., and Landman, W. (2011). Coccidiosis in poultry: anticoccidial products, vaccines and other prevention strategies. *Vet. Q.* 31, 143–161. doi: 10.1080/01652176.2011.605247
- Qi, N., Liao, S., Mohiuddin, M., Abuzeid, A. M., Li, J., Wu, C., et al. (2020). Autophagy induced by monensin serves as a mechanism for programmed death in *Eimeria tenella*. *Vet. Parasitol.* 287, 109181. doi: 10.1016/j.vetpar.2020.109181
- Smith, C. K., and Galloway, R. B. (1983). Influence of monensin on cation influx and glycolysis of *Eimeria tenella* sporozoites *in vitro*. *J. Parasitol.* 69, 666–670.
- Smith, I. L., and Strout, R. (1980). *Eimeria tenella*: effect of narasin, a polyether antibiotic on the ultrastructure of intracellular sporozoites. *Exp. Parasitol.* 50, 426–436.
- Tchouakui, M., Mugenzi, L. M. J., Wondji, M. J., Tchoupo, M., Njiokou, F., and Wondji, C. S. (2021). Combined over-expression of two cytochrome P450 genes exacerbates the fitness cost of pyrethroid resistance in the major African malaria vector *Anopheles funestus*. *Pestic. Biochem. Physiol.* 173, 104772. doi: 10.1016/j.pestbp.2021.104772
- Thabet, A., Honscha, W., Dauschies, A., and Bangoura, B. (2017). Quantitative proteomic studies in resistance mechanisms of *Eimeria tenella* against polyether ionophores. *Parasitol. Res.* 116, 1553–1559. doi: 10.1007/s00436-017-5432-z
- Thabet, A., Schmidt, J., Baumann, S., Honscha, W., Von Bergen, M., Dauschies, A., et al. (2018). Resistance towards monensin is proposed to be acquired in a *Toxoplasma gondii* model by reduced invasion and egress activities, in addition to increased intracellular replication. *Parasitology* 145, 313–325. doi: 10.1017/S0031182017001512
- Vivekanandhan, P., Thendralmanikandan, A., Kweka, E., and Mahande, A. (2021). Resistance to temephos in *Anopheles stephensi* larvae is associated with increased cytochrome P450 and α -esterase genes overexpression. *Int. J. Trop. Insect Sci.* 41, 2543–2548. doi: 10.1007/s42690-021-00434-6

- Wang, Z., Shen, J., Suo, X., Zhao, S., and Cao, X. (2006). Experimentally induced monensin-resistant *Eimeria tenella* and membrane fluidity of sporozoites. *Vet. Parasitol.* 138, 186–193. doi: 10.1016/j.vetpar.2006.01.056
- Weedall, G. D., Mugenzi, L. M. J., Menze, B. D., Tchouakui, M., Ibrahim, S. S., Amvongo-Adjia, N., et al. (2019). A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. *Sci. Transl. Med.* 11, eaat7386. doi: 10.1126/scitranslmed.aat7386
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). ClusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* 2, 100141. doi: 10.1016/j.xinn.2021.100141
- Zhai, B., He, J., Elsheikha, H. M., Li, J.-X., Zhu, X., and Yang, X. (2020). Transcriptional changes in *Toxoplasma gondii* in response to treatment with monensin. *Parasit. Vectors* 13, 84. doi: 10.1186/s13071-020-3970-1
- Zhang, X., Zhang, T., Liu, J., Li, M., Fu, Y., Xu, J., et al. (2017). Functional characterization of a unique cytochrome P450 in *Toxoplasma gondii*. *Oncotarget* 8, 115079–115088. doi: 10.18632/oncotarget.23023

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Zhang, Si, Liu, Suo and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

EDITED BY

Ibrahim Bitar,
Charles University,
Czechia

REVIEWED BY

Balig Panossian,
Queen Mary University of London,
United Kingdom
Marc Finianos,
Charles University,
Czechia

*CORRESPONDENCE

Elisabetta Caselli
csb@unife.it

[†]These authors have contributed equally to
this work

SPECIALTY SECTION

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 15 June 2022

ACCEPTED 12 July 2022

PUBLISHED 28 July 2022

CITATION

Cason C, D'Accolti M, Soffritti I,
Mazzacane S, Comar M and Caselli E (2022)
Next-generation sequencing and PCR
technologies in monitoring the hospital
microbiome and its drug resistance.
Front. Microbiol. 13:969863.
doi: 10.3389/fmicb.2022.969863

COPYRIGHT

© 2022 Cason, D'Accolti, Soffritti,
Mazzacane, Comar and Caselli. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Next-generation sequencing and PCR technologies in monitoring the hospital microbiome and its drug resistance

Carolina Cason^{1†}, Maria D'Accolti^{2,3†}, Irene Soffritti^{2,3},
Sante Mazzacane³, Manola Comar^{1,4} and Elisabetta Caselli^{2,3*}

¹Department of Advanced Translational Microbiology, Institute for Maternal and Child Health, IRCCS "Burlo Garofolo", Trieste, Italy, ²Department of Chemical, Pharmaceutical and Agricultural Sciences, Section of Microbiology and LTITA, University of Ferrara, Ferrara, Italy, ³CIAS Research Centre, University of Ferrara, Ferrara, Italy, ⁴Department of Medical Sciences, University of Trieste, Trieste, Italy

The hospital environment significantly contributes to the onset of healthcare-associated infections (HAIs), which represent one of the most frequent complications occurring in healthcare facilities worldwide. Moreover, the increased antimicrobial resistance (AMR) characterizing HAI-associated microbes is one of the human health's main concerns, requiring the characterization of the contaminating microbial population in the hospital environment. The monitoring of surface microbiota in hospitals is generally addressed by microbial cultural isolation. However, this has some important limitations mainly relating to the inability to define the whole drug-resistance profile of the contaminating microbiota and to the long time period required to obtain the results. Hence, there is an urgent need to implement environmental surveillance systems using more effective methods. Molecular approaches, including next-generation sequencing and PCR assays, may be useful and effective tools to monitor microbial contamination, especially the growing AMR of HAI-associated pathogens. Herein, we summarize the results of our recent studies using culture-based and molecular analyses in 12 hospitals for adults and children over a 5-year period, highlighting the advantages and disadvantages of the techniques used.

KEYWORDS

NGS, microbiome, resistome, healthcare-associated infections, hospital environment

Introduction

It is generally recognized that built environments (BE) can be considered super-organisms, with their own microbiome that, in more confined environments, show more anthropic origin, less biodiversity, and higher antimicrobial resistance (AMR; [Mahnert et al., 2019](#)).

In the hospital environment, these features are very important because the hospital microbiome represents the reservoir for the so-called healthcare-associated infections (HAIs), which are a concern worldwide (Green et al., 1998; Dancer, 2008; Mahnert et al., 2019) and are the most frequent and severe complications of hospitalization (Dancer, 2014). Healthcare-associated infections affect more than four million people per year in the European Community alone, causing over 37,000 deaths (ECDC, 2013, 2015; EARS, 2017).

The hospital microbiome, derived from patients, healthcare workers, and visitors, persists in the hospital environment and thus becomes a reservoir of pathogens and a source of infection (Riggs et al., 2007). Accordingly, the risk of acquiring an infection is increased if a patient is hospitalized in a room previously occupied by an infected patient (Huang et al., 2006; Drees et al., 2008; Datta et al., 2011). This has been reported for *Staphylococcus aureus* (Boyce et al., 1997), *Enterococcus* spp. (Huang et al., 2006), *C. difficile* (Samore et al., 1996), and *Acinetobacter* spp. (Getchell-White et al., 1989). Microorganisms can persist for extended periods on inanimate surfaces, increasing the risk of colonization for hospitalized patients (Kramer et al., 2006). In particular, the ESKAPE pathogens, responsible for severe HAIs (*Enterococcus faecium*, *S. aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) and *Clostridium difficile* are all able to survive for particularly extended periods of time in the healthcare environment (Pendleton et al., 2013).

Healthcare-associated infection severity is tightly associated with AMR concern, which has been continuously increasing in previous decades. Increased antimicrobial resistance of pathogens is particularly frequent in the hospital environment, where the fast rise of microbial resistance is both favored by the selective pressure exerted by the continuous and increased use of antimicrobial drugs and disinfectants (Kampf, 2018); and is also driven by the transfer of AMR genes between taxa through lateral gene transfer (LGT), which represents one of the most dramatic and detrimental consequences of anthropogenic impacts on the evolution of other species (Lax and Gilbert, 2015).

Furthermore, HAIs can be caused by fungi as well as bacteria (Boyce et al., 1997; Weber et al., 2010, 2013; Dancer, 2014), which also start to exhibit drug-resistance, demonstrating that all these microbial components should be controlled to counteract AMR and HAI concerns (D'Accolti et al., 2022). Surveillance appears to be a key component of control strategies, because a deep knowledge of the hospital microbiome may enable real-time countermeasures to prevent AMR spread and possible HAI outbreaks, as well as enable monitoring of the effectiveness of sanitation strategies.

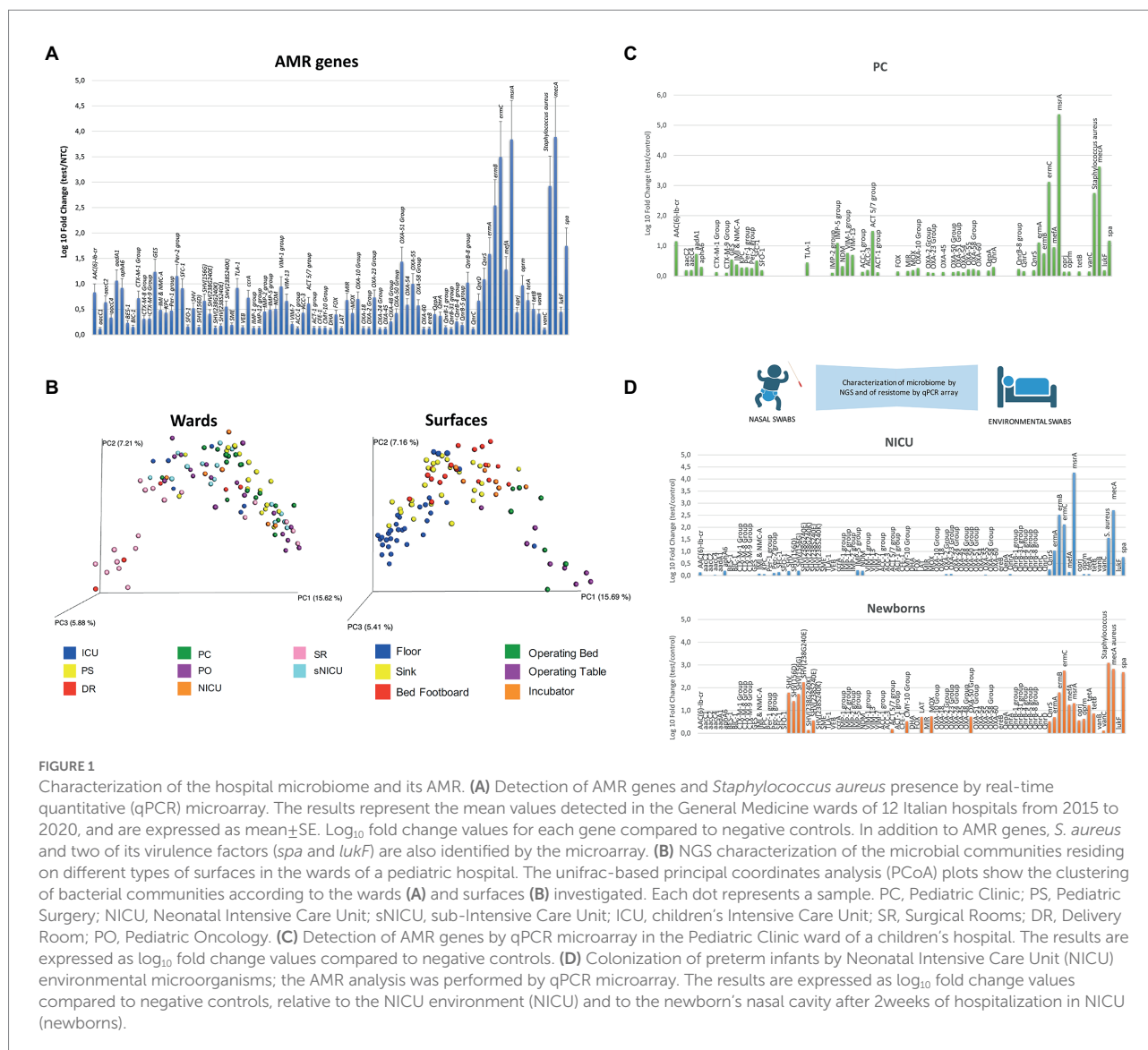
Monitoring the environmental microbial contamination: Culture-based and molecular approaches

Microbiological contamination can be assessed by different methods, one of which is adenosine triphosphate (ATP)

bioluminescence (D'Accolti et al., 2019). Although it is used often due to its low cost and rapidity, it has low sensitivity and is unspecific, detecting not only microbes but also organic materials (Dancer, 2014). Thus, this method may be useful for determining the cleanliness of surfaces but it is not effective for assessing the degree of sanitation and the microbiological composition of the hospital microbiome, which is important to guide AMR and HAI prevention (Mitchell et al., 2013; Nante et al., 2017). Other widely used methods are based on culture isolation and CFU (colony forming unit) count. These methods have been used for decades in healthcare settings, though mostly for high-risk area monitoring (Dancer, 2014; D'Accolti et al., 2019). Sampling is usually performed by RODAC plates (Replicate Organism Detection and Counting plates, containing general or specific selective media), dip slides, nitrocellulose membranes, swabs, sponges or wipes able to collect microbes from surfaces. The samples are incubated for the appropriate time (2–7 days) at the appropriate temperature that allows the development of colonies, which are finally counted and provide a quantitative measurement of the contamination degree of specific species. Although effective, these methods are very time-consuming and are only effective for living microorganisms whose growth conditions can be reproduced in the laboratory. Thus, culture-based methods do not provide detection of unsampled or non-cultivable microbes, consequently underestimating the level of microbiological contamination and complexity (Christoff et al., 2019). Moreover, further tests are needed for species identification (i.e., biochemical or immunological tests) and for the assessment of drug susceptibility/resistance profile of the isolates, which is usually performed by manual or automated diffusion or dilution tests.

Molecular methods, being based on the analysis of the genetic material of the sample, are characterized by high rapidity, sensitivity, and specificity, allowing the detection and quantification of even a very low number of targets in a few hours, and providing an in-depth characterization of the microbial community (D'Accolti et al., 2019). These methods include qualitative and quantitative real-time polymerase chain reaction (PCR, qPCR) and qPCR microarray and can provide the microbial profile of the sample and its AMR characterization due to the presence of AMR genes that can predict the phenotypic AMR in the contaminating microbiome (Hendriksen et al., 2019).

In recent years, many specific qualitative PCR, qPCR, and qPCR microarrays have been developed and commercialized, which have been used to describe surface bioburden in a very rapid and specific manner, compared to culture-based approaches (Caselli et al., 2016, 2018, 2019a). We have previously reported the use of qPCR microarrays to assess the AMR profile of the microbiome of several Italian hospitals, providing information on AMR and its fluctuations over time or as a result of the application of specific sanitation procedures (Caselli et al., 2016, 2019a, 2020; Comar et al., 2019; Cason et al., 2021). Figure 1A summarizes the microbiome detected in the general medicine wards of 12 Italian hospitals enrolled in studies from 2015 to 2020, including public and private settings with over 100 beds, located in different



geographical regions and sampled during different periods of the year. The results, independent of the location and sampling time, showed the presence of persistent contamination by ESKAPE pathogens together with the presence of several AMR genes, conferring resistance to almost all classes of available antibiotics. In particular, *Staphylococcus* spp. was identified as a prevalent contaminant, followed by Gram-positive *Enterococcus* spp., Gram-negative *Enterobacteriaceae*, and fungi (*Candida* and *Aspergillus* spp.; Caselli et al., 2016, 2018). There was a high prevalence detected of the gene conferring the resistance to methicillin (*mecA*) in all sampled settings. The high prevalence of *mecA* was likely due to the high prevalence of *S. aureus* that was also detected on the sampled surfaces (Figure 1A). This was identified by the specific qPCR which was included in the microarray due to previous results, showing that *Staphylococcus* genus was the most prevalent among hospital surface contaminants (Caselli et al., 2016, 2018). The microarray also included two virulence genes

characterizing *S. aureus* among the other *Staphylococcus* species with coding for Staphylococcal protein A (*spa*), which is an important virulence factor enabling *S. aureus* to evade host immune responses (Votintseva et al., 2014), and Leucocidin F (*lukF*), a toxin leading to trans-membrane pore formation (Pedelacq et al., 1999). Both of these genes have been consistently detected together with *S. aureus* and the *mecA* gene. Other prevalent AMR genes detected on hospital surfaces included *msrA* (coding for a macrolide efflux protein present in a wide bacterial host range), *ermA-B-C* (coding for resistance to erythromycin, lincosamides, and streptogramin), *OXA-51* group (conferring resistance against beta-lactams and carbapenems), and *QnrS/QnrB-8* group (coding for resistance against quinolones). By using a sanitation procedure able to modulate the stability of the hospital microbiome compositions—based on the competitive exclusion exerted by *Bacillus* probiotics included in the cleaning system—a significant decrease (80%) in pathogens was observed in the

treated environments, accompanied by a halving of the HAI incidence (Caselli et al., 2018). This highlighted the contribution of the hospital environment to HAI acquisition. The same environmental strains decreased due to the introduction of probiotic sanitation and also decreased as etiological agents isolated from HAI patients, confirming the relationship between the persistence of the resistant hospital microbiome and HAI occurrence (Caselli et al., 2018, 2019a).

This approach can be introduced to control, in a timely manner, any change/acquisition of new AMR genes by the hospital microbiome, thus inhibiting the spread of AMR and difficult to treat HAIs. However, PCR-based methods cannot describe the whole microbiome and/or resistome, being targeted to specific microorganisms and genes. An efficient monitoring system should provide a deep characterization of the environmental bioburden, providing rapid and detailed information on the microbial population, to enable prompt monitoring of any eventual changes in microbial communities and in their AMR.

Based on these considerations, metagenomics studies involving deep-sequencing approaches have been recently identified as powerful tools. They have the potential to provide an in-depth characterization of the microbial community, and to obtain an exhaustive picture of the environmental microbiome and its interaction with the human beings inhabiting that environment.

The deep-sequencing approach: NGS and WGS

Recent advancements in DNA sequencing technologies have dramatically improved microbiological diagnosis as well as microbiological surveillance. In early metagenomic studies, the use of Sanger sequencing technology provided important advances in the knowledge of microorganisms and of their functions (Sanger et al., 1977; Gillespie et al., 2002; Breitbart et al., 2003). However, the advent of next-generation sequencing (NGS) technologies has enormously strengthened the field, enabling simultaneous sequencing of millions of DNA fragments at a lower cost than Sanger sequencing (Kircher and Kelso, 2010). Consequently, in the last decades, NGS has not only provided a deeper knowledge of the human microbiome but has greatly increased our understanding of the microbiome of different habitats including built environments. It has enabled the detection of non-cultivable microbes and has revealed new insights into indoor microbial spread. Moreover, being not limited to targeted microorganisms, NGS has provided a thorough picture of the sampled microbial population, revolutionizing our ability to study hospital-associated bacteria due to the availability of several accessible bioinformatics resources, both for microbiome and AMR studies (Hendriksen et al., 2019). Next-generation sequencing has been recently introduced for the characterization of the microbiome of hospital surfaces, identification of novel pathogens, tracking of disease outbreaks, and the study of AMR

evolution (Zhong et al., 2021). Despite the existence of different NGS platforms and sequencing technologies, all NGS procedures are able to simultaneously sequence high numbers of DNA fragments, which are subsequently mapped by bioinformatic analyses through comparison with reference genomes (Behjati and Tarpey, 2013). Next-generation sequencing can be used to sequence entire genomes or target specific genes. This enables it to reveal and trace specific pathogens and overcome the limitations of conventional characterization of pathogens by its morphological, staining, and metabolic features. The genome of microbes also contains information on drug susceptibility/resistance; therefore, the NGS technique can be used to detect possible sources of difficult-to-treat infections, allowing timely interventions and possible benefits for patient care. This approach shows strong potential for accurate and rapid identification of the transmission pathways between hospital and community settings, and has been used to monitor MRSA outbreaks, for example (Harris et al., 2013).

Deep-sequencing applications can be applied to hospital surveillance and can include amplicon sequencing, whole metagenome shotgun analyses, and RNA sequencing. RNA sequencing is a valid and promising method for microbial characterization, and can be used to determine the transcriptome of a microbial population. This is a further step toward defining microbiome functions, the two main NGS methodologies currently in use are amplicon sequencing and shotgun metagenomic sequencing (Wensel et al., 2022).

Amplicon metagenomic sequencing is an ultra-deep DNA sequencing method that focuses on specific target regions (amplicons). Short hypervariable regions of <500 base pairs (bp) harbored in conserved genes are amplified by PCR using universal primers targeting bacterial 16S ribosomal RNA (rRNA), fungal 18S rRNA, or Internal Transcribed Spacer (ITS) rRNA. The subsequent sequencing of the amplified fragments can provide the characterization of complex microbial communities, as well as the identification of individual species and detection of microorganisms of interest among many others. Furthermore, besides the identification and tracking of microorganisms of interest, 16S/18S/ITS rRNA sequencing data can be used to describe and compare the diversity of multiple complex environments through Alpha (α) and Beta (β) diversity analyses, providing useful parameters for comparing complex microbial communities.

This sequence-based approach, when applied to the study of AMR, can provide information on all AMR genes and their precursors in a microbial species or population, thus profiling the whole resistome (Zankari et al., 2013). In recent years, several studies have been published describing the use of NGS-based methods for AMR profiling in clinical, food, and environmental contexts (Sherry et al., 2013; Bengtsson-Palme et al., 2014; Bradley et al., 2015; Hasman et al., 2015; Noyes et al., 2016; Votintseva et al., 2017), and were acknowledged by the European Commission coordinating the action plan against AMR (Angers et al., 2017). Next-generation sequencing panels

for targeted sequencing are now available commercially, mainly based on amplicon sequencing techniques (including Ion AmpliSeq by Thermo Fisher, AmpliSeq by Illumina, and others).

In the shotgun metagenomic approaches, the extracted DNA is first fragmented by shotgun, generating shotgun libraries that are sequenced, thus providing unrestricted sequencing of all the microbial genomes present in a sample. Shotgun metagenome sequencing is currently performed for taxonomic profiling (diversity and abundance), allowing for parallel sequencing of DNA from all organisms within a community, with high coverage for species-level detection. The resultant sequencing reads are then aligned to a reference database to identify taxa. Thus, being independent from amplification, shotgun metagenomics allow the identification—in a unique sample and with improved genus and species level classification—of bacteria, archaea, eukaryotes, and DNA viruses, with broad taxonomic coverage, accurate functional profiling, and the possibility of detecting previously unknown species and strains of microbes.

Shotgun metagenomics have also been proposed for the detection of AMR-related genes (Bengtsson-Palme et al., 2017; Crofts et al., 2017), since this technique is able to simultaneously quantify thousands of transmissible AMR genes in a single sample, without any prior selection of specific genes to look for. Furthermore, it can provide information about the presence of specific bacterial species and virulence genes, and later, the data can be re-analyzed if needed. In recent years, the reduction in cost and increase in the quality of whole-genome sequencing (WGS) has rendered this technique affordable and feasible as a routine tool, not only in diagnostics but also for surveillance (Aarestrup et al., 2012). Genotyping using WGS to define antimicrobial susceptibility has been extensively reported for bacterial isolates, showing high concordance with phenotypic susceptibility assays (Stoesser et al., 2013; Zankari et al., 2013; Shelburne et al., 2017). However, in 2017, the EUCAST Committee reported a lack of sufficient evidence to support the use of WGS to guide clinical decisions, also identifying the high-cost and limited rapidity of inferring AMR from WGS data (Ellington et al., 2017). Moreover, an obstacle to its routine use is the generation of high amounts of bioinformatics data to interpret, rendering the provision of real-time data time-consuming for clinical decisions (Jovel et al., 2016; Wensel et al., 2022).

Taking into consideration the advantages and disadvantages of the different metagenomics approaches, most of the studies aimed at characterizing the hospital microbiome and its AMR for HAI prevention are currently based on amplicon sequencing (Poza et al., 2012; Bokulich et al., 2013; Hewitt et al., 2013; Oberauner et al., 2013; Tang et al., 2015; Rampelotto et al., 2019; Pochtovyi et al., 2021; Perry-Dow et al., 2022), whose main objective is to identify the microorganisms circulating in a certain environment at a certain time, limiting costs, and bioinformatics concerns (Jovel et al., 2016). In particular, the 16S rRNA sequencing has been mostly used for hospital monitoring, due to the high prevalence of bacteria in the hospital microbiome. The

16S rRNA gene is in fact conserved in bacteria but also contains hyper-variable regions (V1–V9) that allow differentiation between bacterial genera and species, by comparison against reference databases such as GenBank and RDP (Christoff et al., 2019). Metagenomics has been broadly applied for the study of the resistome of the human and animal microbiome (Afridi et al., 2020, 2021; Caselli et al., 2020), and of that from different habitats, including agricultural and urban soils (Crofts et al., 2017). It has also been applied for local and international AMR surveillance of specific strains, mostly associated with HAI onset, such as carbapenem and quinolone-resistance Gram-negative bacteria, multi-resistant *S. aureus* and *Enterobacterales*, and *M. tuberculosis* (Soliman et al., 2020; Lebeaux et al., 2021; Vazquez-Lopez et al., 2021; Waddington et al., 2022). Interestingly, by using a metagenomic approach, it was possible to show that more confined and cleaned environments had reduced microbial diversity and increased resistance, suggesting the implementation of strategies to restore bacterial diversity in certain built environments (Mahnert et al., 2019). Less data are available on the whole resistome of the microbial population colonizing the healthcare facilities, despite the high prevalence of drug-resistant isolates in this kind of environment. This is due to omnipresent selective pressure that is correlated with the extensive use of antibiotics and antimicrobials, which may co-select for antibiotic resistance (Crofts et al., 2017; Kampf, 2018). Interestingly, an AMR analysis of hospital and community wastewater, conducted in 23 countries, demonstrated a higher amount of AMR genes in hospital wastewater, suggesting an important role for hospital effluent as a source of AMR spread in the environment (Hassoun-Kheir et al., 2020).

In recent years, our group has applied the 16S rRNA NGS approach in various studies, to analyze the hospital microbiome in the different types of wards of a pediatric hospital. We compared NGS data with those obtained by conventional culture-based methods and PCR-based techniques (Comar et al., 2019; Cason et al., 2021; Soffritti et al., 2022). In the first study, over 100 surface critical points were analyzed by sequencing the V3 hypervariable region of the 16S rRNA, and in parallel by culture-based techniques (RODAC contact plates) and by qPCR microarray, to characterize the advantages and disadvantages of every method. The results showed that the higher sensitivity of the molecular methods compared to culture-based ones (Comar et al., 2019), provided a more accurate quantification of the actual microbial contamination and enabled the identification of different microbial communities in the diverse types of wards and surfaces (Figure 1B). The 16S rRNA NGS method was uniquely characterized by its ability to detect unsearched bacteria; however, it did not detect mycetes, which were identified and quantified by the culture-based and targeted PCR methods. The analysis of the AMR genes harbored by the complex microbial population colonizing the sampled surfaces, performed by a qPCR-microarray simultaneously detecting 81 AMR genes, allowed the sketching of the main AMR features of the microbial communities colonizing the different types of wards (Comar et al., 2019). A higher presence of AMR

genes was detected in general medicine wards (such as the Pediatric Clinic; [Figure 1C](#)) compared to more controlled areas with less microbial contamination (Surgical and Delivery Rooms). This confirmed the results of studies performed in adult hospitals, where general medicine wards were identified as the most contaminated areas. This approach also identified significant qualitative and quantitative differences between the pediatric and the adult wards that were likely due to the different drugs used, which exerted a different selective pressure on the hospital microbiome. Based on these results, the same approach was used to analyze the evolution of the nasal microbiome of preterm newborns recovering in the neonatal intensive care unit (NICU; [Figure 1D](#); [Cason et al., 2021](#)). The infants' nasal microbiome was analyzed by NGS and qPCR microarray at delivery as well as during NICU stay. This was compared with the microbiological composition of the mother's vaginal tract and from the surfaces of the delivery room and NICU environment. This analytical approach identified a prompt colonization of the newborns' nasal cavities by the NICU environment microbes (including AMR strains). These were absent just after birth but increased with the duration of NICU stay (up to 2 weeks of hospitalization), suggesting that the hospital surface microbiota was not only responsible for contact-transmitted HAIs, but it may also have been transported by air particles to reach the respiratory tract ([Cason et al., 2021](#)).

Recently, the combined NGS/qPCR analysis was used to characterize the microbiome of the Emergency Room during the COVID-19 pandemic. The results showed that the type of sanitation adopted changed the microbiome of the hospital environment, in terms of both the amount of pathogen and the prevalence of AMR ([Soffritti et al., 2022](#)), thus prompting a re-think of the sanitation approach in order to limit AMR spread and HAI onset. Taken together, the collected data highlight the importance of active environmental monitoring to define AMR features of the hospital microbiome in depth in order to counteract more efficiently AMR spread and HAI concerns.

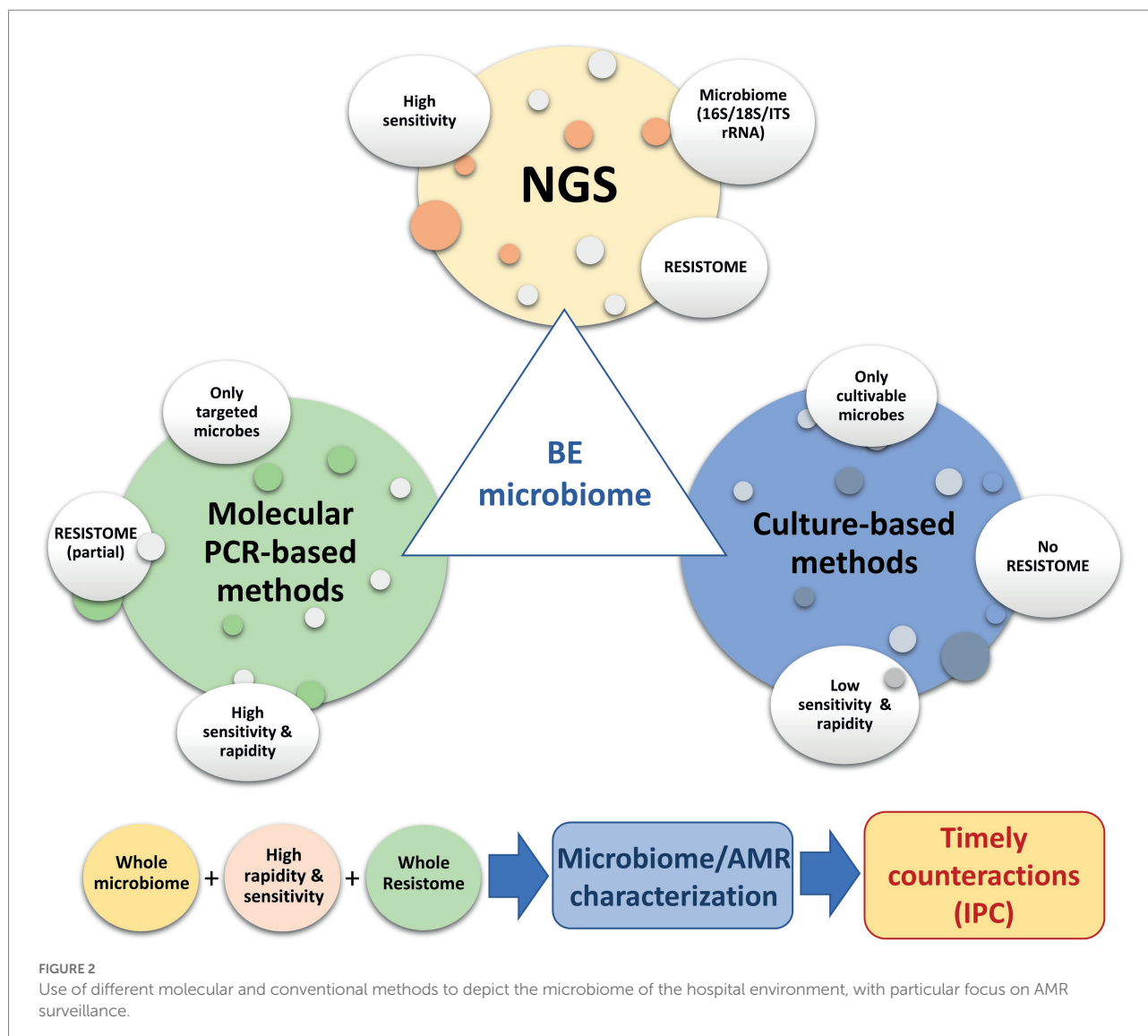
Compared to the results obtained by qPCR microarray, the use of NGS may deepen the resistome analysis, allowing more exhaustive profiling of the AMR features of the analyzed microbiome. The microarray-based results obtained by analyzing the whole contaminant population were in accord with the published NGS/WGS studies on hospital isolates, showing high amounts and prevalence of AMR genes coding for resistance against beta-lactams (including methicillin and carbapenems), quinolones, macrolides, erythromycin, lincosamides, and streptogramin ([Waddington et al., 2022](#)). This strengthens the notion that the hospital environment can be a reservoir of resistant HAI pathogens that can be transmitted to patients. Limitations to the use of NGS include the resolution of NGS being lower than WGS, thereby limiting the taxonomic resolution of closely related species ([Jenkins et al., 2012](#)), and that different studies usually report only a subset of targeted variable regions, thereby requiring standardized protocols to

avoid differing data interpretation ([Chakravorty et al., 2007](#)). However, metagenomics represents a culture-unbiased tool to obtain information on the genes responsible for AMR and on the microorganism harboring such genes ([Kristiansson et al., 2011](#); [Snitkin et al., 2012](#); [Bengtsson-Palme et al., 2016b](#)). Assembling the sequences into longer stretches of DNA to reconstruct complete genes allows the detection of AMR genes and provides information on the clonal lineage of different isolates, leading to the discovery of new resistance patterns and completely novel resistance plasmids, as well as detecting changes in taxonomic composition and other functional genes. The ability of a gene to confer resistance and to spread to other bacteria varies according to the gene's context, for example, if it is located in regions involved in horizontal gene transfer ([Kristiansson et al., 2011](#); [Bengtsson-Palme et al., 2016a](#)).

Metagenomics may result in low sensitivity due to sequencing all of the DNA present in the sample. By contrast, quantitative PCR methods may provide higher sensitivity. A direct comparison of these methods, with respect to sensitivity and specificity, is needed. Next-generation sequencing represents an invaluable tool for the in-depth characterization of the resistome of hospital surfaces. However, for daily surveillance practice, where it is necessary to have results in a short period of time and at reduced costs, the applications of molecular methods such as qPCR microarray—that identify and quantify selected panels of resistance genes associated with HAIs—have been shown to be effective and may be easier to apply ([Caselli et al., 2016, 2018, 2019b](#); [Cason et al., 2021](#)). The higher detection limit of shotgun metagenomics compared to qPCR, particularly when only a few million reads are generated per sample, should also be considered ([Bengtsson-Palme et al., 2017](#)). The main limitation of molecular methods, such as qPCR microarray, is the lack of information provided on the bacterial hosts of the resistance genes; however, the detection of these genes appears to be valuable information and should not be underestimated, for example, the phenomena of LGT. A combined approach may provide a complete picture, allowing timely countermeasures to be taken, if necessary ([Figure 2](#)).

Conclusion

During the last decade, metagenomics approaches have significantly improved analysis capacities, and the cost of new options available to scientists and clinicians is constantly decreasing. The high-throughput sequencing of 16S rRNA gene is uniquely able to identify unsearched bacteria and AMR genes, compared to targeted PCR and conventional microbiological methods, rendering it an effective first-step tool for monitoring the hospital microbiome and evaluating the safety of the environment, in terms of associated infectious risk. Next-generation sequencing can identify all the genetic material in a sample even if present in trace amounts. Due to the LGT of AMR among bacteria ([Lax and Gilbert, 2015](#)) and the low infectious



dose necessary to trigger an infection (Dancer, 2014), the presence of traces of microorganisms or resistance genes should not be underestimated in this type of monitoring. The control of AMR can benefit from NGS at various levels; it has rapidly become a reference method for studies on genomic epidemiology, clinical metagenomics, and environmental analyses, where it can help in tracking multidrug-resistant microbes and their transfer to humans.

The presence of traces that do not necessarily belong to viable microbes at the time of sampling may present a bias. This could be overcome by treating the sample with PMA (propidium monoazide) to block the amplification of dead cells' DNA, for example (Vaishampayan et al., 2013; Emerson et al., 2017).

Furthermore, 16S rRNA-based NGS cannot detect eukaryotic cells, such as fungi, or distinguish virulent from non-virulent bacterial species. Therefore, the addition of other molecular tools, such as WGS or specifically targeted PCR, should be considered

in order to obtain sensitive and accurate identification of pathogens and their AMR.

Currently, qPCR microarrays seem to be a useful tool for monitoring the hospital environment, due to their rapid results on the AMR genes of interest, less complex data processing requirement compared to NGS or WGS, and less expensive instruments (thermal cyclers; Caselli et al., 2016; Comar et al., 2019; Cason et al., 2021; Soffritti et al., 2022). In the studies of several Italian hospitals, this monitoring approach was successfully used to characterize the hospital microbiome and its AMR before and after the introduction of a sanitation system. It was able to stably shape them, resulting in a significant decrease in the environmental resistant pathogens and in their isolation from HAI patients. This confirmed the important contribution of the hospital environment to HAI occurrence (Caselli et al., 2018, 2019a).

In conclusion, currently available data suggest that NGS bacteriome analysis together with a similar evaluation of the

mycobiome (by 18S rRNA analysis) and of the resistome of the microbial population, may provide a deeper knowledge of the microbiome contaminating the hospital environment. This will lead to consistent improvements in protocols and interventions to counteract the increasing AMR diffusion. Molecular analyses can be conveniently included in surveillance programs, providing standardization of pipelines and interpretation of the genomic data, which is urgently needed. In the near future, using these molecular investigations, our understanding of the resistome may become sufficiently mature to enable the development of strategies that actively fight resistance mechanisms and stop the spread of AMR.

Author contributions

CC, MD'A and EC performed literature search, drafted the manuscript, and prepared the figures. IS, MC, and EC edited and revised the manuscript. EC initiated, supervised, and finalized the work for submission. All authors contributed to the article and approved the submitted version.

References

- Aarestrup, F. M., Brown, E. W., Detter, C., Gerner-Smidt, P., Gilmour, M. W., Harmsen, D., et al. (2012). Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response. *Emerg. Infect. Dis.* 18:e1. doi: 10.3201/eid1811.120453
- Afridi, O. K., Ali, J., and Chang, J. H. (2020). Next-generation sequencing based gut Resistome profiling of broiler chickens infected with multidrug-resistant *Escherichia coli*. *Animals (Basel)* 10, 2350–2361. doi: 10.3390/ani10122350
- Afridi, O. K., Ali, J., and Chang, J. H. (2021). Fecal microbiome and Resistome profiling of healthy and diseased Pakistani individuals using next-generation sequencing. *Microorganisms* 9, 616–631. doi: 10.3390/microorganisms9030616
- Angers, A., Petrillo, M., Patak, A., Querci, M., and Van den Eede, G. (2017). The role and implementation of next-generation sequencing Technologies in the Coordinated Action Plan against antimicrobial resistance. *JRC Workshp.* 1–16. doi: 10.2760/745099
- Behjati, S., and Tarpey, P. S. (2013). What is next generation sequencing? *Arch. Dis. Child. Educ. Pract. Ed.* 98, 236–238. doi: 10.1136/archdischild-2013-304340
- Bengtsson-Palme, J., Boulund, F., Edström, R., Feizi, A., Johnning, A., Jonsson, V. A., et al. (2016a). Strategies to improve usability and preserve accuracy in biological sequence databases. *Proteomics* 16, 2454–2460. doi: 10.1002/pmic.201600034
- Bengtsson-Palme, J., Boulund, F., Fick, J., Kristiansson, E., and Larsson, D. G. (2014). Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. *Front. Microbiol.* 5:648. doi: 10.3389/fmicb.2014.00648
- Bengtsson-Palme, J., Hammarén, R., Pal, C., Östman, M., Björleinius, B., Flach, C. F., et al. (2016b). Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. *Sci. Total Environ.* 572, 697–712. doi: 10.1016/j.scitotenv.2016.06.228
- Bengtsson-Palme, J., Larsson, D. G. J., and Kristiansson, E. (2017). Using metagenomics to investigate human and environmental resistomes. *J. Antimicrob. Chemother.* 72, 2690–2703. doi: 10.1093/jac/dkx199
- Bokulich, N. A., Mills, D. A., and Underwood, M. A. (2013). Surface microbes in the neonatal intensive care unit: changes with routine cleaning and over time. *J. Clin. Microbiol.* 51, 2617–2624. doi: 10.1128/jcm.00898-13
- Boyce, J. M., Potter-Burns, G., Chenevert, C., and King, T. (1997). Environmental contamination due to methicillin-resistant *Staphylococcus aureus*: possible infection control implications. *Infect. Control Hosp. Epidemiol.* 18, 622–627. doi: 10.2307/30141488
- Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., et al. (2015). Rapid antibiotic-resistance predictions from genome sequence data for

Acknowledgments

We would like to acknowledge the excellent technical assistance of Francesca Bini and Eleonora Mazziga.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Staphylococcus aureus and *Mycobacterium tuberculosis*. *Nat. Commun.* 6:10063. doi: 10.1038/ncomms10063

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J. M., Nulton, J., Salamon, P., et al. (2003). Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 185, 6220–6223. doi: 10.1128/JB.185.20.6220-6223.2003

Caselli, E., Arnoldo, L., Rognoni, C., D'Accolti, M., Soffritti, I., Lanzoni, L., et al. (2019a). Impact of a probiotic-based hospital sanitation on antimicrobial resistance and HAI-associated antimicrobial consumption and costs: a multicenter study. *Infect Drug Resist* 12, 501–510. doi: 10.2147/IDR.S194670

Caselli, E., Brusaferrero, S., Coccagna, M., Arnoldo, L., Berloco, F., Antonioli, P., et al. (2018). Reducing healthcare-associated infections incidence by a probiotic-based sanitation system: a multicentre, prospective, intervention study. *PLoS One* 13:e0199616. doi: 10.1371/journal.pone.0199616

Caselli, E., D'Accolti, M., Soffritti, I., Lanzoni, L., Bisi, M., Volta, A., et al. (2019b). An innovative strategy for the effective reduction of MDR pathogens from the nosocomial environment. *Adv. Exp. Med. Biol.* 1214, 79–91. doi: 10.1007/5584_2019_399

Caselli, E., D'Accolti, M., Vandini, A., Lanzoni, L., Camerada, M. T., Coccagna, M., et al. (2016). Impact of a probiotic-based cleaning intervention on the microbiota ecosystem of the hospital surfaces: focus on the Resistome Remodulation. *PLoS One* 11:e0148857. doi: 10.1371/journal.pone.0148857

Caselli, E., Fabbri, C., D'Accolti, M., Soffritti, I., Bassi, C., Mazzacane, S., et al. (2020). Defining the oral microbiome by whole-genome sequencing and resistome analysis: the complexity of the healthy picture. *BMC Microbiol.* 20:120. doi: 10.1186/s12866-020-01801-y

Cason, C., D'Accolti, M., Campisciano, G., Soffritti, I., Ponis, G., Mazzacane, S., et al. (2021). Microbial contamination in hospital environment has the potential to colonize preterm Newborns' nasal cavities. *Pathogens* 10, 615–632. doi: 10.3390/pathogens10050615

Chakravorty, S., Helb, D., Burday, M., Connell, N., and Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* 69, 330–339. doi: 10.1016/j.mimet.2007.02.005

Christoff, A. P., Sereia, A. F., Hernandez, C., and de Oliveira, L. F. (2019). Uncovering the hidden microbiota in hospital and built environments: new approaches and solutions. *Exp. Biol. Med. (Maywood)* 244, 534–542. doi: 10.1177/1535370218821857

Comar, M., D'Accolti, M., Cason, C., Soffritti, I., Campisciano, G., Lanzoni, L., et al. (2019). Introduction of NGS in environmental surveillance for healthcare-associated infection control. *Microorganisms* 7, 708–728. doi: 10.3390/microorganisms7120708

- Crofts, T. S., Gasparrini, A. J., and Dantas, G. (2017). Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* 15, 422–434. doi: 10.1038/nrmicro.2017.28
- D'Accolti, M., Soffritti, I., Bini, F., Mazziga, E., Mazzacane, S., and Caselli, E. (2022). Pathogen control in the built environment: a probiotic-based system as a remedy for the spread of antibiotic resistance. *Microorganisms* 10, 225–237. doi: 10.3390/microorganisms10020225
- D'Accolti, M., Soffritti, I., Mazzacane, S., and Caselli, E. (2019). Fighting AMR in the healthcare environment: microbiome-based sanitation approaches and monitoring tools. *Int. J. Mol. Sci.* 20, 1535–1548. doi: 10.3390/ijms20071535
- Dancer, S. J. (2008). Importance of the environment in methicillin-resistant *Staphylococcus aureus* acquisition: the case for hospital cleaning. *Lancet Infect. Dis.* 8, 101–113. doi: 10.1016/S1473-3099(07)70241-4
- Dancer, S. J. (2014). Controlling hospital-acquired infection: focus on the role of the environment and new technologies for decontamination. *Clin. Microbiol. Rev.* 27, 665–690. doi: 10.1128/CMR.00020-14
- Datta, R., Platt, R., Yokoe, D. S., and Huang, S. S. (2011). Environmental cleaning intervention and risk of acquiring multidrug-resistant organisms from prior room occupants. *Arch. Intern. Med.* 171, 491–494. doi: 10.1001/archinternmed.2011.64
- Drees, M., Snyderman, D. R., Schmid, C. H., Barefoot, L., Hansjosten, K., Vue, P. M., et al. (2008). Prior environmental contamination increases the risk of acquisition of vancomycin-resistant enterococci. *Clin. Infect. Dis.* 46, 678–685. doi: 10.1086/527394
- EARS (2017). Surveillance of antimicrobial resistance in Europe. Annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net).
- ECDC (2013). Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2011–2012. (Stockholm).
- ECDC (2015). European Surveillance of Healthcare-Associated Infections in Intensive care Units. HAI-Net ICU Protocol Protocol Version 1.02. (Stockholm).
- Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., et al. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST subcommittee. *Clin. Microbiol. Infect.* 23, 2–22. doi: 10.1016/j.cmi.2016.11.012
- Emerson, J. B., Adams, R. I., Román, C. M. B., Brooks, B., Coil, D. A., Dahlhausen, K., et al. (2017). Schrödinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome* 5, 86. doi: 10.1186/s40168-017-0285-3
- Getchell-White, S. I., Donowitz, L. G., and Gröschel, D. H. (1989). The inanimate environment of an intensive care unit as a potential source of nosocomial bacteria: evidence for long survival of *Acinetobacter calcoaceticus*. *Infect. Control Hosp. Epidemiol.* 10, 402–407. doi: 10.1086/646061
- Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., et al. (2002). Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68, 4301–4306. doi: 10.1128/AEM.68.9.4301-4306.2002
- Green, J., Wright, P. A., Gallimore, C. I., Mitchell, O., Morgan-Capner, P., and Brown, D. W. (1998). The role of environmental contamination with small round structured viruses in a hospital outbreak investigated by reverse-transcriptase polymerase chain reaction assay. *J. Hosp. Infect.* 39, 39–45. doi: 10.1016/S0195-6701(98)90241-9
- Harris, S. R., Cartwright, E. J., Torok, M. E., Holden, M. T., Brown, N. M., Ogilvy-Stuart, A. L., et al. (2013). Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect. Dis.* 13, 130–136. doi: 10.1016/S1473-3099(12)70268-2
- Hasman, H., Hammerum, A. M., Hansen, F., Hendriksen, R. S., Olesen, B., Agerso, Y., et al. (2015). Detection of mcr-1 encoding plasmid-mediated colistin-resistant *Escherichia coli* isolates from human bloodstream infection and imported chicken meat, Denmark 2015. *Euro Surveill.* 20, 1–5. doi: 10.2807/1560-7917.ES.2015.20.49.30085
- Hassoun-Kheir, N., Stabholz, Y., Kreft, J. U., de la Cruz, R., Romalde, J. L., Nesme, J., et al. (2020). Comparison of antibiotic-resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: a systematic review. *Sci. Total Environ.* 743:140804. doi: 10.1016/j.scitotenv.2020.140804
- Hendriksen, R. S., Bortolaia, V., Tate, H., Tyson, G. H., Aarestrup, F. M., and McDermott, P. F. (2019). Using genomics to track global antimicrobial resistance. *Front. Public Health* 7, 242. doi: 10.3389/fpubh.2019.00242
- Hewitt, K. M., Mannino, F. L., Gonzalez, A., Chase, J. H., Caporaso, J. G., Knight, R., et al. (2013). Bacterial diversity in two neonatal intensive care units (NICUs). *PLoS One* 8:e54703. doi: 10.1371/journal.pone.0054703
- Huang, S. S., Datta, R., and Platt, R. (2006). Risk of acquiring antibiotic-resistant bacteria from prior room occupants. *Arch. Intern. Med.* 166, 1945–1951. doi: 10.1001/archinte.166.18.1945
- Jenkins, C., Ling, C. L., Ciesielczuk, H. L., Lockwood, J., Hopkins, S., McHugh, T. D., et al. (2012). Detection and identification of bacteria in clinical samples by 16S rRNA gene sequencing: comparison of two different approaches in clinical practice. *J. Med. Microbiol.* 61, 483–488. doi: 10.1099/jmm.0.030387-0
- Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchell, T., et al. (2016). Characterization of the gut microbiome using 16S or shotgun Metagenomics. *Front. Microbiol.* 7, 459. doi: 10.3389/fmicb.2016.00459
- Kampf, G. (2018). Biocidal agents used for disinfection can enhance antibiotic resistance in gram-negative species. *Antibiotics (Basel)* 7, 110–134. doi: 10.3390/antibiotics7040110
- Kircher, M., and Kelso, J. (2010). High-throughput DNA sequencing—concepts and limitations. *BioEssays* 32, 524–536. doi: 10.1002/bies.200900181
- Kramer, A., Schwebke, I., and Kampf, G. (2006). How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infect Dis* 6, 130. doi: 10.1186/1471-2334-6-130
- Kristiansson, E., Fick, J., Janzon, A., Grabic, R., Rutgerström, C., Weijdegård, B., et al. (2011). Pyrosequencing of antibiotic-contaminated river sediments reveals high levels of resistance and gene transfer elements. *PLoS One* 6:e17038. doi: 10.1371/journal.pone.0017038
- Lax, S., and Gilbert, J. A. (2015). Hospital-associated microbiota and implications for nosocomial infections. *Trends Mol. Med.* 21, 427–432. doi: 10.1016/j.molmed.2015.03.005
- Lebeaux, R. M., Coker, M. O., Dade, E. F., Palys, T. J., Morrison, H. G., Ross, B. D., et al. (2021). The infant gut resistome is associated with *E. coli* and early-life exposures. *BMC Microbiol.* 21, 201. doi: 10.1186/s12866-021-02129-x
- Mahnert, A., Moissl-Eichinger, C., Zojer, M., Bogumil, D., Mizrahi, I., Rattei, T., et al. (2019). Man-made microbial resistances in built environments. *Nat. Commun.* 10:968. doi: 10.1038/s41467-019-08864-0
- Mitchell, B. G., Wilson, F., Dancer, S. J., and McGregor, A. (2013). Methods to evaluate environmental cleanliness in healthcare facilities. *Health infect* 18, 23–30. doi: 10.1071/HI12047
- Nante, N., Ceriale, E., Messina, G., Lenzi, D., and Manzi, P. (2017). Effectiveness of ATP bioluminescence to assess hospital cleaning: a review. *J. Prev. Med. Hyg.* 58, E177–E183.
- Noyes, N. R., Yang, X., Linke, L. M., Magnuson, R. J., Cook, S. R., Zaheer, R., et al. (2016). Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems. *Sci. Rep.* 6, 24645. doi: 10.1038/srep24645
- Oberauer, L., Zachow, C., Lackner, S., Högenauer, C., Smolle, K. H., and Berg, G. (2013). The ignored diversity: complex bacterial communities in intensive care units revealed by 16S pyrosequencing. *Sci. Rep.* 3, 1413. doi: 10.1038/srep01413
- Pedelacq, J. D., Maveyraud, L., Prevost, G., Baba-Moussa, L., Gonzalez, A., Courcelle, E., et al. (1999). The structure of a *Staphylococcus aureus* leucocidin component (Luk F-PV) reveals the fold of the water-soluble species of a family of transmembrane pore-forming toxins. *Structure* 7, 277–287. doi: 10.1016/s0969-2126(99)80038-0
- Pendleton, J. N., Gorman, S. P., and Gilmore, B. F. (2013). Clinical relevance of the ESKAPE pathogens. *Expert Rev. Anti-Infect. Ther.* 11, 297–308. doi: 10.1586/eri.13.12
- Perry-Dow, K. A., de Man, T. J. B., Halpin, A. L., Shams, A. M., Rose, L. J., and Noble-Wang, J. A. (2022). The effect of disinfectants on the microbial community on environmental healthcare surfaces using next generation sequencing. *Am. J. Infect. Control* 50, 54–60. doi: 10.1016/j.ajic.2021.08.027
- Pochtovy, A. A., Vasina, D. V., Kustova, D. D., Divisenko, E. V., Kuznetsova, N. A., Burgasova, O. A., et al. (2021). Contamination of hospital surfaces with bacterial pathogens under the current COVID-19 outbreak. *Int. J. Environ. Res. Public Health* 18, doi:10.3390/ijerph18179042. doi: 10.3390/ijerph18179042
- Poza, M., Gayoso, C., Gomez, M. J., Rumbo-Feal, S., Tomas, M., Aranda, J., et al. (2012). Exploring bacterial diversity in hospital environments by GS-FLX titanium pyrosequencing. *PLoS One* 7:e44105. doi: 10.1371/journal.pone.0044105
- Rampelotto, P. H., Sereia, A. F. R., de Oliveira, L. F. V., and Margis, R. (2019). Exploring the hospital microbiome by high-resolution 16S rRNA profiling. *Int. J. Mol. Sci.* 20, doi:10.3390/ijms20123099. doi: 10.3390/ijms20123099
- Riggs, M. M., Sethi, A. K., Zabarsky, T. F., Eckstein, E. C., Jump, R. L. P., and Donskey, C. J. (2007). Asymptomatic carriers are a potential source for transmission of epidemic and non-epidemic *Clostridium difficile* strains among long-term care facility residents. *Clin. Infect. Dis.* 45, 992–998. doi: 10.1086/521854
- Samore, M. H., Venkataraman, L., DeGirolami, P. C., Arbeit, R. D., and Karchmer, A. W. (1996). Clinical and molecular epidemiology of sporadic and clustered cases of nosocomial *Clostridium difficile* diarrhea. *Am. J. Med.* 100, 32–40. doi: 10.1016/S0002-9343(96)90008-X
- Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, C. A., et al. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 265, 687–695. doi: 10.1038/265687a0
- Shelburne, S. A., Kim, J., Munita, J. M., Sahasrabhojane, P., Shields, R. K., Press, E. G., et al. (2017). Whole-genome sequencing accurately identifies resistance

to extended-Spectrum beta-lactams for major gram-negative bacterial pathogens. *Clin. Infect. Dis.* 65, 738–745. doi: 10.1093/cid/cix417

Sherry, N. L., Porter, J. L., Seemann, T., Watkins, A., Stinear, T. P., and Howden, B. P. (2013). Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory. *J. Clin. Microbiol.* 51, 1396–1401. doi: 10.1128/JCM.03332-12

Snitkin, E. S., Zelazny, A. M., Thomas, P. J., Stock, F., Henderson, D. K., Palmore, T. N., et al. (2012). Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci. Transl. Med.* 4, 148ra116. doi: 10.1126/scitranslmed.3004129

Soffritti, I., D'Accolti, M., Cason, C., Lanzoni, L., Bisi, M., Volta, A., et al. (2022). Introduction of probiotic-based sanitation in the emergency ward of a children's hospital during the COVID-19 pandemic. *Infect. Drug Resist.* 15, 1399–1410. doi: 10.2147/IDR.S356740

Soliman, M. S., Soliman, N. S., El-Manakhly, A. R., ElBanna, S. A., Aziz, R. K., and El-Kholy, A. A. (2020). Genomic characterization of methicillin-resistant *Staphylococcus aureus* (MRSA) by high-throughput sequencing in a tertiary care hospital. *Genes (Basel)* 11, 1219–1236. doi: 10.3390/genes11101219

Stoesser, N., Batty, E. M., Eyre, D. W., Morgan, M., Wyllie, D. H., Del Ojo Elias, C., et al. (2013). Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data. *J. Antimicrob. Chemother.* 68, 2234–2244. doi: 10.1093/jac/dkt180

Tang, C. Y., Yiu, S. M., Kuo, H. Y., Tan, T. S., Liao, K. H., Liu, C. C., et al. (2015). Application of 16S rRNA metagenomics to analyze bacterial communities at a respiratory care Centre in Taiwan. *Appl. Microbiol. Biotechnol.* 99, 2871–2881. doi: 10.1007/s00253-014-6176-7

Vaishampayan, P., Probst, A. J., La Duc, M. T., Bargoma, E., Benardini, J. N., Andersen, G. L., et al. (2013). New perspectives on viable microbial communities in low-biomass cleanroom environments. *ISME J.* 7, 312–324. doi: 10.1038/ismej.2012.114

Vazquez-Lopez, R., Solano-Galvez, S., Alvarez-Hernandez, D. A., Ascencio-Aragon, J. A., Gomez-Conde, E., Pina-Leyva, C., et al. (2021). The

Beta-lactam Resistome expressed by aerobic and anaerobic Bacteria isolated from human feces of healthy donors. *Pharmaceuticals (Basel)* 14, 533–548. doi: 10.3390/ph14060533

Votintseva, A. A., Bradley, P., Pankhurst, L., Del Ojo Elias, C., Loose, M., Nilgiriwala, K., et al. (2017). Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J. Clin. Microbiol.* 55, 1285–1298. doi: 10.1128/JCM.02483-16

Votintseva, A. A., Fung, R., Miller, R. R., Knox, K., Godwin, H., Wyllie, D. H., et al. (2014). Prevalence of *Staphylococcus aureus* protein A (spa) mutants in the community and hospitals in Oxfordshire. *BMC Microbiol.* 14, 63. doi: 10.1186/1471-2180-14-63

Waddington, C., Carey, M. E., Boinett, C. J., Higginson, E., Veeraghavan, B., and Baker, S. (2022). Exploiting genomics to mitigate the public health impact of antimicrobial resistance. *Genome Med.* 14, 15. doi: 10.1186/s13073-022-01020-2

Weber, D. J., Anderson, D., and Rutala, W. A. (2013). The role of the surface environment in healthcare-associated infections. *Curr. Opin. Infect. Dis.* 26, 338–344. doi: 10.1097/QCO.0b013e3283630f04

Weber, D. J., Rutala, W. A., Miller, M. B., Huslage, K., and Sickbert-Bennett, E. (2010). Role of hospital surfaces in the transmission of emerging health care-associated pathogens: norovirus, *Clostridium difficile*, and *Acinetobacter* species. *Am. J. Infect. Control* 38, S25–S33. doi: 10.1016/j.ajic.2010.04.196

Wensel, C. R., Pluznick, J. L., Salzberg, S. L., and Sears, C. L. (2022). Next-generation sequencing: insights to advance clinical investigations of the microbiome. *J. Clin. Invest.* 132:e154944. doi: 10.1172/JCI154944

Zankari, E., Hasman, H., Kaas, R. S., Seyfarth, A. M., Agersø, Y., Lund, O., et al. (2013). Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J. Antimicrob. Chemother.* 68, 771–777. doi: 10.1093/jac/dks496

Zhong, Y., Xu, F., Wu, J., Schubert, J., and Li, M. M. (2021). Application of next generation sequencing in laboratory medicine. *Ann. Lab. Med.* 41, 25–43. doi: 10.3343/alm.2021.41.1.25



OPEN ACCESS

EDITED BY

Ibrahim Bitar,
Charles University,
Czechia

REVIEWED BY

Marc Finianos,
Charles University,
Czechia
Adam Valcek,
Vrije University Brussel,
Belgium

*CORRESPONDENCE

Patricia J. Simner
psimner1@jhmi.edu

SPECIALTY SECTION

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 20 June 2022

ACCEPTED 08 July 2022

PUBLISHED 08 August 2022

CITATION

Conzemius R, Bergman Y, Májek P,
Beisken S, Lewis S, Jacobs EB,
Tamma PD and Simner PJ (2022)
Automated antimicrobial susceptibility
testing and antimicrobial resistance
genotyping using Illumina and Oxford
Nanopore Technologies sequencing data
among *Enterobacteriaceae*.
Front. Microbiol. 13:973605.
doi: 10.3389/fmicb.2022.973605

COPYRIGHT

© 2022 Conzemius, Bergman, Májek,
Beisken, Lewis, Jacobs, Tamma and Simner.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Automated antimicrobial susceptibility testing and antimicrobial resistance genotyping using Illumina and Oxford Nanopore Technologies sequencing data among *Enterobacteriaceae*

Rick Conzemius¹, Yehudit Bergman², Peter Májek¹,
Stephan Beisken¹, Shawna Lewis², Emily B. Jacobs²,
Pranita D. Tamma³ and Patricia J. Simner^{2*}

¹Ares Genetics GmbH, Vienna, Austria, ²Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, United States, ³Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Whole-genome sequencing (WGS) enables the molecular characterization of bacterial pathogens. We compared the accuracy of the Illumina and Oxford Nanopore Technologies (ONT) sequencing platforms for the determination of AMR classes and antimicrobial susceptibility testing (AST) among 181 clinical *Enterobacteriaceae* isolates. Sequencing reads for each isolate were uploaded to AREScloud (Ares Genetics) to determine the presence of AMR markers and the predicted WGS-AST profile. The profiles of both sequencing platforms were compared to broth microdilution (BMD) AST. Isolates were delineated by resistance to third-generation cephalosporins and carbapenems as well as the presence of AMR markers to determine clinically relevant AMR classes. The overall categorical agreement (CA) was 90% (Illumina) and 88% (ONT) across all antimicrobials, 96% for the prediction of resistance to third-generation cephalosporins for both platforms, and 94% (Illumina) and 91% (ONT) for the prediction of resistance to carbapenems. Carbapenem resistance was overestimated on ONT with a major error of 16%. Sensitivity for the detection of carbapenemases, extended-spectrum β -lactamases, and plasmid-mediated *ampC* genes was 98, 95, and 70% by ONT compared to the Illumina dataset as the reference. Our results highlight the potential of the ONT platform's use in clinical microbiology laboratories. When combined with robust bioinformatics methods, WGS-AST predictions may be a future approach to guide effective antimicrobial decision-making.

KEYWORDS

antimicrobial susceptibility testing, molecular diagnostics, carbapenem-resistant *Enterobacteriaceae*, whole-genome sequencing, nanopore sequencing, point-of-care testing, WGS-AST

Introduction

The emergence of antimicrobial resistance (AMR) is recognized by leading health organizations as one of the major threats to global health. The accelerated progression of AMR requires not only the development of new therapeutics but also rapid, comprehensive, and accurate diagnostic methods to guide early and effective antimicrobial therapy (Livermore, 2004; Rochford et al., 2018; World Health Organization, 2018, 2020).

