INVESTIGATION OF WHOLE SYSTEM DETERMINANTS OF ANTIBIOTIC RESISTANCE IN *PSEUDOMONAS AERUGINOSA*

A Thesis Submitted for the Degree of Doctor of Philosophy (PhD)

By

Wedad Mohamed Nageeb Ahmed Saleh

College of Health and Life Sciences, Department of Biomedical Sciences

Brunel University London

2020

THANKS TO ALLAH, THE MOST GLORIOUS, THE MOST MERCIFUL

I HAVE BEEN ABLE TO COMPLETE THIS WORK

THIS WORK IS DEDICATED TO MY LITTLE CHILDREN

YOU ARE A WALKING PART OF MY SOUL

AND YOU MAKE A MEANING FOR EVERYTHING I DO IN MY LIFE

CONTENTS

ACKNOWLEDGEMNT

I could not have completed my PhD without the guidance, attention and support of my principal supervisor professor Nigel Saunders. I am really grateful for him for helping me to become a more mature and independent researcher and to become a better scientist. I would like to thank him for providing me with the resources needed to pursue my PhD and for showing enthusiasm in my work. Next comes my second supervisor dr. Christian Rudolph who was ready to help at any time with his office door open and ready to provide any kind of help and support. I am really so much thankful and grateful to him. I also need to express my deepest thanks to dr. Justin O'grady for taking time to read and examine my thesis and to provide an objective and useful discussion and guidance which helped me to consolidate my thinking and to organize the findings in the best way. I thank him for his extensive notes, comments and suggestions which really helped me to make the thesis much better. So, my great thanks to him for active listening, understanding and advice about how to make the work better. I would also like to thank dr. Ronan McCarthy for his advice, his useful discussion and suggestions and for his guidance about the thesis presentation.

It has been a long journey with some challenges, and I could not go through without the constant help, support, love and encouragement of my entire family. Though stressful, these challenges have helped me to become the best I can and to achieve the best results. Special thanks go to my parents who have usually encouraged me all over my life to stick to and to achieve my goals. They were there all the time to keep me going. Thanks for everything you did for me and thanks for your unconditional love and support.

I would like also to thank Brunel university including all the Systems and Synthetic Biology lab members, the Department of Health and Life Sciences members, the graduate school team, the services and the technical team and every person I have met during my journey at Brunel. I have really enjoyed my time at Brunel and have learnt a lot for my entire life and for my future career.

ABSTRACT

Pseudomonas aeruginosa has been declared as one of the "six top priority dangerous microbes" and has also been classified as one of the six ESKAPE organisms with emerging clinical importance. The organism possesses a large plastic genome which is considered the base for its high physiologic diversity and high metabolic adaptability. *Ps. aeruginosa* shows extreme adaptability to colonize different habitats including hospital environments.

Although equipped with extensive intrinsic resistance machinery that leads to basal levels of lower susceptibility to many antibiotics, a complete understanding of core resistance mechanisms is considered challenging. High-throughput methods that explore for system-level resistance including transcriptional profiling, mutant library screening, and experimental evolution have many technical drawbacks which make them unreliable to predict clinical resistance.

In the thesis, a novel strategy has been adopted to mine for the system level antibiotic susceptibility determinants using a sequence-based integrated genomics approach that combines cluster analysis, predictive modelling, and comparative behavioral genomics. Combining the existing body of knowledge about resistance-associated variants and results of comparative behavioral genomics has resulted into a new way to understand resistance and to prioritize system elements contributing to resistance. This new knowledge is expected to offer promising diagnostic and therapeutic potentials.

The approach has interrogated the whole physiologic system resulting in the identification of a new group of candidate resistance predictor markers and new combinations showing up to 90% better performance. The approach has also created a new understanding about the combinatorial quantitative contribution of different resistance mechanism to quinolones and aminoglycoside groups of antibiotics focusing on gentamycin, amikacin, ciprofloxacin, and levofloxacin. It has also highlighted some novel functional groups and genes that contribute to resistance. This knowledge could offer improved genome-based directed antibiotic treatment.

Recent advances in sequencing technologies are expected to provide a rich information resource and a promising diagnostic platform. The new knowledge is capable of providing a base for rapid point-of-care antibiotic resistance diagnostic platforms, thus increasing the spectrum and informative value of some diagnostic panels in current use.

PRESENTATIONS

Poster Presentation, "Investigation of whole system determinants of antibiotic resistance in *Pseudomonas aeruginosa*," 3rd Annual CHLS Research Conference, 14th Dec 2017, Brunel University London.

Poster Presentation, "New Understanding of Quinolone Resistance in *Pseudomonas aeruginosa* (core physiology rather than mere drug target changes)," Microbiology Society Annual Conference, 11-13 April 2018, Birmingham, UK.

Poster Presentation, "Comparative behavioral genomics reveals potential diagnostic genomic markers associated with *Pseudomonas aeruginosa* sensitivity to aminoglycosides and quinolones group of antibiotics.", Antimicrobial Resistance – Genomes, Big Data and Emerging Technologies conference, 27-29 November 2018, Wellcome Genome Campus, Hinxton, Cambridgeshire, UK.

Oral Presentation, "Using genome based diagnostic clinical microbiology to improve health outcomes"**,** 4th Annual CHLS Research Conference, 6th Dec 2018, Brunel University London.

LIST OF ABBREVIATIONS

AES: Australian Epidemic Strain

AFLP: Amplified Fragment Length Polymorphism

AGs: Aminoglycosides

AMEs: Aminoglycosides Modifying Enzymes

AMR: Antimicrobial Resistance

ARDB: Antibiotic Resistance Database

AST: Antimicrobial susceptibility

BLAST: Basic Local Alignment Search Tool

CAMHB: Cation Adjusted Muller Hinton Broth

CBG: Comparative Behavioral Genomics

CC: Clonal Complex

CFPA: Cystic Fibrosis Pseudomonas Aeruginosa

CFU: Colony Forming Units

cgMLST: Core Genome Multi Locus Sequence Typing

DOR: Diagnostic Odds Ratio

ESBLs: Extended Spectrum Beta Lactamases

FMN: Flavin Mononucleotide

GLN: Glutamine

GLT: Glutamate

GWAS: Genome Wide Association Studies

ICUs: Intensive Care Units

ID/AST: Identification/Antimicrobial Susceptibility

LES: Liverpool Epidemic Strain

LPS: Lipopolysaccharides

LR: Likelihood Ratio

MALDI-TOF-MS: Matrix-assisted Laser desorption ionization-time of flight mass spectrometry

MDR: Multi Drug Resistance

MIC: Minimum Inhibitory Concentration

MLEE: Multi Locus Enzyme Electrophoresis

MLST: Multi-locus Sequence Typing

MLVA: Multi-Locus Variable Number Tandem Repeats Analysis

MRSA: Methicillin-resistance staphylococcus aureus

MST: Minimum Spanning Tree

NPV: Negative Predictive Value

nSSNPs: non-synonymous Single Nucleotide Polymorphism

OD: Optical Density

OSA: O-specific antigen

PCR: Polymerase Chain Reaction

PFGE: Pulsed Field Gel Electrophoresis

PMQR: Plasmid Mediated Quinolone Resistance

POC: Point of Care

PPV: Positive Predictive Value

PQS: Pseudomonas Quinolone Signal

QRDR: Quinolone Resistance Determining Region

RAPD: Random Amplified Polymorphic DNA

RFLP: Random Fragment Length Polymorphism

RGI: Resistance Gene Identifier

ROS: Reactive Oxygen Species

SNP: Single Nucleotide Polymorphism

ST: Sequence Type

TCA: Tricarboxylic Acid

VNTR: Variable Number of Tandem Repeats

wgMLST: Whole Genome Multi Locus Sequence Typing

WGS: Whole Genome Sequencing

XDR: Extensive Drug Resistance

LIST OF TABLES

LIST OF FIGURES

APPENDIX I

APPENDIX II

Chapter 1. Research Introduction and Background

1.1. Antibiotic resistance

1.2. Social, economic and health-care related impact of antibiotic resistance

1.3. Pipeline of new drug introduction and emergence of resistance

1.4. Approaches and initiatives to address Antibiotic Resistance

1.4.1. Some WHO initiatives

1.4.2. The CDC report on antimicrobial resistance threat in the United States

1.4.3. The state of the world's antibiotic report by the center for disease dynamics and economics

1.4.4. The review on antimicrobial resistance chaired by Jim O'Neill

1.5. Mechanisms and origins of bacterial resistance to antibiotics

1.6. Why it is important to understand intrinsic antimicrobial resistance?

1.7. *Pseudomonas aeruginosa*

1.7.1. An environmental pathogen with intrinsic resistance

1.7.2. *Pseudomonas aeruginosa*: The clinical pathogen

1.7.3. The challenging nature of *Ps. aeruginosa* species

1.7.3.1. The genome of *Ps. aeruginosa*

1.7.3.2. The extensive metabolic adaptability in *Ps. aeruginosa*

1.8. Different approaches previously used to understand intrinsic antimicrobial resistance

1.8.1. Transcriptional profiling

1.8.2. Mutant library screening

1.8.3. Experimental evolution

1.9. Previous investigation of quinolone resistance in *Ps. aeruginosa*

1.9.1. Mechanism of quinolones action

1.9.2. Mechanisms and Determinants of quinolones resistance and the importance of understanding the global picture

1.9.3. Agents studied from the quinolone group

1.10. Previous investigation of aminoglycoside resistance in *Ps. aeruginosa*

1.10.1. Mechanisms of aminoglycoside action

1.10.2. Molecular mechanisms and system level perspective of aminoglycosides resistance

1.10.3. Agents studied from the aminoglycoside group

1.11. Molecular platforms used for identifying antibiotic resistance markers

1.12. Practical technologic advances addressing some sequence-based diagnostics limitations

1.13. Project

1.13.1. Significance, justification, and outline

1.13.2. Study aim and objectives

1.1. Antibiotic Resistance

Bacterial resistance to antibiotics is an ancient feature in some bacterial populations. Environmental bacteria living in soil or other natural reservoirs with millions of other species must have evolved many components of resistance to naturally occurring antibiotics a long time before antimicrobials were introduced in clinical practice. These bacteria, especially when they are producers of antimicrobial compounds, have acquired or evolved resistance determinants for self-protection (Wright and Poinar, 2012).

There is a growing interest in exploiting the pool of genes composing bacterial pan-genomes for the identification of novel functional determinants underlying antibiotic resistance. Novel functional resistance may arise from proto-resistance genes or silent resistance genes through mutations, changing gene expression, changing the degree of mobility, or changing allele frequency under modern-day selective pressure (Perry, Westman and Wright, 2014). The hypothesis of resistance arising from changes in the gene pool can be supported by several sources of evidence. One example is the ancient DNA analysis collected from the 30,000 years old Beringian Permafrost showing resistance genes for beta lactams, tetracycline, and glycopeptide antibiotics (D'Costa *et al.*, 2011). Another example is the culturable four million years old isolates from isolated cave environments that showed phenotypic resistance to 14 different antibiotics. Diverse resistance genes including novel ones have been identified in these isolates (Bhullar *et al.*, 2012). Thorough investigation of the basic biology underlying evolution and spread of resistance in bacterial populations is fundamental for identifying determinants that may represent predictive potential for clinical resistance.

The current problem of antibiotic resistance appears to be affected by two forces: first is the natural ecologic biologic or environmental part, second is the human-derived clinical force. The selective pressure exerted by human use of antimicrobials can originate from different sources. These include antibiotic use in medicine, its use in food, animal or fish production, and in agriculture. Compared to the natural ecologic force, clinical selection resulting from human use of antibiotics and antimicrobial compounds is considered much more recent (Baquero, Alvarez-Ortega and Martinez, 2009).

Since the golden era of antibiotic discovery of penicillin in 1929 to the current post-antibiotic era, antibiotic drug discovery and development have passed through several stages (Lewis, 2013). The first antimicrobials used preceding the golden era of penicillin relied upon synthetic molecule discovery. This included screening libraries of chemicals for selective antibacterial activity and these efforts have resulted in using salvarsan and the 'sulfa' drugs. The golden era then followed with penicillin discovery, and during that era, the discovery of natural products and whole-cell screens showed high success (Lewis, 2013).

By the mid-sixties, the era of medicinal chemistry came about to form the next stage of antibiotic discovery and innovation. This involved the creation of synthetic versions of the preceding natural scaffolds with the aim of lowering the therapeutic dose or increasing the spectrum of activity. The approach has achieved success through improving the properties of previous compounds and fighting resistance. Although some efforts are currently directed to new agent development, this is considered challenging and is currently facing many hurdles in respect to experimentation, clinical trials and approval. In addition, resistance usually arises and rapidly spreads to most newly developed agents. This can be observed by reviewing most previous agents in the market. Reviewing the timeline of resistance appearance shows that resistance arises shortly after the introduction of each new drug (Lewis, 2013).

Consequently, extending the life of available antimicrobial agents can be considered the easiest approach to tackle the problem. This can be achieved through improving the prescription of the currently available antibiotics which requires giving the right antibiotic to the right patient in the right dose (Liu *et al.*, 2019). In most clinical settings, physicians have long been unaware about the concept of proper prescription especially in patients with critical conditions and those in the intensive care units (ICUs) (Teixeira Rodrigues *et al.*, 2013). They depend on the empiric use of broad-spectrum antibiotics being considered as "the most powerful option that works for all suspected infections". This results in over-prescription which makes the problem more complicated because such a practice would select for resistant strains that otherwise could have remained susceptible. Another example of poor clinical practice is the improper prescription of last line agents or 'reserve agents', a practice that would deplete the antimicrobial reserve needed for proper future and necessary use (García *et al.*, 2011). Data from the CDC Newsroom show that 1 in 3 antibiotic prescriptions is unnecessary and that at least 30% of antibiotics prescribed

in the united states are unnecessary with goals set to reduce unnecessary use by 2020 according to Fleming-Dutra *et al.*, (2016). However, the problem is still of rising concern according to the latest antibiotic resistance CDC report (US CDC, 2019). A recent study from China has also shown that the problem of inappropriate antibiotic use is still existent in pediatrics wards including over-prescription and overuse of parenteral administration (Miao *et al.*, 2020). Another recent study showed that nearly 45% of antibiotics prescribed over a 10-year period were given out for no clear clinical reason (Fischer *et al.*, 2020). Decreasing the widespread superfluous overuse of antibiotics requires an exhaustive policy. This should include improving stewardship programs and regulations in addition to discovering and optimizing well-functioning rapid diagnostics suited for different types of infections.

1.2. Social, economic, and health-care related impact of antibiotic resistance

The WHO report showed that the yearly cost of antimicrobial resistance problem to the US health system alone has been estimated at $21 to $34 billion, with the additional cost of more than 8 million additional days of hospital stay (WHO, 2016). The review on antimicrobial resistance (2014) by Jim O'Neil showed that antibiotic-resistant infections cause at least 50,000 deaths per year in Europe and the US alone, with many hundreds of thousands in other areas of the world which have not been accurately reported. The continued rise in resistance is predicted to kill 10 million people every year by the year 2050. Three hundred million people are expected to die prematurely in the following thirty-five years if the current problem of resistance is not handled effectively (O' Neil, 2014).

It is also estimated that there will be around 60 to 100 trillion USD economic losses if the antimicrobial resistance problem is not tackled effectively. Also, deaths attributable to antimicrobial resistance are estimated to be around 390,000 in Europe, 317,000 in North America, 4,150,000 in Africa and about 4,730,000 in Asia by the year 2050 (O' Neil, 2014). National summary data from the CDC report 2013 in the US have also shown that antibiotic resistance causes at least 2,049,442 illnesses and 23,000 deaths in the US (Centers for Disease Control and Prevention, 2013).

The antibiotic resistance problem is affecting all the modern areas of medicine including cancer chemotherapy, end-stage renal disease patients on dialysis, bone marrow and organ transplant patients, and all types of surgeries. Cancer patients on chemotherapy being immunocompromised

are more susceptible to serious infections and the lack of proper treatment may lead to serious complications and death. Besides, complex surgeries including cardiac bypass, joint replacements, and other major surgeries may need pre-operative antibiotic prophylaxis in some cases or may be complicated with surgical site infections compromising the whole process and leading to post-surgical failure as a result of the infections caused by resistant pathogens. End-stage renal disease patients on dialysis are also more susceptible to serious bloodstream infections. The success of organ or bone marrow transplants also depends greatly on the ability to effectively treat resistant bacterial infections.

1.3. Pipeline of new drug introduction and emergence of resistance

It has been predicted since the discovery of penicillin that resistance will arise to antibiotics as a normal biologic evolutionary mechanism. However, this has been much accelerated through the selective pressure exerted by clinical and environmental misuse of these drugs. Alexander Fleming warned that bacteria could become resistant to that remarkable discovery in his Nobel prize speech in 1945 (Ventola, 2015).

This has later become evident through the timeline of resistance development following the introduction of each antibiotic class into clinical practice. We are now facing an antibiotic discovery void as reported by Fischbach and Walsh, (2009) who showed that no major classes of antibiotics were introduced between 1962 and 2000 referring to this as 'innovation gap' (Fischbach and Walsh, 2009). Except for carbapenem introduction in 1985, all antibiotics introduced between the 1960s and 2000 were synthetic derivatives of the commonly used scaffolds (or core structure). Developing antimicrobial agents was based on few scaffolds with many generations (Fischbach and Walsh, 2009). Even some of the more recently approved drugs show some limitations and are developing resistance or cross resistance to some previous agents. This consequently necessitates directing efforts of new agent development towards the discovery of new classes rather than to chemically modify existing ones. This can be achieved through mining unexplored natural products and through screening for new synthetic molecules. However, this approach is not considered the most effective and does not offer the most rapid or the best practical solutions.

1.4. Approaches and initiatives to address antibiotic resistance (WHO, 2016)

1.4.1. Some WHO Initiatives

In the 2016 WHO global surveillance report, a strategy was proposed to achieve containment of antimicrobial resistance. The strategy includes:

1) Prioritizations of diseases that need intervention and that included bacterial infections other than tuberculosis.

2) Identification of core sets of interventions for national implementation.

3) Effective implementation including a clear action plan, authority delegation, resource assignment with clear mechanisms for intervention and monitoring of the outcome.

4) Commitment to an all-inclusive national plan with the responsibility and the engagement of civil society and this includes:

- Establishment and consolidation of surveillance and laboratory capacity.
- Ensuring the ongoing access to necessary medicines of guaranteed quality.
- Ruling and advancement for the rational use of medicines, including its use in animal husbandry, and safeguarding appropriate patient care.
- Enhancing infection prevention and control.
- Nurturing innovation in research and development.

5) Improving attentiveness and understanding of antimicrobial resistance through operative communication, education, and training.

6) Solidification of knowledge and evidence base through surveillance and research.

7) Reducing the incidence of infection through effective sanitation, hygiene, and infection prevention measures.

8) Improving the use of antimicrobial medicines in human and animal health.

9) Mounting the economic case for maintainable investment that takes account of the needs of all countries and that increase investment in new medicines, diagnostic tools, vaccines, and other interventions.

1.4.2. The CDC report on antimicrobial resistance threat in the United States (2013)

This report has recommended four actions for the problem including:

1) Preventing infection to prevent the spread of resistance which includes following the general and specific prophylaxis including general hygiene, hand washing, safe food preparation, and immunization. This aims at decreasing the incidence of infection and consequently decreasing antibiotic use and development of resistance.

2) Tracking and surveillance. This includes data gathering which can then be used to guide and develop relevant strategies.

3) Improving antibiotic prescription and stewardship which includes stopping the inappropriate and unnecessary use of antibiotics, a step listed as the most important in managing the problem.

4) Developing new drugs and diagnostic tests.

(Centers for Disease Control and Prevention, 2013) (Centers for Disease Control and Prevention, 2018).

1.4.3. The state of the world's antibiotic report by the center for disease dynamics and economics

This report has suggested six strategies for successful national policies which include:

1) Reducing the need for antibiotics by reducing the burden of infectious diseases requiring antibiotic. This can be achieved through improving vaccination coverage, improving access to clean water, safe and healthful food supply and improving sanitation in order to improve the health and well-being of people in general.

2) Improving hospital infection control and antimicrobial stewardship programs.

3) Changing the incentives that encourage antibiotics overuse and misuse which mainly affect hospitals, health care providers and pharmacists.

4) Reducing and abolishing antibiotic use in agriculture and aquaculture.

5) Educating and informing health care professionals and policy makers about rational antibiotic use.

6) Ensuring political commitment to address antibiotic resistance by generating local interest and pressure of the health care professionals and the public so that policy makers can allocate time and resources to design and implement strategies that promote rational use of antibiotics.

(Hellen Gelband et al., 2015).

1.4.4. The review on antimicrobial resistance chaired by Jim O'Neill (May 2016)

This final report has highlighted 10 important approaches towards solving the problem which include:

1. A massive global public awareness campaign.

2. Improving hygiene and preventing the spread of infection.

3. Abolishing the needless use of antimicrobials in agriculture and their dissemination into the environment.

4. Improving universal surveillance of drug resistance and antimicrobial utilization in humans and animals.

5. Promoting new, rapid diagnostics to cut unnecessary use of antibiotics.

6. Promoting the development and use of vaccines and alternatives.

7. Improving the quantitative and qualitative capacity of people working in infectious disease.

8. Launching a Global Innovation Fund for early-stage and non-commercial research.

9. Better motivations to promote developing new drugs and to advance existing ones.

10. Building a global coalition for real action. (O'Neill, 2016).

1.5. Mechanisms and origins of bacterial resistance to antibiotics

Bacterial resistance to antibiotics can be classified into intrinsic or innate resistance and acquired resistance. Intrinsic resistance means that the bacteria have certain inherent structural or functional characteristics that lead to resistance to a single agent or group of agents irrespective of prior antibiotic exposure. On the other hand, acquired resistance can originate from the acquisition of resistance genes or from chromosomal DNA mutations or from both effects (Blair *et al.*, 2015).

Acquisition of resistance genes is mediated by transferable genetic elements including plasmids, transposons, and integrons. Although mutations in bacteria are rare, one in the population of 10^6 to 10^8 microorganisms, these spontaneous mutations, resulting from errors in chromosomal replications or incorrect repair of damaged DNA, can be responsible for clinically functional resistance to antibiotics.

Bacteria can acquire resistance to antibiotics through five basic mechanisms. The first two mechanisms result in decreasing the intracellular concentration of the antibiotic either by decreasing antibiotic penetration inside cells preventing antibiotic access to its targets or by

increasing efflux of the drug outside the cell. Efflux pumps are a group of transporter proteins with several families responsible for pumping out drugs or toxic materials outside the bacterial cell and they vary widely in their substrate specificity (Webber and Piddock, 2003). Other resistance-underlying mechanisms include protection or modification of antibiotic targets, mutations changing antibiotic targets and direct inactivation of antibiotics through hydrolysis or modification (Blair *et al.*, 2015).

1.6. Why it is important to understand intrinsic antimicrobial resistance?

Although antibiotic resistance is traditionally considered an acquired trait in previously susceptible bacteria, it is crucial to differentiate species-specific intrinsic bacterial resistome from acquired resistance genes or mutations in formerly susceptible bacteria. The intrinsic resistome is defined as the group of chromosomal genes involved in intrinsic resistance which presence is independent of previous antibiotic exposure and is not due to horizontal gene transfer (Fajardo *et al.*, 2008).

Being characterized by a large degree of non-specificity, investigating intrinsic resistance is considered challenging especially in *Ps. aeruginosa*. In addition to the specific well-known mechanisms of resistance, intrinsically resistant bacterial species carry what is called "physiology-dependent" resistome (Fajardo *et al.*, 2008). The greater part of research studies on antibiotic resistance lacks clarity of differentiation between intrinsic resistance and acquired resistance. The contribution of intrinsic system properties to overall resistance is usually not explored. Although important and frequently well-studied, acquired resistance cannot be considered an original or a stable robustly predicted character to be considered as an essential contributor to species behavior. A reason why, the thesis has focused on the assessment of intrinsic chromosomal resistance mechanisms rather than acquired mechanisms of resistance.

1.7. *Pseudomonas aeruginosa*

1.7.1. An environmental pathogen with intrinsic resistance

Ps. aeruginosa is a prominent example of intrinsically resistant bacteria with an environmental origin. The physiologic elements underlying resistance are necessary for the fitness of the species in their natural habitat and can be considered an evolutionary ancient phenotype (Alonso, Sánchez and Martínez, 2001). Intrinsic resistance in soil-dwelling bacteria is an ancient

component that predates the clinical use of antibiotics. An example of that are strains of *Ps. aeruginosa* isolated from non-clinical environments dating back before the discovery of the synthetic quinolones group of antibiotics and exhibiting resistance to this group (Alonso, Rojo and Martínez, 1999). This property becomes problematic when transmitted to the clinic. In addition to being pathogenic, they carry additional potential to transmit some resistance-related elements to other clinically important bacterial species (Forsberg *et al.*, 2012). Opportunistic pathogens may show similar characters of pathogenicity, multidrug resistance, and substrate utilization that are similarly observed in both environmental isolates and in clinical isolates (Alonso, Rojo and Martínez, 1999). Strains of *Ps. aeruginosa* from both clinical and environmental origins have shown to be similar at the genomic level in different studies (Wiehlmann *et al.*, 2007). A study investigating the intrinsic resistome in *Ps. aeruginosa* concluded that intrinsic resistance phenotype requires the combined action of several genes encoding basic functions in cell physiology and that the capacity of this organism to acquire higher levels of resistance can result from complex network interactions among several bacterial proteins (Fajardo *et al.*, 2008).

Although the resistome of *Ps. aeruginosa* has been investigated in different studies, there is a lack of satisfactory explanation and complete understanding about system-level determinants of resistance. The research done to date on *Ps. aeruginosa* resistance determinants has not been able to determine the relative importance or the quantitative contribution of different chromosomal mechanisms of resistance. Our knowledge about the chromosomal basis of resistance is still growing. It can significantly advance with the increasing availability of the huge numbers of sequenced genomes that form a new valuable resource for research and discovery. However, the relative contribution and the extent to which both well-established and other yet undiscovered mechanisms contribute to phenotype remains poorly understood and this was a main gap in the knowledge that the thesis aimed to investigate in more detail.

1.7.2. *Pseudomonas aeruginosa*: The clinical pathogen

Ps. aeruginosa has been declared as one of the “six top priority dangerous microbes” by the infectious disease society of America since 2006 and is still among the list of the most worrying pathogens. According to CDC 2013, MDR *Ps. aeruginosa* was classified as a serious public health threat and is still classified among the organisms with serious threats according to the

latest CDC antibiotic resistance threat report 2019. It is estimated that 51,000 cases of infection exist each year. It is not considered among the urgent threats but it may worsen and becomes urgent without ongoing monitoring and prevention activities (Centers for Disease Control and Prevention, 2013) (US CDC, 2019).

Ps. aeruginosa is one of the Gram-negative bacteria currently causing serious hospital-acquired infections with very few treatment options (Oliveira *et al.*, 2015). The organism has been classified as one of the six ESKAPE organisms with emerging clinical importance. This group of organisms has long been known as responsible for the majority of nosocomial infections and are capable of escaping the action of antimicrobial agents (Pendleton, Gorman and Gilmore, 2013).

The organism has been prevalent especially in intensive care units, in patients with critical conditions, and in burn units. It can also cause a variety of life-threatening conditions including complicated urinary tract infections, ventilator associated pneumonia, and bacteremia. It is also associated with other infections including lung cystic fibrosis, eye infections, and ear infections. Epidemiology of MDR Gram-negative bacilli has been changing over the last two decades. There has been a 5% increase from 17% to 22% in the frequency of *Ps. aeruginosa* in intensive care units as reported by some authors in the period between 2000 and 2008 (Kallen and Srinivasan, 2010). A study that reported the organisms isolated from a large tertiary care hospital in the period between 2000 and 2008 showed that 28.7% of isolated organisms were *Ps. aeruginosa* (2,700 cases). It ranked second to *Acinetobacter baumannii* in its frequency of isolation (36.2% of isolates) (Oliveira *et al.*, 2015).

In the United States, several studies have reported the prevalence of MDR *Ps. aeruginosa* to range between 10-14 % (Karlowsky *et al.*, 2003) (Karlowsky *et al.*, 2005). In Europe, the prevalence of resistance varied between 3-50 % as reported in some earlier studies (Gales *et al.*, 2001) (Goossens, 2003). In multiple countries of Asia and the Pacific, prevalence of resistance was lower during the same period and ranged between 1.6-6.9 % (Tsuji *et al.*, 2005) (Gales *et al.*, 2001) (Raja and Singh, 2007). A more recent study has reported an increase in resistance rate to different antibiotic agents from 25 % to reach 37 % in the period between 2013-2015 (Lila *et al.*, 2017). Another study in China reported an increase in the percentage of *Ps. aeruginosa* isolates in the period between 2007 to 2014 from 10 % to 26 % to represent the most commonly isolated type of Gram-negative pathogens coming only second to *Enterococcus fecalis* from the

Gram-positive group of pathogens. The same study also reported a significant increase in the percentage of MDR isolates from 64% to reach 89.87% in the period between 2007-2014 (Dou *et al.*, 2017).

1.7.3. The challenging nature of *Ps. aeruginosa* species

The high diversity of physiologic traits observed in *Pseudomonas* species is mainly due to its high metabolic adaptability, its ability to use peripheral metabolic pathways, and its possession of a large plastic genome that shows frequent recombination and acquisition of accessory elements that allow its adaptation to different ecologic changes. *Ps. aeruginosa* is considered an epitome of opportunistic infections having the ability to infect tissues with compromised defense mechanisms. It can colonize a broad spectrum of habitats and can exploit different nutritional sources with high capability for adaptation to environmental changes. *Ps. aeruginosa* has minimum nutritional requirements and can grow well at 37 ℃ and also at wide range of temperatures ranging between 4 ℃ and 42 ℃ (Jensen *et al.*, 2004). *Ps. aeruginosa* can produce a polysaccharide, alginate, and can develop biofilms which help to protect against phagocytosis and allow it to multiply inside tissues. *Ps. aeruginosa* strains are highly resistant to antimicrobials and can grow in some hospital disinfectants. It is known to be associated with a wide range of nosocomial infections and can colonize various hospital devices such as catheters and bronchoscopes. *Ps. aeruginosa* is also a source of septicemia in burn wound injuries, urinary tract infections in catheterized patients, and pneumonia in patients on respirators. It is also a predominant cause of morbidity and mortality in patients with lung cystic fibrosis.

1.7.3.1. The genome of *Ps. aeruginosa*

Ps. aeruginosa possess a relatively large genome when compared to most other sequenced bacterial species. The genome size of *Ps. aeruginosa* strains falls in the range of 5.5–7 Mb with the average size of 6.52 Mb which is considered one third longer than that of *E. coli* (4.6 Mb) or of *B. subtilis* (4.2 Mb). Genome Sequencing has offered a better understanding about the diversity of the species. The genomes of *Pseudomonas* spp. show a highly mosaic structure composed of a relatively stable core region interspersed with variable regions, probably acquired through horizontal gene transfer, and accounting for plasticity of the genome (Klockgether *et al.*, 2011). *Ps. aeruginosa* has less genome diversity when compared to other species such as *Streptococcus agalactiae* and *Haemophilus influenzae* (Tettelin *et al.*, 2008) which are known to

have more genomic diversity. Among *Ps. aeruginosa* strains, only about 10 % of genes vary and the rest are homologues. Earlier studies have shown the core genome of *Ps. aeruginosa* to be of 5.84 Mb in size representing 89.7 % (range 86.4 % - 93.3 %) of the total genome. Core genome was initially predicted to contain 5316 predicted genes accounting for 90 % out of 5892 total coding sequences in PA14 and 5570 predicted genes accounting for 95 % out of the total coding sequences of PAO1 (Ozer, Allen and Hauser, 2014). In other earlier studies, the core genome was defined at 5233 orthrologs which represented ~88 % of the average genome (Valot *et al.*, 2015). The predicted core genome was shown to vary between 4455 to 5316 genes in some studies (Grosso-Becerra *et al.*, 2014) (Ozer, Allen and Hauser, 2014). In other studies, the pangenome of *Ps. aeruginosa* was estimated to contain about 16,820 non-redundant genes with 15 % (2503 genes) constituting the core genome which is notably lower than previous studies (Mosquera-Rendón *et al.*, 2016). Other studies using clinical strains of *Ps. aeruginosa* from different geographic locations and different types of infections showed a core genome to be composed of 4910 genes (Subedi *et al.*, 2018). Other more recent studies have predicted the core genome of *Ps. aeruginosa* to be composed of 665 genes (constituting only 1 % of the entire pan-genome which was shown to consist of 54272 genes) (Freschi *et al.*, 2019). Authors show that the reason for that is attributable to the number of isolates used in earlier studies which was an order of magnitude lower than the data set used in their study. The isolates have also shown to include all five *Ps. aeruginosa* phylogenetic groups. However, the reason for these different findings may originate from fully automated informatics and non-manual consolidation. This may also result from fixed length/identity criteria, improper selection of strains or poor compensation of sequence errors. Other recent studies have shown the core genome to consist of 5109 protein-coding genes with defining a set of 321 genes as the core essential genes representing only 6.6 % of the genome (Poulsen *et al.*, 2019).

1.7.3.2. Extensive metabolic adaptability in *Ps. aeruginosa*

Ps. aeruginosa is a highly adaptable organism with the ability to use over 80 different organic compounds as energy and carbon sources (Singh, Saini and Kahlon, 2016). The genus *Pseudomonas* containing a gene pool up to 6396 predicted genes has the ability to converge wide array of organic compounds ranging from simple sugars to complex aromatic hydrocarbons using species-specific peripheral metabolic pathways. Using their peripheral metabolic

pathways, *Pseudomonads* can degrade and utilize a broad range of toxic and non-toxic compounds. The genus also has the potential to produce a wide array of secondary metabolites including various biomolecules, biosurfactants and antimicrobial compounds. These properties are considered instrumental in allowing the species to compete with other populations in the ecosystem making members of this genus one of the most successful and abundant organisms on the earth (Singh, Saini and Kahlon, 2016). *Pseudomonads* have a flexible metabolism that is mostly respiratory/aerobic, however, *Ps. aeruginosa* may grow in the absence of O_2 by using nitrate as terminal electron acceptor. Virulence of *Ps. aeruginosa* as an opportunistic pathogen and its ability to form biofilm-like microcolonies in cystic fibrosis lungs has been linked to its denitrification capabilities (Singh, Saini and Kahlon, 2016).

1.8. Different approaches used to globally investigate antimicrobial resistance

New resistance-underlying genes/gene variants and functional pathways modifying susceptibility to antibiotics are being increasingly discovered (Breidenstein *et al.*, 2008). Different functional and genomic methods have been used to mine for the basis of microbial resistance. Functional metagenomics is one of the techniques developed to mine metagenomes directly without the need for culture (Sommer, Dantas and Church, 2009). It involves extracting total metagenomic DNA from a microbial community, shearing genomic DNA to target size and shotgun cloning of the fragments into expression vectors. This is followed by phenotypic testing of the expression library. The method is considered superior to other methods based on alignment to reference databases because it enables the discovery of novel elements from diverse microbiomes and is independent of prior knowledge of resistance function (Pehrsson *et al.*, 2013). However, metagenomic libraries are not capable of identifying combinatorial antimicrobial resistance requiring the combined presence of multiple genes (Schmieder and Edwards, 2012).

“Molecular padlock probes” is another example of a function-based method for resistance discovery (Mezger *et al.*, 2015). The method was used to predict susceptibility profiles in the patients’ clinical samples. In this method, the microbial community is exposed to an antimicrobial of interest for a short duration of time. Metagenomic DNA is then extracted and mixed with oligonucleotides to capture and detect the gene sequence. DNA copy number is compared before and after exposure to antibiotics. If copy number increases following exposure, this indicates resistance. Although this method was effectively used by clinicians to prescribe

antibiotics for some uro-pathogens, it is considered as a targeted method because it can only detect some resistance elements which are carried on the probe and cannot be used to characterize novel resistance genes (Mezger *et al.*, 2015).