The Infectious Diseases Society of America's treatment guidance for infections caused by multidrug-resistant Gram-negative organisms highlights the importance of understanding the mechanisms mediating AMR, as antimicrobial selection may differ by mechanism. As an example, discerning whether a bloodstream infection caused by a third-generation cephalosporin-resistant organism is mediated by extended-spectrum β -lactamase (ESBL) versus AmpC β -lactamase production is important, as the recommended therapy for the former is a carbapenem and cefepime for the latter. Similarly, understanding whether a carbapenem-resistant Enterobacterales (CRE) infection is caused by the presence of a *bla*_{KPC} gene versus a *bla*_{OXA-48-like} gene is of significance, as the former is generally effectively treated by meropenem-vaborbactam whereas the latter is not.

Whole-genome sequencing (WGS) provides the ability to identify the resistome of microorganisms (i.e., all antimicrobial resistance genes harbored) but can also be used for antimicrobial susceptibility testing (WGS-AST) and potentially guide antimicrobial decision making. We have previously shown that machine learning-based WGS-AST has several advantages over rule-based AMR gene detection for determining susceptibility or resistance to antimicrobials (Lüftinger et al., 2021a). The Oxford Nanopore Technologies (ONT) WGS platform offers rapid and easy library preparation, live readout, reduced turn-around time, lower initial investment, and per-run costs compared to standard platforms such as Illumina (Greninger et al., 2015; Jain et al., 2015; Quick et al., 2016; Schmidt et al., 2017; Charalampous et al., 2019). A significant disadvantage of ONT, however, is the higher per-base error rate, with accuracies between 90 and 99%, depending on the chemistry used, compared to Illumina with 99.9% raw read accuracy. The lower fidelity of ONT reads necessitates the use of specialized bioinformatics tools to mitigate these impacts on downstream analysis (Rang et al., 2018; Wang et al., 2021).

We investigated predicted WGS-AST and AMR gene detection of 181 clinical *Enterobacteriaceae* isolates using sequencing data acquired on both the Illumina and ONT platforms. We compared the WGS-AST profiles to phenotypic broth microdilution (BMD) AST results to determine the overall categorical agreement (CA), very major error (VME), and major error (ME) for ONT and Illumina. We further assessed the suitability of both platforms to determine clinically relevant AMR classes associated with third-generation cephalosporin resistance and carbapenem resistance by combining WGS-AST results with detected molecular markers in a decision tree.

Materials and methods

Isolates

A total of 181 clinical *Enterobacteriaceae* isolates recovered from patients at The Johns Hopkins Hospital (Baltimore, MD) between 2016 and 2020 were included in the current study. The included organisms were as follows: *Klebsiella pneumoniae* [n: 151], *K. quasipneumoniae* [n: 3], *Escherichia coli* [n: 14], *Enterobacter cloacae* [n: 7], *E. hormaechei* [n: 3], *E. chengduensis* [n: 1], *E. kobei* [n: 1], and *E. roggenkampii* [n: 1]. Of these, 49 isolates were carbapenem-susceptible and 132 were carbapenem-resistant. Isolates were subcultured from frozen stocks twice on tryptic soy agar with 5% sheep blood (BD Diagnostics, Sparks, MD) for 18–24 h at 37°C.

Identification

Bacterial genus and species were determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS; Bruker Daltonics Inc., Billerica, MA).

Antimicrobial susceptibility testing

Broth microdilution AST (BMD-AST) was performed with the Sensititre GN7F and MDRGN2F (Thermo Fisher Scientific, Waltham, MA, United States) panels for Gram-negative bacteria following Clinical and Laboratory Standards Institute (CLSI) guidelines for a total of 31 antimicrobials. BMD-AST results were interpreted according to CLSI guidelines M100-S31, except fluoroquinolones were interpreted according to M100-S28 due to the restricted dilutions on the panel (CLSI, 2018, 2021). For all BMD-AST studies, quality control organisms were evaluated each day of testing.

DNA extraction

Genomic DNA was extracted from pure cultures using the DNeasy PowerSoil Pro and DNeasy PowerBiofilm kits (QIAGEN, Hilden, Germany), following the manufacturers' guidelines.

Illumina sequencing

The sequencing libraries were prepared using the Nextera DNA Flex Library Preparation Kit (Illumina, San Diego, CA, United States). Library concentrations were verified using a dsDNA fluorescent dye method on the Qubit 3.0 Fluorometer (Thermo Fisher Scientific, MA, Waltham, United States). DNA fragment size and library quality were confirmed on a 4200

TapeStation system (Agilent, Santa Clara, CA, United States). Cluster generation was performed on the cBot System (Illumina, San Diego, CA United States) using the HiSeq Rapid PE Cluster v2 and HiSeq Rapid Duo cBot Sample Loading kits (Illumina, San Diego, CA United States). Libraries were sequenced on HiSeq 2500 (using the rapid run mode) and MiSeq devices (Illumina, San Diego, CA, United States) at 7 pmol with HiSeq Rapid SBS Kit v2 (2 × 250 bp, 500-cycle kit), MiSeq Reagent Kit v2 (2 × 150 bp, 300-cycle kit), or MiSeq Reagent Kit v3 (2 × 300 bp, 600-cycle kit) chemistries, respectively.

Oxford Nanopore Technologies sequencing

Long-read genomic sequencing was performed using the third generation Oxford Nanopore MinION Mk1B and GridION X5 (Oxford, United Kingdom) sequencing instruments. Each Nanopore sequencing library was prepared using the 1D ligation kit (SQK-LSK108, Oxford Nanopore Technologies) and sequenced on R9.4 flow cells (FLO-MIN106). High-accuracy live base calling was done with Guppy 4.3+ (Oxford Nanopore, Oxford, United Kingdom) as released with the MinKNOW (Oxford Nanopore, Oxford, United Kingdom) software.

Data processing

Sequencing reads were uploaded to AREScloud (Ares Genetics) for genome assembly, quality control, identification, sequence typing, AMR marker detection, and WGS-AST (Ares Genetics GmbH, 2021). For Illumina sequencing, isolates were evaluated using FastQC v0.11.5 and fastq-stats v1.01 (Aronesty, 2011; Andrews, 2016). Datasets exceeding 4,000,000 reads were randomly subsampled using seqtk v1.2-r94 (Li, 2016). Adapter removal and trimming of low-quality paired-end reads were done using Trimmomatic v0.39 (Bolger et al., 2014). Reads were *de novo* assembled with SPAdes v3.15.25 (Prijbelski et al., 2020). ONT sequencing quality was evaluated using NanoPlot v1.39.0 and low-quality reads were removed with NanoFilt v2.8 at a quality threshold of 7 (de Coster et al., 2018). Datasets were subsampled to 600 Mb with rasusa v0.6.0, *de novo* assembled with Canu v2.1.1, and iteratively polished with Racon v1.5 (Koren et al., 2017; Vaser et al., 2017; Hall, 2021). For both sequencing platforms, assembly quality and genome completeness were determined using Quast v5.0.2 and BUSCO v5.2.2 (Mikheenko et al., 2018; Manni et al., 2021). Insert size for Illumina datasets was calculated using CollectInsertSizeMetrics v2.25.2 from Picard Tools (Broad institute, 2021). *De novo* assemblies were annotated with Prokka v1.14.1 and ribosomal RNA genes were identified using Barrnap v0.9 (Seemann, 2014, 2018). The genome annotations and ribosomal RNA genes were used for assembly quality evaluation; the Prokka features also constitute part of the input feature space for the WGS-AST machine learning models. Microbial identification was determined with

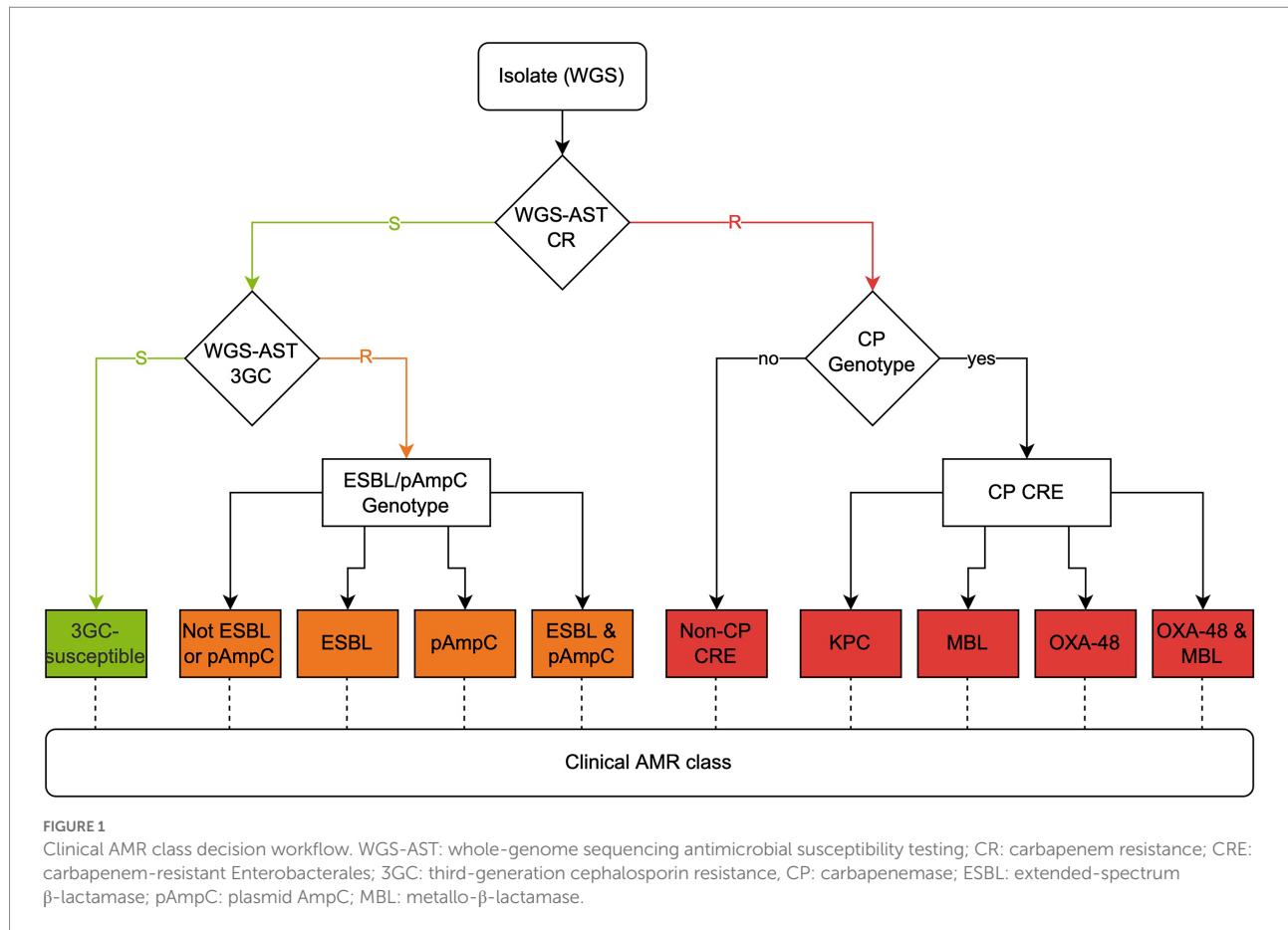
Kraken v2.0.9-beta using the MiniKraken PlusPF-8 database from 2021/05/17 (Wood and Salzberg, 2014). Genome coverage was determined using bwa mem v0.7.17-r1188 and bedtools v2.29.0 (Quinlan and Hall, 2010; Li, 2013). Resistance genes were determined with DIAMOND v1.0.11 via sequence alignment of six-frame translated genome assemblies against ARESdb with a minimum query coverage of 60% and a minimum identity of 90% for Illumina-derived assemblies and a minimum identity of 70% for ONT-derived assemblies (Buchfink et al., 2015; Ferreira et al., 2020). WGS-AST results were generated by susceptibility/resistance (S/R) stacked classification models trained per species-antimicrobial pair on ARESdb (Ferreira et al., 2020). The model stacks combine extreme gradient boosting, elastic net regularized logistic regression (ENLR), and set covering machine models as well as rule-based post-processing routines. ARESdb models were trained on features derived from sequence motifs and variants (Lüftinger et al., 2021b; Májek et al., 2021). Ertapenem was used as a proxy for the prediction of resistance to carbapenems and was optimized from previous publications, to recognize non-CP CREs (Májek et al., 2021) by recognizing changes in the *ompK35*, *ompK36*, and homologs (Doumith et al., 2009; Sugawara et al., 2016). WGS-AST was run for 60 species-antimicrobial pairs listed in Supplementary Table 1. Discrepant species identifications between the Illumina and ONT platforms were confirmed by comparing the assemblies to the reference genomes of the suspected species using FastANI (Jain et al., 2018). Data processing on AREScloud took less than 2 h for Illumina data and up to 5 h for ONT data; the application scales horizontally, i.e., samples are processed in parallel on multiple instances without decrease in computing power.

AMR class definition

Third-generation cephalosporin (3GC) resistance was defined by resistance to ceftriaxone; carbapenem resistance (CR) was defined by resistance to ertapenem. Molecular genotypes (ESBL, plasmid-mediated AmpC beta-lactamase (pAmpC), *Klebsiella pneumoniae* carbapenemase (KPC), metallo- β -lactamase [MBL], OXA-48-like) and WGS-AST results were combined to define AMR classes, as outlined in the decision tree in Figure 1. Briefly, 3GC-resistant, carbapenem-susceptible organisms were delineated by pAmpC and ESBL markers. CP-CREs were characterized by KPC, MBL, and OXA-48-like marker genes. Non-CP CREs were defined by resistance to ertapenem by WGS-AST and the absence of carbapenemase genes. WGS-AST performance and concordance of AMR classes were finally compared between platforms.

Statistical evaluation

Categorical agreement (CA), very major error (VME), and major error (ME) were assessed for WGS-AST. Resistance (R) was considered a positive outcome and susceptibility (S) a negative



outcome. WGS-AST S/R results were classified either as true positives (TP), false positives (FP), true negatives (TN), or false negatives (FN), in comparison to BMD-AST results. Intrinsically resistant (IR) isolates, and isolates with an intermediate (I) phenotype, were treated as resistant. Categorical agreement (accuracy) of S/R designations was compared between WGS-AST and BMD-AST; very major errors (VME) are false-susceptible results (FNR, false-negative rate) by WGS-AST, and major errors (ME) are false-resistant results (FPR, false-positive rate) by WGS-AST. Species-antimicrobial pairs with less than 10 samples were removed: ampicillin (n: 9), ampicillin-sulbactam (n: 9), and minocycline (n: 6). Data handling and analysis was carried out in a Python environment using the Pandas and NumPy libraries. Confusion matrices were generated using the scikit-learn library and statistical tests were calculated using the SciPy library.

Results

Phenotypic antimicrobial susceptibility testing results

Phenotypic antimicrobial susceptibility testing results are summarized in [Supplementary Table 2](#). Of the 181 isolates, 135

were not susceptible (NS; i.e., intermediate or resistant) to at least one carbapenem (ertapenem: 72.9%, doripenem: 55.6%; imipenem: 61.1%; meropenem: 54.5%); 144 isolates were NS to 3GCs, while 136 isolates were NS to fourth-generation cephalosporins. Non-susceptibility to newer β -lactam/ β -lactamase inhibitor combinations was low (imipenem-relebactam: 10.5%, meropenem-vaborbactam: 10.6%, ceftazidime-avibactam: 6.8%). NS to non- β -lactam agents ranged among isolates as follows: 8.3–51.7% for aminoglycosides, 64.1–88.9% for fluoroquinolones, 38.0% for tetracycline, and 49.4% for sulfamethoxazole-trimethoprim.

Sequencing and assembly

ONT assemblies had lower BUSCO completeness scores of $42.59\% \pm 23.59\%$ compared to $98.57\% \pm 0.66\%$ on the Illumina platform. ONT assemblies were less fragmented as determined by BUSCO ($30.48\% \pm 12.65\%$) compared to Illumina assemblies ($0.02\% \pm 0.15\%$); BUSCO missing scores were worse for ONT ($26.92\% \pm 11.36\%$) than for Illumina data ($1.40\% \pm 0.59\%$), highlighting the higher per-base error rates on the ONT platform. All comparisons were significant (two-tailed paired *t*-test, $p < 0.0001$).

Identification

The genus of all 181 isolates (100%) was correctly identified for ONT and Illumina isolates. Species-level concordance was 97.24%; 5 *Enterobacter* isolates were identified on the ONT assemblies as *E. hormaechei* and on the Illumina assemblies as *E. cloacae*. The discrepant isolates were confirmed belonging to the *E. hormaechei* species using the average nucleotide identity (ANI) method at a cutoff of 95% (Supplementary Table 3; Jain et al., 2018).

AMR markers

The number of uniquely identified AMR markers was on average 19.9% higher on the ONT dataset than on the Illumina dataset (Supplementary Table 4). The higher number of retrieved markers on ONT assemblies was due to the lower identity threshold used by DIAMOND v1.0.11 for sequence alignment of AMR markers to ARESdb. The threshold was optimized for ONT data to retrieve all relevant markers needed to assign AMR classes. Due to the higher fragmentation of ONT assemblies and lower base calling fidelity, incorrect alleles were assigned more often on ONT assemblies (two-tailed paired *t*-test, $p < 0.001$).

With the Illumina data as the reference, the sensitivity of AMR markers in the ONT dataset was investigated. The sensitivity of carbapenemases, ESBLs, and pAmpCs was slightly lower on ONT compared to Illumina at 98, 95, and 70%, respectively, with specificities of 100, 70, and 91%, respectively. Among the carbapenemases, accuracy was 98, 100, and 99% for KPCs, MBLs, and OXA-48-like AMR marker genes, respectively. The sensitivity was lowest for pAmpCs (70%), and specificity was lowest for ESBLs (70%). Across all 6 marker classes, the specificity of ONT was 100% for carbapenemases, KPCs, metallo- β -lactamases, and OXA-48-like markers. The lowered sensitivity of ONT led to a high false-negative rate (FNR) in some cases, e.g., 30% for pAmpCs (Table 1).

WGS-AST

Overall, ONT data performed slightly worse than Illumina sequencing data regarding WGS-AST (Table 2), but the difference was not significant (χ^2 test on the vectorized confusion matrices, $p > 0.05$). Categorical agreement (CA) was 88% for ONT and 90% for Illumina, overall ME was higher on the ONT platform with 13% compared to 11% on the Illumina platform, and overall VME was also slightly higher at 11% on ONT compared to 10% on the Illumina platform. On the ONT platform, 25 species-antimicrobial models had a CA between 90 and 100%, 23 between 80 and 90%, and 12 below 80%; on the Illumina platform, 34, 18, and 8 species-antimicrobial models had CAs of 90–100%, 80–90%, and below 80%, respectively. Grouped by antimicrobial, on the ONT platform, 10 species-antimicrobial models had a CA between

90–100%, 11 between 80–90%, and 1 below 80%; on the Illumina platform, 13, 9, and zero species-antimicrobial models had CAs of 90–100%, 80–90%, and below 80%, respectively (Table 3; Supplementary Table 5). Grouped by organism, CA was highest for *Klebsiella* spp. with 89% on the ONT platform and 90% on the Illumina platform. *Enterobacter* reached a CA of 83% on the ONT platform and *Escherichia* of 82% on the ONT platform. *Escherichia* reached a CA of 87% and *Enterobacter* a CA of 85% on the Illumina platform (Supplementary Table 6). Here, two effects were observed: *Escherichia* and *Enterobacter* models had lower CAs than *Klebsiella* models and ONT had a lower CA compared to Illumina as described above. While *Klebsiella* had the highest CA, it also had a higher VME of 11% on both platforms than *Enterobacter* on ONT (6%) and Illumina (8%) platforms, and *Escherichia* on the Illumina platform (5%) but not on the ONT platform (18%). MEs of *Klebsiella* (12% vs. 9%) and *Enterobacter* models (27% vs. 22%) were higher on the ONT platform compared to the Illumina, while *Escherichia* models (19% vs. 32%) had lower MEs on the ONT platform compared to Illumina.

The difference in CA between individual antimicrobials spanned from a 7% lower CA to a 4% higher CA for ONT compared to Illumina sequencing data. Lower performance on ONT sequencing data was, e.g., observed in ciprofloxacin (−7%), levofloxacin (−6%), piperacillin-tazobactam (−7%), and gentamicin (−5%), while better performance of ONT was observed in, e.g., aztreonam (+2%), ticarcillin-clavulanic acid (+3%), amikacin (+3%), and ceftazidime (+4%; Table 3).

WGS-AST models for AMR classes

Categorical agreement for ceftriaxone, which was used as a proxy for the prediction of resistance to third-generation cephalosporins, was 96% on both platforms, with ONT performing better with a lower VME (2% vs. 4%) but higher ME (7% vs. 3%).

Ertapenem—used to indicate carbapenem resistance—reached a CA of 91% on the ONT platform and 94% on the Illumina platform (χ^2 test, $p < 0.001$ as tested on the vectorized confusion matrices). VMEs were comparable on both platforms (7% on ONT and 6% on Illumina), but MEs were higher on the ONT platform (16% vs. 4%; Table 3; Supplementary Table 7). On the ONT platform, MEs of the *E. cloacae* and VMEs of the *E. coli* model, and on the Illumina platform, MEs in the *Enterobacter* and *E. coli* subsets model were noticeable. On both platforms, the *Klebsiella* model performs with a better trade-off between ME and VME (Supplementary Table 5).

Ertapenem WGS-AST models reached 95% CA for CP-CREs (defined by ertapenem BMD-AST and the presence of a carbapenemase gene in the Illumina genome assembly) for data acquired on the ONT platform and 100% for data acquired on the Illumina platform, and 88% CA in the non-CP CRE dataset for ONT and 76% for Illumina platforms. Among ertapenem-susceptible isolates, CA on ONT data was lower (84%) compared to Illumina (96%), due to 8 samples being flagged positive. VMEs

TABLE 1 Performance metrics of the AMR marker identification on ONT data, which is compared to the markers identified on the Illumina data (reference).

Marker class	Accuracy	Sensitivity	Specificity	FNR	FPR	TP	FP	FN	TN	<i>n</i>
ESBL	78%	95%	70%	5%	30%	79	29	4	69	181
pAmpC	96%	70%	91%	30%	9%	16	0	7	158	181
CP	99%	98%	100%	2%	0%	99	0	2	80	181
KPC	98%	96%	100%	4%	0%	74	0	3	104	181
MBL	100%	100%	100%	0%	0%	15	0	0	166	181
OXA-48	99%	93%	100%	7%	0%	13	0	1	167	181

CA: Categorical agreement, FNR: False-negative rate, FPR: False-positive rate, TP: True positive, FP: False positive, FN: False negative, TN: True negative, and *n*: number of evaluated samples.

TABLE 2 Overall performance of the WGS-AST models across all antimicrobials, broken down by sequencing platform.

Platform	CA	VME	ME	TP	FP	FN	TN	<i>n</i>
Illumina	90%	10%	11%	1,646	161	178	1,315	3,300
ONT	88%	11%	13%	1,619	194	205	1,282	3,300

CA: Categorical agreement, VME: Very major error, ME: Major error, TP: True positive, FP: False positive, FN: False negative, TN: True negative, and *n*: number of evaluated species-antimicrobial pairs.

of ertapenem among non-CP CREs were 12% on ONT and 24% on Illumina (Supplementary Table 8).

Combination of WGS-AST and molecular markers for AMR classes

We combined WGS-AST with the identified AMR genotypes, as outlined in Figure 1 in a decision workflow (Table 4; Supplementary Table 9). An isolate was first categorized as carbapenem-resistant or carbapenem-susceptible. Carbapenem-susceptible isolates were further classified as resistant or susceptible to 3GCs based on the ceftriaxone resistance. Error rates, as discussed in the previous section, apply to both carbapenem and 3GC resistance results. The accuracy between AMR classes predicted by ONT data and AMR classes predicted by Illumina data was 87.8% (Figure 2). For the distinction of CP CREs and non-CP CREs from WGS-AST data, data acquired on the Illumina platform was used for the detection of CP genes as the reference. Mainly, non-CP CREs were overpredicted on ONT data due to missed carbapenemase gene detection in the dataset while resistance to ertapenem was correctly predicted.

Discussion

Overall, this work demonstrated that the combination of WGS-AST and molecular identification of AMR markers through WGS can classify isolates into clinically relevant AMR classes regardless of sequencing platform. Second, WGS-AST on AREScloud using ONT sequencing data was comparable to Illumina sequencing data. Overall categorical agreement across all antimicrobials was 88% (ONT) and 90% (Illumina), 96% CA by

both platforms for third-generation cephalosporin resistance, and 91% (ONT) and 94% (Illumina) for carbapenem resistance. VMEs (11% vs. 10%) and MEs (13% vs. 11%) were slightly higher on the ONT platform. The accuracy (89% vs. 90%) of WGS-AST predictions was slightly lower on ONT assemblies compared to Illumina. This difference was likely due to significantly higher fragmentation and higher base calling error rate of the ONT data, translating to an overall difference of 2% CA. Similarly, for carbapenem resistance, there were also differences in VMEs and MEs observed between the two sequencing platforms. The VMEs and MEs on the ONT platform were 7% and 16%, whereas the VMEs and MEs of the Illumina platform were 6% and 4%. The higher ME with ONT might have been due to the higher per-base error rate leading to the interpretation of random sequencing errors as SNPs in the *ompK* genes; however, since the data acquisition, ONT released new base calling models and kit chemistries allowing more complete, less fragmented assemblies, with a lower error rate, potentially lowering the ME had the analysis been repeated with the updated chemistries and base caller.

A total of 22 antimicrobials were tested for the *Enterobacteriaceae*: All antimicrobials for which WGS-AST results were predicted reached CAs above 80% and up to 97% across all species on both sequencing platforms compared to BMD-AST, except for imipenem at 79% on the ONT platform. Of the 60 individual species-antimicrobial WGS-AST models, 8 (Illumina) and 11 (ONT) models had CAs below 80%. All species-antimicrobial pairs below 80% belonged to the *Escherichia* (*n*: 14) or *Enterobacter* (*n*: 13) genera, where sample sizes were significantly lower than for, e.g., *K. pneumoniae*, hampering statistical evaluation. Thirty-three and 25 (out of 60) species-antimicrobial pairs reached CAs above 90% on the Illumina and ONT platforms, respectively (13 and 10 for *Klebsiella* on Illumina and ONT, respectively). Despite WGS-AST models of

TABLE 3 Comparison of the performance metrics of WGS-AST on Illumina and ONT assemblies, by antimicrobial.

Antimicrobial	CA		VME		ME		n
	Illumina	ONT	Illumina	ONT	Illumina	ONT	
Amikacin	80%	83%	23%	20%	19%	16%	163
Aztreonam	91%	92%	9%	6%	12%	14%	180
Cefazolin	84%	84%	22%	20%	0%	6%	68
Cefepime	85%	83%	16%	21%	12%	7%	173
Cefotaxime	96%	95%	4%	4%	0%	10%	107
Ceftazidime	89%	93%	10%	5%	14%	14%	181
Ceftazidime-avibactam	97%	96%	6%	35%	3%	1%	175
Ceftriaxone	96%	96%	4%	2%	3%	7%	79
Ciprofloxacin	95%	88%	3%	12%	10%	12%	181
Doripenem	84%	83%	22%	22%	8%	11%	164
Ertapenem	94%	91%	7%	6%	4%	16%	181
Gentamicin	93%	88%	8%	8%	6%	15%	171
Imipenem	84%	79%	6%	5%	31%	44%	94
Imipenem-relebactam	97%	97%	21%	21%	1%	1%	91
Levofloxacin	90%	85%	9%	14%	11%	18%	181
Meropenem	84%	80%	3%	5%	31%	37%	90
Meropenem-vaborbactam	97%	95%	21%	36%	1%	1%	91
Piperacillin-tazobactam	90%	84%	11%	21%	7%	2%	168
Sulfamethoxazole-trimethoprim	80%	80%	11%	8%	30%	33%	167
Tetracycline	81%	82%	30%	30%	12%	10%	79
Ticarcillin-clavulanic acid	92%	95%	9%	5%	0%	0%	101
Tobramycin	94%	92%	4%	6%	7%	10%	173

CA: Categorical agreement, VME: Very major error, ME: Major error, n: number of evaluated species-antimicrobial pairs.

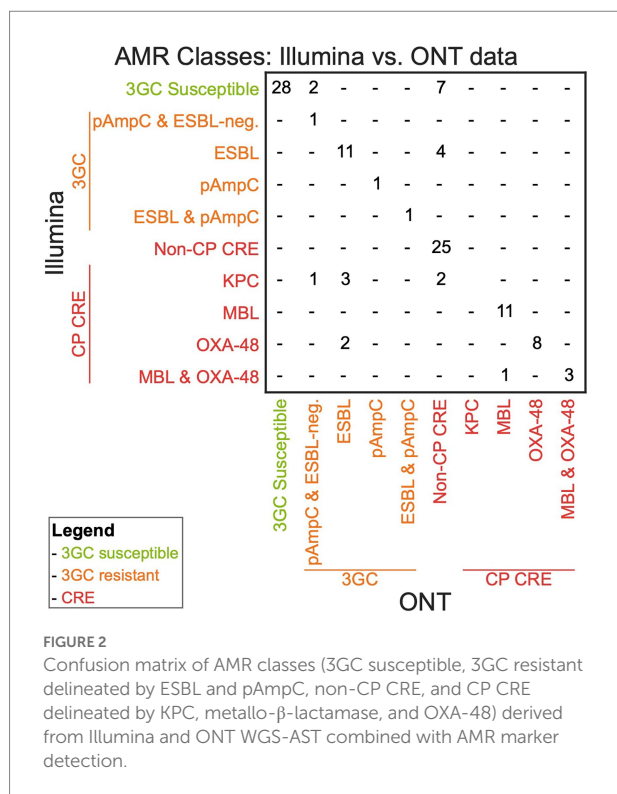
TABLE 4 Breakdown of the isolates by AMR class and sequencing platform at the decision gateways.

Group Level 1			Group Level 2			Group Level 3		
Illumina		ONT	Illumina		ONT	Illumina		ONT
Carbapenem	55	50	3GC Susceptible	37	28	3GC-susceptible	37	28
Susceptible			3GC Resistant	18	22	Not ESBL or pAmpC	1	4
						ESBL	15	16
						pAmpC	1	1
						ESBL & pAmpC	1	1
Carbapenem	126	131	Non-CP CRE	25	38	Non-CP CRE	25	38
Resistant			CP CRE	101	93	KPC	76	70
						MBL	11	12
						OXA-48	10	8
						OXA-48 & MBL	4	3

antimicrobials reaching the 90% CA threshold as defined by the FDA for the clearance of phenotypic AST devices, MEs and VMEs were outside of acceptable limits of 3% and 1.5%, respectively (Food and Drug Administration, 2009). On-going improvements of sequencing chemistries and further refinement of WGS-AST models as more data become available will likely reduce error rates to acceptable limits within the next years.

A possible mitigation strategy to deal with high error rates could be the introduction of areas of technical

uncertainty (ATU), following the lead of EUCAST for phenotypic AST (The European Committee on Antimicrobial Susceptibility Testing, 2019). ATUs would help manage methodological and technical variability as well as variations in interpretation. For example, individual WGS-AST results could be annotated with a confidence level; results within an ATU could then be flagged or excluded from the report if other antimicrobials are available for treatment depending on the clinical scenario.



Regarding AMR marker retrieval, ONT was comparable to Illumina regarding accuracy (98–100%), sensitivity (93–100%), and specificity (100%) for carbapenemases, KPCs, MBLs, and OXA-48-like β -lactamases. For pAmpCs, the sensitivity was lower at 70% due to a high number of false negatives. The low recovery of pAmpCs can possibly be explained by using a ligation library preparation kit, which has a lower recovery of plasmids compared to a rapid detection kit (Wick et al., 2021). For ESBLs, specificity was lower at 70%. Several β -lactamases were misclassified on the ONT platform, such as the broad-spectrum β -lactamase TEM-1 as ESBL TEM-10, or the broad-spectrum β -lactamases SHV-1, SHV-11, or inhibitor-resistant broad-spectrum β -lactamase SHV-26 as ESBLs SHV-12 or SHV-152. The lower fidelity on the allele level for ONT—due to lower quality sequencing data—has been previously observed (Tamma et al., 2019).

The proposed workflow to predict relevant clinical phenotypes (Figure 1) had multiple decision gateways. Early misclassification propagated downstream through the pipeline. More specifically, if an isolate was incorrectly identified on the first grouping level, this error propagated to the final classification. As an example, the phenotypically third-generation cephalosporin-resistant, ESBL-positive isolate CRE596KLPN is classified as such using Illumina sequencing data, but on the ONT platform, ertapenem resistance is overcalled in the absence of a carbapenemase gene. Hence, this isolate is classified as non-CP CRE using ONT due to carbapenem classification occurring upstream of third-generation cephalosporin classification. Nevertheless, the ESBL genes are also identified in this isolate on the ONT platform. As a mitigation strategy, we placed the WGS-AST models at the beginning of the

decision workflow, since we have shown that those models have a higher error tolerance than simple AMR marker detection. This allows generally to compensate for the lower sensitivity of pAmpCs and lower specificity of ESBLs on ONT sequencing data, which is more often the case than vice versa. Co-detection of AMR genes together with WGS-AST is important for tailoring therapy as this discrimination enables selection of the correct downstream treatment more precisely.

The higher performance of ONT compared to Illumina was observed for antimicrobials where identification of markers or marker classes is sufficient for resistance calling. This was, e.g., the case for ceftriaxone and amikacin, where the presence of CTX-M class markers can be sufficient to predict ceftriaxone resistance, or amikacin which is dependent on ribosomal RNA methyltransferases such as *armA* or *rmtA* homologs (Bonnet, 2004; Galimand et al., 2012; Liakopoulos et al., 2016). The lower performance of individual antimicrobials on the ONT platform was more often observed when the resolution of SNPs is crucial for resistance calling, as it is, e.g., the case for *gyrA* and *parC* mutations for fluoroquinolones, or mutations in efflux pumps for piperacillin-tazobactam, or when detection of multiple aminoglycoside-modifying enzymes is required, as it is the case for gentamicin resistance (Chen et al., 2003; Fu et al., 2013).

There are several limitations to this study. First, some species were represented by a relatively small number (*E. coli* [n: 14], *E. cloacae* [n: 7]) of isolates within the 181 clinical isolates. As some AMR markers are specific to individual bacterial species, this work needs to be repeated on a larger sample set. Second, all isolates came from a single site (geographic region) resulting in likely over- or under-representation of individual AMR markers based on the genetic relatedness of isolates. This work is currently in the proof-of-concept stage and should be expanded in the future to include geographically diverse isolates. Third, the Phred quality score of ONT is still inferior to Illumina, although it has increased by an order of magnitude since the beginning of the study, both by new sequencing chemistries and due to improved bioinformatics. WGS-AST models and related bioinformatic pipelines need to be adapted accordingly.

With the availability of sequencing platforms in hospital practice, risk identification based on AMR classes can enhance patient management. For the routine application of WGS-AST, clinical microbiology laboratories need to implement automated approaches for DNA extraction, library preparation, and sequencing, such as fully automated clinical microbial cultivation and identification systems. Once established, methodological control of WGS-AST might be easier to maintain compared to BMD-AST (e.g., daily internal controls) and skills/requirements for routine use might be lower. Additionally, WGS-AST allows drawing retrospective conclusions about antimicrobials that were not available at the time of sequencing, for example, for AST surveillance studies.

The performance gap for WGS-AST between ONT and Illumina has narrowed over the last years, especially with the upcoming of newer base calling models, flow cells, and sequencing chemistries of

higher accuracy, which routinely generate reads at a modal raw read accuracy at a Phred score of 20, which is an improvement from an accuracy of 90–95% ($10 < \text{Phred} < 13$) to 99% ($\text{Phred} = 20$). However, Illumina sequencing reads typically achieve Phred scores of above 30 ($> 99.9\%$ accuracy). Having assemblies of good quality is important for WGS-AST predictions, genotyping, and identification of AMR classes if single-nucleotide polymorphisms are involved. While significant differences still exist, especially for the detection of AMR genes, our results highlight the potential of the ONT platform for consideration in clinical microbiology laboratories if combined with robust bioinformatics methodology. Combining AMR genotypes and WGS-AST predictions can help guide effective therapeutic management decisions.

Data availability statement

The data presented in the study are deposited in the NCBI Sequence Read Archive (SRA) repository, accession number PRJNA825705.

Author contributions

RC: formal analysis, investigation, data curation, software, validation, writing—original draft, writing—reviewing and editing, and visualization. YB, SL, and EJ: data curation, formal analysis, and writing—reviewing and editing. PM: data curation, software, investigation, and validation. SB: software, methodology, data curation, writing—reviewing and editing, supervision, and project administration. PT: conceptualization and writing—reviewing and editing. PS: conceptualization, methodology, writing—reviewing and editing, and supervision. All authors contributed to the article and approved the submitted version.

Conflict of interest

RC, PM, and SB are employees of Ares Genetics GmbH. PS has received grants and personal fees from Accelerate Diagnostics, OpGen Inc., and BD Diagnostics; grants from bioMérieux, Inc., Affinity Biosensors, and Hardy Diagnostics; and personal fees from Roche Diagnostics, GeneCapture, and Shionogi Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Andrews, S. (2016). FastQC: A quality control tool for high throughput sequence data [internet]. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed July 18, 2021).
- Ares Genetics GmbH (2021). AREScLOUD [internet]. Available at: <https://ares-genetics.cloud/> (Accessed March 11, 2021).

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.973605/full#supplementary-material>

SUPPLEMENTARY TABLE 1

List of the species-antimicrobial pairs used for the WGS-AST models.

SUPPLEMENTARY TABLE 2

Summary of BMD-AST results. n: number of evaluated isolates.

SUPPLEMENTARY TABLE 3

Average Nucleotide Identity (ANI) analysis applied to the samples with discrepant species identification results.

SUPPLEMENTARY TABLE 4

List of all AMR markers identified in the Illumina and ONT assemblies.

SUPPLEMENTARY TABLE 5

Performance metrics for all species-antimicrobial pairs by sequencing platform. CA: categorical agreement, VME: very major error, ME: major error, TP: true positive, FP: false positive, FN: false negative, TN: true negative, n: number of evaluated species-antimicrobial pairs.

SUPPLEMENTARY TABLE 6

Performance metrics of WGS-AST on Illumina and ONT assemblies, broken down by taxon. CA: categorical agreement, VME: very major error, ME: major error, TP: true positive, FP: false positive, FN: false negative, TN: true negative, n: number of evaluated species-antimicrobial pairs.

SUPPLEMENTARY TABLE 7

Performance metrics of WGS-AST on Illumina and ONT assemblies, broken down by antimicrobial. CA: categorical agreement, VME: very major error, ME: major error, TP: true positive, FP: false positive, FN: false negative, TN: true negative, n: number of evaluated species-antimicrobial pairs.

SUPPLEMENTARY TABLE 8

WGS-AST performance metrics of ertapenem broken down by BMD-AST phenotype for carbapenem. CA: categorical agreement, VME: very major error, ME: major error, TP: true positive, FP: false positive, FN: false negative, TN: true negative, n: number of evaluated species-antimicrobial pairs.

SUPPLEMENTARY TABLE 9

List of all isolates from the study with their species identification, relevant identified markers, WGS-AST for ertapenem and ceftriaxone, and the resulting predicted phenotype group.

- Aronesty, E. (2011). ea-utils: command-line tools for processing biological sequencing data [internet]. Available at: <https://github.com/ExpressionAnalysis/ea-utils> (Accessed July 18, 2021).

- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

- Bonnet, R. (2004). Growing Group of Extended-Spectrum β -lactamases: the CTX-M enzymes. *Antimicrob. Agents Chemother.* 48, 1–14. doi: 10.1128/AAC.48.1.1–14.2004
- Broad institute. (2021). Picard tools [internet]. Available at: <https://broadinstitute.github.io/picard/> (Accessed July 18, 2021).
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Charalampous, T., Kay, G. L., Richardson, H., Aydin, A., Baldan, R., Jeanes, C., et al. (2019). Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat. Biotechnol.* 37, 783–792. doi: 10.1038/s41587-019-0156-5
- Chen, F. J., Lauderdale, T. L., Ho, M., and Lo, H. J. (2003). The roles of mutations in *gyrA*, *parC*, and *ompK35* in Fluoroquinolone resistance in *Klebsiella pneumoniae*. *Microb. Drug Resist.* 9, 265–271. doi: 10.1089/10762903322286472
- CLSI (2018). *M100: Performance Standards for Antimicrobial Susceptibility Testing*. 28th Edn. Wayne, PA, United States: Clinical and Laboratory Standards Institute.
- CLSI (2021). *M100: Performance Standards for Antimicrobial Susceptibility Testing*. 31st Edn. Wayne, PA, United States: Clinical and Laboratory Standards Institute.
- de Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149
- Doumith, M., Ellington, M. J., Livermore, D. M., and Woodford, N. (2009). Molecular mechanisms disrupting porin expression in ertapenem-resistant *Klebsiella* and *Enterobacter* spp. clinical isolates from the UK. *J. Antimicrob. Chemother.* 63, 659–667. doi: 10.1093/jac/dkp029
- Ferreira, I., Beisken, S., Luefing, L., Weinmaier, T., Klein, M., Bacher, J., et al. (2020). Species identification and antibiotic resistance prediction by analysis of whole-genome sequence data by use of ARESdb: an analysis of isolates from the Unyvero lower respiratory tract infection trial. *J. Clin. Microbiol.* 58:20. doi: 10.1128/JCM.00273-20
- Food and Drug Administration. (2009). Antimicrobial susceptibility test (AST) systems - class II special controls guidance for industry and FDA [internet]. Available at: <https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/antimicrobial-susceptibility-test-ast-systems-class-ii-special-controls-guidance-industry-and-fda> (Accessed June 9, 2022).
- Fu, Y., Zhang, W., Wang, H., Zhao, S., Chen, Y., Meng, F., et al. (2013). Specific patterns of *gyrA* mutations determine the resistance difference to ciprofloxacin and levofloxacin in *Klebsiella pneumoniae* and *Escherichia coli*. *BMC Infect. Dis.* 13:8. doi: 10.1186/1471-2334-13-8
- Galimand, M., Courvalin, P., and Lambert, T. (2012). RmtF, a new member of the aminoglycoside resistance 16S rRNA N7 G1405 Methyltransferase family. *Antimicrob. Agents Chemother.* 56, 3960–3962. doi: 10.1128/AAC.00660-12
- Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., et al. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* 29:99. doi: 10.3389/fmicb.2016.01374
- Hall, M. (2021). Rasusa: randomly subsample sequencing reads to a specified coverage [internet]. Available at: <https://github.com/mbhall88/rasusa> (Accessed September 1, 2021).
- Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., and Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nat. Methods* 12, 351–356. doi: 10.1038/nmeth.3290
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: 1303.3997 (q-bio.GN).
- Li, H. (2016). Seqtk: toolkit for processing sequences in FASTA/Q formats [internet]. Available at: <https://github.com/lh3/seqtk> (Accessed July 18, 2021).
- Liakopoulos, A., Mevius, D., and Ceccarelli, D. (2016). A review of SHV extended-Spectrum β -lactamases: neglected yet ubiquitous. *Front. Microbiol.* 7:374. doi: 10.3389/fmicb.2016.01374
- Livermore, D. M. (2004). The need for new antibiotics. *Clin. Microbiol. Infect.* 10, 1–9. doi: 10.1111/j.1465-0691.2004.1004.x
- Lüftinger, L., Ferreira, I., Frank, B. J. H., Beisken, S., Weinberger, J., von Haeseler, A., et al. (2021a). Predictive antibiotic susceptibility testing by next-generation sequencing for Periprosthetic joint infections: potential and limitations. *Biomedicine* 9:910. doi: 10.3390/biomedicine9080910
- Lüftinger, L., Májek, P., Beisken, S., Rattei, T., and Posch, A. E. (2021b). Learning From limited data: towards best practice techniques for antimicrobial resistance prediction from whole genome sequencing data. *Front. Cell. Infect. Microbiol.* 11:610348. doi: 10.3389/fcimb.2021.610348
- Májek, P., Lüftinger, L., Beisken, S., Rattei, T., and Materna, A. (2021). Genome-wide mutation scoring for machine-learning-based antimicrobial resistance prediction. *Int. J. Mol. Sci.* 22:13049. doi: 10.3390/ijms222313049
- Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. (2021). BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* 38, 4647–4654. doi: 10.1093/molbev/msab199
- Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., and Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34, 1142–1150. doi: 10.1093/bioinformatics/bty266
- Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes De novo assembler. *Curr. Protoc. Bioinformatics* 70:e102. doi: 10.1002/cpbi.102
- Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., et al. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530, 228–232. doi: 10.1038/nature16996
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rang, F. J., Kloosterman, W. P., and de Ridder, J. (2018). From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* 19:90. doi: 10.1186/s13059-018-1462-9
- Rochford, C., Sridhar, D., Woods, N., Saleh, Z., Hartenstein, L., Ahlawat, H., et al. (2018). Global governance of antimicrobial resistance. *Lancet* 391, 1976–1978. doi: 10.1016/S0140-6736(18)31117-6
- Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., et al. (2017). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J. Antimicrob. Chemother.* 72, 104–114. doi: 10.1093/jac/dkw397
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Seemann, T. (2018). Barrnap: BAsic rapid ribosomal RNA predictor [internet]. Available at: <https://github.com/tseemann/barrnap> (Accessed July 18, 2021).
- Sugawara, E., Kojima, S., and Nikaido, H. (2016). *Klebsiella pneumoniae* Major Porins OmpK35 and OmpK36 allow more efficient diffusion of β -lactams than their *Escherichia coli* homologs OmpF and OmpC. *J. Bacteriol.* 198, 3200–3208. doi: 10.1128/JB.00590-16
- Tamma, P. D., Fan, Y., Bergman, Y., Perte, G., Kazmi, A. Q., Lewis, S., et al. (2019). Applying rapid whole-genome sequencing to predict phenotypic antimicrobial susceptibility testing results among Carbapenem-resistant *Klebsiella pneumoniae* clinical isolates. *Antimicrob. Agents Chemother.* 63:18. doi: 10.1128/AAC.01923-18
- The European Committee on Antimicrobial Susceptibility Testing (2019). Area of technical uncertainty (ATU) in antimicrobial susceptibility testing [internet]. Available at: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/ATU/Area_of_Technical_Uncertainty_-_guidance_2019.pdf (Accessed June 9, 2022).
- Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 27, 737–746. doi: 10.1101/gr.214270.116
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y., and Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* 39, 1348–1365. doi: 10.1038/s41587-021-01108-x
- Wick, R. R., Judd, L. M., Wyres, K. L., and Holt, K. E. (2021). Recovery of small plasmid sequences via Oxford Nanopore sequencing. *Microbial Genomics* 7:631. doi: 10.1099/mgen.0.000631
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- World Health Organization (2018). Global tuberculosis report. Available at: <https://apps.who.int/iris/handle/10665/274453> (Accessed June 9, 2022).
- World Health Organization (2020). GLASS whole-genome sequencing for surveillance of antimicrobial resistance. Available at: <https://www.who.int/publications/i/item/9789240011007> (Accessed June 9, 2022).



OPEN ACCESS

EDITED BY
Nagendran Tharmalingam,
Rhode Island Hospital, United States

REVIEWED BY
Charuta Agashe,
Icahn Institute for Genomics
and Multiscale Biology, United States
Debasish Paul,
National Institutes of Health (NIH),
United States

*CORRESPONDENCE
Fabio Arena
fabio.arena@unifg.it

SPECIALTY SECTION
This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 30 June 2022
ACCEPTED 25 July 2022
PUBLISHED 15 August 2022

CITATION
Arena F, Menchinelli G, Di Pilato V,
Torelli R, Antonelli A,
Henrici De Angelis L, Coppi M,
Sanguinetti M and Rossolini GM (2022)
Resistance and virulence features
of hypermucoviscous *Klebsiella*
pneumoniae from bloodstream
infections: Results of a nationwide
Italian surveillance study.
Front. Microbiol. 13:983294.
doi: 10.3389/fmicb.2022.983294

COPYRIGHT
© 2022 Arena, Menchinelli, Di Pilato,
Torelli, Antonelli, Henrici De Angelis,
Coppi, Sanguinetti and Rossolini. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Resistance and virulence features of hypermucoviscous *Klebsiella pneumoniae* from bloodstream infections: Results of a nationwide Italian surveillance study

Fabio Arena^{1,2,3*}, Giulia Menchinelli⁴, Vincenzo Di Pilato⁵,
Riccardo Torelli⁴, Alberto Antonelli^{6,7},
Lucia Henrici De Angelis⁸, Marco Coppi^{6,7},
Maurizio Sanguinetti^{4,9} and Gian Maria Rossolini^{3,6,7} for the
Klebsiella pneumoniae Hypermucoviscous Italian Network
(KHIN) of the Associazione Microbiologi Clinici Italiani (AMCLI)

¹Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy,

²Microbiology and Virology Unit, University Hospital "Riuniti," Foggia, Italy, ³Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS) Don Carlo Gnocchi ONLUS, Florence, Italy, ⁴Dipartimento di Scienze di Laboratorio e Infettivologiche, Fondazione Policlinico Universitario A. Gemelli Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS), Rome, Italy, ⁵Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genoa, Genoa, Italy, ⁶Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy, ⁷Microbiology and Virology Unit, Florence Careggi University Hospital, Florence, Italy, ⁸Department of Medical Biotechnologies, University of Siena, Siena, Italy, ⁹Dipartimento di Scienze Biotechnologiche di Base, Cliniche Intensivologiche e Perioperatorie, Università Cattolica del Sacro Cuore, Rome, Italy

Among Enterobacterales, *Klebsiella pneumoniae* (Kp) is one of the major opportunistic pathogens causing hospital-acquired infections. The most problematic phenomenon linked to Kp is related to the dissemination of multi-drug resistant (MDR) clones producing carbapenem-hydrolyzing enzymes, representing a clinical and public health threat at a global scale. Over the past decades, high-risk MDR clones (e.g., ST512, ST307, ST101 producing *bla*_{KPC}-type carbapenemases) have become endemic in several countries, including Italy. Concurrently, the spread of highly virulent Kp lineages (e.g., ST23, ST86) able to cause severe, community-acquired, pyogenic infections with metastatic dissemination in immunocompetent subjects has started to be documented. These clones, designated as hypervirulent Kp (hvKp), produce an extensive array of virulence factors and are highly virulent in previously validated animal models. While the prevalence and distribution of MDR Kp has been previously assessed at local and national level knowledge about dissemination of hvKp remains scarce. In this work, we studied the phenotypic and genotypic features of hypermucoviscous (HMV, as possible marker of increased virulence) Kp isolates from bloodstream infections (BSI), obtained in 2016–17 from 43 Italian Laboratories. Antimicrobial susceptibility

testing, whole genome sequencing and the use of two animal models (*G. mellonella* and murine) were employed to characterize collected isolates. Over 1502 BSI recorded in the study period, a total of 19 Kp were selected for further investigation based on their HMV phenotype. Results showed that hvKp isolates (ST5, ST8, ST11, ST25) are circulating in Italy, although with a low prevalence and in absence of a clonal expansion; convergence of virulence (yersiniabactin and/or salmochelin, aerobactin, regulators of mucoid phenotype) and antimicrobial-resistance (extended-spectrum beta-lactamases) features was observed in some cases. Conventional MDR Kp clones (ST307, ST512) may exhibit an HMV phenotype, but with a low virulence potential in the animal models. To the best of our knowledge, this work represents the first systematic survey on HMV and hvKp in Italy, employing a functional characterization of collected isolates. Future surveillance programs are warranted to monitor the threatening convergence of virulence and resistance among MDR Kp and the spread of hvKp.

KEYWORDS

genome sequencing, clones, carbapenemases, virulence factor, multi-drug resistant

Introduction

At the beginning of the 80's, bacteria of genus *Klebsiella* had mostly been studied as a cause of community-acquired pneumonia, occurring particularly in chronic alcoholics (Carpenter, 1990). More recently, *Klebsiella* spp., and especially *K. pneumoniae* (Kp) have been recognized as a major public health threat due to their ability to accumulate antimicrobial-resistance determinants and cause severe infections in the healthcare setting. Kp is currently considered the prototype of opportunistic, hospital-acquired, multi-drug resistant (MDR) pathogen (Munoz-Price et al., 2013; Tacconelli et al., 2018).

Some Kp lineages, called hypervirulent *K. pneumoniae* (hvKp), are associated with a more virulent behavior (Shon et al., 2013). hvKp is an evolving pathotype that is more virulent than classical Kp due to the combination of several features including: K1, K2 or K5 capsule polysaccharide; the presence of horizontally acquired virulence factors encoding the siderophores aerobactin (Iuc) and salmochelin (Iro), and the genotoxin colibactin (Clb); a hypermucoid phenotype (often associated with the presence of regulators of the mucoid phenotype (namely Rmp) (Russo and Marr, 2019; Zhu et al., 2021). Although incompletely predictive of a virulent behavior, the presence of the hypermucoid phenotype (HMV) is commonly used as a phenotypic laboratory marker for identification of hvKp strains (Shon et al., 2013; Catalán-Nájera et al., 2017).

hvKp usually cause community-acquired infections in individuals who are previously healthy (Choby et al., 2020). Infections are more common in the Asian Pacific Rim

but are occurring globally (Shon et al., 2013; Russo and Marr, 2019). hvKp infections frequently involve multiple sites or subsequently exhibit metastatic spread, requiring source control. hvKp have an increased ability to cause central nervous system infection and endophthalmitis, which require rapid recognition and site-specific treatment (Choby et al., 2020). hvKp are rarely MDR and most strains retain susceptibility to most of available antibiotics (Shon et al., 2013; Russo and Marr, 2019; Zhu et al., 2021). In fact, MDR Kp strains responsible of hospital acquired infections and hvKp strains causing community acquired infections apparently belong to two distinct and well segregated bacterial populations (Bialek-Davenet et al., 2014; Wyres et al., 2020). However, convergence of virulence and resistance determinants within hybrid lineages is possible and has been increasingly reported worldwide (Tang et al., 2020). Various carbapenemase genes have recently been detected in hvKp isolates, including those encoding enzymes of the OXA-48, KPC, NDM, and VIM lineages, generating great concern at international and European level (European Centre for Disease Prevention and Control (ECDC), 2021). At least one fatal hospital outbreak of MDR hvKp strains has been reported in China, where carbapenemase-producing hvKp are increasingly common (Gu et al., 2018).

In Italy, sporadic cases of infections caused by HMV Kp have been documented (Arena et al., 2016a; Piazza et al., 2022), in some cases the isolates showed a MDR phenotype associated with carbapenemase production (Arena et al., 2017; Scaltriti et al., 2020), but detailed knowledge about the epidemiology of hvKp strains is lacking.

In this work, we undertook a nationwide surveillance to investigate the epidemiology and features of HMV Kp isolated from bloodstream infections in Italy, a setting of endemic dissemination of MRD Kp clones (Di Pilato et al., 2021). The collected isolates were subjected to comprehensive molecular characterization and were also studied for pathogenicity using the *Galleria mellonella* and two different mouse infection models.

Materials and methods

Study design and isolates collection

For a 6 months period (1st December 2016–31st May 2017), 43 laboratories located in different areas of Italy (Figure 1) and belonging to the KHIN (*Klebsiella pneumoniae* Hypermucoviscous Italian Network of AMCLI) were asked to collect and store all consecutive, non-duplicated, Kp isolates from blood cultures showing a positive “string test” (HMV) (Fang et al., 2004). A dedicated slide set for the correct “string test” performance and interpretation was elaborated and distributed to each center together with the participation letter. Bacterial identification was performed, at participating Laboratories, using routine methods (biochemical methods and/or MALDI-ToF). Laboratories were also asked to collect and report the following data: (i) total number of beds in the hospital/s served by the Laboratory; (ii) total number of blood culture sets collected during the study period; (iii) number of Kp bacteremia episodes occurred during the study period (Supplementary Table 1). Suspected HMV Kp were stored at -80°C until shipment. At the end of the collection period all isolates categorized as HMV by satellite centers were centralized to the Microbiology Laboratory of the Department of Experimental and Clinical Medicine, University of Florence (Florence, Italy) for HMV phenotype confirmation and further analyses (central Laboratory).

Antibiotic susceptibility testing

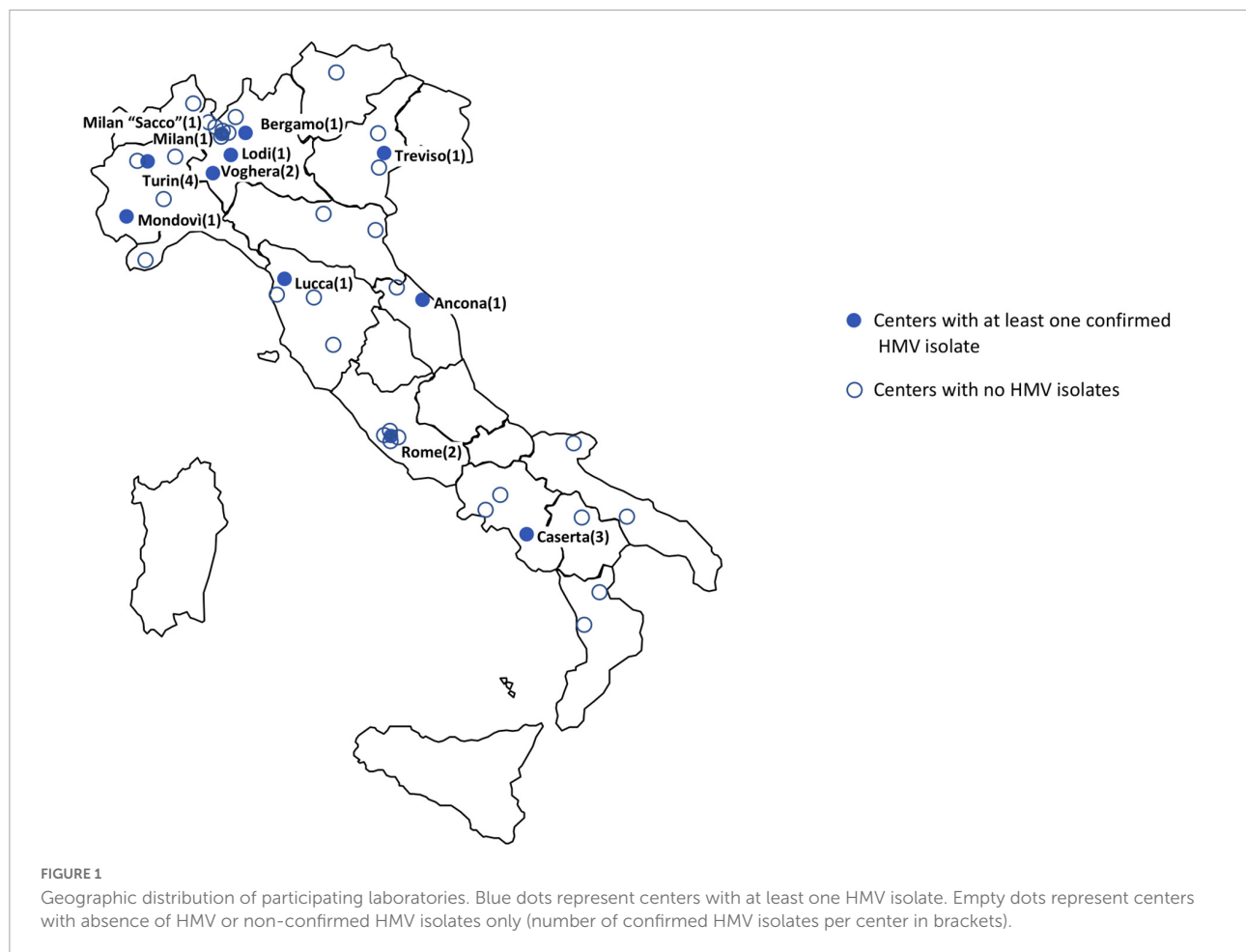
Susceptibility of confirmed HMV isolates to amikacin, cefepime, ceftazidime, ceftazidime/avibactam, ceftolozane/tazobactam, ciprofloxacin, colistin, gentamicin, meropenem, piperacillin/tazobactam, and tigecycline was carried out by broth microdilution (ISO 20776–1, 2019) using lyophilized custom plates (Merlin Diagnostika, Germany). Fosfomycin susceptibility was tested by agar-dilution in the presence of 25 mg glucose-6-phosphate (Camarlinghi et al., 2019). Results were interpreted according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST breakpoints v 12.0). For tigecycline the PK-PD breakpoint of 1 mg/L was adopted.

Whole genome sequencing and bioinformatics

Genomic DNA of confirmed HMV isolates was extracted using the DNeasy PowerLyzer PowerSoil Kit (Qiagen, Hilden, Germany) and subjected to whole genome sequencing (WGS) with the MiSeq platform (Illumina Inc., San Diego, CA), using a 2×250 paired-end approach (Arena et al., 2020). Draft genome assemblies were generated using SPAdes v3.11 (Bankevich et al., 2012).

The PathogenWatch online platform (by Center for Genomic Pathogen Surveillance)¹—was used to generate genotyping report of all sequenced genomes. The software automatically passes assembled genomes to Kleborate and Inctyper pipelines. The Kleborate software provides in depth species, ST (according to the 7 genes MLST scheme), virulence genes, capsule typing, O locus typing, and antimicrobial resistance prediction. Kleborate examines five key acquired virulence loci that are associated with invasive infections and are found at high prevalence among hvKp strains: the siderophores yersiniabactin (ybt), aerobactin (iuc) and salmochelin (iro), the genotoxin colibactin (clb), and the hypermucoidy locus *rmpADC*. The alternative hypermucoidy marker gene *rmpA2* is also searched. Kleborate outputs a simple categorical virulence score, and if resistance screening is enabled, an antimicrobial resistance score as well. These scores provide a rough categorization of the strains to facilitate monitoring resistance-virulence convergence. The virulence score ranges from 0 to 5: 0 = negative for all of yersiniabactin (ybt), colibactin (clb), aerobactin (iuc) 1 = yersiniabactin only 2 = yersiniabactin and colibactin (or colibactin only) 3 = aerobactin (without yersiniabactin or colibactin) 4 = aerobactin with yersiniabactin (without colibactin) 5 = yersiniabactin, colibactin and aerobactin. Kleborate screens input genomes against a curated version of the CARD database of acquired resistance gene alleles, and groups these by drug class for reporting purposes. The resistance score ranges from 0 to 3: 0 = no ESBL, no carbapenemase (regardless of colistin resistance) 1 = ESBL, no carbapenemase (regardless of colistin resistance) 2 = Carbapenemase without colistin resistance (regardless of ESBL genes or OmpK mutations) 3 = Carbapenemase with colistin resistance (regardless of ESBL genes or OmpK mutations) (Wyres et al., 2016; Wick et al., 2018; Lam et al., 2021). Inctyper is an in-house tool that uses the PlasmidFinder database and BLAST to identify the contigs containing a plasmidic Inc reference gene (Carattoli et al., 2014). The presence of core- and accessory virulence factors was investigated using the Virulence Factor Database (VFDB; last access on May 20, 2022) (Chen et al., 2005). For isolates encoding multiple virulence genes, manual analyses

¹ <https://pathogen.watch>



were conducted using NCBI BLAST tool to screen for the presence of additional determinants associated with canonical pK2044-/pLVPK-like virulence plasmids [e.g., genes coding for proteins involved in iron metabolism (*cobW*) and transport (*fecI-fecA*), the hemin and lysine transport system (*shiF*), metabolic transporter (*peg-344*), and transcriptional regulations of virulence gene expression (*luxR*)]. Adjunctive single target analyses were performed using the NCBI BLAST tool with the database BIGSdb curated by Pasteur Institute.²

Lethality in the *Galleria mellonella* model

Bacterial strains were grown in LB broth, aerobically, at $35 \pm 2^\circ\text{C}$, harvested during exponential phase ($\text{OD}_{600} \approx 0.7$), and washed once with 10 mM phosphate-buffered saline (PBS, pH 6.5). Bacteria were then suspended in PBS to an OD_{600} of 1.5, corresponding approximately to 1×10^9 CFU/mL.