On the other hand, "Sequence-based methods" for resistance discovery using antibiotic resistance databases is another alternative approach used to mine for microbial resistomes. This approach is based on BLASTing query sequences of interest against specific databases in order to catalog the antimicrobial resistance gene content of the genome or the metagenome as a general principle. This approach suffers many practical and theoretical limitations as will be addressed in detail later in *Chapter 2*. Different research approaches have been used to globally investigate for antimicrobial resistance (Adu-Oppong, Gasparrini and Dantas, 2017) including transcriptional profiling, experimental evolution, and mutant library screening. These will be briefly described in the following sections.

1.8.1. Transcriptional profiling

Microarray studies can provide a snapshot of the genome-wide response of an organism to the surrounding environment where bacterial transcriptome or genome-wide expression profiles can show bacterial response to antibiotic exposure (Michelle D Brazas and Hancock, 2005). However, these changes are so complex and can include directly related as well as other less related changes. There are different types of expression responses resulting from antibiotic exposure. The first type represents changes resulting as a direct consequence of target inhibition by antibiotic. The second includes other genes that are indirectly related to resistance but can be triggered as a result of target inhibition. The third type of genes are secondary effects which occur downstream of inhibited targets. These have no particular role in antibiotic action and consequently may not affect the fate of antibiotic-treated bacteria. The fourth type is the bystander effect that occurs in generally unrelated genes (Michelle D Brazas and Hancock, 2005). This complexity of expression responses can be significantly challenging for conclusive interpretation. Also, not all cellular functions are regulated at the expression level, the fact that limits the usefulness of this approach when used separately to draw conclusions.

Genome-wide expression studies used to globally investigate antimicrobial resistance did not show to be a good predictor for antimicrobial resistance (Murray *et al.*, 2015). The underpinning principle of using this approach is summarized in examining genes differentially expressed in the

presence of sub-inhibitory levels of antibiotics. This can provide insights into some intrinsic resistance-related factors. However, it has a major drawback of inconsistent physiology which may result from undefined growth media or different growth conditions, the fact that makes it difficult to compare different studies. Exposure to sub-inhibitory concentrations of antimicrobials can slow bacterial growth which may lead to a shift of determinants towards general growth mediated ones (Cirz *et al.*, 2006). An example is the shift in expression pattern on exposure to inhibitory concentrations of ciprofloxacin towards inducible polymerases (Cirz *et al.*, 2006).

Microarray was used to characterize the response of *Ps. aeruginosa* to ciprofloxacin where 941 genes showed statistically significant differential expression to 0.3 MIC (M. D. Brazas and Hancock, 2005). Although hundreds of genes were induced or repressed on exposure to inhibitory and sub-inhibitory concentrations of antibiotics (Morita, Tomida and Kawamura, 2014), the responses resulting from microarray studies can reflect bacterial cell adaptation to stress which may reflect experimental conditions rather than real clinical exposure. Results of microarray analysis examining expression responses to inhibitory, sub-inhibitory, and lethal concentrations of tobramycin support that. Kindrachuk *et al.*, (2011) showed that the drug concentration used has significantly affected the types of genes involved as well as the level of gene expression. Genes involved were either providing short term immediate protection or long term more sustained protection and these findings varied with different drug concentrations used (Kindrachuk *et al.*, 2011). Murray *et al.*, (2015) suggested in their study that fitness profiling in combination with gene expression profiles are considered a better approach towards identifying antimicrobial fitness determinants. The approach offers the advantage of avoiding technical faults related to experimental conditions. Experimental profiles may not reflect real differential profiles because they can result from the adaptive prediction process of the bacteria to growth environment and they can also result from the pooled nature of the experiment leading to interaction between individual mutants and consequently net community-based resistance.

The analysis of global patterns of gene expression has been transformed through high-throughput sequencing which enabled RNA-seq to replace microarray (Davey and Valdivia, 2020). RNA sequencing has been used as a general approach for gene expression profiling and especially for defining transcripts boundaries and operons. It also has the ability to identify novel transcripts

including small non-coding RNA molecules and anti-sense RNA (Sharma and Vogel, 2014). In addition, it provides information on transcriptional start sites, untranslated regions of mRNA genes, open reading frames or sRNA genes which could improve genome annotations. In addition, it has been used to study RNA-protein complexes (König *et al.*, 2012). Multiple methods have been used including RNA-seq with ONT and cDNA sequencing with SMRT. SMRT-Cappable-seq has been developed to isolate unfragmented primary transcript with long read sequencing and showed potential in revealing complex operon variants in bacteria thus enabling detailed analysis of prokaryotic network regulation (Yan *et al.*, 2018). Illumina sequencing of cDNA with higher depth coverage is now considered an important approach to provide detailed transcriptional profiling and gene regulation information. Expression of mRNA and small RNA can also be studied during infection with the ability to study the differential regulation of bacterial and host transcripts using dual RNA-seq (Ritchie and Evans, 2019). Achieving higher sequence depths has enabled the better understanding of the dynamics of infections in multiple organisms (Davey and Valdivia, 2020). Quantifying the expression of transcriptional features has been mainly used to define operons including promotors and terminators. However, precise analysis is a critical step needed to maximize the value of obtaining high-quality information including promotor motif analysis and RNA polymerase binding assay (Creecy and Conway, 2015). Although bacterial transcriptomics offer a new potential of revealing new ways of host-pathogen interaction and for the understanding of networks underpinning metabolic regulation, comparisons of bacterial metabolic pathways in cultured, *in vitro* growth conditions versus *in vivo* animal host models becomes essential (Chan *et al.*, 2019). Although messenger RNA could provide valuable information about the metabolic processes in microorganisms and about microbial community dynamics, many limitations are usually encountered on using that approach in general. These limitations essentially originate from the rapid decay and turn-over time of mRNA, very short RNA half-life times and the poor correlations between mRNA and protein levels. All these factors together with the complex mechanisms governing RNA degradability and instability can raise concerns about its application for activity profiling especially in environmental and also in clinical samples (Steiner *et al.*, 2019).

1.8.2. Mutant library screening

Transposon-Mutant library screening is a systematic approach used to gain a detailed understanding about all single-gene dependent mechanisms underlying bacterial resistance to antibiotics. It has helped to uncover unknown genetic determinants of resistance in *Ps. aeruginosa* with the identification of genes whose inactivation sensitizes bacteria to a broad spectrum of antimicrobial agents (Dötsch *et al.*, 2009). Although this approach is global and is applied at the whole genome level, it is far from being comprehensive. The first reason lies in being limited to identifying non-essential genes. Another important drawback is the probability of missing small mutational events in addition to combinations of mutations leading to resistance (Dötsch *et al.*, 2009). Screening mutant libraries has identified a large ciprofloxacin resistome in *Ps. aeruginosa* including more than 100 genes. It is also unknown whether resistance identified using this approach reflects clinically significant resistance or not. However, identified elements can provide a hypothesized candidate gene pool that may have an influence on antibiotic susceptibility (Breidenstein *et al.*, 2008) with lack of evidence about their practical or clinical significance. Transposon Directed Insertion Sequencing (TraDIS) can be used to analyze the contribution of every gene in the genome to fitness in a particular environment. TraDIS allows the identification of essential genes through allowing large scale competition of millions of single-transposon mutants. First, a library containing thousands of individual transposon insertions at different insertion sites within each gene is constructed and sequenced to determine the specific location and frequency of insertions into each gene. The library is then cultured under the specific condition under study and re-sequenced to determine the positions and frequencies of transposons in the final population (Langridge *et al.*, 2009). Mutants in genes required for growth under particular stress condition will not grow or will poorly grow and thus the gene will be identified as conditionally essential. Other mutants may show relatively improved growth which may indicate that gene inhibition has provided some fitness advantage. This would then help to identify genes that may be required for growth under certain conditions or those that may provide some fitness advantage for the condition tested. However, results from TraDIS can show inconsistencies in the number of genes considered essential and this may result from insertion location or frequency bias. This consequently necessitates determining the difference between truly essential genes and conditionally essential or non-essential genes (Goodall *et al.*, 2018).

1.8.3. Experimental evolution

Experimental evolution is another approach used to understand antibiotic resistance evolution and fitness in *Ps. aeruginosa* (Cabot *et al.*, 2016). It involves exposing the organism to increasing concentrations of antibiotics and using genome sequencing information to compare wild type to mutated strains in controlled laboratory populations (McDonald, 2019). As opposed to transposon-mutant library screening, the approach can give clues about biologic fitness and costs of different genetic events where transposon-mutant library screening shows bias and does not allow studying the functional modifications in essential and non-essential genes. In addition, information obtained from mutant-library screening is restricted to single genetic events and consequently does not have the potential to show competition between emerging variants.

In a study using experimental evolution to perform a comparative analysis of the evolution of antibiotic resistance in *Ps. aeruginosa* mutator phenotypes, the approach has identified mutagenic signatures associated with resistance to different antibiotics which also showed correlations between genotype and antibiotic resistance phenotype. Identified mutations included mutations in some classically established pathways in addition to other newly identified pathways (Cabot *et al.*, 2016).The most important point that needs consideration when using this approach is that evolutionary trajectories may not be the same *in-vivo* when compared to results of *in-vitro* evolutionary experiments. Also, candidate evolutionary trajectories need to be investigated in strains with different genetic backgrounds (Cabot *et al.*, 2016) (McDonald, 2019).

Consequently, using this approach independently is not recommended and should be used as a complement to studying evolution in clinical isolates. Genomics of the evolutionary adaptation on exposure of *Ps. aeruginosa* to ciprofloxacin have been studied. Genome sequencing of evolved isolates showed a parallel evolution of some known resistance genes which correlated with different levels of resistance (Wong, Rodrigue and Kassen, 2012). However, other mutations that co-occurred with other known stereotypic mutations have conferred higher resistance levels suggesting the importance of mutations that compensate for resistance cost. Another limiting factor that also affects conclusions drawn from these experiments is the concentration of antibiotics used and the length of exposure needed before an effect is measured and analyzed. Sub-inhibitory concentrations have led to phenotypic evolutionary changes in *Ps.*

aeruginosa population (Høiby *et al.*, 2016) and this showed to be affected by the length of exposure resulting in different immediate and long-term effects.

1.9. Previous investigation of quinolone resistance in *Ps. aeruginosa*

1.9.1. Mechanism of quinolones action

Quinolones and fluoroquinolone antibiotics are considered one of the most commonly prescribed classes of antibacterial agents for the treatment of a wide range of infections (Correia *et al.*, 2017) (Martínez, 2019). These agents have been increasingly used in clinical as well as veterinary settings since the 1990s (Aldred, Kerns and Osheroff, 2014). Quinolones target mainly DNA gyrase and topisomerase IV with varying efficiencies in different bacterial species (Drlica *et al.*, 2009) (Martínez, 2019). The quinolone cell targets are the essential enzymes involved in DNA replication and supercoiling. When these two enzymes are targeted by the drug, the resulting drug-enzyme-DNA complex results in DNA breaks and chromosomal fragmentation and ultimately into cell death. Two pathways have been linked to the lethal action of quinolones; the protein synthesis-dependent pathway, and the protein synthesis independent pathway (Drlica *et al.*, 2008) (Drlica *et al.*, 2009).

Although there is some research evidence to support that the initial targets for fluoroquinolones action are different in Gram negative from Gram positive species with *par*C considered as the primary target for ciprofloxacin in some studied Gram positive species and *gyr*A as the primary target in Gram negative species (Tankovic *et al.*, 1996) (Ferrero, Cameron and Crouzet, 1995) (Drlica *et al.*, 2008), this topic is still considered controversial and needs to be evaluated on a species-by-species and a drug-by-drug basis (Aldred, Kerns and Osheroff, 2014).

1.9.2. Mechanisms and determinants of quinolones resistance and the importance of understanding the global picture

Resistance to quinolone agents has been reported as multifactorial. Resistance can result from a combination of mechanisms including target site gene mutations, increased production of MDR efflux pumps and modifying enzymes (Redgrave *et al.*, 2014) (Correia *et al.*, 2017).

Although several quinolone resistance mechanisms have been reported in the literature and are quantitatively contributing to increasing quinolones MIC and resistance, it appears that there are

other as yet unknown mechanisms that can be contributing to increased quinolones MIC in clinical isolates of *Ps. aeruginosa* (Bruchmann *et al.*, 2013). The established mechanisms of quinolone resistance reported in the literature will be discussed in detail in *Chapter 3*. Most studies addressing quinolone resistance have focused on target-mediated site mutations as the essential mechanisms underlying resistance. Although highly important, other known and yet undiscovered resistance mechanisms need to be adequately studied in much more detail. There is a growing body of evidence to suggest that resistance-associated mechanisms should not be considered mutually exclusive and that multiple mechanisms can accumulate to exhibit higher levels of resistance in an additive way (Aldred, Kerns and Osheroff, 2014) (Correia *et al.*, 2017).

The individual impact of these mechanisms can also vary among different bacterial species as manifested by the comparative effect of each of these mechanisms on ciprofloxacin MIC in different species. Several research observations have supported that. While a single *gyr*A mutation in *E.coli* exhibited 10-16 fold changes in ciprofloxacin MIC (Piddock, 1999)(Everett *et al.*, 1996), the same mutation showed no impact at all on MIC for *Staphylococcus aureus* (Hooper, 2000)(Hooper, 1999). On the other hand, accumulation of multiple *gyr*A and *par*C mutations showed higher ciprofloxacin MIC fold-changes ranging from 60 in *E. coli* (Morgan-Linnell and Zechiedrich, 2007) to 128 in *Staph. aureus* (Hooper, 2000) (Hooper, 1999). For efflux pump upregulation, fold changes in ciprofloxacin MIC varied between 4 in some Gram positive species (Kaatz, Seo and Foster, 1999) (Hooper, 2000) to 4 to 8 in other Gram negative species (Komp Lindgren *et al.*, 2005). Surprisingly, carriage of *qnr* alleles and *qep*A showed 30-32 ciprofloxacin MIC fold changes in some studies (Briales *et al.*, 2012) (Yamane *et al.*, 2007).

There would therefore seems to be a definite need for the investigation of quinolone resistance mechanisms that takes the system level into consideration rather than target level alone. This requires the inclusion of all the events that lead to resistance which also necessitates considering all the steps of quinolone interaction with different cellular molecules. This starts with quinolone trapping of DNA replication machinery to form cleavage complexes that result in bacteriostasis ending into cell death from chromosomal fragmentation through oxidative DNA damage by reactive oxygen species in the protein synthesis dependent pathway and also through the other protein synthesis independent pathway (Cheng *et al.*, 2013).

Many studies have previously proposed that the main mechanism of fluoroquinolone resistance is mediated through mutations in *gyr*A and *par*C, and that efflux regulatory genes mutations are secondary. However, this theory may not prove to be true in *Ps. aeruginosa*. Some studies have investigated the combined effect of mutations in QRDR and in efflux-pump regulatory genes and have concluded that the cumulative effect of these mutations in conferring resistance phenotype is possible (Higgins *et al.*, 2003). It is likely that efflux-pump regulatory mutations can add up to show equivalence to QRDR mutations towards increasing the level of resistance (Oh *et al.*, 2003). This can be supported by the fact that the presence of one regulatory gene mutation is not usually associated with drug resistance, and that multiple efflux regulatory gene mutations may be required (Kiser *et al.*, 2010). It is also unknown whether individual mutations or changes in protein expression are directly responsible for resistance or they are only markers for the presence of resistance (Kiser *et al.*, 2010).

It has been suggested in some small epidemiologic studies using clinical isolates that chromosomal mutations are not usually required for fluoroquinolones resistance. This observation may indicate that other mechanisms including efflux pumps may represent more essential contributors to resistance (Oh *et al.*, 2003) (Higgins *et al.*, 2003). However, the degree of resistance conferred by efflux pumps is also variable (Suresh *et al.*, 2018). Llanes *et al.*, (2011) have shown that activation of efflux pumps can act as the first step in causing low level resistance to fluoroquinolones. Similarly, the additive effect of multiple mutations in efflux pump regulatory genes on enhancing the pump system expression and consequently on producing multidrug resistant phenotype have been observed in both lab mutants and in clinical isolates of *Ps. aeruginosa* (Sobel *et al.*, 2005) (Llanes *et al.*, 2004).

1.9.3. Agents studied from the quinolone group

levofloxacin and ciprofloxacin were chosen to study in the current analysis because they are the most commonly used agents from the quinolone group of antibiotics in the treatment of different types of clinical infections caused by *Ps. aeruginosa* (Mensa *et al.*, 2018) (Rehman, W. M. Patrick and Lamont, 2019). Both agents show the highest efficacy from the quinolone group of antibiotics against *Ps. aeruginosa* with different reports showing comparable bactericidal activities of these two agents against different clinical infections. The successful treatment of bacterial keratitis and conjunctivitis using ciprofloxacin and ofloxacin and also using the third

generation levofloxacin has been well-established (Leibowitz, 1991) (O'Brien *et al.*, 1995) (LaBorwit *et al.*, 2001). Fluoroquinolone monotherapy is still among the successful options used in the treatment of *Pseudomonas* keratitis (Austin, Lietman and Rose-Nussbaumer, 2017).

Levofloxacin has also been successfully used in the treatment of severe infections caused by *Ps. aeruginosa* including serious respiratory infections and blood stream infections (Marchetti and Viale, 2003) (Schito and Schito, 2004). Levofloxacin is well known in the treatment regimen of ventilator-associated pneumonia and is considered well-tolerated in critically ill patients with fewer adverse events (Álvarez-Lerma, Grau and Álvarez-Beltrán, 2006). On the other hand, ciprofloxacin continues to be the preferred oral agent used for the treatment of urinary tract infection caused by *Ps. aeruginosa* (Shahab Qureshi, Michael Stuart Bronze, 2020). It has also been used for the treatment of osteochondritis, eye infections, ear infections and malignant otitis externa (Mösges, Nematian-Samani and Eichel, 2011). Fluoroquinolones also provide alternative treatment for bacteremia in beta-lactam sensitive patients (Shahab Qureshi, Michael Stuart Bronze, 2020). Ciprofloxacin has long been known in the treatment of *Ps. aeruginosa* and is still considered one of the most effective and most important agents widely used against the bacterium (Mensa *et al.*, 2018) (Rehman, W. M. Patrick and Lamont, 2019). It is delivered orally and intravenously with inhalable formulations developed for the treatment of chronic *Ps. aeruginosa* infections in cystic fibrosis patients (Kłodzińska *et al.*, 2016).

Findings have shown equal *in-vitro* activities of both ciprofloxacin and levofloxacin against 300 *Ps. aeruginosa* isolates from hospitalized patients which were both more active than ofloxacin (Bonfiglio, 2001). The study has concluded that levofloxacin is a good option for the treatment of infections sustained by *Ps. aeruginosa* with excellent bactericidal activity (Bonfiglio, 2001). Similarly, in vitro antibacterial efficacy studies have suggested equivalent activities for both ciprofloxacin and levofloxacin against *Ps. aeruginosa* (MacGowan, Wootton and Holt, 1999). This has also been supported by Kowalski *et al.*, (2001) who showed comparable *in-vitro* activity of the three quinolone agents including levofloxacin, ciprofloxacin and, ofloxacin when applied to keratitis isolates from 200 patients (Kowalski *et al.*, 2001).

On the other hand, other studies have shown that *Ps. aeruginosa* clinical isolates are more susceptible to levofloxacin than to ciprofloxacin with levofloxacin showing greater bactericidal activity than ciprofloxacin using time-kill experiments (Segatore *et al.*, 2000). This has also been

supported by Segatore *et al.,* (1999) who demonstrated the superior activity of levofloxacin when compared to ciprofloxacin as assessed from a national survey including intensive care units from oncology and hematology wards from 13 Italian hospitals (Segatore *et al.*, 1999). Grillon *et al.*, (2016). The study has concluded that the superiority of levofloxacin to ciprofloxacin cannot be affirmed. While ciprofloxacin seems to have good efficiency against susceptible *Ps. aeruginosa,* levofloxacin showed better activity than ciprofloxacin in some tested strains (Grillon *et al.*, 2016).

Although the speed of bacterial killing has been shown to be superior in levofloxacin than in ciprofloxacin, some research evidence has pointed to the equivalence of in-vitro potency of both ciprofloxacin and levofloxacin when calculated on the basis of incidence of resistance (Schito and Schito, 2004). A study has shown that levofloxacin use was more associated with the isolation of quinolone resistant *Ps. aeruginosa* when compared to ciprofloxacin use. This has been attributed to the greater in-vitro activity of ciprofloxacin than that of levofloxacin in *Ps. aeruginosa* (Kaye *et al.*, 2006). Other studies have supported the same observations by demonstrating that emergence of fluoroquinolone resistance was correlated with levofloxacin rather than ciprofloxacin use (Lee *et al.*, 2010). Similarly, the proportion of fluoroquinolone resistant *Ps. aeruginosa* was observed to be correlated with the consumption of all antibiotics and specifically with levofloxacin (Yang *et al.*, 2020).

1.10. Previous investigations of aminoglycoside resistance in *Ps. aeruginosa*

1.10.1. Mechanisms of aminoglycoside action

Aminoglycosides were identified through systematic screening of soil Actinobacteria that started in the 1940s. The first aminoglycoside streptomycin was discovered from *Streptomyces griseus* and was successfully used for the treatment of tuberculosis and then for other infections caused by Gram-negative bacteria in 1944 (Schatz, Bugle and Waksman, 1944). Aminoglycoside agents have been widely used in medicine with the ongoing growth of their use as a result of the increased progress in developing effective derivatives since 1970s (Mingeot-Leclercq, Glupczynski and Tulkens, 1999).

Aminoglycosides have long been considered essential agents in *Ps. aeruginosa* therapy. They are used in the treatment of a wide variety of clinical infections including pulmonary infections,

blood stream infections, urinary infections, eye, wound, and burn infections. Aminoglycosides are bactericidal agents and usually exhibit synergy, hence they are used in combination with other anti-*pseudomonas* agents, most notably with beta-lactams (Edson and Terrell, 1999). They are commonly used in the combination treatment protocols for serious Gram-negative bacterial infections.

While quinolones penetrate cell membrane through porin channels to enter the bacterial cell, aminoglycosides promote their own uptake by interacting with bacterial LPS on the outer membrane of Gram-negative bacteria (Poole, 2005). Although the mechanism by which aminoglycosides antibiotics penetrate Gram negative bacterial cell wall remains incompletely understood, it has been proposed to consist of three different stages. According to the current model, the first stage is simply an electrostatic interaction between the positively charged aminoglycosides (AGs) and the negatively charged lipopolysaccharides (LPS) of the outer bacterial membrane. The two subsequent stages are the energy dependent phase I (EDPI) which is characterized by a slow rate of uptake that is correlated with aminoglycosides concentrations and the energy-dependent phase II (EDPII) that uses energy from electron transport and ATP hydrolysis (Taber *et al.*, 1987).

The main mechanism of bactericidal activity of aminoglycoside antibiotics is related to interfering with various aspects of protein synthesis. Aminoglycosides bind to the decoding region of the 16S ribosomal RNA (rRNA) component of the 30S subunit of bacterial ribosomes. Binding to the A site leads to conformational changes in rRNA which consequently lead to mRNA misreading. This consequently affect translation accuracy, stopping peptide chain elongation which results in the synthesis of defective proteins. Defective proteins accumulated in the cell membrane lead to altered permeability and the resulting secondary increase in intracellular aminoglycosides concentrations. Aminoglycosides also act by inhibiting translocation through immobilizing peptidyl-tRNA at the A-site of the ribosome, consequently inhibiting protein synthesis. Some aminoglycoside agents can also bind to the 23S rRNA component of the 50S subunit. This binding affects the mobility of ribosomal subunits, which consequently interferes with translation and ribosome recycling (Borovinskaya *et al.*, 2007) (Becker and Cooper, 2013).

An important point to consider here is that following the entry of the first few molecules of aminoglycosides which leads to misreading in protein translation, cytoplasmic membrane integrity and function are compromised due to faulty proteins, leading to an autocatalytic cycle of AGs uptake, followed by cell death (Davis, Chen and Tai, 1986).

1.10.2. Molecular mechanisms and system level perspectives of aminoglycosides resistance

Diverse mechanisms have been described as contributors to aminoglycoside resistance in *Ps. aeruginosa*. The most noticeable mechanisms include drug inactivation by aminoglycoside-modifying enzymes and decreased drug accumulation either through active efflux or decreased cell wall permeability. The genes coding for ribosomal protein L25 (*rpl*Y), UDP-glucose pyrophosphorylase (*gal*U), and the *nuo* operon have been linked to aminoglycosides resistance. Inactivation of some chromosomal genes including *mex*Z, *rpl*Y, *gal*U, PA5471, and *nuo*G have been shown to be associated with gradual increase in aminoglycosides MIC in laboratory mutants (El'Garch *et al.*, 2007a) (Islam *et al.*, 2009). El'Garch *et al.*, (2007) showed in their study that double, triple, and quadruple mutants have resulted in a cumulative effect on aminoglycoside resistance. Quadruple mutants have shown 16- to 64-fold increase in MIC when compared to the wild-type strain PAO1. This may demonstrate the ability of *Ps. aeruginosa* to accumulate resistance via intrinsic (i.e. nonenzymatic) mechanisms.

Other studies have also shown that upregulation of MexXY–OprM efflux system is the most important aminoglycoside resistance-conferring system in CF *Ps. aeruginosa* (CFPA) isolates. This usually results from mutations occurring in the repressor regulatory gene *mexZ*, which product downregulates the expression of the operon *mexXY*. MexAB–OprM was also reported to contribute to aminoglycoside resistance in a low ionic strength environment. It was suggested that efflux-mediated aminoglycoside resistance only lead to a moderate two-fold increase in MIC (Vogne *et al.*, 2004).

1.10.3. Agents studied from the aminoglycoside group

According to the British National Formulary (BNF) (Committee, 2020), gentamycin is the aminoglycoside of choice in the UK and is widely used in the treatment of serious infections caused by multiple Gram negative organisms including *Ps. aeruginosa*. On the other hand, amikacin is more stable to enzyme inactivation than gentamycin and is used in the treatment of

serious infections caused by gentamycin-resistant organisms including biliary infections, septicemia, and endocarditis. Aminoglycosides have been considered as a vital component in anti-pseudomonas chemotherapy particularly pulmonary infections in cystic fibrosis patients (Poole, 2005).

Although gentamycin is generally considered an old drug, it possesses a potent bactericidal activity and remains to have a key role in the treatment of some types of infections caused by *Ps. aeruginosa*. Amikacin has the broadest spectrum of activity among the aminoglycosides group of antibiotics and is considered a good treatment candidate to strains showing multiple resistance to other aminoglycosides (Ehsan and Clancy, 2015).

Gentamycin is known to be more active against *Ps. aeruginosa* than amikacin while amikacin is known to exhibit less resistance rates due to its structural nature (Kluge *et al.*, 1974). Both gentamycin and amikacin have been used in the treatment of urinary tract infections caused by *Ps. aeruginosa* with amikacin achieving better peak serum concentrations at lower therapeutic doses (Gilbert, Eubanks and Jackson, 1977). In addition, amikacin also shows activity against gentamycin resistant strains and can achieve high blood levels making it clinically effective in the treatment of *Pseudomonas*-associated pulmonary infections complicating cystic fibrosis (Lau *et al.*, 1977). Recent evidence also shows that amikacin can achieve high *in vitro* potency against *Ps. aeruginosa* respiratory and blood isolates (Sutherland, Verastegui and Nicolau, 2016).

Kim *et al.*, (2018) have recently demonstrated declining trends of gentamycin and amikacin resistance in *Ps. aeruginosa* according to data from the Korean Nationwide Surveillance of Antimicrobial Resistance (KONSAR) program which was attributable to decreased aminoglycosides consumption levels (Kim *et al.*, 2018). This makes it important to investigate the molecular bases of resistance in these agents.

1.11. Molecular platforms used for identifying antibiotic resistance markers

Most molecular platforms for bacterial ID/AST use nucleic acid-based markers indicative of the presence of bacteria and/or antibiotic resistance. Other newly emerging molecular markers also include enzyme, protein, or metabolite markers. Current nucleic-acid based molecular platforms used for bacterial AST include PCR-based assays or other microarray, nanoparticle and microparticle-based nucleic acid assays.

PCR-based assays depend on using culture samples to detect specific sets of pre-identified nucleic acid targets that are species-specific, for bacterial ID, or antibiotic resistance-associated for bacterial AST. While most of these assays are culture-dependent, some other culture-independent assays are available and can greatly reduce the turn-around time (1 to 8 hours).

Some examples of FDA-cleared rapid bacterial ID/AST molecular assays are what is known as "Multiplex Infectious Diseases Syndromic Panels" including; FilmArray®, VERIGNE®, Unyvero, ePlex®, MAX™, ProGastro™ (Gonzalez and McElvania, 2018), (Hanson and Couturier, 2016). Unyvero is a Multiplex Panel used for detection of lower respiratory tract pathogens (Ozongwu *et al.*, 2017). The lower respiratory tract infection panel was FDA cleared April 3, 2018 and can detect 19 Gram-positive and Gram-negative bacterial organisms and 10 genetic markers for antibiotic resistance which include; *mec*A, *tem*, *ctx*-M, *kpc*, *ndm*, *oxa*-48, *oxa*-58, *oxa*-23, *oxa*-24, and *vim*. Another rapid detection system is BIOFIRE® FILMARRAY® Multiplex Real-Time PCR Systems (Poritz *et al.*, 2011). The respiratory panel of this system can detect 20 bacteria and 7 antibiotic resistance markers including; CTX-M, KPC, IM, NDM, VIM, Oxa48-like, *mec*A/C and MREJ. The blood culture identification panel can also test for 24 pathogens and 3 antibiotic-resistance genes. VERIGENE® Gram-positive blood culture test can detect 4 bacterial genera, 9 species, and 3 resistance markers including; *mec*A, *van*A, and *vanB*. VERIGENE® Gram-negative blood culture test can test for 4 bacterial genera, 5 species and 6 resistance genes including; CTX-M, IMP, KPC, NDM, OXA and VIM (Kim *et al.*, 2016). The ePlex® platform by GenMark is an example of another rapid diagnostic testing platforms (Schmitz and Tang, 2018) that uses ferrocene-labeled oligonucleotide probes for multiplex detection of some syndromic panel targets. The platform employs rapid sample-to-answer workflow. The entire steps of nucleic acid extraction, fluidic transport of sample to the amplification, and detection steps occur within the platform cartridge.

Other microarray nucleic acid-based platforms are capable of highly multiplex SNP analysis in a single assay. Pre-culture or nucleic acid amplification primary steps may still be required. Examples for this approach include Check-MDR CT array and the VERIGENE®system. Check-MDR CT103 array (CheckPoints, Wageningen, Netherlands) allows the rapid detection of Extended Spectrum Beta Lactamases (ESBLs), including TEM, SHV, and CTX-M; plasmid-mediated cephalosporinases (CMY-2-like, DHA, FOX, ACC-1, ACT/ MIR, and CMY-1-

like/MOX); and CREs (KPC, OXA-48, VIM, IMP, and NDM) (Cuzon *et al.*, 2012). The VERIGENE® system (Nanosphere, Luminex, Austin, TX) can identify a panel of resistance markers (*mec*A for methicillin; *van*A and *van*B for vancomycin; and CTX-M for the detection of ESBLs, IMP, KPC, NDM, OXA, and VIM for carbapenemases) from positive blood cultures (Ward *et al.*, 2015).

Using multiplex PCR even when combined with DNA microarray hybridization is considered less robust for the detection of less abundant targets and consequently shows lower clinical sensitivity and some discrepancy with culture methods. Also, it may suffer from some technical limitations associated with contamination from non-pathogenic bacteria. The main theoretic challenge here remains in their limited potential to detect a set of pre-identified genes rather than gene variants or SNPs in most cases. I also need to re-emphasize that resistance identification in all currently available rapid "Multiples Syndromic Panels" available is directed to a very limited subset of genes that often includes specific types of acquired resistance rather than intrinsic resistance elements. Most available panels detect carbapenemases-related genes or Extended-spectrum B-lactamases (ESBL). For example, Xpert CarbaR® detects 5 gene families; KPC, NDM, VIM, IMP and OXA-48; Verigene® system detects the most common carbapenemases and CTX-M ESBL and FilmArray® BCID detect KPC (Tuite *et al.*, 2014).

1.12. Practical technologic advances addressing some sequence-based diagnostics limitations

The increasing use of next-generation sequencing technologies has made the analysis of bacterial genomes much more accessible at a lower cost than was previously possible. Introducing benchtop sequencers can be considered a good step towards sequence-based diagnostics and is now considered in the capability of many analytic laboratories (Deurenberg *et al.*, 2017) (Chai *et al.*, 2018) (Morganti *et al.*, 2020). The cost can be specifically reduced for target gene panel identification (van Nimwegen *et al.*, 2016). Using such an approach is superior to PCR-based marker detection as it can detect genetic markers underpinning specific risk profiles and can also provide typing without shipping isolates to reference facilities. In this way, it can provide the "one-solution fits all" approach to deal with different types of infections. An important limitation for some sequencing technologies is the time required for completion of a full sequencing run which can exceed 2 days. One of the approaches used to overcome this situation is GeneSipper which was used for the identification of foodborne pathogens within one working day (Lambert

et al., 2015). It is a genomic tool used for the analysis of single colony isolates by sampling (sipping) of raw data during an early stage of the genome sequencing process to identify a pre-defined set of markers with a quality evaluation to validate the analysis. Another study has also applied a developed bioinformatics analytic approach (Chainmapper) that can be used to analyze microbial reads directly from clinical samples and was used to evaluate sequencing directly from urine samples enabling microbial identification from polymicrobial clinical specimens. It showed a performance comparable to culture-based identification (Hasman *et al.*, 2014). Whole-genome sequencing of *Chlamydia trachomatis* was first achieved directly from clinical samples without culture using an approach called immunomagnetic separation for targeted bacterial enrichment with multiple displacement amplification (IMS-MDA) (Seth-Smith *et al.*, 2013). Other studies have also proved the possibility of metagenome sequencing directly from vaginal swab specimens (Andersson *et al.*, 2013). In other research studies, genome sequencing was applied to identify bacterial outbreak genomes directly from fecal outbreak samples (Loman *et al.*, 2013). All these approaches can prove that direct sequencing and bioinformatics analysis from clinical specimens are now becoming more achievable.

As will be discussed in *Chapter2*, genome sequencing has the potential to provide the information needed to transform the use of the technology into clinical diagnostics. With the increased understanding about the correlation between phenotype and genotype especially in some organisms, there remain some practical limitations that need to be overcome to shorten the turn-around time in case of using different sequencing technologies as a diagnostic tool.

The nucleic acid extraction step is considered a rate-limiting step in sequencing studies. Eliminating inhibitory substances and having high-quality extracts with a rapid turn-around time are essential for successful downstream processing. An example of a recently introduced rapid genomic extraction methodology is PDQeX (Stanton *et al.*, 2019). It is a single-step DNA extraction method that is compatible with high-throughput sequencing and capable of purifying nucleic acid from the sample for downstream applications in under 30 min. It is proposed that similar novel systems can transform and accelerate sample preparation with minimal chance of sample contamination and reduced operator handling which can achieve higher sample quality to be directly used for WGS. Simple and reliable methods for extracting nucleic acids directly from biologic samples are expected to facilitate molecular point of care testing within the next 5 years.

Portable DNA extraction machines and portable sequencers such as MicroGEM's (UK) PDQeX were tested for in-field use. These systems have enabled farmers to whole-genome sequence viruses from plant material in a single day on the farm (Shaffer, 2019).

Another approach is clinical metagenomics using third generation sequencing. MinION (Oxford Nanopore Technologies, Oxford, UK) is a highly portable genome sequencer that weighs less than 100 grams, an advantage that makes it applicable for in-field or point of care use. Data is read off the MinION from a laptop via a Universal Serial Bus (USB) port which adds to the feasibility of its use (Quick, Quinlan and Loman, 2014).

In contrast to Illumina sequencing platforms that utilize optical readings and require precise microscopic alignment with repeated calibration, the MinION (Oxford Nanopore Technologies, Oxford, UK) works by taking high-frequency electrical current measurements as DNA passes through a protein nanopore at 450 bases per second. MinION is a single molecule sequencing approach that can read long DNA molecules which may also lead to a higher error rate. However, generating accurate sequence readings can be achieved by reading the same genomic regions many times and thus eliminating the errors through consensus averaging (Loman and Watson, 2015). However, some errors are systematic which still makes the error rate a challenging point.

The technology has been tried for in-field applications in the diagnosis of some plant viruses and showed successful results (Boykin *et al.*, 2018). It was also used in the field for real-time genomic surveillance and monitoring of the Ebola virus and also showed success in some resource-limited settings (Quick *et al.*, 2016). It has also been used for the diagnosis of Zika virus directly from clinical samples for the purpose of epidemic surveillance and monitoring (Quick *et al.*, 2017) (Faria *et al.*, 2016). In the clinic, Schmidt *et al.*, (2017) show in their study that it is also possible to use the technology for direct metagenome diagnosis of pathogens and acquired antibiotic resistance genes directly from urine samples. In addition, Charalampous *et al.*, (2019) have developed an optimized rapid clinical metagenomics pipeline for the diagnosis of bacterial lower respiratory tract infection and for the identification of resistance genes using nanopore metagenomics which has achieved a rapid turn-around time of 6 hours from sample to answer. Metagenomic sequencing-based approaches show the potential to overcome limitations related to both culture and PCR including the speed and the coverage of wider organism

spectrum (Chiu and Miller, 2019). Although some computational challenges still exist, nanopore sequencing has great potential to be used as point-of-care diagnostic for multiple clinical sequencing applications including patient's bedside, emergency room, local clinic, or field applications (Chiu and Miller, 2019). The use of novel portable sequencing technologies together with clinical metagenomics can also form a foundation for real time pathogen and disease surveillance which helps to direct outbreak responses (Gardy and Loman, 2018).

1.13. Project

1.13.1. Significance, justification and outline

Although *Ps. aeruginosa* is equipped with extensive intrinsic resistance machinery that leads to basal level of lower susceptibility to antibiotics, there is still a lack of comprehensive understanding of resistance-underlying mechanisms. There is still a gap in understanding elements that underpin intrinsic resistance at the whole system level.

Considering the challenging and the worrisome nature of *Ps. aeruginosa* as a species and as an opportunistic clinical pathogen, many targeted comprehensive studies have been performed to survey all known resistance elements and showed that there is still a gap in knowledge and that some undetermined chromosomal mechanisms need to be further investigated (Henrichfreise *et al.*, 2007). The special nature of the species makes AMR prediction in *Ps. aeruginosa* challenging due to the complexity of associating phenotype and AMR genotype, the species large pangenome, lack of understanding about the difference in gene content for virulence and AMR determinants between clinical and environmental isolates (Freschi *et al.*, 2015), and the species-specific complicated regulation of resistance mechanisms (Jeukens, Freschi, Kukavica-Ibrulj, J.-G. Emond-Rheault, *et al.*, 2019). Comparative genomic approaches can be considered superior to other approaches discussed above including genome-wide expression profiling, mutant library screening, and experimental evolution. Sequence-based approaches are considered powerful as they can examine all possible genes and gene variants simultaneously at the population level offering the highest resolution analysis to the single nucleotide variant level. Identifying genetic variants underpinning different phenotypic behavior in the population can help to characterize the genetic background of complex phenotypic traits and to identify genetic signatures related to these traits which can consequently be used to predict and diagnose resistance at the practical level.

1.13.2. Study aim and objectives

- Setting up sequenced strain resources for *Ps. aeruginosa* suitable for comparative genomics.
- Functional analysis to determine functional groups of behavior for antibiotic resistance phenotype.
- Applying comparative behavioral genomics for candidate behavioral determinants.
- Comparing and contrasting findings with prior knowledge and relating it to population structure and wider system determinants.
- Secondary use of genomes to address species genomics and population biology, and ways in which it can otherwise inform the use of genomes in diagnostic microbiology.
- Testing the predictive value of identified determinants and those in literature as a potential basis for genome-based directed antibiotic treatment.
- Identification of DNA markers that can be used for inclusion in rapid diagnostic tests to direct antibiotic selection and treatment.
- Identifying optimized combination of candidate markers from research output and what is known in literature that could act as potential genome-based molecular rapid diagnostics.
- Understanding the population structure of studied bacterial species in relation to studied antibiotic resistance phenotype.

What is this research really about?

Although different approaches have been used to investigate for mechanisms underlying antibiotic resistance at the global level, the thesis tries to explore the topic in a hypothesis free manner using a multi-step approach. Different genome sequencing technologies are emerging with some existing limitations to be addressed before it can be transformed into a practical applied tool. The application of next- generation sequencing technologies for example may require some demanding infrastructure requirements which may limit its use to some settings as will be discussed in *Chapter 2*. On the theoretic level, the performance of some available bioinformatic platforms are not considered comprehensive enough to be used as an applied tool with *Ps. aeruginosa* species as will be evaluated in *Chapter 2*. Comprehensive databases that can be utilized for diagnostic workflows are still lacking. In general, there is still uncertainty about the practicality or the feasibility of applying different sequencing technologies in clinical

diagnostics. Some practical and theoretic limitations still exist as will be discussed further in *chapter 6* and these limitations still need to be handled. There is still a gap in our knowledge about the most comprehensive and most informative set of molecular predictors that can be translated into best potential diagnostic markers for practical application.

The research approach adopted here focuses on the best use of sequence information to explore for the best predictor molecular elements that explain resistance and susceptibility phenotype in the organism *Ps. aeruginosa* at the system level and within the frame of background population structure. This will be achieved by evaluating the performance of some available bioinformatics resources and workflows in predicting for antibiotic resistance phenotype in *Chapter 2*. The next stage will expand more on the findings of the first stage by mining for known genes and gene variants that have been described in the literature in relation to chromosomal resistance of quinolone and aminoglycoside groups of antibiotics. This will be combined with the evaluation of newly encountered variants in the genes assessed as will be shown in *Chapter 3*. The stage that follows will investigate the background population structure in *Ps. aeruginosa* with the aim of exploring the relation of the identified elements to high-risk clones, the topic that will be studied in *Chapter 4*. In *Chapter 5*, comparative behavioral genomics will be applied to fill additional gaps by exploring for new elements that can be added to improve our understanding and knowledge. The ultimate aim is to find the best combination of molecular predictors from what is known and what is newly explored in order to be adapted into an assessment and decisions flows that can improve our toolbox to fight against AMR, the perspective and steps that are going to be discussed in detail in *Chapter 6*.

Adopting this approach, the thesis seeks to create new understanding of the existing knowledge about antibiotic resistance by exploring the relative contribution of different chromosomal mechanisms of quinolone and aminoglycoside resistance in order to prioritize the best performing set of candidate markers for inclusion into a rapid diagnostic platform. This is expected to offer better alternatives to other rapid diagnostics that are currently available in the market. The thesis also suggests a diagnostic algorithm using the investigated markers to be used when sequencing information becomes available in a clinical setting as will be illustrated in examples shown in *Chapter 6.* The following table summarizes the themes that will be presented in the thesis to achieve its over-all aim.

Topic	Objectives
Chapter 2. Genotypic versus phenotypic prediction of antibiotic resistance in *Pseudomonas aeruginosa* using existing tools and databases	• Evaluation of the accuracy of some currently available (WGS)-based tools in predicting antibiotic resistance in *Ps. aeruginosa* and the best ways to overcome current limitations.
Chapter 3. The role of known Quinolones and Aminoglycosides resistance mechanisms in explaining resistance in *Pseudomonas aeruginosa*	• Review of the literature to extract genes and gene variants associated with quinolone and aminoglycoside resistance. • Describing and analyzing the distribution of identified resistance-associated markers in the studied set of *Ps. aeruginosa* isolates. • Testing for predictive values and other measures of diagnostic accuracy for known resistance-associated markers. • Exploring the best predictor combination of markers that could improve diagnostic performance using cluster analysis and multiple regression analysis.
Chapter 4. Background genomic context of Quinolone and Aminoglycoside resistance determinant and molecular markers in *Pseudomonas aeruginosa*	• MLST and serotype profile analysis in relation to quinolone and aminoglycoside resistance phenotype. • Describing the population structure and diversity in *Ps. aeruginosa*. • Identifying quinolone and aminoglycoside resistance markers in relation to high-risk clones.
Chapter 5. Investigation of Whole System determinants of antibiotic resistance to gentamycin and ciprofloxacin in *Ps. aeruginosa* using comparative behavioral genomics	• Identification of system-level functions associated with susceptibility and resistance phenotype to quinolone group of antibiotics (1st round of CBG annotation). • Identification of gene variants showing significant association with susceptibility and resistance to ciprofloxacin and gentamycin (2nd round of CBG annotation). • Testing for the potential practical application of the identified variants to be used as molecular diagnostics by evaluating the diagnostic accuracy of candidate markers. • Examining the functional effect of prioritized gene variants. • Finding the best combination of variants with improved diagnostic performance. • Examining the distribution of candidate alleles in relation to background genomic structure.
Chapter 6. General discussion and conclusions	• Conceptual and practical framework of current research findings into real-life application.

Chapter 2. Genotypic versus phenotypic prediction of antibiotic resistance in *Ps. aeruginosa* using existing tools and databases

2.1. Introduction and background

2.1.1. Applications of Genome-sequencing in clinical microbiology for public health benefit

2.1.1.1. Culture-free microbial identification

2.1.1.2. Molecular epidemiologic investigations and Typing

2.1.1.3. Tracing bacterial transmission pathways

2.1.2. Advantage and value of using genome-sequencing in diagnostic microbiology

2.1.3. Culture-free clinical metagenomics

2.1.4. Available Resistance Prediction Tools and bioinformatics platforms

2.1.5. Antimicrobial resistance prediction in *Ps. aeruginosa*

2.2. Aim: Evaluation of the accuracy of Whole Genome Sequencing (WGS)-based tools in predicting antibiotic resistance in *Ps. aeruginosa*

2.3. Methodology

2.3.1. Collection, culture and preservation of study isolates

2.3.2. Phenotypic antimicrobial susceptibility testing

2.3.3. Whole genome DNA extraction, sequencing and assembly

2.3.4. Assessment of genetic diversity and selection of the most diverse group

2.3.5 In-silico prediction of antimicrobial susceptibility

2.3.6. Comparison of phenotypic testing to genomic prediction

2.4. Results

2.4.1. Description of the genetic background of studied *Ps. aeruginosa* isolates

2.4.2. Comparison of the phenotype distribution of study collection with international population data

2.4.3. Predicted antibiotic sensitivity using in-silico tool

2.4.4. Comparing predictions to measured phenotypes

2.5. Discussion

2.6. Conclusions: Requirements for the implementation of WGS in clinical microbiology laboratories to predict resistance and guide interventions

2.1. Introduction and Background

The use of genome sequencing is expected to offer a key tool to fight against antimicrobial resistance (AMR). The impressive number of sequenced bacterial genomes represents an invaluable resource to investigate the molecular basis underlying AMR. Sequencing as a growing technology is expected to offer personalized infection diagnosis and management that can overcome many of the current limitations encountered using conventional infectious disease diagnostics.

Susceptibility testing based on culture is still considered the current standard of practice in guiding antibiotic selection (Bayot ML, 2020). While the method has an inherent drawback of longer specimen to answer time compared to what sequencing is currently capable of (Charalampous *et al.*, 2019), it is still considered the gold standard method for most pathogens (Belkum and Dunne, 2013). Other limitations of conventional culture include the need for relatively large numbers of viable organisms, several steps of pre-analytic processing, and limited organism detection spectrum (Belkum and Dunne, 2013). Some practical drawbacks can also be encountered in clinical practice which makes using sequencing-based diagnosis and treatment a better alternative in some settings. Sequence-based diagnostics have a great potential to overcome some challenges related to conventional culture yield and organism recovery timelines through the rapid identification and testing of pathogens occurring at very low concentrations (Allcock, Jennison and Warrilow, 2017). This would make it possible to detect pathogens or antibiotic resistance markers from specimens even when collected after initiation of antibiotic therapy at higher sensitivity (Rossen *et al.*, 2018). This gives the opportunity to change empirical therapy into more specific antimicrobials earlier in the treatment course. This would subsequently help to shorten the duration of treatment and its other associated complications (Caliendo *et al.*, 2013).

Over the past decade, there has been a growing interest in developing a rapid molecular diagnostic approach for early detection of antibiotic resistance (Clerc and Greub, 2010) (Bauer *et al.*, 2014) (Carey-Ann D. Burnham *et al.*, 2017). PCR has been the most common nucleic acid-based tool used for microbial identification, typing, and for identification of the genomic basis of resistance. To identify resistance using PCR, commercial PCR kits and in-house tests are used to identify resistance genes and mutations or to identify a panel of markers of interest (Walker *et*

al., 2016) (Torres *et al.*, 2016) (Allcock, Jennison and Warrilow, 2017). The most widely used commercial kits for this purpose include those used for the identification of *mec*A methicillin resistance *Staphylococcus aureus* (Patel *et al.*, 2011) and *van*B *Enterococcus* resistance genes (Bourdon *et al.*, 2010). Other examples include the detection of resistance causing mutations in *Mycobacterium tuberculosis* (Bowles *et al.*, 2011) and macrolides resistance in *Helicobacter pylori* (Zhang *et al.*, 2016). Sequencing has the potential to reveal the entire genomic repertoire of markers linked to resistance including core genes and also acquired resistance elements, thereby enabling the identification of multiple resistance mechanisms simultaneously (Allcock, Jennison and Warrilow, 2017). Identified resistance elements can also be linked to genomic and epidemiologic background giving greater insights into forces that drive resistance spread and transmission. New variants of genes can sometimes preclude their detection using available PCR kits (García-Álvarez *et al.*, 2011), while this is not an obstacle if sequencing is used instead.

2.1.1. Applications of genome-sequencing in clinical microbiology for public health benefit

Genome sequencing has been used for culture free microbial identification, for molecular epidemiologic investigations and for high-resolution typing. It has also been used for tracing bacterial transmission pathways in localized hospital outbreaks and in larger epidemics, for the identification and prediction of antimicrobial resistance and virulence genes, and also to direct infection control policies and procedures. Some of these uses are discussed in the following sections.

2.1.1.1. Culture-free microbial identification

Using genome sequencing to identify microbial species has the advantage of rapid high precision species identification using the same technology, platform, and expertise. This includes fastidious, slowly growing and difficult to culture taxa. It also provides an unbiased approach for identifying a wide range of pathogens in complex polymicrobial samples (Sabat *et al.*, 2017). Wide arguments exist about the possibility of genome sequencing to replace routine microbial culture-based identification. However, genome-based signatures can offer the ultimate resolution and precision compared to conventional morphologic and biochemical-based identification. In the near future, routine bacterial culture may end up as a lost art. This can result from the increasing availability of genome sequence data information together with the accumulating knowledge about the molecular bases of resistance and virulence (Rossen *et al.*, 2018). 16S

rRNA gene Sanger sequencing has been historically used for accurate bacterial identification from biologic samples independent of their cultivability or phenotype (Woo, P.C.; Lau, S.K.; Teng, J.L.; Tse, H.; Yuen, 2008). It has been used as an alternative to conventional culture in some slowly growing, difficult to culture bacteria, or bacteria with uncommon phenotype. Next generation sequencing and third generation sequencing technologies have more recently allowed reliable identification of bacterial genera as a rapid alternative analysis to sanger sequencing without culturing steps (Winand *et al.*, 2020). Due to the high sequencing throughput, NGS technologies have allowed 16S rRNA gene sequences generation from organisms in mixed samples without the need for intermediary culture steps.

2.1.1.2. Molecular epidemiologic investigations and Typing

Next generation sequencing has emerged as a powerful tool with a greater discriminatory power when compared to all other molecular typing methods (Allcock, Jennison and Warrilow, 2017). It can discriminate similarities and differences with evolutionary projections about relatedness that consequently help to frame these differences into background biologic context (Ladner *et al.*, 2019). Sequencing allows the comparison of bacterial isolates down to the resolution of single nucleotide differences, thus providing higher discriminatory resolution which can greatly improve our understanding of different epidemiologic and biologic aspects of hospital pathogens. It can help to monitor the global emergence and dispersal of infections in different settings. This has also enabled pathogen outbreak investigation at the highest possible resolution which also allowed the improved detection of transmission events in addition to other potential cost benefits (Madigan *et al.*, 2018) (Besser *et al.*, 2018). The added value of WGS to routine local TB surveillance systems has been studied by integrating epidemiologic and clinical information with WGS genotypic data of *Mycobacterium tuberculosis* (Zakham *et al.*, 2019). WGS-based SNP typing showed to be useful for routine TB surveillance system. Another study has also shown that molecular typing enabled the detection of unsuspected *M. tuberculosis* transmission as compared to traditional contact tracing (Genestet *et al.*, 2019).

Using genome sequencing data has previously shown to provide some insights about the biology and the epidemiology of persistent isolates of *Ps. aeruginosa* in hospital settings. In a study investigating hospital outbreaks of *Ps. aeruginosa* , a group of strain specific gene clusters and SNP differences that confer phenotypic differences were described as helping in strain

persistence and survival over years in hospital environment (Snyder *et al.*, 2013). Using whole genome sequencing to investigate for *Ps. aeruginosa* hospital outbreaks provided detailed epidemiologic information without the need for further typing. Authors suggest that implementation of WGS in real time typing of hospital outbreaks can become widespread due to reproducibility and lower cost as it also helped to understand the outbreak situation rapidly and with certainty (Parcell *et al.*, 2018).

2.1.1.3. Tracing bacterial transmission pathways

Sequencing data can be used effectively to trace bacterial transmission pathways. This includes direct transmission between individuals, identifying worldwide patterns of spread as well as local outbreak investigations (Croucher and Didelot, 2015). Direct transmission between individuals is often used to determine whether individuals are part of the same transmission chain or not. The genomic investigations of outbreaks help in identifying sources of infection, in identifying super-spreaders and consequently in directing infection control measures. Using genome sequence data combined with epidemiologic date allows outbreaks' recognition and accurate inferences about transmission patterns. This permits more effective interventions and targeting of infection control resources (Price et al., 2013).

Bacterial genome sequencing enables the understanding of the dynamics of infection transmission within hospital settings and the factors governing such an interaction between health-care facilities and the community. Sequencing has superior discriminatory capability to reconstruct pathogen transmission pathways when compared to conventional typing (Nikolayevskyy *et al.*, 2019). This can show great potential when used as part of the surveillance system for the early detection and management of health care-associated infection. Applying epidemiologic surveillance systems internationally will permit the monitoring of newly emerging strains. When these types are linked to local epidemiologic clinical surveillance data, they can provide an effective early warning system.

Harris et al., (2013) have evaluated the advantage of whole genome sequencing for the accurate detection and understanding of outbreak transmission events and concluded that more precise identification of all patients involved in a MRSA outbreak was possible with sequencing offering superior discrimination than other standard techniques in practice. This has consequently enabled more precise detection of the outbreak network and the community sources linked to the

outbreak (Harris et al., 2013). Using sequencing-based approaches offer an advantage for identifying sources of ongoing transmission of hospital outbreaks which can directly inform infection-control interventions. This is considered promising with the declining cost of rapid genome sequencing (Campbell *et al.*, 2018). The discriminatory power of sequencing-based outbreak investigation has been evident in different organisms including *Vibrios cholera* (Chin et al., 2011), *Mycobacterium tuberculosis* (Gardy et al., 2011) and *Escherichia coli* (Rasko et al., 2011). Other more recent studies have also used SNP-based analysis in hospital epidemiologic investigation in *Staphylococcus aureus* in ICU and identified multiple transmission pathways between patients and hospital reservoirs (Dancer *et al.*, 2019). This investigation helps in identifying sources of infection transmission which consequently helps in its control. WGS was also used to investigate outbreak versus non-outbreak inter-facility transmission of carbapenem resistant *Klebsiella pneumoniae* (Spencer *et al.*, 2019). WGS-based typing has also showed an important role in tracing inter-hospital transmission of carbapenemase producing *K. pneumoniae* in other recent studies (van Beek *et al.*, 2019). This has also shown an important role in tracing *Ps. aeruginosa* outbreaks (Parcell *et al.*, 2018).

2.1.2. Advantage and value of using genome-sequencing in diagnostic microbiology

WGS is becoming increasingly used to confirm infection outbreaks and other types of persistent infections, thus enabling the highest discrimination previously unachievable using conventional methods (Humphreys and Coleman, 2019). Köser, Holden, *et al.*, (2012) showed that using conventional methods to infer transmission (including antibiograms) was useful in only minority of cases with genome-sequencing showing at least 136 SNPs in isolates with nearly identical antibiograms and the same sequence type which would suggest absence of any relation or transmission event among these patients as judged by the genetic differences. latest generation benchtop sequencing platforms can provide useful sequence data in less than one day when compared to the standard clinical microbiology practice. It also enables multiple additional analyses with direct and immediate clinical value at no additional costs (Köser, Holden, *et al.*, 2012). The accuracy and depth of information provided enables the discrimination of outbreak and non-outbreak isolates which is faster and superior to other commonly used conventional typing methods (Reuter *et al.*, 2013) (Quainoo *et al.*, 2017).

The need for customized species-based clinical diagnosis is an underestimated drawback encountered in conventional microbiology lab. Cost of reagents, equipment, specialized tests and the need for experienced staff to diagnose and test for uncommon or rare organisms are usually overlooked. This makes some tests limited to reference laboratories which usually have more-prolonged turnaround time and commonly do not provide results with the speed needed to inform patient management (Mosammaparast, Nolte and McAdam, 2012). Genome sequencing can offer a promising solution to this situation in routine diagnostic laboratories. This can be implemented through analytic platforms integrated into the laboratory usual workflow. When sufficient background knowledge and infrastructure becomes readily available, sequence-based analytic platforms can become an automated process applied to a wide spectrum of organisms. This would provide superior quality information, decreased cost and decreased turn-around time (Köser, Ellington, *et al.*, 2012) (Ladner *et al.*, 2019).

Harris *et al.*, (2013) showed in their study that health-care cost attributable to MRSA outbreak was in excess of 10000 sterling pounds while the cost of rapid WGS of one isolate was 95 sterling pounds including sample preparation, library quality control and sequencing (Harris *et al.*, 2013). The current availability of bench top sequencers and data analysis software with price tags comparable to conventional molecular techniques are taking the technology closer to clinical microbiology laboratories (Rossen *et al.*, 2018).

Although next-generation sequencing has improved precision and sensitivity of detection when compared to culture-based molecular identification, turn-around time is sometimes considered a challenge with the use of next-generation sequencing (Loit *et al.*, 2019). This includes the time needed for complex library preparation, other technical steps including post sequencing assembly and interpretation. However, bench top sequencers and third generation sequencing are becoming more accessible and user friendly and would allow sequencing in hours (Schadt, Turner and Kasarskis, 2010) (Watts *et al.*, 2017). Harris *et al.*, (2013) showed that the time needed to do MLST, to detect certain genes and to verify point mutations associated with virulence/risk can outweigh the time needed to perform one sequencing run. The sequencing run took one day to complete and the bioinformatics pipeline required less than 1 hour of analysis time (Harris *et al.*, 2013).

Another study has shown that a rapid turn-around time is achievable using Illumina Miseq benchtop sequencer which is capable of providing a comprehensive report with all possible genomic information and with the highest resolution for outbreak investigation in just two days (McGann *et al.*, 2016). Many challenges exist for the application of genome sequencing both in the wet lab and in the dry lab. These challenges include sample preparation, library construction and sequencing in the wet lab. Bioinformatics support with infrastructure and the need for automation in the dry lab. However, there is evidence that a rapid turnaround time can be achieved in small-scale laboratories and not only in reference laboratories (McGann *et al.*, 2016). Third generation sequencing technologies including MinION (Oxford Nanopore Technologies) allows more rapid identification of organisms following both amplicon-based and PCR-free metagenomics approaches (Loit *et al.*, 2019). MinION showed the potential to be used as a rapid diagnostic achieving turnaround time as short as 2.5 hours (Loit *et al.*, 2019).

2.1.3. Culture-free clinical metagenomics

A faster diagnostic capable of rapid identification of infectious pathogens and antibiotic resistance is needed ('Antibiotic susceptibility diagnostics for the future', 2019). Clinical metagenomics includes the comprehensive analysis of microbial and host genetic material using sequencing (Chiu and Miller, 2019). Metagenomic sequencing enables the identification of a broad range of pathogens including viruses, bacteria, fungi and/or parasites directly from clinical samples on the basis of unique DNA and/or RNA sequences. The approach has high potential to change infectious disease diagnostics especially when combined with portable sequencing technologies such as MinION nanopore. Clinical metagenomics can prove superior to WGS-based single isolate identification when considering the turnaround time of sample to answer (Carey-Ann D Burnham *et al.*, 2017). Metagenomic-based tests has been developed for dual pathogen and antimicrobial resistance identification achieving short turn-around time of 6 hours sample-to-answer (Charalampous *et al.*, 2018). It has been shown that metagenomic-based sequencing could replace "gold standard" culture by providing rapid, sensitive and real time results for the diagnosis of lower respiratory infections (Charalampous *et al.*, 2018). In addition, technical feasibility and proof-of concept clinical validity have also been demonstrated for nanopore metagenomics sequencing use in severe pneumonia diagnosis offering the information needed from the sequencing profiles in culture negative samples (Yang *et al.*, 2019).

Another metagenomic application is the detection of circulating pathogens produced by infections at diverse body sites using the "Karius test". This test allows the detection and quantification of microbial cell free DNA circulating in the bloodstream with the capability of identifying more than 1250 clinical pathogens in blood samples (O'Grady, 2019). The test has been validated for the diagnosis of sepsis and showed promising results. Because detection can be confounded by commensal microorganisms or unrelated minor infections that may lead to false positive results, it is recommended for use as a complement standard diagnosis in patients with suspected sepsis, pneumonia, or fever of unknown origin (O'Grady, 2019).

2.1.4. Available Resistance Prediction Tools and bioinformatics platforms

Some high-throughput tools are currently available to analyze sequence data and to predict for antibiotic resistance. These tools can usually identify AMR associated genes and single nucleotide polymorphism in query sequences based on the databases supported by the tool. This subsequently depend on data curation by software developers. Using this approach has many theoretical and practical limitations.

In theory, our understanding about the mechanisms and diversity of the molecular basis of AMR is still growing with the increasing availability of genome sequence data. Available methods are limited to the known types of AMR determinants and sometimes cannot be applied to the data generated by all sequencing technologies. In addition, the data used can be restricted to a specific set of reference sequences that are sometimes not comprehensively representative of all available knowledge on AMR determinants in the microbial species studied. Another important point is the organism spectrum included in the tool analysis capabilities. Most tools either align sequencing reads to a set of reference genes or search for reference gene matches in *de novo* assembled sequences. Examples for the first include SRST2 (Inouye *et al.,* 2014) and Kmer Resistance (Clausen *et al.,* 2016). A drawback for both SRST2 and Kmer Resistance is their inability to identify or interpret different variants such as some SNP-associated resistance mutations. They are usually limited to the identification of particular genes or a group of pre-defined alleles. Other tools that require the use of pre-assembled sequences can be considered more computationally expensive with the possibility of errors during the assembly step which can lead to missed identification of some resistance associated determinants. Examples of these

tools include ARG-ANNOT (Gupta *et al.*, 2014a), ResFinder (Zankari *et al.*, 2012), SSTAR (de Man and Limbago, 2016), and RAST (Davis *et al.*, 2016).

On the practical side, many limitations exist for using these tools for applied purposes. First of all is the availability of the tool and its accessible use for microbiologists or clinicians without specialist bioinformatics skills. Second practical limitation is their high computational resource requirements. Some tools are only available via web services while others lack the high-throughput analysis stream for a large number of samples. For example, ARIBA tool (Hunt *et al.*, 2017) combines both alignment and targeted local assembly with the opportunity to identify AMR genes and variants precisely. It gives support to a number of public databases including ARG-ANNOT, CARD (Jia *et al.*, 2017), MEGARes (Lakin *et al.*, 2017), and ResFinder. Despite having many functionalities, it requires specific expertise skills for obtaining the input data, for preparing the data for analysis, running, and summarizing the output data which cannot be easily achieved in most clinical settings.

2.1.5. Antimicrobial resistance prediction in *Pseudomonas aeruginosa*

Ps. aeruginosa can be considered a candidate study system for antibiotic resistance prediction, however, the special nature of the species makes AMR prediction in *Ps. aeruginosa* specifically challenging (Jeukens, Freschi, Kukavica-Ibrulj, J.-G. Emond-Rheault, *et al.*, 2019). This is in part attributable to the large pan-genome of the species. J. Jeukens *et al.*, (2017) have highlighted the complexity of associating phenotype and AMR genotype in *Ps. aeruginosa* due to the species-specific complicated regulation of resistance mechanisms. Their results showed that the identified resistome using the CARD database does not correlate with phenotype and that some resistance genes can be found in some sensitive isolates. It is important to consider that even antibiotic susceptible strains of the species can carry some resistance mechanisms due to intrinsic AMR. Another possible source of bias in predicting clinically important resistance determinants is the current lack of clarity of understanding of the difference in gene content for virulence and AMR between clinical and environmental isolates (Freschi *et al.*, 2015).

2.2. Aim: Evaluation of the accuracy of Whole Genome Sequencing (WGS)-based tools in predicting antibiotic resistance in *Ps. aeruginosa*

The ultimate goal of a better molecular microbiology diagnostic is to reduce the traditional microbiology lab turn-around time to be able to impact decision making. This can only be achieved when sequencing is linked to an automated system of data interpretation which can be converted into real time meaningful clinical reports that can be used to direct treatment choice. However, to date, these automated comprehensive diagnostic workflows have not yet been developed for clinical use for most bacterial species. Instead some current bioinformatics tools are publicly available and are used to predict for phenotypic resistance by identifying all known resistance associated genes and mutations. The aim of this chapter is to evaluate the accuracy of some genomic based antibiotic resistance-prediction tools currently available for *Ps. aeruginosa* and consequently their suitability to be used practically to direct treatment.

There is also a lack of comprehensive studies that investigate the concordance between phenotypic antimicrobial susceptibility testing and WGS-based resistance prediction for *Ps. aeruginosa.* Although some studies have been performed for epidemiologic purposes to link background genetic markers of antimicrobial resistance to phenotype (Kos *et al.*, 2015), these studies are directed to identify previously known intrinsic and acquired resistance determinants. Kos *et al.*, (2015) showed in their study a sensitivity and specificity of 94% for genotypic inference of levofloxacin resistance. Genotypic markers for amikacin resistance were identified in only 60% of amikacin non-susceptible isolates. In addition, 30 out of 283 amikacin susceptible isolates were found to carry amikacin resistance-associated genes. However, in that study, resistant strains were over-represented and there was no information about the genetic diversity of tested isolates. In a recent comprehensive study, the CARD database was used to identify the resistome of 390 *Ps. aeruginosa* strains from a wide array of environmental, clinical, and animal sources, but it was not linked to the strains phenotype (Julie Jeukens *et al.*, 2017). To achieve that, a genetically and epidemiologically representative collection of *Ps. aeruginosa* isolates including both sensitive and resistant isolates has been tested in this chapter.

2.3. Methodology

2.3.1. Collection, culture and preservation of study isolates

Strains included in the study were kindly provided by Prof. David Livermore (Livermore, Williams and Williams, 1981) and from the BSAC Bacteraemia Resistance Surveillance Program (Reynolds, Hope and Williams, 2008).

Bacterial strains provided on agar slopes were cultured overnight at 37.5 ℃ on non-selective nutrient agar media obtained from Sigma-Aldrich. Cryogenic bacterial stocks were preserved by collecting 200 µl of pure culture colonies for each isolate grown on nutrient agar in cryopreservation fluid which was preserved in cryogenic tubes and kept at - 80 ℃. The preservation cryogenic fluid was prepared from a mix of 1:1 nutrient broth and 80 % glycerol to a total volume of 3 ml. 80 % glycerol for cryopreserved stocks was prepared by dissolving 200 gm in final volume of 250 ml Milli-Q water, autoclaved at 121 ℃, 15 minutes. Nutrient broth for microbiology was obtained from Sigma-Aldrich and prepared by dissolving 4 g media powder in 250 ml (2X) and autoclaving for 15 minutes at 121 ℃.

2.3.2. Phenotypic antimicrobial susceptibility testing

Eighty- seven highly diverse isolates of *Ps. aeruginosa* were selected from an in-house collection developed at Brunel University London. Antibiotic susceptibility testing was performed using the disc diffusion Kirby-Bauer method (Andrews and Howe, 2011) and the Oxoid M.I.C.Evaluator™ (Thermo Fisher Scientific) to measure MIC values according to the manufacturer's instructions. Each strain was tested for its susceptibility to ciprofloxacin, levofloxacin, gentamycin, and amikacin. Disc-zone diameter and MIC values were used to classify the isolates into sensitive or resistant according to the latest clinical breakpoints defined by EUCAST (EUCAST, 2018). Figures are shown in Appendix I.

Oxoid M.I.C. Evaluator™ (Thermo Fisher Scientific) provides a gradient of antibiotic stabilized on a plastic strip with 30 graduations to give an accurate MIC over the range of 256 µg/ml- 0.015 µg/ml. The antibiotic starts to release from the plastic forming a defined concentration gradient in the area around the strip after the M.I.C. evaluator strip is applied to the pre-inoculated agar. After incubation, a zone of inhibition forms around the M.I.C. evaluator strip. The MIC is read using the graduated scale at the point where the growth of the test organism touches the strip.

Antibiotic susceptibility testing using Oxoid M.I.C. Evaluator™ (Thermo Fisher Scientific) to measure MIC was performed according to the following steps:

- Muller-Hinton or Cation-adjusted Muller Hinton agar media was prepared and poured to a depth of 4mm ±0.5 mm.
- Several colonies of the test isolate from a pure culture were emulsified into saline suspension and compared to 0.5 McFarland Standard to adjust turbidity level. McFarland standard was prepared as shown in Appendix I.
- A sterile cotton swab was then dipped into the suspension and excess moisture was removed by pressing and rotating against the edge of the tube.
- The plate was then inoculated with the bacterial suspension by swabbing in at least three different directions taking care to do this evenly to ensure that no gaps are left in the deposited inoculum.
- The surface of the agar was then allowed to dry completely before applying the M.I.C.E strip since excess moisture can cause distortion of the gradient.
- The strip was applied to the plates within 15 minutes of inoculation to avoid pre-growth of the organism.
- Using sterile forceps, the strip was removed from the sachet by handling the end with logo and antibiotic code.
- The strip was then placed with the scale facing upwards and the antibiotic gradient downwards in contact with the agar.
- The strip was applied by putting the end with the lowest concentration onto the plate first and then carefully rolling the strip onto the agar to ensure good contact with the entire length of the strip.
- Once the strip was applied to the agar, the plates were the incubated immediately to avoid pre-diffusion of the antibiotic.
- After incubation, MIC value was read at the point of intersection of growth inhibition with the strip gradient.

2.3.3. Whole genome DNA extraction, purity check, sequencing, and assembly

Whole genomic DNA extraction for *Ps. aeruginosa* isolates was done using the FastDNA® SPIN Kit and the FastPrep® Instrument (MP Biomedicals, Santa Ana, CA).