² <https://bigsd.bpasteur.fr/klebsiella/>

Larvae weighing 450–600 mg were used for the experiments. For comparative evaluation of virulence, groups of 10 larvae were injected with 5×10^5 CFU of each strain and with sterile PBS as control. Larvae were kept at $35 \pm 2^\circ\text{C}$ in the dark, in humidified atmosphere, with food, and daily examined for pigmentation and mobility. Time of death was recorded at 24, 48, and 72 h. For each strain, data from three independent experiments were combined. GraphPad Prism 6.0 software was used for mean mortality \pm SD calculation (GraphPad Software Inc., La Jolla, CA). *K. pneumoniae* NTUH-H2044, an ST23 hvKp strain and the less virulent ST258 Kp KKBO-1 strain were also included in all experiments as references (Fang et al., 2004; Arena et al., 2016b).

Murine model

Isolates were grown on TSA agar. A single colony was suspended in 10 ml of Brain Heart Infusion broth and incubated over-night at $35 \pm 2^\circ\text{C}$. Overnight cultures were centrifuged, and pellets were resuspended in PBS at final concentrations of 10^9 CFU/mL. A total of 100 μl of each strain were injected into

10 BALB/c mice (Harlan Italy S.r.l., Udine, Italy) through the tail vein. All animals used for experiments were females 10 weeks of age weighing approximately 20–25 g. Forty-eight hours after inoculation, mice were sacrificed by cervical dislocation. Kidneys and livers were removed, weighed, and homogenized using a stomacher (model 80; Pbi International, Milan, Italy). Serial homogenate dilutions were plated onto Brain Heart Infusion agar for CFU count. To confirm their identity and to avoid possible contaminations, different morphologies' isolates were identified using the MALDI-ToF based method. Animal experiments were performed at animal facilities of the Catholic University (Rome, Italy), under a protocol approved by the Institutional Animal Use and Care Committee at Università Cattolica del S. Cuore, Rome, Italy and authorized by the Italian Ministry of Health, according to Legislative Decree 116/92, which implemented the European Directive 86/609/EEC on laboratory animal protection in Italy. Animal welfare was routinely checked by veterinarians of the Service for Animal Welfare, (Protocol n°935/2017-PR).

Statistics

Differences between groups of isolates with different virulence levels in the animal models were assessed using the one-way ANOVA. Analysis was performed using Stata 15 (StataCorp, College Station, TX) or GraphPad Prism 7 (GraphPad Software, San Diego, CA) software, and $P < 0.05$ was considered statistically significant.

Results

During the study period, the 43 participating Laboratories (serving a total of 29,719 hospital beds) processed 191,800 blood culture sets. A total of 1,502 bacteremia episodes sustained by bacteria identified as Kp were reported. Of these, 52 (52/1,502, 3.5%) were reported as “string test positive” according to satellite Laboratory screening. At the central Laboratory, 19 isolates were confirmed as HMV (19/1,502, 1.3% of HMV Kp bacteremia cases) and subjected to further analysis. Most of confirmed HMV isolates were obtained in Laboratories from northern Italy, mainly from Lombardy (6/19 isolates, 31.6%) (Figure 1). Prevalence of HMV strains over the total of Kp BSI episodes at participating centers ranged from 0 to 20% (Supplementary Table 1).

Antimicrobial susceptibility of confirmed hypermucoviscous isolates

Overall, the most active drugs were ceftazidime/avibactam (5.3% of resistant isolates), followed by amikacin, meropenem, gentamicin, colistin and fosfomycin (10.6, 21.0, 36.8, 36.8,

and 47.4% of resistant isolates, respectively). Five (5/19, 26.3%) isolates showed a meropenem MIC > 0.125 mg/L and were therefore considered as suspect carbapenemase producers. Eleven isolates (11/19, 57.9%) had an MIC for ceftazidime > 1 mg/L and were suspect ESBL producers (Table 1, Supplementary Table 2).

Genome sequencing and virulence data

Bacterial identification and clonality

Based on the nature of the housekeeping beta-lactamase gene, 18 isolates were identified as Kp (presence of a *bla*_{SHV}-type gene) and one as *Klebsiella variicola* (presence of a *bla*_{LEN}-type gene). According to the 7 genes MLST scheme, one third of Kp isolates (6/18) belonged to ST307, three isolates were assigned to ST512 and two isolates to ST29. The remaining were singletons of ST5, ST8, ST11, ST25, ST35, and ST198. One isolate represented a new ST: a single locus variant of ST37 (new *phoE* allele).

ST307 isolates

All the six ST307 isolates had a KL102 capsule locus type (associated with *wzi173* allele) with a O1/O2v2 O locus and carried the yersiniabactin siderophore gene cluster *ybt* 9, associated with a ICEKp3 (Figure 2, Supplementary Figure 1, and Supplementary Table 3). With respect to the resistome, two isolates carried a *bla*_{KPC} carbapenemase gene, a *bla*_{KPC-3}, and a *bla*_{KPC-2} variant, respectively. One isolate (GMR153) encoded a CMY-16, class C, beta-lactamase. All isolates encoded a CTX-M-15 ESBL and trimethoprim-sulfamethoxazole resistance determinants. All except one carried genes coding for multiple

TABLE 1 Antimicrobial susceptibility testing results for HMV isolates included in the study.

Antibiotics	MIC range	MIC50	MIC90	% S	% I	% R
Amikacin	=4–32	=4	8	89.4	0.0	10.6
Cefepime	=1– > 8	4	>8	42.1	5.3	52.6
Ceftazidime	=0.25– > 32	32	>32	42.1	0.0	57.9
Ceftazidime/ avibactam	=1/4– > 8/4	=1/4	2/4	94.7	0.0	5.3
Ceftolozane/ tazobactam	=1/4– > 8/4	2/4	>8/4	52.6	0.0	47.4
Ciprofloxacin	=0.0625– > 8	2	>8	42.1	0.0	57.9
Colistin	=1– > 8	=1	>8	63.2	0.0	36.8
Fosfomycin	=4– > 128	16	>128	52.6	0.0	47.4
Gentamicin	=0.25–32	=0.25	>32	63.2	0.0	36.8
Meropenem	=0.125– > 16	=0.125	>16	73.7	5.3	21.0
Piperacillin- tazobactam	2/4– > 128/4	16/4	>128/4	42.1	0.0	57.9
Tigecycline	1– > 8	2	8	36.8	0.0	63.2

MICs are expressed in µg/mL.

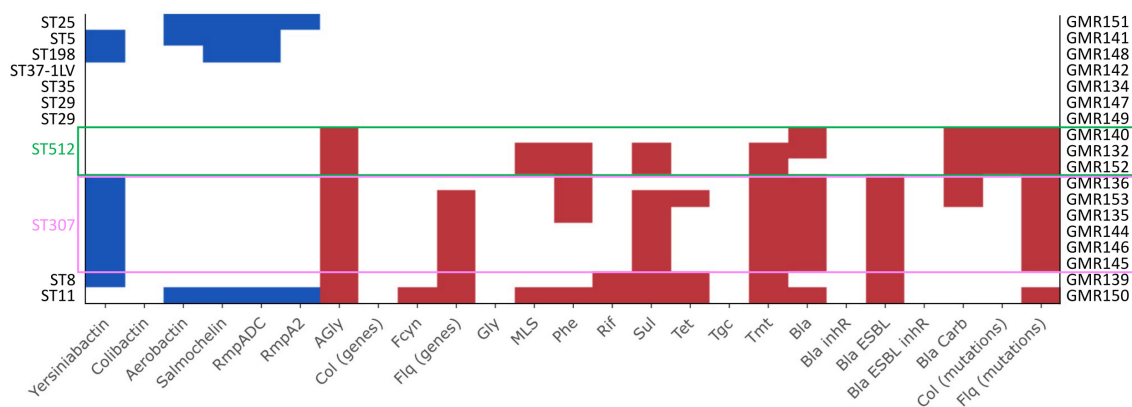


FIGURE 2

Results of Kleborate pipeline visualized with the Kleborate-viz tool. Detected resistance determinants are in red while virulence factors are in blue. Legend: AGly_acquired (aminoglycosides), Bla_acquired (beta-lactamases), Bla_inhR (beta-lactamases with resistance to beta-lactamase inhibitors), Bla_Carb (carbapenemase), Bla_ESBL (extended spectrum beta-lactamases), Bla_ESBL_inhR (extended spectrum beta-lactamases with resistance to beta-lactamase inhibitors), Fcyn (fosfomycin), Fiq (fluoroquinolones), Gly (glycopeptides), MLS (macrolides), Phe (phenicols), Rif (rifampin), Sul (sulfonamides), Tet (tetracyclines), Tmt (trimethoprim), Tgc (tigecycline).

aminoglycoside-modifying enzymes and fluoroquinolones resistance (Figure 2 and Supplementary Table 3).

All isolates were phenotypically resistant to ceftazidime and ciprofloxacin. Five isolates were colistin resistant and two were resistant to carbapenems (the two KPC-type producers) (Supplementary Table 2).

In the *G. mellonella* animal model, all ST307 isolates showed a low/intermediate virulence potential, comparable to that of the control strain KKBO-1 (Figure 3A). These results were confirmed in both murine infection models (Figure 3B).

ST512 isolates

The three isolates belonging to ST512 had a KL107 capsule locus type (associated with *wzi154* allele) and a O1/O2v2 O locus. None of the searched virulence factors was found. All isolates showed a complex content of resistance genes (resistance score: 3), including a *bla*_{KPC-3} carbapenemase, aminoglycosides resistance determinants, mutational alterations previously associated with colistin resistance (alterations in the *mgrB* gene) and fluoroquinolones resistance. Two out of three isolates encoded also for trimethoprim-sulfamethoxazole resistance determinants (Figure 2, Supplementary Figure 1, and Supplementary Table 3).

Coherently, all isolates were resistant to ceftazidime, cefepime, ceftolozane-tazobactam, piperacillin-tazobactam and meropenem. One isolate was resistant to ceftazidime/avibactam also (GMR140). Two isolates out of three were colistin resistant and all were resistant to fosfomycin and ciprofloxacin (Supplementary Table 2).

The virulence potential of these strains was low in two cases and intermediate in one case (GMR152) in both *G. mellonella* and murine infection models (Figures 3A,B).

ST29 isolates

The two ST29 isolates were associated with a O1/O2v2 O locus but carried a different K loci, KL30 (*wzi85*) and KL114 (*wzi354*), respectively. None of the virulence and resistance determinants included in the analysis pipeline was detected (Figure 2, Supplementary Figure 1, and Supplementary Table 3). Both isolates were phenotypically susceptible to all tested drugs, except for GMR149 that was resistant to fosfomycin and showed a tigecycline MIC above the PK/PD breakpoint (MIC > 8 µg/mL) (Supplementary Table 2).

In the animal models, ST29 isolates showed a level of virulence that was intermediate between that of KKBO-1 and NTUH-K2044 strains (Figures 3A,B).

Other sequence types

The ST5 isolate (GMR141) was associated with the KL39 (*wzi112* allele) and the O1/O2v1 O locus. The isolate carried an *ybt 2* locus associated with a ICE*Kp1* (Ybt ST324), together with synthesis loci for salmochelin (*iro3* lineage—*SmST 3*), aerobactin (*iuc1*), and regulator of hypermucoidy (*rmpADC* lineage 3). Sequence analysis revealed that the acquisition of the former virulence factors was not likely associated with the presence of a pK2044-like element, the archetypal virulence plasmid from the hypervirulent *K. pneumoniae* NTUH-K2044, suggesting that a novel or recombined virulence element could have mediated the acquisition of these traits. Concerning the resistome, the housekeeping SHV-11 beta-lactamase only (a non-ESBL) was found (Figure 2, Supplementary Figure 1, and Supplementary Table 3), coherently with the susceptibility profile (Supplementary Table 2). In the animal models, GMR141 was the most virulent strain of the collection, with lethality and ability to infect mouse liver and kidneys higher than the virulent control NTUH-K2044 (Figures 3A,B).

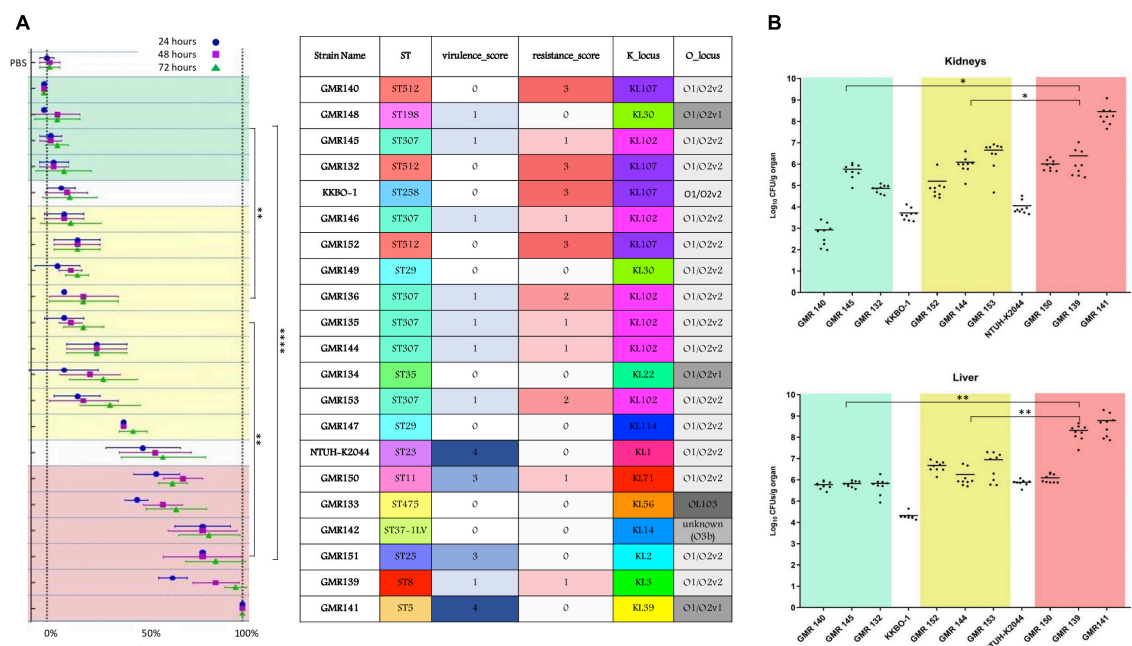


FIGURE 3 Summary of most relevant features of studied strains (ST, virulence score, resistance score, K type and O type) and results obtained in the animal infection models. **(A)** Results of experiments in the *G. mellonella* model expressed as mortality rate (percentage of dead over the total of inoculated larvae). Means and SD of three replicate experiments are shown at different time points for each strain (after 24, 48, and 72 h). **(B)** Results of virulence experiments, for representative HMV isolates, in the murine model expressed as Log10CFU/g organ of the studied strains, recovered from kidneys and livers. Each dot represents data obtained from one mouse specimen and the solid line represents the mean value. In both the models, the NTUH-K2044 virulent strain and the KKBO-1 low-virulence strain were included and used as reference for classification of studied isolates in three different virulence categories: low (green zone), intermediate (yellow) and high (red). Mortality rate at 72 h **(A)** and Log10 counts **(B)** for the three groups were compared. *P*-values < 0.05 were considered significant; * < 0.05, ** < 0.005, **** < 0.0001.

The ST8 isolate (GMR139) included in the study was associated with a KL3 (*wzi59* allele) and O1/O2v2 capsular locus and O locus, respectively. The genome was characterized by the presence of *ybt 7* locus associated with a ICEKp7. By contrast, the content of resistance determinants was relevant with a *bla*_{SHV-2} (ESBL phenotype), multiple aminoglycosides modifying enzymes, trimethoprim-sulfamethoxazole resistance genes and an acquired *qnr*-type gene associated with fluoroquinolones resistance (Figure 2, Supplementary Figure 1, and Supplementary Table 3). The isolate was resistant to ceftazidime, cefepime, ceftolozane-tazobactam, piperacillin-tazobactam and ciprofloxacin and was highly virulent in the *G. mellonella* and the murine models (Figures 3A,B).

The ST11 isolate (GMR150) was associated with a KL71 capsular locus (*wzi72*) and a O3b locus. The aerobactin (*iuc 1*), salmochelin (*iro 1*) and the regulators of hypermucoidy (*rmpADC* locus) were detected. Sequence analysis revealed that a plasmid element resembling pK2044 (query coverage: 84%, BLAST identity: 99.6%) likely contributed to the acquisition of the virulence factors (i.e., aerobactin, salmochelin, *rmp* cluster). Concerning the resistome, genes coding for an acquired class C (*bla*_{DHA-1}) and class A (*bla*_{CTX-M-65}) beta-lactamases, as well as for

multiple acquired aminoglycoside modifying enzymes and for fosfomycin resistance (*fosA3*) were detected (Figure 2, Supplementary Figure 1, and Supplementary Table 3). The isolate was phenotypically resistant to ceftazidime, cefepime, ceftolozane-tazobactam, ciprofloxacin, fosfomycin and piperacillin-tazobactam and was highly virulent in the *G. mellonella* and murine models, with lethality higher than NTUH-K2044 (Figures 3A,B).

The ST25 isolate (GMR151) was associated with a KL2 capsular locus (*wzi72*) and a O1/O2v2 O locus. The content of virulence genes was relevant (virulence score: 3) including synthesis loci for salmochelin (*iro1* lineage), aerobactin (*iuc1*), and regulators of the mucoid phenotype (*rmpADC*). As for the ST11 isolate, sequence analysis revealed the presence of a pK2044 derivative plasmid (query coverage: 81%, BLAST identity: 99.5%) that likely contributed to the acquisition of former virulence factors (i.e., *iuc1*, *iro1*, *rmpADC*).

Except the housekeeping SHV-11 beta-lactamase, none of the screened resistance determinants was found (Figure 2 and Supplementary Table 3), coherently with the susceptibility profile (Supplementary Table 2). A high lethality in the *G. mellonella* model was shown (Figure 3A).

The ST198 isolate (GMR148) was associated with a KL30 capsular locus (*wzi85*) and a O1/O2v1O locus, carried an *ybt* 2 locus associated with a ICE*Kp*1 (Ybt ST324); it carried a variant of the salmochelin synthesis locus (*iro* 3 lineage) and regulator of hypermucoidy (*rmpADC* lineage 3). The acquisition of the former virulence factors was not likely associated with the presence of a pK2044-like element. No acquired resistance genes were detected, and the isolate was phenotypically susceptible to tested antibiotics (Supplementary Table 2). The virulence level in the *G. mellonella* model was extremely low (Figure 3A).

The ST35 (GMR134), the new ST single locus variant of ST37 (GMR142), and the *K. variicola* (GMR133) isolates were associated with KL22 (*wzi37*)/O1/O1v1, KL14 (*wzi14*)/unknown (closest match to O3b) and KL56/OL103, capsular/O locus types, respectively. No resistance genes nor virulence determinants were found and the isolates were susceptible to all tested drugs, except for two isolates (GMR134, GMR142) that were resistant to fosfomycin (Figure 2 and Supplementary Tables 2, 3). The three isolates showed a variable level of virulence: intermediate (GMR134) or high (GMR142 and GMR 133).

Discussion

In this work we report on the results of the first Italian nationwide surveillance (more than 40 hospitals involved) study for HMV *Kp* bacteremia. Overall, the prevalence of isolates with an HMV phenotype (positive string test), over the total of *Kp* obtained from blood cultures, at the involved centers, was low (19/1502, 1.3%). A trend trough an overestimation of the HMV phenotype was observed in most of the participating centers. In fact, satellite centers stored 52 suspected HMV isolates but 19 only were confirmed HMV at the central Laboratory. This observation underscore that need of a certain level of Laboratory skills for the correct interpretation of the “string test.”

Of the 19 HMV confirmed isolates, only six were highly virulent in the animal models and could be therefore considered hv*Kp* (overall prevalence of highly virulent HMV lineages: 0.4%), confirming that HMV phenotype and hypervirulence can't be considered synonymous in *Kp* (Catalán-Nájera et al., 2017). Systematic surveillance studies are lacking but similar low prevalence data were obtained from other countries (Peirano et al., 2013; Örstén et al., 2020) with the notable exception of China, Taiwan and South Korea. Indeed, HMV *Kp* circulation is historically more relevant in these areas (Lee et al., 2006, 2017; Li et al., 2014; Zhang et al., 2016).

In our surveillance study, despite the low prevalence, at least one HMV isolate was found in 11/43 centers located in different geographic areas of the Country, mainly clustered in Northern Italy, underscoring a significant geographic dissemination of the phenomenon.

The antimicrobial susceptibility profile of collected HMV strains reflected the situation reported for Italy in the same period. In particular, the 26.3% of HMV strains included in our surveillance was meropenem non-susceptible (suspect carbapenemase-producers) and the 57.9% were ceftazidime non-susceptible (suspect ESBL producers). In 2017, the EARS-NET European Surveillance System (European Center for Disease Prevention and Control), reported for Italy similar prevalence data with: 32.1 and 56.1% of carbapenems non-susceptible and 3rd generation cephalosporins non-susceptible *Kp* isolates, respectively.³

Our study also provides the first comprehensive overview of genetic, phenotypic and virulence features of systematically collected HMV strains, from Italy.

Analyzing the data, it was possible to identify three different patterns of isolates. Pattern i) including isolates that accumulated multiple resistance determinants (including carbapenemase genes and colistin resistance mutations) but with absence or a low content of virulence genes and a low/intermediate virulence behavior in animal models. These isolates belong to two major high-risk clones, namely ST512 and ST307.

A multi-national large genomic surveillance study revealed that the population of carbapenem-resistant *Kp*, from invasive infections, in Italy was mainly contributed by ST258/512, ST101, and ST307 nosocomial clones producing the KPC-type carbapenemase, during 2013–14 (David et al., 2019).

Consistently, a more recent Italian survey performed in the 2016–2017 period (Di Pilato et al., 2021) showed that ST512 was largely prevalent among carbapenemase producers, followed by ST307, both producing the KPC-type beta-lactamases. Therefore, we can speculate that occasional acquisition events of the HMV phenotype may be recognized within the endemically disseminated nosocomial clones circulating in Italy. A case describing an HMV *Kp* belonging to the ST515 epidemic clone has been reported from Northern Italy. In that case, despite the low virulence potential in the *G. mellonella*, the isolate caused a fatal systemic infection with liver abscesses in an immunocompromised patient (Arena et al., 2017).

Interestingly, similarly to what previously described, none of the isolates in this pattern was HMV due to the presence of the classic regulators of the mucoid phenotype (*rmpADC*). Consistently, none of these isolates carried classic or recombined versions of most common virulence plasmids.

As such, the presence of *rmp* genes as molecular marker predictive of the HMV phenotype should be carefully considered for epidemiological purposes. The pathogenetic significance of the HMV phenotype in these cases remain uncertain. However, none of pattern (i) isolates was highly virulent in animal models. Differently from what previously

³ <https://www.ecdc.europa.eu/en/surveillance-atlas-infectious-diseases>

described by other authors (Diago-Navarro et al., 2014), in the HMV population included in this study, the mortality in *G. mellonella* model predicted high virulence in mice models. Our study supports the use of the *G. mellonella* model as a feasible and predictive approach for high throughput screening of virulence behavior in Kp.

Pattern (ii) including isolates coupling the absence of major resistance and virulence determinants with a low or intermediate virulence potential in the animal model. These strains belonging to ST29 and ST35 were phenotypically multi-susceptible and the HMV phenotype wasn't linked to *rmpADC* mediated mechanism.

Pattern (iii) including isolates characterized by a relevant content of virulence genes and, in some cases, presence of multiple resistance determinants and an MDR phenotype with retained susceptibility to carbapenems. These strains, belonging to ST5, ST8, ST11, and ST25, were all highly virulent in animal models.

According to previously published extensive genomic data, ST5 and ST8 isolates are extremely infrequent (David et al., 2019).

GMR141, an ST5 strain, had the highest virulence potential of the collection and showed multi-susceptible phenotype. It carried the most extensive array of virulence factors (virulence score 4) including: ICEKp1, the aerobactin, the salmochelin siderophores and a *rmpADC* hypermucoidy lineage 3 (Figure S1). In the genomic population study conducted in 2019 by David et al. (2019), three ST5 strains only, over a total of 1,717 Kp, were described and none of these strains displayed a such relevant content of virulence genes. Therefore, we can conclude that this strain represents a potentially emerging new virulent clone.

The same can be asserted for GMR139, an ST8 ESBL strain with the *ybt7* cluster and a virulent behavior in the animal models. ST8 isolates are quite uncommon (one isolate over 1,717 only in the David et al. study). Interestingly, GMR139 has a KL3 *cps* type that has been previously identified as a relevant virulence factor for *Klebsiella pneumoniae* subsp. *rhinoscleromatis* (Corelli et al., 2018).

Differently from the previous two strains included in this cluster, GMR150 belongs to a well-known MDR high-risk clone, namely ST11. ST11 has been previously associated with virtually all carbapenemase classes, mainly *bla*_{KPC-2}, *bla*_{NDM-1}, and *bla*_{OXA-48} (Liao et al., 2020). Some ST11 lineages have been previously associated with an aggressive behavior and hospital outbreaks of carbapenem-resistant ST11 have been described. A KPC-2-producing hvKp ST11 has been linked to a cluster of severe lung infection with fatal multi-organ failure or septic shock in two Chinese hospitals (Gu et al., 2018; Xu et al., 2019). Similar to previously characterized ST11 hvKp MDR strains, GMR150 carries a plasmid highly correlated to pK2044, found in NTUH-K2044, that likely contributed to the acquisition of the former virulence factors

(i.e., aerobactin, salmochelin, *rmp* cluster). As other ST11 MDR strains from China, GMR150 was highly virulent in the *G. mellonella* and murine models (Zhang et al., 2020). Differently from previously described ST11 hvKp MDR strains, our isolate carried an unusual capsular type: the KL71. KL71 has been previously linked to *Klebsiella variicola* (de Campos et al., 2021). Even if the strain described in this work was carbapenem-susceptible and none of most common carbapenemase classes was detected, the content of acquired resistance mechanisms was notable, including a *bla*_{CTX-M-65} a *bla*_{DHA-1} beta-lactamases and the *fosA3* fosfomicin resistance gene. Interestingly, these two latter resistance mechanisms are strongly epidemiologically linked to the Asiatic continent (Taniguchi et al., 2017; Hennequin et al., 2018; Hao et al., 2021; Wang et al., 2022). To our best knowledge, this is the first description of a FosA3 producing Kp reported from Italy. Therefore, we can conclude that we documented the presence of a highly virulent ST11 MDR Kp clone with several genetic features all supporting a link with the Asiatic continent. Analyzing the features of this strain, the classic proposed paradigm depicting a situation with a clear segregation between populations of MDR and hvKp isolates seems to be surpassed (Bialek-Davenet et al., 2014).

The last strain included in this pattern belongs to the ST25 (GMR151). ST25, associated with the KL2 capsule, is a known hypervirulent strain initially linked to outbreaks of septicemia cases in pigs in England (Bidewell et al., 2018; Wanford et al., 2021). Recently, an expansion of hypermucoviscous MDR ST25 (production of *bla*_{KPC} carbapenemase) has been reported from Argentina (Cejas et al., 2019). As for the ST11 isolate, the presence of a pK2044 derivative plasmid was likely the explanation of the virulence potential of this strain.

Finally, some isolates were not assimilable to the proposed patterns. For example, the ST198 was avirulent in the animal models but carries an array of virulence genes resembling that of the ST5 strain with the notable exception of the aerobactin siderophore. The new ST single locus variant of ST37 (GMR142) was multi-susceptible too but highly virulent and presented the new O locus (O3b-like).

The only *K. variicola* (GMR133) isolate in our collection was highly virulent in the *G. mellonella* model and presented a multi-susceptible phenotype. *K. variicola* has been associated with severe infections (Rodríguez-Medina et al., 2019) and hypervirulent strains have been described (Lu et al., 2018). Interestingly, our strain didn't carry classic regulators of the mucoid phenotype nor plasmids similar to those previously associated with the HMV phenotype in *K. variicola* (Rodríguez-Medina et al., 2020).

Concluding, in this work we describe the genetic features of HMV Kp isolates circulating in Italy using an approach based on a first phenotypic screening with the "string test" followed by genome sequencing and final confirmation of the virulent behavior in the *G. mellonella* and mouse models. The approach

enabled us to demonstrate the presence in the Italian territory of MDR (mainly associated with an ESBL phenotype), highly virulent clones (such as ST11) characterized by the carriage of derivatives of known virulence plasmid. Furthermore, we identified some new highly virulent clones belonging to ST8 and ST5. Adjunctively, we confirmed the presence of carbapenem-resistant ST512 and ST307 HMV isolates with a high content of resistance genes and a low/intermediate virulence level. A possible limitation of our study is that, due to the study design, patients demographic information, and comorbidities data were not collected and analyzed.

Overall, our data can be considered reassuring. In fact, the prevalence of hvKP was extremely low and all highly virulent and MDR strains in our collection were singletons of different sequence types (STs). Apparently, a clonal expansion of most problematic lineages was not present. Furthermore, the classic hvKP clones (ST23 and ST86) were not found.

However, the prevalence of hvKP was higher in some centers, distributed in northern Italy, that could become hot spots of dissemination. Continued efforts in identifying and tracking the dissemination of hvKP are warranted to deliver timely public health responses aimed at containing their spread within the healthcare system.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA781811.

Ethics statement

The animal study was reviewed and approved by the Institutional Animal Use and Care Committee at Università Cattolica del S. Cuore (Protocol n°935/2017-PR).

The AMCLI *Klebsiella pneumoniae* hypermucoviscous Italian network

Maria Grazia Cusi (Siena University Hospital, Siena, Italy), Anesi Adriano (Maggiore Hospital, Lodi, Italy), Maria Labonia (IRCCS “Casa Sollievo della Sofferenza,” San Giovanni Rotondo, Foggia, Italy), Ivana Vada (Cardinal Massaia Hospital, Asti, Italy), Francesca Maldera (Della Murgia F. Perinei Hospital, Altamura, Bari, Italy), Sandra Carpicci (S. Pietro Fatebenefratelli Hospital, Rome, Italy), Chiara Vettori, Romano Mattei (USL North Western Tuscany, Lucca, Italy), Tamara Brunelli (USL Center Tuscany, Prato, Italy), Antonietta Sinno (Madonna delle Grazie Hospital, Matera, Italy), Sara

Rimoldi (Sacco Hospital, Milan, Italy), Barbara Pieretti (Santa Croce Hospital, AO Riuniti Hospital Marche Nord, Ancona, Italy), Richard Aschbacher (Bolzano Central Hospital, Azienda Sanitaria dell’Alto Adige, Bolzano, Italy), Vincenzo Minasi, Loredana Delflorio (Multimedica S.p.a., Sesto San Giovanni, Milan, Italy), Andrea Bartolini, Margherita Scapaticci (San Camillo Hospital, Treviso, Italy), Maria Zoppelletto (ASL 3 Bassano de Grappa, Vicenza, Italy), Cinzia Rossi, Claudia Canale (Castelli Hospital -Asl VCO, Verbania, Italy), Pier Andrea Dusi (USL 1 Imperiese, San Remo, Imperia, Italy), Maria Maddalena Grossi (S. Francesco Hospital, Paola, Cosenza, Italy), Maria Federica Pedna (AUSL della Romagna, Pievesestina, Cesena, Italy), Bruno Mariani (San Camillo Forlanini Hospital, Rome, Italy), Chiara Vincenzi (Riuniti Hospital Ancona, Ancona, Italy), Gerardino Amato, Linda Degl’Innocenti (Cardarelli Hospital, Naples, Italy), Erminio Torresani, Elisabetta Cesana (Istituto Auxologico Italiano, Milan, Italy), Domenico Salamone (USL North Western Tuscany, Pontedera, Pisa, Italy), Maria Paola Ferrero (Koelliker Hospital, Turin, Italy), Daniela Rossi, (Villa Salus Hospital, Mestre, Venice, Italy), Fulvia Milano (ASL Vercelli, Vercelli, Italy), Luigi Principe (A. Manzoni Hospital, Lecco, Italy), Roberto D’Angelo, Loredana Vizzini (A.S.P. Golgi-Redaelli, Milan, Italy), Marco Passera (ASST Papa Giovanni XXIII, Bergamo, Italy), Claudia Venturelli (AOU Policlinico di Modena, Modena, Italy), Floriana Gona (San Raffaele Scientific Institute, Milan, Italy), Erminia Casari (Humanitas Research Hospital, Milan, Italy), Rita Greco, Vittorio Panetta (AORN Sant’Anna e San Sebastiano, Caserta, Italy), Lara Ines Bellazzi (ASST di Pavia—Voghera Civile Hospital, Pavia, Italy), Maria Carmela Cava, Carmen Luciana Bonanno (Sandro Pertini Hospital-ASL Roma 2, Rome, Italy), Alfredo Focà, Angela Quirino (AOU Mater Domini, Catanzaro, Italy), Maria Rosaria Catania (AOU Federico II di Napoli, Naples, Italy), Davide Carcione (Centro cardiologico Monzino IRCCS, Monza Brianza, Italy), Lucina Fossati, Gabriele Bianco, Rossana Cavallo (A.O. Città della Salute e della Scienza, Presidio Molinette, Turin, Italy), Carla Fontana (Tor Vergata University Hospital, Rome, Italy), Michela Quatela (Mondovì Hospital, Cuneo, Italy).

Author contributions

FA: conceptualization, formal analysis, investigation, data curation, writing—original draft, writing—review and editing, and project administration. GM and VD: investigation, data curation, and writing—review and editing. RT, LH, and MC: investigation. AA: investigation and writing—review and editing. MS: resources, writing—review and editing, and supervision. GR: resources, draft, writing—review and editing, and supervision. All authors contributed to the article and approved the submitted version.

Acknowledgments

We are grateful to the members of the Italian Working Group for infections in Critically ill Patient (GLIPaC) of the Italian Association of Clinical Microbiologists for the assistance with the surveillance study organization. We are also grateful to Jin-Town Wang for providing the *Klebsiella pneumoniae* NTUH-K2044 strain as control for virulence experiments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.983294/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Distribution of core and acquired factors involved in host-pathogen interaction and virulence (source: The virulence factor database, VFDB; last access on May 20, 2022). Details of gene products (right panel) were also reported in alphabetical order.

SUPPLEMENTARY TABLE 1

List of included centers with information on number of hospital beds, blood cultures sets, Kp bacteremia episodes, suspected and confirmed HMV bacteremia per year.

SUPPLEMENTARY TABLE 2

Antimicrobial susceptibility testing results of confirmed HMV isolates.

SUPPLEMENTARY TABLE 3

Analytic Kleborate output of confirmed HMV isolates subjected to WGS.

References

- Arena, F., Di Pilato, V., Vannetti, F., Fabbri, L., Antonelli, A., Coppi, M., et al. (2020). Population structure of KPC carbapenemase-producing *Klebsiella pneumoniae* in a long-term acute-care rehabilitation facility: Identification of a new lineage of clonal group 101, associated with local hyperendemicity. *Microb. Genom.* 6:e000308. doi: 10.1099/mgen.0.000308
- Arena, F., Spanu, T., Henrici De Angelis, L., Liotti, F. M., D'Andrea, M. M., Menchinelli, G., et al. (2016a). First case of bacteremic liver abscess caused by an ST260-related (ST1861), hypervirulent *Klebsiella pneumoniae*. *J. Infect.* 73, 88–91. doi: 10.1016/j.jinf.2016.04.006
- Arena, F., Henrici De Angelis, L., Cannatelli, A., Di Pilato, V., and Amorese, M. (2016b). Colistin resistance caused by inactivation of the MgrB regulator is not associated with decreased virulence of sequence type 258 KPC carbapenemase-producing *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* 60, 2509–2512. doi: 10.1128/AAC.02981-15
- Arena, F., Henrici De Angelis, L., D'Andrea, M. M., Cannatelli, A., Fossati, L., Di Pilato, V., et al. (2017). Infections caused by carbapenem-resistant *Klebsiella pneumoniae* with hypermucoviscous phenotype: A case report and literature review. *Virulence* 8, 1900–1908. doi: 10.1080/21505594.2017.1286439
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., and Dvorkin, M. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Bialek-Davenet, S., Criscuolo, A., Ailloud, F., Passet, V., Jones, L., Delannoy-Vieillard, A. S., et al. (2014). Genomic definition of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* clonal groups. *Emerg. Infect. Dis.* 20, 1812–1820. doi: 10.3201/eid2011.140206
- Bidewell, C. A., Williamson, S. M., Rogers, J., Tang, Y., Ellis, R. J., Petrovska, L., et al. (2018). Emergence of *Klebsiella pneumoniae* subspecies *pneumoniae* as a cause of septicemia in pigs in England. *PLoS One* 13:e0191958. doi: 10.1371/journal.pone.0191958
- Camarlinghi, G., Parisio, E. M., Antonelli, A., Nardone, M., and Coppi, M. (2019). Discrepancies in fosfomicin susceptibility testing of KPC-producing *Klebsiella pneumoniae* with various commercial methods. *Diagn. Microbiol. Infect. Dis.* 93, 74–76. doi: 10.1016/j.diagmicrobio.2018.07.014
- Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). In silico detection and typing of plasmids using plasmid finder and plasmid multilocus sequence typing. *Antim. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14
- Carpenter, J. L. (1990). *Klebsiella* pulmonary infections: Occurrence at one medical center and review. *Rev. Infect. Dis.* 4, 672–682. doi: 10.1093/clinids/12.4.672
- Catalán-Nájera, J. C., Garza-Ramos, U., and Barrios-Camacho, H. (2017). Hypervirulence and hypermucoviscosity: Two different but complementary *Klebsiella* spp. phenotypes? *Virulence* 8, 1111–1123. doi: 10.1080/21505594.2017.1317412
- Cejas, D., Elena, A., Guevara Nuñez, D., Sevillano Platero, P., De Paulis, A., Magariños, F., et al. (2019). Changing epidemiology of KPC-producing *Klebsiella pneumoniae* in Argentina: Emergence of hypermucoviscous ST25 and high-risk clone ST307. *J. Glob. Antim. Resist.* 18, 238–242. doi: 10.1016/j.jgar.2019.06.005
- Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., et al. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* 33, D325–D328. doi: 10.1093/nar/gki008
- Choby, J. E., Howard-Anderson, J., and Weiss, D. S. (2020). Hypervirulent *Klebsiella pneumoniae* - clinical and molecular perspectives. *J. Int. Med.* 287, 283–300. doi: 10.1111/joim.13007
- Corelli, B., Almeida, A. S., Sonego, F., Castiglia, V., Fevre, C., Brisse, S., et al. (2018). Rhinoscleroma pathogenesis: The type K3 capsule of *Klebsiella rhinoscleromatis* is a virulence factor not involved in mikulicz cells formation. *PLoS Negl. Trop. Dis.* 12:e0006201. doi: 10.1371/journal.pntd.0006201
- David, S., Reuter, S., Harris, S. R., Glasner, C., Feltwell, T., Argimon, S., et al. (2019). Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat. Microbiol.* 4, 1919–1929. doi: 10.1038/s41564-019-0492-8
- de Campos, T. A., de Almeida, F. M., de Almeida, A. P. C., Nakamura-Silva, R., Oliveira-Silva, M., de Sousa, I. F. A., et al. (2021). Multidrug-resistant (MDR) *Klebsiella variicola* strains isolated in a Brazilian hospital belong to new clones. *Front. Microbiol.* 12:604031. doi: 10.3389/fmicb.2021.604031

- Di Pilato, V., Errico, G., Monaco, M., Giani, T., Del Grosso, M., Antonelli, A., et al. (2021). The changing epidemiology of carbapenemase-producing *Klebsiella pneumoniae* in Italy: Toward polyclonal evolution with emergence of high-risk lineages. *J. Antim. Chemother.* 76, 355–361. doi: 10.1093/jac/dkaa431
- Diago-Navarro, E., Chen, L., Passet, V., Burack, S., Ulacia-Hernando, A., Kodiyanplakkal, R. P., et al. (2014). Carbapenem-resistant *Klebsiella pneumoniae* exhibit variability in capsular polysaccharide and capsule associated virulence traits. *J. Infect. Dis.* 210, 803–813. doi: 10.1093/infdis/ji u157
- European Centre for Disease Prevention and Control (ECDC) (2021). Risk Assessment: Emergence of Hypervirulent *Klebsiella pneumoniae* ST23 Carrying Carbapenemase genes in EU/EEA Countries. Available online at: <https://www.ecdc.europa.eu/en/publications-data/risk-assessment-emergence-hypervirulent-klebsiella-pneumoniae-eu-eea> (accessed March 17, 2021).
- Fang, C. T., Chuang, Y. P., Shun, C. T., Chang, S. C., and Wang, J. T. (2004). A novel virulence gene in *Klebsiella pneumoniae* strains causing primary liver abscess and septic metastatic complications. *J. Exp. Med.* 199, 697–705. doi: 10.1084/jem.20030857
- Gu, D., Dong, N., Zheng, Z., Lin, D., Huang, M., Wang, L., et al. (2018). A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: A molecular epidemiological study. *Lancet Infect Dis.* 18, 37–46. doi: 10.1016/S1473-3099(17)30489-9
- Hao, Y., Zhao, X., Zhang, C., Bai, Y., Song, Z., Lu, X., et al. (2021). Clonal dissemination of clinical carbapenem-resistant *Klebsiella pneumoniae* isolates carrying fosA3 and bla KPC-2 cohabiting plasmids in shandong, china. *Front. Microbiol.* 12:771170. doi: 10.3389/fmicb.2021.771170
- Hennequin, C., Ravet, V., and Robin, F. (2018). Plasmids carrying DHA-1 β -lactamases. *Eur. J. Clin. Microbiol. Infect. Dis.* 37, 1197–1209. doi: 10.1007/s10096-018-3231-9
- ISO 20776-1 (2019). *Clinical laboratory testing and in vitro diagnostic test systems – Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices—Part 1: Broth micro-dilution reference method for testing the in vitro activity of antimicrobial agents against rapidly growing aerobic bacteria involved in infectious diseases*. Geneva: Switzerland.
- Lam, M. M. C., Wick, R. R., Watts, S. C., Cerdeira, L. T., Wyres, K. L., and Holt, K. E. (2021). A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat. Commun.* 12:4188. doi: 10.1038/s41467-021-24448-3
- Lee, C. R., Lee, J. H., Park, K. S., Jeon, J. H., Kim, Y. B., Cha, C. J., et al. (2017). Antimicrobial resistance of Hypervirulent *Klebsiella pneumoniae*: Epidemiology, hypervirulence-associated determinants, and resistance mechanisms. *Front. Cell Infect. Microbiol.* 7:483. doi: 10.3389/fcimb.2017.00483
- Lee, H. C., Chuang, Y. C., Yu, W. L., Lee, N. Y., Chang, C. M., Ko, N. Y., et al. (2006). Clinical implications of hypermucoviscosity phenotype in *Klebsiella pneumoniae* isolates: Association with invasive syndrome in patients with community-acquired bacteraemia. *J. Int. Med.* 259, 606–614. doi: 10.1111/j.1365-2796.2006.01641.x
- Li, W., Sun, G., Yu, Y., Li, N., Chen, M., Jin, R., et al. (2014). Increasing occurrence of antimicrobial-resistant hypervirulent (hypermucoviscous) *Klebsiella pneumoniae* isolates in China. *Clin. Infect. Dis.* 58, 225–232. doi: 10.1093/cid/cit675
- Liao, W., Liu, Y., and Zhang, W. (2020). Virulence evolution, molecular mechanisms of resistance and prevalence of ST11 carbapenem-resistant *Klebsiella pneumoniae* in China: A review over the last 10 years. *J. Glob. Antim. Resist.* 23, 174–180. doi: 10.1016/j.jgar.2020.09.004
- Lu, Y., Feng, Y., McNally, A., and Zong, Z. (2018). Occurrence of colistin-resistant hypervirulent *Klebsiella variicola*. *J. Antim. Chemother.* 73, 3001–3004. doi: 10.1093/jac/dky301
- Munoz-Price, L. S., Poirel, L., Bonomo, R. A., Schwaber, M. J., Daikos, G. L., Cormican, M., et al. (2013). Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis.* 13, 785–796. doi: 10.1016/S1473-3099(13)70190-7
- Örsten, S., Demirci-Duarte, S., Ünalán-Altıntop, T., Çakar, A., Sancak, B., Ergünay, K., et al. (2020). Low prevalence of hypervirulent *Klebsiella pneumoniae* in anatolia, screened via phenotypic and genotypic testing. *Acta Microbiol. Immunol. Hung.* 67, 120–126. doi: 10.1556/030.2020.01143
- Peirano, G., Pitout, J. D., Laupland, K. B., Meatherall, B., and Gregson, D. B. (2013). Population-based surveillance for hypermucoviscosity *Klebsiella pneumoniae* causing community-acquired bacteremia in calgary, alberta. *Can. J. Infect. Dis. Med. Microbiol.* 24, e61–e64. doi: 10.1155/2013/828741
- Piazza, A., Perini, M., Mauri, C., Comandatore, F., Meroni, E., Luzzaro, F., et al. (2022). Antimicrobial susceptibility, virulence, and genomic features of a hypervirulent serotype K2, ST65 *Klebsiella pneumoniae* causing meningitis in Italy. *Antibiotics (Basel)* 11:261. doi: 10.3390/antibiotics11020261
- Rodríguez-Medina, N., Barrios-Camacho, H., Duran-Bedolla, J., and Garza-Ramos, U. (2019). *Klebsiella variicola*: An emerging pathogen in humans. *Emerg. Microbes Infect.* 8, 973–988. doi: 10.1080/22221751.2019.1634981
- Rodríguez-Medina, N., Martínez-Romero, E., De la Cruz, M. A., Ares, M. A., Valdovinos-Torres, H., and Silva-Sánchez, J. (2020). A *klebsiella variicola* plasmid confers hypermucoviscosity-like phenotype and alters capsule production and virulence. *Front. Microbiol.* 11:579612. doi: 10.3389/fmicb.2020.579612
- Russo, T. A., and Marr, C. M. (2019). Hypervirulent *Klebsiella pneumoniae*. *Clin. Microbiol. Rev.* 32:e00001–e19. doi: 10.1128/CMR.00001-19
- Scaltriti, E., Piccinelli, G., Corbellini, S., Caruso, A., Latronico, N., and De Francesco, M. A. (2020). Detection of a hypermucoviscous *Klebsiella pneumoniae* co-producing NDM-5 and OXA-48 carbapenemases with sequence type 383, Brescia, Italy. *Int. J. Antim. Agents* 56:106130. doi: 10.1016/j.ijantimicag.2020.106130
- Shon, A. S., Bajwa, R. P., and Russo, T. A. (2013). Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: A new and dangerous breed. *Virulence* 4, 107–118. doi: 10.4161/viru.22718
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis.* 18, 318–327. doi: 10.1016/S1473-3099(17)30753-3
- Tang, M., Kong, X., Hao, J., and Liu, J. (2020). Epidemiological characteristics and formation mechanisms of multidrug-resistant hypervirulent *Klebsiella pneumoniae*. *Front. Microbiol.* 20:581543. doi: 10.3389/fmicb.2020.581543
- Taniguchi, Y., Maeyama, Y., Ohsaki, Y., Hayashi, W., Osaka, S., Koide, S., et al. (2017). Co-resistance to colistin and tigecycline by disrupting mgrB and ramR with IS insertions in a canine *Klebsiella pneumoniae* ST37 isolate producing SHV-12, DHA-1 and FosA3. *Int. J. Antimicrob. Agents* 50, 697–698. doi: 10.1016/j.ijantimicag.2017.09.011
- Wanford, J. J., Hames, R. G., Carreno, D., Jasiunaite, Z., Chung, W. Y., Arena, F., et al. (2021). Interaction of *Klebsiella pneumoniae* with tissue macrophages in a mouse infection model and ex-vivo pig organ perfusions: An exploratory investigation. *Lancet Microbe*. 2, e695–e703. doi: 10.1016/S2666-5247(21)00195-6
- Wang, Y. P., Chen, Y. H., Hung, I. C., Chu, P. H., Chang, Y. H., Lin, Y. T., et al. (2022). Transporter genes and fosA associated with fosfomycin resistance in carbapenem-resistant *Klebsiella pneumoniae*. *Front. Microbiol.* 13:816806. doi: 10.3389/fmicb.2022.816806
- Wick, R. R., Heinz, E., Holt, K. E., and Wyres, K. L. (2018). Kaptive web: User-friendly capsule and lipopolysaccharide serotype prediction for *Klebsiella* genomes. *J. Clin. Microbiol.* 56:e00197–18. doi: 10.1128/JCM.00197-18
- Wyres, K. L., Lam, M. M. C., and Holt, K. E. (2020). Population genomics of *Klebsiella pneumoniae*. *Nat. Rev. Microbiol.* 18, 344–359. doi: 10.1038/s41579-019-0315-1
- Wyres, K. L., Wick, R. R., Gorrie, C., Jenney, A., Follador, R., Thomson, N. R., et al. (2016). Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb. Genom.* 2:e000102. doi: 10.1099/mgen.0.000102
- Xu, M., Fu, Y., Fang, Y., Xu, H., Kong, H., Liu, Y., et al. (2019). High prevalence of KPC-2-producing hypervirulent *Klebsiella pneumoniae* causing meningitis in Eastern China. *Infect Drug Resist.* 12, 641–653. doi: 10.2147/IDR.S191892
- Zhang, Y., Jin, L., Ouyang, P., Wang, Q., Wang, R., Wang, J., et al. (2020). Evolution of hypervirulence in carbapenem-resistant *Klebsiella pneumoniae* in China: A multicentre, molecular epidemiological analysis. *J. Antim. Chemother.* 75, 327–336. doi: 10.1093/jac/dkz446
- Zhang, Y., Zhao, C., Wang, Q., Wang, X., Chen, H., Li, H., et al. (2016). High prevalence of hypervirulent *Klebsiella pneumoniae* infection in china: geographic distribution, clinical characteristics, and antimicrobial resistance. *Antimicrob. Agents Chemother.* 60, 6115–6120. doi: 10.1128/AAC.01127-16
- Zhu, J., Wang, T., Chen, L., and Du, H. (2021). Virulence factors in hypervirulent *Klebsiella pneumoniae*. *Front. Microbiol.* 8:642484. doi: 10.3389/fmicb.2021.642484



OPEN ACCESS

EDITED BY

Costas C. Papagiannitsis,
University of Thessaly,
Greece

REVIEWED BY

Marc Finianos,
Charles University,
Czechia
Zoi Florou,
University of Thessaly,
Greece

*CORRESPONDENCE

Wing-Cheong Yam
wcyam@hku.hk

[†]These authors have contributed equally to this study

SPECIALTY SECTION

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 21 June 2022

ACCEPTED 02 August 2022

PUBLISHED 09 September 2022

CITATION

Leung KS-S, Tam KK-G, Ng TT-L, Lao H-Y,
Shek RC-M, Ma OCK, Yu S-H, Chen J-X,
Han Q, Siu GK-H and Yam W-C (2022)
Clinical utility of target amplicon
sequencing test for rapid diagnosis of
drug-resistant *Mycobacterium tuberculosis*
from respiratory specimens.
Front. Microbiol. 13:974428.
doi: 10.3389/fmicb.2022.974428

COPYRIGHT

© 2022 Leung, Tam, Ng, Lao, Shek, Ma, Yu,
Chen, Han, Siu and Yam. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Clinical utility of target amplicon sequencing test for rapid diagnosis of drug-resistant *Mycobacterium tuberculosis* from respiratory specimens

Kenneth Siu-Sing Leung^{1†}, Kingsley King-Gee Tam^{1†},
Timothy Ting-Leung Ng^{2†}, Hiu-Yin Lao², Raymond Chiu-
Man Shek¹, Oliver Chiu Kit Ma³, Shi-Hui Yu⁴, Jing-Xian Chen⁵,
Qi Han⁵, Gilman Kit-Hang Siu² and Wing-Cheong Yam^{1*}

¹Department of Microbiology, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong SAR, China, ²Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China, ³KingMed Diagnostics, Science Park, Kwun Tong, Hong Kong SAR, China, ⁴Guangdong-Hong Kong-Macao Joint Laboratory of Respiratory Infectious Disease, Guangzhou, China, ⁵Guangzhou KingMed Diagnostics Group, Guangzhou, China

An in-house-developed target amplicon sequencing by next-generation sequencing technology (TB-NGS) enables simultaneous detection of resistance-related mutations in *Mycobacterium tuberculosis* (MTB) against 8 anti-tuberculosis drug classes. In this multi-center study, we investigated the clinical utility of incorporating TB-NGS for rapid drug-resistant MTB detection in high endemic regions in southeast China. From January 2018 to November 2019, 4,047 respiratory specimens were available from patients suffering lower respiratory tract infections in Hong Kong and Guangzhou, among which 501 were TB-positive as detected by in-house IS6110-qPCR assay with diagnostic sensitivity and specificity of 97.9 and 99.2%, respectively. Preliminary resistance screening by GenoType MTBDR_{plus} and MTBDR_{sl} identified 25 drug-resistant specimens including 10 multidrug-resistant TB. TB-NGS was performed using MiSeq on all drug-resistant specimens alongside 67 pan-susceptible specimens, and demonstrated 100% concordance to phenotypic drug susceptibility test. All phenotypically resistant specimens with dominating resistance-related mutations exhibited a mutation frequency of over 60%. Three quasiespecies were identified with mutation frequency of less than 35% among phenotypically susceptible specimens. They were well distinguished from phenotypically resistant cases and thus would not complicate TB-NGS results interpretations. This is the first large-scale study that explored the use of laboratory-developed NGS platforms for rapid TB diagnosis. By incorporating TB-NGS with our proposed diagnostic algorithm, the workflow would provide a user-friendly, cost-effective routine diagnostic solution for complicated TB cases with an average turnaround time of 6 working days. This is critical for timely management of drug resistant TB patients and expediting public health control on the emergence of drug-resistant TB.

KEYWORDS

Mycobacterium tuberculosis, next-generation sequencing, drug resistance, diagnosis, MDR-TB

Introduction

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (MTB) accounts for an estimated 10 million new cases and 1.5 million TB-related deaths annually (World Health Organization, 2021). This situation was further complicated with drug-resistant TB emergence. Globally, around 3.4% of new TB cases and 18% of previously treated cases were classified as multidrug-resistant TB (MDR-TB) with resistance to both rifampicin (RIF) and isoniazid (INH), or RIF mono-resistant (RR-TB) strains. Approximately 6.2% of global MDR-TB/RR-TB cases could be further classified as extensively drug-resistant (XDR-TB) strains with additional resistance to fluoroquinolones and at least one additional Group A anti-TB drug (bedaquiline or linezolid; World Health Organization, 2021). According to a national survey back in 2007, approximately 10% of TB cases in China were classified as MDR-TB (Zhao et al., 2012). To ensure a better control of the global pandemic, a rapid, accurate and cost-effective diagnostic platform for TB detection and subsequent drug resistance identification is required.

Previous studies have explored the use of next-generation sequencing (NGS) for TB drug resistance determination, and the platform has proved its capability as a diagnostic tool for TB outbreak control and drug susceptibility predictions in clinical isolates (Walker et al., 2017; Satta et al., 2018; Comin et al., 2020). While these studies have demonstrated the feasibility of NGS in TB drug resistance diagnosis, the sequencing platforms are either obsoleted (Colman et al., 2016), or the study only focused on AFB positive specimens with limited variability in drug resistance patterns (Smith et al., 2020). More importantly, as most platform focused on whole genome sequencing data, high-quality genomic DNA prepared from MTB isolates are usually required (Lam et al., 2021; Olawoye et al., 2021). The long turnaround time of MTB culture would not be in favor of rapid MTB resistance analysis. Recently, several retrospective studies have demonstrated the efficacy of deep amplicon sequencing for culture-free MTB drug resistance detection using Deeplex Myc-TB platform with excellent performance among selected AFB-smear-positive clinical specimens (Kayomo et al., 2020; Bonnet et al., 2021; Jouet et al., 2021). However, with the low sensitivity of AFB smear test, a large proportion of MTB positive specimens with lower bacterial load could be missed, thus reducing the feasibility for rapid diagnosis of most drug-resistant specimens in actual clinical settings. A robust platform which is suitable for both AFB smear-positive and negative respiratory specimens is therefore in need to implementing the platform into routine clinical practice.

Our previous study introduced a highly sensitive and specific in-house IS6110-qPCR protocol for routine TB

diagnosis with comparable performance to commercial platforms (Leung et al., 2018). In our practice, DNA extracts from IS6110-qPCR positive sputum were further tested for drug resistance-related mutation by Genotype MTBDR_{plus} and MTBDR_{sl} v2 (Hain Lifescience, Germany; MTBDR assay) for resistance-related mutation detection against RIF, INH, fluoroquinolones, and aminoglycosides. Sanger sequencing was performed on indeterminate results, such as specimens with absence of both wild-type probes and mutant probes (Lau et al., 2011; Siu et al., 2011a,b; Tam et al., 2019). Due to the high MDR-TB burden in our locality, false-resistant and false-susceptible cases could periodically be reported by line probe assay as a result of the vast variety of mutations in drug resistance-related genes (Javed et al., 2018). Although the discrepancies could be resolved by Sanger sequencing, the process is usually slow and labor intensive. To circumvent the problem, we recently developed a target amplicon sequencing protocol (TB-NGS) specifically targeting drug resistance-related genes in MTB isolates, which demonstrated complete concordance to conventional drug susceptibility testing (Tafess et al., 2020).

In order to reduce the turnaround time for drug-resistant MTB diagnosis, we extended TB-NGS into our rapid diagnostic platform to respiratory specimens. A total of 17 genes covering drug-resistant-related mutations of 8 classes of commonly used anti-TB drugs were selected, amplified and sequenced. As whole-genome data was not required by TB-NGS, data per sample could be greatly reduced and multiple specimens could be simultaneously analyzed in a single flow cell. The overall cost and turnaround time could be significantly reduced with the capability of direct application on respiratory specimen. In addition, we combined TB-NGS with our previously published IS6110-qPCR MTB diagnostic assay (Leung et al., 2018). The new routine diagnostic pipeline allows medium- to large-scale sample screening in daily clinical practice. For complicated TB cases such as retreatment cases and patients refractory to first-line treatment, TB-NGS could be performed to provide extra guidance.

In this study, we focused on the clinical utility of incorporating TB-NGS platform into routine settings. We performed a large-scale evaluation of TB-NGS workflow among respiratory specimens. Using respiratory specimens collected in Hong Kong and Guangzhou, the diagnostic performance of TB-NGS was assessed and compared with reference to the best available reference standard in our study cohort.

Materials and methods

Specimen collection and processing

From January 2018 to November 2019, respiratory specimens were collected from patients suffering from lower respiratory infections from 13 local chest clinics and 3 public hospitals in Hong Kong. Similarly, respiratory specimens were available from KingMed Diagnostic Centre for patients

Abbreviations: TB, Tuberculosis; MTB, *Mycobacterium tuberculosis*; MDR-TB, Multidrug-resistant TB; RIF, Rifampicin; INH, Isoniazid; RR-TB, RIF mono-resistant; TB-NGS, target amplicon sequencing protocol; NALC, N-Acetyl-L-Cysteine; LJ, Lowenstein-Jensen; MTBC, *Mycobacterium tuberculosis* complex; RRDR, RIF resistance determining region; DOTS, Directly observed short course treatment; NGS, Next-generation sequencing.

suspected of chest infections in Guangzhou, China. Each sputum specimen was decontaminated by N-Acetyl-L-Cysteine (NALC)-NaOH method with a final NaOH concentration of 2%. Approximately 500 µl processed sediments were inoculated onto Lowenstein-Jensen (LJ) slant (BioMérieux, Marcy l'Etoile, France) and BACTEC™ MGIT™ 960 Mycobacterial Detection System (Becton Dickinson, Baltimore, MD, United States), and specimen DNA were extracted from 500 µl processed sediments as previously described (Chan et al., 1996), and the identity of the positive AFB culture were confirmed by 16S rRNA sequencing (Yam et al., 2006).

Crude DNA extraction, MTBC identification and line probe assay

For each specimen, crude DNA samples were extracted by AMPLICOR® Respiratory Specimen Preparation kit (Roche Lifescience, Germany). *Mycobacterium tuberculosis* complex (MTBC) detection was performed by IS6110-qPCR assay as previously described (Leung et al., 2018). Positive IS6110-qPCR was regarded when Ct value was less than 24.14. IS6110-qPCR positive specimens were further screened by MTBDR assay according to manufacturer's instructions.

Target amplification and preparation

Suspected genotypic resistant specimens were selected for TB-NGS. Specimens were purified by Agencourt AMPure XP PCR kit (Beckman Coulter, United States) in crude DNA extract to bead solution ratio of 1:1.8. Purified DNA was resuspended in 25 µl DNase-free water and stored at −80°C prior to usage. A total of 17 drug resistance-related gene targets (Table 1) were selected and amplified in five multiplex reaction mixes with primers listed in our previous study (Supplementary material; Tafess et al., 2020). A DNase-free water sample was added to each run as a negative control to minimize cross contamination.

Amplicons were purified again by AMPure XP PCR kit according to manufacturer's instructions. DNA quantity was measured by Qubit dsDNA HS assay kit. Equal mass of DNA from each reaction mix was added to a new microcentrifuge tube.

Library preparation and next-generation sequencing

Library was prepared using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs, United States) according to manufacturer's instructions. Approximately 100 ng of the mixed purified amplicon was inputted, tagged and barcoded. MiSeq was performed using MiSeq Reagent Kit v2 Nano kit, which generates 500Mbp output

TABLE 1 Gene targets for target amplicon sequencing.

Drug	Gene	Amplicon size (bp)	Major mutation covered*
Rifampicin	<i>rpoB</i> -RRDR ¹	288	<i>rpoB</i> RRDR
	<i>rpoB</i> -full ¹	1,311	<i>rpoB</i> RRDR, V146F, I572F
Isonizid	<i>katG</i>	435	<i>katG</i> S315T
	<i>inhA</i>	454	<i>inhA</i> C-15 T
	promoter		
	<i>inhA</i>	922	<i>inhA</i> codon 94 and 95
	<i>katG</i> - <i>furA</i>	892	<i>furA</i> codon 4, −134 bp upstream deletion
	intergenic region		
Ethambutol	<i>embB</i>	955	<i>embB</i> M306V/I, G406A/D/S
	<i>ubiA</i>	1,119	Compensatory mutation
Pyrazinamide	<i>pncA</i>	813	Entire gene
	<i>rpsA</i>	1,601	Entire gene
Fluoroquinolones	<i>gyrA</i>	751	90–94 QRDR
	<i>gyrB</i>	1,054	N538D/E540V
Aminoglycosides	<i>eis</i> promoter	593	C-14 T, C-12 T, G-10A, G-10C
	<i>rpsL</i>	472	K43R/K88Q
	<i>rrs</i>	1,211	C1400, A1401/ C1483
			R3*; Q22*
Capreomycin	<i>tylA</i>	945	Loss of <i>tlyA</i> expression
Linezolid	<i>rplC</i>	710	G2061T/G2576T
	<i>rrl</i>	1,102	T460C

¹Two sets of primers were designed for *rpoB* to ensure a better coverage at 81 bp RIF resistance determining regions.

*In accordance with 03–2018 literature review database.

in 2×250bp format. Each flow cell generates sequences for 24 samples.

Bioinformatic analysis

To identify mutations from target amplicon sequencing data, fastq files were uploaded onto Bacteriochek (Advanced Biological Laboratories S.A., Luxembourg), a CE-IVD online platform for resistance detection. Minor variant mutations cutoffs were set at 20% and resistance-related mutation patterns were identified in accordance with 03–2018 literature review database. Successful sequencing results were defined according to the default criteria set by Bacteriochek followed by sufficient sequencing depth (≥500X) in *rpoB*, *katG*, *mabA*, *inhA*, *furA*, *gyrA*, *gyrB*, *rrs* and *eis* genes. These genes were selected as the surrogate markers for TB-NGS due to their importance in identifying MDR-TB and XDR-TB cases.