The following shows whole genome DNA extraction steps performed:

1) Cell Lysis:

►Lysing matrix containing garnet matrix and one ¼ inch ceramic bead was used with the cell lysis solution and homogenized in the FastPrep instrument. The orange-capped tubes containing Lysing Matrix A were appropriately labeled and filled with 1 mL of CLS-TC solution (from the kit). A 10 µl blue inoculating loop was used to generously scoop approx. 200 µl of bacteria from the plate. The bacteria on the loop were then rubbed against the tube wall and matrix to dislodge into the tube. The caps were then screwed tightly making sure the rubber seal was in place, and processed in the FastPrep Instrument at speed 6.0 m/s for 40 sec.

► The tubes were then centrifuged at 13,000 rpm for 10 min at room temperature. The centrifuge step was done to pellet the debris including all membranes and proteins.

►DNA released in solution is then added to the binding matrix step, mixed, and incubated. The Binding Matrix (from the kit) was thoroughly mixed to resuspend silica particles, and 750 µl was dispensed in fresh marked 2 mL Eppendorf tubes. After centrifugation, the supernatant (approx.750 µl) was transferred into the tube with Binding Matrix, avoiding the pellet of cell debris and lysing matrix. The tube was thoroughly mixed by inverting several times, and the tubes were the incubated with agitation on a rotating wheel for 5 minutes at room temperature. The tubes were then centrifuged at 13,000 rpm for 1 minute at room temperature to pellet the binding matrix. The supernatant was carefully discarded to preserve the pellet, subsequently 500 µl of SEWS-M solution (from the kit) was added and the pellet was completely resuspended by tapping the tube.

2) Solid phase purification of DNA:

►DNA was then transferred to the spin column module and after centrifugation, it was bound to the silica membrane and all other contaminants are removed by two wash steps. To achieve that, the resuspended Binding Matrix was applied to a SPIN module and centrifuged at 13,000 rpm for

1 minute. The contents of the catch tube were discarded and replaced. The SPIN modules were centrifuged a second time at 13,000 rpm for 1 minute at room temperature, and the catch tube was replaced with a recovery tube.

►DNA was then eluted in TE buffer and collected in the tube. The DNA was eluted by gently re-suspending Binding Matrix above the SPIN filter in 100 µl TE. This was then incubated for 5 minutes at 55 ℃ in a heat block. The SPIN module was then centrifuged at 13,000 rpm for 1 minute to bring the eluted DNA into the recovery tube, and the SPIN filter was discarded.

3) Purification and precipitation of DNA:

This stage includes RNase treatment, washing DNA pellets in 70 % ethanol solution to remove trapped solutes and finally DNA was eluted and stored in TE buffer. The DNA was RNase treated by adding 2 µl RNase A (10 mg/mL) and incubated for 30-45 minutes at 37 ℃.1/10 volume of 3M Na-Ac pH 5.2 was added and mixed thoroughly; subsequently 2.5 volume of cold 100 % ethanol was added. The tubes were inverted several times to mix and stored at - 20 ºC for 30 minutes. The tubes were then centrifuged at 13,000 rpm for 30 minutes at room temperature. After the presence of the pellet was carefully inspected, the supernatant was carefully removed and 500 µl of cold 70 % ethanol was added and inverted several times. The tubes were centrifuged at 13,000 rpm for 5 minutes at room temperature. The ethanol was carefully removed, and the pellet was air-dried for 5 minutes at room temperature. The DNA pellet was finally dissolved in 50-100 µl TE buffer, by gently tapping the tube and the tubes were incubated tubes in heat block at 55 ℃ for 5 minutes or room temperature for 1 hour (Bej *et al.,* 1996).

Checking DNA quantity and purity:

This step was carried out using the NanoDrop microvolume sample retention system (Thermo Scientific NanoDrop). The software automatically calculates the nucleic acid concentration and purity ratios after sample measurement. Sample quality can also be assessed from spectral image assessment. Sample quality can be accurately assessed by assessing the overall spectral quality together with 260/280 or 260/230 ratios. Pure nucleic acids typically yield a 260/280 ratio of ~1.8 for DNA. The 260/230 purity ratio is a second measure of DNA purity with values for a "pure" nucleic acid commonly in the range of 1.8 - 2.2 (Desjardins and Conklin, 2010). Graphs showing assessment of sample quality using NanoDrop are shown in Appendix I.

The Qubit® dsDNA HS Assay Kits (Invitrogen by life technologies) and the Qubit® 2.0 Fluorometer instrument (Invitrogen by life technologies) were used according to manufacturer instructions to quantify dsDNA in extracted DNA samples and the final DNA concentration was normalized to 30 ng/µl before being sent for sequencing. Detailed steps are shown in Appendix I.

Sequencing and assembly:

Isolates were sequenced by the Wellcome Trust Centre for Human Genetics using Illumina next generation 150-bp paired-end sequencing with 192 multiplexed libraries to generate 14-45x coverage on Illumina HiSeq 2500. Each genome was assembled using the *de novo* sequence assembly program SPAdes (Bankevich *et al.*, 2012). Assembly quality and downstream sequence analyses were carried out using MUMmer (Delcher, 2002), BLAST (Altschul *et al.*, 1990), and in-house perl scripts.

2.3.4. Assessment of genetic diversity and selection of the most diverse group

The isolates used in this study were chosen as a set of strains best representing species diversity from a larger collection of isolates developed at Brunel University London. Strains with greater than 13,500 SNPs in the core genome separating any strain from its nearest neighbor in the collection were selected for use in this study.

2.3.5. In-silico prediction of antimicrobial susceptibility

Three methods were used to predict microbial sensitivity based upon the genome sequence data. Depending upon the requirements of the analysis pipeline, Raw sequence reads, or assembled sequences were submitted for analysis.

MICRA: FASTQ files of sequenced strains were uploaded for analysis by the core part of MICRA pipeline before running the antibiotic susceptibility and resistance prediction module (www.pegasebiosciences.com/MICRA/micra.php). The post-analysis module of MICRA pipeline was used to identify potential antibiotic susceptibility and resistance as described (Caboche *et al.*, 2017).

CARD: The Resistance Gene Identifier (RGI) analytic tool provided by CARD (https://card.mcmaster.ca/analyze/rgi) was used to analyze uploaded sequence files in FASTA format as described (McArthur *et al.*, 2013). The RGI tool generated a detailed output report

showing all antibiotic resistance associated genes and SNPs that were identified in submitted sequences. Those associated with resistance to aminoglycosides and quinolone groups of antibiotics were selected for further analysis in this study.

ResFinder: Assembled sequences in FASTA format were used as the input to the web-based tool ResFinder provided by the Centre for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/ResFinder/) to identify acquired antimicrobial resistance genes related to aminoglycosides and quinolones groups of antibiotics (Zankari *et al.*, 2012). Default parameter for gene identification were used (90% as the threshold for %ID and 60% as minimum length).

2.3.6. Comparison of phenotypic testing to genomic prediction

The phenotypically determined susceptibilities were compared with the results of the WGS-based prediction tools. The accuracy of the genome-based methods was assessed by calculating sensitivity, specificity, and predictive values using IBM SPSS Statistics for Windows (SPSS.V21).

2.4. Results

2.4.1. Description of the genetic background of studied *Ps. aeruginosa* isolates

The dataset included in this study showed a total of 877,218 SNPs in the comparable core genomes. The isolate distance measure used to ascertain diversity was the pairwise SNP distance of an isolate from its nearest neighbor in the core genome. The isolates had an average distance of 18,213 SNPs, and the isolates with the minimum and maximum distance being 13,731 and 63,011 SNPs, respectively (Figure 2.1).

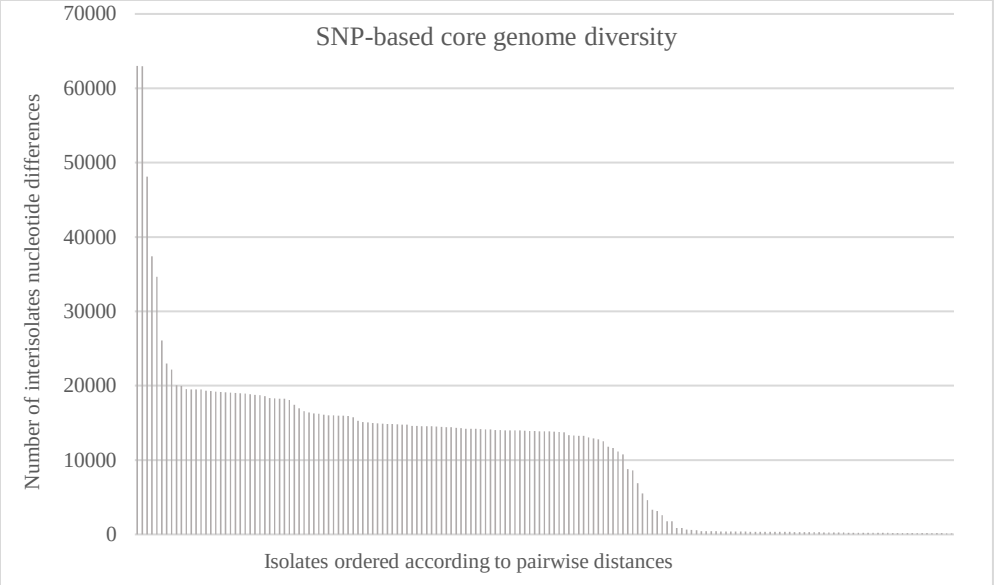


Figure 2.1. Estimates of Evolutionary Divergence in sequences shown as pairwise distances to the next nearest strain in the collection

2.4.2. Comparison of antibiotic susceptibility phenotype distribution of study collection with international population data

The distribution of ciprofloxacin MICs measured in tested isolates was compared to the international MIC distribution available from EUCAST databases at http://www.eucast.org/mic_distributions_and_ecoffs/ (Figure 2.2). Compared distributions show that the distribution of susceptibility of the strains used in this study is representative of that observed in the normal clinically isolated population; though it does not include strains with very high resistance phenotypes. The international MIC distribution of ciprofloxacin sensitivity from EUCAST reference database is based on 27,967 observations (82 data sources). Similarly, Figure 2.3 shows the compared distribution of ciprofloxacin sensitivity of study isolates to the international distribution from the EUCAST reference database based on 4,096 observations (9 data sources) as expressed by the diameter of the growth inhibition zone.

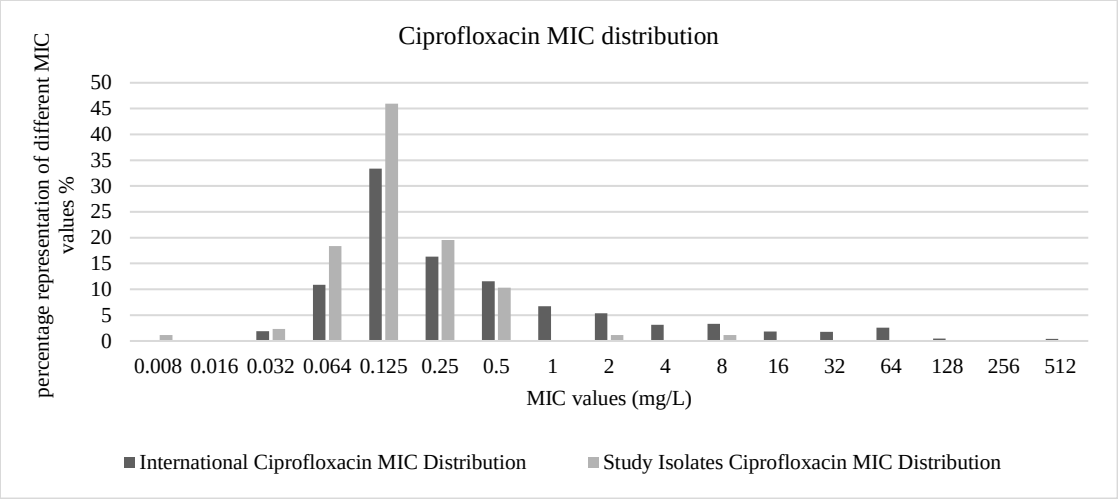


Figure 2.2. Distribution of ciprofloxacin susceptibility among international Ps. aeruginosa population and study isolates (MIC)

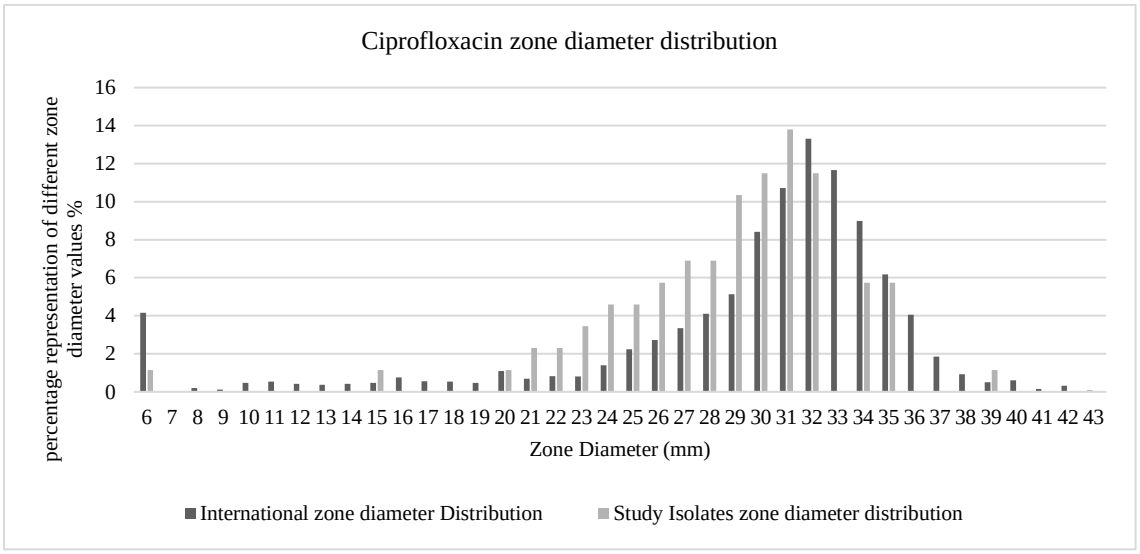


Figure 2.3. Distribution of ciprofloxacin susceptibility among international Ps. aeruginosa population and study isolates (zone diameters)

The distribution of levofloxacin sensitivity as measured by both MIC and the diameter of the growth inhibition zone was also compared to the international distribution available from EUCAST databases as shown in Figure 2.4 and Figure 2.5, respectively. The international distribution of levofloxacin MIC is based on 14,871 observations (11 data sources) and that of diameter of inhibition zone is based on 363 observations (4 data sources). Similar to ciprofloxacin, the distribution of sensitivity is highly concordant with that observed in the

worldwide clinical population, though slightly shifted towards smaller zone diameters as observed in Figure 2.5.

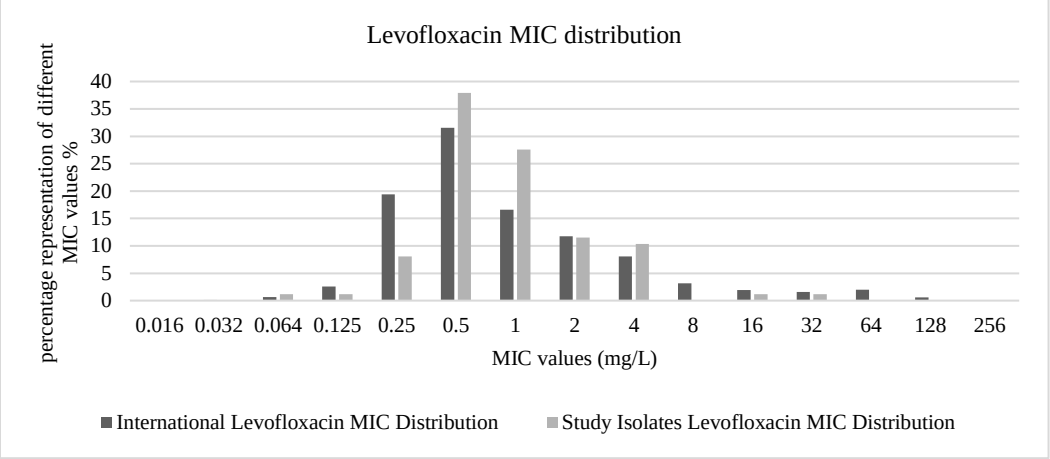


Figure 2.4. Distribution of levofloxacin susceptibility among international Ps. aeruginosa population and tested isolates (MIC)

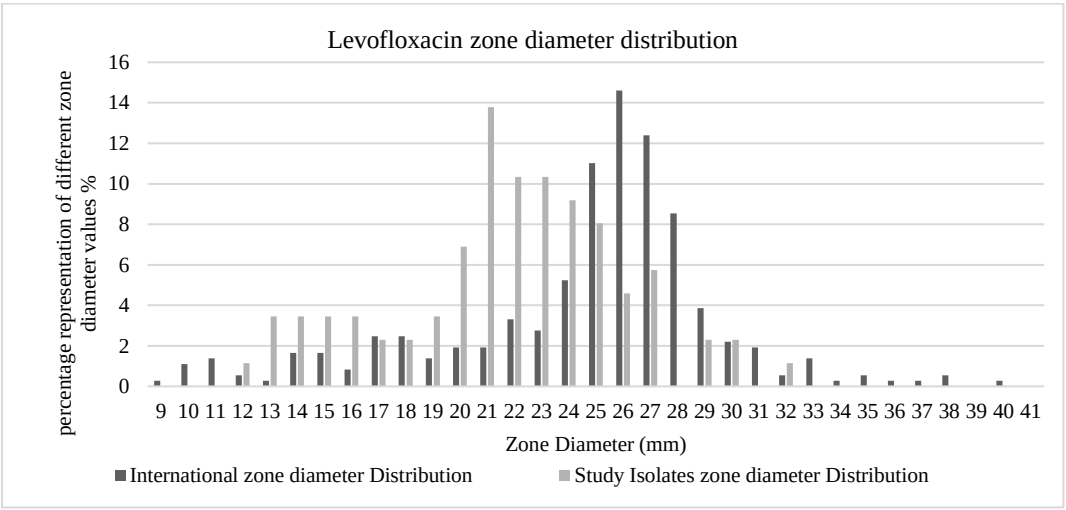


Figure 2.5. Distribution of levofloxacin susceptibility among international Ps. aeruginosa population and tested isolates (zone diameters)

The same distribution data for amikacin sensitivity using MIC and growth inhibition zones are shown in Figure 2.6 and Figure 2.7, respectively. International distribution of amikacin MIC is based on 17,369 observations (13 data sources) and that of diameter of inhibition zone is based on 538 observations (3 data sources). Again, the distribution of susceptibility and resistance is highly concordant with that observed in worldwide clinical population. This similar distribution

is expected because the strains in the collection are selected for diversity and include clinical strains that pre-date much of the selective pressures for evolution to resistance.

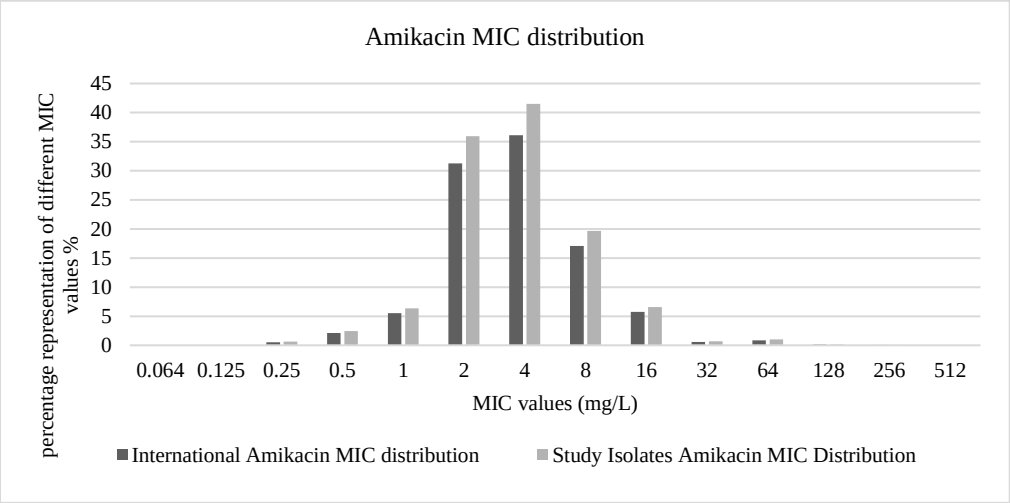


Figure 2.6. Distribution of amikacin susceptibility among international Ps. aeruginosa population and tested isolates (MIC)

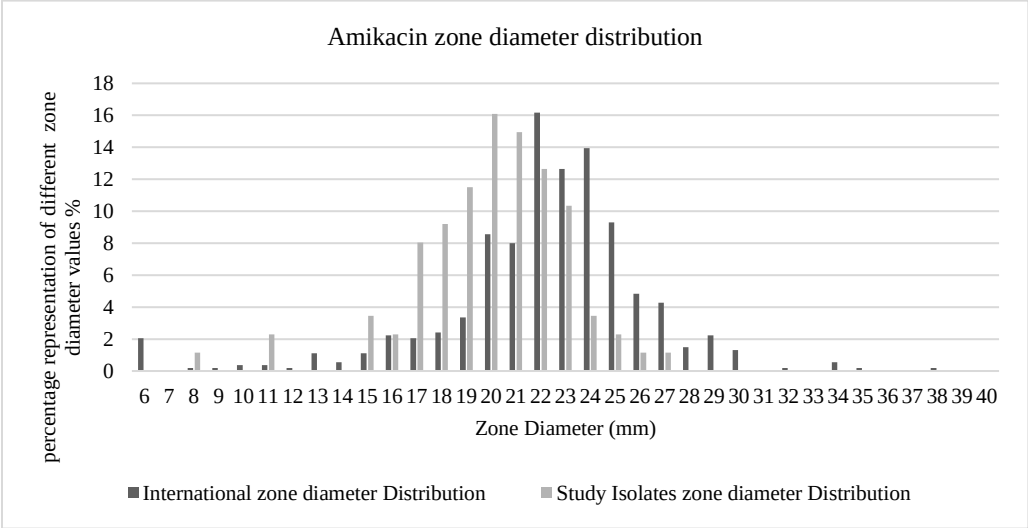


Figure 2.7. Distribution of amikacin susceptibility among international Ps. aeruginosa population and tested isolates (zone diameter)

The distribution of gentamycin susceptibility as measured by both MIC and the diameter of the growth inhibition zone were also compared to the international distribution available from EUCAST databases as shown in Figure 2.8 and Figure 2.9, respectively. The international distribution of gentamycin MIC is based on 24,490 observations (72 data sources) and that of

diameter of inhibition zone is based on 4095 observations (9 data sources). The distribution of studied isolates was shifted towards higher gentamycin MICs and towards lower zone diameters.

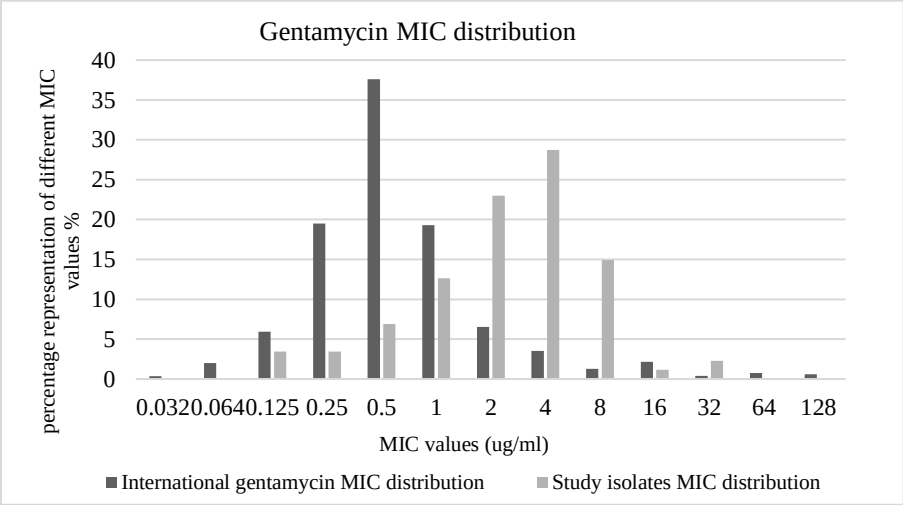


Figure 2.8. Distribution of gentamycin susceptibility among international Ps. aeruginosa population and tested isolates (MIC)

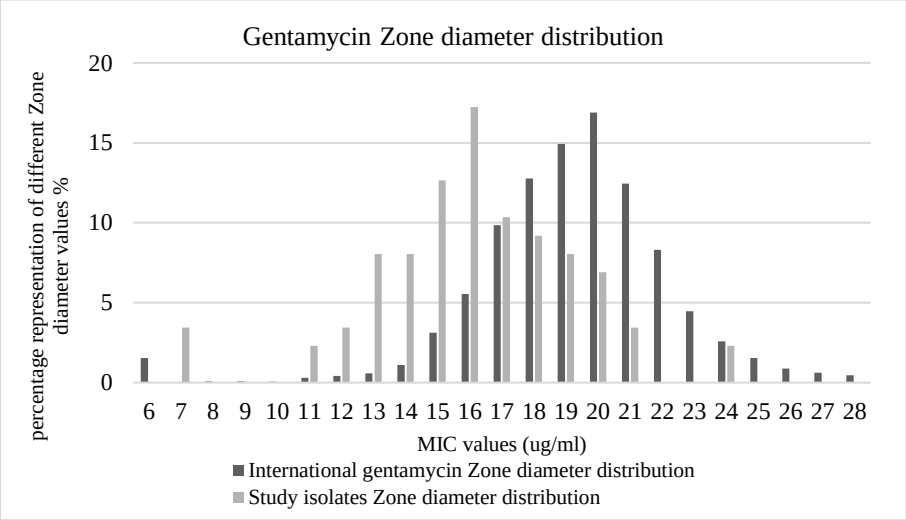


Figure 2.9. Distribution of gentamycin susceptibility among international Ps. aeruginosa population and tested isolates (Zone diameter)

2.4.3. Predicted antibiotic susceptibility using in-silico tools

WGS data were used to compare identified genotypic resistance with measured resistance and accuracy of prediction was evaluated. Resistance profiles identified to quinolones and aminoglycosides using the (RGI) tool of CARD database (McArthur *et al.*, 2013) are shown in

Table 2.1. Although the CARD database does not provide predicted e-antibiogram, identified elements of resistance (mutations and genes) do not correlate with or predict for the experimentally determined phenotype. Most of the known resistance genes and mutations were identified in nearly all isolates irrespective of their phenotype and were similarly identified in both resistant and sensitive isolates.

The creation and curation of a database that describes all point mutations associated with antibiotic resistance is considered one of the main challenges for culture-independent susceptibility testing (Gordon *et al.*, 2014) (Stoesser *et al.*, 2013). ResFinder cannot be used to address quinolone resistance because it can only identify acquired resistance elements and not nucleotide changes associated with resistance for *Ps. aeruginosa*. The species included in the platform identifying chromosomal mutations by ResFinder are *E. coli*, *Enterococcus faecalis*, *Enterococcus faecium*, *H. pylori*, *Klebsiella*, *M. tuberculosis*, *N. gonorrhoeae*, *Salmonella* and *Staphylococcus aureus*. In addition, it did not identify any of the acquired resistance genes that are associated with transferable quinolone resistance. ResFinder does not distinguish between specific agents within the same antibiotic class. The results reported here are therefore only for the aminoglycoside group (Table 2.2). Even then it can only identify the presence or absence of three aminoglycosides resistance genes: *aph* (3')-IIb, *aad*A13 and *ant* (2")-Ia. The acquired aminoglycosides resistance gene *aph* (3')-IIb which confers resistance to some aminoglycosides but not to amikacin was identified in almost all studied isolates (86 isolates) except for one susceptible isolate which lacked the gene. This means that the gene cannot differentiate resistance from susceptibility because it has been identified in 42 gentamycin resistant isolates and in 44 gentamycin susceptible isolates. The two other genes for acquired aminoglycosides resistance; *aad*A13 and *ant* (2")-Ia were absent in all 86 isolates except for one susceptible isolate that showed presence of the gene *aad*A13. Acquired resistance genes as distributed in the studied set of isolates do not appear to be the main underlying basis of aminoglycosides resistance or at least they do not appear to differentiate susceptibility from resistance for both studied agents (gentamycin and amikacin).

MICRA is another publicly available automatic pipeline for microbial identification and characterization (Caboche *et al.*, 2017). The antibiotic resistance computational post-analysis module for prediction was used. The predicted e-antibiograms are shown in (Table 2).

MICRA predicts the e-antibiogram by identifying 'susceptibility genes' through BLASTing the query sequences against a local version of Drug Bank (Wishart *et al.*, 2006). It also identifies 'resistance genes' by BLASTing the sequences against a modified version of the antibiotic resistance database (ARDB) (Liu and Pop, 2009). In that respect, MICRA has an added function of identifying 'susceptibility elements' in addition to 'resistance elements' that can be identified using the other tools. The final MICRA in-silico report showed details for the complete set of genes identified as match from ARDB in addition to matches from Drug Bank. All study isolates were identified as 'not susceptible' as shown in (Table 2.2) because all had matches from the ARDB and no matches from Drug Bank. In addition, we tested for gentamycin and norfloxacin *in-silico* susceptibility using MICRA, but the final report showed the result as 'not defined'. This means that prediction data may not be available for all agents even from the same antibiotic class.

Table 2.1. Resistance profiles identified using the resistance gene Identifier tool by CARD database

Resistance genes or mutations	***mexA***	***mexB***	***mexC***	***mexD***	***mexE***	***mexF***	***mexG***	***mexH***	***mexI***	***mexR***	***mexS***	***mexT***	***mexV***	***mexW***	***mexY***	***mexZ***	***soxR***	***cpxR***	***emrE***	***oprM***	***oprN***	***oprJ***	***nalD***	***nalC* S209R**	***nalC* G71E**	***nalC* A186T**	***nfxB***	**APH (3')-IIb**	***opmD***	***armR***	***parE* A473V**	***gyrA* T83I**	***Genomic profile***
Antibiotic class tested	**Q**	**Q**	**Q-AG**	**Q-AG**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q**	**Q-AG**	**Q-AG**	**Q**	**Q-AG**	**AG**	**Q-AG**	**Q**	**Q-AG**	**Q**	**Q**	**Q**	**Q**	**Q-AG**	**AG**	**Q**	**Q**	**Q**	**Q**	
PAE0002	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■				1
PAE0005	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■		11
PAE0006	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■				1
PAE0007	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■			4
PAE0008	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■		■		12
PAE0010	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■	■	■	■	■	■	■		■	■	■	■		■	13
PAE0011	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■			4
PAE0012	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■				8
PAE0014	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■	■	■				9
PAE0018	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■		■	■	■				2
PAE0020	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■		■	■	■				2
PAE0021	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■				1
PAE0024	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			5
PAE0025	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■			6
PAE0026	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■			14
PAE0029	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			15
PAE0030	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			4
PAE0032	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■				1
PAE0035	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■				1
PAE0036	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■				1
PAE0039	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■				8

Resistance genes or mutations	*mex*A	*mex*B	*mex*C	*mex*D	*mex*E	*mex*F	*mex*G	*mex*H	*mex*I	*mex*R	*mex*S	*mex*T	*mex*V	*mex*W	*mex*Y	*mex*Z	*sox*R	*cpx*R	*emr*E	*opr*M	*opr*N	*opr*J	*nal*D	*nal*C S209R	*nal*C G71E	*nal*C A186T	*nfx*B	APH (3')-IIb	*opm*D	*arm*R	*par*E A473V	*gyr*A T83I	*Genomic profile*
PAE0040																																	1
PAE0041																																	8
PAE0043																																	9
PAE0044																																	1
PAE0046																																	9
PAE0047																																	1
PAE0048																																	1
PAE0051																																	1
PAE0055																																	2
PAE0057																																	17
PAE0059																																	4
PAE0060																																	1
PAE0061																																	2
PAE0062																																	3
PAE0063																																	9
PAE0068																																	9
PAE0070																																	10
PAE0071																																	2
PAE0072																																	1
PAE0076																																	1
PAE0079																																	3
PAE0081																																	18
PAE0082																																	1
PAE0083																																	2
PAE0084																																	1
PAE0085																																	10
PAE0087																																	1
PAE0088																																	1
PAE0090																																	16
PAE0091																																	1
PAE0092																																	1
PAE0093																																	9
PAE0095																																	19
PAE0096																																	1
PAE0098																																	1
PAE0099																																	4
PAE0106																																	1
PAE0109																																	7
PAE0111																																	2
PAE0112																																	1
PAE0113																																	1
PAE0115																																	1
PAE0116																																	1
PAE0119																																	1
PAE0124																																	1
PAE0125																																	8
PAE0136																																	1
PAE0142																																	4
PAE0144																																	1
PAE0145																																	1
PAE0147																																	2
PAE0148																																	1

Resistance genes or mutations	*mex*A	*mex*B	*mex*C	*mex*D	*mex*E	*mex*F	*mex*G	*mex*H	*mex*I	*mex*R	*mex*S	*mex*T	*mex*V	*mex*W	*mex*Y	*mex*Z	*sox*R	*cpx*R	*emr*E	*opr*M	*opr*N	*opr*J	*nal*D	*nal*C S209R	*nal*C G71E	*nal*C A186T	*rfx*B	APH (3')-IIb	*opm*D	*arm*R	*par*E A473V	*gyr*A T83I	*Genomic profile*
PAE0149																																	2
PAE0151																																	2
PAE0154																																	1
PAE0156																																	1
PAE0157																																	2
PAE0158																																	1
PAE0160																																	8
PAE0161																																	4
PAE0167																																	1
PAE0168																																	1
PAE0171																																	1
PAE0172																																	14
PAE0174																																	3
PAE0175																																	1

Grey cell: gene present, white cell: gene absent Q: Quinolone, AG: Aminoglycoside

Table 2.2. Results of susceptibility testing using both MICRA and ResFinder compared to measured susceptibility using both zone inhibition diameter and MIC

Tested isolate ID	MICRA prediction			Res finder prediction			Measured Phenotypic behavior															
	ciprofloxacin	levofloxacin	amikacin	*aph* (3')-IIb	*aad*A13	*ant* (2'')-Ia	Gentamycin zone	zone breakpoint	gentamycin MIC	MIC breakpoint	levofloxacin zone	zone breakpoint	levofloxacin MIC	MIC breakpoint	Ciprofloxacin zone	zone breakpoint	ciprofloxacin MIC	MIC breakpoint	amikacin zone	zone breakpoint	amikacin MIC	MIC breakpoint
PAE0002	NS	NS	NS				8	R	14	R	24	S	1	S	28	S	0.1	S	21	S	4	S
PAE0005	NS	NS	NS				8	R	14	R	21	R	1	S	25	R	0.1	S	18	S	4	S
PAE0006	NS	NS	NS				16	R	18	S	23	S	1	S	30	S	0.1	S	23	S	4	S
PAE0007	NS	NS	NS				4	S	16	S	21	R	1	S	30	S	0.1	S	19	S	2	S
PAE0008	NS	NS	NS				4	S	19	S	27	S	0.5	S	33	S	0.1	S	23	S	4	S
PAE0010	NS	NS	NS				16	R	14	R	6	R	>32	R	6	R	8	R	17	I	8	S
PAE0011	NS	NS	NS				8	R	17	S	28	S	0.25	S	35	S	0.1	S	22	S	4	S
PAE0012	NS	NS	NS				8	R	15	S	15	R	2	R	22	R	0.3	S	19	S	4	S
PAE0014	NS	NS	NS				8	R	14	R	20	R	1	S	25	R	0.1	S	17	I	4	S
PAE0018	NS	NS	NS				1	S	19	S	23	S	0.5	S	31	S	0.1	S	22	S	2	S
PAE0020	NS	NS	NS				4	S	15	S	23	S	0.5	S	28	S	0.1	S	19	S	2	S
PAE0021	NS	NS	NS				4	S	16	S	23	S	0.5	S	31	S	0.1	S	21	S	2	S
PAE0024	NS	NS	NS				16	R	12	R	12	R	4	R	20	R	0.5	S	15	I	8	S
PAE0025	NS	NS	NS				4	S	17	S	27	S	0.5	S	33	S	0.1	S	23	S	2	S
PAE0026	NS	NS	NS				4	S	17	S	23	S	1	S	28	S	0.3	S	20	S	4	S
PAE0029	NS	NS	NS				8	R	15	S	21	R	0.5	S	27	S	0.1	S	18	S	4	S
PAE0030	NS	NS	NS				4	S	16	S	24	S	1	S	31	S	0.1	S	22	S	2	S
PAE0032	NS	NS	NS				8	R	16	S	27	S	0.5	S	34	S	0.1	S	20	S	4	S
PAE0035	NS	NS	NS				16	R	11	R	17	R	2	R	24	R	0.5	S	15	R	16	I
PAE0036	NS	NS	NS				4	S	18	S	19	R	2	R	27	S	0.3	S	22	S	2	S
PAE0039	NS	NS	NS				8	R	15	S	21	R	1	S	27	S	0.1	S	19	S	4	S
PAE0040	NS	NS	NS				4	S	17	S	23	S	0.5	S	31	S	0.1	S	20	S	4	S

PAE0041	NS	NS	NS				1	S	21	S	27	S	0.5	S	3	S	0.1	S	25	S	0.5	S
PAE0043	NS	NS	NS				2	S	19	S	26	S	0.5	S	31	S	0.1	S	23	S	2	S
PAE0044	NS	NS	NS				8	R	15	S	22	S	0.5	S	25	R	0.1	S	21	S	4	S
PAE0046	NS	NS	NS				16	R	12	R	17	R	4	R	26	S	0.3	S	15	I	8	S
PAE0047	NS	NS	NS				8	R	17	S	20	R	1	S	21	R	0.1	S	17	I	4	S
PAE0048	NS	NS	NS				8	R	14	R	20	R	2	R	32	S	0.1	S	21	S	4	S
PAE0051	NS	NS	NS				8	R	14	R	21	R	0.5	S	27	S	0.1	S	18	I	4	S
PAE0055	NS	NS	NS				8	R	15	S	22	S	1	S	30	S	0.3	S	20	S	4	S
PAE0057	NS	NS	NS				1	S	20	S	25	S	0.5	S	32	S	0.1	S	23	S	2	S
PAE0059	NS	NS	NS				16	R	13	R	13	R	4	R	22	R	0.5	S	16	I	8	S
PAE0060	NS	NS	NS				0.5	S	20	S	24	S	0.5	S	32	S	0.1	S	23	S	2	S
PAE0061	NS	NS	NS				16	R	11	R	14	R	4	R	24	R	0.5	S	16	I	8	S
PAE0062	NS	NS	NS				2	S	19	S	30	S	0.5	S	32	S	0.1	S	24	S	1	S
PAE0063	NS	NS	NS				8	R	15	S	18	R	4	R	27	S	0.3	S	19	S	4	S
PAE0068	NS	NS	NS				4	S	17	S	31	S	0.25	S	34	S	0	S	22	S	4	S
PAE0070	NS	NS	NS				2	S	18	S	21	R	1	S	32	S	0.1	S	23	S	2	S
PAE0071	NS	NS	NS				4	S	16	S	16	R	2	R	28	S	0.3	S	21	S	4	S
PAE0072	NS	NS	NS				8	R	16	S	24	S	0.5	S	30	S	0.1	S	21	S	2	S
PAE0076	NS	NS	NS				2	S	18	S	25	S	0.5	S	31	S	0.1	S	23	S	2	S
PAE0079	NS	NS	NS				4	S	17	S	15	R	2	R	21	R	0.5	S	17	I	4	S
PAE0081	NS	NS	NS				0.5	S	21	S	26	S	0.25	S	34	S	0.1	S	24	S	1	S
PAE0082	NS	NS	NS				4	S	16	S	28	S	0.25	S	33	S	0.1	S	20	S	2	S
PAE0083	NS	NS	NS				4	S	16	S	21	R	0.5	S	29	S	0.1	S	21	S	2	S
PAE0084	NS	NS	NS				2	S	20	S	27	S	0.5	S	35	S	0.1	S	24	S	1	S
PAE0085	NS	NS	NS				8	R	15	S	21	R	0.5	S	27	S	0.1	S	19	S	4	S
PAE0087	NS	NS	NS				4	S	16	S	15	R	4	R	21	R	0.5	S	21	S	4	S
PAE0088	NS	NS	NS				16	R	12	R	32	S	0.12	S	32	S	0	S	17	I	8	S
PAE0090	NS	NS	NS				2	S	19	S	23	S	0.5	S	32	S	0.1	S	22	S	2	S
PAE0091	NS	NS	NS				8	R	15	S	21	R	1	S	25	R	0.1	S	19	S	4	S
PAE0092	NS	NS	NS				2	S	20	S	20	R	1	S	31	S	0.1	S	22	S	2	S
PAE0093	NS	NS	NS				4	S	18	S	31	S	0.25	S	39	S	0.1	S	21	S	4	S
PAE0095	NS	NS	NS				0.25	S	24	S	14	R	4	R	25	R	0.5	S	26	S	0.5	S
PAE0096	NS	NS	NS				4	S	17	S	25	S	0.5	S	31	S	0.1	S	22	S	2	S
PAE0098	NS	NS	NS				2	S	18	S	26	S	0.5	S	35	S	0.1	S	23	S	1	S
PAE0099	NS	NS	NS				1	S	20	S	24	S	0.5	S	33	S	0.1	S	24	S	1	S
PAE0106	NS	NS	NS				2	S	20	S	25	S	0.5	S	35	S	0.1	S	25	S	2	S
PAE0109	NS	NS	NS				8	R	13	R	21	R	0.5	S	24	R	0.1	S	18	S	4	S
PAE0111	NS	NS	NS				8	R	16	S	13	R	16	R	15	R	2	R	19	S	4	S
PAE0112	NS	NS	NS				4	S	17	S	23	S	1	S	31	S	0.1	S	21	S	2	S
PAE0113	NS	NS	NS				8	R	16	S	23	S	1	S	31	S	0.1	S	22	S	2	S
PAE0115	NS	NS	NS				2	S	18	S	14	R	2	R	24	R	0.3	S	23	S	2	S
PAE0116	NS	NS	NS				2	S	19	S	26	S	0.5	S	32	S	0.1	S	23	S	1	S
PAE0119	NS	NS	NS				0.5	S	21	S	25	S	0.5	S	35	S	0.1	S	25	S	0.5	S
PAE0124	NS	NS	NS				16	R	13	R	16	R	4	R	23	R	0.5	S	17	I	8	S
PAE0125	NS	NS	NS				8	R	16	S	24	S	0.5	S	30	S	0.1	S	21	S	4	S
PAE0136	NS	NS	NS				8	R	16	S	29	S	0.25	S	32	S	0.1	S	20	S	4	S
PAE0142	NS	NS	NS				16	R	13	R	19	R	2	R	26	S	0.3	S	18	I	4	S
PAE0144	NS	NS	NS				8	R	14	R	23	S	0.5	S	30	S	0.1	S	19	S	4	S
PAE0145	NS	NS	NS				16	R	14	R	22	S	1	S	28	S	0.1	S	19	S	4	S
PAE0147	NS	NS	NS				4	S	17	S	26	S	0.5	S	34	S	0.1	S	21	S	4	S
PAE0148	NS	NS	NS				4	S	16	S	21	R	1	S	30	S	0.3	S	21	S	4	S
PAE0149	NS	NS	NS				0.25	S	22	S	28	S	0.25	S	34	S	0.1	S	26	S	0.5	S
PAE0151	NS	NS	NS				0.25	S	24	S	31	S	0.06	S	32	S	0	S	27	S	0.5	S
PAE0154	NS	NS	NS				8	R	15	S	22	R	2	R	27	S	0.3	S	20	S	8	S
PAE0156	NS	NS	NS				16	R	15	S	22	S	1	S	28	S	0.3	S	19	S	4	S
PAE0157	NS	NS	NS				32	R	7	R	25	S	1	S	32	S	0.1	S	8	R	32	R
PAE0158	NS	NS	NS				4	S	18	S	16	R	4	R	25	R	0.5	S	22	S	2	S
PAE0160	NS	NS	NS				8	R	13	R	18	R	2	R	25	R	0.3	S	18	S	4	S
PAE0161	NS	NS	NS				1	S	19	S	25	S	0.5	S	32	S	0.1	S	23	S	1	S

PAE0167	NS	NS	NS				16	R	13	R	20	R	1	S	26	S	0.3	S	17	I	8	S
PAE0168	NS	NS	NS				64	R	7	R	23	S	1	S	31	S	0.3	S	11	R	32	R
PAE0171	NS	NS	NS				16	R	13	R	18	R	1	S	25	R	0.3	S	17	I	8	S
PAE0172	NS	NS	NS				4	S	16	S	25	S	0.5	S	33	S	0.1	S	21	S	4	S
PAE0174	NS	NS	NS				1	S	19	S	20	R	1	S	25	R	0.1	S	20	S	2	S
PAE0175	NS	NS	NS				64	R	7	R	21	R	1	S	26	S	0.3	S	11	R	32	R

NS: Not sensitive, R: resistant, S: sensitive, I: intermediate, MIC values are shown in mg/L, Zone diameters are shown in mm.

Grey cell: gene present, white cell: gene absent

2.4.4. Comparing predictions to measured phenotypes

Sensitivity, specificity, and predictive values were used to evaluate accuracy of prediction. The results are shown in Table 2.3 and 2.4. Predictive values presented are rounded to the nearest percent. Only three genes identified using CARD showed different distribution among isolates. These were *mex*Y, *mex*F, and *arm*R. The other twenty resistance genes were identified in all eighty-seven study isolates as shown in (Table 2.1). Five of the previously known quinolone resistances associated mutations were also identified. Genes and mutations showing different distributions among isolates were evaluated. This is shown in Table 2.3.

To evaluate the performance of the set of genes/gene variants included in ResFinder and CARD as predictors to phenotypic resistance in the studied quinolones and aminoglycosides agents, repeated genetic profiles shown in last column of Table 2.1 are assessed here. This was done because no one set of predictors is seen with all tested isolates, instead profiles of different combinations are observed. In addition, ResFinder and CARD do not give a final predicted resistance/susceptibility (positive/negative) output to evaluate. Assessment of the observed profiles based on the known set of markers (as extracted from ResFinder and CARD) is shown in Table 2.5. Profiles that occurred only once or twice were not assessed. The MICRA-predicted e-antibiogram was compared to the experimentally determined phenotype and the results are shown in Table 2.4. MICRA predicted all isolates as ‘not susceptible’, so only sensitivity and positive predictive values can be calculated because there are no true negative cases. The MICRA prediction output gave a detailed table (in HTML format) showing antibiotic susceptibility and resistance prediction for each tested isolate (tables are not shown here). This table showed a detailed list of all hits the pipeline has identified for each drug by BLASTing the query sequences against modified local versions for both ARDB and Drug bank as well as the detailed BLAST results. The tool then identified its final prediction based on this detailed report for each isolate and this final prediction is shown here (Table 2.2). The MICRA pipeline

validation has been performed with different organisms including *E. coli str. K12 substr. DH10B, Bordetella pertussis*, *E. coli* O104:H4, *Staphylococcus aureus*, and *Clostridium autoethanogenum*. However, the authors and developers of the tool did not show that the tool application is restricted to certain species or to a specific spectrum of organisms. There did not seem to be any issues related to the prediction module functionalities because a detailed output table showing the hits for all the tested sequences was generated using the tool. A possible reason for the false prediction (generated using the tool) in susceptible isolates perhaps originates from the databases that the tool uses to retrieve "antibiotic resistance genes" hits. The set of genes retrieved from the commonly used databases and especially for (ARDB) appear to be not comprehensive enough to predict and differentiate the antibiotic resistance/susceptibility phenotype. This can be essentially related to the nature of *Ps. aeruginosa* species that carry many resistant determinants even in wild type strains and clinically susceptible isolates. It seems that these resistance determinants are not sufficient to predict clinical resistance, and this is the main topic that needs more evaluation and understanding. Although MICRA identified all resistant strains correctly (high sensitivity) false positive rate (Type 1 error) was high as it identified 'phenotypically susceptible' as 'not susceptible' (Table 2.4). This is a situation that we need to minimize because the cost of false positives can lead to antibiotic over prescription.

Table 2.3. Performance of some known genomic markers (genes-mutations) identified by CARD database for in-silico antibiotic resistance prediction

Genomic marker	**Sensitivity**		**Specificity**		**Positive predictive value**		**Negative predictive value**	
	Ciprofloxacin zone	Levofloxacin zone	Ciprofloxacin zone	Levofloxacin zone	Ciprofloxacin zone	Levofloxacin zone	Ciprofloxacin zone	Levofloxacin zone
*par*E A473V	5%	3%	**97%**	**96%**	33%	33%	75%	55%
*gyr*A T83I	9%	5%	**100%**	**100%**	**100%**	**100%**	77%	57%
*nal*C S209R	**82%**	77%	25%	23%	27%	45%	**80%**	55%
*nal*C G71E	**86%**	**90%**	8%	8%	24%	44%	63%	50%
*nal*C A186T	14%	10%	**94%**	**94%**	43%	57%	76%	56%
*mex*Y	73%	80%	19%	21%	23%	45%	67%	56%
*mex*F	73%	74%	14%	10%	22%	40%	60%	33%
*arm*R	23%	21%	**83%**	**83%**	31%	50%	76%	56%

Table 2.4. Evaluation of post-analysis module of MICRA pipeline as an in-silico antibiotic resistance prediction tool

Measured phenotype as determined by	MICRA prediction		
	Sensitivity	**Positive Predictive Value**	**False positive rate (Type 1 error)**
Ciprofloxacin zone inhibition diameter	**100%**	25%	75%
Ciprofloxacin MIC	**100%**	2%	98%
Levofloxacin zone inhibition diameter	**100%**	45%	55%
Levofloxacin MIC	**100%**	24%	76%
Amikacin zone inhibition diameter	**100%**	21%	79%
Amikacin MIC	**100%**	5%	95%

Table 2.5. Evaluation of the genetic profiles identified using ResFinder and CARD as possible resistance predictors

ResFinder and CARD predictor genetic profiles	Parameters of performance	Phenotype under evaluation							
		Ciprofloxacin zone inhibition diameter	Ciprofloxacin MIC	Levofloxacin zone inhibition diameter	Levofloxacin MIC	Amikacin zone inhibition diameter	Amikacin MIC	Gentamycin zone inhibition diameter	Gentamycin MIC
Genetic profile 1	Sensitivity	45.5%	-----	43.6%	38.1%	75%	66.7%	54.5%	54.8%
	Specificity	52.3%	51.8%	50%	50%	55.1%	54.2%	55.6%	59.1%
	PPV	24.4%	----	41.5%	19.5%	8.8%	5%	30%	56.1%
	NPV	73.9%	95.7%	52.2%	71.7%	97.4%	97.8%	77.8%	57.8%
Genetic profile 2	Sensitivity	9.1%	50%	10.3%	14.3%	25%	33.3%	9.1%	9.5%
	Specificity	86.2%	88.2%	85.4%	87.9%	87%	88%	85.7%	84.1%
	PPV	18.2%	9.1%	36.4%	27.3%	10%	9.1%	18.2%	36.4%
	NPV	73.7%	98.7%	53.9%	76.3%	95.2%	97.3%	73%	49.3%
Genetic profile 3	Sensitivity	9.1%	-----	5.1%	4.8%	-----	-----	-------	------
	Specificity	98.5%	96.5%	97.9%	97%	97.1%	96.4%	96.8%	95.5%
	PPV	66.7%	-----	66.7%	33.3%	-----	-----	-------	-------
	NPV	76.2%	97.6%	56%	76.2%	94.4%	96.4%	73.5%	50%

Genetic profile 4	Sensitivity	4.5%	-------	7.7%	9.55%	-------	-----	9.1%	7.1%
	Specificity	90.8%	91.8%	91.7%	92.4%	92.8%	91.6%	92.1%	90.9%
	PPV	14.3%	-------	42.9%	28.6%	--------	------	28.6%	42.9%
	NPV	73.8%	97.5%	55%	76.3%	94.1%	96.2%	74.4%	50.6%
Genetic profile 8	Sensitivity	9.1%	------	7.7%	9.5%	-------	------	4.5%	9.5%
	Specificity	95.4%	94.1%	95.8%	95.5%	92.8%	94%	93.7%	97.7%
	PPV	40%	-------	60%	40%	-------	------	20%	80%
	NPV	75.6%	97.6%	56.1%	76.8%	94.1%	96.3%	73.8	53.1%
Genetic profile 9	Sensitivity	-------	-------	5.1%	9.5%	-------	-----	4.5%	4.8%
	Specificity	92.3%	94.1%	93.8%	95.5%	94.2%	94%	93.7%	93.2%
	PPV	-------	-------	40%	40%	--------	-----	20%	40%
	NPV	73.2%	97.6%	54.9%	76.8%	94.2%	96.3%	73.8%	50.6%

2.5. Discussion

The ability to predict antibiotic resistance phenotypes from genotypes is still in development. It was reported that the presence or absence of AMR genes does not always guarantee a respective phenotypic profile in all bacterial species (Rossen *et al.*, 2018), and this is born-out in this study. While this can be established for some species, it is more challenging for *Ps. aeruginosa* as will be discussed in more detail. Consequently, it is not reliable to build a predicted e-antibiogram that can be used in the clinic based on the identification of known resistance genes and mutations for these antibiotics because there are clearly many other unknown determinants of resistance that are still inadequately understood. In addition, tools used for *in silico* prediction depend on either BLAST or reference mapping to known AMR genetic determinants without considering the background gene pool or other cellular mechanisms that contribute to resistance. These may include modification of endogenous genes by insertion sequences, regulatory or post-translational modifications, and alteration of gene expression levels which can all play a combinatorial role in determining susceptibility. Based on the results of this chapter, the set of known resistance determinants is currently insufficient to predict for the net behavior of resistance to aminoglycosides and quinolones groups of antibiotics in *Ps. aeruginosa*.

Acceptable standards for positive and negative predictive values that can translate into changes in patient management are still lacking (Caliendo *et al.*, 2013). This makes it important to consider all the parameters that can be used to assess accuracy of the diagnostic test on a setting-

relevant basis and according to the condition under consideration. Some studies have compared sequence-based prediction to phenotypic resistance in different bacterial pathogens including *Staph. aureus* (Gordon *et al.*, 2014) (Köser, Holden, *et al.*, 2012), *E. coli*, and *K. pneumoniae* (Stoesser *et al.*, 2013) and have shown high predictive values and high concordance between *in silico* prediction and measured phenotype. However, validation studies and more extensive evaluation using larger and diverse sets of isolates are still needed before it can be reliably used to support clinical decision making more widely (Stoesser *et al.*, 2013) (Köser, Holden, *et al.*, 2012). It has also been shown in several later studies that antimicrobial resistance can be accurately predicted in some bacterial species based on genome-derived sequence information (Bradley *et al.*, 2015) (Juarez *et al.*, 2018).

Ps. aeruginosa is an environmental pathogen with extensive metabolic adaptability and is considered a potentially challenging species for antimicrobial resistance prediction using genomics-based approaches. Even antibiotic-susceptible strains of *Ps. aeruginosa* have natural defenses due to intrinsic AMR (Julie Jeukens *et al.*, 2017). The organism poses a specific challenge in carrying a very large pool of genes encoding transcriptional regulators and two-component regulatory systems (Silby *et al.*, 2011). The mosaic structure of *Ps. aeruginosa* chromosome which can be continuously modified by acquisition of new DNA, larger or smaller mutational deletion events, mutations, and chromosomal inversions can continuously modify the phenotype (Klockgether *et al.*, 2011). Although the prediction of phenotypic susceptibility using a pre-defined set of genomic markers might give high sensitivity and specificity for some bacterial species (M. S. Wright *et al.*, 2015) (Gupta *et al.*, 2014b), this can be a major challenge when applied to *Ps. aeruginosa*. To the best of our knowledge, there are no studies available to date that evaluate the accuracy of WGS-based prediction for this organism. In a recent review, Ellington *et al.*, (2017) concluded that there is not sufficient evidence for many bacterial species to support the use of sequence-inferred antibiotic susceptibility in clinical practice. A conclusion that the current study would support. Many sources of error can be encountered when WGS is used for prediction of antibiotic resistance which often leads to lack of concordance between phenotypic AMR and genotypic AMR (Ellington *et al.*, 2017). A more recent study investigating the use of machine learning approach in combination with transcriptomics for predicting antibiotic resistance in *Ps. aeruginosa* has shown that gene expression information can improve the diagnostic performance for all drugs tested except for ciprofloxacin (Khaledi *et al.*, 2020). A

major drawback of the currently available genomic-based tools is that the set of markers used as determinants to differentiate susceptibility from resistance appears to be very limited. In most cases these are gene presence or absence information. An expanded set of markers including different gene variants is needed to give more complete understanding and prediction of phenotype. Although detailed investigation of the genetic elements underlying resistance is considered a promising step towards building better diagnostics, combining transcriptomic profiles with genomic-based resistance profiles have generated improved predictive models for antibiotic resistance in *Ps. aeruginosa* for tobramycin, ceftazidime, and meropenem but not for ciprofloxacin (Khaledi *et al.*, 2020). However, some transcriptional responses due to mutations in some gene expression negative regulators have shown to be fixed in clinical *Ps. aeruginosa* isolates (Frimodt-Møller *et al.*, 2018). This consequently mean that although beneficial, combining both quantitative gene expression information and genome information may not reliably predict the corresponding phenotype in all cases.

It is concluded that genome-sequencing can provide rich source of information on the correlation between genotype and phenotype, however, more research is still needed to find a reliable approach that can be converted into a practical and informative tool within the time frame of the clinic and the routine health care needs. To do that, predictive interpretive tools need to be available to interpret sequencing-derived information for the bacterial species addressed. Such tools need to be highly accurate, rapid, comprehensive, informative, and capable of providing clinically trusted interpretations (Judge *et al.*, 2015). More data and validation studies are still needed to test for the reliability and practicality of the currently available tools before they can be used in clinical practice to guide patient interventions. This is important to avoid misuse of antibiotics and over-prescription of unnecessary medications that can lead to loss of the current available treatment options. In order to transform the use of NGS technologies from a research tool into a clinically useful tool that can be used in routine clinical practice and diagnostic setting, several important requirements need to be fulfilled.

2.6. Conclusions: Requirements for the implementation of sequencing technologies in clinical microbiology laboratories to predict resistance and to guide interventions

- The need to improve sample preparation techniques that enable sequencing from primary isolation plates and single colonies without the need to subculture.

- The need to have knowledge about the population structure and the biologic background context for significant pathogenic species and lineages at national and individual patient specimen levels.
- The need to develop central database for comparison of sequence data generated in real time with the previously available local, national, and global isolates' data.
- The need to develop a system for automated sequence data interpretation. This will help to decrease turnaround times and to disseminate the use of the technology to settings with less expertise.
- The need for information technology infrastructure for data storage and exchange.
- The need for expertise in technical issues of sequencing in addition to collaboration with specialist bioinformaticians and analysts' teams.
- The need for automated data interpretation into real time meaningful clinical reports easily understood by clinicians and health care providers.

Different genome sequencing technologies can provide clinically relevant data within a time frame that would influence patient care. It has the potential to replace multiple diagnostic and reference tests. Genome sequencing can make important contributions to infection-control investigations and practice and can also impact different other aspects of clinical practice. This can subsequently reduce infection-associated morbidity and cost. It can confirm the outbreak close to its start point and can also draw the link between outbreak cases and the community. It can detect genotypes and other information related to important virulence genes and mutations in a short time with greater discriminatory power (Besser *et al.*, 2018) .

However, it is still considered relatively expensive and technically difficult. Using the technology has some practical limitations and is currently preserved to the research community. Despite that, it is expected that genome sequencing will revolutionize and replace many other sets of diagnostic techniques. This is attributable to the decline in the cost of sequencing machines and to the growing interest among clinical microbiologists in using the technology. The increasing availability of user-friendly analysis pipelines and workflows has greatly enforced that interest. In this chapter, some current limitations that exist to transform sequencing technologies from a research tool into a clinically useful tool have been discussed suggesting a framework for implementation at the practical level. This can be achieved once a valid theory

becomes available on the genotypic-phenotypic correlation for antibiotic resistance behavior. The chapter that follows reviews the literature related to the molecular bases of quinolone and aminoglycoside resistance and establishes a framework for the detailed assessment of all known genes and gene variants that can act as molecular diagnostic markers.

Declaration of contribution:

I would like to thank prof. Nigel Saunders for giving me access to use the set of 162 completely sequenced *Ps. aeruginosa* isolates through Brunel Systems and Synthetic Biology lab resources. I would like also to thank Mr. Arshad Khan for performing genome assembly, quality checks and analysis and for running the diversity analysis and estimates of evolutionary divergence to select for the most diverse group of isolates as described in Methods sections 2.3.3 and 2.4.1.

Chapter 3. The role of known Quinolones and Aminoglycosides resistance mechanisms in explaining resistance in *Pseudomonas aeruginosa*

3.1. Molecular basis of quinolone resistance

3.1.1. Quinolone Resistance Determining Region (QRDR) Mutations

3.1.2. Efflux pumps' regulators and global transcription regulators

3.1.3. SOS response (stress)-mediated pathways

3.1.4. The importance of understanding the global picture

3.2. Molecular basis of aminoglycosides resistance

3.2.1. Aminoglycosides-Modifying Enzymes (enzymatically catalyzed antibiotic-inactivation)

3.2.2. Ribosomal Mutations or Modification

3.2.3. Cell membrane modifications (Changes in cellular permeability)

3.2.4. Efflux pumps and SOS response Mediated pathways

3.3. Objectives

3.4. Methodology: Assessment of the predictive potential of established resistance mechanisms and markers separately and in combination for prediction of ciprofloxacin, levofloxacin, gentamycin, and amikacin resistance/susceptibility phenotype

3.4.1. Literature search and identification of variants

3.4.2. Choosing a set of completely sequenced genomes for performing the analysis

3.4.3. Identifying the distribution and performance of each of the resistance/susceptibility-associated genes and mutations

3.4.4. Finding combinations with improved performance

3.5. Results Section 1 (ciprofloxacin)

3.5.1. Description of mutations distribution

3.5.2. Predictive values and measures of diagnostic accuracy for known resistance-associated genes and mutations

3.5.3. Finding the best possible predictor combinations

3.1. Molecular Basis of Quinolone Resistance

3.1.1. Quinolone Resistance Determining Region (QRDR) mutations

Quinolone resistance determining region or "QRDR" is the name given to key sites of the enzymes acting as targets for quinolones action. Mutations affecting these key sites lead to quinolone resistance. Four genes commonly show theses mutations including genes encoding DNA gyrase (*gyr*A and *gyr*B) and/or topisomerase IV (*par*C and *par*E) (Correia *et al.*, 2017). These mutations lead to decreasing the binding affinity of quinolones to their targets (Yoshida *et al.*, 1990). It has been shown in different studies that these mutations can show variable effects on different quinolone agents' resistance in different organisms (Piddock, 1999) (Kaatz, Seo and Foster, 1999) (Komp Lindgren *et al.*, 2005) (Morgan-Linnell and Zechiedrich, 2007) (Hooper, 1999) (Hooper, 2000) (Everett *et al.*, 1996).

A study including 100 clinical isolates of *Ps. aeruginosa* concluded that mutations in QRDR adds to pre-existing specific resistance levels of isolates including efflux pumps overexpression. The study showed that these mutations can separately contribute to low level resistance and may lead to higher-level resistance in an additive way when combined with other mechanisms of resistance (Bruchmann *et al.*, 2013). The additive effect has also been shown in multiple other studies (Hooper and Jacoby, 2015) (Rehman, W. Patrick and Lamont, 2019).

Although some other studies have shown that even a single amino acid alteration in *gyr*A can cause clinically important levels of resistance to fluoroquinolones in *Ps. aeruginosa* (Nouri *et al.*, 2016), comprehensive review of the literature does not support that. It appears from more careful analysis of studies investigating QRDR mutations and its relation to resistance that these mutations can be necessary but may not be sufficient in all cases to lead to a full-blown clinical resistance. Supplementary Table 5 summarizes the results of a review of research articles studying quinolone resistance and its relation to mutations in the QRDR; 1) (Matsumoto *et al.*, 2012), 2) (Salma *et al.*, 2013), 3) (Nouri *et al.*, 2016), 4) (Wang, Lee and Peng, 2014), 5) (Yonezawa *et al.*, 1995), 6) (Jalal *et al.*, 2000), 7) (Akasaka *et al.*, 2001), 8) (Nakano *et al.*, 1997), 9) (Mouneimné *et al.*, 1999).

3.1.2. Efflux pump regulators and global transcriptional regulators

Efflux pumps are membrane associated porin-regulated systems that enable resistance to a wide range of drugs with different structural and functional features. Four types of MDR efflux pumps (from the RND family) have been reported as the main contributing types to extruding a broad range of antibiotic agents in *Ps. aeruginosa* including quinolones. These efflux pumps show varying degrees of effect with the most important one being MexEF-OprN. Efflux pumps can either be constitutively expressed, such as (MexAB-OprM) and (MexXY-OprM); or inducible, such as (MexCD-OprJ) and (MexEF-OprN) (Yordanov and Strateva, 2009).

These four active efflux-pump systems can be responsible for increasing resistance to quinolones agents which can vary in the range of 2-fold up to 16-fold (Mima *et al.*, 2005). Mutations in different efflux pump-associated regulatory genes can induce pump over-expression and consequently increase resistance.

A) <u>Mex AB-OprM</u>

This efflux pump was the first one to be described in *Ps. aeruginosa*. It is described to target a wide range of drug classes as well as other dyes and chemicals. It is constitutively expressed in wild-type organisms and is considered a contributor to intrinsic resistance to a wide range of antimicrobial agents (Zhao *et al.*, 1998). In addition, the system can be hyper-expressed in different types of mutants affecting its regulatory elements. *nal*B mutant is the first type of regulator mutant detected in some clinical isolates and can also be selected through *in vivo* and *in vitro* treatment by quinolones. The *nal*B mutant affects the *mex*R gene which is a pump repressor located upstream of MexAB-OprM (Srikumar, Paul and Poole, 2000). Strains with the *nal*B phenotype also occur as a result of total absence of *mex*R, decreased amounts of MexR or normal amounts with loss of its ability to dimerize (Boutoille *et al.*, 2004). *nal*C and *nal*D were also linked to MexAB-OprM hyper-expression. The protein products of these two genes are considered as negative regulators for the expression of MexAB-OprM operon. (Srikumar, Paul and Poole, 2000). Different types of mutations in *nal*C and *nal*D genes showed association with different resistance phenotypes (Pan *et al.*, 2016). Mutations in *mex*R, *nal*C, *nal*D, and *arm*R that were reported in the literature as linked to increased expression of MexAB-OprM are summarized in Table 3.1

B) MexCD-OprJ

This type of efflux pump exhibits a high degree of sequence similarity to MexAB-OprM with a similar range of substrate activity. However, it is not expressed in wild type cells, so it is not considered as a contributor to intrinsic resistance of the organism (Srikumar, Li and Poole, 1997). The pump can also export a wide range of other substrates. Expression of MexCD-OprJ is controlled by *nfx*B which is a negative regulator of the pump (Poole and Srikumar, 2001). Mutations in *nfx*B can alter its repressor activity leading to overexpression of the pump which can significantly affect quinolone resistance. This results in increased MIC that is not interpreted by topisomerase mutations alone (Oh *et al.*, 2003). Two types of *nfx*B mutants were described including Type A which leads to a moderate level of efflux system expression and Type B with high level of expression (Masuda *et al.*, 1996). MexCD-OprJ efflux pump regulatory mutations reported in the literature are summarized in Table 3.1

C) Mex XY-OprM

This type of efflux pump has a narrower substrate specificity and its substrate profile includes fluoroquinolones, specific B-lactams, aminoglycosides, tetracyclines, chloramphenicol, and erythromycin (Schweizer, 2003). However, this efflux-pump is recognized as one of the primary determinants of aminoglycosides resistance (Nikaido and Pagès, 2012). The system is known to be expressed in wild type strains in the presence of aminoglycosides, tetracyclines, and erythromycin which makes it an element for intrinsic resistance for these agents in *Ps. aeruginosa*. On the other hand, its expression is not known to be induced by quinolones in wild type strains. However, mutants hyper-expressing the pump demonstrate enhanced resistance to fluoroquinolones (Masuda *et al.*, 2000). The system is under the control of the *mex*Z gene which encodes a MexXY repressor. *mex*Z deletion has been linked to increased MexXY transcription and consequently aminoglycosides, quinolones, and cefepime resistance (Hay *et al.*, 2013).

D) MexEF-OprN

The fourth type of efflux pump that shows significant relation to quinolone resistance is the MexEF-OprN. This type of efflux pump is quiescent in wild type strains and therefore is not considered a cause of intrinsic resistance. It is known to be under regulation of MexT and MexS in addition to other regulators. The *nfx*C type mutant is a multidrug resistant mutant type known

to be originally selected by fluoroquinolones (Köhler *et al.*, 1997). *nfx*-C type mutant in *Ps. aeruginosa* is a quinolone resistant mutant with over-expressed efflux-pump MexEF-OprN. This over-expression phenotype was linked to *mex*S mutation that makes it non-functional and consequently releasing its repression on MexT. Different types of mutations were linked to the *nfx*C type mutant including mutations in *mex*S, *mex*T, *mva*T, *amp*R, and *mxt*R genes (Llanes *et al.*, 2011). MexEF-OprN over-expression mutant (*nfx*C phenotype) was observed to be prevalent in hospital settings where different types of *mex*S and *mex*T as well as other possible unexplored mutations are thought to contribute to the phenotype (Richardot *et al.*, 2016).

In addition, transcription of the *mex*EF-*opr*N operon has been shown to be regulated by both MexT (PA2492) positive regulator and MexS (PA2491) acting as a negative regulator (Köhler et al., 1999). The MexEF-OprN system of operon is known to be quiescent in wild-type cells. An active MexT transcriptional regulator is required to produce the pump MexEF-OprN with the presence of an upper negative regulator MexS that inactivate MexT in wild type strains and prevent pump expression. This occurs through two independent pathways (Uwate et al., 2013). MexEF-OprN efflux pump regulatory mutations are listed in Table 3.1

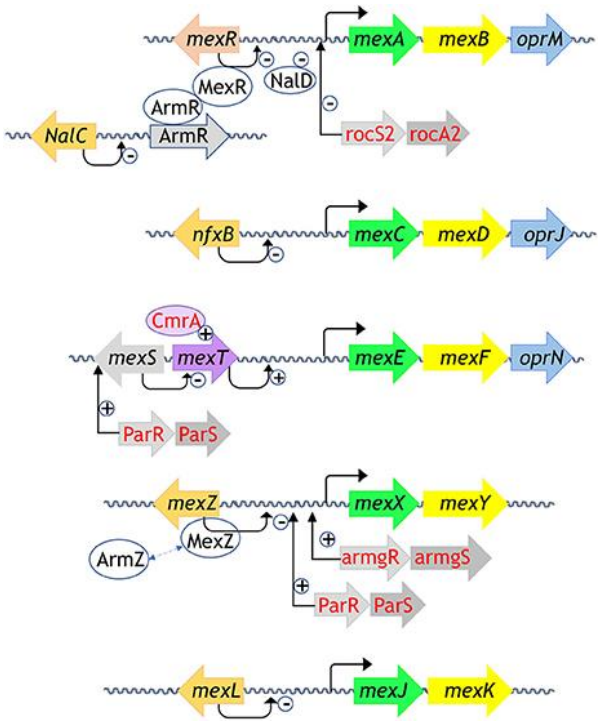
In addition to both *mex*S and *mex*T, *mva*T (PA4315) mutation was also shown to modify the expression of MexEF-OprN independent of the MexT or MexS pathway (Westfall et al., 2006). Mutations in genes that encode for the global regulators MvaT and AmpR have also been reported in other studies to activate MexEF-OprN operon in the *in vitro* mutants. It was reported that *nfx*C mutant overexpressing MexEF-OprN shows mutations in either *mex*T, *mex*S or *mva*T genes (Llanes et al., 2011). However, it was also shown more recently that there are still other unknown loci that appear to be implicated in the pump overproduction in clinical settings (Richardot et al., 2016).