Phenotypic DST

All MTBC culture-positive specimens were subjected to phenotypic drug susceptibility test (DST) using MGIT tubes with recommended critical concentration according to World Health Organization guidelines (World Health Organization, 2018). The critical concentrations for RIF, INH, ethambutol, streptomycin, pyrazinamide (PZA), kanamycin, amikacin, capreomycin, levofloxacin, moxifloxacin, linezolid are 1.0 µg/ml, 0.1 µg/ml, 5.0 µg/ml, 1.0 µg/ml, 100 µg/ml at pH 5.6, 2.5 µg/ml, 1.0 µg/ml, 2.5 µg/ml, 1.0 µg/ml, 0.25 µg/ml and 1.0 µg/ml, respectively. In brief, a loopful of mycobacterial colonies was suspended in 4 ml of BBL Middlebrook 7H9 Broth (Becton Dickinson, Baltimore, MD, United States) containing 8–10 glass beads. The suspension was vortexed for 2–3 min to ensure a homogenized suspension. The suspension was then adjusted to a 0.5 McFarland turbidity standard. A total of 1 ml of the adjusted suspension was diluted into 4 ml sterile saline (referred as DST inoculum). Subsequently, a 1:10 growth control suspension for PZA and a 1:100 Growth control suspension for other antibiotics was prepared. Standard inoculum of both GC and DST inoculum was aseptically pipetted into a drug free MGIT tube and a drug containing MGIT tubes, respectively. The tubes were mixed thoroughly by gentle inversion for 3 to 4 times. They were then being inserted into a MGIT DST set carrier and placed into MGIT 960 instrument to incubate at 37°C for a maximum of 21 days with the exception of PZA being 14 days.

Statistical analysis

IS6110-qPCR assay results were primarily compared with AFB culture results. Specimens with positive IS6110-qPCR assay results but negative AFB cultures were resolved by additional culture result from specimens collected on consecutive days. Indeterminate results were classified as cases with the absence of both IS6110 and IC amplification. These indeterminate samples were enumerated but were excluded from sensitivity and specificity calculations. Sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV) were calculated with reference to the diagnostic results based on bacteriological and clinical information. TB-NGS results were compared primarily with reference to phenotypic DST. For minor variant cases, patients were closely monitored for another 6 months. Sequencing successful rate, result concordance and associated mutations were analyzed accordingly. Turnaround time of this study was calculated from sample collection date to report date.

Results

Specimen collection and preliminary screening result

From January 2018 to November 2019, 4,047 respiratory specimens from 2,627 patients were collected in Hong Kong and

Guangzhou. Among these specimens, 473/4047 (11.7%) specimens showed positive MTBC culture, with 93/473 (19.7%) being AFB smear-positive. Mean turnaround time from sample collection to phenotypic DST was 50 days.

IS6110-qPCR

Our previous study confirmed the cutoff Ct value for IS6110-qPCR assay was optimal at 24.14. Using Ct = 24.14 as cut off, 501/4047 (12.3%) specimens were identified as MTBC positive by IS6110-qPCR assay with Ct ranged from 4.17–24.13. The remaining 3454/4047 (87.6%) were reported as MTB negative by IS6110-qPCR assay. A total of 92 specimens were reported as indeterminate result with both negative IS6110 and IC probe.

Using AFB culture as reference standard, IS6110-qPCR successfully detected 463/473 (97.9%) specimens with positive MTBC culture. For the remaining MTBC culture-positive specimens with negative IS6110-qPCR results, they showed positive IS6110 amplification with Ct ranged from 25.71 to 29.48. They were still reported as MTBC negative as the Ct value were greater than the cutoff Ct value. It is noted that 28 specimens showed positive IS6110-qPCR amplification but with negative AFB culture results. While these 28 specimens could be resolved with additional clinical information such as chest radiograph abnormalities compatible with pulmonary TB or other demographic features such as responses to anti-TB treatment and past TB history, they were still regarded as false positive cases due to negative AFB culture and the unavailability of culture isolated for further confirmation by 16S rRNA sequencing. Sensitivity and specificity of IS6110-qPCR assay were hence calculated at 100 and 100% for AFB specimens, and 97.4 and 99.2% for AFB smear negative specimens, respectively (Table 2).

MTBDR assay

All IS6110-qPCR positive specimens were tested by MTBDR assay with 438/501 (87.4%) specimens achieving interpretable results. The remaining 63/501 (12.6%) specimens showed uninterpretable result due to the absence of control TB probe. Among the positive specimens, 413/438 (94.5%) were reported with wild-type genotype, while mutant probe was detected in 25/438 (5.5%) specimens (Supplementary material).

TB-NGS and concordance

TB-NGS was performed on all 25 mutant specimens detected by MTBDR assay. To investigate the efficacy of TB-NGS on wild-type genotypes, 67 wild-type specimens were randomly selected with IS6110-qPCR Ct ranged from 6.35–24.1, yielding a total of 92 specimens enrolled for TB-NGS analysis. Using the criteria, we defined, 83/92 (90.2%) specimens were successfully sequenced. The average sequencing depth for each specimen is 22,500X (95%

TABLE 2 Diagnostic performance of IS6110-qPCR among 4,047 respiratory specimens using cutoff Ct value at 24.14.

AFB Smear	MTBC culture	IS6110 qPCR			Resolved performance (%) [95% CI]			
		Positive	Negative	Indeterminate	Sensitivity	Specificity	PPV	NPV
Positive (<i>n</i> = 97)	MTBC Positive (<i>n</i> = 93) ^a	93	0	0	100	100	100	100
	MTBC Negative (<i>n</i> = 4)	0	0	4				
Negative (<i>n</i> = 3,950)	MTBC Positive (<i>n</i> = 370) ^a	360	10	0	97.3 [95.1–98.7]	99.2 [98.8–99.5]	92.8 [89.9–94.9]	99.7 [99.5–99.8]
	MTBC Negative (<i>n</i> = 3,580)	28	3,464	88				
	Clinical TB cases (<i>n</i> = 28) ^b	28	0	0				
	Culture negative (<i>n</i> = 3,552)	0	3,376	88 ^c				

^aMTBC culture-positive specimens from original or subsequent specimens collected in consecutive days.

^bClinical TB cases were defined as patients with abnormal chest radiographs compatible with pulmonary TB, or supported by other demographic features such as response to anti-TB treatment or patients with past TB history.

^cIndeterminate cases were yielded due to the absence of both IS6110 and internal control amplification.

confidence interval: 20660X, 24,400X). Resistant-related mutations were found in 30 specimens with mutation patterns listed in [Table 3](#) and [Supplementary material](#).

TB-NGS demonstrated 100% concordance among all 25 drug-resistant strains. Among these strains, one INH mono-resistant specimen (WC14) defined by MTBDR assay was confirmed as MDR-TB by TB-NGS with *rpoB* I572F mutation located outside RIF resistance determining region (RRDR). The results were further confirmed by phenotypic DST as MDR-TB. One genotypic FQs mono-resistant specimen (WC10) was also identified with *gyrA* D94G missense mutation and had concordant result between MTBDR assay, TB-NGS and phenotypic DST. Meanwhile, 5 additional wild-type strains defined by MTBDR assay were found with drug resistance-related mutations in TB-NGS. Among these five specimens, two (WC26 and WC27) were genotypically STR mono-resistant with *rpsL* K43R. The remaining three specimens were defined as minor variants by TB-NGS but were considered as wild-type strains by MTBDR assay. Two *rpoB* mutations D516V (WC30) and H526R (WC29) were separately identified in two specimens with a low mutation frequency of 31.55 and 20.94%, respectively ([Table 3](#)). One specimen (WC28) was found to carry *gyrA* mutation A74S with mutation frequency of 35.90%. Consistent results were obtained in technical replicates, which confirmed the presence of the mutation despite at a low frequency.

Phenotypic DST

A total of 68/92 specimens enrolled in TB-NGS study yielded MTBC culture-positive specimens and were subjected to phenotypic DST ([Supplementary material](#)). Among these 68 specimens, 6 phenotypically pan-susceptible strains yielded

indeterminate TB-NGS results and were excluded from subsequent analysis. In the remaining 62 strains, 12 were drug-resistant TB with resistance to one or more anti-TB drugs, 4 were MDR-TB, 6 were pre-XDR-TB, and 40 were pan-susceptible strains ([Table 4](#)). All remaining 408 MTBC culture-positive specimens were defined as pan-susceptible by phenotypic DST.

Using phenotypic DST as reference standard, TB-NGS showed concordant results in 59/62 specimens. The three discordant cases were pan-susceptible by phenotypic DST ([Table 4](#), Footnote a–c) and was identified by TB-NGS as minor variant cases with a mutation frequency less than 35%. These patients exhibited no signs of drug resistance development during 6 month follow-up sessions, and were subsequently discharged after directly observed short course treatment (DOTS).

Discussion

Conventional TB diagnosis is confirmed fundamentally on AFB culture and subsequent phenotypic DST. However, the entire procedures from sample collection to phenotypic DST require at least 5–12 weeks. Due to the long turnaround time of MTB culture, results are usually unavailable to clinicians for treatment guidance when starting empirical treatments. The long turnaround time is more evident among drug-resistant specimens due to the even slower growth of resistant MTB strains. Molecular approaches such as line-probe assays are timely and reliable genotypic DST methods, but increasing number of studies showed that false resistance could sometimes be reported due to the presence of naturally occurring polymorphisms ([Coll et al., 2015; Hu et al., 2019](#)). In addition, the presence of uncommon mutation might also result in discordant genotypic and phenotypic drug susceptibility results. For instance, MTB strains with *rpoB* L533P

TABLE 3 Summary of drug-resistant mutations detected by TB-NGS.

Drugs	Gene	Mutation patterns	No. of specimens	Mutation frequency (±SD)
Rifampicin	<i>rpoB</i>	D516V	1	99.22
		D516Y ^a	1	31.55
		S531L	7	84.97 (±17.28)
		H526N ^c	2	95.91 (±1.50)
		H526L	1	100
		H526R	1	66.96
		H526R ^b	1	20.94
		I572L ^c	2	98.28 (±1.69)
		I572F	1	98.5
Isoniazid	<i>katG</i>	S315T	16	98.67 (±1.96)
		S315N	2	98.70 (±0.01)
	<i>mabA</i>	c-15 ^d	2	99.70 (±0.43)
	<i>inhA</i>	I194T ^d	1	98.5
	<i>furA-inhA</i>	None	N/A	N/A
Ethambutol	<i>embB</i>	M306V	3	99.69 (±0.26)
		M306I	3	89.05 (±15.35)
		G406S	3	99.45 (±0.95)
		G406C	1	98.87
	<i>ubiA</i>	None	N/A	N/A
Pyrazinamide	<i>pncA</i>	V139G	2	98.665 (±0.66)
	<i>rpsA</i>	None	N/A	N/A
Fluoroquinolones	<i>gyrA</i>	A74S ^e	1	35.90
		A90V	2	98.55 (±0.32)
		D94G	3	80.10 (±16.66)
		D94N	2	100.0 (±0.01)
	<i>gyrB</i>	None	N/A	N/A
Streptomycin	<i>rpsL</i>	K43R	6	98.36
		K88R	2	98.87 (±0.04)
	<i>rrs</i>	c513t	1	99.63
Aminoglycosides	<i>rrs</i>	a1401g	1	100.0
	<i>eis</i>	None	N/A	N/A
Capreomycin	<i>tlyA</i>	None	N/A	N/A
Linezolid	<i>rrl</i>	None	N/A	N/A
	<i>rplC</i>	None	N/A	N/A

Abbreviation: SD: standard deviation; N/A, not available.

^aOne strain (WC30) carried a minor variant *rpoB* D516Y with low mutation frequency at 31.55% as detected by TB-NGS.

^bOne strain (WC29) carried a minor variant *rpoB* H526R with low mutation frequency at 20.94% as detected by TB-NGS.

^cTwo strains carried both *rpoB* H526N and I572L mutation. Phenotypic DST result later showed that the two strains were phenotypically resistant to RIF.

^dOne strain carried both *mabA* c-15t and *inhA* I194T mutation. Phenotypic DST result later revealed that the strain was phenotypically resistant to INH.

^eOne strain (WC28) carried a minor variant *gyrA* A74S with low mutation frequency at 35.9% as detected by TB-NGS.

mutation might exhibit slightly elevated RIF Minimum Inhibitory concentration compared to fully susceptible strains but remain phenotypically susceptible when RIF critical concentration at 1 µg/ml was performed in MGIT SIRE as recommended by WHO guidelines (Tam et al., 2017; Shea et al., 2021). Similarly, while

mutation at promoter region of *inhA* is generally believed to cause low-level INH resistance, concomitant mutation at *inhA* I194T could drastically increase INH resistance, rendering the drug useless even when a higher INH dosage was used (Machado et al., 2013). A rapid and reliable diagnostic method that reveals the identity of drug resistance-related mutation is thus urgently needed to fully explore the identity of the mutation.

The basis of TB-NGS was developed according to our previous study, which has demonstrated excellent reliability in identifying drug resistance-related mutations among culture isolates (Tafess et al., 2020). To extend the test to respiratory specimens, several modifications were implemented. First of all, instead of multiplex PCR of 17 gene regions in a single-tube, target amplification was separated into five multiplex PCR with each tube amplifying 4–5 regions of similar sizes to improve the amplification efficiency. To improve the sensitivity of TB-NGS especially on RIF, an additional target at *rpoB* covering RRDR and codon I572 was included. With these modifications, TB-NGS successfully sequenced 83/92 (90.2%) of the specimens enrolled in this study with 100% concordance when compared with phenotypic DST. For the remaining 9 specimens with unsuccessful TB-NGS results, they were all AFB smear-negative specimens with a Ct value ranged from 19.57–24.1. Meanwhile, most of the sequenced specimens have IS6110 Ct value of less than 18, indicating that the detection limit of TB-NGS would be equivalent to approximately 250 CFU/ mL according to the LOD analysis performed in our previous study (Leung et al., 2018). The higher demand in bacterial load might also explain the nine specimens with positive IS6110-qPCR but indeterminate TB-NGS results in this study.

Based on the TB-NGS result several uncommon mutation patterns were identified. One specimen (WC14) exhibited concurrent INH-resistance-related mutations at both *inhA* regulatory region c-15t and structural region I194T with mutation frequency of 99.39 and 98.50%, respectively. Phenotypic DST revealed that the specimen was resistant to INH at 0.4 µg/ml, indicating high-level INH resistance instead of the low-level resistance commonly reflected by single mutation at *inhA* regulatory site. In addition, two MDR-TB specimens from the same patient (WC11 and WC12) exhibited concurrent *rpoB* H526N and I572L mutations with mutation frequency of >95%. Mutation at *rpoB* H526N is normally regarded as a natural polymorphism (Jamieson et al., 2014), while *rpoB* I572L is a RIF resistant-related mutation located outside RRDR and is commonly not recognized by GeneXpert MTB/RIF Ultra and MTBDR_{plus} assay. Similarly, one INH mono-resistant specimen defined by MTBDR assay was confirmed as MDR-TB by TB-NGS with *rpoB* I572F mutation. For PZA resistance detection, two PZA resistant specimen from one patient were identified in our study cohort. TB-NGS showed that the strain carried *pncA* V139G mutation which is a PZA resistance-related mutation commonly found in our locality as previously reported (Tam et al., 2019). The resistance profiles identified by TB-NGS in this study highlighted the strength of our diagnostic platform. Being an unbiased

TABLE 4 Phenotypic DST profiles of MTBC culture-positive strains from Hong Kong and Guangzhou by phenotypic DST in this study.

Specimen type	Resistance pattern defined by TB-NGS	No. of specimens	Phenotypic drug susceptibility patterns									
			culture (+)	RIF	INH	EMB	PZA	STR	FLQ	CAP	AMI	LZD
Genotypic drug resistance defined by MTBDR assay (<i>n</i> = 25)	Drug-resistant TB (<i>n</i> = 15)	2	1	R	S	S	S	S	S	S	S	S
		1	1	R	S	S	S	R	S	S	S	S
		5	2	S	R	S	S	S	S	S	S	S
		1	1	S	R	R	S	S	S	S	S	S
		1	1	S	R	R	S	R	S	S	R	S
		4	3	S	R	S	S	R	S	S	S	S
		1	1	S	S	S	S	S	R	S	S	S
	MDR-TB (<i>n</i> = 4)	2	2	R	R	R	R	S	S	S	S	S
		2	2	R	R	S	S	R	S	S	S	S
	Pre-XDR-TB (<i>n</i> = 6)	2	2	R	R	R	S	S	R	S	S	S
Drug resistance not detected by MTBDR assay (<i>n</i> = 476)	Drug-resistant TB (<i>n</i> = 2)	4	4	R	R	R	S	R	R	S	S	S
	Minor variant (<i>n</i> = 3) ^{a, b, c}	2	2	S	S	S	S	R	S	S	S	S
	Pan-susceptible TB (<i>n</i> = 53)	3	3	S	S	S	S	S	S	S	S	S
		53	37	S	S	S	S	S	S	S	S	S
	Unsuccessful sequencing (<i>n</i> = 9)	9	6	S	S	S	S	S	S	S	S	S
	TB-NGS not performed (<i>n</i> = 409)	409	405	S	S	S	S	S	S	S	S	S

Abbreviation: RIF, rifampicin; INH, isoniazid; EMB, ethambutol; PZA, pyrazinamide; FLQ, fluoroquinolones; STR, streptomycin; AMIs, aminoglycosides; CAP, capreomycin; LZD, linezolid^aWC29 carried a minor variant at *rpoB* H526R with mutation frequency of 20.94% by TB-NGS. Phenotypic DST result showed that the strain was susceptible to RIF.

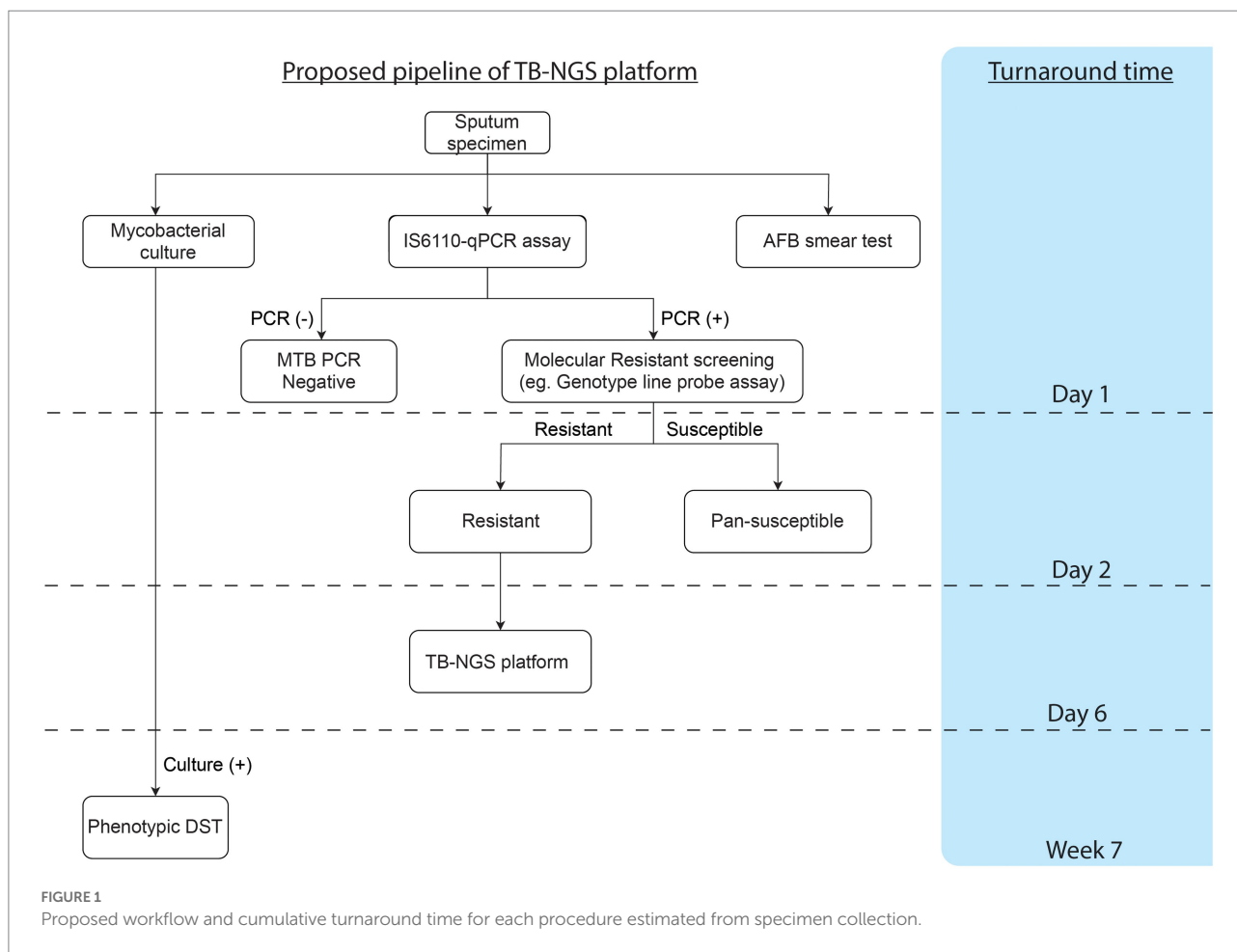
^bWC30 carried a minor variant at *rpoB* D516Y with mutation frequency of 31.55% by TB-NGS. Phenotypic DST result showed that the strain was susceptible to RIF.

^cWC28 carried a minor variant at *gyrA* A74S with mutation frequency of 35.90% by TB-NGS. Phenotypic DST result showed that the strain was susceptible to FQs.

platform for identifying drug resistance mutations, TB-NGS could provide a much more comprehensive analysis on the resistance pattern of each specimen with a much shorter turnaround time. This would facilitate a judicious use of antibiotics especially in the treatment of MDR-TB.

In this study, three quasispecies were identified among wildtype specimens, of which two were *rpoB* mutations that could be correlated to phenotypic RIF resistance. These minor variants were consistent in multiple technical replicates, suggesting that these minor variants were unlikely to be sequencing artifacts. Phenotypic DST on these three quasispecies revealed pan-susceptible phenotypes, and the patients fully recovered after 6-months DOTS follow-up without complications. With a mutation frequency of less than 35%, these minor variants could be well distinguished from dominant resistant mutations which usually carried a mutation frequency of >60%. Therefore, these minor variants would not complicate empirical clinical judgment. Nevertheless, to fully elucidate the effect of these minor variants, a large-scale longitudinal study could be required to determine the progression of minor variant resistance developments.

TB-NGS generally requires a minimum turnaround time of 6 working days. Depending on the Illumina flow cell being used, TB-NGS can accommodate 12–96 samples per single run, which would fit low to high throughput analysis. Furthermore, with Bacteriochek having CE-IVD systems, the assay would be well suited in most clinical laboratories. While the platform is suitable for routine surveillance in low endemic regions in developed countries, a cost-effective protocol is highly recommended for high endemic regions as demonstrated in our study (Figure 1). Preliminary screening can first be performed for MTBC detection, followed by molecular genotypic resistance detection. For cases with resistant mutation or upon clinical suspicions, such as patients unresponsive to treatment or requiring additional support among treatment defaulters and retreatment cases, TB-NGS can be performed to provide a comprehensive drug resistance profile. Applying this algorithm to our study, only 25 out of the 4,000 specimens would require additional aid from TB-NGS. More importantly, our diagnostic algorithm missed no drug-resistant case as indicated by phenotypic DST results on all MTBC culture-positive specimens. The implementation of our proposed algorithm



would therefore greatly reduce the workload and financial burden in clinical microbiology laboratories, especially in high endemic regions with limited resources.

Despite the potential of TB-NGS for better MTB management support, the assay has its limitations. The current platform is based on the knowledge of resistance-related mutation. For novel mutations located outside targeted sites they will not be detected. Nevertheless, the current test has included all well-established resistance-conferring mutations which would have been suitable for normal routine clinical practice. On the other hand, while the guideline updated in late 2020 included bedaquiline as one of the medications for XDR-TB definition, our current TB-NGS protocol did not include bedaquiline genotypic resistance test due to several reasons. As a new anti-TB drug recommended by WHO in 2016, the exact mechanism for bedaquiline resistance pattern in clinical isolates are still far from fully elucidated (Nguyen et al., 2018; Nieto Ramirez et al., 2020; Gomez-Gonzalez et al., 2021). In addition, the occurrence of bedaquiline resistance TB strains in our locality is still extremely rare. Among 4,043 specimens collected in this study, only 6 pre-XDR-TB specimens were identified while still susceptible to PZA, LZD and aminoglycosides. Genotypic resistance test would still

be necessary in further development of TB-NGS platform due to increasing use of bedaquiline in complicated TB cases as well as the emergence of bedaquiline resistant strains (Chesov et al., 2021). Lastly, while other sequencing platforms such as Nanopore might be a more rapid and budget-friendly approach, we chose Illumina as the primary sequencing platform due to its high single read accuracy for the investigation of minor variants in our study cohort. Evaluation of TB-NGS by Nanopore in a multi-center analysis is therefore warranted to further extend the utility of TB-NGS for more flexible and cost-effective routine service.

To conclude, TB-NGS is an accurate platform for TB drug resistance detection. When accompany with our proposed diagnostic algorithm, the workflow would provide a cost-effective diagnostic solution for complicated TB cases and allow early detection of specimens with potential drug resistance developments.

Data availability statement

The data presented in this study are deposit in NCBI database, Accession Number SAMN30429057-SAMN30429139.

Ethics statement

This study has been approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (Ref. number: UW 12-309).

Author contributions

KL, KT, TN, and GS were responsible for designing the study, performing the experiments and writing the manuscript. KL was responsible for bioinformatic and statistical analysis of this study. H-YL and RS were responsible for doing the experiments. OM, S-HY, J-XC, and QH were responsible for providing processing specimens from mainland China. W-CY is the corresponding author and was responsible of supervising the study. All authors contributed to the article and approved the submitted version.

Funding

This study is supported by the Science and Technology Planning Project of Guangdong Province, China (grant number: 2019B121205010).

References

- Bonnet, I., Enouf, V., Morel, F., Ok, V., Jaffre, J., Jarlier, V., et al. (2021). A comprehensive evaluation of GeneLEAD VIII DNA platform combined to Deeplex Myc-TB(R) assay to detect in 8 days drug resistance to 13 Antituberculous drugs and transmission of mycobacterium tuberculosis complex directly From clinical samples. *Front. Cell. Infect. Microbiol.* 11:707244. doi: 10.3389/fcimb.2021.707244
- Chan, C. M., Yuen, K. Y., Chan, K. S., Yam, W. C., Yim, K. H., Ng, W. F., et al. (1996). Single-tube nested PCR in the diagnosis of tuberculosis. *J. Clin. Pathol.* 49, 290–294. doi: 10.1136/jcp.49.4.290
- Chesov, E., Chesov, D., Maurer, F. P., Andres, S., Utpatel, C., Barilar, I., et al. (2021). Emergence of bedaquiline-resistance in a high-burden country of tuberculosis. *Eur. Respir. J.* 57:2002544. doi: 10.1183/13993003.00621-2021
- Coll, F., McNeerney, R., Preston, M. D., Guerra-Assuncao, J. A., Warry, A., Hill-Cawthorne, G., et al. (2015). Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* 7:51. doi: 10.1186/s13073-015-0164-0
- Colman, R. E., Anderson, J., Lemmer, D., Lehmkuhl, E., Georgiou, S. B., Heaton, H., et al. (2016). Rapid drug susceptibility testing of drug-resistant mycobacterium tuberculosis isolates directly from clinical samples by use of amplicon sequencing: a proof-of-concept study. *J. Clin. Microbiol.* 54, 2058–2067. doi: 10.1128/JCM.00535-16
- Comin, J., Chaure, A., Cebollada, A., Ibarz, D., Vinuelas, J., Vitoria, M. A., et al. (2020). Investigation of a rapidly spreading tuberculosis outbreak using whole-genome sequencing. *Infect. Genet. Evol.* 81:104184. doi: 10.1016/j.meegid.2020.104184
- Gomez-Gonzalez, P. J., Perdigo, J., Gomes, P., Puyen, Z. M., Santos-Lazaro, D., Napier, G., et al. (2021). Genetic diversity of candidate loci linked to mycobacterium tuberculosis resistance to bedaquiline, delamanid and pretomanid. *Sci. Rep.* 11:19431. doi: 10.1038/s41598-021-98862-4
- Hu, P., Zhang, H., Fleming, J., Zhu, G., Zhang, S., Wang, Y., et al. (2019). Retrospective analysis of false-positive and disputed rifampin resistance Xpert MTB/RIF assay results in clinical samples from a referral Hospital in Hunan, China. *J. Clin. Microbiol.* 57:4. doi: 10.1128/JCM.01707-18
- Jameson, F. B., Guthrie, J. L., Neemuchwala, A., Lastovetska, O., Melano, R. G., and Mehaffy, C. (2014). Profiling of rpoB mutations and MICs for rifampin and

Conflict of interest

OM, S-HY, J-XC and QH were employed by KingMed Diagnostics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.974428/full#supplementary-material>

rifabutin in mycobacterium tuberculosis. *J. Clin. Microbiol.* 52, 2157–2162. doi: 10.1128/JCM.00691-14

Javed, H., Bakula, Z., Plen, M., Hashmi, H. J., Tahir, Z., Jamil, N., et al. (2018). Evaluation of genotype MTBDRplus and MTBDRsl assays for rapid detection of drug resistance in extensively drug-resistant mycobacterium tuberculosis isolates in Pakistan. *Front. Microbiol.* 9, 2265. doi: 10.3389/fmicb.2018.02265

Jouet, A., Gaudin, C., Badalato, N., Allix-Beguec, C., Duthoy, S., Ferre, A., et al. (2021). Deep amplicon sequencing for culture-free prediction of susceptibility or resistance to 13 anti-tuberculous drugs. *Eur. Respir. J.* 57:2002338. doi: 10.1183/13993003.02338-2020

Kayomo, M. K., Mbula, V. N., Aloni, M., Andre, E., Rigouts, L., Boutachkourt, F., et al. (2020). Targeted next-generation sequencing of sputum for diagnosis of drug-resistant TB: results of a national survey in Democratic Republic of the Congo. *Sci. Rep.* 10:10786. doi: 10.1038/s41598-020-67479-4

Lam, C., Martinez, E., Crichton, T., Furlong, C., Donnan, E., Marais, B. J., et al. (2021). Value of routine whole genome sequencing for mycobacterium tuberculosis drug resistance detection. *Int. J. Infect. Dis.* 113, S48–S54. doi: 10.1016/j.ijid.2021.03.033

Lau, R. W., Ho, P. L., Kao, R. Y., Yew, W. W., Lau, T. C., Cheng, V. C., et al. (2011). Molecular characterization of fluoroquinolone resistance in mycobacterium tuberculosis: functional analysis of gyrA mutation at position 74. *Antimicrob. Agents Chemother.* 55, 608–614. doi: 10.1128/AAC.00920-10

Leung, K. S., Siu, G. K., Tam, K. K., Ho, P. L., Wong, S. S., Leung, E. K., et al. (2018). Diagnostic evaluation of an in-house developed single-tube, duplex, nested IS6110 real-time PCR assay for rapid pulmonary tuberculosis diagnosis. *Tuberculosis (Edinb.)* 112, 120–125. doi: 10.1016/j.tube.2018.08.008

Machado, D., Perdigo, J., Ramos, J., Couto, I., Portugal, I., Ritter, C., et al. (2013). High-level resistance to isoniazid and ethionamide in multidrug-resistant mycobacterium tuberculosis of the Lisboa family is associated with inhA double mutations. *J. Antimicrob. Chemother.* 68, 1728–1732. doi: 10.1093/jac/dkt090

Nguyen, T. V. A., Anthony, R. M., Banuls, A. L., Nguyen, T. V. A., Vu, D. H., and Alffenaar, J. C. (2018). Bedaquiline resistance: its emergence, mechanism, and prevention. *Clin. Infect. Dis.* 66, 1625–1630. doi: 10.1093/cid/cix992

Nieto Ramirez, L. M., Quintero Vargas, K., and Diaz, G. (2020). Whole genome sequencing for the analysis of drug resistant strains of mycobacterium tuberculosis.

System. Rev. Bedaquiline and Delamanid. Antibiotics (Basel). 9:133. doi: 10.3390/antibiotics9030133

Olawoye, I. B., Uwanibe, J. N., Kunle-Ope, C. N., Davies-Bolorunduro, O. F., Abiodun, T. A., Audu, R. A., et al. (2021). Whole genome sequencing of clinical samples reveals extensively drug resistant tuberculosis (XDR TB) strains from the Beijing lineage in Nigeria, West Africa. *Sci. Rep.* 11:17387. doi: 10.1038/s41598-021-96956-7

Satta, G., Lipman, M., Smith, G. P., Arnold, C., Kon, O. M., and McHugh, T. D. (2018). Mycobacterium tuberculosis and whole-genome sequencing: how close are we to unleashing its full potential? *Clin. Microbiol. Infect.* 24, 604–609. doi: 10.1016/j.cmi.2017.10.030

Shea, J., Halse, T. A., Kohlerschmidt, D., Lapierre, P., Modestil, H. A., Kearns, C. H., et al. (2021). Low-level rifampin resistance and rpoB mutations in mycobacterium tuberculosis: an analysis of whole-genome sequencing and drug susceptibility test data in New York. *J. Clin. Microbiol.* 59:20. doi: 10.1128/JCM.01885-20

Siu, G. K., Tam, Y. H., Ho, P. L., Lee, A. S., Que, T. L., Tse, C. W., et al. (2011a). Direct detection of isoniazid-resistant mycobacterium tuberculosis in respiratory specimens by multiplex allele-specific polymerase chain reaction. *Diagn. Microbiol. Infect. Dis.* 69, 51–58. doi: 10.1016/j.diagmicrobio.2010.08.021

Siu, G. K., Zhang, Y., Lau, T. C., Lau, R. W., Ho, P. L., Yew, W. W., et al. (2011b). Mutations outside the rifampicin resistance-determining region associated with rifampicin resistance in mycobacterium tuberculosis. *J. Antimicrob. Chemother.* 66, 730–733. doi: 10.1093/jac/dkq519

Smith, C., Halse, T. A., Shea, J., Modestil, H., Fowler, R. C., Musser, K. A., et al. (2020). Assessing Nanopore sequencing for clinical diagnostics: a comparison of next-generation sequencing (NGS) methods for mycobacterium tuberculosis. *J. Clin. Microbiol.* 59:20. doi: 10.1128/JCM.00583-20

Tafess, K., Ng, T. T. L., Lao, H. Y., Leung, K. S. S., Tam, K. K. G., Rajwani, R., et al. (2020). Targeted-sequencing workflows for comprehensive drug resistance profiling of mycobacterium tuberculosis cultures using two commercial sequencing platforms: comparison of analytical and diagnostic performance, turnaround time, and cost. *Clin. Chem.* 66, 809–820. doi: 10.1093/clinchem/hvaa092

Tam, K. K., Leung, K. S., Siu, G. K., Chang, K. C., Wong, S. S., Ho, P. L., et al. (2019). Direct detection of pyrazinamide resistance in mycobacterium tuberculosis by use of pncA PCR sequencing. *J. Clin. Microbiol.* 57:19. doi: 10.1128/JCM.00145-19

Tam, K. K., Leung, K. S., To, S. W., Siu, G. K., Lau, T. C., Shek, V. C., et al. (2017). Direct detection of mycobacterium tuberculosis and drug resistance in respiratory specimen using Abbott Realtime MTB detection and RIF/INH resistance assay. *Diagn. Microbiol. Infect. Dis.* 89, 118–124. doi: 10.1016/j.diagmicrobio.2017.06.018

Walker, T. M., Merker, M., Kohl, T. A., Crook, D. W., Niemann, S., and Peto, T. E. (2017). Whole genome sequencing for M/XDR tuberculosis surveillance and for resistance testing. *Clin. Microbiol. Infect.* 23, 161–166. doi: 10.1016/j.cmi.2016.10.014

World Health Organization (2018). *Technical Manual for Drug Susceptibility Testing of Medicines Used in the Treatment of Tuberculosis*. Geneva:39 WHO.

World Health Organization (2021). *Global tuberculosis report*. Geneva: Global Tuberculosis Programme.

Yam, W. C., Yuen, K. Y., Kam, S. Y., Yiu, L. S., Chan, K. S., Leung, C. C., et al. (2006). Diagnostic application of genotypic identification of mycobacteria. *J. Med. Microbiol.* 55, 529–536. doi: 10.1099/jmm.0.46298-0

Zhao, Y., Xu, S., Wang, L., Chin, D. P., Wang, S., Jiang, G., et al. (2012). National survey of drug-resistant tuberculosis in China. *N. Engl. J. Med.* 366, 2161–2170. doi: 10.1056/NEJMoa1108789



OPEN ACCESS

EDITED BY
Daniel Golparian,
Örebro University, Sweden

REVIEWED BY
Enass Abdel-Hameed,
University of Cincinnati, United States
Lei Xiong,
Wuhan University, China

*CORRESPONDENCE
Francesco Comandatore
francesco.comandatore@unimi.it

SPECIALTY SECTION
This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 31 May 2022
ACCEPTED 26 July 2022
PUBLISHED 15 September 2022

CITATION
Papaleo S, Alvaro A, Nodari R, Panelli S,
Bitar I and Comandatore F (2022) The
red thread between methylation and
mutation in bacterial antibiotic
resistance: How third-generation
sequencing can help to unravel this
relationship.
Front. Microbiol. 13:957901.
doi: 10.3389/fmicb.2022.957901

COPYRIGHT
© 2022 Papaleo, Alvaro, Nodari,
Panelli, Bitar and Comandatore. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction
in other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

The red thread between methylation and mutation in bacterial antibiotic resistance: How third-generation sequencing can help to unravel this relationship

Stella Papaleo¹, Alessandro Alvaro², Riccardo Nodari¹,
Simona Panelli¹, Ibrahim Bitar^{3,4} and
Francesco Comandatore^{1*}

¹Romeo ed Enrica Invernizzi Pediatric Research Center, Department of Biomedical and Clinical Sciences, University of Milan, Milan, Italy, ²Romeo ed Enrica Invernizzi Pediatric Research Center, Department of Bioscience, University of Milan, Milan, Italy, ³Department of Microbiology, Faculty of Medicine and University Hospital in Pilsen, Charles University, Pilsen, Czechia, ⁴Biomedical Center, Faculty of Medicine, Charles University, Pilsen, Czechia

DNA methylation is an important mechanism involved in bacteria limiting foreign DNA acquisition, maintenance of mobile genetic elements, DNA mismatch repair, and gene expression. Changes in DNA methylation pattern are observed in bacteria under stress conditions, including exposure to antimicrobial compounds. These changes can result in transient and fast-appearing adaptive antibiotic resistance (AdR) phenotypes, e.g., strain overexpressing efflux pumps. DNA methylation can be related to DNA mutation rate, because it is involved in DNA mismatch repair systems and because methylated bases are well-known mutational hotspots. The AdR process can be the first important step in the selection of antibiotic-resistant strains, allowing the survival of the bacterial population until more efficient resistant mutants emerge. Epigenetic modifications can be investigated by third-generation sequencing platforms that allow us to simultaneously detect all the methylated bases along with the DNA sequencing. In this scenario, this sequencing technology enables the study of epigenetic modifications in link with antibiotic resistance and will help to investigate the relationship between methylation and mutation in the development of stable mechanisms of resistance.

KEYWORDS

adaptive antibiotic resistance, DNA methylation, mutations, third-generation sequencing, bacterial epigenetics, DNA repair systems

Antibiotic resistance

The discovery of antibiotic compounds has revolutionized the way in which we face infectious diseases: antibiotic drugs have saved several lives worldwide, fighting infections that nowadays seem easy to cure, but were not so simple in the past (Aminov, 2010; Ventola, 2015). However, the misuse of antibiotics led to the emergence and

spread of antibiotic-resistant bacterial clones (Nature, 2013). Indeed, it is clear that the development of a novel antibiotic drug is often followed by the discovery of a novel resistance mechanism (Ventola, 2015).

Genetic bases of antibiotic resistance

The word “antibiotic” defines the application of a compound used to kill or inhibit bacterial growth, more than a specific class of molecules (Waksman, 1947). Antibiotic drugs can be produced naturally by microorganisms or artificially through chemical synthesis. Each of these molecules affects a particular target and blocks a specific pathway, allowing classification in accordance with its mechanism of action: there are antibiotics targeting the cell wall (beta-lactams and glycopeptides), inhibitors of protein biosynthesis that block the 30S subunit (aminoglycosides and tetracyclines) or the 50S one (macrolides, chloramphenicol, oxazolidinones), inhibitors of DNA replication (quinolones), and folic acid metabolism inhibitors (Sulfonamides and trimethoprim) (Kapoor et al., 2017). Due to the differences between antibiotic mechanisms of action, the resistant strains withstand antibiotic treatment in several ways. The main resistance mechanisms can be summarized in three ways to face antibiotic attack: mutating the target site of the antibiotic or modifying it with the help of specific enzymes (Lambert, 2005; Schaenzer and Wright, 2020), inducing a modification of the antibiotic, also operated by specific enzymes (Ramirez and Tolmasky, 2010; Wilson, 2014), or increasing the efflux pump activity, which transport the drug out of the cell (Webber and Piddock, 2003).

Thus, it seems clear that a resistant strain needs a genetic determinant to overcome antibiotic attack: resistance can originate from mutations or it could be acquired through horizontal gene transfer of mobile genetic elements (Windels et al., 2020). Most of the Gram-positive organisms have a single resistant trait, making resistance detection through molecular investigation quite accurate. The situation is different for Gram-negative organisms, where antibiotic resistance is more heterogeneous: identifying it from the molecular analysis is tough and the absence of a resistance gene doesn't always mean that the strain is susceptible to a specific antibiotic (Yee et al., 2021).

Furthermore, adaptive antibiotic resistances (AdR) have been observed: resistance phenotypes not clearly associated with specific genetic determinants, such as mutations or gene presence/absence.

AdR: Adaptive antibiotic resistance

AdR is defined as the ability of the strain to temporarily withstand antibiotic presence, achieved by changes in gene

or protein expression (Fernández and Hancock, 2013). It can emerge when a strain is subjected to sub-inhibitory increasing antibiotic concentrations for a small amount of time (George and Levy, 1983; Barclay et al., 1992; Toprak et al., 2011). The AdR phenotype appears quickly and is reversible, although it is inheritable and transmissible to subsequent generations. Once the treatment with the antibiotic stops, the resistant cells revert to the susceptible phenotype, as opposed to intrinsic or acquired resistance, which is stable among the generations (Fernández and Hancock, 2013) (Figure 1).

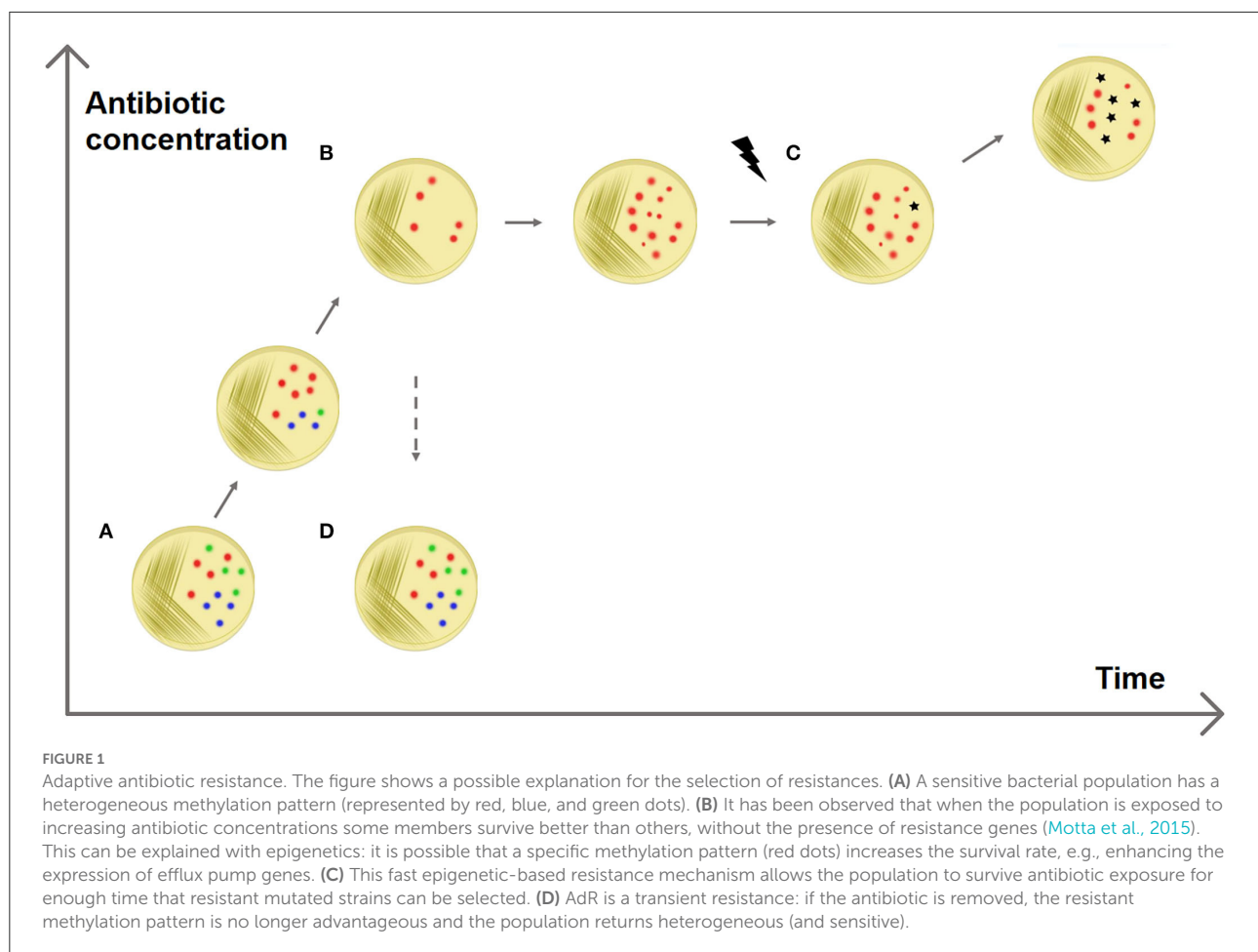
In this regard, Adam et al. (2008) investigated the emergence of antibiotic resistance in *Escherichia coli* when exposed to low concentrations of antibiotics and they found out that the different gene expression patterns were the basis of the development of AdR. The authors excluded the possibility of a mutation-mediated resistance for three reasons: (i) the survival rates were too high; (ii) the resistance MIC increased with the antibiotic concentration; (iii) the reverting frequency after antibiotic removal (around 50%) was too high for a genetic-mediated mechanism (i.e. mutation or resistance genes). These results and the investigation of differential gene expression on resistant strains led the authors to propose that the observed resistance phenotype was due to epigenetic regulation (Adam et al., 2008).

Bacterial epigenetics

The term “Epigenetics” was coined by Conrad Waddington to define the branch of biology that aimed to unravel the mechanisms under the development of a specific phenotype given a certain genotype (Goldberg et al., 2007). Epigenetic modifications are stable, inheritable, and reversible modifications of the DNA or of the histones, among which the most known and studied is methylation. This phenomenon was firstly investigated in eukaryotic organisms, where methylation of the cytosines (5-methylcytosine) in particular regions of the genome, CpG islands, was found to be involved in gene expression regulation (Moore et al., 2013). The first studies on bacterial epigenetics were published in 1955, when the presence of N6-Methyladenosine base was discovered in *E. coli* (Dunn and Smith, 1955).

In eukaryotic genomes, it is possible to observe two different methylated bases: 5-Methylcytosine and N6-Methyladenosine. Bacterial genomes can instead harbor N6-Methyladenosine, 5-Methylcytosine, and 4-Methylcytosine. The most-represented modified base in bacteria is N6-Methyladenosine, and 4-Methylcytosine is exclusive to bacteria and archaea (Sánchez-Romero and Casadesús, 2020) (Supplementary Table S2).

Another mechanism of bacterial epigenomics, whose importance is rising nowadays, is phosphorothioation, which is the first physiological modification of the DNA backbone to be discovered (Wang et al., 2007) (Supplementary Table S2).



DNA methylation

DNA base methylation is operated by a class of enzymes called DNA methyltransferases, able to catalyze the addition of a methyl group to a DNA base from a donor molecule, e.g., the S-adenosylmethionine (SAM) (Cheng, 1995). The methyltransferases recognize and methylate specific motifs on the DNA double strand. This reaction takes place immediately after the replication, when the DNA is hemimethylated: the parental strand is methylated and the daughter strand is not. The methyltransferases recognize the already-methylated motif on the parental strand and methylate the corresponding position on the other strand (Adhikari and Curtis, 2016). Methyltransferases are not error-free and they can fail to correctly recognize and methylate their motifs. It can happen for different reasons: randomly, for steric hindrance (Casadesús and Low, 2006) or for reduced processivity due to sequences flanking the recognition motifs (Peterson and Reich, 2006).

When the methyltransferase fails to methylate a motif, it produces a hemimethylated position. This can cause the unmethylation of that site in the daughter cells, after

DNA replication. This effect occurs because during DNA replication methyltransferases use each strand as a stamp for the methylation of the newly synthesized DNA daughter strands. Thus, the motif on the methylated strand will be inherited as methylated, while the other as unmethylated (Sánchez-Romero and Casadesús, 2021). This phenomenon can generate bacterial populations genetically identical but with heterogeneous methylation patterns.

Methyltransferases are mainly classified into two classes: those coupled to cognate restriction endonucleases (as part of Restriction-Modification, RM, systems) and the other lacking these paired enzymes (called orphan methyltransferases) (Adhikari and Curtis, 2016).

Restriction-modification systems

RM systems involve a methyltransferase that recognizes and methylates a particular motif, and a cognate restriction endonuclease that cleaves the same motif if it is unmethylated (Figure 2) (Vasu and Nagaraja, 2013). The main role of bacterial

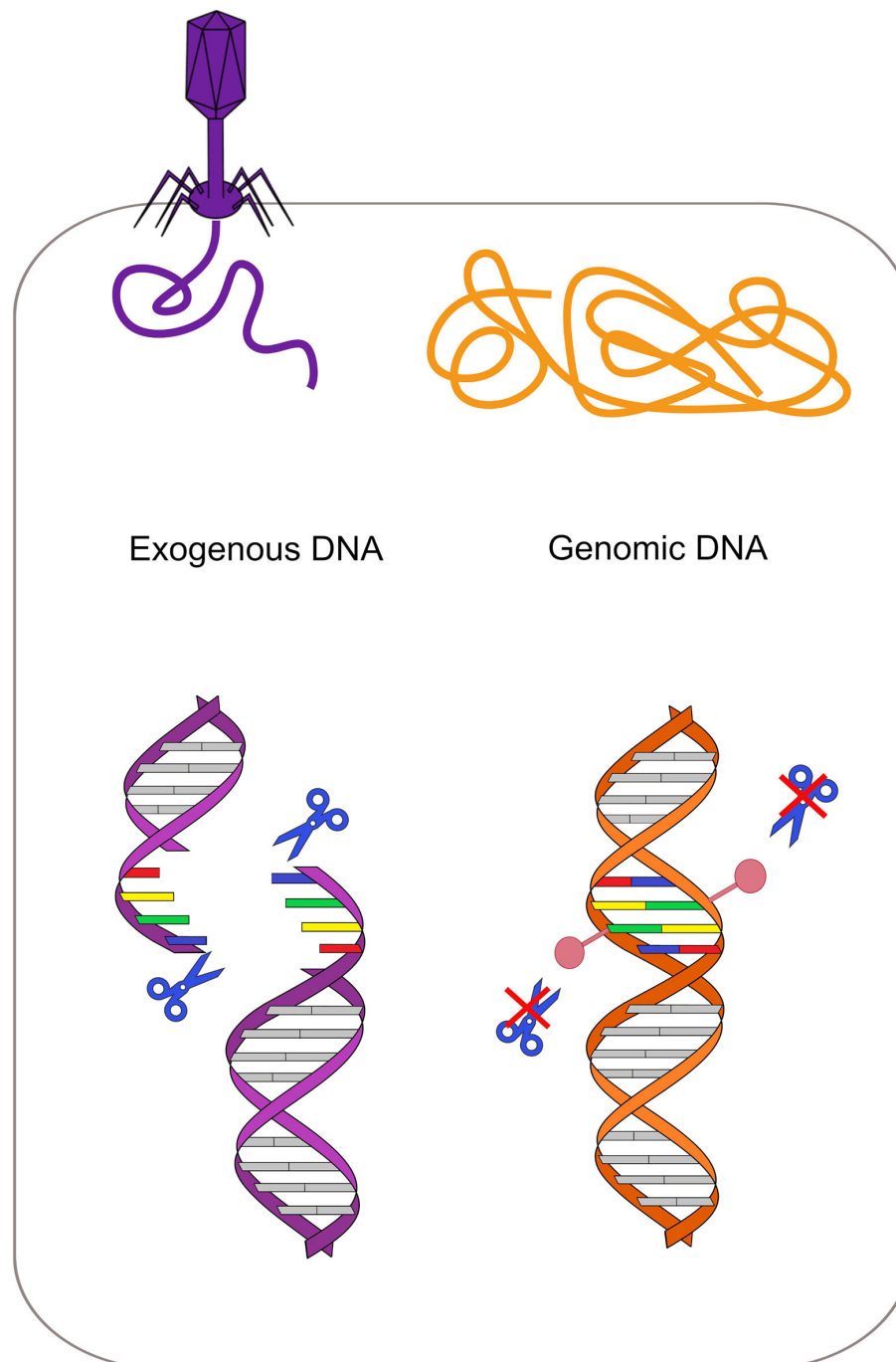


FIGURE 2

Restriction-modification system protects from exogenous DNA. Each restriction enzyme recognizes a specific motif and cleaves it if it is not methylated. In the bacterial genome (orange on the right), this specific motif is methylated by the cognate methyltransferase, while the same motif is unmethylated in the phage DNA (violet on the left). Restriction enzymes can thus cleave the phage DNA only, protecting the bacterium.

RM systems is to protect them from exogenous DNA, as a sort of primitive immune system (Jensen et al., 2019). However, recent studies revealed that methyltransferases can be also involved in transcriptional regulation (Vasu and Nagaraja, 2013).

The role of RM systems in genome protection makes them a driver for the acquisition and maintenance of mobile genetic elements (Spadar et al., 2021), directly affecting the transformation efficiency (Nye et al., 2019).

There are bacterial plasmids that, exploiting this system, increase their spreading capability. For instance, experimental evidence shows that the IncA/C plasmid, which encodes three methyltransferases genes, is able to increase its conjugation success rate in *Vibrio cholerae* modifying the bacterium methylation pattern. Indeed, the silencing of these three methyltransferases genes blocks the ability to transfer the plasmid among *Vibrio cholerae* strains and from *Vibrio cholerae* to *Escherichia coli* (Wang et al., 2019).

RM systems are involved in plasmid maintenance as they work like toxin–antitoxin (TA) systems (Kulakauskas et al., 1995). TA systems are usually composed of two proteins: a stable and lethal toxin and an unstable antitoxin, which binds and stabilizes the toxin, making it harmless. When this couple of enzymes is located on a plasmid, bacterial cells that lose that plasmid will die. Indeed, immediately after the plasmid loss, the toxin is still active while the antitoxin is rapidly degraded. This mechanism makes the plasmid indispensable for bacterial cell survival (Unterholzner et al., 2013). When the genes of methyltransferase and its cognate endonuclease are localized on the same plasmid, the functioning of the RM system can recall a toxin–antitoxin system. The restriction endonuclease acts as a toxin while the methyltransferase acts as an antitoxin, protecting the chromosomal DNA from cleavage. Indeed, plasmid-free cells can't methylate anymore the target DNA, but they still have their counterpart restriction enzyme that cleaves the target DNA and leads to cellular death. This system, leading to postsegregational killing of plasmid-free cells, increases plasmid stability (Mruk and Kobayashi, 2014) (Figure 3).

Orphan methyltransferases

Orphan methyltransferases are a group of methyltransferases highly conserved among Bacteria that lack the restriction part (Oliveira and Fang, 2021). These enzymes are involved in cellular processes such as the initiation of DNA replication, DNA repair mechanisms, and gene expression regulation (Adhikari and Curtis, 2016). The most investigated methyltransferase is deoxyadenosine methyltransferase (Dam) in *E. coli* (Schlagman et al., 1986; Messer and Noyer-Weidner, 1988; Palmer and Marinus, 1994; Calmann and Marinus, 2003). This enzyme is conserved in most of the Gammaproteobacteria, and it is an orphan adenine methyltransferase that targets the adenine of the palindromic motif 5'-GATC-3' (Ghosh et al., 2020). The GATC motif is involved in the control of DNA replication initiation. Indeed, until the genomic DNA replication starts, the regulatory protein SeqA binds several of the GATC motifs present in the OriC locus, preventing the starting of the replication phase. DNA replication can then start only when SeqA is removed and all the GATC motifs

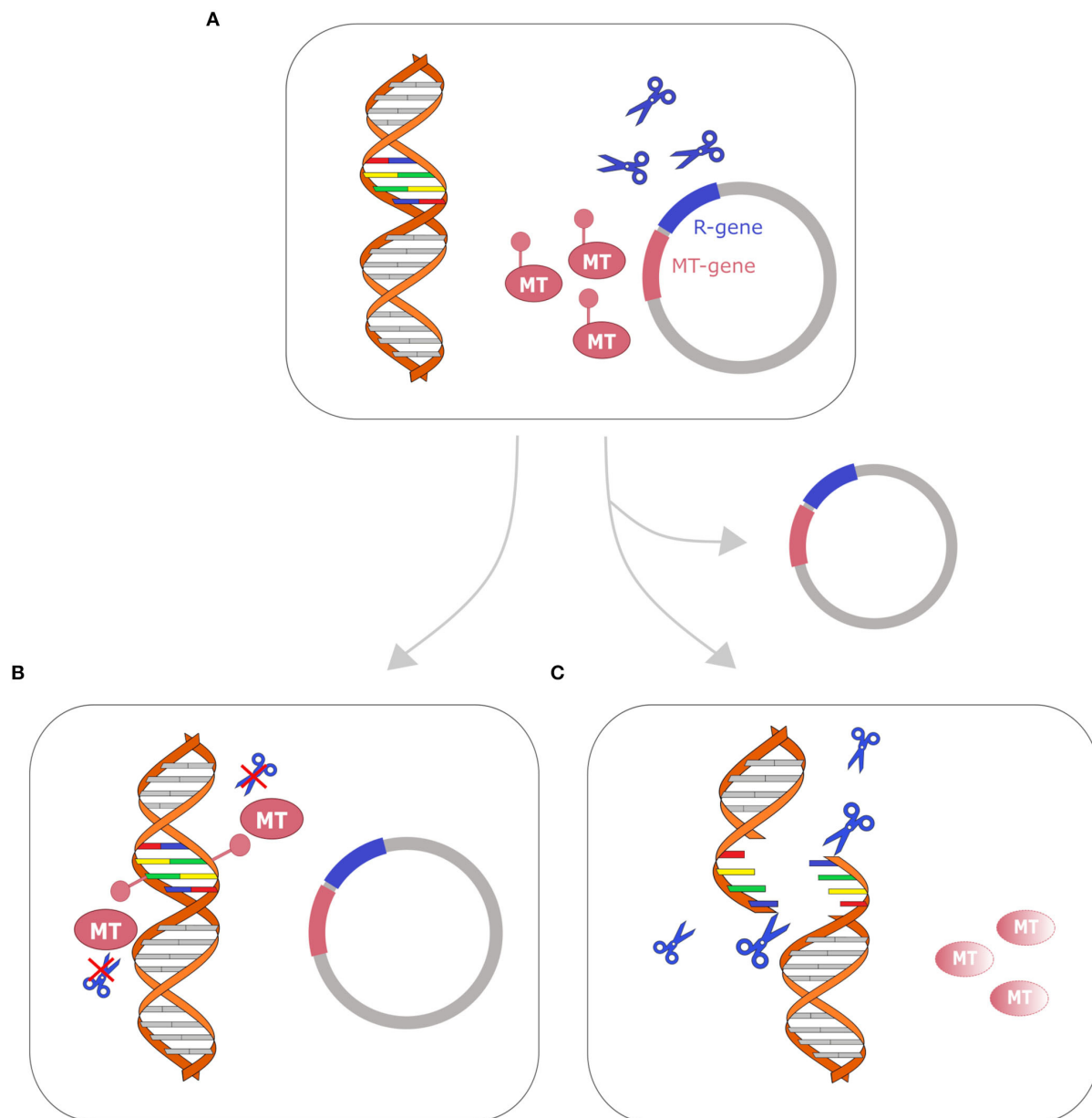
in the OriC locus are fully methylated by Dam (Kang et al., 1999). Orphan methyltransferases are also involved in the regulation of gene expression: methylation of particular motifs can enhance or inhibit the binding of regulatory proteins, such as transcriptional activators or repressors, therefore, influencing gene expression (Figure 4) (Oshima et al., 2002). Other effects of methylation on DNA are the changing of DNA curvature, the reduction of DNA thermostability, and the competition with DNA-binding proteins (Marinus and Casadesus, 2009).

Phosphorothioation

Phosphorothioation (PT) is a chemical modification of the DNA that occurs not on bases, like methylation, but on DNA backbone: it is the substitution of a non-bridging oxygen on the phosphodiester bond with a sulfur atom (Wang et al., 2007). This oxygen–sulfur exchange is catalyzed by the cooperation of the products of the *dnd* gene cluster (*dndABCDE*) that involve a cysteine desulfurase (DndA), an iron–sulfur cluster protein (DndC), a protein with ATPase activity (DndD), a protein that binds nicked dsDNA (DndE) and DndB, that seems to be not essential for PT modification (Jian et al., 2021). This modification occurs on 5'-GAAT3'/5'-GTTTC3'/5'-GATC3' and similar motifs (Wu et al., 2020). Phosphorothioation is linked with some important functions: (i) it confers protection to DNA against oxidative damage; (ii) it is often coupled with *dndFGHI* genes, which cleave unmodified DNA motifs (likewise RM systems) (Tong et al., 2018); (iii) influence or inhibit restriction enzymes that cleave close to modified sites; (iv) it changes the affinity of regulatory proteins affecting gene expression regulation (Jian et al., 2021). Interestingly, both PT modification and DNA methylation systems can recognize the 5'-GATC-3' motif, and thus a hybrid 5'-GPS6mATC-3' can be produced (Chen et al., 2017).

AdR as a bridge to stable antibiotic resistances

AdR has the potential to be the first line of defense against antibiotic exposure, unlike intrinsic and acquired resistance, which are achieved slowly by mutation or acquisition of resistance genes. This fast-appearing resistance is transient, and the resistant phenotype is easily reverted by the removal of the inducing condition. However, exploring possibilities in the survival scenario through the stochastic creation of different epigenetic lineages can be a fast way to withstand the antibiotic presence while searching for a more stable mechanism of resistance.