E) Global transcriptional regulators

MvaT is a global transcriptional regulator that appears to be involved in several aspects of bacterial physiology in *Ps. aeruginosa*. It was originally described as a global regulator of virulence gene function in *Ps. aeruginosa* (Diggle *et al.*, 2002). *mva*T mutation has shown to modify MexEF-OprN efflux pump expression (Westfall *et al.*, 2006) and it was concluded that *mva*T mutants modify the resistance of *Ps. aeruginosa* to several antibiotics by increasing the expression of MexEF-OprN operon.

*amp*R is an additional non-specific transcriptional regulator showing more global roles in the regulation of several pathways in *Ps. aeruginosa*. It was established that *amp*R has a critical regulatory role in antibiotic resistance, virulence and general metabolism in *Ps. aeruginosa* (Balasubramanian et al., 2012). It is also involved in transcriptional regulation of hundreds of genes from diverse pathways (Balasubramanian, Kumari and Mathee, 2014). Gene transcriptions and other phenotypic assays have shown that *amp*R negatively regulates the transcription and function of MexEF-OprN efflux system through modulating the expression of MexT which encodes a positive regulator. Several *amp*R mutations in clinical isolates were reported in the literature as associated with multidrug resistance in *Ps. aeruginosa* (Balasubramanian, Kumari and Mathee, 2014) (Cabot et al., 2012). Some of these mutations are summarized in Table 3.1

The following diagram summarizes the different efflux pump-regulatory mechanisms described above.



Reference: (Housseini B Issa, Phan and Broutin, 2018)

Table 3.1. Summary of mutations in transcriptional regulator genes related to hyperexpression of different types of efflux pumps

Mutant	Description	Type/position of mutation	Reference
*nal*B mutants	Mutation in *mex*R leading to overexpression in efflux pump	Arg70Trp, Leu80Pro, Arg91His, Arg83His, Leu13Met, Ala110Thr Leu57Arg, Arg59Cys, Gly58Glu, Leu95Phe, Thr69Ile, Arg21Trp Thr130Pro	(Adewoye *et al.*, 2002)
		Asp8Gly, Ala66His, Lys44Ser	(Choudhury *et al.*, 2016)
		Arg70Trp, Thr130Pro Gln94Pro, Leu57Arg	(Srikumar, Paul and Poole, 2000)
		Pro7leu, Pro143Thr, Val126Glu	(Suresh *et al.*, 2018)
		Asp8Glu, Ala66val	(Ziha-Zarifi *et al.*, 1999)
		Lys44Met, Val126Glu, Ala66Pro	(Llanes *et al.*, 2004)
		His107Pro, Ala103Thr, Gln106His, Asn53Asp, Asn53Tyr, Arg21Gly Ser26Gly, Asn79Gly, Met10Arg, Ser88Cys	(Higgins *et al.*, 2003)
		Val-126-Glu	(Vaez *et al.*, 2014)
*nal*C mutants		Gly71Glu, Ser46Ala, Glu153Gln, Ser209Arg, Asp76Glu, Leu61Pro Met151Thr	(Llanes *et al.*, 2004)
		Ser127Pro, Thr50Pro	(Cao, Srikumar and Poole, 2004)
*nal*D mutants	*nal*D gene absence or disruption leads to overexpression of efflux pump	Ser32Asn Change at nucloetideT410 Change at nucleotide G433	(Sobel *et al.*, 2005)
	*nal*D mutations that interfere with NalD binding to its promoter	Phe175Ala	(Chen *et al.*, 2016)
		Leu22Pro, Thr158Ile	(Jorth *et al.*, 2017)
	Mutation in ligand binding domain	Leu153Gln	(Suresh *et al.*, 2018)

*nfx*B	Mutation leading to loss of repressor activity of *nfx*B which results in hyperexpression of MexCD-OprJ	Ala124Glu, Glu111Lys, Arg21His, Asp56Gly, Glu8Lys, Gln64His Gln52His, Ala141Gly, Ser36Gly, Ala38Gly, Glu75Gln	(Higgins *et al.*, 2003)
		Gly166Asp, Gly192Asp, Phe147Ser	(Purssell and Poole, 2013)
		Arg42Gly, Arg42His	(Chuanchuen *et al.*, 2001)
		Glu124Ala	(Vaez *et al.*, 2014)
		Arg82Leu	(Jalal *et al.*, 2000)
*nfx*C	Laboratory and clinical mutants overproducing M*ex*EF-O*pr*N due to derepressed *mex*T or non-functional *mex*S.	*mex*S Asn249Asp, *mex*S Val104Ala, *mex*S Phe253Leu, *mex*S Leu263Gln	(Richardot *et al.*, 2016)
		*mex*S Glu54Gly, *mex*S Gly78Ser *mex*S Ala75Val, *mex*S Thr152Ala *mex*S Ala175Val, *mex*S Glu181Asp *mex*S Val308Ile, *mex*S Cys269Yyr	(Llanes *et al.*, 2011)
		*mex*S Asp244Asn, *mex*S Val333Gly *mex*S Ser124Arg	(Sobel, Neshat and Poole, 2005)
		*mex*T Gly258Asp, *mex*T Yyr138Asp	(Richardot *et al.*, 2016)
		*mex*T Gly257Ser, *mex*T Arg166His	(Llanes *et al.*, 2011)
*mva*T		*mva*T Ala115Thr	(Llanes *et al.*, 2011)
*amp*R	A global transcriptional regulator -its deletion or some types of mutations are associated with XDR/MDR	Gly154Arg, Glu114Ala, Gly283Glu Met288Arg, Ala51Thr	(Cabot *et al.*, 2012)
		Asp135Asn, Gly102Glu	(Caille *et al.*, 2014)

3.1.3. SOS response (stress) mediated pathways

Resistance mechanisms listed above appear to contribute additively to determine quinolone resistance. Although each mechanism can be identified separately in individual isolates and can generally contribute to resistance, it appears that each is not enough separately for high level resistance. Mutations in genes encoded in the SOS system were linked to quinolone resistance and these mutations appear to show a more remarkable effect on quinolone resistance with a possible cross-interaction with other resistance mechanisms.

(Breidenstein, Bains and Hancock, 2012). The two most commonly studied and reported mutations were *Lex*A and *Lon* protease related mutations. The SOS response is considered a classic bacterial stress response that is induced by DNA damage which results from a wide range of stressful exposures including antibiotics (Miller *et al.*, 2004). Fluoroquinolones cause dsDNA breaks as an essential step of its mechanism of action and thus is considered as a potent inducer of SOS response (Blázquez *et al.*, 2006).

In *Ps. aeruginosa*, the global transcriptional response to clinical doses of ciprofloxacin was studied in a microarray analysis comparing *Lex*A mutant to wild type. This showed that *Lex*A regulates the expression of 15 genes with specialized DNA recombination and replication function (Cirz *et al.*, 2006)(Erill, Campoy and Barbé, 2007). In addition, other complex coordinated networks of *lex*A-dependent responses to ciprofloxacin related to downregulation of many aspects of metabolism, motility and permeability were also identified (Cirz *et al.*, 2006). The same study showed that sixty-four of the upregulated genes in the response to ciprofloxacin are in regulons controlled by *Lex*A-like repressor. *Lon* protease also affects the SOS response through acting on *sul*A as a substrate. It showed to act by cleaving *rec*A repressors leading to SOS induction. Results of microarray studies comparing *Lon* mutants to wild type showed that *lon* mutants exhibit increased susceptibility to ciprofloxacin as a result of suppression of the SOS response triggered by DNA damaging agents (Breidenstein, Bains and Hancock, 2012).

3.1.4. The importance of understanding the global picture

Understanding the global picture and the core physiology of resistance is crucial for designing better diagnostics and also for identifying more effective drug-targets. It is important to understand the relative contribution of each resistance-associated element to the overall resistance in order to fully predict what it really takes the cell to be resistant. Research investigating the effect of efflux pumps on function is usually performed on experimental mutants under controlled laboratory conditions. This means that such studies may not be a true reflection of the real clinical conditions. For that reason, a larger scale analysis is needed and should include both resistant and susceptible clinical isolates. This would help in evaluating the combined effect and the co-contribution of all resistance elements to overall resistance.

Doing that would increase the knowledge needed for the practical use of genome sequencing data in the development of rapid diagnostics that can direct treatment in an effective way.

Although there are currently some bioinformatics tools that are available to predict for resistance using WGS data, these pipelines have major drawbacks as evaluated and explained in detail in *Chapter 2*. They are not evaluated for their practical use and they only include a limited set of molecular markers as a basis for their prediction. Most of these markers include gene presence/absence information. Based on the assessment of these tools in *Chapter 2*, the set of molecular markers in these databases appears to be insufficient to predict for the correct phenotype. Performance criteria of the diagnostic test needs to be optimized to offer the most accurate and reliable diagnostic. Although reported in different studies that QRDR mutations are not the main contributor to overall resistance, this observation was not advanced into more detailed exploration of the relative contribution of each system element into resistance. The idea of cumulative mutational events, whether in target enzymes or in transcriptional regulators, that add up to cause resistance was observed before. However, to the best of my knowledge, no particular specific predictive model has been proposed to link different mutational patterns to resistance phenotypes.

The aim of this chapter is to test the theory of additive and combinatorial effect of resistance-conferring genes and mutations and to build a simple predictive model that shows the relative contribution and the significance of each mechanism in explaining resistance. This helps to find the best combination of changes that can predict resistance and can subsequently be used as a better molecular diagnostic set of markers.

3.2. Molecular basis of aminoglycoside resistance

3.2.1. Aminoglycoside-modifying enzymes (enzymatically- catalyzed antibiotic-inactivation)

Inactivation of aminoglycosides antibiotics by resistant *Ps. aeruginosa* isolates has been recognized since the 1960s and 1970s. Modifying enzymes have been extensively described in the literature (Smith and Baker, 2002) (Azucena and Mobashery, 2001). Many enzymes exist to perform the role of enzymatic inactivation through acetylation, adenylation or, phosphorylation (Wright, 1999). This mechanism has been considered the most widespread cause of aminoglycosides resistance. Some of these enzymes are usually detected using publicly available tools for resistance prediction and were assessed in *Chapter 2*.Although these enzymes are commonly encountered and reported in the literature, chromosomal mutations need to be

evaluated in detail to understand their contribution to overall resistance, and this is one of the study points addressed in this chapter.

3.2.2. Ribosomal Mutation or Modification

Mutations or enzymatic modifications of the ribosome at aminoglycosides binding sites are two possible ways bacteria can evade aminoglycoside access to its target. Ribosomal modifying methyltransferases (16SrRNA methylases) perform the role of enzymatic modifications at the aminoglycosides ribosomal binding site. *rmtA*, that encodes a 16S rRNA methylase has been recognized as a cause of high-level pan-aminoglycosides resistance (Yokoyama *et al.*, 2003). Highly drug resistant bacteria that produce enzymes capable of modifying the active site of the 16S rRNA through methylation has become prominent in recent years (Doi, Wachino and Arakawa, 2016). Although resistance mediated by acquired 16S-RMTase was first described in the early 2000s, not much data are available on the impact of acquired 16S-RMTase production and clinical outcome of patients when they are treated with aminoglycosides (Doi, Wachino and Arakawa, 2016). Mutations in *rrs* gene, which code for 16S rRNA, hinders aminoglycoside binding to its ribosomal target leading to resistance. These mutations are not very common because changes to this vital cellular machinery are often lethal. However, some viable mutations in *rrs* have been reported including A1401G, C1402T, and G1484T. These mutations have been identified in some clinically isolated strains of resistant *M. tuberculosis* (Maus, Plikaytis and Shinnick, 2005). Loss-of-function mutations in rRNA methylase *tlyA* and the *gidB* gene have also been identified in resistant *M. tuberculosis* (Georghiou *et al.*, 2012). Mutations in the *rplY (*PA14_61780) gene which encodes a ribosome associated protein (L25) with a role in protein synthesis ensuring accurate ribosomal translation under stress conditions have been recognized to affect aminoglycosides susceptibility. The *rpl*B gene that codes for the L2 protein associated with a 50S ribosome has also been linked to aminoglycosides resistance (Feng, Jonker, Moustakas, Brul and Ter Kuile, 2016). Mutations in *fusA1*, coding for the elongation factor EF-G1A, has also been linked to aminoglycoside resistance in both clinical isolates and *in-vitro* mutants (López-Causapé *et al.*, 2018). Mutations associated with aminoglycosides resistance as extracted from the literature are summarized in Table 3.2

3.2.3. Cell membrane modifications (changes in cellular permeability)

The role of cell membrane permeability changes in aminoglycosides-resistant clinical isolates of *Ps. aeruginosa* (Kettner *et al.*, 1995) has been early recognized. Impermeability-related resistance has been linked to changes in outer membrane composition of *Ps. aeruginosa*, including alterations in the structure of lipopolysaccharide, overexpression of outer membrane protein (OMP) OprH and changes in the electron transport chain. Mutations associated with outer membrane porin changes (OprH-PhoP/PhoQ) have been linked to aminoglycosides resistance including resistance in clinical isolates (Hasegawa *et al.*, 1997) (Yoneyama, Sato and Nakae, 1991).(Shearer and Legakis, 1985).

Inactivation of *galU* and *nuoG* genes is also associated with impaired outer membrane uptake in addition to reduced active transport (El'Garch *et al.*, 2007). The inactivation of some genes, including *nuo*G, *rpl*Y, and *gal*U, has shown to gradually increase AG resistance by reducing proton motif force, modifying the AGs target and impairing AG binding and uptake in laboratory strains, respectively (Dean and Goldberg, 2002). Combined activation of *Ps. aeruginosa* two-component systems *amg*RS and *pmr*AB has also been linked to aminoglycoside resistance phenotype (Schniederjans, Koska and Häussler, 2017). Expression levels of *pmr*A PA4776 (PA14_63150) two component sensor-regulator and *pmr*B PA4777 (PA14_63160) two component sensor kinase has been reported to increase in clinical *Ps. aeruginosa* isolates in association with aminoglycosides resistance and the small colony variant phenotype.

Genes associated with lipid biosynthesis or metabolism (*lpt*A, *fao*A, *arn* genes) and phosphate uptake (*pst*B) are another group of genes showing important link to aminoglycosides resistance. Disruption of the genes *pst*B, *fao*AB, and *lpt*A showed association with increased aminoglycosides susceptibility (Krahn *et al.*, 2012). Arn (PA3552-PA3559) LPS modification genes have also been linked to aminoglycosides resistance. Expression of the *arn*BCADTEF operon is known to decrease the interaction and uptake of polycationic antibiotics (Schniederjans, Koska and Häussler, 2017).

3.2.4. Efflux-pumps and SOS response -mediated pathways

MexXY-OprM in *Ps. aeruginosa* has an important role in conferring resistance to the aminoglycoside (AG) class of antimicrobials studied (Morita, Tomida and Kawamura, 2012)..

The expression of the MexXY-OprM efflux pump has also shown to be inducible by exposure to reactive oxygen species, contributing to pan-AG resistance (Fraud and Poole, 2011). Overall, the contribution of efflux to aminoglycosides resistance is low, however, MexXY overexpression due to a mutation in the repressor gene *mex*Z is the most common mechanism of AG resistance in lung isolates from cystic fibrosis patients with chronic *Ps. aeruginosa* infections (Poole, 2011). Regulatory mutations related to this type of efflux-pump are shown in Table 3.1

SOS-response stress mediated pathways are known to be induced by both aminoglycosides and quinolones as shown above in section 3.1.3

Table 3.2. Summary of mutations extracted from the literature in relation to aminoglycosides resistance

Gene	Variant	Reference
*rpl*Y (PA4671)	Gly367Thr, Ala123Ser	(Islam *et al.*, 2009) (Poonsuk, Tribuddharat and Chuanchuen, 2013)
*rps*L	G524C	(Springer *et al.*, 2001)
Rrs gene	A1401G, A514C, C517T, A513C G1484T C1402T, G1158T, A1338C, A907C A1408G, T1406A, C1409T, G1491T G524C, C526T, C522T	(Feuerriegel *et al.*, 2009) (Jugheli *et al.*, 2009) (Perdigao *et al.*, 2010) (Leung *et al.*, 2010) (Campbell *et al.*, 2011) (Engstrom *et al.*, 2011) (Sirgel *et al.*, 2012) (Maus, Plikaytis and Shinnick, 2005) (Nessar *et al.*, 2011) (Springer *et al.*, 2001)
*tly*A	Gly196Glu, G223T, T220C, T708G	
*gid*B	T230C, C286T, T104G, A254G	
*amg*S	Arg182Cys, Val121Gly, Asp106Asn Ala28Glu, Cys28Ala	(Lau *et al.*, 2015) (Lau *et al.*, 2013) (Schniederjans, Koska and Häussler, 2017)
*pmr*A	Leu71Arg, Asp92Tyr	(Schniederjans, Koska and Häussler, 2017)
*pmr*B	Thr4Ala, Leu323His, Ser420Arg Gly423Cys, Leu243Gln, Ala248Val	(Schniederjans, Koska and Häussler, 2017)

		(Moskowitz, Ernst and Miller, 2004)
fusA1 (PA4266)	Yyr552Cys, Arg371Cys, Thr456Ala Arg680Cys, Val93Ala, Ala555Glu Thr671Ala	(Bolard, Plésiat and Jeannot, 2018) (López-Causapé *et al.*, 2018)

The aim of this chapter is to test the combinatorial effect of different quinolone and aminoglycoside resistance mechanisms to understand the significance and relative contribution of some previously reported variants as well as newly identified variants in genes known to have a role through the quinolone and aminoglycoside mechanisms of action pathway. I also seek to build a simple predictive model for quinolones and aminoglycoside resistance phenotype.

3.3. Objectives

The work detailed in this chapter aims to:

- Review of the literature to extract genes and gene variants associated with quinolone and aminoglycoside resistance.
- Describe and analyze the distribution of previously identified resistance-associated markers in the studied set of *Ps. aeruginosa* isolates.
- Test predictive values and other measures of diagnostic accuracy for established resistance- associated markers.
- Explore the best predictor combinations of markers that can improve diagnostic performance using cluster analysis and multiple regression analysis.

3.4. Methodology: Assessment of the predictive potential of established resistance mechanisms and markers separately and in combination for predicting ciprofloxacin, levofloxacin, gentamycin, and amikacin resistance/susceptibility phenotypes

3.4.1. Literature search and identification of variants

- Primary literature review to extract genes and gene variants associated with quinolone and aminoglycoside resistance was carried out on each of PMC PubMed, ACADEMIC SEARCH COMPLETE (EBSCO host) and ScienceDirect using search criteria:

"*Pseudomonas aeruginosa*"[title/abstract] AND "aminoglycosides resistance"[title/abstract]
"*Pseudomonas aeruginosa*"[title/abstract] AND "Quinolone resistance"[title/abstract]

- Secondary more specific searches were also conducted using search criteria 'Efflux pumps OR Target mutations AND *Pseudomonas aeruginosa*'.
- All search results were analyzed to extract variants and genes with function related to antibiotic resistance to studied antibiotic groups. These are briefly discussed in section 3.1 and section 3.2 and summarized in Table 3.1 and Table 3.2

3.4.2. Choosing a set of completely sequenced genomes for performing the analysis

- Completely sequenced genomes with associated laboratory measured phenotypic data for ciprofloxacin, levofloxacin, gentamycin, and amikacin were downloaded from the Patric database (Wattam *et al.*, 2017). The list of genomes included in the study are shown in supplementary Table 1A.
- Sources of bias originating from possible deviation in disease spectrum was decreased by including the whole spectrum of the conditions under evaluation (resistance and susceptibility to antibiotics).
- For ciprofloxacin, the analysis included the lab group of isolates in addition to 144 genomes for which ciprofloxacin lab measured susceptibility data are available from the Patric database (37 susceptible isolates and107 resistant isolates). For levofloxacin the analysis included the lab group of isolates in addition to 532 genomes for which levofloxacin lab measured phenotypic data are available from the Patric database (342 resistant and 190 susceptible). For amikacin, the analysis included a total of 690 *Ps. aeruginosa* isolates; 162 lab group of isolates (6 are amikacin resistant isolates and 156 amikacin susceptible isolates) and 528 genomes from Patric database for which phenotypic data are available (142 amikacin resistant isolates and 386 amikacin susceptible isolates). For gentamycin the analysis included a total of 301 *Ps. aeruginosa* isolates: 162 lab group of isolates (49 are gentamycin resistant) and 139 genomes from the Patric database for which phenotypic data are available (57 are gentamycin resistant).
- Breakpoints for analysis of sensitivity and resistance were defined according to the latest EUCAST recommendations (Rules, 2018). The list of genomes included in the study is

shown in supplementary Table 1A and Table 1B. Primary data of distribution of studied genes and mutations are shown in supplementary Table 2A, Table 2B, Table 3A and Table 3B for quinolones and in supplementary Table 4A and Table 4B for aminoglycosides.

3.4.3. Identifying the distribution and performance of each of the resistance/susceptibility-associated genes and mutations

- The distribution of resistance-associated mutations previously described in the literature in both completely sequenced genomes (lab group) and genomes selected from Patric was identified using NCBI BLAST (Delcher, 2002) or integrated CARD resistance genes features and mutation annotations when available for genomes from Patric (Wattam *et al.*, 2017). Detailed results are shown in supplementary Table 2A and Table 2B.
- The sequence of each of the genes shown in the literature was extracted by searching the *Pseudomonas* genome database (Winsor *et al.*, 2016) available at https://www.pseudomonas.com/ using the known gene identifier ID and/or the gene name as shown in the review detailed above and in case the gene was not found, the gene was alternatively extracted from NCBI database at https://www.ncbi.nlm.nih.gov/gene. The available gene sequence was download from *Pseudomonas* genome database or from NCBI by constructing a FASTA file using the sequence available on either of the two databases.
- The NCBI BLAST+ BLASTN tool available at https://usegalaxy.org/ was used to search nucleotide database with nucleotide query sequence(s) (Galaxy Version 0.3.3) (Cock *et al.*, 2015). The nucleotide query sequence of the gene sequences extracted above was used to search the constructed nucleotide BLAST database using the megaBLAST option and the default Set expectation value cutoff at 0.001. The BLAST tabular output of the tool was then converted into FASTA format using the convert formats tool Tabular-to-FASTA which converts tabular file to FASTA format (Galaxy Version 1.1.0).
- The aligned part of both subject and query sequences in FASTA format was then visualized and explored using MEGA 7 software (Kumar, Stecher and Tamura, 2016)

- Multiple gene copies from the same genome and partial hits were excluded from the alignment. All genes identified, included and used in the current analyses showed > 80% percentage identity and >95% query coverage.
- Genes of interest were manually explored in detail to extract variants of interest (that were previously reported in the literature) as well as novel variants (amino acids or nucleotide changes showing specific differential pattern of distribution).
- A matrix showing the distribution of each of the variants of interest was manually generated in an excel data sheet which was then used to conduct further analyses.
- Where novel variants were identified, these were tested for their predicted functional effect using PROVEAN (Choi and Chan, 2015).
- PROVEAN (**Pro**tein **V**ariation **E**ffect **An**alyzer) is a software tool used to predict whether an amino acid substitution has an effect on the protein biologic function. PROVEAN Protein tool from the PROVEAN web server functions using PROVEAN v1.1.3.provides PROVEAN prediction for a protein sequence from any organisms. The tool available at http://provean.jcvi.org/seq_submit.php was used to study the effect of different amino acid variants of interest. Protein query sequences of interest were extracted from the Pseudomonas genome database or from NCBI; and the protein sequence file in FASTA format together with the amino acid variant of interest were used as input to the tool using the input format as specified by the tool at http://provean.jcvi.org/help.php#protein_variation_input_format. The output showed PROVEAN scores for each of variants entered and this was then used to predict the functional effect of the variant based on either the (-2.5) default cutoff point or a less stringent cutoff of (-1.3). Scores less than -2.5 or -1.3 are predicted as deleterious based on the cutoff chosen for prediction.
- The distribution of the mutations was then tested for its correlational pattern with phenotype and different measures of diagnostic accuracy were evaluated to test their diagnostic benefit and potential use as molecular predictive markers in the whole collection.
- To do so, the variant distribution matrix generated was used to construct a 2*2 contingency table for each variant in relation to resistance/susceptibility phenotype. “Cross Tab” function was used to generate these contingency tables using SPSS

(SPSS.V21) with checking all the options to calculate chisquare test for independence or Fischer exact, significance value, phi coefficient, cramer V and likelihood ratios.

- Parameters of performance for each single variant including, sensitivity, specificity, NPV, PPV, Likelihood Ratio (LR), Likelihood Ratio positive (LHR+), Likelihood Ratio negative (LHR-), diagnostic odds ratio (DOR), Youden index and diagnostic accuracy for each contingency table were calculated according to the following equations; Sensitivity = TP/TP+FN, Specificity =TN/TN+FP, PPV= TP/TP+FP, NPV= TN/TN+FN, LHR+ = sensitivity / (1-specificity), LHR - = (1-sensitivity) / specificity, DOR= LHR+/ LHR- or DOR= sensitivity* specificity/ [(1-sensitivity)*(1-specificity)]. Diagnostic accuracy= (TP+TN)/(TP+TN+FP+FN).

3.4.4. Finding combinations with improved performance

Common guides of test performance include sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios. None of these indicators in itself can represent a full discriminatory performance.

Sensitivity can be considered only one part of the discriminatory evidence (Shaughnessy, 2007). There is no simple combination rule to associate sensitivity and specificity into one operational measure. Also, there are no guidelines to rank pairs of indicators with better performance. For all these reasons, a single indicator showing high discriminatory performance needs to be used (Glas *et al.*, 2003). Examples of these types of single indicators include accuracy, Youden's index, and diagnostic odds ratio. Accuracy can show the percentage of correct classification by the test under evaluation and it depends on the prevalence of the target condition evaluated when sensitivity and specificity are not equal. It also has the advantage of weighing false positive and false negative findings equally (Linnet, 1988). Youden's index is another single indicator derived from sensitivity and specificity, but its values are difficult to interpret (Hilden and Glasziou, 1996). Diagnostic odds ratio can be considered as a global single measure of test performance and is not prevalence dependent (Glas *et al.*, 2003). An important point to consider is that for all indicators of test performance, the spectrum of disease severity has a great effect (G. M. Moons *et al.*, 1997) and this has been considered in the current analysis.

Sensitivity and specificity cannot be considered as the sole predictive measures because they can only describe the ability of the diagnostic test used to correctly identify all cases with the disease or specific diagnosis (sensitivity) or to correctly exclude cases without the disease (specificity). Both sensitivity and specificity have the advantage of being not influenced by disease prevalence which means that results can be transferred from one study to others. However, sensitivity and specificity can vary greatly depending on the disease spectrum which was considered in the current analysis by including a wide range of sensitivity and resistance phenotypes (MIC ranges).

Predictive values can provide useful information about the usefulness of the diagnostic test; however, it is highly affected by the prevalence of the disease or condition under investigation. Predictive values give information about the probability of having the disease when test result is positive (PPV) or the probability of being healthy with a negative test result (NPV). Likelihood ratio should be an optimal choice in reporting diagnostic accuracy because it considers both sensitivity and specificity and is not dependent on disease prevalence (Eusebi, 2013).

For all of the reasons described above, a new approach was applied here to evaluate the quantitative individual and combined relative contribution of variants to the phenotype under evaluation and this was implemented in two stages:

Stage 1:

Selecting the best performing individual markers that showed the highest values among the combination of the parameters listed above. These included: Diagnostic accuracy > 0.65, LHR+ > 5, DOR >20, Youden index > 0.5 and any value greater than 90% or 95% for sensitivity, specificity, PPV, and NPV. Where these values are not met in a specific dataset, the highest values found in the evaluated sets are used. For example, for DOR and LR, any value > 1 is accepted for differentiating cases from controls, however, the higher the value the better the performance is. This means that for each set of data evaluated, standards from literature are applied first and if not found, the best values among each specific data set are used.

Stage 2:

This stage aimed at the evaluation of the combinatorial and quantitative contribution of different genomic elements to phenotype in order to find the most informative individual markers or combination of markers that can add to the interpretation of the phenotype. This was done

through applying two statistical techniques including step- wise multiple regression analysis and cluster analysis.

- Stepwise multiple regression was conducted to examine the extent of variance in phenotype as explained by different molecular markers under assessment. It can also give an indication about the relative contribution of each marker to antibiotic resistance phenotype. The method is used to develop a subset of independent variables (genes/gene variants) that are useful in predicting the dependent variable (susceptibility/resistance phenotype) and eliminate those independent variables that do not provide additional prediction. The method was chosen because we have a group of independent variables (genomic markers) and there is no theoretical background about which variables are contributing more towards determining the phenotype. Using this method, blocks of variables are assessed step-wise without base line hypotheses with all variables being entered into the equation , then some variables can be included based on a set of statistical criteria for selection (variables with smaller correlations and multi-collinearity are removed) to find the best model. Variables are assessed step-wise and once all the variables are entered, the overall model is assessed in terms of its ability to predict variance in the dependent variable and the relative contribution of each block of variables.
- Cluster analysis (Bacher, Wenzig and Vogler, 2004) was then used to explore the variants distribution to find previously hidden but useful groups or combination of mutations that can improve diagnostic accuracy.
- Cluster analysis is an exploratory multivariate method used for the classification or grouping of subjects under study (in our case; isolates genomes) to find connectivity based on a set of measured parameters. Cluster analysis classifies cases into relative groups which is based on similarities of different attributes among members of the same group. Clusters are examined to determine characteristics that are unique to each cluster. The attributes examined here include: The S/R phenotype, the MIC values and the marker presence/absence information.
- Clustering is not an automatic process but an iterative process of knowledge discovery. It involves several trials of iterative optimization. Data entered into the cluster analysis

needs to be pre-processed based on the desired tested parameters till the best model is found.

- To apply that, different individual markers are assessed for their diagnostic performance using the parameters shown above, then those showing the highest possible values or according to what is shown in ***stage 1*** above are evaluated through the clustering algorithm. Because there are no standard criteria or guidelines about what are considered the best values for some parameters, different combinations are therefore tested till better combinations are achieved. The number and the size of clusters are automatically determined by the clustering algorithm.
- The final output tables or graphs from the cluster analysis show the following:
    - The size (number of isolates and percentage to total) for each set of isolates defined as a group.
    - The comparative pattern in the evaluated set based on MIC distribution and average MIC.
    - The percentage of susceptibility/resistance observed in each group according to clinical breakpoint classification.
    - The relative difference in the distribution of observed variants on which the classification was based (The differential percentage of marker presence/absence).
- After applying cluster analysis which showed output classification based on 3 attributes (S/R phenotype, MIC values and marker presence/absence information), all the strains for which phenotypic data were available were tested for each antibiotic as entry at the same time (there was no prior classification into resistant and sensitive). From the output, each strain from all those tested then should belong to one cluster as defined by the algorithm. Each of all tested strains will then belong to one of the predicted clusters with improved performance and is given "new cluster or group membership ID", here we do not have missing cases. Each of the tested strains have a phenotype which is either sensitive or resistant. In basic definition, sensitivity, specificity and predictive values cannot be calculated except when divided by the total number of strains assessed. So, for each strain of all those assessed, we only have 2 probabilities either to belong to a certain predicted cluster or not, either positive or negative and this may be true or false (based on phenotypic data). This forms the basic 2*2 table which is used for calculation of

predictive values. In the end we have true positive, true negative, false positive and false negative. Each of assessed strains must then belong to one category of those four (True Positive [TP]- True Negative [TN]- False Positive [FP]- False Negative [FN]). The equations for the predictive values are then applied as shown above.

- All statistical analyses including predictive values, cluster analysis and stepwise multiple regression analysis were conducted using IBM SPSS (SPSS.V21) according to the instructions shown in the IBM SPSS manual guide at https://www-01.ibm.com/support/docview.wss?uid=swg27021213

3.5. Results Section.1 (ciprofloxacin)

3.5.1. Description of mutations distribution

1) QRDR

*gyr*A G83I was the most frequently observed in this group. This mutation was identified in 29 genomes (28 resistant and 1 susceptible) out of 310 genomes studied for ciprofloxacin sensitivity. A total of 84 ciprofloxacin resistant isolates did not show *gyr*A G83I mutation. *gyr*A D87N was identified in only 4 genomes (all resistant to ciprofloxacin). *gyr*B E468D was identified in only one resistant isolate.

*par*E mutations were less frequently observed. *par*E A473V was identified only in 3 genomes out of 310 (all three are ciprofloxacin susceptible). *par*E V460G was found only in one susceptible isolate. *par*E S457G was identified in only one resistant isolate. The same was true for both, *par*C E91K and *par*C E91L. *par*C S87W was identified in only four isolates out of 310 (all are ciprofloxacin resistant MIC=8). Interestingly, *par*C S87L was more frequently observed being present in14 out of 310 total genomes (all of which are ciprofloxacin resistant).

2) Efflux pumps related operons and regulators

1) *nal*D *transcriptional regulator (PA3574)*

This gene was searched for in the whole collection using the CARD annotation available at the Patric genome group linked special features (Wattam *et al.*, 2017) or NCBI blast (Cock *et al.*, 2015). The gene was absent in 18 genomes out of 310 (all were ciprofloxacin resistant).

Mutations previously reported in the literature as shown in Table 3.1 were explored in the studied collection. *nal*D ser32Asn variant identified in two isolates by Sobel *et al.*, (2005) was identified in one susceptible isolate in the current study set. The mechanism of *nalD* induced regulation of MexAB-OprM is not fully understood. One of the studied pathways of MexAB-OprM *nal*D- induced regulation showed that PAO1Δ*nal*D complemented with *nal*D F175A is capable of exhibiting increased resistance and thus concluded that this mutation leads to decreased DNA binding affinity to its promotor leading to the pump overexpression (Chen *et al.*, 2016). Suresh *et al.*, (2018) showed that this is one hypothesis to explain one mechanism of *nal*D-related regulation but also showed that there is a diversity of other regulatory mechanisms that can be involved in resistance regulation. When the current set of isolates were scanned, the reported mutation was not identified. A possible reason is that although the experimentally induced mutation showed functional effect on its binding domain, this does not necessarily mean that the mutation must be encountered in the clinical population of the organism but alternatively, other mutations can be encountered to produce the same function or the same effect on DNA binding affinity. On the other hand, another variant (*nal*D L153Q) previously reported to affect the ligand binding domain (Suresh *et al.*, 2018) was identified in 2 isolates among the current study set and both were ciprofloxacin resistant. *nal*D Thr158Ile and *nal*D Leu 22 Pro are two other mutations that showed to be selected experimentally by multiple passages on exposure to aztreonam (Jorth *et al.*, 2017). *nal*D is a regulator that affects MexAB-OprM expression and hence *nal*D mutations should affect resistance to a wide range of antibiotics that are substrates of the pump. However, Jorth *et al.*, (2017) showed that although these mutations were selected by continuous experimental passage with aztreonam, there was no corresponding increase in resistance phenotype to all antibiotics exported by the pump, the fact that may be expected because no single variant in one regulatory gene would necessarily confer resistance to all different antibiotic classes even if all are substrates of the pump. This would also explain why these two mutations were not encountered in the current analysis. In addition, Jorth *et al.*, (2017) showed that the same mutations can have a role in modifying virulence but concluded that such *mex*R and *nal*D selected mutations can affect both resistance and virulence, however this needs to be further tested.