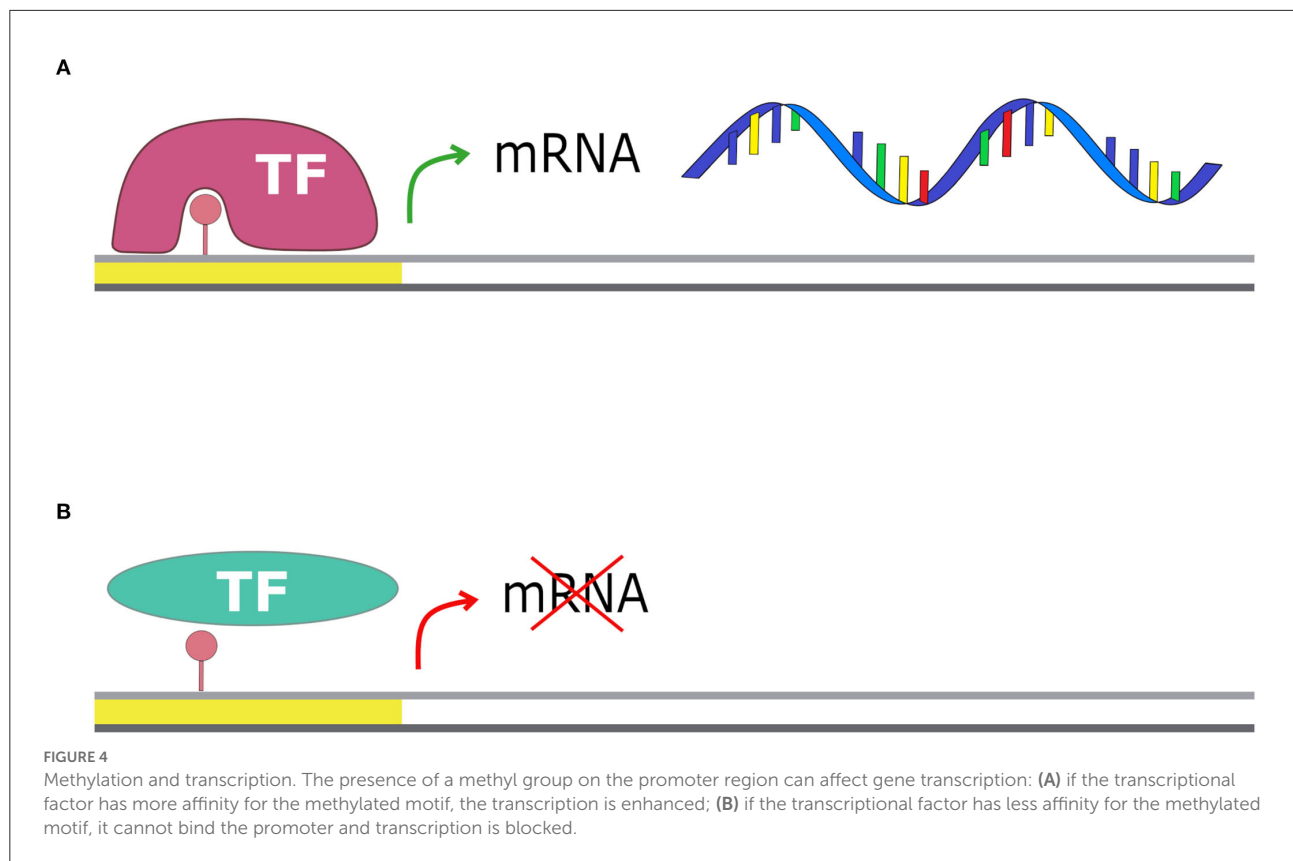
**FIGURE 3**

Restriction-modification system as a toxin-antitoxin system. Cell with plasmid that encodes R-M genes toxin-antitoxin system is represented on the top **(A)**. After replication, two fates are possible: **(B)** if the cell retains the plasmid (bottom left), the genomic DNA is methylated by the methyltransferase (MT) and the restriction enzyme (represented with blue scissors) cannot cleave the DNA; **(C)** if the cell loses the plasmid, the methyltransferase (MT) is fastly degraded, while the restriction enzyme remains active and cleaves the DNA, leading to cellular death. A similar mechanism has been described in several toxin-antitoxin systems (Kulakauskas et al., 1995; Unterholzner et al., 2013; Mruk and Kobayashi, 2014).

Adapting mechanisms: Non-conventional methods of resistance

Epigenetic adaptive resistance is mainly related to changes in gene expression regulation by the enhancement or inhibition of certain cellular features (Supplementary Table S3). For instance,

epigenetics regulation can lead to the enhancement of the expression of efflux pumps belonging to the resistance-nodulation-division (RND) superfamily able to extrude a wide range of toxins and antibiotics outside the cell (Motta et al., 2015). In fact, different expression levels of efflux pump genes such as *acrD*, *marR*, *rpoS*, *fabI*, and *lrhA* in adaptive



antibiotic-resistant strains are related to Dam methylation (Hughes et al., 2021). The overproduction of these MDR efflux pumps is also related to a decreased permeability of the cell, controlled by the underexpression of membrane porins, such as OmpC, that reduce the intake of toxic compounds (Sánchez-Romero and Casadesús, 2014).

Another AdR mechanism concerns chaperonins, enzymes involved in protein folding (Beissinger and Buchner, 1998). It is known that resistance to antibiotics that interfere with translation can emerge by the overexpression of chaperonins (Carvalho et al., 2021). Indeed, the overexpression of chaperonins GroEL/GroES expand the mutational space, because they guarantee the correct folding of proteins despite the presence of potentially lethal mutations (Goltermann et al., 2015). There is experimental evidence that chaperonins are regulated by methylation in *Vibrio cholerae*: the deletion of the orphan methyltransferase *vchM* (5'-RCCGGY-3' motif) is associated with groESL-2 upregulation in *V. cholerae* and a greater survival rate under aminoglycoside stress. The same effect has been observed in wt *V. cholerae* strains, when the 5'-RCCGGY-3' motifs in the groESL-2 region are unmethylated (Carvalho et al., 2021). Antibiotics affect main bacterial processes, depending on the mechanisms of action. For this reason, they create an imbalance in normal cellular functions and alter the cellular redox state. Therefore, the

lethality of the antibiotics is also linked to the oxidative stress that they induce in the cell (Dwyer et al., 2014). As stated above, phosphorothioation modification has both redox and nucleophilic properties that are likely to have effects on bacterial fitness in stressful environments (Kellner et al., 2017) such as antibiotic treatment.

From fast and transient mechanisms (methylation) to slow and stable ones (mutation)

When a sensitive bacterium is exposed to an antibiotic molecule, the first-line reaction can be a fast modulation of gene expression by epigenetic mechanisms. During the time period of this epigenetic-based reaction, resistant strains could emerge and a novel antibiotic-resistant bacterial population can be selected (Olofsson and Cars, 2007).

For instance, antibiotic exposure can lead to a higher expression of genes associated with efflux pumps (Fernández and Hancock, 2013). This transient phenotype can be stabilized by acquiring DNA mutations that increase the efficiency and specificity of efflux mechanisms (Sandoval-Motta and Aldana, 2016).

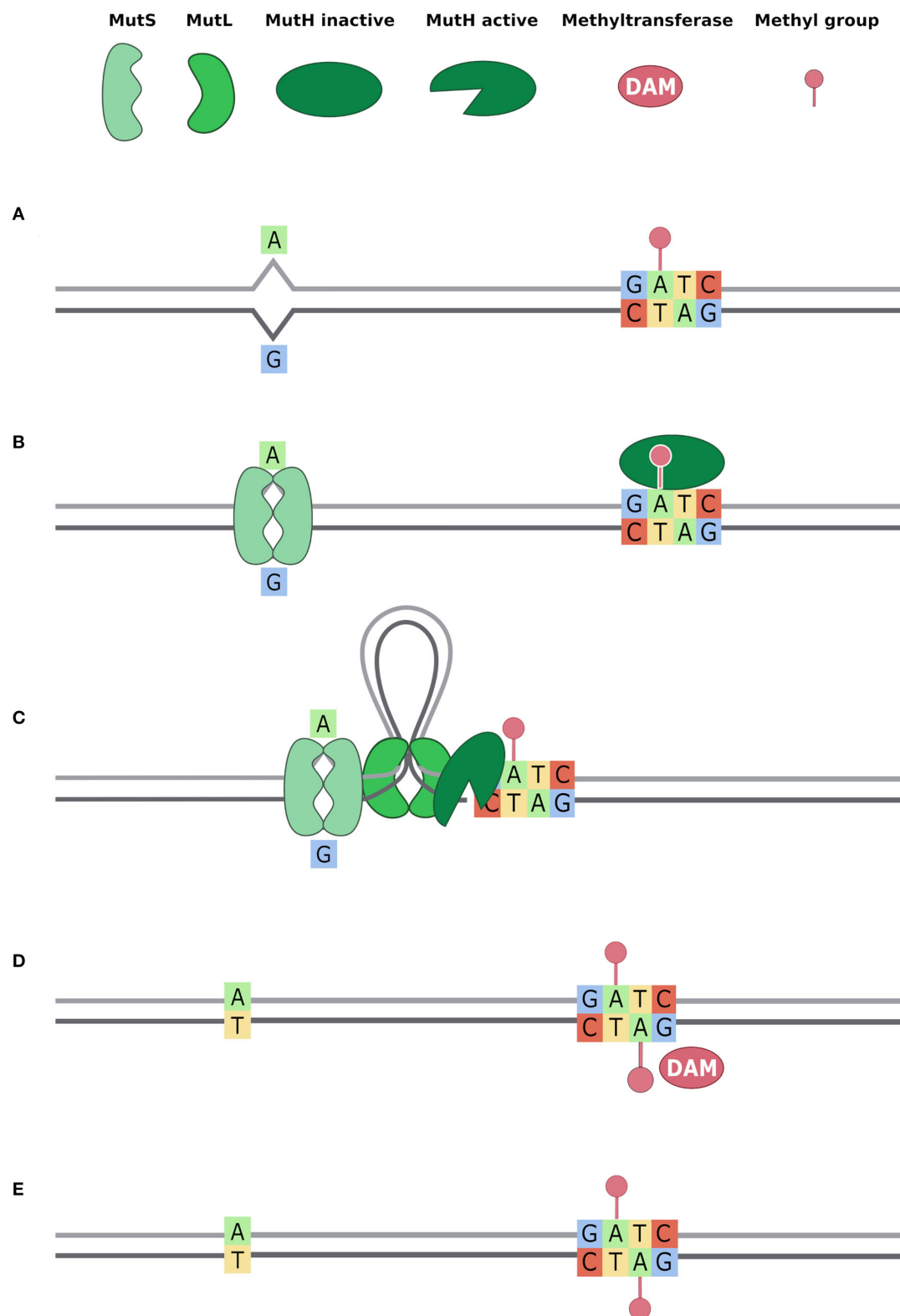


FIGURE 5

Methyl Mismatch Repair system (MMR) (also known as MutSLH). Dam methyltransferase methylates the GATC motifs right after the action of DNA polymerase and DNA repair systems (Acharya et al., 2003; Li, 2008; Sandoval-Motta and Aldana, 2016). Thus, when DNA repair systems bind (Continued)

FIGURE 5 (Continued)

the DNA, GATC motifs are methylated on the parental strand only. This allows the DNA repair system to distinguish, in a mismatch, the original base from the mutated one. In the figure, we describe the activity of MutSLH system in five steps (A–E): (A) a T → G mutation is represented and, downstream the mismatch, the nearest GATC site is methylated only on the parental strand; (B) the MutS dimer binds the mismatch and MutH binds hemimethylated GATC; (C) MutL dimer binds MutS dimer and forms a loop in the DNA, searching for the nearest GATC site to bind MutH. This activates MutH endonuclease activity that creates a nick in the unmethylated newly synthesized strand; (D) MutSLH complex recruits DNA helicase, exonuclease, DNA Pol III, and DNA ligase to fix the mismatch. Once the mismatch is fixed, Dam enzyme can methylate the GATC motif also on the other strand; (E) Both strands are methylated on the GATC site.

In this scenario, there is a switch from a Fast and Transient Mechanism (FTM), where phenotypic changes are still reversible, to a Slow and Stable Mechanism (SSM) (Sandoval-Motta and Aldana, 2016), which may overcome the transient resistance with compensatory modifications (El'Garch et al., 2007), giving a permanent resistance against specific antibiotics.

Moreover, DNA methylation affects the mutation rate, thus one of the mechanisms behind the first fast response to the antibiotics may be the one that promotes a more stable adaptation (Sandoval-Motta and Aldana, 2016).

How methylation is linked to mutation?

Bacteria can exploit adapting resistance mechanisms at the expense of other cellular functions.

Directing energies toward a high expression level of efflux pumps has a cost that bacteria pay with decreased fitness (Andersson and Levin, 1999; Ebbensgaard et al., 2020). For instance, it has been demonstrated that *E. coli* with high levels of efflux pumps show decreased activity in mismatch repair systems (Foster, 2005; El Meouche and Dunlop, 2018). Bacteria under stress conditions have a less functioning repairing system, such as for lower expression of *mutS* gene (El Meouche and Dunlop, 2018). MutS participates in two of the main DNA repair systems, in *E. coli* and other bacteria (Lieb et al., 2001): the methyl mismatch repair system (MMR) and the very short patch (VSP) repair system. The first one, also called the MutSLH pathway, is active during the exponential growth phase, in which it repairs replication errors (Figure 5). It is composed of MutS that recognizes mismatches and binds MutL, which recruits the endonuclease MutH. Cooperatively with other proteins, they restore the correct nucleotide sequence, repairing mismatch bases on the newly synthesized strand (Acharya et al., 2003). After DNA replication, there is a phase in which methylated Guanine, Adenine, Thymine and Cytosine (GATC) sites (with the adenine methylated to 6 mA) are present on the parental strand only because the Dam still has to methylate the newly synthesized strand. MutSLH is particularly active in this phase: after the MutS mismatch recognition, MutH binds the hemimethylated GATC sites and discriminates the parental strand from the newly synthesized strand for the presence of the methylated adenine (6 mA). Mismatch is thus repaired

using the parental strand as a template and replacing the incorrect base on the newly synthesized one (Li, 2008; Sandoval-Motta and Aldana, 2016). When MutS is down expressed, this system can't work properly, therefore leading to a high mutation rate due to uncorrected mismatched bases caused by DNA polymerase errors (El Meouche and Dunlop, 2018). MutS is also involved in the VSP mismatch repair system, which fixes T/G mismatches in non-dividing cells (Figure 6). Indeed, methylated cytosines spontaneously deaminate at thymines, creating T/G mismatches. The VSP system is able to recognize and repair these mismatches with the cooperative work of MutS, which binds the mismatch, of the VSR endonuclease, which cuts the DNA, and MutL, which recruits the helicase (Drotschmann et al., 1998). VSP prevents the fixation of the mutation due to methylated cytosine deamination that results in a C to T transition. VSP works on the 5'-CCWGG-3' motif (Marinus, 2012), where the inner cytosine is methylated by Dcm, an orphan methyltransferase. Thus, if MutS is lacking, VSP can't restore T/G mismatches. In this scenario, uncorrected deamination of methylated cytosine has a strong effect on the mutational rate.

Therefore, under antibiotic treatment, the emerged adaptive-resistant bacteria have higher mutation rate as consequence of the lower efficiency of DNA mismatch repair systems. This leads to two options: the bacteria can acquire lethal mutations that kill the cell or it can store advantageous mutations in order to reach a stable antibiotic resistance.

Furthermore, methylated bases are shown to be directly linked with mutation because they are mutational hotspots. A genomic study, performed by combining short reads with SMRT sequencing, revealed that methylated adenine (6 mA) are mutational hotspots in *Neisseria meningitidis* (Sater et al., 2015). The study of thousands of *E. coli* and *Salmonella* spp. genome assemblies revealed that there is mutation bias in methylation motifs (Cherry, 2018). As stated above, Dcm methylates the inner C residual of the motif 5'-CCWGG-3', and the authors found that this C residual has a C to T mutation rate 8 fold higher than the other C residuals in the genome.

The authors also found a similar mutation bias for the motifs of the other RM methyltransferases. In another work, Cherry (2021) showed how 4 mC methylation caused by *Salmonella enterica* Type III RM system in the motif 5'-CAC4mCGT-3' increases the transversion rate from the methylated cytosine to adenine by 500-fold.

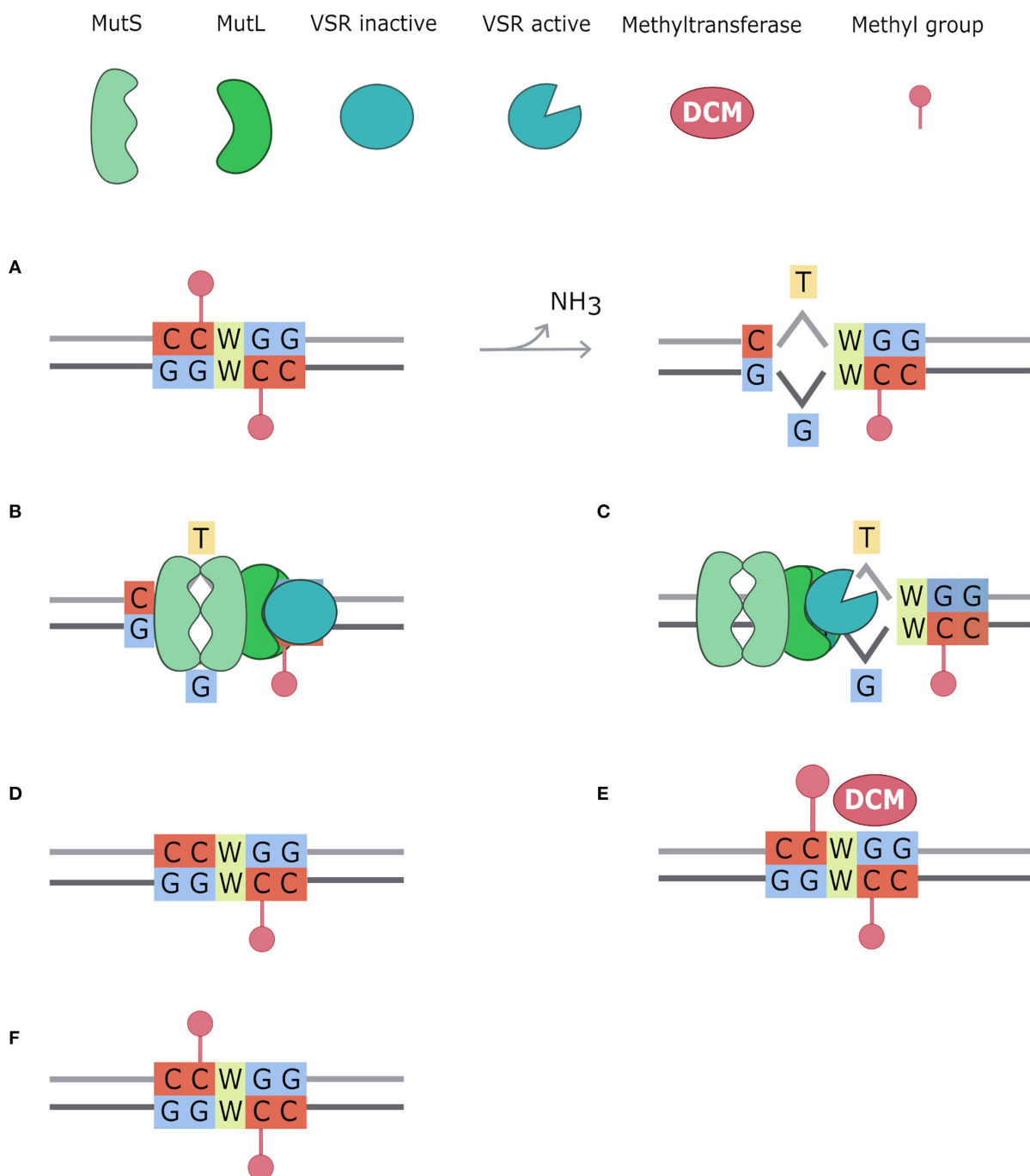


FIGURE 6

Very Short Patch repair system (VSP). Methylated cytosines spontaneously deaminate to Thymine creating T/G mismatches. The very short patch repair system (VSP) acts repairing these mismatches, always in favor of the Guanine (Drotschmann et al., 1998; Marinus, 2012). In this figure, the VSP system mechanism is divided into six steps: **(A)** CCWGG motifs are methylated on both strands by the Dcm enzyme. Methylated cytosines spontaneously deaminate to thymine, creating a mismatch with guanine; **(B)** MutS dimer recognizes the mismatch and binds MutL dimer which acts as a bridge with VSR protein; **(C)** VSR is activated by the MutSL complex and creates a nick in the strand containing the incorrect base, using as parental stand that containing the Guanosine; **(D)** the MutSL-VSR complex recruits DNA helicase, exonuclease, DNA polymerase, and DNA ligase to fix the mismatch; **(E)** CCWGG motif is restored and thus DCM methyltransferase can methylate it; **(F)** CCWGG is methylated on both strands.

Methods used to investigate epigenetics in bacteria

Despite the importance of AdR and epigenetics in the emergence of antibiotic-resistant strains, these mechanisms remain poorly investigated. Among the reasons, the lack of suitable techniques to detect methylated bases is likely one of the main. Recently, the development of third-generation sequencing machines revolutionized the field allowing the precise identification of methylated bases. The main sequencing-based methods available for the analysis of DNA methylation are, namely, bisulfite sequencing, restriction enzymes-based mapping and third-generation sequencing (SMRT and Nanopore) (Supplementary Table S4).

Bisulfite sequencing technology (Beaulaurier et al., 2019) is able to detect 5 mC and 4 mC sites but is blind to m6A. Sodium bisulfite treatment converts the unmethylated cytosines of genomic DNA in uracil, while methylated cytosines remain intact. During library preparation uracil is converted to thymine; therefore, the resulting sequence can be compared with the sequence of untreated DNA to obtain the 5 mC positions. Four mC is detected by adding the TET (ten-eleven translocation) enzyme to the standard bisulfite sequencing protocol. This enzyme catalyzes the oxidation of 5 mC to 5-carboxylcytosine (5caC), which is read as thymine in the final sequence, remaining only 4 mCs as cytosine.

Restriction enzymes-based mapping is a sequencing protocol able to detect methylation of known motifs: it is based on the use of a couple of restriction enzymes that recognize a sequence if methylated or unmethylated. The resulting fragments will be analyzed with Next Generation Sequencing (NGS). This method is reliable but useful only if the restriction motif is known, and if methyl sensitive and insensitive specific restriction enzymes exist for that motif (Beaulaurier et al., 2019).

The recent advances in long-read sequencing technologies provided the means to directly investigate bacterial DNA methylation patterns. The emergence of third-generation sequencing machines allowed the detection of the N6-methyladenosine (m6A) and of the other modified bases of the DNA, along with the simultaneous detection of the nucleotide sequence. Two platforms are available for the detection of modified bases on a genomic scale (Payelleville and Brillard, 2021): Single Molecule, Real-Time (SMRT) sequencing (Eid et al., 2009) and the Oxford Nanopore technology (ONT). These approaches don't involve a replication step during the library preparation, reducing the time required before sequencing, and allowing the preservation of all the modifications on the genome.

Despite the great advantages provided by third-generation sequencing, at the state-of-the-art, some limits remain: (i) DNA has to be extracted from fresh samples; (ii) the high error rate of SMRT and Nanopore sequencing could include errors in the methylation analysis; and (iii) gold standard data analysis algorithms and software still lack.

SMRT

Single Molecule, Real-Time (SMRT) sequencing can detect, with different sensitivity, all the main modifications of the DNA: it allows an optimal detection of 4 mC and m6A, but for 5 mC it requires additional steps, such as TET conversion or very deep sequencing coverage (Beaulaurier et al., 2019). In addition to that, SMRT sequencing also allows the detection of the phosphotioroation of the backbone of the DNA (Cao et al., 2014). The fundamental unity of this sequencing technology is composed by a zero-mode waveguide (ZMW), a small chamber in which the light converges, on which a DNA polymerase is immobilized. The template is a DNA double-stranded fragment, circularized and ligated with hairpin adaptors to each end (Eid et al., 2009), and it is anchored to the DNA polymerase at the bottom of the chamber. The enzyme proceeds multiple times on the template, adding fluorescently nucleotides (dNTP) complementary to the template base and labeled with four different fluorophores. Each base is added and immobilized for a short amount of time in the ZMW, its fluorescence pulse captured by a camera, and the set of all the pulses is used to construct the nucleotide sequence. There is an interval of time between every incorporating event that is called inter-pulse duration (IPD) and describes the polymerase kinetics. IPD is eventually modified by chemical modifications of the template DNA, so methylated bases or modification of the DNA backbone can be identified with the changes in polymerase kinetics (Ardui et al., 2018).

SMRT sequencing can be coupled with microarray techniques to link DNA methylation profiles with gene expression, allowing correlations between transcriptional gene levels and different methylation patterns in antibiotic resistance strains (Chen et al., 2018). This sequencing technique has shown to be useful to detect new methylation motifs (Blow et al., 2016), and it has been used to assess also the possible indirect correlations between the methylation status of a specific motif and the expression of a resistance gene (Spadar et al., 2021). It has been established that the absence of methylation in a motif downstream of a gene that is not transcribed may be a sign of a DNA conformation that prevents methyltransferase binding but also RNA polymerase binding. Thus, methylation analysis could provide information useful to identify distant regulatory regions or secondary DNA structure that can affect gene transcription (Spadar et al., 2021).

Oxford Nanopore technologies

Similar to SMRT, Nanopore sequencing doesn't require an amplification step and allows the sequencing of very long reads (around 10 kb). This system works on a membrane filled with nanopore proteins, immersed in an electrolyte solution on which a voltage current is applied. When the single-strand DNA

template passes through the nanopore protein, a nucleotide-specific electric alteration is produced and registered (Branton et al., 2008). The system can also detect methylated bases because they generate a distinctive current pattern, different from the unmethylated bases (Ciuffreda et al., 2021).

Epigenetic analysis with third-generation sequencing technologies will represent a boost for antibiotic resistance studies

Bacterial epigenetics has a role in antibiotic resistance, and third-generation sequencing platforms are pivotal tools for the investigation of this phenomenon. Third-generation sequencing, and in particular SMRT, can be used for the detection of all the main epigenetic features of bacterial DNA, including the identification of methylated bases and modification of the DNA backbone. The use of these sequencing platforms could allow the discovery of novel mechanisms for the emergence of antibiotic-resistant strains. It is also possible that, in the future, the prediction of resistance phenotype based on bacterial genetics will also take into account epigenetic aspects. Furthermore, third-generation sequencing allows the comprehension of the mechanisms of genetic transfer, in order to monitor transfer and acquisition of resistance genes and mobile elements with resistance determinants. On the one hand, the combination of second- and third-generation sequencing platforms (short and long reads) allow us to investigate the link between methylation and mutation. On the other hand, third-generation sequencing data could be combined with RNA-seq or microarray experiment results to study the role of fast gene expression modulation in the emergence of antibiotic-resistant strains. Methylation analysis can be also useful to identify distant regulatory regions or secondary DNA structure that can prevent gene transcription and to correlate methylation patterns with different gene expression also in resistant strains. Exploring the epigenetic mechanism beyond the antibiotic resistance could also lead to the discovery of novel pharmacological targets, perhaps not subject to selective pressure and that could not lead to the development of new antibiotic resistance mechanisms.

Conclusion

Epigenetic mechanisms in bacteria are still not fully understood, and the studies performed until now have only scratched the surface of the problem. This knowledge gap in such an important process can strongly limit our ability to understand the mechanisms beyond the emergence of antibiotic-resistant nosocomial pathogens.

Most of the experimental studies on DNA methylation have been carried out on *E. coli* by the knockout of methylation-associated genes, rather than focusing on the identification of methylated DNA bases. The study of the expression of genes involved in DNA methylation in *E. coli* during the early stages of exposure to antibiotics could help to highlight the pathways involved in this mechanism. Third-generation sequencing technologies directly collect methylation information during the DNA sequencing process, allowing us to investigate the relationship between methylation and mutation. In the future, improvements in standardization of methylation experiments and in sequencing data analysis will be pivotal to obtain more robust and comparable results.

Author contributions

SPap performed the review of the literature, wrote the manuscript, and prepared the figures. AA, RN, and SPan performed the review of the literature. IB conceived the idea. FC conceived the idea and wrote the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors acknowledge support from the University of Milan through the APC initiative.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.957901/full#supplementary-material>

References

- Acharya, S., Foster, P. L., Brooks, P., and Fishel, R. (2003). The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol. Cell* 12, 233–246. doi: 10.1016/S1097-2765(03)00219-3
- Adam, M., Murali, B., Glenn, N. O., and Potter, S. S. (2008). Epigenetic inheritance based evolution of antibiotic resistance in bacteria. *BMC Evol. Biol.* 8, 52. doi: 10.1186/1471-2148-8-52
- Adhikari, S., and Curtis, P. D. (2016). DNA methyltransferases and epigenetic regulation in bacteria. *FEMS Microbiol. Rev.* 40, 575–591. doi: 10.1093/femsre/fuw023
- Aminov, R. I. (2010). A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* 1, 134. doi: 10.3389/fmicb.2010.00134
- Andersson, D. I., and Levin, B. R. (1999). The biological cost of antibiotic resistance. *Curr. Opin. Microbiol.* 2, 489–493. doi: 10.1016/S1369-5274(99)00005-3
- Ardui, S., Ameer, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res.* 46, 2159–2168. doi: 10.1093/nar/gky066
- Barclay, M. L., Begg, E. J., and Chambers, S. T. (1992). Adaptive resistance following single doses of gentamicin in a dynamic in vitro model. *Antimicrob. Agents Chemother.* 36, 1951–1957. doi: 10.1128/AAC.36.9.1951
- Beaulaurier, J., Schadt, E. E., and Fang, G. (2019). Deciphering bacterial epigenomes using modern sequencing technologies. *Nat. Rev. Genet.* 20, 157–172. doi: 10.1038/s41576-018-0081-3
- Beissinger, M., and Buchner, J. (1998). How chaperones fold proteins. *Biol. Chem.* 379, 245–259.
- Blow, M. J., Clark, T. A., Daum, C. G., Deutschbauer, A. M., Fomenkov, A., Fries, R., et al. (2016). The epigenomic landscape of prokaryotes. *PLoS Genet.* 12, e1005854. doi: 10.1371/journal.pgen.1005854
- Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., et al. (2008). The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 26, 1146–1153. doi: 10.1038/nbt.1495
- Calmann, M. A., and Marinus, M. G. (2003). Regulated expression of the Escherichia coli dam gene. *J. Bacteriol.* 185, 5012–5014. doi: 10.1128/JB.185.16.5012-5014.2003
- Cao, B., Chen, C., DeMott, M. S., Cheng, Q., Clark, T. A., Xiong, X., et al. (2014). Genomic mapping of phosphorothioates reveals partial modification of short consensus sequences. *Nat. Commun.* 5, 1–13. doi: 10.1038/ncomms4951
- Carvalho, A., Mazel, D., and Baharoglu, Z. (2021). Deficiency in cytosine DNA methylation leads to high chaperonin expression and tolerance to aminoglycosides in Vibrio cholerae. *PLoS Genet.* 17, e1009748. doi: 10.1371/journal.pgen.1009748
- Casadesús, J., and Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* 70, 830–856. doi: 10.1128/MMBR.00016-06
- Chen, C., Wang, L., Chen, S., Wu, X., Gu, M., Chen, X., et al. (2017). Convergence of DNA methylation and phosphorothioation epigenetics in bacterial genomes. *Proc. Natl. Acad. Sci. USA* 114, 4501–4506. doi: 10.1073/pnas.1702450114
- Chen, L., Li, H., Chen, T., Yu, L., Guo, H., Chen, Y., et al. (2018). Genome-wide DNA methylation and transcriptome changes in Mycobacterium tuberculosis with rifampicin and isoniazid resistance. *Int. J. Clin. Exp. Pathol.* 11, 3036–3045.
- Cheng, X. (1995). Structure and function of DNA methyltransferases. *Annu. Rev. Biophys. Biomol. Struct.* 24, 293–318. doi: 10.1146/annurev.bb.24.060195.001453
- Cherry, J. L. (2018). Methylation-induced hypermutation in natural populations of bacteria. *J. Bacteriol.* 200, e00371–e00418. doi: 10.1128/JB.00371-18
- Cherry, J. L. (2021). Extreme c-to-a hypermutation at a site of cytosine-N4 methylation. *mBio* 12, e00172–21. doi: 10.1128/mBio.00172-21
- Ciuffreda, L., Rodríguez-Pérez, H., and Flores, C. (2021). Nanopore sequencing and its application to the study of microbial communities. *Comput. Struct. Biotechnol. J.* 19, 1497–1511. doi: 10.1016/j.csbj.2021.02.020
- Drotschmann, K., Aronshtam, A., Fritz, H. J., and Marinus, M. G. (1998). The Escherichia coli MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res.* 26, 948–953. doi: 10.1093/nar/26.4.948
- Dunn, D. B., and Smith, J. D. (1955). Occurrence of a new base in the deoxyribonucleic acid of a strain of Bacterium coli. *Nature* 175, 336–337. doi: 10.1038/175336a0
- Dwyer, D. J., Belenky, P. A., Yang, J. H., MacDonald, I. C., Martell, J. D., Takahashi, N., et al. (2014). Antibiotics induce redox-related physiological alterations as part of their lethality. *Proc. Natl. Acad. Sci. USA* 111, E2100–E2109. doi: 10.1073/pnas.1401876111
- Ebbensgaard, A. E., Løbner-Olesen, A., and Frimodt-Møller, J. (2020). The role of efflux pumps in the transition from low-level to clinical antibiotic resistance. *Antibiotics* 9, 855. doi: 10.3390/antibiotics9120855
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986
- El Meouche, I., and Dunlop, M. J. (2018). Heterogeneity in efflux pump expression predisposes antibiotic-resistant cells to mutation. *Science* 362, 686–690. doi: 10.1126/science.aar7981
- El'Garch, F., Jeannot, K., Hocquet, D., Llanes-Barakat, C., and Plésiat, P. (2007). Cumulative effects of several nonenzymatic mechanisms on the resistance of Pseudomonas aeruginosa to aminoglycosides. *Antimicrob. Agents Chemother.* 51, 1016–1021. doi: 10.1128/AAC.00704-06
- Fernández, L., and Hancock, R. E. W. (2013). Adaptive and mutational resistance: Role of porins and efflux pumps in drug resistance. *Clin. Microbiol. Rev.* 26, 163–163. doi: 10.1128/CMR.00094-12
- Foster, P. L. (2005). Stress responses and genetic variation in bacteria. *Mutation Res. Fund. Mol. Mech. Mutagenesis* 569, 3–11. doi: 10.1016/j.mrfmmm.2004.07.017
- George, A. M., and Levy, S. B. (1983). Amplifiable resistance to tetracycline, chloramphenicol, and other antibiotics in Escherichia coli: involvement of a non-plasmid-determined efflux of tetracycline. *J. Bacteriol.* 155, 531–540. doi: 10.1128/jb.155.2.531-540.1983
- Ghosh, D., Veeraraghavan, B., Elangovan, R., and Vivekanandan, P. (2020). Antibiotic resistance and epigenetics: more to it than meets the eye. *Antimicrob. Agents Chemother.* 64, e02225–e02239. doi: 10.1128/AAC.02225-19
- Goldberg, A. D., Allis, C. D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell* 128, 635–638. doi: 10.1016/j.cell.2007.02.006
- Goltermann, L., Sarusie, M. V., and Bentin, T. (2015). Chaperonin GroEL/GroES over-expression promotes aminoglycoside resistance and reduces drug susceptibilities in Escherichia coli following exposure to sublethal aminoglycoside doses. *Front. Microbiol.* 6, 1572. doi: 10.3389/fmicb.2015.01572
- Hughes, L., Roberts, W., and Johnson, D. (2021). The impact of DNA adenine methyltransferase knockout on the development of triclosan resistance and antibiotic cross-resistance in. *Access Microbiol* 3, acmi000178. doi: 10.1099/acmi.0.000178
- Jensen, T. Ø., Tellgren-Roth, C., Redl, S., Mauray, J., Jacobsen, S. A. B., Pedersen, L. E., et al. (2019). Genome-wide systematic identification of methyltransferase recognition and modification patterns. *Nat. Commun.* 10, 3311. doi: 10.1038/s41467-019-11179-9
- Jian, H., Xu, G., Yi, Y., Hao, Y., Wang, Y., Xiong, L., et al. (2021). The origin and impeded dissemination of the DNA phosphorothioation system in prokaryotes. *Nat. Commun.* 12, 6382. doi: 10.1038/s41467-021-26636-7
- Kang, S., Lee, H., Han, J. S., and Hwang, D. S. (1999). Interaction of SeqA and Dam methylase on the hemimethylated origin of Escherichia coli chromosomal DNA replication. *J. Biol. Chem.* 274, 11463–11468. doi: 10.1074/jbc.274.17.11463
- Kapoor, G., Saigal, S., and Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: a guide for clinicians. *J. Anaesthesiol. Clin. Pharmacol.* 33, 300–305. doi: 10.4103/joacp.JOACP_349_15
- Kellner, S., DeMott, M. S., Cheng, C. P., Russell, B. S., Cao, B., You, D., et al. (2017). Oxidation of phosphorothioate DNA modifications leads to lethal genomic instability. *Nat. Chem. Biol.* 13, 888–894. doi: 10.1038/nchembio.2407
- Kulakauskas, S., Lubys, A., and Ehrlich, S. D. (1995). DNA restriction-modification systems mediate plasmid maintenance. *J. Bacteriol.* 177, 3451–3454. doi: 10.1128/jb.177.12.3451-3454.1995
- Lambert, P. A. (2005). Bacterial resistance to antibiotics: modified target sites. *Adv. Drug Deliv. Rev.* 57, 1471–1485. doi: 10.1016/j.addr.2005.04.003
- Li, G.-M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Res.* 18, 85–98. doi: 10.1038/cr.2007.115
- Lieb, M., Rehmat, S., and Bhagwat, A. S. (2001). Interaction of MutS and Vsr: some dominant-negative mutS mutations that disable methyladenine-directed mismatch repair are active in very-short-patch repair. *J. Bacteriol.* 183, 6487–6490. doi: 10.1128/JB.183.21.6487-6490.2001
- Marinus, M. G. (2012). DNA mismatch repair. *EcoSal Plus.* 5, doi: 10.1128/ecosalplus.7.2.5

- Marinus, M. G., and Casadesus, J. (2009). Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol. Rev.* 33, 488–503. doi: 10.1111/j.1574-6976.2008.00159.x
- Messer, W., and Noyer-Weidner, M. (1988). Timing and targeting: the biological functions of Dam methylation in *E. coli*. *Cell* 54, 735–737. doi: 10.1016/S0092-8674(88)90911-7
- Moore, L. D., Le, T., and Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology* 38, 23–38. doi: 10.1038/npp.2012.112
- Motta, S. S., Cluzel, P., and Aldana, M. (2015). Adaptive resistance in bacteria requires epigenetic inheritance, genetic noise, and cost of efflux pumps. *PLoS ONE* 10, e0118464. doi: 10.1371/journal.pone.0118464
- Mruk, I., and Kobayashi, I. (2014). To be or not to be: regulation of restriction-modification systems and other toxin-antitoxin systems. *Nucleic Acids Res.* 42, 70–86. doi: 10.1093/nar/gkt711
- Nature, E. (2013). The antibiotic alarm. *Nature* 495, 141. doi: 10.1038/495141a
- Nye, T. M., Jacob, K. M., Holley, E. K., Nevarez, J. M., Dawid, S., Simmons, L. A., et al. (2019). DNA methylation from a Type I restriction modification system influences gene expression and virulence in *Streptococcus pyogenes*. *PLoS Pathog.* 15, e1007841. doi: 10.1371/journal.ppat.1007841
- Oliveira, P. H., and Fang, G. (2021). Conserved DNA Methyltransferases: A Window into Fundamental Mechanisms of Epigenetic Regulation in Bacteria. *Trends Microbiol.* 29, 28–40. doi: 10.1016/j.tim.2020.04.007
- Olofsson, S. K., and Cars, O. (2007). Optimizing drug exposure to minimize selection of antibiotic resistance. *Clin. Infect. Dis.* 45, S129–36. doi: 10.1086/519256
- Oshima, T., Wada, C., Kawagoe, Y., Ara, T., Maeda, M., Masuda, Y., et al. (2002). Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*. *Mol. Microbiol.* 45, 673–695. doi: 10.1046/j.1365-2958.2002.03037.x
- Palmer, B. R., and Marinus, M. G. (1994). The dam and dcm strains of *Escherichia coli*—a review. *Gene* 143, 1–12. doi: 10.1016/0378-1119(94)90597-5
- Payelleville, A., and Brillard, J. (2021). Novel Identification of Bacterial Epigenetic Regulations Would Benefit From a Better Exploitation of Methylomic Data. *Front. Microbiol.* 12, 685670. doi: 10.3389/fmicb.2021.685670
- Peterson, S. N., and Reich, N. O. (2006). GATC flanking sequences regulate Dam activity: evidence for how Dam specificity may influence pap expression. *J. Mol. Biol.* 355, 459–472. doi: 10.1016/j.jmb.2005.11.003
- Ramirez, M. S., and Tolmashy, M. E. (2010). Aminoglycoside modifying enzymes. *Drug Resist. Updat.* 13, 151–171. doi: 10.1016/j.drug.2010.08.003
- Sánchez-Romero, M. A., and Casadesús, J. (2014). Contribution of phenotypic heterogeneity to adaptive antibiotic resistance. *Proc. Natl. Acad. Sci. USA* 111, 355–360. doi: 10.1073/pnas.1316084111
- Sánchez-Romero, M. A., and Casadesús, J. (2020). The bacterial epigenome. *Nat. Rev. Microbiol.* 18, 7–20. doi: 10.1038/s41579-019-0286-2
- Sánchez-Romero, M. A., and Casadesús, J. (2021). Waddington's landscapes in the bacterial world. *Front. Microbiol.* 12, 1167. doi: 10.3389/fmicb.2021.685080
- Sandoval-Motta, S., and Aldana, M. (2016). Adaptive resistance to antibiotics in bacteria: a systems biology perspective. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 8, 253–267. doi: 10.1002/wsbm.1335
- Sater, M. R. A., Lamelas, A., Wang, G., Clark, T. A., Röltgen, K., Mane, S., et al. (2015). DNA methylation assessed by SMRT sequencing is linked to mutations in *Neisseria meningitidis* Isolates. *PLoS ONE* 10, e0144612. doi: 10.1371/journal.pone.0144612
- Schaenzer, A. J., and Wright, G. D. (2020). Antibiotic resistance by enzymatic modification of antibiotic targets. *Trends Mol. Med.* 26, 768–782. doi: 10.1016/j.molmed.2020.05.001
- Schlagman, S. L., Hattman, S., and Marinus, M. G. (1986). Direct role of the *Escherichia coli* dam DNA methyltransferase in methylation-directed mismatch repair. *J. Bacteriol.* 165, 896–900. doi: 10.1128/jb.165.3.896-900.1986
- Spadar, A., Perdigão, J., Phelan, J., Charleston, J., Modesto, A., Elias, R., et al. (2021). Methylation analysis of *Klebsiella pneumoniae* from Portuguese hospitals. *Sci. Rep.* 11, 6491. doi: 10.1038/s41598-021-85724-2
- Tong, T., Chen, S., Wang, L., Tang, Y., Ryu, J. Y., Jiang, S., et al. (2018). Occurrence, evolution, and functions of DNA phosphorothioate epigenetics in bacteria. *Proc. Natl. Acad. Sci. USA* 115, E2988–E2996. doi: 10.1073/pnas.1721916115
- Toprak, E., Veres, A., Michel, J.-B., Chait, R., Hartl, D. L., and Kishony, R. (2011). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.* 44, 101–105. doi: 10.1038/ng.1034
- Unterholzner, S. J., Poppenberger, B., and Rozhon, W. (2013). Toxin-antitoxin systems: Biology, identification, and application. *Mob. Genet. Elements* 3, e26219. doi: 10.4161/mge.26219
- Vasu, K., and Nagaraja, V. (2013). Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol. Mol. Biol. Rev.* 77, 53–72. doi: 10.1128/MMBR.00044-12
- Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *Pharmacy Therapeutics* 40, 277–283.
- Waksman, S. A. (1947). What is an antibiotic or an antibiotic substance? *Mycologia* 39, 565–569. doi: 10.1080/00275514.1947.12017635
- Wang, L., Chen, S., Xu, T., Taghizadeh, K., Wishnok, J. S., Zhou, X., et al. (2007). Phosphorothioation of DNA in bacteria by dnd genes. *Nat. Chem. Biol.* 3, 709–710. doi: 10.1038/nchembio.2007.39
- Wang, R., Lou, J., and Li, J. (2019). A mobile restriction modification system consisting of methylases on the IncA/C plasmid. *Mob. DNA* 10, 26. doi: 10.1186/s13100-019-0168-1
- Webber, M. A., and Piddock, L. J. V. (2003). The importance of efflux pumps in bacterial antibiotic resistance. *J. Antimicrob. Chemother.* 51, 9–11. doi: 10.1093/jac/dkg050
- Wilson, D. N. (2014). Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* 12, 35–48. doi: 10.1038/nrmicro.3155
- Windels, E. M., Van den Bergh, B., and Michiels, J. (2020). Bacteria under antibiotic attack: Different strategies for evolutionary adaptation. *PLoS Pathog.* 16, e1008431. doi: 10.1371/journal.ppat.1008431
- Wu, X., Cao, B., Aquino, P., Chiu, T.-P., Chen, C., Jiang, S., et al. (2020). Epigenetic competition reveals density-dependent regulation and target site plasticity of phosphorothioate epigenetics in bacteria. *Proc. Natl. Acad. Sci. USA* 117, 14322–14330. doi: 10.1073/pnas.2002933117
- Yee, R., Dien Bard, J., and Simner, P. J. (2021). The genotype-to-phenotype dilemma: how should laboratories approach discordant susceptibility results? *J. Clin. Microbiol.* 59, e00138–e00220. doi: 10.1128/JCM.00138-20



OPEN ACCESS

EDITED BY
Alberto Antonelli,
University of Florence, Italy

REVIEWED BY
Dolla Karam Sarkis,
Saint Joseph University of Beirut,
Lebanon
Giulia Menchinelli,
Catholic University of the Sacred
Heart, Rome, Italy

*CORRESPONDENCE
Tsegaye Sewunet
tsegaye.sewunet@ki.se;
tsegishs2010@gmail.com

SPECIALTY SECTION
This article was submitted to
Antimicrobials, Resistance
and Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 24 May 2022
ACCEPTED 17 August 2022
PUBLISHED 20 September 2022

CITATION
Sewunet T, Asrat D, Woldeamanuel Y,
Aseffa A and Giske CG (2022)
Molecular epidemiology
and antimicrobial susceptibility
of *Pseudomonas* spp.
and *Acinetobacter* spp. from clinical
samples at Jimma medical center,
Ethiopia.
Front. Microbiol. 13:951857.
doi: 10.3389/fmicb.2022.951857

COPYRIGHT
© 2022 Sewunet, Asrat,
Woldeamanuel, Aseffa and Giske. This
is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Molecular epidemiology and antimicrobial susceptibility of *Pseudomonas* spp. and *Acinetobacter* spp. from clinical samples at Jimma medical center, Ethiopia

Tsegaye Sewunet^{1*}, Daniel Asrat²,
Yimtubezinash Woldeamanuel², Abraham Aseffa³ and
Christian G. Giske^{1,4}

¹Division of Clinical Microbiology, Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden, ²Department of Microbiology, Immunology and Parasitology, Addis Ababa University, Addis Ababa, Ethiopia, ³Armauer Hansen Research Institute, Addis Ababa, Ethiopia, ⁴Department of Clinical Microbiology, Karolinska University Hospital, Stockholm, Sweden

Introduction: *Pseudomonas aeruginosa* (*P. aeruginosa*) and *Acinetobacter baumannii* (*A. baumannii*) can cause difficult-to-treat infections. We characterized molecular epidemiology of ceftazidime-resistant *P. aeruginosa* and carbapenem-resistant *A. baumannii* at a tertiary hospital in Ethiopia.

Materials and methods: Non-fermenting gram-negative bacilli ($n = 80$) isolated from admitted patients were subjected for species identification by MALDI-TOF. *Pseudomonas* species resistant to ceftazidime or meropenem, and *Acinetobacter* species resistant to meropenem, or imipenem were selected for whole genome sequencing. DNA extracted with EZ1 Advanced XL instrument (Qiagen, Hilden, Germany) was sequenced on Illumina (HiSeq2500) using libraries prepared by NEXTRA-kits (Illumina). Raw reads were assembled using SPAdes 3.13.0, and assembled genomes were used to query databases for resistome profile and sequence types.

Result: Among *Pseudomonas* species isolated, 31.7% (13/41), and 7.3% (3/41) were non-susceptible to ceftazidime, and meropenem, respectively. Carbapenem-resistance was 56.4% (22/39) among *Acinetobacter* species. Moreover, 92% (12/13) of *Pseudomonas* species non-susceptible to ceftazidime and/or meropenem, and 89.4% (17/19) of *Acinetobacter* species encoded multiple resistance genes for at least three classes of antimicrobials. The prevalent β -lactamase genes were *bla*_{OXA-486} (53.8%, 7/13), *bla*_{CTX-M-15} (23.0%, 3/13) among *Pseudomonas*, and *bla*_{GES-11} (57.8%, 11/19) among *Acinetobacter*. The *bla*_{OXA-51-like} β -lactamase, *bla*_{OXA-69} (63.1%, 12/19) was the most prevalent carbapenemase gene among *Acinetobacter* isolates. Single isolates from both *P. aeruginosa*, and *A. baumannii* were detected with the *bla*_{NDM-1}. Sequence type (ST)1 *A. baumannii* and ST274 *P. aeruginosa* were the prevalent sequence types. A cgMLST analysis of the ST1 *A. baumannii*

isolates showed that they were closely related and belonged to the international clonal complex one (ICC1). Similarly, ST274 *P. aeruginosa* isolates were clonally related.

Conclusion: The prevalence of MDR isolates of *Pseudomonas* and *Acinetobacter* spp. was high. *A. baumannii* isolates were clonally spreading in the admission wards at the hospital. Emergence of *bla*_{NDM-1} in the intensive care, and surgical wards of the hospital is a severe threat that requires urgent intervention.

KEYWORDS

ESBLs, carbapenemase, *bla*_{CTX-M-15}, *bla*_{GES-11}, *bla*_{NDM-1}, *P. aeruginosa*, *A. baumannii*, Ethiopia

Introduction

Pseudomonas aeruginosa (*P. aeruginosa*) and *Acinetobacter baumannii* (*A. baumannii*) are among the main causes of nosocomial infections (De Oliveira et al., 2020). They belong to the group of bacteria known as “ESKAPE” (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *A. baumannii*, *P. aeruginosa*, and *Enterobacter* species). Moreover, these group of bacteria are difficult to treat in most cases (Bergogne-Bérézin and Towner, 1996; De Oliveira et al., 2020; Ma et al., 2020).

The prevalence of antimicrobial resistant sub-populations of these strains has rapidly increased over the last few decades (Wong et al., 2017; Horcajada et al., 2019). Carbapenemase-producing *A. baumannii* (CRAB) and *P. aeruginosa* were listed top two of the three critical priority pathogens for which new antimicrobials are urgently needed (WHO, 2017). Several *Acinetobacter* and *Pseudomonas* species were previously reported from different clinical samples from both animal, and human infections (Chusri et al., 2014; Wong et al., 2017; Agnese et al., 2018). Many of them were resistant to multiple classes of antibiotics primarily by several intrinsic resistance mechanism they encode, and secondly by acquired resistance mechanisms (Bello-López et al., 2020; Meng et al., 2020). Studies have also shown that multidrug-resistant strains of *P. aeruginosa* and *A. baumannii* were the main drivers of hospital-acquired infections (Djahmi et al., 2014; Eichenberger and Thaden, 2019; Kazmierczak et al., 2020). A recent review of global epidemiology of carbapenemase-producing Gram-negative bacteria reported that carbapenemase-producing *P. aeruginosa* (CRPA) were associated with high mortality and morbidity among hospitalized patients with pneumonia and bloodstream infections in the United States (Brink, 2019). Though regional variations are common, colonization by carbapenem-resistant *A. baumannii* increased the risk of acquisition of bloodstream infection four-fold (Munoz-Price et al., 2016;

Bassetti et al., 2017). In low-income countries like Ethiopia, comprehensive microbiological data is lacking.

The classical phenotyping methods commonly used in low-income countries cannot reliably define mechanism of resistance in both *Pseudomonas* and *Acinetobacter* species. Lack of sufficient standardized genotypic methods for detection and tracking of multidrug-resistant or extensively drug-resistant isolates was one of the challenges to understand epidemiology of antimicrobial resistance in sub-Saharan African countries (Eichenberger and Thaden, 2019). In most cases, global reports on antimicrobial resistance lack data from African countries. Despite discrepancies in availability of data, there is sufficient overall evidence that carbapenemase-producing Gram-negative bacilli have become a threat to global health. Rapid detection and tracking of any ongoing spread of resistant strains is necessary.

We aimed to analyze the phenotypic and molecular characteristics of *Pseudomonas* species and *Acinetobacter* species isolated from clinical samples at Jimma Medical Center (JMC), a tertiary hospital in Ethiopia.

Materials and methods

Isolation, identification, and selection of strains

As part of a large epidemiological study, a total of 1,087 clinical samples (urine, stools, wound secretions, and sputum) were collected from patients with suspected infections seeking medical care from June to October 2016 at JMC, Ethiopia. *Pseudomonas* species and *Acinetobacter* species were isolated on MacConkey agar and sheep blood agar. Species identification was performed by MALDI-TOF (Bruker Daltonik GmbH, Bremen, Germany) at Karolinska University Hospital (KUH), Clinical Microbiology laboratory, and a full panel of

TABLE 1A Socio-demographic and clinical characteristics patients and isolation *Pseudomonas* species.

Patient ID	Age	Sex	Inpatient/ outpatient	Current diagnosis	*Current antibiotic	Specimen	Underlying disease	<i>Pseudomonas</i> species
I020	28	M	Inpatient	Surgical site infection	CRO, MET	Wound swab	Surgical incision	<i>P. aeruginosa</i>
I032	28	M	Inpatient	Urinary tract infection	CRO, MET	Urine	Trauma	<i>P. aeruginosa</i>
I038	30	M	Inpatient	Urinary tract infection	CRO	Urine	Severe head injury	<i>P. putida</i>
I043	17	M	Inpatient	Urinary tract infection	CRO	Urine	Aspiration pneumonia	<i>P. aeruginosa</i>
M019	70	M	Inpatient	COPD	CRO, VAN	Sputum	Cor pulmonale	<i>P. aeruginosa</i>
M030	60	M	Inpatient	Community- acquired pneumonia	CRO	Sputum	**COPD	<i>P. aeruginosa</i>
M074	50	M	Inpatient	COPD	VAN, CIP	Sputum	Asthma	<i>P. aeruginosa</i>
M119	40	M	Inpatient	Pneumonia	CRO	Sputum	Post TB fibrosis	<i>P. aeruginosa</i>
M304	40	M	Inpatient	Community- acquired pneumonia	No	Sputum	Post TB fibrosis	<i>P. aeruginosa</i>
M334	35	F	Inpatient	Severe community-acquired pneumonia	CRO	Sputum	No	<i>P. aeruginosa</i>
M521	60	F	Outpatient	Community- acquired pneumonia	CRO	Sputum	**T2DM	<i>P. aeruginosa</i>
P014	14	M	Inpatient	Wound infection	AUG	Wound swab	No	<i>P. aeruginosa</i>
P109	4	F	Inpatient	Diarrhea	AMOX, GENT	Stool	SAM, pneumonia	<i>P. aeruginosa</i>
S007	32	M	Inpatient	Necrotizing fasciitis	No	Wound swab	No	<i>P. aeruginosa</i>
S010	5	M	Inpatient	Wound infection	CAF, CLO	Wound swab	Trauma	<i>P. aeruginosa</i>
S011	7	M	Inpatient	Wound infection	AMP, CAF	Wound swab	Trauma	<i>P. aeruginosa</i>
S017	18	M	Inpatient	Necrotizing fasciitis	CRO, MET	Wound swab	No	<i>P. aeruginosa</i>
S019	60	M	Inpatient	Foot ulcer	CRO, MET	Wound swab	Diabetes mellitus	<i>P. aeruginosa</i>
S020	30	M	Inpatient	Wound infection	CRO, MET	Wound swab	Trauma	<i>P. aeruginosa</i>
S036	17	F	Inpatient	Wound infection	No	Wound swab	Unstable pelvis fracture	<i>P. aeruginosa</i>
S047	10	M	Inpatient	Wound infection	AMP	Wound swab	Chronic osteomyelitis	<i>P. aeruginosa</i>
S048	15	M	Inpatient	Wound infection	AMP, CAF	Wound swab	Fracture of femoral shaft	<i>P. aeruginosa</i>
S077	3	M	Inpatient	Wound infection	AMP, GENT	Wound swab	Colostomy, imperforation	<i>P. aeruginosa</i>
S116	50	F	Inpatient	Wound infection	AMOX, MET	Wound swab	Uterine cancer	<i>P. aeruginosa</i>
S129	77	M	Inpatient	Wound infection	CRO, VAN	Wound swab	Amputation of leg	<i>P. aeruginosa</i>
S133	45	M	Inpatient	Wound infection	CRO, MET	Wound swab	Neck injury trauma	<i>P. aeruginosa</i>
S114	25	F	Inpatient	Acute kidney infection	CRO, MET	Urine	Rib fracture	<i>P. aeruginosa</i>
S155	21	F	Inpatient	Wound infection	AMP, CAF	Wound swab	Infected skin graft	<i>P. aeruginosa</i>
S174	60	M	Inpatient	Wound infection	CAF, AMP	Wound swab	Infected fracture site	<i>P. aeruginosa</i>
S192	75	M	Inpatient	Urinary tract infection	CRO	Urine	**BOO	<i>P. putida</i>
S195	21	F	Inpatient	Wound infection	CRO, MET	Wound swab	Skin graft infection	<i>P. aeruginosa</i>
S198	57	M	Outpatient	Wound infection	CLO, CAF	Wound swab	Left femoral fracture	<i>P. aeruginosa</i>

(Continued)

TABLE 1A (Continued)

Patient ID	Age	Sex	Inpatient/ outpatient	Current diagnosis	*Current antibiotic	Specimen	Underlying disease	<i>Pseudomonas</i> species
S209	18	F	Inpatient	Wound infection	CRO	Wound swab	Burn wound	<i>P. aeruginosa</i>
S248	20	F	Inpatient	Contaminated wound		Wound swab	No	<i>P. aeruginosa</i>
S288	50	M	Inpatient	Pneumonia	No	Sputum	Abdominal mass	<i>P. aeruginosa</i>
S319	40	M	Inpatient	Urinary tract infection	No	Urine	BPH	<i>P. aeruginosa</i>
S325	37	M	Inpatient	Wound infection	CAF, AMP	Wound swab	Compound distal fracture	<i>P. aeruginosa</i>
S328	60	M	Inpatient	Pneumonia	CRO, MET	Sputum	3rd degree burn	<i>P. aeruginosa</i>
S332	30	M	Inpatient	Wound infection	CRO, MET	Wound swab	2nd degree burn	<i>P. fulva</i>
S356	30	M	Inpatient	Wound infection	CRO, MET	Wound swab	Infected palate of right knee	<i>P. aeruginosa</i>
S371	64	M	Inpatient	Surgical site infection	CLO, CAF	Wound swab	Laparotomy	<i>P. aeruginosa</i>

*Current antibiotics: CRO, ceftriaxone; Met, metronidazole; VAN, vancomycin; CLO, cloxacillin; CAF, chloramphenicol; CIP, ciprofloxacin; AMP, ampicillin; GENT, gentamicin.

**Underlying diseases: COPD, congestive obstructive pulmonary disease; T2DM, type-2 diabetes mellitus; BOO, Bladder outlet obstruction; BPH, Benign prostatic hyperplasia.

antimicrobial susceptibility testing was performed by using the EUCAST 2021 v11 guideline.¹

Antimicrobial susceptibility testing

All *Pseudomonas* species and *Acinetobacter* species isolated were subjected to disk-diffusion susceptibility testing. Antibiotic discs of ceftazidime, meropenem, piperacillin-tazobactam, gentamicin, amikacin, ciprofloxacin was used for *Pseudomonas* species. Similarly, all *Acinetobacter* isolates were tested by using meropenem, imipenem, gentamicin, amikacin, ciprofloxacin, trimethoprim-sulfamethoxazole. Then, isolates with reduced susceptibility to ceftazidime and/or meropenem for *Pseudomonas* spp., and meropenem or imipenem for *Acinetobacter* spp. were selected for antimicrobial susceptibility testing using the newer antimicrobials (cefiderocol, ceftazidime-avibactam, ceftolozane-tazobactam, and imipenem-relebactam for *Pseudomonas* spp., and cefiderocol, meropenem, imipenem, and imipenem-relebactam for *Acinetobacter* spp. using microbroth dilution technique), and whole genome sequencing (WGS). Patients' clinical data like admission, presence of underlying chronic illnesses, current use of antibiotics, and other factors was collected using a structured questionnaire.

DNA extraction, whole genome sequencing and analysis of genomic data

Genomic DNA was extracted using Qiagen kits on EZ1 automated DNA extractions system. The extracted

DNA was quantified using Qubit™ 3.0 (Massachusetts, United States) and library preps were performed using NEXTRA-kit (Illumina) and sequenced using HiSeq2500 (Illumina). Raw reads were assembled using SPAdes ver. 3.13.0, and the assembled draft genomes were used for querying different databases, MLST-typing 2.0, hosted at center for genomic epidemiology, and detection resistome profile by using ResFinder 4.1.^{2,3} Epidemiologic analysis of relatedness between the isolates, and to other international isolates was performed by the minimum spanning tree using the isolate genomes deposited at the public domain for *A. baumannii* at⁴, and *P. aeruginosa* at.⁵ The Genome sequences were deposited at the NCBI, SRA database (Bioproject number: PRJNA593604, Biosample accession: SUB11593554).

Results

Clinical and demographic characteristics of the patient

From a total of 1,087 non-repeat clinical samples collected during the study period, non-duplicate, non-fermenting Gram-negative bacilli that belong to either *Pseudomonas* spp. or *Acinetobacter* spp. were isolated from 80 patients. Most of these patients, 73.7% (59/80) were male, and 26.3% (21/80) were female. Ninety percent (72/80) of these patients were

1 www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_11.0_Breakpoint_Tables.pdf

2 <https://cge.food.dtu.dk/services/MLST/>

3 <https://cge.food.dtu.dk/services/ResFinder/>

4 https://pubmlst.org/bigsdb?db=pubmlst_abaumannii_isolates&page=query&genomes=1

5 <https://pubmlst.org/organisms/pseudomonas-aeruginosa>

TABLE 1B Socio-demographic and clinical characteristics patients and isolation *Acinetobacter* species.

Patient ID	Age	Sex	Inpatient/ outpatient	Current diagnosis	*Current antibiotic	Specimen	Underlying disease	Bacterial species
I027	45	F	Inpatient	Aspiration pneumonia	CRO, MET	Sputum	Stroke/hemiparalysis	<i>A. baylyi</i>
I030	28	F	Inpatient	Skin infection	No	Wound swab	No	<i>A. baumannii</i>
M029	22	M	Inpatient	Urinary tract infection	CRO	Urine	Retroviral infection	<i>A. baumannii</i>
M057	60	F	Inpatient	COPD	No	Sputum	No	<i>A. baumannii</i>
M135	70	F	Outpatient	Community-acquired pneumonia	CRO	Sputum	No	<i>A. junii</i>
M212	25	M	Inpatient	Pneumonia	CRO, VAN	Sputum	Disseminated tuberculosis	<i>A. calcoaceticus</i>
M217	35	M	Inpatient	Community-acquired pneumonia	No	Sputum	**COPD	<i>A. schindleri</i>
M328	50	F	Outpatient	Severe community-acquired pneumonia	No	Sputum	**T2DM	<i>A. baumannii</i>
M344	17	M	Outpatient	Pneumonia	No	Sputum	Electrical burn	<i>A. junii</i>
M431	50	F	Inpatient	Pneumonia	CRO	Sputum	Rt&Lt femoral fracture	<i>A. baumannii</i>
P015	30	M	Inpatient	Wound infection	CLO, CAF	Wound swab	Compound fracture of right tibia	<i>A. baumannii</i>
P038	35	M	Outpatient	Wound infection	AMP, CAF	Wound swab	Infected fracture site	<i>A. baumannii</i>
S073	15	M	Inpatient	Wound infection	CRO	Wound swab	Chronic osteomyelitis	<i>A. baumannii</i>
S082	45	M	Inpatient	Wound infection	CIP	Wound swab	Femoral fracture	<i>A. baumannii</i>
S126	40	M	Inpatient	Wound infection	CRO, MET	Wound swab	Surgical site infection	<i>A. baumannii</i>
S130	30	M	Inpatient	Wound infection	CRO, MET	Wound swab	3rd deg. burn	<i>A. baumannii</i>
S147	30	F	Inpatient	Wound infection	CRO	Wound swab	Colostomy	<i>A. baumannii</i>
S161	50	M	Inpatient	Wound infection	AMP, CAF	Wound swab	No	<i>A. parvus</i>
S165	8	M	Inpatient	Wound infection	CAF, CLO	Wound swab	No	<i>A. baumannii</i>
S167	40	M	Inpatient	Wound infection	CRO, MET	Wound swab	Scalp abscess	<i>A. baumannii</i>
S170	27	M	Inpatient	Wound infection	CRO, MET	Wound swab	No	<i>A. baumannii</i>
S171	32	M	Inpatient	Wound infection	CRO, MET	Wound swab	Retroviral infection	<i>A. baumannii</i>
S176	25	F	Inpatient	Necrotic wound infection	CRO, MET	Wound swab	No	<i>A. baumannii</i>
S209	50	M	Outpatient	Pneumonia	CRO, MET	Sputum	No	<i>A. baumannii</i>
S210	25	M	Inpatient	Wound infection	CRO	Wound swab	No	<i>A. baumannii</i>
S212	40	M	Inpatient	Wound infection	CRO, MET	Wound swab	Compound fracture left leg	<i>A. baumannii</i>
S217	48	M	Inpatient	Pneumonia	CRO, MET	Sputum	colostomy/post-operation	<i>A. baumannii</i>
S219	39	M	Inpatient	Wound infection	No	Wound swab	Tibia-fibular fracture	<i>A. baumannii</i>
S226	24	M	Inpatient	Urinary tract infection	No	Urine	Urethral stricture	<i>A. haemolyticus</i>
S247	40	M	Outpatient	Wound infection	AMP, CAF	Wound swab	Infected incision site	<i>A. baumannii</i>
S260	18	F	Inpatient	Wound infection	CRO	Wound swab	Burn wound	<i>A. baumannii</i>
S267	30	F	Inpatient	Wound infection	CRO, MET	Wound swab	Surgical site infection	<i>A. baumannii</i>
S270	26	M	Inpatient	Wound infection	CRO, MET	Wound swab	Chronic osteomyelitis	<i>A. baumannii</i>
S275	48	M	Inpatient	Wound infection	No	Wound swab	Compound fracture of femur	<i>A. baumannii</i>
S294	70	M	Inpatient	Urinary tract infection	CRO	Urine	Bladder outlet obstruction	<i>A. baumannii</i>

(Continued)

TABLE 1B (Continued)

Patient ID	Age	Sex	Inpatient/ outpatient	Current diagnosis	*Current antibiotic	Specimen	Underlying disease	Bacterial species
S296	10	M	Inpatient	Wound infection	AMP, CAF	Wound swab	Left femoral fracture	<i>A. baumannii</i>
S315	2	M	Inpatient	Diarrhea	AMP, GENT, CLO	Stool	No	<i>A. lwoffii</i>
S327	29	M	Inpatient	Severe community-acquired pneumonia	CRO, MET	Sputum	No	<i>A. baumannii</i>
S347	18	F	Inpatient	Urinary tract infection	AMP, CAF	Urine	Cervical Spine fracture	<i>A. baumannii</i>

*Current antibiotics: CRO, ceftriaxone; Met, metronidazole; VAN, vancomycin; CLO, cloxacillin; CAF, chloramphenicol; CIP, ciprofloxacin; AMP, ampicillin; GENT, gentamicin.

**Underlying diseases: COPD, congestive obstructive pulmonary disease; T2DM, type-2 diabetes mellitus.

admitted to different wards in the hospital [surgical ward ($n = 52$), intensive care unit ($n = 6$), pediatric ward ($n = 3$), and medical ward ($n = 11$)]. Clinical samples collected include urine for urinary tract infections, sputum for lower respiratory tract infections, wound swab for wound infections/abscess, and stools for diarrhea. Most of these patients were admitted to the hospital for other underlying diseases and developed infection after admission, and most of them were treated with locally available antimicrobial therapy. Among *Pseudomonas* isolates 68.2% (28/41) were from surgical ward, followed by 17.0% (7/41) from medical ward, and 9.7% (4/41) from the intensive care unit (Table 1A). Similarly, 69.2% (27/39) of the *Acinetobacter* species were isolated from the surgical ward and 20.5% (8/39) from the medical ward (Table 1B).

Species diversity and antimicrobial susceptibility

Species diversity

From a total of 41 isolates of *Pseudomonas* spp., 92.6% (38/41) were *P. aeruginosa*, and 7.3% (3/41) were another *Pseudomonas* spp. Among 39 isolates of *Acinetobacter* spp. ($n = 39$), 79.5% (31/39) were *A. baumannii*, and 20.5% (8/39) were other *Acinetobacter* spp. [*A. junii* which account for 5.1% (2/39), and *A. baylyi*, *A. lwoffii*, *A. calcoaceticus*, *A. haemolyticus*, *A. parvus*, and *A. schindleri* each account for 2.5% (1/39)].

Antimicrobial susceptibility pattern

Among isolates of *Pseudomonas* spp., 31.7% (13/41) were non-susceptible to ceftazidime, and 7.3% (3/41) to meropenem. Among isolates of *Acinetobacter* spp., 56.4% (22/39) were carbapenem-resistant. Most of the ceftazidime-resistant *Pseudomonas* spp. were also resistant to piperacillin-tazobactam 84.6% (11/13), ciprofloxacin 30.7% (4/13), ceftazidime-avibactam 46.1% (6/13) and ceftolozane-tazobactam 53.8%

(7/13). However, most of these isolates were susceptible to cefiderocol 84.7% (11/13), imipenem-relebactam 92.3% (12/13), and all *Pseudomonas* isolates were susceptible to amikacin (Table 2A). Most of the *Acinetobacter* isolates were resistant to meropenem, imipenem, and imipenem-relebactam 54.5%, (12/22). However, a lower rate of resistance was detected to ciprofloxacin 18.1% (4/22), gentamicin 27.2% (6/22), amikacin 18.1% (4/22), and cefiderocol 9.1% (2/22) (Table 2B).

Resistome profile of *Pseudomonas* and *Acinetobacter* species

Resistome profiling showed that both *Pseudomonas* and *Acinetobacter* spp. encoded multiple β - lactamase genes. The *bla*_{OXA-486}, *bla*_{OXA-50} and *bla*_{CTX-M-15} genes were most common among *Pseudomonas* isolates, and the *bla*_{OXA-51}-like *bla*_{OXA-69}, *bla*_{OXA-66}, *bla*_{OXA-91}, *bla*_{OXA-180} and *bla*_{GES-11} were the most common among *Acinetobacter* isolates (Table 3A).

Moreover, two isolates, *P. aeruginosa* ($n = 1$) and *A. baumannii* ($n = 1$) carrying the *bla*_{NDM-1} carbapenemase gene were detected. Furthermore, the resistome profile shows that resistance genes to the antimicrobial classes of aminoglycosides, fluoroquinolones, and trimethoprim were prevalent among these two isolates (Tables 3A,B).

Molecular epidemiology using cgMLST

Epidemiologic typing using the seven gene multi-locus sequence typing demonstrated that ST274 (23.0%, 3/13) among *P. aeruginosa* and ST1 (MLST-Pasteur) *A. baumannii* (63.1%, 12/19) were the most prevalent sequence types (Tables 3A,B). A cgMLST analysis showed that the ST1 *A. baumannii* isolates were highly similar with no allelic differences between them. Most of the *Acinetobacter* isolates belonged to the international

TABLE 2A Antimicrobial susceptibility pattern of ceftazidime resistant *Pseudomonas* spp. isolated at JMC, Ethiopia.

Bacterial species		Type of antimicrobials used for antimicrobial susceptibility testing									
		Ceftazidime	Meropenem	Imipenem	Piperacillin-tazobactam	Ciprofloxacin	Amikacin	Cefiderocol	Ceftazidime-avibactam	Ceftolozane-tazobactam	Imipenem-relebactam
<i>Pseudomonas aeruginosa</i> (n = 10)	S (%)	0 (0.0%)	7 (70)	8 (80)	0 (0)	7 (70)	10 (100)	9 (90)	5 (50)	6 (60)	9 (90)
	R (%)	10 (100)	3 (30)	2 (20)	10 (100)	3 (30)	0 (0)	1 (10)	5 (50)	4 (40)	1 (10)
Other <i>Pseudomonas</i> species (n = 3)	S (%)	1 (32.4)	1 (32.4)	3 (100)	1 (32.4)	2 (66.6)	3 (100)	2 (66.6)	2 (66.6)	0 (0)	3 (100)
	R (%)	2 (66.6)	2 (66.6)	0 (0)	2 (66.6)	1 (32.4)	0 (0)	1 (32.4)	1 (32.4)	3 (100)	0 (0)

S, susceptible; R, resistant, other *Pseudomonas* species [*Pseudomonas putida* (n = 2), *Pseudomonas fulva* (n = 1)].

TABLE 2B Antimicrobial susceptibility pattern of carbapenem resistant *Acinetobacter* species isolated at JMC, Ethiopia.

Bacterial species		Type of antimicrobials used for antimicrobial susceptibility testing							
		Meropenem	Imipenem	Imipenem-relebactam	Cefiderocol	Ciprofloxacin	Trimethoprim/sulfamethoxazole	Gentamicin	Amikacin
<i>Acinetobacter baumannii</i> (n = 22)	S (%)	5 (27.8)	9 (50)	9 (50)	17 (94.4)	14 (77.8)	12 (66.7)	12 (66.7)	15 (83.3)
	R (%)	13 (72.2)	9 (50)	9 (50)	1 (5.6)	4 (22.2)	6 (33.3)	6 (33.3)	3 (16.7)
Other <i>Acinetobacter</i> spp.	S (%)	1 (25)	1 (25)	1 (25)	4 (100)	1 (25)	0 (0)	4 (100)	4 (100)
	R (%)	3 (75)	3 (75)	3 (75)	0 (0)	3 (75)	4 (100)	0 (0)	(100)

S, susceptible; R, resistant; Other *Acinetobacter* spp. (n = 4). (*A. calcoaceticus*, *A. baylyi*, *A. junii*, and *A. lwoffii*).

clonal complexes, ICC1 (includes ST1) ($n = 12$) and ICC2 ($n = 2$) (Table 3A).

Discussion

In this study, nearly all isolates of both *Pseudomonas* and *Acinetobacter* spp. were from patients admitted to the hospital for more than 72 h. Nosocomial acquisition of MDR isolates of *Pseudomonas* spp. and *Acinetobacter* spp. is worrisome. The situation is complicated by limited availability of antimicrobial agents, lack of prescription guidelines, and insufficient standard routine microbiology laboratory services to support antibiotic selection. In such cases, the safety of patients admitted to the hospital can be severely compromised.

The prevalence of ESBL-producing strains among *P. aeruginosa*, and carbapenem non-susceptible isolates among *A. baumannii* strains was high in the present study. Previous phenotypic studies from Ethiopia also show that MDR strains of *Pseudomonas* and *Acinetobacter* were prevalent (Motbainor et al., 2020; Birhane Fiseha et al., 2021). However, most of these studies did not describe the genotypes and resistome profile of the isolates, and mechanism of resistance is difficult to compare between studies. So far, to our knowledge, there is no report of a genotypic study on the prevalence of ESBL-producing *P. aeruginosa* in Ethiopia. Generally, there are limited studies conducted on carbapenemase-producing *P. aeruginosa* in Africa. The few reports available are from northern Africa and mainly from Egypt (Osei Sekyere and Reta, 2020; Rizk et al., 2021). In northern Africa, the prevalence ranges from 0 to 96% (Gaballah et al., 2018). A finding from Uganda (7.4%) was comparable to the present study (Aruhomukama et al., 2019). But, one study from South Africa (51.0%) reported a higher prevalence of ceftazidime resistant *P. aeruginosa* compared to the present study (Hosu et al., 2021). The phenotypic studies from Ethiopia, and other African countries showed higher prevalence as compared to the present study.

In Ethiopia, a high prevalence of carbapenem-resistant *Acinetobacter* spp. was reported from one previous phenotypic study (Ayenew et al., 2021), which is comparable to the prevalence of carbapenem resistance herein. On the other hand, a systematic review and meta analysis on carbapenemase producing *P. aeruginosa* and *A. baumannii* in Africa showed that the lowest prevalence of carbapenemase-producing *A. baumannii* was 4.7% ($n = 21$), and the highest prevalence was 100% ($n = 7$) (Kindu et al., 2020). Studies conducted on *P. aeruginosa* and *A. baumannii* strains from Africa were limited to small sample sizes and were mainly phenotypic studies, which makes the comparison with genotypic studies difficult (Kindu et al., 2020; Osei Sekyere and Reta, 2020; Ayenew et al., 2021;

Mekonnen et al., 2021). However, most available studies including a study for hospital environment (Solomon et al., 2017) reported higher prevalence of the MDR *P. aeruginosa* and *A. baumannii* in Ethiopia, calling for the application of genotypic methods to studies on mechanisms of resistance and spread.

The multiple genetic variants of antibiotic resistance observed among both *Pseudomonas* and *Acinetobacter* spp. pose a huge challenge on the limited therapeutic options available to low-income countries. Among *Acinetobacter* spp., the presence of the *bla*_{GES-11} (ESBL-genotype and weak carbapenemase) and the OXA-51-like (*bla*_{OXA-66} and *bla*_{OXA-69}) carbapenemases encoding genes is a serious threat. The OXA-51-like intrinsic carbapenemase encoding ST1 *A. baumannii* Isolates were reported from India (Rose et al., 2021), but isolates from the present study encoded additionally the weakly carbapenem hydrolyzing *bla*_{GES-11} gene. The *bla*_{GES-11} ESBL-genotypes, and the *bla*_{OXA-51}-like carbapenemases have not been previously reported from Ethiopia, and the present study is to our knowledge the first report of *bla*_{OXA-51} like and *bla*_{GES-11} from Ethiopia. Generally from Africa, only one study from Tunisia has documented clinical isolates of *A. baumannii* encoding the *bla*_{GES-11} (Chihi et al., 2016). The emergence of *bla*_{NDM-1} encoding isolates of *A. baumannii* at surgical ward, and *P. aeruginosa* at intensive care unit can severely compromise the safety of vulnerable patients admitted to the hospital. Moreover, detection of two isolates encoding the *bla*_{NDM-1} which showed resistance to the newer antimicrobials, *A. baumannii* for (cefiderocol and imipenem-relebactam), and *P. aeruginosa* for (cefiderocol, ceftazidime-avibactam, ceftolozane-tazobactam) compromises the already limited treatment alternatives for vulnerable groups of patients at this hospital.

Sequence typing showed that both *P. aeruginosa* and *A. baumannii* strains were polyclonal. But, since a large proportion of *A. baumannii* isolates were ST1, the spread of *A. baumannii* isolates at this hospital might also be to some extent clonal. A previous study conducted at the same hospital had identified three strains of *A. baumannii* encoding *bla*_{NDM-1}, and all of them belonged to the ST597. Furthermore, cgMLST analysis of *A. baumannii* with other international isolates in pubmlst showed that isolates in the current study were distinct from isolates from other African countries (Figure 1).⁶ However, these isolates were clustered with isolates from other countries like United States, and Brazil. Though isolates from this study were polyclonal, the most prevalent isolates were those that belong to the international cluster, CC1 and CC2.

Although *A. baumannii* is a well-known major cause of nosocomial infections, knowledge of its genomic epidemiology

⁶ <https://pubmlst.org/organisms/acinetobacter-baumannii>

TABLE 3A Sequence types and resistance genes observed among carbapenem-resistant *Acinetobacter* species isolated at Jimma Medical Center, Ethiopia ($n = 19$).

Isolate ID	STs	Number and type of antimicrobial resistance genes						
		Carbapenemase	ESBL and other β -lactamases	Aminoglycosides	Trimethoprim	Sulfonamides	Tetracycline	Macrolides
I027AB*	ST2	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
I030AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
M217AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A1,	<i>sul</i> 1		
M328AB	164	<i>bla</i> _{OXA-91}	<i>bla</i> _{CARB-5} , <i>bla</i> _{CARB-49}				<i>tet</i> (39)	
P015AL	NA		<i>bla</i> _{CTX-M-15} , <i>bla</i> _{OXA-1} ,	<i>aac</i> (6')-Ib-cr			<i>tet</i> (A)	<i>mdf</i> (A), <i>mph</i> (A), <i>mph</i> (D)
S126AB	ST2090	<i>bla</i> _{OXA-259/180}						
S130AB	ST85	<i>bla</i> _{NDM-1} , <i>bla</i> _{OXA-94}		<i>ant</i> (2'')-Ia, <i>aph</i> (3')-VI		<i>sul</i> 2		<i>mph</i> (E) <i>msr</i> (E)
S161AB	ST2	<i>bla</i> _{OXA-66}	<i>bla</i> _{TEM-1D} ,	<i>aac</i> (3)-Ia, <i>aph</i> (6)-Id		<i>sul</i> 1		
S167AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib-cr, <i>aac</i> (6')-Ib3	<i>dfr</i> A7	<i>sul</i> 1		
S170AB	ST164	<i>bla</i> _{OXA-91}	<i>bla</i> _{CARB-49} , <i>bla</i> _{CARB-16}				<i>tet</i> (39)	
S171AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S176AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S209AB	ST2	<i>bla</i> _{OXA-66}		<i>aac</i> (3)-Ia, <i>aad</i> A1, <i>aph</i> (6)-Id		<i>sul</i> 1 <i>sul</i> 2	<i>tet</i> (B)	
S212AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr,	<i>dfr</i> A7	<i>sul</i> 1		
S270AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S275AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S296AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S315AB	ST1	<i>bla</i> _{OXA-69}	<i>bla</i> _{GES-11}	<i>aac</i> (6')-Ib3, <i>aac</i> (6')-Ib-cr	<i>dfr</i> A7	<i>sul</i> 1		
S327AB	ST1	<i>bla</i> _{OXA-69}						

TABLE 3B Sequence types and resistance genes observed among ceftazidime-resistant *Pseudomonas* spp. isolated at Jimma Medical Center, Ethiopia ($n = 13$).

Strain ID	STs	Number and type of antimicrobial resistance genes						
		Carbapenemase	β -lactamase genes	Aminoglycosides	Fluoroquinolones	Sulfonamides	Phenicol	Fosfomycin
I020PA*	11		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO}	<i>aph</i> (3')-IIb			<i>catB7</i>	<i>fosA</i>
I032PA	2948	<i>bla</i> _{NDM-1}	<i>bla</i> _{OXA-10} , <i>bla</i> _{OXA-50}	<i>aadA1</i> , <i>ant</i> (2'')-Ia, <i>aph</i> (3')-IIb	<i>crpP</i>	<i>sul1</i>	<i>catB7</i> , <i>cmlA1</i>	<i>fosA</i>
I038PP*	NA		<i>bla</i> _{CTX-M-15} , <i>bla</i> _{OXA-1}	<i>aac</i> (6')-Ib-cr,				
M019PA	1228		<i>bla</i> _{VEB-1} , <i>bla</i> _{OXA-486} <i>bla</i> _{OXA-10}	<i>aadA1</i> , <i>aadA2</i> <i>ant</i> (2'')-Ia, <i>aph</i> (3')-IIb	<i>crpP qnrD1</i>	<i>sul1</i>	<i>catB7 cmlA1</i>	<i>fosA</i>
M119PA	274		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO} ,				<i>catB7</i>	<i>fosA</i>
M304PA	274		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO}				<i>catB7</i>	<i>fosA</i>
S116PA	500		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO}		<i>crpP</i>		<i>catB7</i>	<i>fosA</i>
S114PP*	NA		<i>bla</i> _{CTX-M-15} , <i>bla</i> _{OXA-10} <i>bla</i> _{OXA-1}	<i>aac</i> (3)-Iia, <i>aac</i> (3)-Ib <i>aac</i> (6')-Ib, <i>aac</i> (6')-Ib-cr, <i>ant</i> (3'')-Ia, <i>aph</i> (3'')-Ib	<i>oqxP qnrVC1</i>	<i>sul1</i>		
S155PA	274		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO}	<i>aph</i> (3')-IIb			<i>catB7</i>	<i>fosA</i>
S248PA	840		<i>bla</i> _{OXA-486} , <i>bla</i> _{PAO}	<i>aph</i> (3')-IIb	<i>crpP</i>		<i>catB7</i>	<i>fosA</i>
S332PF**	NA		<i>bla</i> _{PER-1} ,	<i>aph</i> (6)-Id	<i>qnrVC6</i>	<i>sul1</i>		
S356PA	646		<i>bla</i> _{CTX-M-15} , <i>bla</i> _{OXA-494} <i>bla</i> _{OXA-50} , <i>bla</i> _{TEM-1B}	<i>ant</i> (2'')-Ia, <i>aph</i> (3')-IIb <i>aph</i> (3')-Ia, <i>ph</i> (6)-Id		<i>sul1</i>	<i>catB7, floR</i>	<i>fosA</i>
S047PA	244		<i>bla</i> _{OXA-50}	<i>aph</i> (3')-IIb			<i>catB7</i>	<i>fosA</i>

*A. baumannii; PA, *Pseudomonas aeruginosa*; PP*, *Pseudomonas putida*; PF*, *Pseudomonas fulva*.

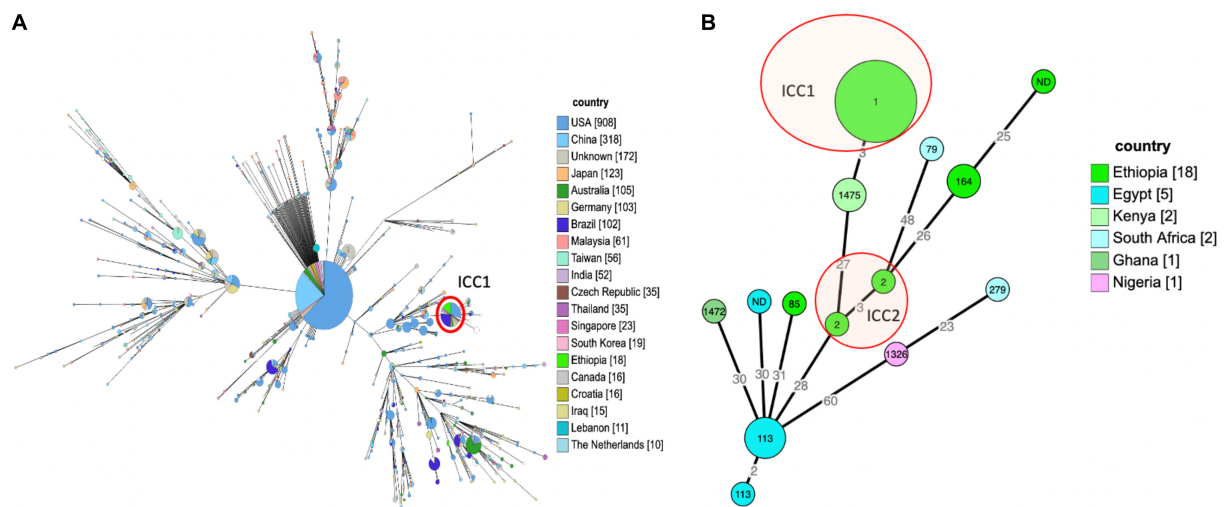


FIGURE 1

(A) Epidemiologic relatedness of *A. baumannii* isolates to the isolates collected from different geographical regions and deposited in the public database (pubmlst <https://pubmlst.org/organisms/acinetobacter-baumannii> accessed on 26-09-2021), the isolates from this study (bright green) were clustered with the global ICC1 clone marked with red circle. The numbers in the circle show the sequence type of the isolates, the size of the circle is proportional to the number of isolates that belongs to that sequence type, and the color shows the country of origin for the isolates. The minimum spanning tree was constructed for isolates from different country of origin having ≥ 10 isolates recorded in the database. (B) Minimum spanning tree of 29 *A. baumannii* isolates from Africa, the tree was constructed based on core genome multi-locus sequence typing. There were no previous isolates from Ethiopia in PubMLST, all isolates labeled Ethiopia in the tree are from this study.

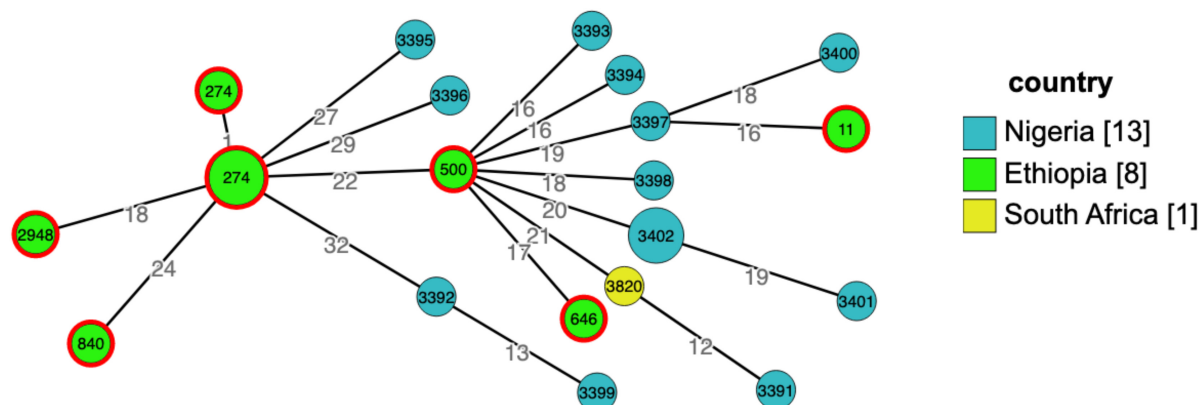


FIGURE 2

Grape tree generated using cgMLST from genomic collection of *Pseudomonas aeruginosa* isolates from Africa ($n = 22$), Ethiopia, Nigeria, and South Africa, pubmlst database for microorganisms (https://pubmlst.org/bigsdb?db=pubmlst_paeruginosa_isolates&page=query&genomes=1) accessed on 29-11-2021. The numbers in circles/nodes indicate sequence types of the isolates, size of the nodes is proportional to the number of isolates in a particular sequence type and the numbers on the branches are the number of allelic differences between the two neighboring nodes. From 10 isolates sequenced in this study, eight were included in the tree, two of the sequences failed quality required to upload to the pubmlst database.

and availability of reliable data regarding the genetic basis of antibiotic resistance is limited in low-income countries. Similarly, *P. aeruginosa* isolates were found to be polyclonal, and different from a collection of isolates other African countries found in pubmlst (see text footnote 5) (Figure 2).

The current study may serve as a baseline regarding local spread of international clones and alert clinicians and other health workers, researchers, and public health

policy makers to the problem. Implementation of strict infection prevention and control strategies, and antimicrobial stewardship programs are highly desirable in the admission wards where the international clones are spreading. Furthermore, despite limitation of resources, the added value of next generation sequencing is in understanding the dynamics and mechanisms of spread of MDR bacterial clones.

Conclusion

The prevalence of MDR isolates is high among both in clinical isolates of *Pseudomonas* species and *Acinetobacter* species at Jimma medical center. Emergence of the *bla*_{NDM-1} in clinical isolates of *P. aeruginosa* and *A. baumannii* strains is worrisome. However, the susceptibility of *P. aeruginosa* strains to amikacin, cefiderocol, imipenem-relebactam and ceftolozane-tazobactam, and *A. baumannii* strains to amikacin and cefiderocol is important to consider as alternative options to carbapenems. The use of next generation sequencing is important to understand the mechanism of resistance and spread of resistant clones such as ICC1, and ICC2 *A. baumannii* strains detected at this hospital.

Data availability statement

The genome sequences were deposited at the NCBI, SRA database (PRJNA593604, Biosample accession: SUB11593554).

Ethics statement

The study obtained ethical approval from Addis Ababa University Institutional Review Board (AAU-IRB), Armauer Hansen Research Institute – ALERT Hospital Institutional Review Board (AHRI-ALERT-IRB), and Ethiopian National Ethics Review Committee (NERC). Patients were informed about the study and given written consent to participate in the study.

Author contributions

TS contributed to design, data acquisition, data analysis, and drafting of the manuscript. DA and YW contributed

to data acquisition and write-up. AA contributed to the design, data acquisition, and revision of the manuscript. CG contributed to the overall design, data acquisition, supervision, drafting, and writing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Swedish International Development Agency (SIDA).

Acknowledgments

We acknowledge Isak Sylvén for technical support for genome assembly for both *P. aeruginosa* and *A. baumannii* isolates.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Agnese, L., Marisa, H. J., and Ean-Yves, M. (2018). Antimicrobial resistance in *Acinetobacter* spp. and *Pseudomonas* spp. *Microbiol. Spectr.* 6. doi: 10.1128/microbiolspec.ARBA-0007-2017
- Aruhomukama, D., Najjuka, C. F., Kajumbula, H., Okee, M., Mboowa, G., Sserwadda, I., et al. (2019). *Bla*VIM- and *bla*OXA-mediated carbapenem resistance among *Acinetobacter baumannii* and *Pseudomonas aeruginosa* isolates from the Mulago hospital intensive care unit in Kampala, Uganda. *BMC Infect. Dis.* 19:853. doi: 10.1186/s12879-019-4510-5
- Ayenew, Z., Tigabu, E., Syoum, E., Ebrahim, S., Assefa, D., and Tsige, E. (2021). Multidrug resistance pattern of *Acinetobacter* species isolated from clinical specimens referred to the Ethiopian Public Health Institute: 2014 to 2018 trend analysis. *PLoS One* 16:e0250896. doi: 10.1371/journal.pone.0250896
- Bassetti, M., Carnelutti, A., and Peghin, M. (2017). Patient specific risk stratification for antimicrobial resistance and possible treatment strategies in gram-negative bacterial infections. *Expert Rev. Anti Infect. Ther.* 15, 55–65. doi: 10.1080/14787210.2017.1251840
- Bello-López, E., Rocha-Gracia, R. D. C., Castro-Jaimes, S., Cevallos, M. Á., Vargas-Cruz, M., Verdugo-Yocupicio, R., et al. (2020). Antibiotic resistance mechanisms in *Acinetobacter* spp. Strains isolated from patients in a paediatric hospital in Mexico. *J. Glob. Antimicrob. Resist.* 23, 120–129. doi: 10.1016/j.jgar.2020.08.014
- Bergogne-Bérézin, E., and Towner, K. J. (1996). *Acinetobacter* spp. As nosocomial pathogens: Microbiological, clinical, and epidemiological features. *Clin. Microbiol. Rev.* 9, 148–165. doi: 10.1128/cmr.9.2.148-165.1996
- Birhane Fiseha, S., Mulatu Jara, G., Azerefeñ Woldetsadik, E., Belayneh Bekele, F., and Mohammed Ali, M. (2021). Colonization rate of potential neonatal disease-causing bacteria, associated factors, and antimicrobial susceptibility profile among pregnant women attending government hospitals

- in Hawassa, Ethiopia. *Infect. Drug Resist.* 14, 3159–3168. doi: 10.2147/idr.s326200
- Brink, A. J. (2019). Epidemiology of carbapenem-resistant Gram-negative infections globally. *Curr. Opin. Infect. Dis.* 32, 609–616.
- Chihi, H., Bonnin, R. A., Bourouis, A., Mahrouki, S., Besbes, S., Moussa, M. B., et al. (2016). GES-11-producing *Acinetobacter baumannii* clinical isolates from Tunisian hospitals: Long-term dissemination of GES-type carbapenemases in North Africa. *J. Glob. Antimicrob. Resist.* 5, 47–50. doi: 10.1016/j.jgar.2016.03.005
- Chusri, S., Chongsuvivatwong, V., Rivera, J. I., Silpapojakul, K., Singkhamanan, K., McNeil, E., et al. (2014). Clinical outcomes of hospital-acquired infection with *Acinetobacter nosocomialis* and *Acinetobacter pittii*. *Antimicrob. Agents Chemother.* 58, 4172–4179. doi: 10.1128/AAC.02992-14
- De Oliveira, D. M. P., Forde, B. M., Kidd, T. J., Harris, P. N. A., Schembri, M. A., Beatson, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* 33:e00181-19. doi: 10.1128/CMR.00181-19
- Djahmi, N., Dunyach-Remy, C., Pantel, A., Dekhil, M., Sotto, A., and Lavigne, J. P. (2014). Epidemiology of Carbapenemase-producing *Enterobacteriaceae* and *Acinetobacter baumannii* in Mediterranean Countries. *Biomed. Res. Int.* 2014:305784. doi: 10.1155/2014/305784
- Eichenberger, E. M., and Thaden, J. T. (2019). Epidemiology and Mechanisms of resistance of extensively drug resistant gram-negative bacteria. *Antibiotics* 8:37. doi: 10.3390/antibiotics8020037
- Gaballah, A., Elbaradei, A., and Elbaradei, A. (2018). Emergence of bla_{VEB} and bla_{GES} among VIM-producing *Pseudomonas aeruginosa* clinical isolates in Alexandria, Egypt. *Acta Microbiol. Immunol. Hung.* 66, 131–142. doi: 10.1556/030.65.2018.044
- Horcajada, J. P., Montero, M., Oliver, A., Sorlí, L., Luque, S., Gómez-Zorrilla, S., et al. (2019). Epidemiology and treatment of multidrug-resistant and extensively drug-resistant *Pseudomonas aeruginosa* infections. *Clin. Microbiol. Rev.* 32:e00031-19. doi: 10.1128/CMR.00031-19
- Hosu, M. C., Vasaikar, S. D., Okuthe, G. E., and Apalata, T. (2021). Detection of extended spectrum b-lactamase genes in *Pseudomonas aeruginosa* isolated from patients in rural Eastern Cape Province, South Africa. *Sci. Rep.* 11, 1–8. doi: 10.1038/s41598-021-86570-y
- Kazmierczak, K. M., de Jonge, B. L. M., Stone, G. G., and Sahm, D. F. (2020). Longitudinal analysis of ESBL and carbapenemase carriage among Enterobacterales and *Pseudomonas aeruginosa* isolates collected in Europe as part of the International Network for Optimal Resistance Monitoring (INFORM) global surveillance programme, 2013–17. *J. Antimicrob. Chemother.* 75, 1165–1173. doi: 10.1093/jac/dkz571
- Kindu, M., Derseh, L., Gelaw, B., and Moges, F. (2020). Carbapenemase-Producing non-glucose-fermenting gram-negative Bacilli in Africa, *Pseudomonas aeruginosa* and *Acinetobacter baumannii*: A systematic review and meta-analysis. *Int. J. Microbiol.* 2020:9461901. doi: 10.1155/2020/9461901
- Ma, Y. X., Wang, C. Y., Li, Y. Y., Li, J., Wan, Q. Q., Chen, J. H., et al. (2020). Considerations and caveats in combating ESKAPE pathogens against nosocomial infections. *Adv. Sci.* 7:1901872. doi: 10.1002/advs.201901872
- Mekonnen, H., Seid, A., Molla Fenta, G., and Gebrecherkos, T. (2021). Antimicrobial resistance profiles and associated factors of *Acinetobacter* and *Pseudomonas aeruginosa* nosocomial infection among patients admitted at Dessie comprehensive specialized Hospital, NorthEast Ethiopia: A cross-sectional study. *PLoS One* 16:e0257272. doi: 10.1371/journal.pone.0257272
- Meng, L., Liu, H., Lan, T., Dong, L., Hu, H., Zhao, S., et al. (2020). Antibiotic resistance patterns of *Pseudomonas* spp. isolated from raw milk revealed by whole genome sequencing. *Front. Microbiol.* 11:1005. doi: 10.3389/fmicb.2020.01005
- Motbainor, H., Bereded, F., and Mulu, W. (2020). Multi-drug resistance of blood stream, urinary tract and surgical site nosocomial infections of *Acinetobacter baumannii* and *Pseudomonas aeruginosa* among patients hospitalized at Felegehiwot referral hospital, Northwest Ethiopia: A cross-sectional study. *BMC Infect. Dis.* 20:92. doi: 10.1186/s12879-020-4811-8
- Munoz-Price, L. S., Rosa, R., Castro, J. G., Laowansiri, P., Latibeaudiere, R., Namias, N., et al. (2016). Evaluating the impact of antibiotic exposures as time-dependent variables on the acquisition of carbapenem-resistant *Acinetobacter baumannii*. *Crit. Care Med.* 44, e949–e956. doi: 10.1097/CCM.0000000000001848
- Osei Sekyere, J., and Reta, M. A. (2020). Genomic and resistance epidemiology of gram-negative bacteria in Africa: A systematic review and phylogenomic analyses from a one health perspective. *mSystems* 5:e00897-20. doi: 10.1128/mSystems.00897-20
- Rizk, S. S., Elwakil, W. H., and Attia, A. S. (2021). Antibiotic-resistant *Acinetobacter baumannii* in low-income countries (2000–2020): Twenty-one years and still below the radar, is it not there or can they not afford to look for it? *Antibiotics* 10:764. doi: 10.3390/antibiotics10070764
- Rose, S., Shamanna, V., Underwood, A., Nagaraj, G., Prasanna, A., Govindan, V., et al. (2021). Molecular dissection of carbapenem-resistant *Acinetobacter baumannii* circulating in Indian hospitals using whole genome sequencing. *bioRxiv* [Preprint]. 451. doi: 10.1101/2021.07.30.454432
- Solomon, F. B., Wadilo, F., Tufa, E. G., and Mitiku, M. (2017). Extended spectrum and metallo b-lactamase producing airborne *Pseudomonas aeruginosa* and *Acinetobacter baumannii* in restricted settings of a referral hospital: A neglected condition. *Antimicrob. Resist. Infect. Control* 6:106. doi: 10.1186/s13756-017-0266-0
- WHO (2017). *WHO publishes list of bacteria for which new antibiotics are urgently needed*. Available online at: <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed> (accessed September 3, 2021).
- Wong, D., Nielsen, T. B., Bonomo, R. A., Pantapalangkoor, P., Luna, B., Spellberg, B., et al. (2017). Clinical and pathophysiological overview of *Acinetobacter* infections: A century of challenges. *Clin. Microbiol. Rev.* 30, 409–447. doi: 10.1128/CMR.00058-16



OPEN ACCESS

EDITED BY

Costas C. Papagiannitsis,
University of Thessaly,
Greece

REVIEWED BY

Tamara Salloum,
Harvard Medical School,
United States
Anna Skalova,
University Hospital in Pilsen, Czechia
Katerina Tsilipounidaki,
University of Thessaly,
Greece

*CORRESPONDENCE

Susanne A. Kraemer
susanne.kraemer@mail.concordia.ca

SPECIALTY SECTION

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 15 July 2022

ACCEPTED 02 September 2022

PUBLISHED 06 October 2022

CITATION

Kraemer SA, Barbosa da Costa N, Oliva A,
Huot Y and Walsh DA (2022) A resistome
survey across hundreds of freshwater
bacterial communities reveals the impacts
of veterinary and human antibiotics use.
Front. Microbiol. 13:995418.
doi: 10.3389/fmicb.2022.995418

COPYRIGHT

© 2022 Kraemer, Barbosa da Costa, Oliva,
Huot and Walsh. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

A resistome survey across hundreds of freshwater bacterial communities reveals the impacts of veterinary and human antibiotics use

Susanne A. Kraemer^{1,2,3*}, Naila Barbosa da Costa⁴,
Anais Oliva⁵, Yannick Huot⁵ and David A. Walsh³

¹Aquatic Contaminants Research Division, Environment and Climate Change Canada, Montreal, QC, Canada, ²Department of Microbiology and Immunology, McGill University, Montreal, QC, Canada, ³Department of Biology, Concordia University, Montreal, QC, Canada, ⁴Department of Biological Sciences, Université de Montréal, Montreal, QC, Canada, ⁵Department of Applied Geomatics, Université de Sherbrooke, Sherbrooke, QC, Canada

Our decreasing ability to fight bacterial infections is a major health concern. It is arising due to the evolution of antimicrobial resistance (AMR) in response to the mis- and overuse of antibiotics in both human and veterinary medicine. Lakes integrate watershed processes and thus may act as receptors and reservoirs of antibiotic resistance genes (ARGs) introduced into the watershed by human activities. The resistome – the diversity of ARGs – under varying anthropogenic watershed pressures has been previously studied either focused on few select genes or few lakes. Here, we link the resistome of ~350 lakes sampled across Canada to human watershed activity, trophic status, as well as point sources of ARG pollution including wastewater treatment plants and hospitals in the watershed. A high percentage of the resistance genes detected was either unimpacted by human activity or highly prevalent in pristine lakes, highlighting the role of AMR in microbial ecology in aquatic systems, as well as a pool of genes available for potential horizontal gene transfer to pathogenic species. Nonetheless, watershed agricultural and pasture area significantly impacted the resistome. Moreover, the number of hospitals and the population density in a watershed, the volume of wastewater entering the lake, as well as the fraction of manure applied in the watershed as fertilizer significantly impacted ARG diversity. Together, these findings indicate that lake resistomes are regularly stocked with resistance genes evolved in the context of both veterinary and human antibiotics use and represent reservoirs of ARGs that require further monitoring.

KEYWORDS

resistome, microbial ecology, next generation sequencing, antibiotic resistance, freshwater

Introduction

The emergence of antimicrobial resistance (AMR) is a major health challenge of global concern threatening the effectiveness of known antibiotics to treat infectious diseases. Over the past decades we have been facing an antibiotic crisis, as AMR increases faster than the discovery of new antibiotics (Lewis, 2020). As a consequence, a recent systematic analysis estimated that 1.27 million deaths worldwide in 2019 were directly attributable to bacterial AMR in 88 pathogen-drug combinations (Murray et al., 2022). Moreover, the treatment associated with resistant infections results in substantial financial burden to countries' economy: for example an estimated increase in Canadian healthcare costs of \$6–\$8 billion per year by 2050 (Council of Canadian Academics, 2019).

Even though it is a current health concern, AMR and the antibiotic resistance genes (ARGs) causing it are ancient phenomena and readily detected in pristine environments (Martínez, 2008; D'Costa et al., 2011; Bhullar et al., 2012; Perron et al., 2015), where they likely serve complex functions including communication and antagonism in competing and cohabiting bacteria (Goh et al., 2002; Skinders et al., 2008). Bacterial AMR results from intrinsic or acquired resistance mechanisms of defense. Intrinsic molecular mechanisms, such as permeability barriers and broad-spectrum efflux pumps, evolved as a general response to nonspecific toxic compounds, and often confer resistance against multiple antibiotics (multidrug resistance; Surette and Wright, 2017). Alternatively, acquired resistance mechanisms evolved as a targeted response to particular antibiotics, such as compound-specific efflux pumps and enzymes that modify antibiotic molecules (Surette and Wright, 2017). Besides its ancient and environmental origin, the spread and evolution of ARGs in pathogenic and non-pathogenic bacterial populations is directly and indirectly affected by human activities (Martínez, 2008; Gillings and Stokes, 2012).

Antibiotic resistant bacteria (ARBs) and ARGs enter the environment in Canada primarily *via* two routes. The first is through human-associated antibiotics use both in hospitals, as well as in the general population, causing ARBs and ARGs to accumulate in wastewater, which is collected and treated in wastewater treatment plants (WWTPs; Reinthaler et al., 2003; Hocquet et al., 2016). Importantly, WWTPs are not specifically equipped to remove ARGs and ARBs, and may in fact constitute hotspots for ARG evolution and spread, due to sub-minimum inhibitory concentrations (MIC) of antibiotics in digestors as well as conditions promoting horizontal gene transfer (Rizzo et al., 2013). While many ARBs and ARGs are reduced in concentration in WWTP effluent compared to the influent, their relative abundance and their genetic context can shift before entry into surface water systems (Reinthaler et al., 2003; Szczepanowski et al., 2009; Rizzo et al., 2013; Hocquet et al., 2016). In addition, biosolids from WWTPs have been increasingly used as agricultural fertilizers due to their high nutrient content such as phosphorous and nitrogen (Lu et al., 2012). Biosolids are enriched with ARGs and runoff from biosolid-fertilized fields can thus

introduce those genes into neighboring environments (McLain et al., 2017).

The second route of ARG introduction into the environment is *via* agricultural and specifically animal feeding operations (AFOs). Despite recent attempts to curb the use of antibiotics in agricultural settings in many countries, for example *via* bans on feeding antibiotics as growth promoters in Europe and Canada (Kraemer et al., 2019), vastly more antibiotics are used in veterinary than in a human medicine settings; 80% of antibiotics sold in the United States and 78% of those sold in Canada are for animals (Ventola, 2015; Public Health Agency of Canada, 2022). ARBs, ARGs, as well as excreted antibiotics are present in animal manure and urine, which is subsequently degraded and used as fertilizer. While manure fermentation decreases the abundance of some ARBs and ARGs, ARGs associated with veterinary antibiotics use can be readily detected in soil after fertilization (Graham et al., 2016).

Large scale human land use such as mining and agriculture can also indirectly increase the presence of ARGs in natural bacterial populations. Contamination with heavy metals, biocides and pesticides may increase ARG frequency in bacterial communities due to co- or cross-resistance of genes involved in the defense mechanism against these contaminants and antibiotics (Baker-Austin et al., 2006; Pal et al., 2015; Yang et al., 2017; Liao et al., 2021).

Although *a priori* AMR of environmental, non-pathogenic bacteria is not a concern to human health, it increases the risk of spreading ARGs to pathogens through horizontal gene transfer (Gillings and Stokes, 2012). For this reason, the surveillance of natural environments is one of the goals of the World Health Organization global action plan on AMR (World Health Organisation, 2017) as it contributes to the assessment of AMR dissemination risk (Berendonk et al., 2015). Freshwater environments in particular may act as ARGs reservoirs (Zhang et al., 2009), as surface waters receive runoff or effluent from important AMR hotspots, such as wastewater treatment plants and manure-fertilized fields (Rizzo et al., 2013; Kraemer et al., 2019).

Therefore, contamination of freshwater system with ARGs and the antibiotic resistant bacteria carrying them has received increased interest. For example, ARGs of interest have been tracked along rivers and in lakes receiving point sources of pollution (e.g., WWTPs or AFOs; Storteboom et al., 2010; Pruden et al., 2012; Thevenon et al., 2012; Czekalski et al., 2014, 2015; Chu et al., 2018). The absolute abundance of ARGs (and specifically *sul1*; Pruden et al., 2012; Czekalski et al., 2015) often decreased as distance from a potential pollution source increased. In addition to such specific point sources of pollution, lakes may receive ARGs and ARBs due to general human activity in the watershed. For example, higher concentrations of all or specific ARGs may be found in lakes with highly urbanized watershed (Yang et al., 2017) due to stormwater runoff containing ARGs (Białasek and Miłobędzka, 2020). Similarly, runoff from fields fertilized with manure or pastures could have similar effects in highly agricultural watersheds (Chee-Sanford et al., 2009). Anthropogenic activity is often

related to the trophic status of a lake. Recent work has linked nutrient status to lake resistomes (Huang et al., 2019; Pan et al., 2020; Rajasekar et al., 2022), including studies showing a correlation between increasing eutrophication and ARG abundance (Thevenon et al., 2012; Wang et al., 2020).

Most studies on natural resistomes in lakes have been limited to one (Chen et al., 2019; Stange et al., 2019; Wang et al., 2020) or few (Pal et al., 2015; Wang et al., 2020) lakes. A notable exception being a study by Czekalski et al. (2015), which assessed human impact on six antibiotic resistance genes across 21 lakes. Moreover, studied lakes are often highly eutrophic and impacted by multiple potential ARG pollution sources at the same time, making generalization difficult. Here, we conducted a comprehensive profiling of relative ARGs abundance and diversity in bacterial communities of ~350 lakes under varying pressure by human activities across Canada using metagenomic high throughput sequencing data. Lakes are the source of ~90% of drinking water in Canada (Environment Canada, 2011) and occupy 10% of the country's territory (Huot et al., 2019). We sought to understand how human impact affects the presence of ARGs in these environments. We found that, even though there was a vast diversity of natural ARGs independent of human impact, agriculture, and specifically manure applied as fertilizer in a watershed, as well as effluent from hospitals and WWTPs had a measurable input on lake resistomes across Canada.

Materials and methods

Sampling, DNA extraction and sequencing

We sampled 366 lakes across 11 Canadian ecozones during the summer of 2017–2019 at the height of summer stratification in the context of the NSERC Canadian Lake Pulse Network (referred to as LakePulse below) project as described previously (Huot et al., 2019; Kraemer et al., 2020). Briefly, an integrated epilimnion sample was collected at the deepest point of the lake as determined using a depth sounder and stored in a cooler until further processing on the same day. Subsequently, ~500 ml of water were filtered on site onto a 0.22 µm filter after pre-filtration through a 100 µm mesh. Filters were immediately frozen at –80°C and transferred to the lab. DNA was extracted using the DNeasy PowerWater kit (QIAGEN, United States; see Kraemer et al. (2020) for details). DNA was visually inspected for intactness using agarose gels and quantified using a Qubit assay (Lumiprobe, United States).

Extracted DNA was sequenced using the NovaSeq Illumina technology with an insert size of 250 base pairs and read lengths of 150 base pairs at Genome Quebec. Sequenced reads were demultiplexed and trimmed to remove low quality bases and adapters using Trimmomatic v.0.38 with default settings (Bolger et al., 2014). Datasets are available on EBI MGnify under project number PRJEB29238, accession number ERP111525.

Environmental parameters

The investigated lakes are a subset of ~660 lakes which were analyzed in the context of the LakePulse project (Huot et al., 2019). Briefly, lakes were selected for sampling across 11 Canadian ecozones with lake size and average human impact index (HII) within each ecozone as stratifying factors. Lakes further than 1 km from the closest road were excluded. To calculate a synthetic HII variable, the watershed of each lake was delineated based on geographic information system (GIS) data and land use classes assigned to each pixel within the watershed. Land use classes included natural landscapes, pasture, forestry (recent clear-cuts), agriculture, urban (buildings and roads) and mines. To calculate HII, land use classes were weighted differentially (urban, agriculture, mines: 1; pasture, forestry: 0.5; natural landscapes: 0) for each lake.

Population density (pop_wshd) in the watershed was based on WorldPop,¹ taking into account the respective year of sampling (2017, 2018, or 2019). We summed all the pixel values (population counts) falling inside each lake watershed to approximate population counts for each lake. For shared watersheds between United States and Canada, we used WorldPop population data from both territories. The number of hospitals (nbr_hospitals) and the number of hospital beds (nbr_beds) in a watershed were based on a hospital census provided by the Public Health Agency of Canada (data available upon request).

Reported WWTPs (nbr_wwtp) was based on the Government of Canada's list of Wastewater Systems Effluent Regulations as of the 20th of June 2022.² The average effluent volume per day (in m³) was normalized to the total lake volume in million cubic meters (effluent_per_volume). WWTPs and hospital locations falling in a 1 km buffer area around the watershed boundaries were visually inspected. For those points, we corrected GPS locations when necessary to make sure that they fell on the correct side of the boundary. After GPS corrections, we excluded the hospitals and WWTPs outside the watershed areas. When falling within several watershed areas, we attributed the closest lake to each WWTPs and hospitals.

Animal density (anim_den) and the application of manure-based fertilizer onto fields within a watershed were calculated based on the 2016 Agriculture and Pasture Census published by Agriculture Canada.³ Total_manure represent the theoretical fraction of the watershed area where any type of manure was applied in the year prior to the census (Oliva et al., 2022). The mean density of animals in the watershed was calculated as follow: (1) we extracted the agriculture and pasture pixels from the census areas and we redistributed the animal census counts per pixel; (2)

1 <https://www.worldpop.org/doi/10.5258/SOTON/WP00645>, Accessed March 10, 2022

2 <https://open.canada.ca/data/en/dataset/9e11e114-ef0d-4814-8d93-24af23716489>

3 <https://open.canada.ca/data/dataset/ae77c7bc-5289-47fe-ad48-5edd1d11f3d5>

we summed the pixel values (i.e., animal unit counts) falling inside each lake watershed to approximate the total number of animal units per watershed; (3) the total number of animal units was divided by the watershed area. All types of livestock were taken into account and each type of animal was normalized to animal units using a formula based on Beaulieu and Bédard (2003), to ensure that animals with different environmental impact are weighted appropriately.

The trophic state of each lake was defined based on the concentration of TP based on the thresholds for Canadian freshwater systems as follows: ultraoligotrophic (TP < 4 µg/L), oligotrophic (4–10 µg/L), mesotrophic (10–20 µg/L), mesoeutrophic (20–35 µg/L), eutrophic (35–100 µg/L) and hypereutrophic (>100 µg/L; Garner et al., 2022). We removed 17 hyper-saline lakes from the analysis following Garner et al. (2022), leaving a dataset of 349 lake.

Screening for antibiotic resistance genes in unassembled reads

To screen unassembled reads for the presence of resistance genes, all trimmed forward reads were blasted against a combined, translated version of the ARDB and CARD databases as published within the args-oap software (SARG v1.0; args-oap is a software package for the detection and quantification of resistance genes; Yang et al., 2016) using diamond blast with a minimum id of 80 and a minimum hit length of 25 amino acids (Buchfink et al., 2015). Only the best hit for each read was subsequently analyzed. To estimate the number of cells per sample, we also blasted all forward reads against a database of 30 single copy marker genes as implemented in args-oap (Yang et al., 2016). We calculated the total coverage of each marker gene by summing up the length of the blast hits and normalized this by dividing by the total length of the marker gene and then calculated the estimated cell number of each sample as the average coverage across the 30 marker genes. The normalized relative abundance of an ARG was calculated as the total hits, divided by the length of the gene and the estimated cell number in a metagenomic sample.

Screening for antibiotic resistance genes in assembled metagenomes

Trimmed reads for each lake were individually assembled using the spades assembler v. 3.13.0 with kmers 21, 33, 55, 77, 99, 127 and the assembler_only option (Bankevich et al., 2012). Subsequently, contigs were screened for the presence of ARGs using abricate v.1.0.1 against the CARD database (Seemann, 2020). The output was filtered for a minimum gene coverage of 70 and all scaffolds smaller than 1 kb excluded from further analysis. We investigated the genomic context of the remaining scaffolds using the plasClass software package with a probability cut-off of

50% (Pellow et al., 2020). PlasClass uses standard classifiers and performs specifically well on datasets with shorter contigs, such as assembled metagenomic data (Pellow et al., 2020). The taxonomy of ARG-containing scaffolds was assigned using Kraken 2 (Wood et al., 2019).

Statistical analyses

We utilized base R v.4.1.2 to test for interactions between the total normalized relative abundance of ARGs and land use and source variables (R Core Team, 2022). *p*-Values were corrected for multiple testing using the Hochberg method. The R packages vegan (Oksanen et al., 2016) and ape (Paradis and Schliep, 2019) were used for analyses of beta-diversity and ggplot2 for visualizations (Wilkinson, 2011).

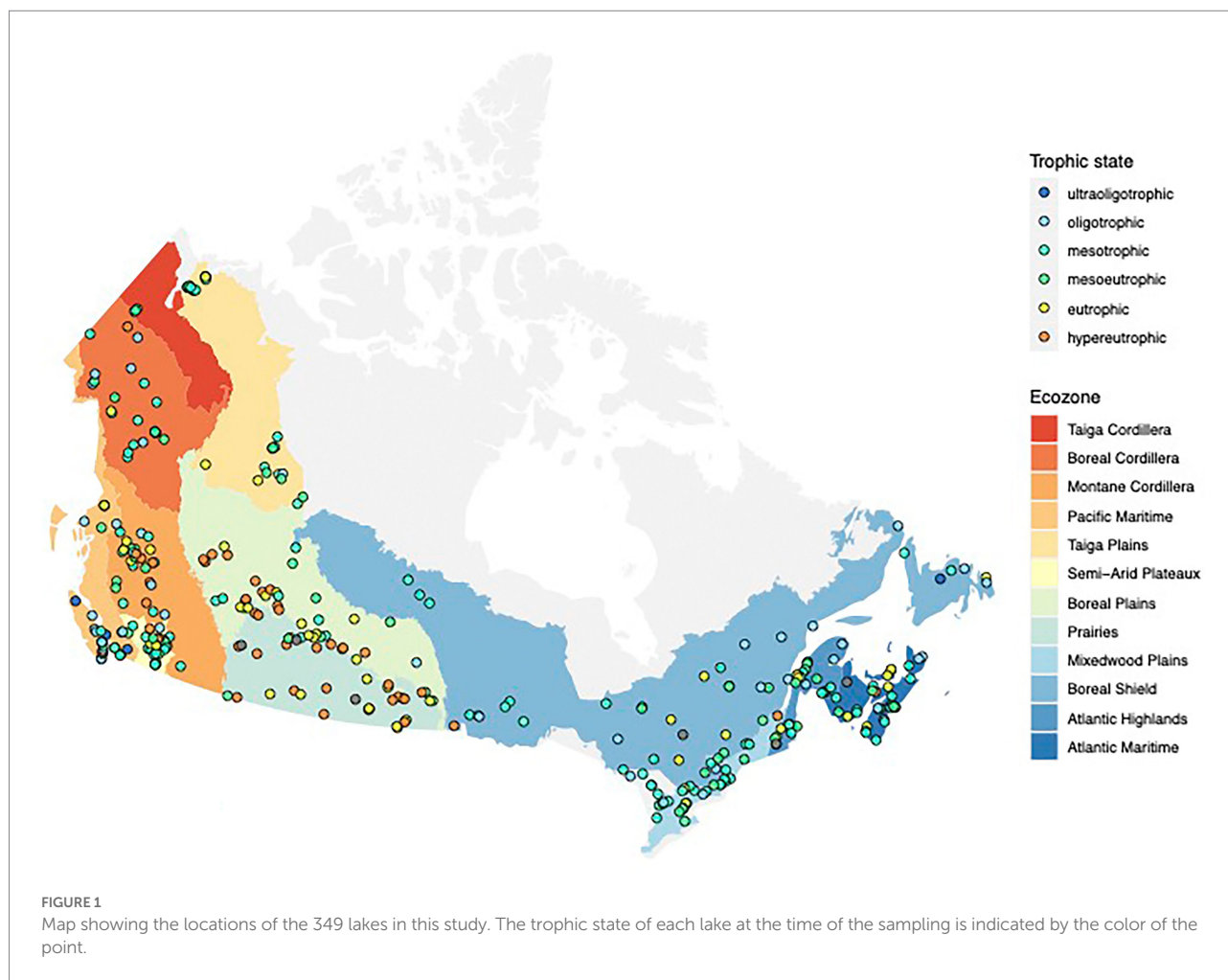
Results

Lake physicochemical parameters

In this study, we focused on 349 Canadian lakes which were sampled in the context of a large survey of lakes across Canada conducted across 3 years (2017–2019; Huot et al., 2019; Figure 1). Lakes were sampled between Canada's three coasts (43–68°N, 53–141°W) and represent a broad range of lake surface areas (0.005–100 km²) and maximum depths (0.7–150 m). They are located within 12 Canadian ecozones, which are terrestrial regions defined by characteristic environmental conditions following the Canadian Council on Ecological Areas (2014). Following their diverse geographical and ecological settings, the chosen lakes have a wide variety of human activity within their watersheds (0%–98% of the watershed impacted by human activity) and range from hyperoligotrophic to hypereutrophic. We identified 32 WWTPs and 21 hospitals in 17 and 12 watersheds, respectively. Details on environmental parameters for all lakes in this study are available in Supplementary Table 1. Details on the full survey of 660 Canadian lakes, including the 349 discussed lakes, are deposited at <https://lakepulse.ca/lakeportal/>.

Antibiotic resistance genes in unassembled read data

Unassembled metagenomic read data offer the opportunity to study the relative abundance and distribution of resistance genes in our samples. Individual lake metagenomes consisted of 17,500–91,200 K paired reads. We screened the resistomes of 349 lakes by matching them against the CARD database. In total, we detected 468 ARGs, belonging to 20 different resistance gene classes. The ARG class with the highest average relative normalized abundance conferred resistance against bacitracin (average normalized



abundance: 7.7×10^{-2} , most commonly detected genes: *bacA* [349/349], *bcrA* [349/349]), followed by multidrug (MDR) (conferring resistance against at least three different classes of antibiotics): 3×10^{-2} , most common genes detected: *mdtB*, *mdtC*, *mexF*, *mexW*, *ompR* [all genes 349/349]) and Macrolide-Lincosamide-Streptogramin (MLS) resistance genes (4.1×10^{-3} , most common gene detected: *macB* [349/349]). We also detected resistance genes against synthetic antibiotics including sulfonamides (most common gene: *sul1* [53/349]), quinolones (most common gene: *qepA* [37/349]) and trimethoprim (most common gene: *dhfrB6* [49/349]), which had lower average relative normalized abundances between 1×10^{-5} and 1×10^{-4} (Supplementary Tables 2, 3; Supplementary Figure 1 shows the prevalence of all genes across 349 lakes). Supplementary Table 4 shows the class and specific resistance conferred by the 6 most prevalent MDR genes.

Among the resistance genes that we sporadically detected in a small subset of lakes were those associated with resistance to carbapenems, which are considered antibiotics of last resort (Cooper et al., 2021; e.g., *blaOXA-48* [detected in two lakes], *blaKPC* [detected in eight lakes] and *blaIMP* beta-lactamases [detected in 12 lakes]).

Land use impact on the resistome

Across the scales investigated in this study, anthropogenic impacts within the watershed are complex and can originate from a number of factors including: agricultural activity such as fertilization with manure; urban stormwater runoff; and mining. To normalize these complex pressures on a continental scale, we utilized a synthetic variable called human impact index, which summarizes the relative proportion of different watershed land use. Subsequently, we defined a human impact index between 0 and 0.1 as low, between 0.1 and 0.5 as moderate and above 0.5 as high (Kraemer et al., 2020). Overall, the total normalized relative abundance of resistance genes in lake metagenomes was highest in lakes with the smallest anthropogenic impact in their watershed (average normalized abundance 0.134), followed by moderately (0.119) and highly (0.1) impacted lake metagenomes (Kruskal-Wallis test, $p < 0.001$). This result was heavily driven by the highly abundant resistance gene classes bacitracin and multidrug ARGs, which decreased from low to highly impacted lakes (Figure 2). However, for some individual resistance gene classes we observed an increase of the normalized relative abundance. This was the case for aminoglycosides, puromycin, vancomycin, as well as resistance against the synthetic antibiotic trimethoprim (Figure 2).