Although the *nal*D regulator appeared to be conserved in the studied set, a group of *nal*D variants were observed in the current study set which includes *nal*D D187A identified in 6

isolates (2 were ciprofloxacin susceptible and 4 were resistant), *nal*D D187H identified in 4 isolates (all were susceptible to ciprofloxacin) and *nal*D D185H identified in 1 isolate which was also susceptible. These types of observations are listed here for further detailed functional and epidemiologic testing.

*2) nal*C *transcriptional regulator (PA3721)*

The *nal*C gene was absent in 27 genomes out of 310 studied in total (all are ciprofloxacin resistant). *nal*C G71E identified by Llanes *et al.*, (2004) was a common variant in the current study set and was identified in 257 isolates. *nal*C S209R identified in the same study was also commonly encountered in 207 isolates. In addition, both *nal*C S46A and *nal*C E153Q reported in the same study were identified in 116 isolates and 16 isolates respectively. However, three other variants including D76E, L61P and M151T were not encountered in the current study set. This may be expected because each of these three variants were identified in only one isolate each by Llanes *et al.*, (2004). The other three variants which were frequently observed in the current study set also appeared to be a recurrent substitution as shown by Llanes *et al.*, (2004). Each of G71E, E153Q and S209R were identified in 11 out of 12 isolates which support the distribution encountered in the current study set. *nal*C A186T was identified in 23 isolates. *nal*C Thr50pro shown by Cao, Srikumar and Poole, (2004) using transposon insertion mutants displaying the *nal*C phenotype was identified in 2 isolates among the current set while Ser127Pro shown in the same study was not identified.

3) nalB mutant phenotype: mexR multidrug resistance operon repressor MexR (PA0424)

*Mex*R R79S previously reported in the literature (Higgins *et al.*, 2003) was identified in 7 isolates (4 are resistant and 3 are susceptible). Another substitution *mex*R R79N at the same position was more frequently observed in 297 isolates (102 are resistant and 195 are susceptible). Four other substitutions in *mex*R; *mex*R G97L, *mex*R L29D, *mex*R E70R and; *mex*R L130T were frequently observed in 306, 304, 303, 302 isolates respectively. All were more frequently observed in susceptible isolates. Other single amino acid changes previously shown by Adewoye *et al.*, (2002) and Choudhury *et al.*, (2016) to affect the stability of *mex*R or its ability to dimerize were not identified in the studied set possibly because the other identified mutations are performing this function or are inactivating the gene in an alternative mechanism. The detailed study of the possible effect of each variant was not the aim of the current analysis.

4) armR (PA3719)

This gene was one of the important genes observed in the current analysis to have a differential effect on susceptibility and resistance and was present in 182 isolates out of 306 (102 resistants to ciprofloxacin and 80 susceptible to ciprofloxacin).

5) nfxB mutant

The *nfx*B gene was absent in 6 genomes, 5 of which are ciprofloxacin resistant and one is ciprofloxacin susceptible. The mutation *nfx*B Ala124Glu which was reported by Higgins *et al.*, (2003) was not found in the studied group of genomes but instead another substitution at the same position A124T was found in 5 isolates; one of which is susceptible to ciprofloxacin while the other 4 are resistant. This would support the importance of mutation observed at that position in compromising the activity of *nfx*B repressor. *nfx*B Arg82Leu was found in 6 isolates out of 310 (2 are ciprofloxacin resistant and 4 are susceptible). *nfx*B Arg 21 His was found in 13 isolates (10 are susceptible and 3 are resistant). *nfx*B Asp56Gly was also found in 15 isolates (12 are ciprofloxacin susceptible and 3 are resistant). Some other mutations reported in one study (Higgins *et al.*, 2003) including; *nfx*B ser36Gly (reported in only one isolate in combination with other *gyr*A and *par*C mutations), *nfx*B Ala38Gly (reported only in one isolate in combination with another *gyr*A mutation) and *nfx*B glu111Lys (which was similarly reported in one isolate in combination with three other mutations in *gyr*A, *par*C and *mex*R) were not identified in the current studied group of the strains. This may be expected as the study reporting these mutations has included 58 *Ps. aeruginosa* isolates from a certain geographic locality (European hospitals) at certain time period (1998-1999). The study has identified each of these three mutations only in a single isolate and in combination with other mutations. The study has showed the observation of these variant in combination with other variants which indicates a possible additive role of these mutations. However, this is not necessarily generalizable to other settings and also gives no confirmation about the possible role of each of these mutations separately considering that these three mutations were very infrequent. Each of these variants was identified only in 1 out of 58 isolates and was also observed in association with other mutations. The current analysis which included a more diverse and comprehensive set shows agreement with the findings from the same study by reporting exactly the same mutations excluding the three variants shown above. This may indicate that these mutations are not so frequent or important. The same study has also

concluded that these regulatory mutations were not found separately but were identified in combination with other *gyr*A or *par*C mutations.

Taken together, these findings support conclusions drawn from the literature about the association of some mutations in DNA binding domain of *nfx*B regulator with resistance (Okazaki and Hirai, 1992), and that spectra of mutations can also occur in the entire length of the gene leading to compromised activity of *nfx*B repressor (Monti et al., 2013). Both experimental and clinical isolates expressing the efflux system and exhibiting high levels of quinolone resistance invariably show mutations in *nfx*B (Purssell and Poole, 2013). It has also been reported that the MIC of different antibiotic classes can show up to 500-fold increase in *nfx*B mutants including the MIC of ciprofloxacin which showed 94-fold increase (Chuanchuen et al., 2001). *nfx*B mutations are considered particularly important because it is the only known basis for over-expression of MexCD-OprJ operon which is normally silent in wild-type cells (Li, Plésiat and Nikaido, 2015).

6) *mexS (PA2491) mutations*

*mex*S gene was absent in a total 12 genomes in the whole studied group (11 are ciprofloxacin resistant and 1 is susceptible). The most common mutation reported in the literature was *mex*S N249D, however, this mutation was not identified among the smaller collection of isolates (tested with ciprofloxacin) with an alternative substitution at the same position which was identified in the larger collection (tested with levofloxacin as will be discussed in levofloxacin section below). Similarly, *mex*S Ser124Arg was not identified in the smaller collection but was identified in 88 isolates of the larger collection (shown in levofloxacin section below). *mex*S Val333Gly was identified in only 2 resistant isolates out of 310. *mex*S A75V was identified in 12 out of 310 genomes (9 are susceptible and 3 are resistant). *mex*S E181D was identified in 12 out of 310 genomes (7 are resistant and 5 are susceptible). *mex*S V308I was identified in 13 genomes (8 are resistant) and *mex*S G78S in 6 genomes (5 of which are susceptible). Both *mex*S E54G and *mex*S T152A were not identified in any isolate among the smaller collection evaluated with ciprofloxacin but were identified in the larger collection (shown in levofloxacin section). *mex*S V104A was identified in only 1 resistant isolate among those evaluated for ciprofloxacin resistance. *mex*S F253L, *mex*S L263Q, and *mex*S C269Y previously reported in the literature (Richardot *et al.*, 2016) (Llanes *et al.*, 2011) were not identified in any isolate among the smaller

set but alternative mutations were identified at the same position in the larger study set (shown in levofloxacin section).

7) mexT (PA2492) mutations:

None of the mutations previously identified in two studies (Richardot *et al.*, 2016) (Llanes *et al.*, 2011) for *mex*T gene including *mex*TG257S, *mex*TR166H, *mex*TG258D, and *mex*TY138D were identified in the studied set of isolates.

3) Global transcriptional regulators

1) ampR

The gene was absent in 4 resistant isolates. The most frequently observed mutation in the gene was *amp*R G283E. It was identified in 132 isolates (78 are susceptible). Other frequently observed mutations included *amp*R M288R and *amp*R E114A which were identified in 99 (62 susceptible) and 73 (50 susceptible) isolates respectively. Both *amp*R A51T and *amp*R D135N were less frequently identified in 8 and 5 isolates, respectively. Both G154R and G102E substitutions previously reported in the literature (Cabot *et al.*, 2012)(Caille *et al.*, 2014) were not identified in any single isolate.

2) mvaT (PA4315)

mvaT R80A was identified only in 6 isolates (2 are susceptible and 4 are resistant).

4) SOS-response regulation and related mutants

The *lex*A mutant S125A previously reported in the literature as associated with uncleavable *lex*A (Cirz *et al.*, 2006) and consequently hyper-susceptibility was not detected in any isolate. Other studied mutants that are associated with non-cleavable *lex*A including those at essential sites V88, G91, A90 were similarly not identified. Alternatively, *lon*A499S variant was frequently present in the studied collection and was identified in a total of 83 isolates (65 of which were susceptible).

3.5.2. Predictive values and measures of diagnostic accuracy for known resistance-associated genes and mutations

Table 3.3. Summary of different measures of diagnostic accuracy for genes and mutations in relation to ciprofloxacin resistance

Molecular Marker	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Likelihood Ratio positive (LR+)**	**Likelihood Ratio negative (LR-)**	**Likelihood Ratio (LR)**	**Diagnostic odds ratio (DOR)**	**Youden index**	**Diagnostic Accuracy/effectiveness**	**Frequency of gene/mutation (%)**
*arm*R	**91.1**	59.6	56	**92.2**	**2.25**	0.15	**85.8**	**15.05**	**0.507**	**0.71**	**58.7**
*mex*Z	**97.3**	1.5	35.9	**50**	0.99	1.8	0.5	0.56	-0.012	0.36	**98.1**
*nal*CS46A	**71.4**	81.8	69	**97**	**3.92**	0.35	**217**	**72**	**0.532**	**0.86**	**37.4**
*nal*CG71E	**70.5**	10.1	30.7	**76.9**	0.78	2.92	**60.3**	1.48	-0.194	0.35	**82.9**
*nal*CS209R	**52.7**	25.3	28.5	**65.8**	0.71	1.87	**60.5**	0.77	-0.22	0.39	**66.8**
*amp*R G283E	**48.2**	60.2	40.9	**68.6**	1.211	0.86	11.128	1.51	0.084	0.57	**42.9**
*gyr*A T83I	25	**99.5**	96.6	70.1	**50**	0.75	**54.08**	**65.67**	0.245	**0.73**	9.4
*nfx*B A124T	3.6	**99**	80	65.6	**3.6**	0.97	10.07	**7.61**	0.026	0.66	1.6
*nal*C E153Q	9.8	**97.5**	68.8	72.3	**3.92**	0.93	**70.5**	**5.74**	0.073	**0.72**	5.2
*nal*C Thr50pro	0.9	**99.5**	50	70.1	**1.8**	0.99	**60**	**2.35**	0.004	**0.69**	0.6
*nal*D D187A	3.6	**99**	66.7	68.5	**3.6**	0.97	41.7	**4.35**	0.026	0.68	1.9
*mex*SVal333Gly	1.8	**99.5**	100	66.8	**3.6**	0.99	23.4	ND	0.013	0.67	0.6
*mex*S A175V	6.3	**97.5**	63.6	67.5	**2.52**	0.96	23.34	**3.63**	0.038	0.67	3.5
*mex*S E181D	6.3	**97**	58.3	67.4	**2.1**	0.97	22.24	**2.89**	0.033	0.67	3.9
*mex*S V308I	7.1	**97**	61.5	67.6	**2.37**	0.96	23.5	**3.34**	0.041	0.67	4.2
*mex*R R79S	3.6	**98.5**	57.1	64.6	**2.4**	0.98	3.4	**2.43**	0.021	0.64	2.3
*amp*R A51T	5.4	**99**	75	66	**5.4**	0.96	13.7	**5.8**	0.044	0.66	2.6
*amp*R D135N	3.6	**99.5**	80	65.7	**7.2**	0.97	12.6	**7.65**	0.031	0.66	1.6
*mva*T R80A	3.6	**99**	66.7	64.7	**3.6**	0.973	2.4	**3.66**	0.026	0.65	1.9
*mex*S V104G	0.9	**99.5**	100	66.6	**1.8**	0.995	21.25	ND	0.004	0.67	0.3
*mex*S T152E	0.9	**99.5**	100	66.6	**1.8**	0.995	21.25	ND	0.004	0.67	0.3
*nfx*B Arg82Leu	1.8	97.5	33.3	64.8	0.72	1.007	5.8	0.92	-0.007	0.64	1.9
*nfx*B Arg21His	2.7	94.4	23.1	64.3	0.48	1.03	6.7	0.54	-0.029	0.625	4.2
nfxB Asp56Gly	2.7	93.4	20	64	0.41	1.041	7.5	0.44	-0.039	0.62	4.8
*nal*C A186T	8	92.9	39.1	70.8	1.13	0.99	**60.6**	1.56	0.009	0.68	7.4
*mex*S G78S	0.9	97	16.7	66	0.3	1.021	20	0.39	-0.021	0.65	1.9
*mex*S A75V	2.7	94.9	25	66	0.53	1.025	19.5	0.65	-0.024	0.64	3.9
*nal*D I153Q	1.8	100	100	68.3	ND	0.982	**43.2**	ND	0.018	0.68	0.6
*par*E V460G	0.9	100	100	64.1	ND	0.991	2.04	ND	0.009	0.64	0.3

*par*E S457G	0.9	100	100	64.1	ND	0.991	2.04	ND	0.009	0.64	0.3
*par*C S87W	3.6	100	100	64.7	ND	0.964	8.24	ND	0.036	0.65	1.3
*par*C S87L	12.5	100	100	66.9	ND	0.875	29.7	ND	0.125	0.68	4.5
*par*C E91K	0.9	100	100	64.1	ND	0.991	2.042	ND	0.009	0.64	0.3
*par*C E91L	0.9	100	100	64.1	ND	0.991	2.04	ND	0.009	0.64	0.3
*gyr*A D87N	3.6	100	100	64.7	ND	0.964	8.24	ND	0.036	0.65	1.3
*gyr*B E468D	0.9	100	100	64.3	ND	0.991	2.05	ND	0.009	0.64	0.3
*amp*R M288R	33.6	68	37.4	65.7	1.05	0.98	8.5	1.14	0.016	0.56	32.6
lon A499S (PA1803)	16.2	57.1	21.7	71.5	0.38	1.47	**45.5**	0.69	-0.267	0.54	26.9
*nal*D gene	**100**	16.1	67.8	**100**	1.19	**0**	38.6	ND	0.161	**0.69**	**94.2**
*nal*C gene	**100**	24.1	70	**100**	1.32	**0**	**59.7**	ND	0.241	**0.73**	**91.3**
*mex*S gene	**99.5**	9.8	66.1	**91.7**	1.10	**0.05**	17	21.46	0.093	0.67	**96.1**
*nfx*B gene	**99.5**	4.5	64.8	**83.3**	1.04	**0.11**	5.8	9.21	0.04	0.65	**98.1**
*amp*R gene	**100**	3.6	64.7	**100**	1.04	**0**	8.23	ND	0.036	0.65	**98.7**
*mex*R R79N	**98.5**	8	65.7	**75**	1.07	0.19	10	5.74	0.065	0.66	**95.8**
*mex*R E70R	**99.5**	4.5	65	**83.3**	1.04	**0.11**	7.9	9.29	0.04	0.65	**97.7**
*mex*R L130T	**100**	6.3	65.6	**100**	1.07	**0**	16.7	ND	0.063	0.66	**97.4**
*mex*R G97L	**100**	2.7	64.7	**100**	1.03	**0**	8.23	ND	0.027	0.65	**98.7**
*mex*R L29D	**100**	4.5	65.1	**100**	1.05	**0**	12.43	ND	0.045	0.66	**98.1**
*nal*D D187H	2	**83.9**	100	32.6	0.12	1.17	**41.8**	ND	-0.141	0.34	1.3
*nal*D ser32Asn	0.5	**83.9**	100	32.3	0.03	1.19	39.42	ND	-0.16	0.33	0.3
*par*E A473V	1.5	**100**	100	36.5	ND	0.99	2.7	ND	0.015	0.37	1
*Amp*R E114A	25.3	**75.9**	68.5	36.5	1.049	0.98	8.9	1.25	0.012	0.44	23.5

Numbers in bold show markers with best performance parameters. These were used as input to be tested through the clustering model in an iterative multi-step process. Multiple combinations were tested based on the results shown in this table.

3.5.3. Finding the best possible predictor combinations

Applying cluster analysis using 12 of the previously tested molecular markers showing the best performance (based on the combination of parameters shown in Table 3.3) among the whole group of 48 tested markers as an input to the predictive model revealed five clusters (combination of molecular markers) with improved diagnostic performance. Twelve markers were used as an input and eight of them showed to differentiate the clusters as an output of the analysis (These are shown in Table 3.4). The clusters are based on the probabilities of occurrence of the markers in different categories of behavior according to clustering algorithm (Chiu *et al.*, 2001). These probabilities are shown in Table 3.4. The number of isolates classified as belonging to each cluster are shown with the respective MIC for each cluster in Table 3.5 while the total number of susceptible and resistant isolates belonging to each cluster is shown in Figure 3.1.

Both combinations 1 and 2 showed a tendency to cluster at lower ciprofloxacin MICs while combinations 3, 4, and 5 showed tendency to cluster towards higher ciprofloxacin MICs. The three markers (*nal*CE153Q, *nal*CThr50pro, and *arm*R) showed the highest importance among all predictors in the model. The new clusters showed very high likelihood ratio in differentiating sensitivity from resistance phenotype (LR=283.721, p<0.0005) with strong effect sizes (Phi=0.862, p<0.0005). A summary of all combinations or markers in observed clusters are shown in Table 3.4.

Table 3.4. combination of molecular markers forming new clusters in relation to ciprofloxacin susceptibility and resistance

Cluster	**1**	**5**	**2**	**4**	**3**
Size	**123 (39.8%)**	**61 (19.7%)**	**60 (19.4%)**	**39 (12.6%)**	**26 (8.4%)**
Cipro sensitivity (breakpoint)	**95.9% susceptible**	**67.2% resistant**	**100% susceptible**	**100% resistant**	**100% resistant**
Average ciprofloxacin MIC	0.22	1.39	0.22	4.82	4.62
*nal*C E153Q	Mutation absent 95.9%	Mutation absent 93.4%	Mutation absent 100%	Mutation absent 82.1%	Gene absent 100%
*nal*C Thr50pro	Mutation absent 99.2%	Mutation absent 100%	Mutation absent 100%	Mutation absent 97.4%	Gene absent 100%
arm R	Gene absent 100%	Gene present 98.4	Gene present 100%	Gene present 97.4%	Gene present 92.3%
*nal*D	Gene present 100%	Gene present 98.4%	Gene present 100%	Gene present 97.4%	Gene absent 57.7%
*gyr*A T83I	Mutation absent 99.2%	Mutation absent 100%	Mutation absent 100%	Mutation absent 56.4%	Mutation absent 57.7%
*mex*S	Gene present 99.2%	Gene present 98.4%	Gene present 100%	Gene present 100%	Gene present 61.5%
*mex*Z	Gene present 98.4%	Gene present 98.4%	Gene present 100%	Gene present 100%	Gene present 92.3%
*nfx*B A124T	Mutation absent 95.9%	Mutation absent 96.7%	Mutation absent 100%	Mutation absent 97.4%	Mutation absent 92.3%

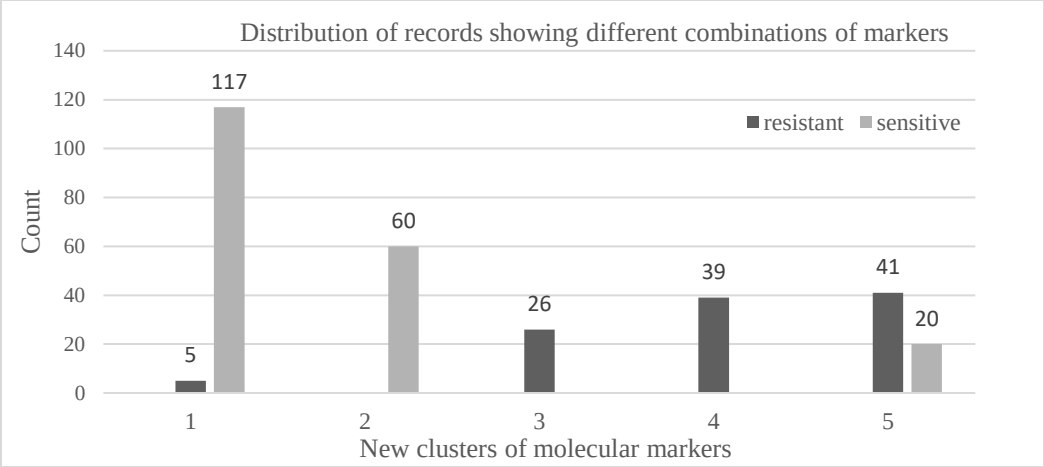


Figure 3.1. Clusters of molecular markers in relation to ciprofloxacin susceptibility and resistance according to clinical breakpoints

Table 3.5. Distribution of observed ciprofloxacin clusters among different categories of behavior in the studied set of isolates

MIC	0.012	0.03	0.06	0.08	0.12	0.25	0.5	1	2	4	8	Total
Cluster 1	1	8	26	4	45	26	7	1	4	0	0	122
Cluster 2	0	4	16	1	15	6	18	0	0	0	0	60
Cluster 3	0	0	0	0	0	0	0	4	10	0	12	26
Cluster 4	0	0	0	0	0	0	0	4	14	3	18	39
Cluster 5	0	0	0	0	0	0	20	7	34	0	0	61
Total	1	12	42	5	60	32	45	16	62	3	30	308

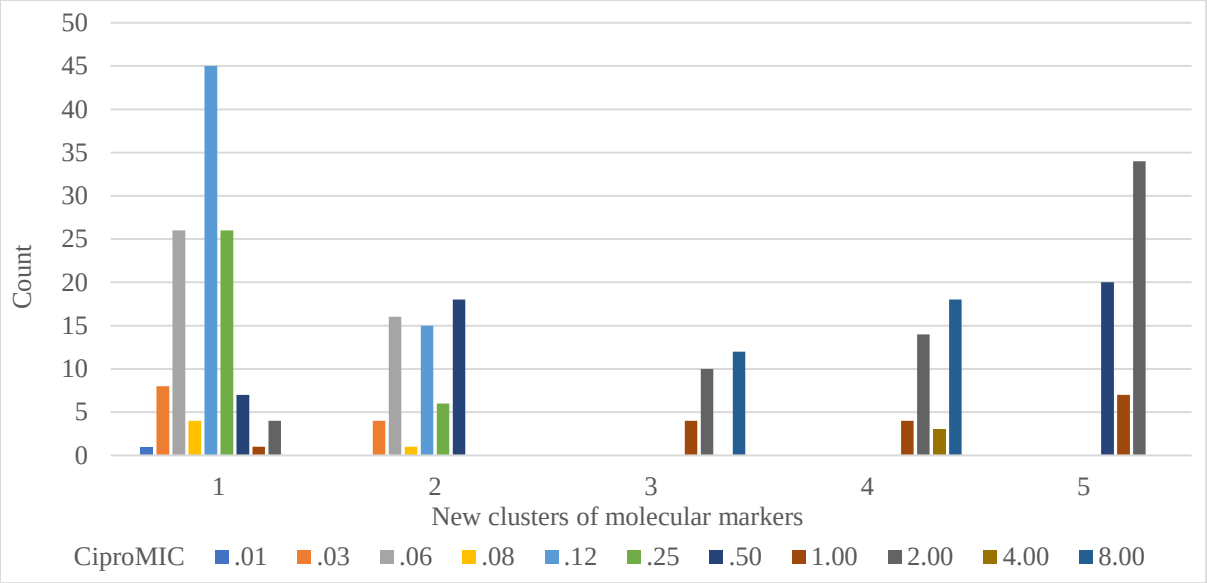


Figure 3.2. Distribution of different ciprofloxacin susceptibility levels within new clusters of molecular markers

3.5.4. Statistical and practical significance of individual molecular markers and new combinations

Table 3.6. Statistical significance and effect sizes for the studied ciprofloxacin molecular markers

	chi square	**significance**	**Phi coefficient**	**significance**	**Cramer' s V**	**significance**
*arm*R	75.8	<0.005	0.494	<0.005	0.494	<0.005
*mex*Z	0.5	0.475	0.041	0.475	0.041	0.475
*nal*C S46A	181.4	<0.005	0.765	<0.005	0.765	<0.005
*nal*C G71E	52.9	<0.005	0.413	<0.005	0.413	<0.005
*nal*C S209R	53.07	<0.005	0.414	<0.005	0.414	<0.005
*amp*R G283E	10.013	0.007	0.18	0.007	0.18	0.007
*gyr*A T83I	50.6	<0.005	0.404	<0.005	0.404	<0.005
*nfx*B A124T	10.33	0.006	0.183	0.006	0.183	0.006
*nal*C E153Q	63.3	<0.005	0.452	<0.005	0.452	<0.005
*nal*C Thr50pro	52.6	<0.005	0.412	<0.005	0.412	<0.005
*nal*D D187A	37	<0.005	0.345	<0.005	0.345	<0.005
*mex*S Val333Gly	22.4	<0.005	0.269	<0.005	0.269	<0.005
*mex*S A175V	23	<0.005	0.272	<0.005	0.272	<0.005
*mex*S E181D	21.9	<0.005	0.266	<0.005	0.266	<0.005
*mex*S V308I	23.13	<0.005	0.273	<0.005	0.273	<0.005
*mex*R R79S	3.17	0.205	0.101	0.205	0.101	0.205
*amp*R A51T	12.9	0.002	0.204	0.002	0.204	0.002

	chi square	**significance**	**Phi coefficient**	**significance**	**Cramer' s V**	**significance**
*amp*R D135N	11.7	0.003	0.194	0.003	0.194	0.003
*mva*T R80A	2.5	0.113	0.09	0.113	0.09	0.113
*mex*S V104G	20.5	<0.005	0.257	<0.005	0.257	<0.005
*mex*S T152E	20.5	<0.005	0.257	<0.005	0.257	<0.005
*nfx*B Arg82Leu	5.9	0.052	0.138	0.052	0.138	0.052
*nfx*B Arg21His	6.8	0.034	0.148	0.034	0.148	0.034
*nfx*B Asp56Gly	7.5	0.024	0.155	0.024	0.155	0.024
*nal*C A186T	53.18	<0.005	0.414	<0.005	0.414	<0.005
*mex*S G78S	19.32	<0.005	0.25	<0.005	0.25	<0.005
*mex*S A75V	19	<0.005	0.247	<0.005	0.247	<0.005
*nal*D I153Q	37.8	<0.005	0.345	<0.005	0.345	<0.005
*par*E V460G	1.8	0.183	0.076	0.183	0.076	0.183
*par*E S457G	1.8	0.183	0.076	0.183	0.076	0.183
*par*C S87W	7.16	0.007	0.152	0.007	0.152	0.007
*par*C S87L	25.92	<0.005	0.289	<0.005	0.289	<0.005
*par*C E91K	1.8	0.183	0.076	0.183	0.076	0.183
*par*C E91L	1.8	0.183	0.076	0.183	0.076	0.183
*gyr*A D87N	7.16	0.007	0.152	0.007	0.152	0.007
*gyr*B E468D	1.8	0.181	0.076	0.181	0.076	0.181
*amp*R M288R	7.4	0.025	0.156	0.025	0.156	0.025
lon A499S (PA1803)	46.6	<0.005	0.388	<0.005	0.388	<0.005
*nal*D gene	33.8	<0.005	0.33	<0.005	0.33	<0.005
*nal*C gene	52.3	<0.005	0.411	<0.005	0.411	<0.005
*mex*S gene	16.7	<0.005	0.232	<0.005	0.232	<0.005
*nfx*B gene	5.9	0.015	0.138	0.015	0.138	0.015
*amp*R gene	7.16	0.007	0.152	0.007	0.152	0.007
*mex*R R79N	10.03	0.007	0.18	0.007	0.018	0.007
*mex*R E70R	7.8	0.021	0.158	0.021	0.158	0.021
*mex*R L130T	14.5	0.001	0.216	0.001	0.216	0.001
*mex*R G97L	0.716	0.028	0.152	0.028	0.152	0.028
*mex*R L29D	10.816	0.004	0.187	0.004	0.187	0.004
*nal*D D187H	35.6	<0.005	0.339	<0.005	0.339	<0.005
*nal*D ser32Asn	34.23	<0.005	0.332	<0.005	0.332	<0.005
*par*E A473V	1.7	0.191	0.074	0.191	0.074	0.191
*amp*R E114A	7.8	0.021	0.158	0.021	0.158	0.021
New clusters	228.88	<0.005	**0.862**	<0.005	**0.862**	<0.005

3.5.5. Multiple regression for best molecular predictors

Stepwise multiple regression was used to assess the ability of all molecular markers assessed above to predict the MIC, after excluding markers with high multicollinearity, the most important predictors in the model were *mex*S N249L, *gyr*A T83I, *mva*T R80A, *mex*Z, *amp*R A51T, *nal*C S46A, and *mex*R R79S.

In the proposed model, 77.2 % of variance in the dependent variable (the MIC) is explained by the predictors in the model ($p<0.0005$). Variables that make significant unique contributions to the prediction of the dependent variable (MIC level) when statistical effect of overlapping variables is excluded includes *gyr*AT83I (*beta*= -0.654, p <0.0005), *nal*CS46A (*beta*= -0.433, p <0.0005), *mex*RR79S (*beta*= 0.176, p <0.0005), and *mva*TR80A (*beta*= -0.111, p <0.011). In this model, *gyr*AT83I uniquely explains 28.8% of the variance in MIC, *nal*C S46A uniquely explains 8.8% of the variance in MIC and *mex*R R79S uniquely explains 1.18 % of the variance in MIC.

3.6. Results Section 2 (levofloxacin)

3.6.1. Description of mutations distribution

1) QRDR

Among 696 genomes for which levofloxacin susceptibility data are available, *gyr*A G83I was also the most frequently observed mutation. The mutation was identified in 186 genomes (two are levofloxacin susceptible). *gyr*A D87N was identified in 31 genomes (all are resistant) and *gyr*B E468D was identified in 8 genomes (all are resistant).

The *par*E mutations were less frequently observed. *par*E A473V was present in 13 genomes out of 696 (8 are resistant and 5 are susceptible to levofloxacin). *par*E V460G was present in four genomes out of 696 (3 are levofloxacin resistant). *Par*E S457G in 5 genomes out of 696 (all are levofloxacin resistant). Similarly, *par*C E91K was rare and identified in only 5 isolates out of 696 in total (4 are resistant). *par*C S87W was identified in 20 isolates (all are resistant). Interestingly, *Par*C S87L was more frequently observed; in 133 out of 698 (all are resistant).

2) Efflux pumps related operons and regulators

1) nalD transcriptional regulator (PA3574)

The gene was absent in 47 isolates out of 696 (all are resistant). Two of the previously identified mutations were not identified in any isolate. These are *nal*D Ser32Asn and the *nal*D Thr158Ile. However, other mutations identified in the tested group included *nal*D D187A which was identified in 9 isolates (5 are levofloxacin susceptible and 4 are resistant). The other mutation was the *nal*D D187H which was identified in 7 isolates and (5 isolates are levofloxacin susceptible and 2 are resistant). Other significant mutations include: *nal*D Ser32Asn which was identified in 9 isolates (all are levofloxacin resistant) and *nal*D I153Q which was identified in 10 isolates (all are levofloxacin resistant).

2) nalC transcriptional regulator (PA3721)

The *nal*C gene was absent in 35 isolates out of 696 (33 are resistant). Three *nal*C mutations annotated by the CARD database were identified and studied. *nal*C G71E was a relatively frequent mutation that was identified in 348 isolates out of 696 (50%). *nal*C S209R was also frequent and was identified in 267 isolates out of 696 (38.36%). In contrast, *nal*C A186T was only identified in 38 isolates (5.46%). Another frequently observed mutation *nal*C S46A was identified in 396 isolates out of 696 while *nal*C E153Q was less frequently observed only in 21 isolates.

3) nalB mutant phenotype: mexR multidrug resistance operon repressor MexR (PA0424)

*mex*R R79S as previously reported in the literature was identified only in 11 isolates (5 are resistant and 6 are susceptible). A novel mutation *mex*R R79N was more frequently observed at the same position in 306 isolates. Both *mex*R G97L and *mex*R L29D were frequently observed variants identified in 325 and 322 isolates respectively. Similarly, each of the two variants; *mex*R E70R and *mex*R L130T was identified in 311 isolates. The *mex*R gene was absent in 7 resistant isolates out of 696 total isolates studied.

4) nfxB mutant

The *nfx*B gene was absent in 37 genomes out of 696. *nfx*B Ala124Glu, which was previously reported in the literature, was not identified in the studied group of genomes, but instead a *nfx*B

A124T variant was identified in 7 isolates; four are levofloxacin resistant and three are susceptible. *nfx*B Arg82leu was identified in 18 isolates (10 are resistant and 8 are susceptible). *nfx*B Arg21His was identified in 29 isolates (21 are susceptible and 8 are resistant). Similarly, *nfx*B Asp56Gly was identified in 31 isolates (23 are susceptible and 8 are resistant). Other mutations including *nfx*B Ser36Gly, *nfx*B Ala38Gly, and *nfx*B Glu111Lys which were previously reported by Higgins *et al.*, (2003) were not identified in the current data set as these appear to be infrequently occurring mutations (detailed explanation shown in ciprofloxacin section).

5) mexS (PA2491) mutations

Many types of *mex*S mutations have been reported as associated with MexEF-OprN efflux pump hyperexpression. These mutations were explored in the studied set of strains. In addition, novel types of substitutions were also observed at the same positions as those previously reported in the literature, and these new types of substitutions were more frequent. The gene was absent in a total 34 isolates in the whole studied group of 696 isolates.

The most common mutation previously reported in the literature was *mex*S N249D. This mutation was not identified among the studied collection of isolates and alternatively *mex*S N249L variant was identified at the same position in 72 isolates. *mex*S Val333Gly was only identified in 3 isolates. *mex*S Ser124Arg was identified in a total of 88 isolates. *mex*S A175V was identified in 18 isolates out of 696 (10 are resistant and 8 are susceptible). *mex*S E181D was similarly identified in 18 isolates (12 are resistant and 6 are susceptible). *mex*S V308I was identified in 11 isolates, *mex*S A75V in 14 isolates, *mex*S G78S in 21 isolates. *mex*S E54G was a more frequent variant which was identified in 77 isolates. Similarly, *mex*S T152A was also a frequent variant identified in 79 isolates. Similarly, *mex*S V104G was identified in 76 isolates. *mex*S F253L, *mex*S L263Q, and *mex*S C269Y previously reported in one study by Richardot *et al.*, (2016) were not encountered in the current study set but alternatively, other variants were identified at the same position including *mex*S F253D variant which was identified in 70 isolates and *mex*S C269D variant which was identified in 73 isolates, a finding that supports the importance of these positions.