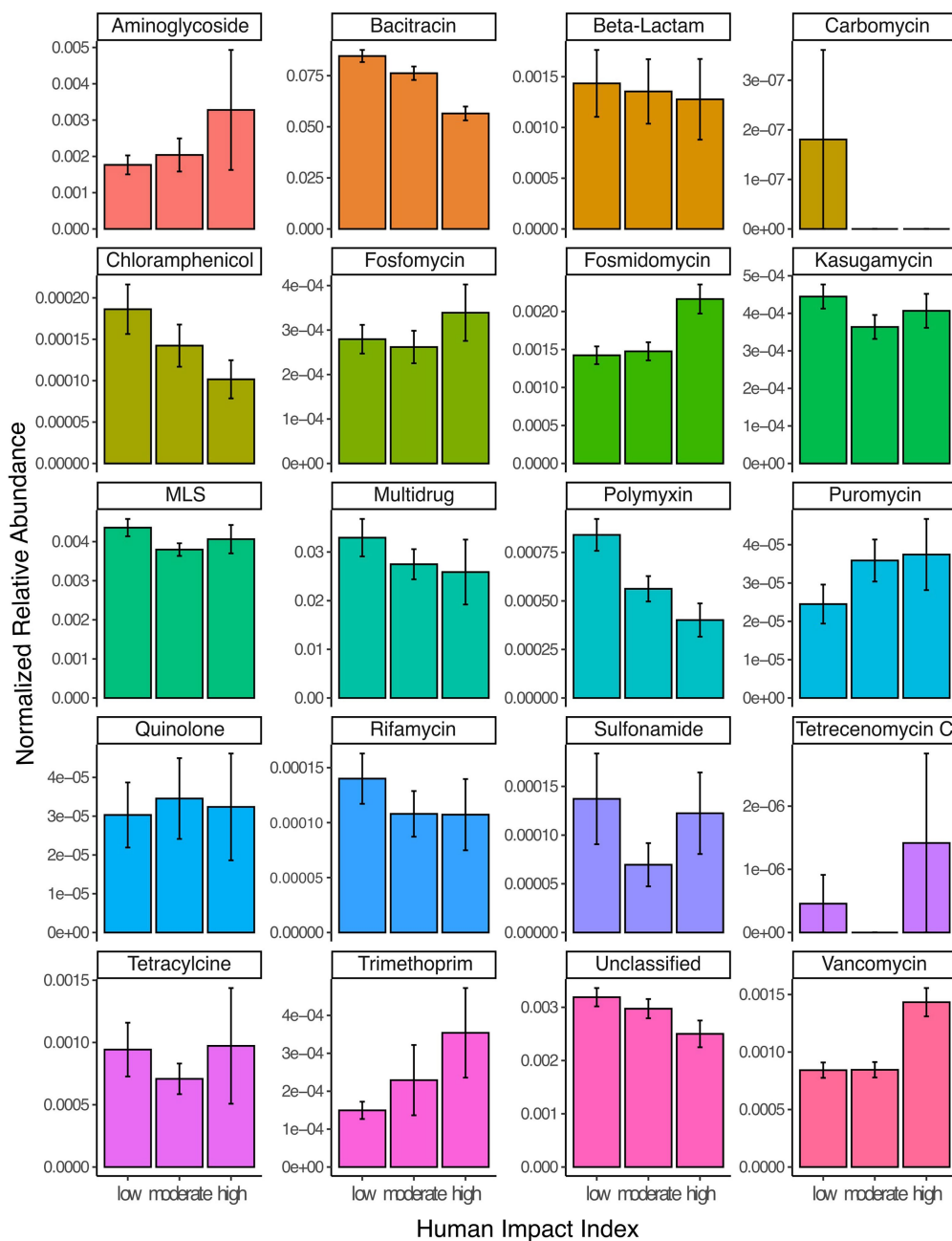
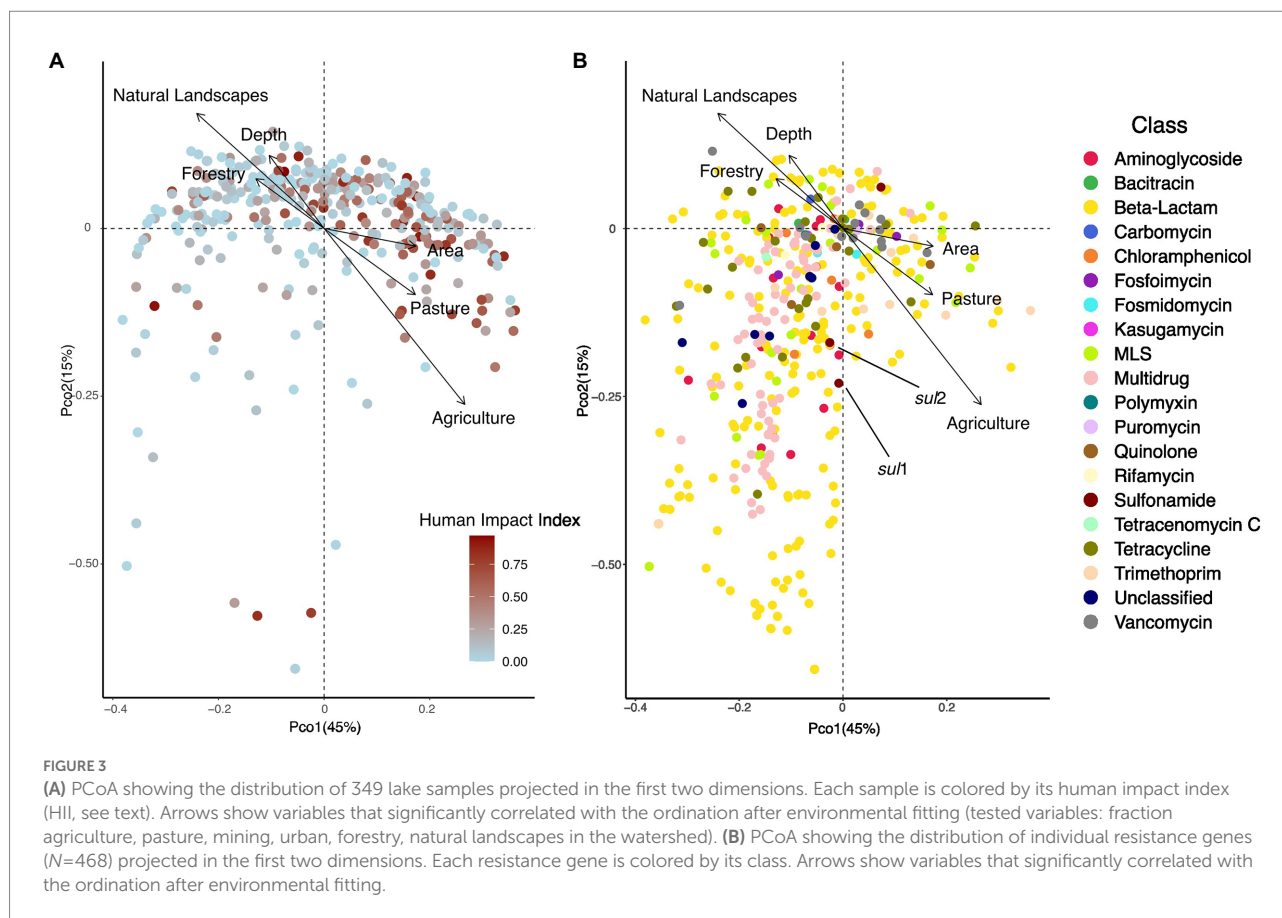


FIGURE 2

Read abundances normalized by gene length and estimated number of genomes per sample for 20 different classes of ARGs grouped by samples from lakes with low (HII: 0–0.1), moderate (HII: 0.1–0.5) or high (HII: 0.5–1) anthropogenic watershed impact. Error bars represent 95% confidence intervals. 'Unclassified' refers to genes where no specific resistance classification was available in the database.

Next, we investigated how HII impacted ARG beta-diversity. We detected a significant impact of HII, as well as lake area and ecozone on the diversity of resistance genes (Permanova, all $p < 0.05$). We explored the structure of lake resistomes further using PCoA. Environmental variables that significantly covaried with the resistome structure were the percentage of agriculture and pasture within the watershed, which contrasted with lakes whose watersheds were characterized by high percentages of forestry (recent clear-cuts) and natural landscapes (Figure 3A). In contrast,

we did not observe a significant structuring of ARG classes with land use. In fact, the vast diversity of beta-lactam and multidrug efflux pump resistance genes was largely independently distributed from any of the land use variables measured (Figure 3B). The sulfonamide resistance genes *sul1* and *sul2* have been previously investigated as markers of anthropogenic pollution (Czekalski et al., 2015; Pruden et al., 2012). Here, we found that these genes covaried with the second PCoA axis but were not correlated directly with agriculture or pasture impact on the lake resistome (Figure 3B).



Trophic state impact on the resistome

The HII variable represents a relatively static description of the potential state of a lake. A more direct way to assess lake state is to assign trophic status based on the available total phosphorous measured *in situ* at the time of sampling. We were able to assign a trophic status to 343 out of the 349 lakes. Four lakes were ultraoligotrophic, 46 oligotrophic, 122 mesotrophic, 67 mesoeutrophic, 58 eutrophic and 46 hypereutrophic at the time of sampling. In accordance with the HII impact on relative ARG abundance, we found that trophic status correlated with the total normalized abundance of resistance genes, such that oligotrophic (and even the under-sampled ultraoligotrophic) lakes had a higher normalized abundance which steadily decreased up to hypereutrophic lakes (Kruskal–Wallis test, $p < 0.001$). This pattern was driven by bacitracin ARGs (Figure 4). Other resistance gene families showed varying patterns in correlation with lake trophic status. For example, beta-lactam, quinolone and tetracycline ARGs appear to peak in mesotrophic lakes, whereas trimethoprim, vancomycin, fosmidomycin and puromycin increase with increasing lake trophic status (Figure 4).

While lake trophic status significantly correlated with lake resistome beta diversity (Permanova, $p < 0.001$), it explained only ~10% of between-lake variation (distance-based redundancy analysis, $p = 0.001$).

Potential sources of resistance genes

Resistance genes in natural environments can be either autochthonous, or allochthonous and derived from sources of pollution such as animal production or human antibiotics usage. To investigate whether we could detect the impact of distinct pollution sources on lake resistomes, we considered two variables related to veterinary antibiotics use and five variables related to human influence. Veterinary use-related variables are the density of farm animals per census agriculture area as well as the fraction of the watershed to which manure was applied. Variables pertaining to human antibiotics usage are the number of hospitals and hospital beds, the number of WWTPs, the wastewater effluent normalized to lake volume, as well as the total mean density of the population within the watershed. We did detect a significant impact of manure applied in the watershed onto ARG beta-diversity (Permanova, $p < 0.05$), but did not detect any significant correlations between total normalized abundance and the source variables tested.

Next, we investigated whether the source variables impacted resistance gene diversity by converting the normalized relative abundances to presence-absence data. Watershed population density, number of hospitals, wastewater effluent per lake volume, as well as the total watershed fraction of manure applied were all found to significantly structure resistome diversity (Permanova, all $p < 0.05$). Moreover, we detected significant positive correlations between the

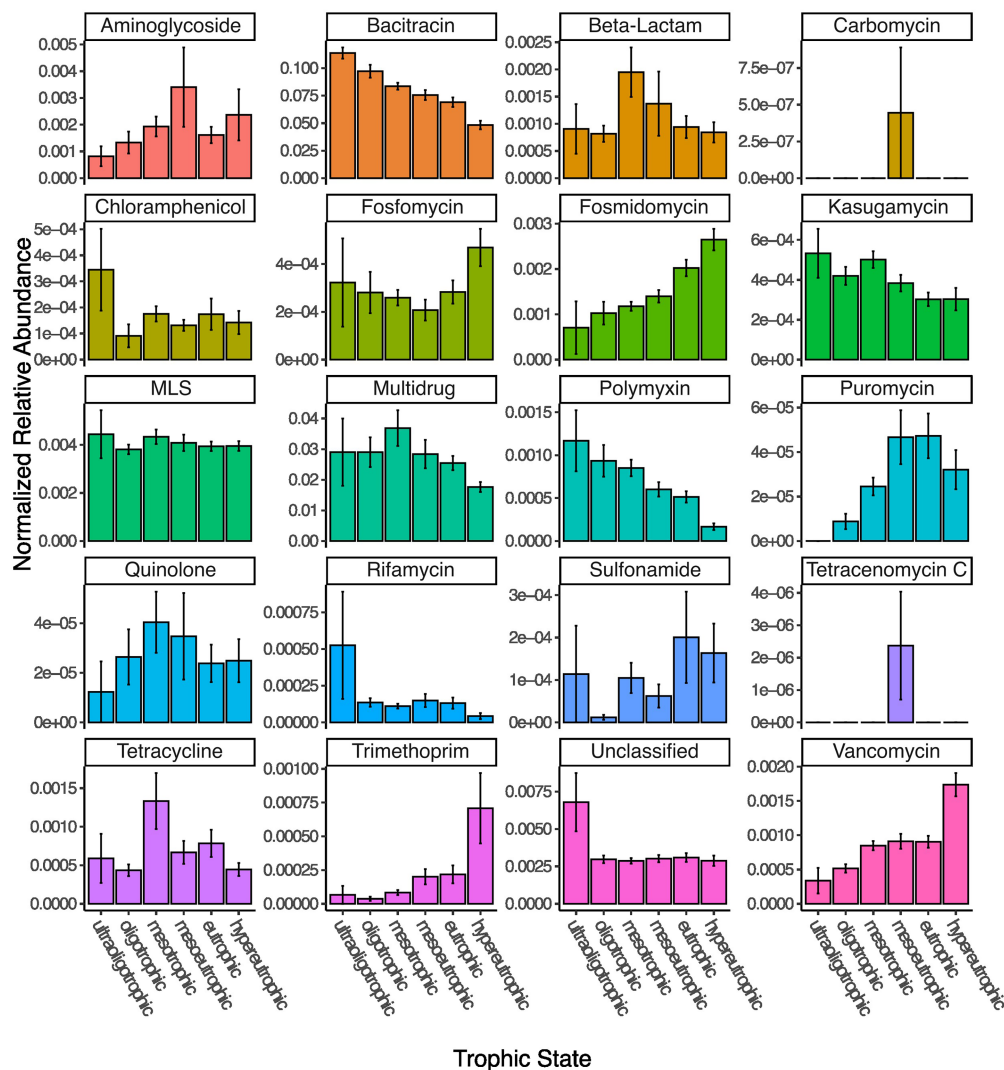


FIGURE 4

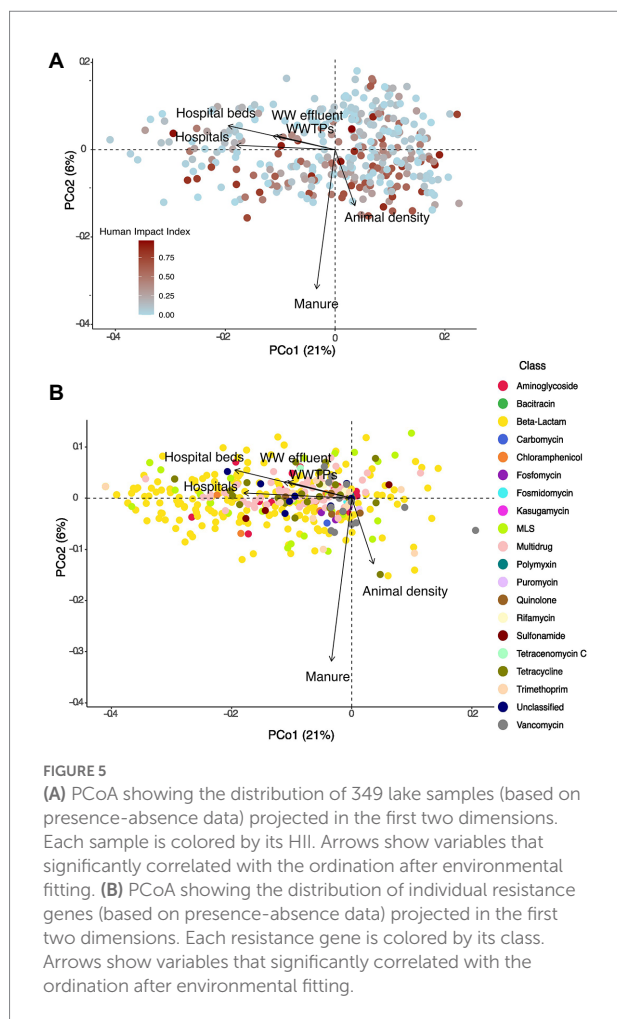
Normalized read abundances (by gene length and estimated number of genomes per sample) for 20 different classes of ARGs grouped by lake trophic status.

number of hospitals and hospital beds within a watershed and the number of different resistance genes detected in a lake metagenome (Pearson's product-moment correlations, $\text{cor}=0.21$ and 0.23 , all $p<0.001$), as well as the normalized effluent volume ($\text{cor}=0.16$, $p=0.015$), while the number of WWTPs was marginally correlated ($\text{cor}=0.13$, $p=0.08$). In total, sources of pollution explained only 5% of variation in ARG diversity (Permanova: unexplained variation 95%). Fitting of environmental variables after PCoA revealed a significant fit to the resistome diversity for the number of hospitals ($p=0.003$), hospital beds ($p=0.001$), animal density ($p=0.04$), and manure used in the watershed ($p=0.001$) as well as a marginal fit for the volume of wastewater effluent ($p=0.06$) and the number of WWTPs ($p=0.098$; Figures 5A,B). Human antibiotics use-associated variables were oriented perpendicularly to the fit for the density of animals in the watershed, indicating how these different sources of pollution influence the resistome in distinct ways. Moreover, fitted

arrows for the number of hospitals/hospital beds and the volume of wastewater are associated with the first PCoA axis and are associated with many of the beta-lactam and multidrug efflux pump genes (Figure 5), which were previously not linked with any of the land use variables tested (Figure 3).

Detection of antibiotic resistance genes in assembled data

Overall, we detected resistance gene-containing scaffolds in 246 out of the 349 lakes for a total of 1820 scaffolds. Of these, 661 scaffolds were longer than 1 kb, and we assigned them to either plasmid or chromosomal origin. Most scaffolds containing resistance genes were assigned as chromosomal (598), whereas only 63 were characterized as plasmids. 98% (587) of chromosomal



scaffolds and 84% (53) of plasmid-derived scaffolds could be taxonomically assigned to the genus-level using kraken2. The same taxa *Mycobacterium* (35% of chromosomal scaffolds/17% of plasmid scaffolds), *Acinetobacter* (26%/24%) and *Mycobacterium* (26%/17%) contributed the most ARG-containing scaffolds on chromosomes or plasmids. Both *Acinetobacter* and *Mycobacterium* are widespread in the environment and sewage and have received attention as potential opportunistic pathogens (Doughari et al., 2011; Pfyffer, 2015). We subsequently tested whether the plasmid load (the count of ARG-containing scaffolds characterized as plasmid-originating divided by the total ARG-containing scaffolds in the sample) was significantly correlated with trophic state, watershed land use or resistance gene source variables, but found no indication (anova and Pearson's product moment correlation tests, all $p > 0.05$).

Impact of land use and trophic status on the number of genomes

We tested whether trophic status or human impact index significantly correlated with the number of genomes detected in

the samples. We did not detect a significant correlation between the average genome number estimated in a sample and the human impact index (Pearson's product moment correlation test, $p = 0.6$), nor did we detect a significant impact of trophic status or human impact class (high, moderate, low) on the number of genomes detected (anova, all $p > 0.05$).

Discussion

Lakes integrate watershed-scale processes and are thus considered sentinels of processes such as climate change (Adrian et al., 2009), contaminant use (Schindler, 2009) and pathogen presence (Oliva et al., 2022). Lakes may thus constitute environmental receptors and reservoirs for ARGs, while at the same time providing critical ecosystem services, including 90% of Canadian drinking water (Environment Canada, 2000). Here, we take advantage of a continental-scale dataset of ~350 lakes to investigate whether we can detect signatures of human or veterinary use-induced antibiotic resistance across large spatial scales.

In contrast to previous work which used qPCR to detect absolute abundances of a small number of resistance genes (e.g., Pruden et al., 2012; Thevenon et al., 2012; Czekalski et al., 2015), we utilize metagenomic data, which is wider in scope, but produces relative abundances of ARGs. Thus, shifts in the total abundance of cells in a sample, for example during a bloom event under eutrophic conditions, could lead to a decrease in relative abundance of an ARG (if the blooming organisms are outcompeting the organisms containing the ARG), but could nonetheless represent a total increase of the ARG if measured *via* qPCR. We did not detect any indication that metagenomic data sets from lakes with higher human impact index or trophic status were characterized by a higher number of genomes, indicating that the relative abundances we report here may correlate well with absolute abundances of resistance genes.

Overall, the total relative abundance of ARGs was inversely correlated with total human watershed impact. This relationship was largely driven by the most abundant classes of genes conferring resistance against bacitracin and multidrug resistance genes. High relative abundances of genes conferring resistance to bacitracin have been previously identified in pristine lake sediments from a Tibetan lake (Chen et al., 2016), and in less-disturbed river samples (Lee et al., 2021). In contrast, Pan et al. (2020) detected an increase in the relative abundance of bacitracin ARGs with increasing lake nutrient status, but focused only on highly polluted lakes. Overall, the role of bacitracin resistance in aquatic bacterial ecology in the absence of direct anthropogenic pressures is unknown, but this resistance class may be unsuitable for environmental source tracking of ARG pollution (Lee et al., 2021).

The second most common class of resistance genes, multidrug resistance (MDR), is often associated with efflux pumps. Antibiotic resistance-related efflux pumps are prevalent both in

antibiotic-resistant and sensitive bacteria, share homologies with other efflux pumps and resistance to an antibiotic may often be conferred *via* point mutations. Thus, multidrug efflux pumps may not be reliable markers of functional resistance (Nesme et al., 2014). We repeated the analyses after exclusion of all genes categorized with the mechanisms 'efflux pumps' or 'membrane permeability change', but found similar results, indicating that the misidentification of ARGs is not driving the patterns observed (data available upon request).

The total relative abundance of resistance genes to other classes of antibiotics including aminoglycosides, fosmidomycin, puromycin, trimethoprim and vancomycin increased the higher human impact in the watershed. Trimethoprim, a synthetic antibiotic mostly used in human medicine, has been regularly detected in wastewater effluent and is present in lakes (Yang et al., 2018), where it is persistent and resistant to break-down (Bai and Acharya, 2017). However, the concentration in lakes is usually a magnitude lower (10–100 ng/L) than the MIC (100 µg/L), making *in situ* evolution of resistance unlikely (Yang et al., 2018). Therefore, ARGs likely enter the system *via* contamination with human wastewater. Genes conferring resistance to aminoglycosides have been previously found to be prevalent in interconnected lake and river systems (Zhang et al., 2021) and their prevalence in sediments has been previously linked to WWTP effluent (Devarajan et al., 2015). Vancomycin resistance is often associated with the use of veterinary antibiotics in animal production, from where it has spread to resistant pathogens such as vancomycin-resistant enterococci, which have been previously isolated from surface water (Messi et al., 2006). The genes detected in association with resistance to fosmidomycin and puromycin resistance all represent efflux pumps, and thus the same caveats may apply to them as to multidrug efflux pump resistance genes. However, they have been previously detected in stormwater (Białasek and Miłobędzka, 2020), and may thus be associated with urban impact on aquatic ecosystems. In summary, our findings support the notion that source tracking of anthropogenic pollution should be based on a multiple antibiotic resistance framework, rather than individual ARGs (Krumperman, 1983; Kelsey et al., 2003).

The trophic state of a lake is often taken as an indicator of human activity. We observed highly similar trends of resistance gene abundances whether we used human impact index or lake trophic state as an explanatory variable, indicating that GIS-derived land use data offers a valid proxy for *in situ* measurements to assess lake trophic states during summer stratification. Overall, resistance gene abundances were significantly different between trophic states, and significantly decreased from ultraoligotrophic to hypereutrophic lakes, due to the decrease of both bacitracin and multidrug resistance gene abundances with increasing trophic status. The overall relationship between lake nutrient status and the abundance of ARGs is unclear. Previous studies have reported positive correlations (Thevenon et al., 2012; Czekalski et al., 2015; Wang et al., 2020; Rajasekar et al., 2022) between select ARG absolute abundances and TP, but negative correlations have been reported as well (Huang et al., 2019). Our results indicate that with increasing human activity and nutrient status the resistome shifts, with major ARG classes declining in relative abundance.

In correspondence with those shifts, we found the overall composition of resistance genes to be significantly impacted by agricultural activity and pasture land use within the watershed, contrasting with the effects of natural landscapes and forestry. However, we did not detect an association of these significant land use variables with a specific class of resistance genes. Instead, the vast diversity of beta-lactam and multidrug resistance genes detected in the lake dataset was unassociated with any explanatory variable tested. Individual genes conferring resistance to trimethoprim, beta-lactams and MLS were associated with agricultural and pasture land use. This indicates that, while a large fraction of the lake resistome appears non-impacted by watershed land use, shifts from pristine or forestry-dominated (e.g., recent clear-cuts) watersheds to those associated with animal production do significantly alter resistome composition, likely in association with the veterinary use of antibiotics.

We specifically linked census-derived source variables regarding the origin of ARBs and ARGs to the composition of the resistome. Total population density, number of hospitals and hospital beds, the fraction of watershed with applied manure and the volume of wastewater effluent all significantly impacted ARG diversity, and the number of hospitals, hospital beds, and the volume of WWTP discharge were positively correlated with the total number of different ARGs detected. While this is an intriguing finding, we caution against over-interpretation, since only a small minority of lake watersheds contained any hospitals or WWTPs, and they often co-occur in the same watershed. Nonetheless, these correlations are in line with previous studies that found that WWTP effluent, and specifically effluent from hospitals, was enriched for resistance genes associated with clinical antibiotics usage (Reinthal et al., 2003; Hocquet et al., 2016). Moreover, WWTP effluent has been identified as a point source for the pollution with ARGs and ARBs in previous work, including lakes in Switzerland (Czekalski et al., 2015) and lake Michigan (Chu et al., 2018), as well as river systems (Pei et al., 2006). The overall influence of WWTP discharge does not appear to be significant enough to cause shifts in overall ARG abundances in lakes but acts as a source for novel resistance genes in the system. Even though the overall impact of ARGs involved in human-associated antibiotics use is much smaller than that associated with animal production in the watershed, human-associated ARGs are of special concern for human health, due to a higher risk of reinfection, as well as the evolution of novel pathogens due to the dissemination of ARGs into new genetic backgrounds (Kraemer et al., 2019).

Previously, a highly impacted lake ecosystem has been shown to be enriched for mobile genetic elements, leading us to hypothesize that increasing human impact may select for an increased abundance of ARGs located on plasmids (Bengtsson-Palme et al., 2014). However, we did not detect any indication that land use, trophic status or resistance gene pollution sources altered the propensity of resistance genes to be located on chromosomes versus plasmids. This may be because the ecosystems studied are much less impacted than

the polluted lake in the original study, both by direct pollution with antibiotics from antibiotics-producing industry, as well as with ARBs that may enter the system *via* wastewater. It is thus likely that the range of environmental conditions of the lakes described here is not extreme enough to exert the selective pressures necessary to shift toward a more mobile resistome (Bengtsson-Palme et al., 2017).

In summary, we present a survey of the resistome of 349 lakes ranging from pristine to highly impacted. To our knowledge, this is the largest systematic study of aquatic resistomes to date, which we aim to link to watershed land use, as well as sources of ARG point pollution. We detected a vast ‘natural’ aquatic resistome, which likely corresponds to ARGs used in inter-bacterial communication and competition. However, we also detected that the resistome was structured by watershed-scale land use, with pasture and agriculture-dominated lakes contrasting with those in pristine or forestry-dominated watersheds. Correspondingly, we did detect an impact of pollution sources for veterinary-origin ARGs on the structure of lake resistomes. Moreover, human antibiotics use within the watershed was found to have a positive correlation with ARG diversity. Taken together, these findings support the monitoring of ARG loading by both agricultural and clinical pollution sources in sink environments such as lakes.

Data availability statement

The data presented in the study are deposited in the EBI MGnify depository, accession number PRJEB29238.

Author contributions

SK, YH, and DW conceived the study. AO provided geomatic analyses on the sources of resistance genes. SK

conducted the analyses. SK and NB wrote the manuscript. SK, YH, DW, AO, and NB edited and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by NSERC through the Canadian Lake Pulse Network Strategic Partnership Network program (NETGP 479720).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.995418/full#supplementary-material>

References

- Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., et al. (2009). Lakes as sentinels of climate change. *Limnol. Oceanogr.* 54, 2283–2297. doi: 10.4319/lo.2009.54.6_part_2.2283
- Bai, X., and Acharya, K. (2017). Algae-mediated removal of selected pharmaceutical and personal care products (PPCPs) from Lake Mead water. *Sci. Total Environ.* 581–582, 734–740. doi: 10.1016/j.scitotenv.2016.12.192
- Baker-Austin, C., Wright, M. S., Stepanauskas, R., and McArthur, J. V. (2006). Co-selection of antibiotic and metal resistance. *Trends Microbiol.* 14, 176–182. doi: 10.1016/j.tim.2006.02.006
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Beaulieu, M. S., and Bédard, F. (2003). A geographic profile of Canadian livestock, 1991–2001. Agricultural and rural working Paper Series. Working paper no. 62. statistics Canada, agriculture division.
- Bengtsson-Palme, J., Boulund, F., Fick, J., Kristiansson, E., and Larsson, D. G. J. (2014). Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. *Front. Microbiol.* 5:648. doi: 10.3389/fmicb.2014.00648
- Bengtsson-Palme, J., Kristiansson, E., and Larsson, D. G. J. (2017). Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol. Rev.* 42. doi: 10.1093/femsre/fux053
- Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., et al. (2015). Tackling antibiotic resistance: the environmental framework. *Nat. Rev. Microbiol.* 13, 310–317. doi: 10.1038/nrmicro3439
- Bhullar, K., Wagelchner, N., Pawlowski, A., Koteva, K., Banks, E. D., Johnston, M. D., et al. (2012). Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One* 7:e34953. doi: 10.1371/journal.pone.0034953
- Białasek, M., and Miłobędzka, A. (2020). Revealing antimicrobial resistance in stormwater with min ION. *Chemosphere* 258:127392. doi: 10.1016/j.chemosphere.2020.127392
- Bolger, A., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible read trimming tool for Illumina NGS data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Buchfink, B., Xie, C., and Huson, D. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Canadian Council on Ecological Areas (2014). Available at: <https://ccea-ccae.org/ecozones-download>
- Chee-Sanford, J. C., Mackie, R. I., Koike, S., Krapac, I. G., Lin, Y.-F., Yannarell, A. C., et al. (2009). Fate and transport of antibiotic residues and antibiotic resistance genes following land application of manure waste. *J. Environ. Qual.* 38, 1086–1108. doi: 10.2134/jeq2008.0128
- Chen, H., Jing, L., Yao, Z., Meng, F., and Teng, Y. (2019). Prevalence, source and risk of antibiotic resistance genes in the sediments of Lake tai (China) deciphered

- by metagenomic assembly: a comparison with other global lakes. *Environ. Int.* 127, 267–275. doi: 10.1016/j.envint.2019.03.048
- Chen, B., Yuan, K., Chen, X., Yang, Y., Zhang, T., Wang, Y., et al. (2016). Metagenomic analysis revealing antibiotic resistance genes (ARGs) and their genetic compartments in the Tibetan environment. *Environ. Sci. Technol.* 50, 6670–6679. doi: 10.1021/acs.est.6b00619
- Chu, B. T. T., Petrovich, M. L., Chaudhary, A., Wright, D., Murphy, B., Wells, G., et al. (2018). Metagenomics reveals the impact of wastewater treatment plants on the dispersal of microorganisms and genes in aquatic sediments. *Appl. Environ. Microbiol.* 84. doi: 10.1128/AEM.02168-17
- Cooper, A. L., Carter, C., McLeod, H., Wright, M., Sritharan, P., Tamber, S., et al. (2021). Detection of carbapenem-resistance genes in bacteria isolated from wastewater in Ontario. *FACETS*. 6, 569–591. doi: 10.1139/facets-2020-0101
- Council of Canadian Academics (2019). When Antibiotics Fail: The Expert Panel on the Potential Socio-Economic Impacts of Antimicrobial Resistance in Canada. Ottawa, ON: Council of Canadian Academies.
- Czekalski, N., Gascón Díez, E., and Bürgmann, H. (2014). Wastewater as a point source of antibiotic-resistance genes in the sediment of a freshwater lake. *ISME J.* 8, 1381–1390. doi: 10.1038/ismej.2014.8
- Czekalski, N., Sigdel, R., Birtel, J., Matthews, B., and Bürgmann, H. (2015). Does human activity impact the natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss lakes. *Environ. Int.* 81, 45–55. doi: 10.1016/j.envint.2015.04.005
- D'Costa, V. M., Osta, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., et al. (2011). Antibiotic resistance is ancient. *Nature* 477, 457–461. doi: 10.1038/nature10388
- Devarajan, N., Laffite, A., Graham, N. D., Meijer, M., Prabakar, K., Mubedi, J. I., et al. (2015). Accumulation of clinically relevant antibiotic-resistance genes, bacterial load, and metals in freshwater lake sediments in central Europe. *Environ. Sci. Technol.* 49, 6528–6537. doi: 10.1021/acs.est.5b01031
- Doughari, H. J., Ndkidemi, P. A., Human, I. S., and Benade, S. (2011). The ecology, biology and pathogenesis of *Acinetobacter* spp.: an overview. *Microbes Environ.* 26, 101–112. doi: 10.1264/jsm.2.ME10179
- Environment Canada. (2000). 2011 municipal water use report: Municipal water use 2009 statistics.
- Environment Canada (2011). 2011 Municipal Water Use Report.
- Garner, R. E., Kraemer, S. A., Onana, V., Huot, Y., Gregory-Eaves, I., and Walsh, D. A. (2022). Protist diversity and metabolic strategy in freshwater lakes are shaped by trophic state and watershed land use on a continental scale. *mSystems* 7:e0031622. doi: 10.1128/mSystems.00316-22
- Gillings, M. R., and Stokes, H. (2012). Are humans increasing bacterial evolvability? *Trends Ecol. Evol.* 27, 346–352. doi: 10.1016/j.tree.2012.02.006
- Goh, E.-B., Yim, G., Tsui, W., McClure, J., Surette, M. G., and Davies, J. (2002). Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc. Natl. Acad. Sci.* 99, 17025–17030. doi: 10.1073/pnas.252607699
- Graham, D. W., Knapp, C. W., Christensen, B. T., McCluskey, S., and Dolfing, J. (2016). Appearance of β -lactam resistance genes in agricultural soils and clinical isolates over the 20th century. *Sci. Rep.* 6. doi: 10.1038/srep21550
- Hocquet, D., Muller, A., and Bertrand, X. (2016). What happens in hospitals does not stay in hospitals: antibiotic-resistant bacteria in hospital wastewater systems. *J. Hosp. Infect.* 93, 395–402. doi: 10.1016/j.jhin.2016.01.010
- Huang, Z., Zhao, W., Xu, T., Zheng, B., and Yin, D. (2019). Occurrence and distribution of antibiotic resistance genes in the water and sediments of Qingcaosha Reservoir, Shanghai, China. *Environ. Sci. Eur.* 31:81. doi: 10.1186/s12302-019-0265-2
- Huot, Y., Brown, C. A., Potvin, G., Antoniadou, D., Baulch, H. M., Beisner, B. E., et al. (2019). The NSERC Canadian Lake pulse network: a national assessment of lake health providing science for water management in a changing climate. *Sci. Total Environ.* 695:133668. doi: 10.1016/j.scitotenv.2019.133668
- Kelsey, R. H., Scott, G. I., Porter, D. E., Thompson, B., and Webster, L. (2003). "Using multiple antibiotic resistance and land use characteristics to determine sources of fecal coliform bacterial pollution," in *Coastal Monitoring through Partnerships* eds. B. D. Melzian, V. Engle, M. McAlister, S. Sandhu, and L. K. Eads (Dordrecht: Springer).
- Kraemer, S. A., Barbosa da Costa, N., Shapiro, B. J., Fradette, M., Huot, Y., and Walsh, D. A. (2020). A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial communities. *ISME J.* 14, 3011–3023. doi: 10.1038/s41396-020-0733-0
- Kraemer, S. A., Ramachandran, A., and Perron, G. G. (2019). Antibiotic pollution in the environment: from microbial ecology to public policy. *Microorganisms* 7. doi: 10.3390/microorganisms7060180
- Krumpalman, P. H. (1983). Multiple antibiotic resistance indexing of *Escherichia coli* to identify high-risk sources of fecal contamination of food. *Appl. Environ. Microbiol.* 46, 165–170. doi: 10.1128/aem.46.1.165-170.1983
- Lee, J., Ju, F., Maile-Moskowitz, A., Beck, K., Maccagnan, A., McArdell, C. S., et al. (2021). Unraveling the riverine antibiotic resistome: the downstream fate of anthropogenic inputs. *Water Res.* 197:117050. doi: 10.1016/j.watres.2021.117050
- Lewis, K. (2020). The science of antibiotic discovery. *Cells* 181, 29–45. doi: 10.1016/j.cell.2020.02.056
- Liao, H., Li, X., Yang, Q., Bai, Y., Cui, P., Wen, C., et al. (2021). Herbicide selection promotes antibiotic resistance in soil microbiomes. *Mol. Biol. Evol.* 38, 2337–2350. doi: 10.1093/molbev/msab029
- Lu, Q., He, Z. L., and Stoffella, P. J. (2012). Land application of biosolids in the USA: a review. *Appl. Environ. Soil Sci.* 2012, 1–11. doi: 10.1155/2012/201462
- Martínez, J. L. (2008). Antibiotics and antibiotic resistance genes in natural environments. *Science (80-)* 321, 365–367. doi: 10.1126/science.1159483
- McLain, J., Rock, C., and Gerba, C. (2017). "Environmental antibiotic resistance associated with land application of biosolids," in *Antimicrobial Resistance in Wastewater Treatment Processes* eds. P. L. Keen and R. Fugère (Hoboken, NJ, USA: John Wiley & Sons, Inc.), 241–252.
- Messi, P., Guerrieri, E., de Niederhäusern, S., Sabia, C., and Bondi, M. (2006). Undefined vancomycin-resistant enterococci (VRE) in meat and environmental samples. *Int. J. Food Microbiol.* 107, 218–222. doi: 10.1016/j.ijfoodmicro.2005.08.026
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 399, 629–655. doi: 10.1016/S0140-6736(21)02724-0
- Nesme, J., Cécillon, S., Delmont, T. O., Monier, J. M., Vogel, T. M., and Simonet, P. (2014). Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr. Biol.* 24, 1096–1100. doi: 10.1016/j.cub.2014.03.036
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al. (2016). Vegan: Community Ecology Package. Available at: <https://cran.r-project.org/>; <https://github.com/vegandevs/vegan>
- Oliva, A., Garner, R., Walsh, D. A., and Huot, Y. (2022). The occurrence of potentially pathogenic fungi and protists in Canadian lakes predicted using geomatics, in situ and satellite-derived variables: towards a tele-epidemiological approach. *Water Res.* 209:117935. doi: 10.1016/j.watres.2021.117935
- Pal, C., Bengtsson-Palme, J., Kristiansson, E., and Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* 16:964. doi: 10.1186/s12864-015-2153-5
- Pan, X., Lin, L., Zhang, W., Dong, L., and Yang, Y. (2020). Metagenome sequencing to unveil the resistome in a deep subtropical lake on the Yunnan-Guizhou plateau. *China. Environ. Pollut.* 263:114470. doi: 10.1016/j.envpol.2020.114470
- Paradis, E., and Schliep, K. (2019). Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633
- Pei, R., Kim, S. C., Carlson, K. H., and Pruden, A. (2006). Effect of river landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Res.* 40, 2427–2435. doi: 10.1016/j.watres.2006.04.017
- Pellow, D., Mizrahi, I., and Shamir, R. (2020). PlasClass improves plasmid sequence classification. *PLoS Comput. Biol.* 16:e1007781. doi: 10.1371/journal.pcbi.1007781
- Perron, G. G., Whyte, L., Turnbaugh, P. J., Goordial, J., Hanage, W. P., Dantas, G., et al. (2015). Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS One* 10:e0069533. doi: 10.1371/journal.pone.0069533
- Pfiffer, G. E. (2015). "Mycobacterium: General characteristics, laboratory detection and staining procedures," in *Manual of Clinical Microbiology*. eds. J. Versalovic, K. C. Carroll, G. Funke, J. H. Jorgensen, M. L. Landry, and D. W. Tenover (Washington, DC, USA: ASM Press).
- Pruden, A., Arabi, M., and Storteboom, H. N. (2012). Correlation between upstream human activities and riverine antibiotic resistance genes. *Environ. Sci. Technol.* 46, 11541–11549. doi: 10.1021/es302657r
- Public Health Agency of Canada. (2022). Canadian Antimicrobial Resistance Surveillance System Report.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.R-project.org/>
- Rajasekar, A., Qiu, M., Wang, B., Murava, R. T., and Norgbey, E. (2022). Relationship between water quality, heavy metals and antibiotic resistance genes among three freshwater lakes. *Environ. Monit. Assess.* 194, 64–10. doi: 10.1007/s10661-021-09704-9
- Reinthal, F. F., Posch, J., Feierl, G., Wüst, G., Haas, D., Ruckebauer, G., et al. (2003). Antibiotic resistance of *E. coli* in sewage and sludge. *Water Res.* 37, 1685–1690. doi: 10.1016/S0043-1354(02)00569-9
- Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., et al. (2013). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Sci. Total Environ.* 447, 345–360. doi: 10.1016/j.scitotenv.2013.01.032

- Schindler, D. W. (2009). Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnol. Oceanogr.* 54, 2349–2358. doi: 10.4319/lo.2009.54.6_part_2.2349
- Seemann, T. (2020). Abricate. Available at: <https://github.com/tseemann/abricate>.
- Skindersoe, M. E., Alhede, M., Phipps, R., Yang, L., Jensen, P. O., Rasmussen, T. B., et al. (2008). Effects of antibiotics on quorum sensing in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 52, 3648–3663. doi: 10.1128/AAC.01230-07
- Stange, C., Yin, D., Xu, T., Guo, X., Schäfer, C., and Tiehm, A. (2019). Distribution of clinically relevant antibiotic resistance genes in Lake tai. *China. Sci. Total Environ.* 655, 337–346. doi: 10.1016/j.scitotenv.2018.11.211
- Storteboom, H., Arabi, M., Davis, J. G., Crimi, B., and Pruden, A. (2010). Identification of antibiotic-resistance-gene molecular signatures suitable as tracers of pristine river, urban, and agricultural sources. *Environ. Sci. Technol.* 44, 1947–1953. doi: 10.1021/es902893f
- Surette, M. D., and Wright, G. D. (2017). Lessons from the environmental antibiotic Resistome. *Annu. Rev. Microbiol.* 71, 309–329. doi: 10.1146/annurev-micro-090816-093420
- Szczepanowski, R., Linke, B., Krahn, I., Gartemann, K. H., Gützkow, T., Eichler, W., et al. (2009). Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiology* 155, 2306–2319. doi: 10.1099/mic.0.028233-0
- Thevenon, F., Adatte, T., Wildi, W., and Poté, J. (2012). Antibiotic resistant bacteria/genes dissemination in lacustrine sediments highly increased following cultural eutrophication of Lake Geneva (Switzerland). *Chemosphere* 86, 468–476. doi: 10.1016/j.chemosphere.2011.09.048
- Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *Pharm. Ther. J.* 40, 277–283.
- Wang, Z., Han, M., Li, E., Liu, X., Wei, H., Yang, C., et al. (2020). Distribution of antibiotic resistance genes in an agriculturally disturbed lake in China: their links with microbial communities, antibiotics, and water quality. *J. Hazard. Mater.* 393:122426. doi: 10.1016/j.jhazmat.2020.122426
- Wilkinson, L. (2011). Ggplot 2: elegant graphics for data analysis by WICKHAM, H. *Biometrics* 67, 678–679. doi: 10.1111/j.1541-0420.2011.01616.x
- Wood, D., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome Biol.* 20:257. doi: 10.1186/s13059-019-1891-0
- World Health Organisation (2017). Global Action Plan on Antimicrobial Resistance. WHO, Geneva, Switzerland: WHO Press.
- Yang, Y., Jiang, X., Chai, B., Ma, L., Li, B., Zhang, A., et al. (2016). ARGs-OAP: online analysis pipeline for antibiotic resistance gene detection from metagenomic data using an integrated structured ARG-database. *Bioinformatics* 32, 2346–2351. doi: 10.1093/bioinformatics/btw136
- Yang, Y., Song, W., Lin, H., Wang, W., Du, L., and Xing, W. (2018). Antibiotics and antibiotic resistance genes in global lakes: a review and meta-analysis. *Environ. Int.* 116, 60–73. doi: 10.1016/j.envint.2018.04.011
- Yang, Y., Xu, C., Cao, X., Lin, H., and Wang, J. (2017). Antibiotic resistance genes in surface water of eutrophic urban lakes are related to heavy metals, antibiotics, lake morphology and anthropic impact. *Ecotoxicology* 26, 831–840. doi: 10.1007/s10646-017-1814-3
- Zhang, M. S., Li, W., Zhang, W. G., Li, Y. T., Li, J. Y., and Gao, Y. (2021). Agricultural land-use change exacerbates the dissemination of antibiotic resistance genes via surface runoffs in Lake Tai Basin, China. *Ecotoxicol. Environ. Saf.* 220:112328. doi: 10.1016/j.ecoenv.2021.112328
- Zhang, X. X., Zhang, T., and Fang, H. H. P. (2009). Antibiotic resistance genes in water environment. *Appl. Microbiol. Biotechnol.* 82, 397–414. doi: 10.1007/s00253-008-1829-z



OPEN ACCESS

EDITED BY

Alberto Antonelli,
University of Florence,
Italy

REVIEWED BY

Sezer Okay,
Hacettepe University,
Turkey
Ilaria Baccani,
University of Florence,
Italy

*CORRESPONDENCE

Alastair I. H. Murchie
aih@fudan.edu.cn
Dongrong Chen
drchen@fudan.edu.cn

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to
Antimicrobials, Resistance and
Chemotherapy,
a section of the journal
Frontiers in Microbiology

RECEIVED 06 May 2022

ACCEPTED 10 November 2022

PUBLISHED 29 November 2022

CITATION

Chang Y, Zhang X, Murchie AH and
Chen D (2022) Transcriptome profiling in
response to Kanamycin B reveals its wider
non-antibiotic cellular function in
Escherichia coli.
Front. Microbiol. 13:937827.
doi: 10.3389/fmicb.2022.937827

COPYRIGHT

© 2022 Chang, Zhang, Murchie and Chen.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Transcriptome profiling in response to Kanamycin B reveals its wider non-antibiotic cellular function in *Escherichia coli*

Yaowen Chang[†], Xuhui Zhang[†], Alastair I. H. Murchie* and
Dongrong Chen*

Fudan University Pudong Medical Center, and Institute of Biomedical Sciences, Shanghai Medical College, Key Laboratory of Medical Epigenetics and Metabolism, Fudan University, Shanghai, China

Aminoglycosides are not only antibiotics but also have wider and diverse non-antibiotic cellular functions. To elucidate the understanding of non-antibiotic cellular functions, here we report transcriptome-profiling analysis of *Escherichia coli* in the absence or presence of 0.5 and 1 μ M of Kanamycin B, concentrations that are neither lethal nor inhibit growth, and identified the differentially expressed genes (DEGs) at two given concentrations of Kanamycin B. Functional classification of the DEGs revealed that they were mainly related to microbial metabolism including two-component systems, biofilm formation, oxidative phosphorylation and nitrogen metabolism in diverse environments. We further showed that Kanamycin B and other aminoglycosides can induce reporter gene expression through the 5' UTR of *napF* gene or *narK* gene (both identified as DEG) and Kanamycin B can directly bind to the RNA. The results provide new insights into a better understanding of the wider aminoglycosides cellular function in *E. coli* rather than its known antibiotics function.

KEYWORDS

Kanamycin B, transcriptome, aminoglycosides, differentially expressed genes, binding

Introduction

Aminoglycosides are antibiotics that are known to inhibit translation by binding to the A site of the ribosome during the translocation process (Davies and Davis, 1968; Fourmy et al., 1996; Carter et al., 2000), including Kanamycin B, Gentamicin, Amikacin, Tobramycin, Neomycin, and Streptomycin etc. Aminoglycosides are broad-spectrum antibiotics that are potent against various Gram-positive and Gram-negative organisms including members of the *Enterobacteriaceae* family such as *Escherichia coli* (*E. coli*) (Krause et al., 2016).

There have been reports on the response and effect of aminoglycosides in *E. coli* and in other bacteria. Biofilms are bacterial aggregates that resist antibiotic treatment. Subinhibitory concentrations of aminoglycosides were shown to regulate biofilm formation in *E. coli* and also in *Pseudomonas aeruginosa* (Hoffman et al., 2005; Khan et al., 2020). Gentamycin at

sub-inhibitory concentrations can inhibit the swarming motility of *E. coli* (Zhuang et al., 2016). In *E. coli*, it is known that the uptake of aminoglycosides requires proton motive force generated by electron flow through the respiratory chain (Complexes I and II containing Fe-S clusters). Fe-S proteins play an important role in aminoglycoside killing through their effect on aminoglycoside uptake (Ezraty et al., 2013). In some organisms the aminoglycosides are linked to the bacterial SOS response. For example, aminoglycosides were found to induce SOS responses in *Vibrio cholerae*. However, in *E. coli*, aminoglycosides was shown not to induce SOS (Baharoglu and Mazel, 2011).

Genome-wide analysis of transcriptome profiling of *E. coli* and other bacteria upon treatment with specific aminoglycosides have been reported. Upon treatment with subinhibitory dose of gentamicin (1/2 MIC) that aims to find genes that may be involved in the development of adaptive resistance, transcriptome data of *E. coli* has revealed that genes associated with membrane protein and transporter were highly regulated. The transcriptome data identified yhjX (a putative transporter) as most upregulated gene and it regulates the cell growth of *E. coli* at subinhibitory dose of gentamicin and associate with the adaptive resistance to gentamicin (Zhou et al., 2019). Transcriptome profiling of *Pseudomonas aeruginosa* upon treatment with Tobramycin revealed expression changes in genes involved in a wider cellular function such as amino acid catabolism, tricarboxylic acid cycle (TCA) or bacterial motility and attachment etc. (Cianciulli Sesso et al., 2021). RNA-Seq analysis of *Pseudomonas aeruginosa* swarming cells at subinhibitory concentrations of tobramycin identified expression changes in gene encoding multidrug efflux pump and virulence factors (Coleman et al., 2020). In addition, transcriptional profiling changes in *Mycobacterium tuberculosis* in response to Kanamycin have been observed (Habib et al., 2017).

The use of aminoglycosides in the clinic has led to the development of resistance. Aminoglycoside resistance occurs commonly through modification of the aminoglycosides, methylation of their target rRNA or overexpression of efflux pumps (Nikaido, 2009). Modification of the aminoglycosides is achieved by aminoglycoside acetyltransferases (AAC) for acetylation, aminoglycoside adenylyltransferase (AAD) for adenylation, or phosphotransferases for phosphorylation (Mingeot-Leclercq et al., 1999). Aminoglycoside resistance is known to be inducible (Lovett and Rogers, 1996). One mechanism of induction of aminoglycosides resistance has been identified. The aminoglycosides can bind to riboswitch RNA in *Pseudomonas fluorescens* to regulate translation of AAC or AAD expression (He et al., 2013; Jia et al., 2013; Chen and Murchie, 2014; Wang et al., 2019; Zhang et al., 2020). The aminoglycoside sensing riboswitch

is located at the 5' untranslated region of AAC or AAD. Reporter assays in *E. coli* showed that aminoglycosides specifically induce expression of a reporter gene through the riboswitch RNA (Jia et al., 2013). Direct binding of aminoglycosides and the riboswitch RNA was measured. Chemical probing showed that aminoglycoside binding to the riboswitch RNA induces a structural change in Shine-Delgarno sequence (bacterial ribosomal binding site) that in turn regulate the translation of AAC or AAD.

These evidences indicate that the aminoglycoside has wider and diverse cellular functions as well as act as antibiotics. Here we have carried out transcriptome analysis by RNA-sequencing (RNA-seq) to investigate the genome-wide effect of Kanamycin B on *E. coli*, focusing on founding non-antibiotic cellular functions and kanamycin binding RNA. We identified the differentially expressed genes (DEGs) that are mainly related to microbial metabolism in diverse environments, two-component systems, nitrogen metabolism upon Kanamycin B treatment. We further showed that Kanamycin B could bind to 5' UTR of the DEG and induce reporter gene expression. The results provide an insight into a better understanding of the wider aminoglycosides function in *E. coli*.

Materials and methods

Bacterial strains and growth condition

The strain used in this study was *E. coli* K-12 derived JM109 strain (Sangong, China). A single colony of JM109 was inoculated in 5 ml of rich medium LB and cultured overnight at 37°C under the condition of shaking speed at 200 rpm. The overnight bacterial culture was diluted 1: 1000 in fresh LB at 0, 0.5, 1, and 2 µM Kanamycin B. OD600 was measured using a SpectraMax® M5 every 1 h for 12 h. Each experiment was replicated three times.

RNA extraction and library preparation for transcriptome sequencing

E. coli JM09 was cultured with 0, 0.5, and 1 µM Kanamycin B for 7 h to an OD600 ≈ 0.6. Total RNA of harvested cells was extracted using Trizol (Invitrogen, United States). Total RNA of each sample was quantified and qualified by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States), NanoDrop (Thermo Fisher Scientific Inc.), and 1% agarose gel. The NEBNext Ultra II Directional RNA Library Prep Kit was used for the preparation of the library. The library preparations were performed following the manufacturer's protocol. Three biological replicates of each library were sequenced on an Illumina HiSeq X Ten (Illumina, San Diego, CA, United States) with 150-bp paired-ends by AZENTA (Suzhou, China). RNA-seq data were uploaded to the Sequence Read Archive of the National Center for Biotechnology Information (accession number: PRJNA756617).

Abbreviations: PCA, principal component analysis; DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene ontology; STEM, Short Time-series Expression Miner; RT-qPCR, real-time quantitative PCR; *E. coli*, *Escherichia coli*; MST, MicroScale thermophoresis; MIC, minimal inhibitory concentration.

Sequence reads mapping and assembly

The raw reads of fastq format were initially filtered by removing reads containing adaptors, poly-N and low-quality reads. The high-quality clean reads were obtained (reads contain 20% base quality lower than Q20). Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. The analyses were performed by using high quality clean reads. The clean reads were subsequently mapped to the *E. coli* str. K-12 substr. MG1655, genome assembly (NCBI: NC_000913.3) by using bowtie2 v2.2.6.

Principal component analysis

Principal Component Analysis (PCA) was performed to investigate the reproducibility of the biological repeats. DESeq2 in the `deseq2_qc.r` script was used for PCA calculation. The scores of the first and second principal components were plotted into two-dimensional space to represent the spatial relationships within the repeat samples for visualization (Long et al., 2020).

Differential expression analysis and functional annotations of DEGs

The expression levels of genes were counted by using HTseq in each sample (Anders et al., 2015). The DEGs are filtered to meet the threshold level with a $|\text{fold change}| \geq 2$, and false discovery rate value of $p < 0.05$ in each pairwise comparison by using edgeR (Robinson et al., 2010). Functional annotations of DEGs are performed by R package, ClusterProfiler (Yu et al., 2012) with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. KEGG pathways with $p < 0.05$ were considered as significantly enriched.

Hierarchical clustering analysis

Hierarchical clustering analysis of DEGs was carried out by using the function heatmap. Two of gplots (version 3.0.1) in R version 3.3.1. The heatmaps are the outcome of this analysis to represent the reproducibility of the biological repeats (Khan et al., 2019).

Series-cluster analysis and functional annotation of the clusters

Short Time-series Expression Miner, version 1.3.13 (STEM) as a non-parametric clustering algorithm is designed to analyze short timeseries or concentration series expression data (Ernst and Bar-Joseph, 2006). STEM is a new clustering method that can identify real patterns from patterns. It clusters genes based on a

series of pre-defined patterns (expression profiles). A profile is considered meaningful if the number of genes assigned to it surpasses the number of genes that are estimated to occur by chance with $p < 0.05$. The functional annotations of the clusters were analyzed by Gene ontology (GO) that is imbedded in the STEM. GO enrichment analysis is considered significant if the value of $p < 0.05$.

Validation of DEGs by real-time quantitative PCR

The 9 RNA samples for transcriptome sequencing were also used to perform real-time quantitative PCR (RT-qPCR) to assess the reliability of the sequencing data. RNA were reverse transcribed to cDNA using by PrimeScript™ RT reagent Kit (Takara, Japan) by reverse transcription. Obtained cDNAs were further used for RT-qPCR. The RT-qPCR reactions were carried out on a 7,500 Real-Time PCR System (Applied Biosystems) using the TB Green® Premix Ex Taq™ II (Takara, Japan) according to the manufacturer's instructions. For absolute quantification of DEGs copy number, a gene-specific PCR of known concentration was prepared as a standard. The primers were obtained from AZENTA (Suzhou, China; Supplementary Table S1). Melting curve analysis was utilized to verify specificity of all PCR products. Error bars are the standard deviation of three independent experiments.

Reporter plasmid construction

The DNA sequences corresponding to the 5' UTR of *napF* or *narK* were cloned into the reporter vector pGEX-leaderRNA-aac/aadlacZα as described before (Jia et al., 2013) to generate reporter plasmid pGEX-*napF* 5' UTR-lacZα or pGEX-*narK* 5' UTR-lacZα. The JM109 competent cells are placed on ice for 30 min. 1 μl of pGEX-*napF* 5' UTR-lacZα plasmid was added to 50 μl of competent cells for transformation. The mixture was incubated on ice for 30 min. The cells were heat-shocked at 42°C in a heating block for 90 s followed by cooling down on ice for 2 min. Transfer the transformed competent cells onto LB plates containing Ampicillin (100 μg/ml). Spread with aseptic spreader uniformly until it gets resistant to move. Invert the plates and incubate at 37°C overnight and transformed colonies should appear in 12–16 h.

The reporter assays: Agar diffusion and β-galactosidase assays

Five milliliters culture of JM109 cells transformed with the pGEX-*napF* 5' UTR-lacZα reporter plasmid was incubated overnight in LB broth containing 100 μg/ml Ampicillin. The culture was diluted 1: 100 in fresh LB broth containing Ampicillin (100 μg/

ml) and IPTG (0.5 mM) and incubated for 2 h. A total of 1 ml of cultured cells (an $OD_{600} \approx 0.3$) was then mixed well with 10 ml of 0.6% LB agar at 45°C. After mixing of components, the cells were poured on a 1.5% LB agar plate containing 100 µg/ml Ampicillin, 0.5 mM IPTG, and 200 µg/ml X-Gal. For the control sample IPTG was not added in the culture or on the agar plate. After the agar had solidified, discs of filter paper spotted with 2 µl different antibiotics stock solutions were placed onto agar plates and incubated at 37°C for at least 18 h (Bailey et al., 2008). Stock solutions included Kanamycin B (10 mM), Sisomicin (10 mM), Neamine (150 mM), Amikacin (10 mM), Gentamycin (10 mM), Tobramycin (10 mM), Ribostamycin (10 mM), Paromomycin (10 mM), Levofloxacin (5 mM), Tetracycline (10 mM), Erythromycin (100 mM), Trimethoprim (5 mM). β -Galactosidase assays were performed as previously described (Zhang and Bremer, 1995). Error bars are the standard deviation of three independent experiments.

MicroScale thermophoresis

The 5' UTR of *napF* RNAs were prepared by *in vitro* transcription using T7 RNA polymerase. After PAGE purification, RNA was labeled with fluorescein-5-thiosemicarbazide as previously described (Wu et al., 1996). The fluorescein-RNA was annealed separately with 40 mM HEPES (pH 7.4), 100 mM KCl, the mixture was heated at 95°C for 3 min followed by subsequently placed on ice for 3 min. $MgCl_2$ was added to a final concentration of 2.5 mM. Series of aminoglycosides were prepared with a final concentration from 125 nM to 128 µM. After incubation at 25°C for 20 min. The samples were added to premium coated capillaries (NanoTemper Technologies, GmbH, Munich, Germany) and subsequently subjected to MST analysis (20% MST power, 30% LED power) on a Monolith NT.115 pico instrument at 25°C. And data analysis was performed by MO. Affinity Analysis software

(Entzian and Schubert, 2016; Moon et al., 2018). Error bars are the standard deviation of three independent experiments. The signal-to-noise ratio reflects the quality of the binding data and a ratio more than 12 reflects an excellent assay.

Results

Overview of the RNA-seq data

To investigate the effect of Kanamycin B on *E. coli* JM109, we performed transcriptome-profiling analysis by RNA-seq in the absence or presence of Kanamycin B. Since Kanamycin B is a bactericidal antibiotic, initial cultures in the absence or presence of a broad range of Kanamycin B were cultivated to select concentrations that are neither lethal nor inhibit growth (Figure 1A). 0.5 or 1 µM of Kanamycin B showed minimal effects on cell growth, while 2 µM of Kanamycin B clearly inhibited growth. As a result, 0.5 and 1 µM of Kanamycin B were chosen for transcriptome profiling analysis. Total RNA was extracted from cells at mid-log phase that had been treated with 0, 0.5, and 1 µM of Kanamycin B for RNA-seq analysis. The transcriptome-profiling data was uploaded to the National Center for Biotechnology Information Sequence Read Archive (accession number PRJNA756617). With three biological repeats of each sample, a total of 9 cDNA libraries were constructed containing 340.20 million raw reads; 339.64 million clean reads (accounting for 99.83% of raw reads) were recorded after removing adapter sequences and reads of low quality and those with more than 5% N bases. The average number of clean reads per sample was about 37.74 million and the clean Q20 (sequencing error rate < 1%) base rate was >97.86% for each sample. Ultimately, 336.74 million high-quality reads (accounting for 99.14% of clean reads) were mapped to the *E. coli* str. K-12 substr. MG1655 genome.

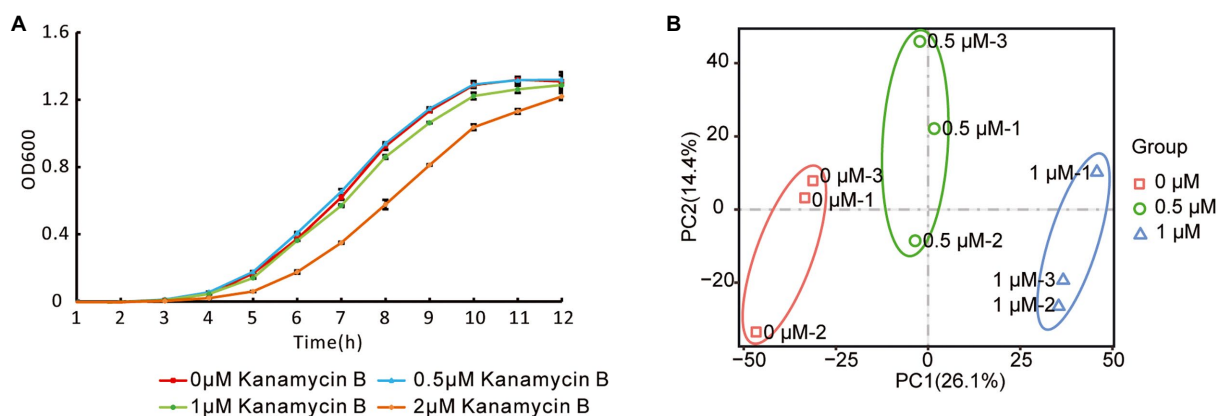


FIGURE 1
Escherichia coli growth curve and reproducibility of RNA-seq data. (A) Growth curve of JM109 treated with or without Kanamycin B. JM109 cells were grown in the presence of 0, 0.5, 1, and 2 µM Kanamycin B. Cell density was measured hourly. Error bars are the standard deviation of three independent experiments. (B) Principal component analysis of the RNA-Seq data. The small icon represents the original samples, colors are used to differentiate between treatments: red (0 µM), green (0.5 µM), and blue (1 µM).

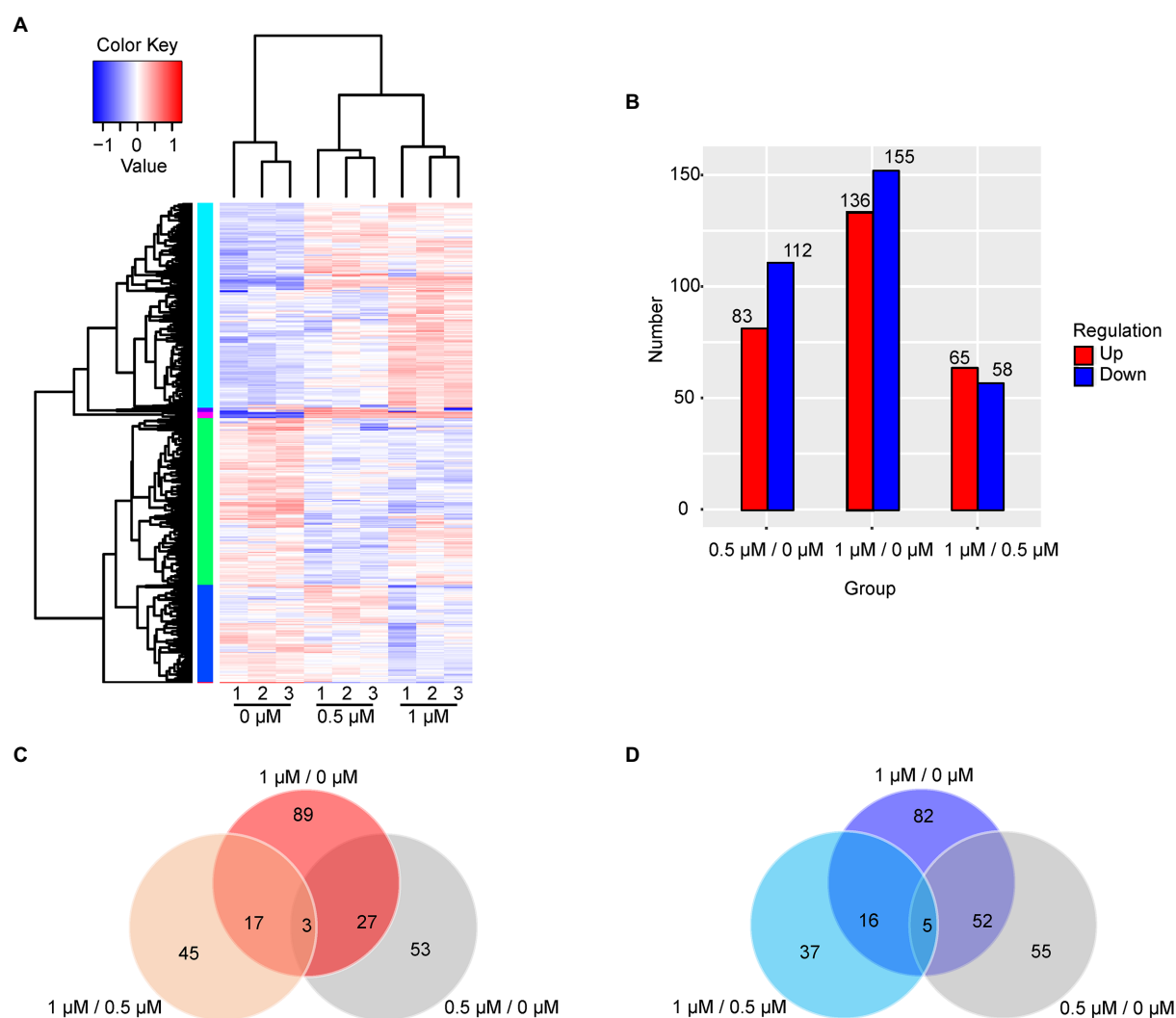


FIGURE 2

Analysis of DEGs. (A) The hierarchical cluster analysis of 463 DEGs. Rows represent gene expression levels and columns represent concentration of Kanamycin B, treatment and repeats. The relative quantitative changes within the row are shown in color: red suggests a relative higher expression whereas blue suggests a relative lower expression level. (B) The number of DEGs in the comparison groups is presented in histogram. Upregulated DEGs (red), and downregulated DEGs (blue) were presented by histogram. (C) The Venn diagram displays the number of the shared and unique genes of upregulated DEGs. (D) The Venn diagram displays the number of the shared and unique genes of downregulated DEGs.

To investigate their reproducibility, the biological repeats were analyzed by PCA. The results revealed that PC1 and PC2 had values of 26.1% and 14.4%, respectively, and accounted for 40.5% of the principal components (Figure 1B). PCA showed consistency between the three replicates at each Kanamycin B concentration; samples from the biological repeat clustered together, reflecting minimal differences between them. PCA suggested that the three biological repeats in this study were reasonably reproducible (Figure 1B).

Gene expression profile in response to Kanamycin B treatment

The cellular response to Kanamycin B was revealed by the changes in the levels of gene expression. The DEGs in response

to treatment by Kanamycin B were identified by selecting gene expression levels with $|\text{fold changes}| \geq 2$ and significant differences in p -values of <0.05 at 0.5 or 1 μ M. Figure 2A shows a cluster analysis of the DEGs (for a full list of DEGs; see Supplementary Table S2). There are two main clusters: one with genes that were induced and one with genes that were repressed in response to Kanamycin B treatment. We identified 83 or 136 induced DEGs and 112 or 155 repressed DEGs upon treatment with 0.5 or 1 μ M of Kanamycin B, respectively, compared to no treatment (Figure 2B). There are 65 upregulated and 58 downregulated DEGs in the 1/0.5 μ M (comparisons between sample treated with 1 and 0.5 μ M Kanamycin B) group (Figure 2B). Fewer numbers of DEGs were observed in the 1/0.5 μ M group compared to the 0.5/0 μ M and 1/0 μ M (comparisons between sample treated with 0.5 or

1 μ M Kanamycin B and the control) groups suggesting that 0.5 or 1 μ M Kanamycin B causes similar cellular responses. We next compared the similarities and differences of the DEGs in response to 0.5 or 1 μ M Kanamycin B treatment. The numbers of overlapping genes between the groups (0.5/0 μ M, 1/0 μ M, and 1/0.5 μ M) are shown in the Venn diagram for upregulated or downregulated DEGs (Figures 2C,D) and overlapping groups listed (Supplementary Table S3). There is considerable overlap between 0.5/0 μ M and 1/0 μ M DEG groups; 30 genes and 57 genes were co-induced or co-repress in 0.5 and 1 μ M of Kanamycin B treatment. All of the DEGs represent 8.8% of the whole genome transcripts. In addition, the results show that the magnitude of the transcriptional responses in DEGs varied significantly from gene to gene. For example, some genes are upregulated over 10-fold upon Kanamycin B treatment while the expression of many other genes are induced only by about 2-fold (Table 1; Supplementary Table S2).

Functional classification of DEGs

To investigate the function of the DEGs, we carried out a KEGG pathway analysis of DEGs. Table 1 (1/0 μ M), Supplementary Table S4 (0.5/0 μ M), and Figure 3 show the DEGs classified into known or predicted functional groups. These genes encode proteins involved in the following cellular functions: microbial metabolism in diverse environments, two-component systems, nitrogen metabolism, butanoate metabolism, arginine and proline metabolism and inositol phosphate metabolism. Genes that encode proteins that participate in cellular degradation, including chloroalkane and chloroalkene degradation, naphthalene degradation, chlorocyclohexane and chlorobenzene degradation, fluorobenzoate degradation, toluene degradation and degradation of aromatic compounds. Genes that encode proteins involved in oxidative phosphorylation, aminoacyl-tRNA biosynthesis or the TCA cycle. KEGG pathway analysis of DEGs showed that many genes are classified in more than one functional group (Table 1). Further analysis of the genes that belong to the five relatively large functional classes (two-component systems, oxidative phosphorylation, nitrogen metabolism, microbial metabolism in diverse environments and butanoate metabolism) revealed the linkage network of these DEGs (Figure 4; Table 1), suggesting an overlap between these main functional classes through these DEGs.

Series-cluster analysis and functional annotation of the clusters

In addition to the DEGs analysis of the cellular response to Kanamycin B, we observed diverse and complex gene expression patterns over the Kanamycin B concentration courses. To dissect the gene expression patterns and further understand the gene

functions, we performed a series-cluster analysis and functional annotation of the clusters. The series-cluster analysis can group genes that have similar expression patterns to form clusters of profiles and GO enrichment analysis further reveals functional annotations. A total of 4,534 genes were partitioned into 16 profiles of which 5 profiles were statistically significant ($p < 0.05$; Figure 5; Supplementary Table S5). STEM calculated the five profiles into two clusters. Each gene cluster exhibited a distinctive expression pattern. Cluster 1 includes profile 11 (115 genes), 12 (98 genes), 13 (107 genes), and 15 (116 genes). Genes in profile 11 show that expression levels increase rapidly after 0.5 μ M Kanamycin B treatment and reach a plateau with 1 μ M Kanamycin B, while the genes in profile 12, 13, and 15 present a gradual upregulation pattern. Cluster 2 consists of only profile 8 with 86 genes that exhibit similar expression from 0 to 0.5 μ M Kanamycin B treatment and an increase in transcription with 1 μ M Kanamycin B. STEM also supports GO enrichment analysis for each cluster and profile and includes functional classification of genes in the profiles in Figure 5 (Supplementary Table S6). Profile 11 includes 115 genes of which 18 GO terms are associated with nitrogen compound metabolism and oxidation–reduction processes. The profile 12 cluster contains a number of membrane genes, and the bacterial outer membrane contributes to drug efflux resistance mechanisms. The profile 13 includes 107 genes enriched in 68 GO terms that are linked to DNA binding and the regulation of RNA biosynthesis. Profile 15 has 116 genes enriched in 10 GO Terms of which 8 are related to biofilm formation and cell adhesion. Profile 8 in cluster 2 is associated with genes involved in RNA biosynthetic processes, regulation of primary metabolic process and cellular biosynthetic process.

Validation by RT-qPCR

To verify the accuracy and reproducibility of the transcriptome results by an alternative method, we performed quantitation by RT-qPCR. We selected 8 genes that were differentially expressed upon treatment with Kanamycin B. A comparative analysis of all the selected genes showed a similar expression pattern in the RT-qPCR analysis as observed in RNAseq data (Figure 6). Thus, the RT-qPCR experiments confirm the reliability of the transcriptome sequencing data.

The Kanamycin B and other aminoglycosides induce reporter gene expression through the UTR of the identified DEG

The *napF* gene is one of the identified DEGs (Figure 6; Table 1) and the transcription start site of the *napF* gene has been mapped (Stewart et al., 2003). There are 80 nucleotides (nt) upstream of the transcription start site to the coding sequence of the *napF* gene. It has been shown that aminoglycosides can bind to the 5' UTR of

TABLE 1 Functional classification of DEGs (1/0 μ M).

Gene symbol	Fold change	Regulation	Location	Gene annotation
Microbial metabolism in diverse environments				
<i>nirD*</i>	219.45	Up	Cytosol	Nitrite reductase subunit NirD
<i>nirB</i>	5.38	Up	Cytosol	Nitrite reductase catalytic subunit NirB
<i>narI</i>	11.30	Up	Inner membrane	Nitrate reductase A subunit γ
<i>narG</i>	9.58	Up	Inner membrane	Nitrate reductase A subunit α
<i>narH</i>	7.48	Up	Inner membrane	Nitrate reductase A subunit β
<i>ydhX</i>	3.61	Up	Periplasmic space	Putative 4Fe-4S ferredoxin-like protein YdhX
<i>glcD</i>	3.09	Up	Cytosol	Glycolate dehydrogenase, putative FAD-linked subunit
<i>glpE</i>	2.75	Up	Cytosol	Thiosulfate sulfurtransferase GlpE
<i>yqeF</i>	2.39	Up	Cytosol	Putative acyltransferase
<i>lysC</i>	2.14	Up	Cytosol	Aspartate kinase III
<i>frdD</i>	-9.37	Down	Inner membrane	Fumarate reductase membrane protein FrdD
<i>frdB</i>	-9.33	Down	Inner membrane, cytosol	Fumarate reductase iron-sulfur protein
<i>frdA</i>	-6.35	Down	Inner membrane, cytosol	Fumarate reductase flavoprotein subunit
<i>frdC</i>	-3.00	Down	Inner membrane	Fumarate reductase membrane protein FrdC
<i>hyaB</i>	-4.88	Down	Periplasmic space, inner membrane	Hydrogenase 1 large subunit
<i>hyaA</i>	-3.41	Down	Inner membrane	Hydrogenase 1 small subunit
<i>adhE</i>	-2.86	Down	Cytosol	Alcohol dehydrogenase/aldehyde-dehydrogenase
<i>adhP</i>	-2.67	Down	Cytosol	Ethanol dehydrogenase / alcohol dehydrogenase
<i>gabT</i>	-3.19	Down	Cytosol	4-aminobutyrate aminotransferase GabT
<i>gabD</i>	-2.68	Down	Cytosol	Succinate-semialdehyde dehydrogenase (NADP+) GabD
<i>gadA</i>	-2.43	Down	Cytosol	Glutamate decarboxylase A
<i>gadB</i>	-2.10	Down	Cytosol, membrane	Glutamate decarboxylase B
<i>hcaE</i>	-2.86	Down	No annotation	Putative 3-phenylpropionate/cinnamate dioxygenase subunit α
<i>paaK</i>	-2.73	Down	Cytosol	Phenylacetate-coa ligase
<i>yghX</i>	-2.57	Down	No annotation	Putative hydrolase fragment
<i>aldB</i>	-2.43	Down	Cytosol	Aldehyde dehydrogenase B
<i>fbaB</i>	-2.19	Down	Cytosol	Fructose-bisphosphate aldolase class I
<i>tktB</i>	-2.19	Down	Cytosol	Transketolase 2
<i>allB</i>	-2.16	Down	Cytosol	Allantoinase
Two-component system				
<i>fdnG</i>	19.98	Up	Periplasmic space	Formate dehydrogenase n subunit α
<i>fdnH</i>	11.67	Up	Inner membrane	Formate dehydrogenase n subunit β
<i>fdnI</i>	5.38	Up	Inner membrane	Formate dehydrogenase n subunit γ
<i>narI</i>	11.30	Up	Inner membrane	Nitrate reductase a subunit γ
<i>narG</i>	9.58	Up	Inner membrane	Nitrate reductase a subunit α
<i>narH</i>	7.48	Up	Inner membrane	Nitrate reductase a subunit β
<i>narJ</i>	6.84	Up	Cytosol	Nitrate reductase 1 molybdenum cofactor assembly chaperone
<i>narX</i>	2.24	Up	Periplasmic space, inner membrane	Sensory histidine kinase NarX
<i>rstB</i>	2.50	Up	Inner membrane	Sensory histidine kinase RstB
<i>uhpB</i>	2.16	Up	Inner membrane	Sensory histidine kinase UhpB
<i>baeS</i>	2.20	Up	Inner membrane	Sensor histidine kinase BaeS
<i>rcaA</i>	3.35	Up	Cytosol	DNA-binding transcriptional activator RcaA
<i>ompF</i>	2.98	Up	Outer membrane	Outer membrane porin F
<i>yqeF</i>	2.39	Up	Cytosol	Putative acyltransferase
<i>uhpT</i>	2.19	Up	Inner membrane	Hexose-6-phosphate:phosphate antiporter
<i>frdD</i>	-9.37	Down	Inner membrane	Fumarate reductase membrane protein FrdD
<i>frdB</i>	-9.33	Down	Inner membrane, cytosol	Fumarate reductase iron-sulfur protein
<i>frdA</i>	-6.35	Down	Inner membrane, cytosol	Fumarate reductase flavoprotein subunit
<i>frdC</i>	-3.00	Down	Inner membrane	Fumarate reductase membrane protein FrdC

(Continued)

TABLE 1 (Continued)

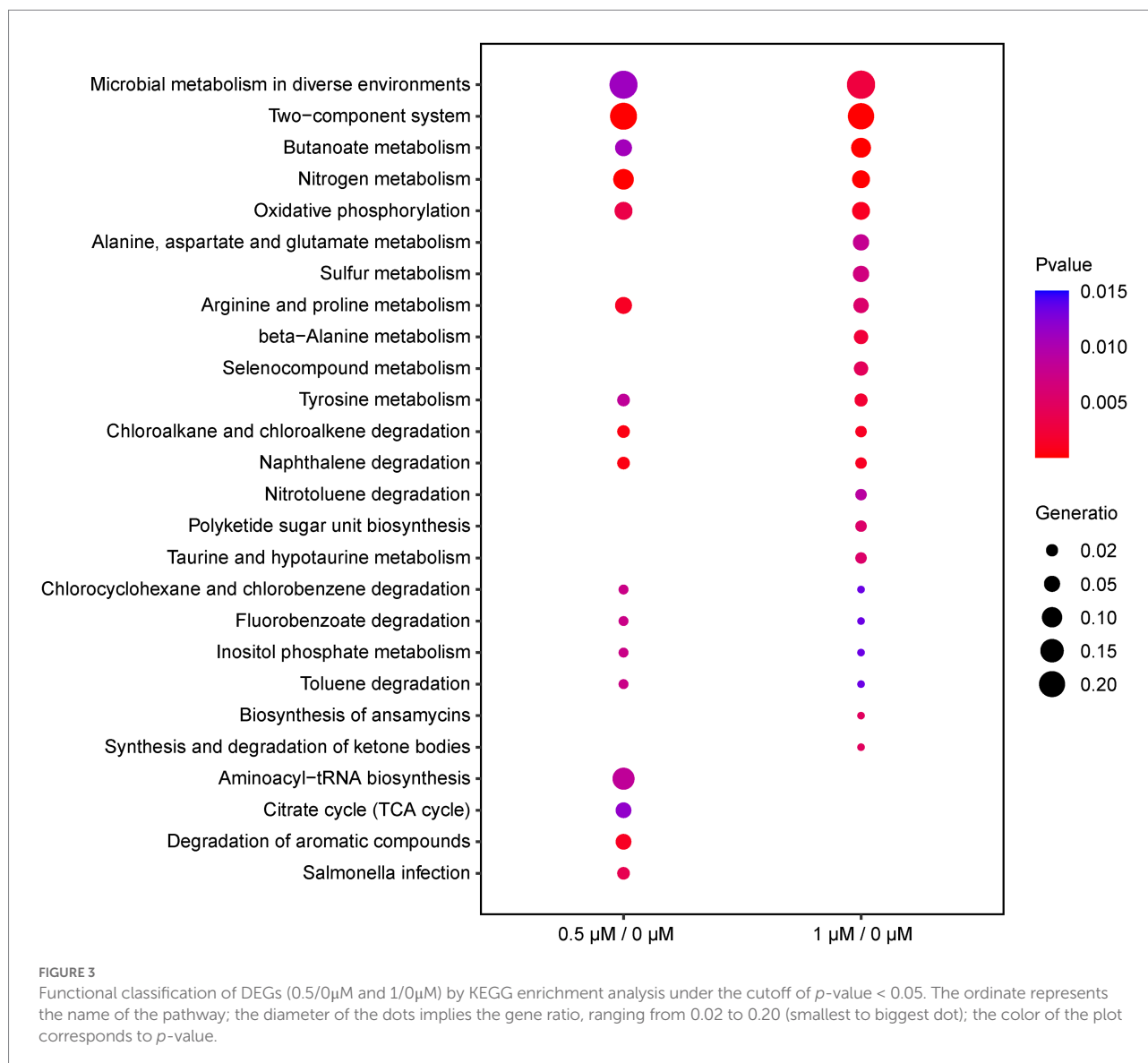
Gene symbol	Fold change	Regulation	Location	Gene annotation
<i>appB</i>	−6.19	Down	Inner membrane	Cytochrome bd-ii ubiquinol oxidase subunit ii
<i>appC</i>	−4.33	Down	Inner membrane	Cytochrome bd-ii ubiquinol oxidase subunit i
<i>hyaC</i>	−7.15	Down	Inner membrane	Hydrogenase 1 cytochrome subunit
<i>cusA</i>	−2.99	Down	Inner membrane	Copper/silver export system rnd permease
<i>mdtC</i>	−2.58	Down	Inner membrane	Multidrug efflux pump rnd permease subunit MdtC
Butanoate metabolism				
<i>yqeF</i>	2.39	Up	Cytosol	Putative acyltransferase
<i>frdD</i>	−9.37	Down	Inner membrane	Fumarate reductase membrane protein FrdD
<i>frdB</i>	−9.33	Down	Inner membrane, cytosol	Fumarate reductase iron–sulfur protein
<i>frdA</i>	−6.35	Down	Inner membrane, cytosol	Fumarate reductase flavoprotein subunit
<i>frdC</i>	−3.00	Down	Inner membrane	Fumarate reductase membrane protein FrdC
<i>gabT</i>	−3.19	Down	Cytosol	4-aminobutyrate aminotransferase GabT
<i>gabD</i>	−2.68	Down	Cytosol	Succinate-semialdehyde dehydrogenase (NADP+) GabD
<i>gadA</i>	−2.43	Down	Cytosol	Glutamate decarboxylase a
<i>gadB</i>	−2.10	Down	Cytosol, membrane	Glutamate decarboxylase b
<i>adhE</i>	−2.86	Down	Cytosol	Alcohol dehydrogenase/aldehyde-dehydrogenase
<i>dmlA</i>	−2.01	Down	Cytosol	D-malate/3-isopropylmalate dehydrogenase (decarboxylating)
Nitrogen metabolism				
<i>nirD*</i>	219.45	Up	Cytosol	Nitrite reductase subunit NirD
<i>nirB</i>	5.38	Up	Cytosol	Nitrite reductase catalytic subunit NirB
<i>narK</i>	16.98	Up	Inner membrane	Nitrate:nitrite antiporter NarK
<i>narI</i>	11.30	Up	Inner membrane	Nitrate reductase a subunit γ
<i>narG</i>	9.58	Up	Inner membrane	Nitrate reductase a subunit α
<i>narH</i>	7.48	Up	Inner membrane	Nitrate reductase a subunit β
<i>hcp</i>	4.03	Up	Cytosol	Proteins-nitrosylase
<i>can</i>	2.41	Up	Cytosol	Carbonic anhydrase 2
Oxidative phosphorylation				
<i>cyoA</i>	2.48	Up	Inner membrane	Cytochromebo3ubiquinol oxidase subunit 2
<i>ppa</i>	2.11	Up	Cytosol	Inorganic pyrophosphatase
<i>frdD</i>	−9.37	Down	Inner membrane	Fumarate reductase membrane protein FrdD
<i>frdB</i>	−9.33	Down	Inner membrane, cytosol	Fumarate reductase iron–sulfur protein
<i>frdA</i>	−6.35	Down	Inner membrane, cytosol	Fumarate reductase flavoprotein subunit
<i>frdC</i>	−3.00	Down	Inner membrane	Fumarate reductase membrane protein FrdC
<i>appB</i>	−6.19	Down	Inner membrane	Cytochrome bd-ii ubiquinol oxidase subunit ii
<i>appC</i>	−4.33	Down	Inner membrane	Cytochrome bd-ii ubiquinol oxidase subunit i
Alanine, aspartate, and glutamate metabolism				
<i>gabT</i>	−3.19	Down	Cytosol	4-aminobutyrate aminotransferase GabT
<i>gabD</i>	−2.68	Down	Cytosol	Succinate-semialdehyde dehydrogenase (NADP+) GabD
<i>gadA</i>	−2.43	Down	Cytosol	Glutamate decarboxylase a
<i>gadB</i>	−2.10	Down	Cytosol, membrane	Glutamate decarboxylase b
<i>aspA</i>	−2.49	Down	Cytosol	Aspartate ammonia-lyase
<i>glsA</i>	−2.40	Down	No annotation	Glutaminase 1
Sulfur metabolism				
<i>ydhX</i>	3.61	Up	Periplasmic space	Putative 4fe-4s ferredoxin-like protein YdhX
<i>glpE</i>	2.75	Up	Cytosol	Thiosulfate sulfurtransferase GlpE
<i>sbp</i>	2.05	Up	Periplasmic space	Sulfate/thiosulfate abc transporter periplasmic binding protein Sbp
<i>dmsB</i>	−4.23	Down	Inner membrane	Dimethyl sulfoxide reductase subunit b
<i>dmsC</i>	−3.93	Down	Inner membrane	Dimethyl sulfoxide reductase subunit c
<i>dmsA</i>	−2.49	Down	Inner membrane	Dimethyl sulfoxide reductase subunit a

(Continued)

TABLE 1 (Continued)

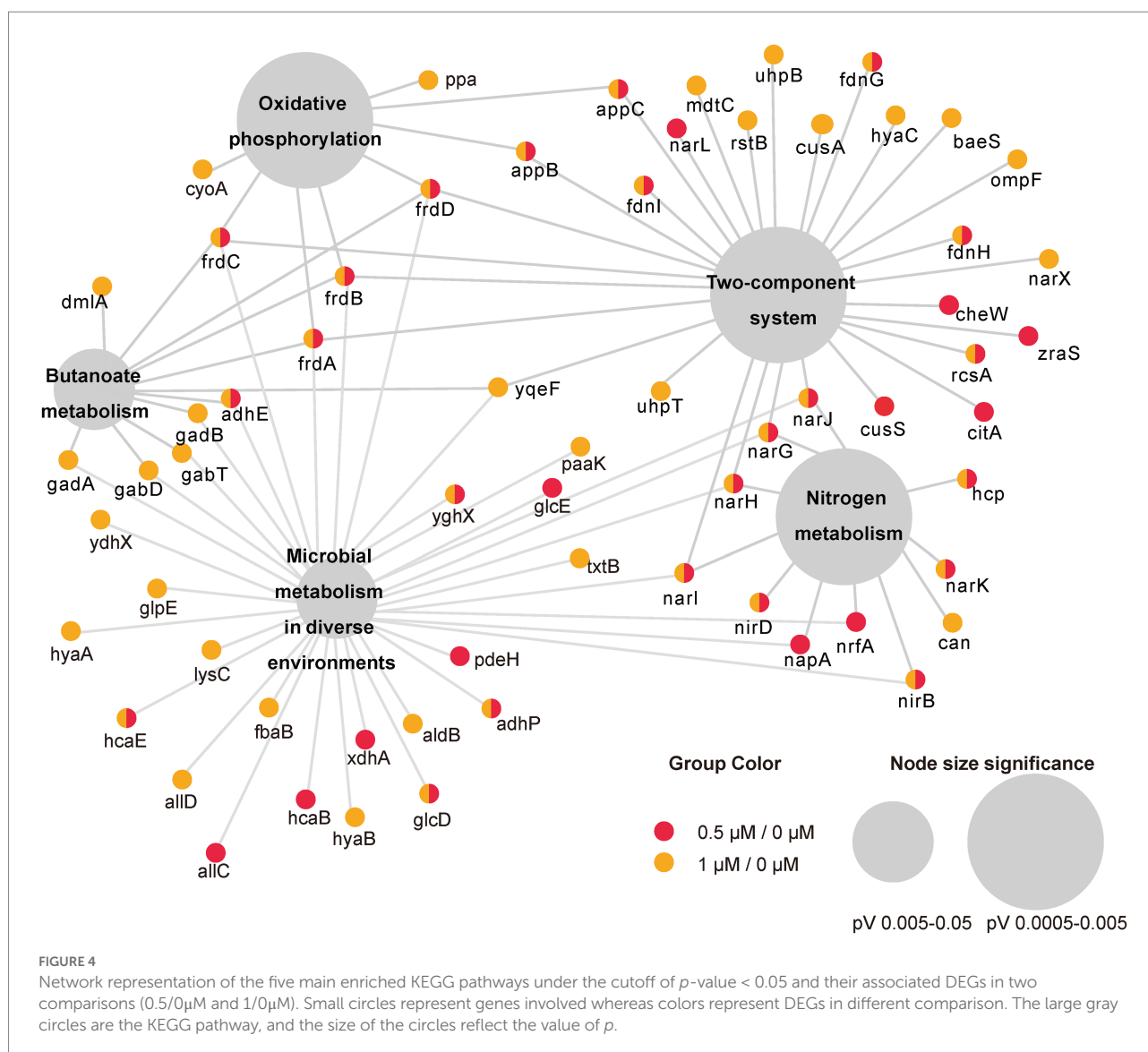
Gene symbol	Fold change	Regulation	Location	Gene annotation
Arginine and proline metabolism				
<i>puuD</i>	3.81	Up	Cytosol	γ -glutamyl- γ -aminobutyrate hydrolase
<i>puuC</i>	3.22	Up	Cytosol	γ -glutamyl- γ -aminobutyraldehyde dehydrogenase
<i>puuA</i>	2.96	Up	Cytosol	Glutamate-putrescine ligase
<i>speG</i>	2.31	Up	Cytosol	spermidine N-acetyltransferase
<i>patA</i>	−2.16	Down	Cytosol	Putrescine aminotransferase
beta-alanine metabolism				
<i>gabT</i>	−3.19	Down	Cytosol	4-aminobutyrate aminotransferase GabT
<i>gabD</i>	−2.68	Down	Cytosol	Succinate-semialdehyde dehydrogenase (NADP+) GabD
<i>gadA</i>	−2.43	Down	Cytosol	Glutamate decarboxylase a
<i>gadB</i>	−2.10	Down	Cytosol, membrane	Glutamate decarboxylase b
Selenocompound metabolism				
<i>ynfE</i>	−9.15	Down	Periplasmic space, inner membrane	Putative selenate reductase YnfE
<i>ynfF</i>	−5.26	Down	Periplasmic space, inner membrane	Putative selenate reductase YnfF
<i>sufS</i>	−2.31	Down	Cytosol	L-cysteine desulfurase
<i>metE</i>	−2.15	Down	Cytosol	Cobalamin-independent homocysteine transmethylease
Tyrosine metabolism				
<i>adhE</i>	−2.86	Down	Cytosol	Alcohol dehydrogenase/aldehyde-dehydrogenase
<i>adhP</i>	−2.67	Down	Cytosol	Ethanol dehydrogenase / alcohol dehydrogenase
<i>gabD</i>	−2.68	Down	Cytosol	Succinate-semialdehyde dehydrogenase (NADP+) GabD
Chloroalkane and chloroalkene degradation				
<i>adhE</i>	−2.86	Down	Cytosol	Alcohol dehydrogenase/aldehyde-dehydrogenase
<i>adhP</i>	−2.67	Down	Cytosol	Ethanol dehydrogenase/alcohol dehydrogenase
Naphthalene degradation				
<i>adhE</i>	−2.86	Down	Cytosol	Alcohol dehydrogenase/aldehyde-dehydrogenase
<i>adhP</i>	−2.67	Down	Cytosol	Ethanol dehydrogenase/alcohol dehydrogenase
Nitrotoluene degradation				
<i>hyaB</i>	−4.88	Down	Periplasmic space, inner membrane	Hydrogenase 1 large subunit
<i>hyaA</i>	−3.41	Down	Inner membrane	Hydrogenase 1 small subunit
Polyketide sugar unit biosynthesis				
<i>rfbD</i>	2.97	Up	Cytosol	Dtdp-4-dehydrorhamnose reductase
<i>rfbA</i>	2.48	Up	Cytosol	Dtdp-glucose pyrophosphorylase
Taurine and hypotaurine metabolism				
<i>gadA</i>	−2.43	Down	Cytosol	Glutamate decarboxylase a
Chlorocyclohexane and chlorobenzene degradation				
<i>yghX</i>	−2.57	Down	No annotation	Putative hydrolase fragment
<i>gadB</i>	−2.10	Down	Cytosol, membrane	Glutamate decarboxylase b
Fluorobenzoate degradation				
<i>yghX</i>	−2.57	Down	No annotation	Putative hydrolase fragment
Inositol phosphate metabolism				
<i>appA</i>	−3.96	Down	Periplasmic space	Periplasmic phosphoanhydride phosphatase/multiple inositol-polyphosphate phosphatase
Toluene degradation				
<i>yghX</i>	−2.57	Down	No annotation	Putative hydrolase fragment
Biosynthesis of ansamycins				
<i>tktB</i>	−2.19	Down	Cytosol	Transketolase 2
Synthesis and degradation of ketone bodies				
<i>yqeF</i>	2.39	Up	Cytosol	Putative acyltransferase

*The counts of the *nirD* gene is 0.



aac/aad gene to regulate downstream gene expression (Jia et al., 2013). To investigate whether Kanamycin B and other aminoglycoside have effects on downstream gene expression through the 5' UTR of the *napF* gene, we constructed reporter plasmids pGEX-*napF* 5' UTR-lacZα in which the 80 nt UTR of the *napF* gene and 9 nt into the coding sequence was under the control of the IPTG-inducible tac promoter (Ptac) and placed upstream of a β-galactosidase (β-gal) reporter gene. The reporter plasmid was transformed into *E. coli* JM109 and β-gal activity was tested in the absence and presence of Kanamycin B or other aminoglycoside antibiotics (Sisomycin, Tobramycin, Gentamycin, Amikacin) or control molecules (Ribostamycin, Neamine or Paromomycin and Levofloxacin, Tetracycline, Erythromycin, Trimethoprim) by agar diffusion assays. We observed induction of the reporter gene expression with Kanamycin B and Sisomycin, Amikacin, Gentamycin, and Tobramycin but not for the control molecules Ribostamycin, Neamine, Paromomycin, Levofloxacin, Tetracycline,

Erythromycin or Trimethoprim (Figure 7A). No induction by Kanamycin B was seen in cells transformed with the reporter plasmid without IPTG in which Ptac promoter is inactive and the 5' UTR of the *napF* gene is not made (Supplementary Figure S1A). No induction was observed on plates without Kanamycin B (Supplementary Figure S1B). No induction by Kanamycin B was seen in cells transformed with the reporter plasmid containing a control sequence with IPTG (Wang et al., 2019). To further quantify the agar diffusion assay, we measured β-gal activity in solution (Zhang and Bremer, 1995) on titration of Kanamycin B and other aminoglycosides (Figures 7B,C; Supplementary Figure S1C). Kanamycin B, Amikacin and Gentamycin induce the reporter gene expression by more than 3-fold. These results suggest that Kanamycin B and other aminoglycosides can induce the reporter gene expression through the 5' UTR of the *napF* gene. The reporter gene expression requires both Kanamycin B or other aminoglycosides and the 5' UTR of the *napF* gene. In parallel,



we also observed that Kanamycin B and other aminoglycosides can induce the reporter gene expression through the 5' UTR of the *narK* gene (another DEG; [Supplementary Figure S2](#)).

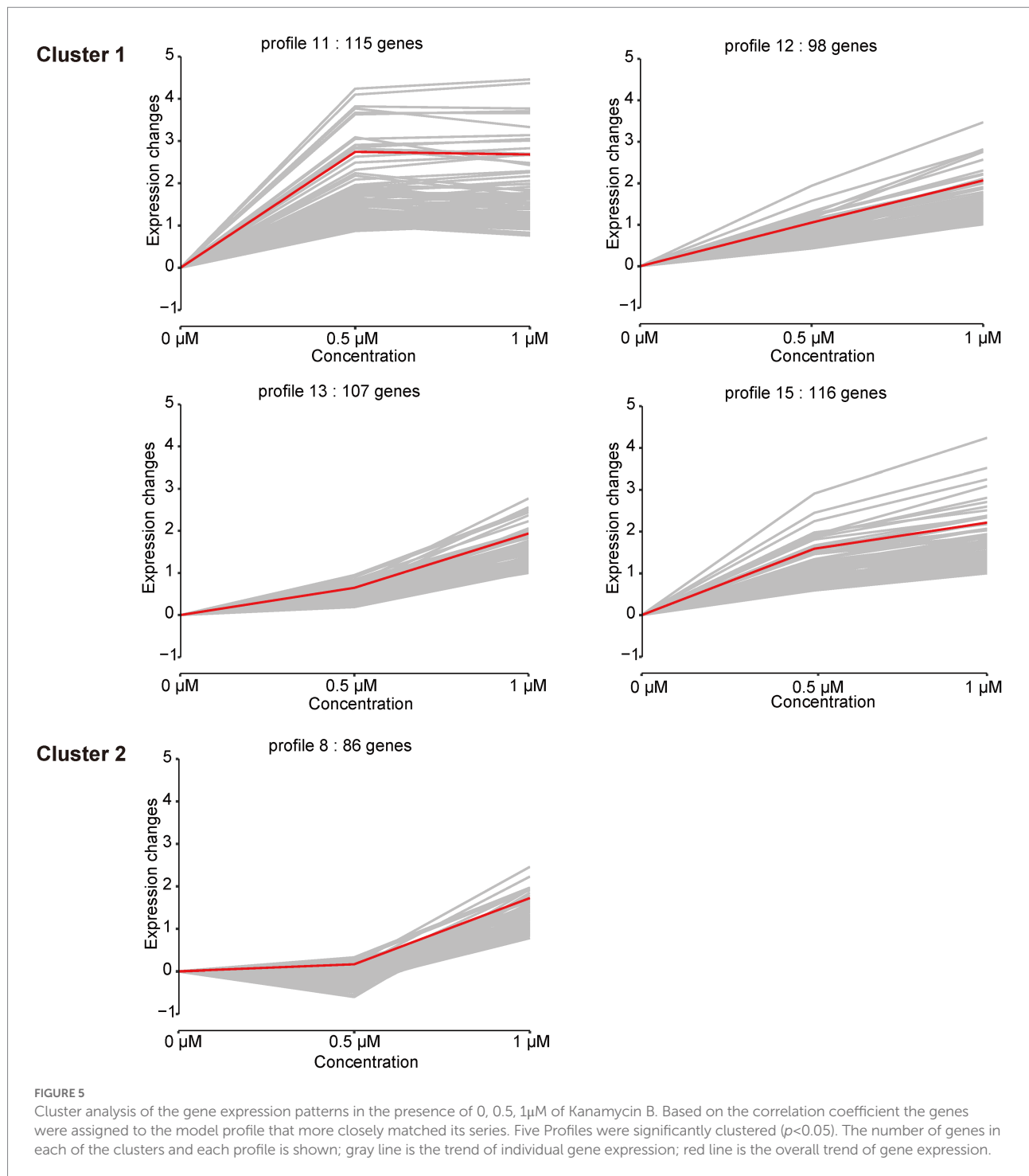
Kanamycin B binds to 5' UTR RNA

To investigate if Kanamycin B directly bind to the 5' UTR of the *napF* gene to induce reporter gene expression, we used MST to measure the binding of Kanamycin B or control molecule (Ribostamycin or Neamine) to the 5' UTR of the *napF* gene or Kanamycin B to a random RNA. The 5' UTR of the *napF* gene was prepared by *in vitro* transcription using in-house purified T7 RNA polymerase and was labeled with fluorescein-5-thiosemicarbazide as previously described ([Wu et al., 1996](#)). Binding measurements was made on a Monolith NT.115 system by NanoTemper Technologies ([Entzian and Schubert, 2016](#); [Moon et al., 2018](#)). An increase in the measured response upon titration of the

Kanamycin B was observed, indicating the formation of a Kanamycin B-RNA complex. Kanamycin B binds to the RNA with affinity at 4.5 μM. In contrast, Ribostamycin, Neamine have weaker binding and lower affinities ([Figure 8A](#)) and no binding between Kanamycin B and the control RNA (*ilvL*) was detected ([Figure 8B](#)). Thus, Kanamycin B that induces reporter gene expression can bind to the 5' UTR of the *napF* gene.

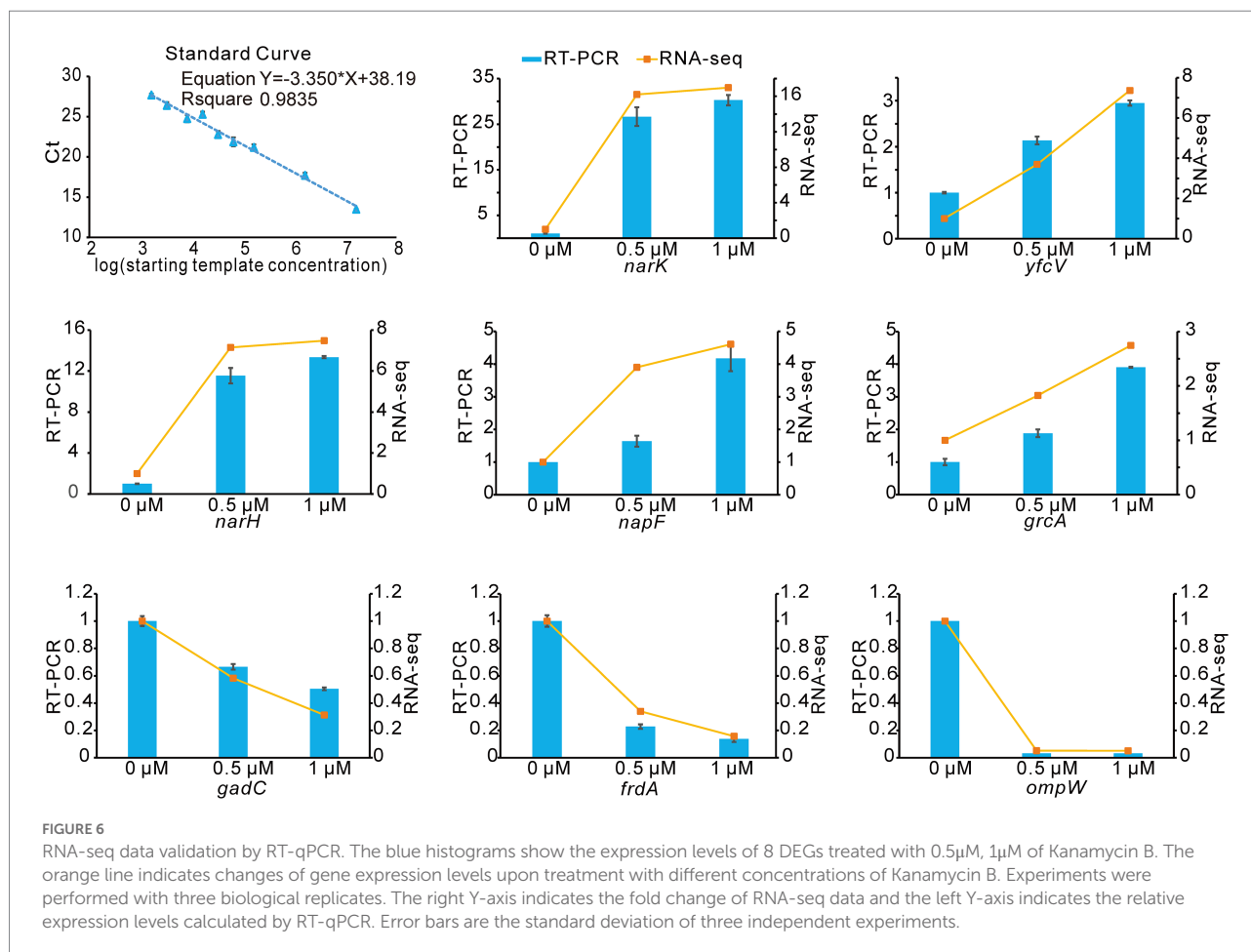
Discussion

Transcriptome analysis of *E. coli* JM109 treated with Kanamycin B uncovered gene expression networks that respond to Kanamycin B. We set out to investigate the extra non ribosomal inhibitory cellular (i.e., the "off target") function of Kanamycin B in *E. coli* through transcriptome data analysis ([Figures 1–6](#)). DEGs and STEM analysis of transcriptome data revealed that Kanamycin B treatment effects the expression of genes that are involved in a



variety of cellular functions including oxidative phosphorylation, nitrogen metabolism, microbial metabolism in diverse environments, biofilm formation and cell adhesion (Table 1; Figures 3–5). Kanamycin B binds to the 5' UTR of the *napF* or *narK* gene (one of the DEGs) and induces reporter gene expression (Figure 8; Supplementary Figure S3). The results provide insights into the cellular effects of Kanamycin B and are useful for establishing the non-antibiotic function of Kanamycin B.

We compare and contrast our study with the genome-wide transcriptome analysis of *E. coli* in response to gentamycin (Zhou et al., 2019). There are several differences between the two studies in terms of growth conditions, methods of treatment and concentration of aminoglycosides. *E. coli* cell was pretreated in the presence of 1 μg/ml gentamycin (1/2 MIC) for 1 h in MHB growth medium before harvesting for RNA extraction and RNA-seq. In comparison, 0.5 μM (1/8 MIC) or 1 μM (1/4 MIC) of Kanamycin

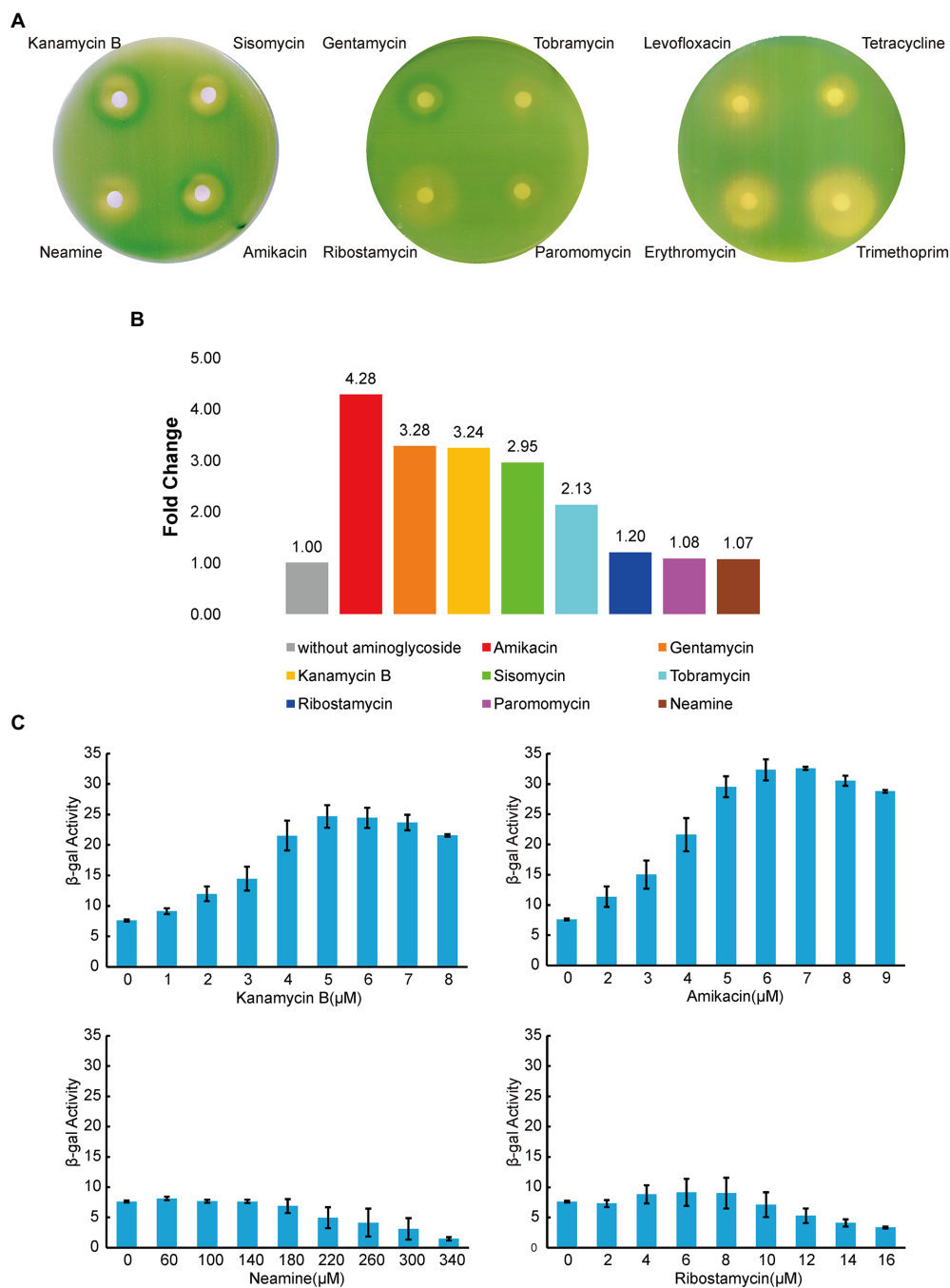


B was used in our study to treat *E. coli* cell for 7 h in LB medium prior to transcriptome analysis. These differences would expect to result in different DEGs in the two studies. However, further analysis showed some complementary findings. A high number of genes related to membrane protein and transporter functions were strongly regulated in response to gentamycin treatment. Similarly, our data (Supplementary Table S2) also showed that a considerable number of transmembrane transport genes were differentially expressed. These data from two independent studies suggest that membrane protein and transporter may have key roles in response to aminoglycosides in *E. coli*. In particular, *narK* is membrane protein and we showed biochemically that Kanamycin B may effect gene expression through binding to the 5' UTR of *narK* (Supplementary Figure S3). In addition, both studies have identified DEGs that are involved in biological process including ribosome and translation, TCA cycle, glycolysis or carbohydrate metabolism (Supplementary Tables S2, S6; Zhou et al., 2019).

Genome-wide transcriptome profiling of *Mycobacterium tuberculosis* upon treatment with Kanamycin has been reported (Habib et al., 2017). Comparison with the functional classification and pathways of the DEGs identified in the *M. tuberculosis* study with those identified in this study in *E. coli* indicated more differences than similarities between the two studies. Notably, the diverse responses

are probably due to the fundamental differences in the biology between *E. coli* and *M. tuberculosis*. However, other factors may also account for some of the different responses. The study in *M. tuberculosis* used Kanamycin instead of Kanamycin B which necessitates different growth conditions in comparison with our study.

It has been reported that the uptake of aminoglycosides into bacterial cells needs the proton motive force that is produced by electron flow through the respiratory chain of oxidative phosphorylation. The proton motive force is mainly generated by the respiratory complex I that contains membrane proteins and/or Fe-S clusters and oxidoreductases (Ezraty et al., 2013). The mechanism of aminoglycoside uptake and the bacterial cell response to the aminoglycosides is still unclear. Aminoglycoside binding riboswitches have been characterized (Jia et al., 2013) and a randomly selected Kanamycin B riboswitch has been reported (Kwon et al., 2001). The transcription levels of some genes associated with oxidative phosphorylation change upon treatment with Kanamycin B in this study. These observations together raise the possibility that Kanamycin B may regulate the expression with oxidative phosphorylation genes through riboregulatory interactions with the transcripts. Further studies will be required to examine this speculation.



In this study, treatment by Kanamycin B induces transcription of nitrate reductase (*narK*, *narG*, *narI*, *narH*, and *napA*) and nitrite reductase (*nirB*, *nirD*, and *nrfA*) genes (Table 1; Supplementary Table S4), which are key genes that are involved in nitrogen metabolism. These reductases

participate in the conversion of NO_3^- to NO_2^- to NO in cells and are associated with wider essential metabolic processes such as energy production, amino acid metabolism, biofilm formation, antibiotic resistance and bacterial pathogenesis. Nitrogen metabolism is also closely interlinked with biofilm

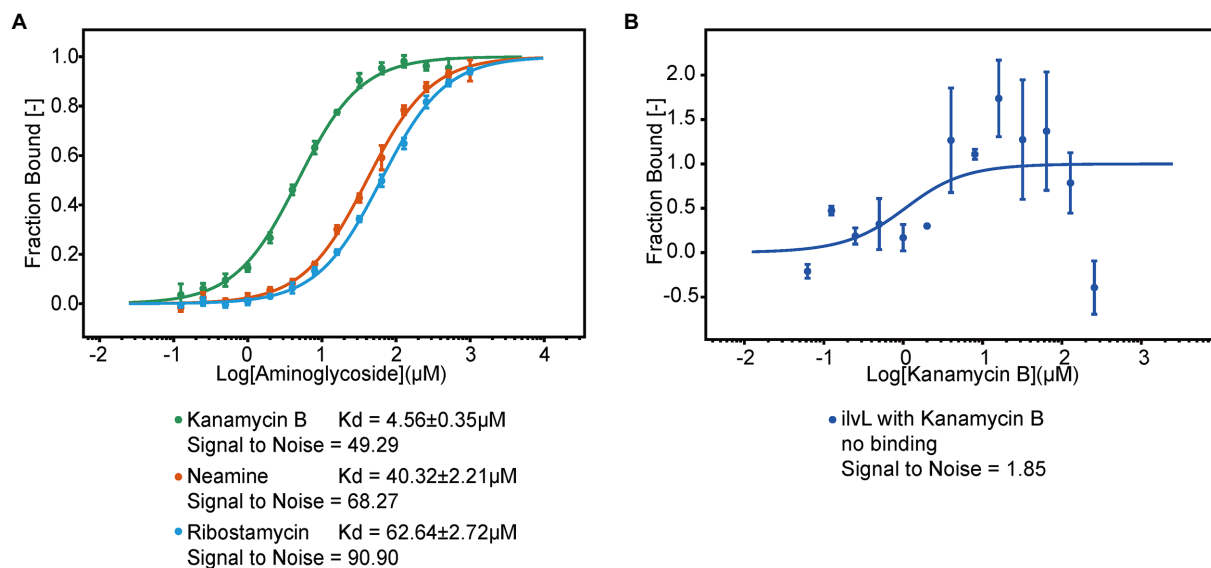


FIGURE 8

The MST binding of the 5' UTR of *napF* with Kanamycin B and other aminoglycosides. (A) Binding curve generated by MST for binding of aminoglycosides to the 5' UTR of *napF*. Error bars are standard deviations of at least three independent experiments. (B) Binding curve generated by MST for binding of Kanamycin B to the control RNA *ilvL*. Error bars are the standard deviation of three independent experiments.

formation that is also linked with the metabolism of glutamic acid, glutamine and arginine. It is widely recognized that NO plays an important role in modulating the architecture of biofilms (Vázquez-Torres and Bäumlér, 2016; Rinaldo et al., 2018). Other antibiotics have been reported to be able to induce biofilm formation (Linares et al., 2006). Therefore, it is possible that Kanamycin B may be linked to wider biological functions through its association with nitrogen metabolism.

We have further analyzed the identified DEGs by studying the association and effect of Kanamycin B and the UTR of the DEG gene on reporter gene expression. The result showed that Kanamycin B and other aminoglycosides induced reporter gene expression through the 5' UTR of the *napF* (involved in nitrogen metabolism) and further direct binding of Kanamycin B to the RNA was detected (Figures 7, 8; Supplementary Figure S1), suggesting that Kanamycin B could directly bind to the 5' UTR of the *napF* gene to effect expression of the *napF* gene in *E. coli*. Although how the binding of Kanamycin B to the 5' UTR of the *napF* gene regulate gene expression remains unclear, the result of this finding revealed that Kanamycin B can bind to the UTR RNA of a gene and effect the expression of its downstream gene. In this case, Kanamycin B may effect cellular nitrogen metabolism through binding to the 5' UTR of the *napF* gene and regulate expression of the gene. Taken together, this study has collectively revealed new insights into the non-antibiotic function of Kanamycin B and its wider cellular effect on *E. coli*.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

AM and DC wrote the main manuscript text. YC and XZ prepared Figures 1–8; Table 1; Supplementary Tables 1–6; Supplementary Figures S1–S3. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Natural Science Foundation grants 32271345 and 32250710140 to AM.

Acknowledgments

We thank members of the Murchie laboratories for discussion.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.937827/full#supplementary-material>

References

- Anders, S., Pyl, P., and Huber, W. (2015). HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Baharoglu, Z., and Mazel, D. (2011). *Vibrio cholerae* triggers SOS and mutagenesis in response to a wide range of antibiotics: a route towards multi-resistance. *Antimicrob. Agents Chemother.* 55, 2438–2441. doi: 10.1128/aac.01549-10
- Bailey, M., Chettiath, T., and Mankin, A. S. (2008). Induction of erm (C) expression by noninducing antibiotics. *Antimicrob. Agents Chemother.* 52, 866–874. doi: 10.1128/AAC.01266-07
- Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., and Ramakrishnan, V. (2000). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407, 340–348. doi: 10.1038/35030019
- Chen, D., and Murchie, A. (2014). An aminoglycoside sensing riboswitch controls the expression of aminoglycoside resistance acetyltransferase and adenylyltransferases. *Biochim. Biophys. Acta* 1839, 951–958. doi: 10.1016/j.bbagg.2014.02.019
- Cianciulli Sesso, A., Lilić, B., Amman, F., Wolfinger, M., Sonnleitner, E., and Bläsi, U. (2021). *Pseudomonas aeruginosa* gene expression profiling of upon exposure to Colistin and tobramycin. *Front. Microbiol.* 12:626715. doi: 10.3389/fmicb.2021.626715
- Coleman, S., Blimkie, T., Falsafi, R., and Hancock, R. (2020). Multidrug adaptive resistance of *Pseudomonas aeruginosa* swarming cells. *Antimicrob. Agents Chemother.* 64, e01999–19. doi: 10.1128/aac.01999-19
- Davies, J., and Davis, B. (1968). Misreading of ribonucleic acid code words induced by aminoglycoside antibiotics. The effect of drug concentration. *J. Biol. Chem.* 243, 3312–3316. doi: 10.1016/S0021-9258(18)93308-9
- Entzian, C., and Schubert, T. (2016). Studying small molecule–aptamer interactions using micro scale thermophoresis (MST). *Methods* 97, 27–34. doi: 10.1016/j.ymeth.2015.08.023
- Ernst, J., and Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* 7:191. doi: 10.1186/1471-2105-7-191
- Ezraty, B., Vergnes, A., Banzhaf, M., Duverger, Y., Huguenot, A., Brochado, A. R., et al. (2013). Fe-S cluster biosynthesis controls uptake of aminoglycosides in a ROS-less death pathway. *Science* 340, 1583–1587. doi: 10.1126/science.1238328
- Fourmy, D., Recht, M., Blanchard, S., and Puglisi, J. (1996). Structure of the site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* 274, 1367–1371. doi: 10.1126/science.274.5291.1367
- Habib, Z., Xu, W., Jamal, M., Rehman, K., and Gang, C. (2017). Adaptive gene profiling of mycobacterium tuberculosis during sub-lethal kanamycin exposure. *Microb. Pathog.* 112, 243–253. doi: 10.1016/j.micpath.2017.09.055
- He, W., Zhang, X., Zhang, J., Jia, X., Zhang, J., Sun, W., et al. (2013). Riboswitch control of induction of aminoglycoside resistance acetyl and adenylyl-transferases. *RNA Biol.* 10, 1266–1273. doi: 10.4161/rna.25757
- Hoffman, L., D'Argenio, D., Maccoss, M., Zhang, Z., Jones, R., and Miller, S. (2005). Aminoglycoside antibiotics induce bacterial biofilm formation. *Nature* 436, 1171–1175. doi: 10.1038/nature03912
- Jia, X., Zhang, J., Sun, W., He, W., Jiang, H., Chen, D., et al. (2013). Riboswitch control of aminoglycoside antibiotic resistance. *Cell* 152, 68–81. doi: 10.1016/j.cell.2012.12.019
- Khan, F., Lee, J., Pham, D., Lee, J., Kim, H., Kim, Y., et al. (2020). Streptomycin mediated biofilm inhibition and suppression of virulence properties in *Pseudomonas aeruginosa* PAO1. *Appl. Microbiol. Biotechnol.* 104, 799–816. doi: 10.1007/s00253-019-10190-w
- Khan, S., Wu, S. B., and Roberts, J. (2019). RNA-sequencing analysis of shell gland shows differences in gene expression profile at two time-points of eggshell formation in laying chickens. *BMC Genomics* 20:89. doi: 10.1186/s12864-019-5460-4
- Krause, K., Serio, A., Kane, T., and Connolly, L. (2016). Aminoglycosides: an overview. *Cold Spring Harb. Perspect. Med.* 6. doi: 10.1101/cshperspect.a027029
- Kwon, M., Chun, S., Jeong, S., and Yu, J. (2001). In vitro selection of RNA against kanamycin B. *Mol. Cells* 11, 303–311. doi: 10.1002/yea.750
- Linares, J., Gustafsson, I., Baquero, F., and Martinez, J. (2006). Antibiotics as intermicrobial signaling agents instead of weapons. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19484–19489. doi: 10.1073/pnas.0608949103
- Long, J., Zhang, J., Zhang, X., Wu, J., and Du, C. (2020). Genetic diversity of common bean (*Phaseolus vulgaris* L.) germplasm resources in Chongqing, evidenced by morphological characterization. *Front. Genet.* 11:697. doi: 10.3389/fgenet.2020.00697
- Lovett, P. S., and Rogers, E. J. (1996). Ribosome regulation by the nascent peptide. *Microbiol. Rev.* 60, 366–385. doi: 10.1128/mr.60.2.366-385.1996
- Mingeot-Leclercq, M. P., Glupczynski, Y., and Tulkens, P. M. (1999). Aminoglycosides: activity and resistance. *Antimicrob. Agents Chemother.* 43, 727–737. doi: 10.1128/AAC.43.4.727
- Moon, M., Hilimire, T., Sanders, A., and Schneekloth, J. (2018). Measuring RNA-ligand interactions with Microscale thermophoresis. *Biochemistry* 57, 4638–4643. doi: 10.1021/acs.biochem.7b01141
- Nikaido, H. (2009). Multidrug resistance in bacteria. *Annu. Rev. Biochem.* 78, 119–146. doi: 10.1146/annurev.biochem.78.082907.145923
- Rinaldo, S., Giardina, G., Mantoni, F., Paone, A., and Cutruzzola, F. (2018). Beyond nitrogen metabolism: nitric oxide, cyclic-di-GMP and bacterial biofilms. *FEMS Microbiol. Lett.* 365. doi: 10.1093/femsle/fny029
- Robinson, M., McCarthy, D., and Smyth, G. (2010). Edge R: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Stewart, V., Bledsoe, P. J., and Williams, S. B. (2003). Dual overlapping promoters control nap F (periplasmic nitrate reductase) operon expression in *Escherichia coli* K-12. *J. Bacteriol.* 185, 5862–5870. doi: 10.1128/JB.185.19.5862-5870.2003
- Vázquez-Torres, A., and Bäuml, A. (2016). Nitrate, nitrite and nitric oxide reductases: from the last universal common ancestor to modern bacterial pathogens. *Curr. Opin. Microbiol.* 29, 1–8. doi: 10.1016/j.mib.2015.09.002
- Wang, S., He, W., Sun, W., Zhang, J., Chang, Y., Chen, D., et al. (2019). Integron-derived aminoglycoside-sensing riboswitches control aminoglycoside acetyltransferase resistance gene expression. *Antimicrob. Agents Chemother.* 63, e00236–19. doi: 10.1128/aac.00236-19
- Wu, T., Ruan, K., and Liu, W. (1996). A fluorescence-labeling method for sequencing small RNA on polyacrylamide gel. *Nucleic Acids Res.* 24, 3472–3473. doi: 10.1093/nar/24.17.3472
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). Cluster profiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, X., and Bremer, H. (1995). Control of the *Escherichia coli* rrn B P 1 promoter strength by pp Gpp. *J. Biol. Chem.* 270, 11181–11189. doi: 10.1074/jbc.270.19.11181
- Zhang, J., Liu, G., Zhang, X., Chang, Y., Wang, S., He, W., et al. (2020). Aminoglycoside riboswitch control of the expression of integron associated aminoglycoside resistance adenylyltransferases. *Virulence* 11, 1432–1442. doi: 10.1080/21505594.2020.1836910

Zhou, S., Zhuang, Y., Zhu, X., Yao, F., Li, H., Li, H., et al. (2019). Yhj X regulates the growth of in the presence of a subinhibitory concentration of gentamicin and mediates the adaptive resistance to gentamicin. *Front. Microbiol.* 10:1180. doi: 10.3389/fmicb.2019.01180

Zhuang, Y., Chen, W., Yao, F., Huang, Y., Zhou, S., Li, H., et al. (2016). Short-term pretreatment of sub-inhibitory concentrations of gentamycin inhibits the swarming motility of *Escherichia Coli* by Down-regulating the succinate dehydrogenase gene. *Cell. Physiol. Biochem.* 39, 1307–1316. doi: 10.1159/000447835

Frontiers in Microbiology

Explores the habitable world and the potential of microbial life

The largest and most cited microbiology journal which advances our understanding of the role microbes play in addressing global challenges such as healthcare, food security, and climate change.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