6) mexT (PA2492) mutations

Some *mex*T mutations previously shown in the literature including *mex*T G257S, *mex*T R166H, *mex*T G258D, and *mex*T Y138D were not identified in the current set. Both *mex*T Gly257Ser and *mex*T Arg166His reported by Llanes *et al.*, (2011) were only detected in 2 isolates out of 10 isolates showing upregulation of MexEF-OprN with each variant being reported only in one isolate. Llanes *et al.*, (2011) have also shown that these substitutions map outside both the helix-turn-helix motif and the DNA binding domain in *mex*T which would not be expected for these mutations should they affect the MexT function. This gives a possible explanation for not observing these mutations in the studied set of isolates. In addition, Llanes *et al.*, (2011) showed that the extent to which each of the amino acid substitutions in MexS, MexT and/or MvaT may account for the *nfx*C resistance profile is still unclear and thus concluded that regulation of MexEF-OprN is more complex than anticipated and requires further investigation which may indicate that these two mutations were only a bystander finding. *mex*T G258D and *mex*T Y138D were reported by Richardot *et al.*, (2016). The two variants were observed in MexT in only two isolates and were tested for the effect associated with the substitution on MexT activity and showed that complemented strains did not show any effect on MexT function and also concluded that *nfx*C phenotype can be caused by other unknown mutations. None of these mutations have been identified in the current set of isolates. This supports the conclusion about the complex regulatory elements affecting MexT function and activity and the unclear role of these changes on resistance profiles. An important point to consider here is that the *nfx*C mutant overexpressing MexEF-OprN may show mutations in either *mex*T, *mex*S, or *mva*T genes (Llanes et al., 2011) and not necessarily all at the same time because pump overexpression occurs through two pathways; either the *mex*S and *mex*T pathway or the *mva*T dependent pathway. *mva*T dependent pathway can also modify the expression of MexEF-OprN independent of the MexT or MexS-related pathway (Westfall et al., 2006). This all taken together means that observing one variant and not the other is normally expected as two alternative mechanisms/pathways affecting same pump-system regulation. In addition, Richardot et al., (2016) have shown that there are still other unknown loci that appear to be implicated in the pump overproduction in clinical settings which makes it expected to find mutations in either *mex*S, *mex*T, *mva*T, and *amp*R each one separately or in combination with others, the notion that is being explored here.

3) Global transcriptional regulators

1) ampR

The gene was absent in 33 isolates (all are resistant). The most frequently observed mutation in the gene was *amp*R G283E. It was identified in 213 isolates (127 are resistant and 86 are susceptible). Other frequently observed mutations included *amp*R M288R which was identified in 159 isolates (92 are resistant and 67 are susceptible). Both *amp*R A51T and *amp*R D135N were less frequently identified in 11 and 6 isolates respectively. Both G154R reported by Cabot *et al.*, (2012) and G102E reported by Caille *et al.*, (2014) were not identified in the studied set of isolates. This may be an expected finding for different reasons. Caille *et al.*, (2014) has performed functional studies for the effect of both *amp*R D135N and G102E substitution on the activity of AmpR. Both variants were previously implicated in the repression state of AmpR from the *Enterobacteriaceae*. *amp*R D135N was identified in 6 isolates among the current study set (5 of which are levofloxacin resistant and 4 are ciprofloxacin resistant), a finding that supports the findings of Caille *et al.*, (2014) who identified the same substitution in link to inducer-independent increase in the transcriptional activity. In the same study, it has been postulated that D135N substitution in the effector binding domain appears to stabilize the active conformation turning AmpR into an activator of *amp*C transcription. On the other hand, the same study has shown Gly102 substitution to play a different role which was suggested to be a structural role that leads to destabilization of the protein and loss of activity. This substitution was not encountered in the current studied set of isolates probably because the function of AmpR is jeopardized through the other types of variants reported in this study. The *amp*R variant G154R reported by Cabot *et al.*, (2012) can convert AmpR into a transcriptional activator, but was not identified in the current set of isolates. However, the other three variants shown in the same study including *amp*RG283E, *amp*RM288R, and *amp*RA51T were identified with variable frequencies as shown above. A possible explanation is that G154R was more frequently observed in XDR while both *amp*RG283E and *amp*RM288R were shown to co-occur in MDR and moderately resistant isolates and the set included here has variable degrees of susceptibility. An important point to keep in mind while doing such an exploratory analysis that it is not necessary to find every single mutation reported in certain settings in all other settings within the same gene and that was one purpose of doing the current study to find the possible co-contribution of

different regulatory mutations. In addition, for AmpR as a global transcriptional regulator affecting different metabolic functions, different variants can perform different functional roles.

2) mvaT (PA4315)

mvaT R80A was identified only in 9 isolates (4 are resistant and 5 are susceptible).

4) SOS-response regulation and related mutants

The *lex*A S125A reported in the literature to be associated with uncleavable *lex*A and consequently hyper-susceptibility was not detected in any isolate. Other studied mutants that are associated with non-cleavable *lex*A including those at essential sites V88, G91, A90 were similarly not identified. On the other hand, *Lon* A499S variant that is related to SOS response regulation of resistance was identified in 184 isolates (103 susceptible and 81 resistant).

3.6.2. Predictive values and measures of diagnostic accuracy for known resistance-associated genes and mutations

Table 3.7. Summary of different measures of diagnostic accuracy for genes and mutations related to levofloxacin resistance

Molecular Marker	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Likelihood Ratio positive (LR+)**	**Likelihood Ratio negative (LR-)**	**Likelihood Ratio (LR)**	**Diagnostic odds ratio (DOR)**	**Youden index**	**Diagnostic Accuracy/effectiveness**	**Frequency of gene/mutation (%)**
*nfx*B gene	**99.7**	9.4	47.8	**97.3**	1.10	0.032	34.48	**32.97**	0.091	0.50	**94.7**
*mex*R gene	**100**	1.8	45.9	**100**	1.02	0	ND	ND	0.018	0.46	**99**
*mex*S gene	**99.7**	8.7	47.6	**97.1**	1.092	0.034	31.67	**29.97**	0.084	0.5	**95.1**
*nal*C gene	**99.4**	8.7	47.5	**94.3**	1.089	0.069	15.79	**14.94**	0.081	0.49	**95**
*nal*D gene	**100**	12.3	48.7	**100**	1.14	0	ND	ND	0.123	0.52	**93.3**
*Amp*R gene	**100**	8.7	47.7	**100**	1.095	0	ND	ND	0.087	0.50	**95.3**
*mex*Z	**99.1**	9.7	47.7	**92.5**	1.097	0.093	11.83	**11.26**	0.088	0.50	**94.3**
*gyr*T83I	48.3	**99.4**	98.9	61.5	**80.5**	0.52	154.77	**147.1**	0.477	0.71	26.6
*nfx*BA124T	1	**98.7**	57.1	47.9	0.769	1.003	0.77	1.22	-0.003	0.48	1
*nfx*BArg82Leu	2.6	**97.2**	55.6	47.9	0.93	1.002	0.93	1.149	-0.002	0.48	2.6
*nal*CE153Q	4.5	**98.4**	81	48.5	**2.81**	0.97	**2.89**	**4.006**	0.029	0.49	3
*nal*Dser32Asn	2.4	**100**	100	49.4	ND	0.976	ND	ND	0.024	0.50	1.3
*nal*DI153Q	2.6	**100**	100	49.5	ND	0.974	ND	ND	0.026	0.50	1.4

*mex*RR79S	1.3	**98.1**	45.5	45.7	0.68	1.006	0.68	0.70	-0.006	0.46	1.6
*mex*SA75V	2.4	**98.1**	64.3	47.8	1.26	0.995	1.27	1.651	0.005	0.48	2
*mex*SA175V	2.6	**97.2**	55.6	47.7	0.929	1.002	0.927	1.139	-0.002	0.48	2.6
*mex*SE181D	3.1	**97.8**	66.7	48	1.41	0.99	1.422	1.85	0.009	0.48	2.6
*mex*SV308I	1	**97.5**	36.4	47.3	0.4	1.015385	0.393	0.513	-0.015	0.47	1.6
*nfx*BArg21His	2.1	**93.1**	27.6	46.7	0.304	1.05	0.289	0.33	-0.048	0.46	4.2
*nfx*B Asp56Gly	2.1	**92.4**	25.8	46.5	0.28	1.059	0.26	0.30	-0.055	0.46	4.4
*par*EA473V	2.1	98.4	61.5	45.5	1.31	0.995	1.32	1.34	0.005	0.46	1.9
*par*EV460G	0.8	**99.7**	75	45.5	**2.67**	0.995	**2.68**	2.51	0.005	0.46	0.6
*par*ES457G	1.3	**100**	100	45.7	ND	0.987	ND	ND	0.013	0.46	0.7
*par*CS87W	5.2	**100**	100	46.8	ND	0.948	ND	ND	0.052	0.48	2.9
*par*CS87L	34.9	**100**	100	55.8	ND	0.651	ND	ND	0.349	0.64	19.2
*par*CE91K	1	**99.7**	80	45.6	**3.33**	0.99	**3.36**	3.35	0.007	0.46	0.7
*par*CE91V	0.8	**100**	100	45.6	ND	0.992	ND	ND	0.008	0.46	0.4
*gyr*AD87N	8.1	**100**	100	47.5	ND	0.919	ND	ND	0.081	0.498	4.4
*gyr*BE468D	2.1	**100**	100	46	ND	0.979	ND	ND	0.021	0.47	1.1
*amp*RA51T	1.6	98.4	54.5	47.9	1	1	1	1.10	0	0.48	1.6
*amp*RD135N	1.3	**99.7**	83.3	48	**4.33**	0.99	**4.38**	**4.62**	0.01	0.48	0.9
*mex*SVal333Gly	0.5	99.4	66.7	47.7	0.83	1.001	0.83	1.82	-0.001	0.48	0.4
*mex*SSer124Arg	13.4	88	58	48.4	1.12	0.98	1.13	1.29	0.014	0.49	12.6
*mex*SG78S	2.1	95.6	38.1	47.1	0.48	1.024	0.47	0.55	-0.023	0.47	3
*mex*SN249L	9.7	88.6	51.4	47.5	0.85	1.019	0.83	0.96	-0.017	0.48	10.3
*mex*SV104G	11.1	89	55.3	48	1.009	0.99	1.01	1.14	0.001	0.49	10.9
*mex*SF253D	9.4	89	51.4	47.5	0.85	1.018	0.84	0.96	-0.016	0.48	10
*mex*SC269D	10.5	89.3	54.8	47.9	0.98	1.002	0.98	1.11	-0.002	0.49	10.5
*mex*SE54V	11.5	89.3	57.1	48.2	1.075	0.9	1.08	1.24	0.008	0.49	11
*mex*SG78A	11.5	89.3	57.1	48.2	1.075	0.99	1.08	1.24	0.008	0.49	11
*mex*ST152E	11.5	88.6	55.7	48	1.008	0.99	1.009	1.16	0.001	0.49	11.3
*amp*RG283E	33.3	72.7	59.6	50.9	1.22	0.92	1.33	1.53	0.06	0.54	30.6
*amp*RM288R	24.3	78.7	57.9	49.3	1.14	0.96	1.19	1.34	0.03	0.51	22.9
*nal*CG71E	44.9	43.5	49.1	43.8	0.79	1.27	0.63	0.75	-0.116	0.47	49.9
*nal*CA186T	6	94.6	60.5	47.9	1.11	0.99	1.12	1.41	0.006	0.49	5.5
*nal*C S209R	33.9	55.8	48.3	44.7	0.77	1.19	0.65	32	-0.103	0.46	38.3
*mex*R R79N	50.8	60.1	52.6	59.5	1.27	0.82	1.56	18.6	0.109	0.56	43.8
*mex*R E79R	52.4	60.1	53.4	60.3	1.31	0.79	1.66	21.4	0.125	0.57	44.6
*mex*R E70R	52.4	60.1	53.4	60.3	1.31	0.79	1.66	21.4	0.125	0.57	44.6
*mex*R L130T	53	60.6	54	60.8	1.35	0.78	1.73	23.7	0.136	0.58	44.6
*mex*R G97L	53	57	51.7	59.3	1.23	0.82	1.49	17	0.1	0.56	46.6
*mex*R L29D	53	58	52.2	59.7	1.26	0.81	1.56	17.2	0.11	0.56	46.1
*amp*R E114A	16.4	80.3	55.3	53.6	0.83	1.04	0.79	43.9	-0.033	0.54	13.5
*mva*T R80A	1.6	98.9	55.6	54.6	1.45	0.99	1.46	0.366	0.005	0.55	1.3
*lon*A499S	32.5	58.5	56	53.6	0.78	1.15	0.68	33	-0.09	0.54	26.4
*arm*R gene	**82.4**	29.7	58.5	**58.4**	1.17	0.59	1.98	14.2	0.121	0.58	**76.9**
*nal*C S46A	**64.3**	52.1	61.9	**61.6**	1.34	0.69	1.96	**69.2**	0.164	0.62	**56.7**

Numbers in bold represent markers with best performance parameters. These were used as an input to test using the clustering model in an iterative multi-step process. Multiple combinations were tested based on the results shown in the table above.

3.6.3. Finding the best possible predictor combinations

Applying cluster analysis using 13 of the previously tested molecular markers showing the best performance among the whole group of 57 tested markers as an input to the predictive model revealed five clusters (combination of molecular markers) with improved diagnostic performance. The markers differentiating these clusters included 10 markers as an output (These markers are shown in Table 3.8) Cluster 1 and cluster 5 showed tendency to cluster at lower levofloxacin MICs (susceptible isolates) while cluster 2, cluster 3, and cluster 4 showed tendency to cluster towards higher levofloxacin MICs (resistant isolates). The five markers (*nalC S46A, nalC E153Q, ampR D135N, nfxB, mexZ*) showed the highest importance among all predictors in the model.

The new clusters showed very high likelihood ratio in differentiating sensitivity from resistance phenotype (LR=931.639, p<0.0005) with strong effect sizes (Phi=0.991, p<0.0005). Summary of all combination or markers in observed clusters are shown in Table 3.8.

Table 3.8. Combinations of molecular markers forming the new clusters in relation to levofloxacin susceptibility and resistance

Cluster	**4**	**1**	**5**	**2**	**3**
Size	**214 (30.7%)**	**164 (23.5%)**	**150 (21.5%)**	**138 (19.8%)**	**31 (4.4%)**
Levofloxacin sensitivity (breakpoint)	**100% resistant**	**100% susceptible**	**100% susceptible**	**97.8% resistant**	**100% resistant**
Average levofloxacin MIC	16.88	0.56	0.54	11.05	4.13
*nal*C E153Q	Mutation absent 100%	Mutation absent 97.6%	Mutation absent 100%	Mutation absent 85.5%	Gene absent 100%
nalC gene	Gene present 100%	Gene present 99.4%	Gene present 100%	Gene present 97.8%	Gene absent 100%
*amp*R D135N	Mutation absent 100%	Mutation absent 100%	Mutation absent 99.3%	Mutation absent 94.2%	Gene absent 93.5%
*nfx*B gene	Gene present 100%	Gene present 100%	Gene present 100%	Gene present 94.9%	Gene absent 96.8%
mexZ gene	Gene present 100%	Gene present 100%	Gene present 99.3%	Gene present 94.2%	Gene absent 100%
armR gene	Gene present 100%	Gene absent 56.1%	Gene present 99.3%	Gene present 72.5%	Gene absent 96.8%
*mex*S gene	Gene present 100%	Gene present 100%	Gene present 99.3%	Gene present 91.3%	Gene absent 64.5%
*mex*R gene	Gene present 100%	Gene present 100%	Gene present 100%	Gene present 96.4%	Gene present 93.5%

*par*C E91K	Mutation absent 100%	Mutation absent 100%	Mutation absent 100%	Mutation absent 97.1%	Mutation absent 96.8%
*par*E V460G	Mutation absent 100%	Mutation absent 100%	Mutation absent 99.3%	Mutation absent 97.8%	Mutation absent 100%

It is important to note here that the size of the clusters identified by the clustering algorithm is automatically determined in the two-step method, because markers do not occur in the population with the same frequency. Some markers are rare and not commonly encountered in the population. The algorithm aims at finding observable or significant patterns based on the difference in the percentage of presence or absence of specific markers. When the variant is more frequent in the population, the difference in distribution is more observable.

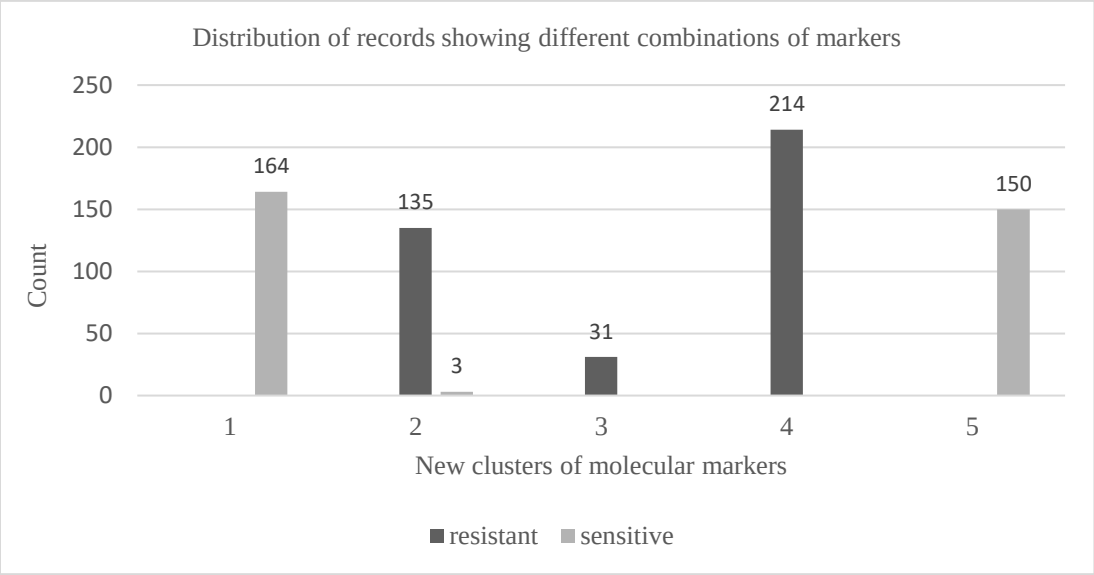


Figure 3.3. Clusters of molecular markers in relation to levofloxacin susceptibility and resistance according to clinical breakpoints

Table 3.9. Distribution of observed levofloxacin clusters in different categories of behavior among the studied set of isolates

MIC	0.012	0.03	0.06	0.12	0.25	0.5	1	2	4	8	16	32	128	Total
Cluster 1	0	0	2	9	37	68	48	0	0	0	0	0	0	164
Cluster 2	0	0	0	0	0	1	2	33	40	24	7	31	0	138
Cluster 3	0	0	0	0	0	0	0	0	30	1	0	0	0	31
Cluster 4	0	0	0	0	0	0	0	36	35	43	15	84	1	214
Cluster 5	3	3	3	6	42	48	45	0	0	0	0	0	0	150
Total	3	3	5	15	79	117	95	69	105	68	22	115	1	697

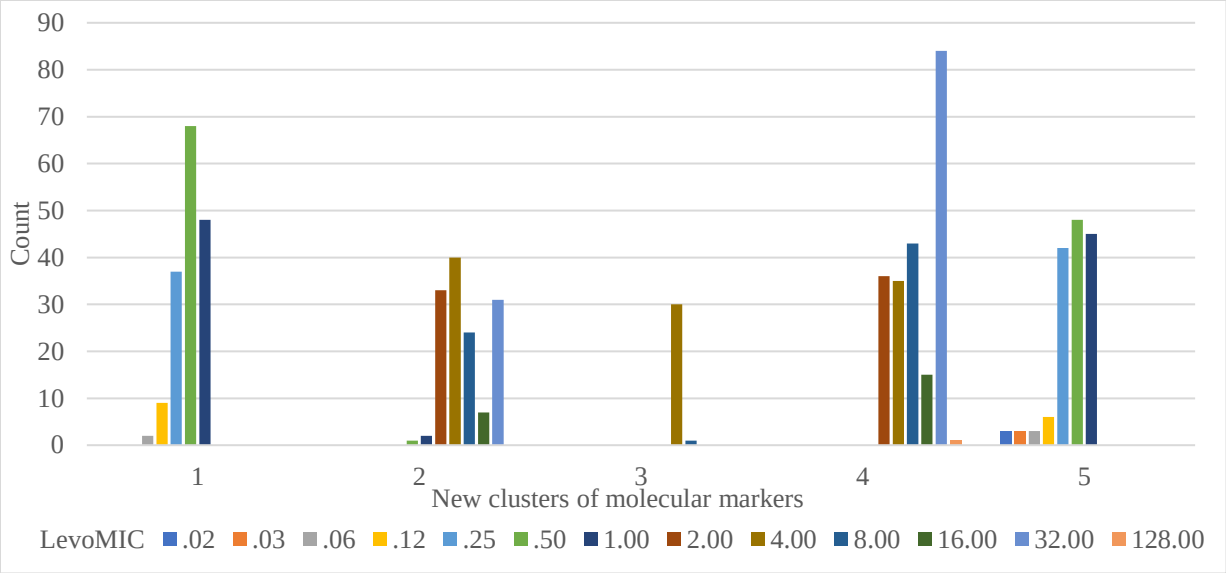


Figure 3.4. Distribution of different levofloxacin susceptibility levels within new clusters of molecular markers

3.6.4. Statistical and practical significance of individual molecular markers and new combinations

Table 3.10. Statistical significance and effect sizes for different studied levofloxacin molecular marker

	chi square	**significance**	**Phi coefficient**	**significance**	**Cramer' s V**	**significance**
*nfx*B gene	28.8	<0.005	0.203	<0.005	0.203	<0.005
*mex*R gene	5.9	0.015	0.092	0.015	0.092	0.015
*mex*S gene	26.01	<0.005	0.0193	<0.005	0.193	<0.005
*nal*C gene	23.4	<0.005	0.183	<0.005	0.183	<0.005
*nal*D gene	41.9	<0.005	0.245	<0.005	0.245	<0.005
*amp*R gene	28.8	<0.005	0.203	<0.005	0.203	<0.005
*mex*Z	24.6	<0.005	0.188	<0.005	0.188	<0.005
*gyr*T83I	201	<0.005	0.537	<0.005	0.537	<0.005
*nfx*BA124T	29	<0.005	0.203	<0.005	0.203	<0.005
*nfx*BArg82Leu	28.8	<0.005	<0.005	<0.005	<0.005	<0.005
*nal*CE153Q	33.1	<0.005	0.218	<0.005	0.218	<0.005
*nal*Dser32Asn	51	<0.005	0.269	<0.005	0.269	<0.005
*nal*DI153Q	52	<0.005	0.272	<0.005	0.272	<0.005
*mex*R R79S	6.2	0.045	0.094	0.045	0.094	0.045
*mex*SA75V	27	<0.005	0.196	<0.005	0.196	<0.005
*mex*S A175V	26	<0.005	0.193	<0.005	0.193	<0.005
*mex*SE181D	27.5	<0.005	0.199	<0.005	0.199	<0.005

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*mex*SV308I	27.2	<0.005	0.197	<0.005	0.197	<0.005
*nfx*BArg21His	36.2	<0.005	0.228	<0.005	0.228	<0.005
*nfx*B Asp56Gly	38	<0.005	0.233	<0.005	0.233	<0.005
*par*EA473V	0.26	0.611	0.019	0.611	0.019	0.611
*par*EV460G	0.676	0.411	0.031	0.0411	0.031	0.411
*par*ES457G	4.2	0.041	0.077	0.041	0.077	0.041
*par*CS87W	17.13	<0.005	0.157	<0.005	0.157	<0.005
*par*CS87L	135.2	<0.005	0.441	<0.005	0.441	<0.005
*par*CE91K	1.3	0.252	0.043	0.252	0.043	0.252
*par*CE91V	2.5	0.113	0.06	0.113	0.06	0.113
*gyr*AD87N	27	<0.005	0.197	<0.005	0.197	<0.005
*gyr*BE468D	6.8	0.009	0.098	0.009	0.098	0.009
*amp*RA51T	29	<0.005	0.204	<0.005	0.204	<0.005
*amp*RD135N	31.3	<0.005	0.212	<0.005	0.212	<0.005
*mex*SVal333Gly	26.2	<0.005	0.194	<0.005	0.194	<0.005
*mex*S ser124Arg	27.2	<0.005	0.198	<0.005	0.198	<0.005
*mex*SG78S	28	<0.005	0.2	<0.005	0.2	<0.005
*mex*SN249L	26	<0.005	0.193	<0.005	0.193	<0.005
*mex*SV104G	26.4	<0.005	0.195	<0.005	0.195	<0.005
*mex*SF253D	26	<0.005	0.193	<0.005	0.193	<0.005
*mex*SC269D	26.2	<0.005	0.194	<0.005	0.194	<0.005
*mex*SE54V	27	<0.005	0.196	<0.005	0.196	<0.005
*mex*SG78A	27	<0.005	0.196	<0.005	0.196	<0.005
*mex*ST152E	26.4	<0.005	0.194	<0.005	0.194	<0.005
*amp*RG283E	35.09	<0.005	0.225	<0.005	0.225	<0.005
*amp*RM288R	31.2	<0.005	0.212	<0.005	0.212	<0.005
*nal*CG71E	27	<0.005	0.196	<0.005	0.196	<0.005
*nal*CA186T	24.4	<0.005	0.187	<0.005	0.187	<0.005
*nal*CS209R	27	<0.005	0.195	<0.005	0.195	<0.005
*mex*R R79N	16	<0.005	0.151	<0.005	0.151	<0.005
*mex*RE79R	18.7	<0.005	0.164	<0.005	0.164	<0.005
*mex*RE70R	18.7	<0.005	0.164	<0.005	0.164	<0.005
*mex*RL130T	21	<0.005	0.174	<0.005	0.174	<0.005
*mex*RG97L	14.3	<0.005	0.143	0.001	0.143	0.001
*mex*RL29D	14.9	0.001	0.146	0.001	0.0146	0.001
*amp*RE114A	31.4	<0.005	0.212	<0.005	0.212	<0.005
*mva*T R80A	0.37	0.544	0.023	0.544	0.023	0.544
*lon*A499S	31.2	<0.005	0.211	<0.005	0.211	<0.005
*arm*R gene	14.2	<0.005	0.143	<0.005	0.143	<0.005
*nal*CS46A	61.4	<0.005	0.297	<0.005	0.297	<0.005

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*nfx*B gene	28.8	<0.005	0.203	<0.005	0.203	<0.005
New clusters	685.164	<0.005	**0.991**	<0.005	**0.991**	<0.005

3.6.5. Multiple regression for best molecular predictors

Stepwise multiple regression was used to assess the ability of all molecular markers assessed above to predict level of MIC, after excluding markers with high multicollinearity, the most important predictors in the model included *Predictors gyr*A T83I, *par*C S87L, *par*C S87W, *mex*R L130T, *par*C E91K, *nal*C G71E, *gyr*A D87N, *mex*R E70R, *mex*S Ser124Arg, and *mex*S G78S. In the proposed model, 60.9 % of variance in the dependent variable (MIC level) is explained by the predictors in the model ($p<0.0005$). Variables that make significant unique contribution to the prediction of the dependent variable (MIC level) includes *gyr*A T83I (*beta*= -0.388, *p* <0.0005), *par*C S87L (*beta*= -0.341, *p* <0.0005), *mex*R L130T (*beta*= 0.254, *p* <0.009), *mex*R E70R (*beta*= -0.208, *p* <0.031), and *par*C S87W (*beta*= -0.126, *p* <0.0005). In this model, *gyr*A T83I uniquely explains 4.2 % of the variance in MIC and *par*C S87L uniquely explains 3.4 % of the variance in MIC.

3.7. Results Section 3 (Amikacin and Gentamycin)

3.7.1. Description of mutation distribution

The *rrs* gene in *Ps. aeruginosa* was explored for mutations similar to those reported in the literature with *M. tuberculosis*. Among those variants previously reported (shown in Table 3.2), five mutations were identified in the studied set of isolates, however, these were very infrequent. Both A907C and A514C were identified only in one isolate. Both A514C and G1491T were identified in two isolates. Variant C522T was identified in four isolates.

However, *fus*A1 gene coding for elongation factor appeared to be an important determinant of resistance. Among seven mutations previously reported in the literature (shown in Table 3.2), the variant *fus*A1Y552C was identified in 7 isolates, 2 of them are gentamycin resistant, 5 are amikacin resistant, and 2 are amikacin susceptible while the *fus*A1 Ala555Glu mutation which is

located just 2 nucleotides at the start of its containing beta strand was identified only in one amikacin resistant isolate. This may indicate that *fus*A1Y552C is probably performing a more important functional roles being located just at the bend between the previous alpha helix and the subsequent beta strand (Nyfeler *et al.*, 2012).

Bolard, Plésiat and Jeannot, (2018) have shown that three in vitro mutants in *fus*A1 including Arg371Cys in domains II, Thr456Ala in domain III, and Arg680Cys in domain V have resulted in increased MIC to some tested aminoglycosides. By scanning the current studied set of isolates, two of these *in vitro* mutations have been identified in the clinical set of isolates including *fus*A1 Thr456Ala which was identified in 2 amikacin resistant isolates and *fus*A1 Arg680Cys which was identified in one isolate that showed resistance to both gentamycin and amikacin. The mutation mapped just at the end of the turn between the previous alpha helix and the subsequent beta strand. These results combined together confirm the role of multiple mutations in EF-G1A in conferring aminoglycoside resistance.

*fus*A1 Thr671Ala mutation located at domain V at the start of an alpha helix is probably affecting the gene function (Bolard, Plésiat and Jeannot, 2018). In the current set, the variant was identified in one gentamycin resistant but amikacin susceptible isolate. Interestingly, a novel variant, *fus*A1 D588G which map 2 nucleotides before the end of its containing beta strand at position 590 (Nyfeler *et al.*, 2012) was identified in 5 isolates; 3 are resistant to both gentamycin and amikacin and 2 are resistant to amikacin. This variant was predicted as deleterious at - 2.5 cutoff (PROVEAN score = -6.198).

*rps*L appeared to be conserved in the studied set of isolates with no significant variation observed. The *gid*B nucleotide variants T230C, C286T, T104G, and A254G reported in the literature were not identified among the studied isolates. Instead, other novel amino acid variants were observed. Three AA substitutions predicted as deleterious at -1.3 cutoff were identified. These included E126G (PROVEAN score= -2.779), E97Q (PROVEAN score= -2.240), and Q28K (PROVEAN score= -1.442). *gid*B E126G was identified in 8 isolates, all are amikacin susceptible, 3 are gentamycin susceptible, and 1 is gentamycin resistant. The position of this variant when mapped to the secondary structure of *gid*B methyltransferase from *Bacillus Subtilis* showed to occur at the position of the bend between the end of the third beta strand and the following alpha helix which may indicate its possible associated functional role. *gid*B Q28K

which occurs at the end of the second alpha helix (Zhang, R., Wu, R., Collart, F., Joachimiak, 2004) was identified in 9 isolates, all are amikacin susceptible and 4 are gentamycin resistant. *gid*B E97Q was identified in 54 isolates, 11 are gentamycin resistant and 3 are gentamycin susceptible, 22 are amikacin resistant and 32 are amikacin susceptible. *gid*B E186A was identified in 54 isolates, 9 are gentamycin resistant and 7 are gentamycin susceptible, 22 are amikacin resistant and 30 are amikacin susceptible.

Among the *amg*S variants previously reported in the literature, *amg*S D106N was identified in 3 isolates; 1 isolate was resistant to both gentamycin and amikacin, 1 isolate was susceptible to both gentamycin and amikacin, and 1 isolate was susceptible to amikacin. *amg*S V121G was identified in 1 isolate which was resistant to both gentamycin and amikacin. The other three variants extracted from the literature (shown in Table 3.2) were not identified. Interestingly, the novel variant E108Q was identified in 11 isolates; all are amikacin susceptible (3 of which are susceptible to both gentamycin and amikacin and 5 are resistant to gentamycin but susceptible to amikacin). This novel variant was predicted as deleterious at a cutoff of -1.3, PROVEAN score = -1.774.

In contrast to the genes explored and summarized above, *pmr*A-*pmr*B aminoglycosides resistance-related genes showed more frequent variations. *pmr*B showed frequent occurrence of the same type of variants reported in the literature. The five variants Leu323His, Ser420Arg, Gly423Cys, L243Q, and A248V occurred together in 112 isolates. Thirty isolates were susceptible to both gentamycin and amikacin. Thirty-eight isolates were susceptible to amikacin, eleven were susceptible to amikacin and resistant to gentamycin. Twenty-two were resistant to amikacin and ten were resistant to both gentamycin and amikacin. *pmr*B Ala4Thr was identified in 117 isolates; 27 were amikacin resistant and 12 were gentamycin resistant. *pmr*A Asp92Tyr previously reported in the literature was not identified in the studied set of isolates, while *pmr*A Leu71Arg was identified in 178 isolates (37 are amikacin resistant and 30 are gentamycin resistant). Similarly, *rpl*Y Ala123Ser was identified in 217 isolates (47 are amikacin resistant and 34 are gentamycin resistant). Novel variant *rpl*Y Q41L was identified in 6 isolates; all are amikacin and gentamycin susceptible and only 1 isolate is gentamycin resistant. Q41L was predicted as deleterious at -2.5 cutoff (PROVEAN score= -3.098).

By exploring other aminoglycosides resistance-related genes, some novel variants were identified in *pho*P, *pho*Q, *nuo*G, *pst*P, *lpt*A, *fao*A, and *arn*ABCD. *pho*P P31Q was identified in 9 isolates (8 are amikacin susceptible, 1 is amikacin resistant and 3 are gentamycin resistant). *pho*Q Y85F was identified in 69 isolates (42 are amikacin susceptible, 27 are amikacin resistant, 13 are gentamycin resistant, and 15 are gentamycin susceptible). Two other variants in *pho*Q S300R and L331Q predicted to be deleterious (PROVEAN score: -1.562 and -2.182) were identified in 13 isolates and 4 isolates respectively. Among 13 isolates with *pho*Q S300R, 10 were amikacin susceptible (5 of which were susceptible to both amikacin and gentamycin). *nuo*G S468A was identified in 118 isolates (82 are amikacin susceptible and 29 of which are susceptible to both amikacin and gentamycin). *nuo*G A574T variant predicted as deleterious (PROVEAN score= -1.940) was identified in 36 isolates (34 are amikacin susceptible and 10 of which were susceptible to both amikacin and gentamycin). *nuo*G A890T was identified in 19 isolates (4 are susceptible to both amikacin and gentamycin, 18 are amikacin susceptible and 1 isolate is amikacin resistant). Two novel variants were identified in *pst*B and both were predicted as deleterious. These include *pst*B R87C (PROVEAN score= -6.874) and *pst*B E89Q (PROVEAN score= -1.865). *pst*B R87C was identified only in 4 isolates (all are gentamycin and amikacin susceptible). On the other hand, *pst*B E89Q was identified in 20 isolates (14 are amikacin resistant, 6 are amikacin susceptible, 5 are gentamycin resistant, and 2 are gentamycin susceptible).

*lpt*A T55A was identified in 53 isolates (6 are amikacin resistant, 47 are amikacin susceptible and 18 were susceptible to both amikacin and gentamycin). *lpt*A R62S predicted as deleterious (PROVEAN score= -3.392) was identified in 28 isolates (27 are amikacin susceptible while 12 were susceptible to both amikacin and gentamycin).

*fao*A T385A predicted as deleterious (PROVEAN score= -3.272) was identified in 16 isolates (8 are susceptible to both gentamycin and amikacin, 15 are amikacin susceptible and 1isolate is amikacin resistant). Two other novel variants observed in *arn*ABCD operon included: *arn*A A170T (PROVEAN score= -2.065) and *arn*D G206C (PROVEAN score= -8.374). Both were predicted as deleterious and were identified in 24 isolates. *arn*A A170T was identified in 9 amikacin susceptible isolates, 15 amikacin resistant isolates, 8 gentamycin resistant isolates and 2 gentamycin susceptible isolates. *arn*D G206C was identified in 10 amikacin susceptible

isolates, 14 amikacin resistant isolates, 7 gentamycin resistant isolates and 2 gentamycin susceptible isolates.

3.7.2. Predictive values and measures of diagnostic accuracy for known resistance-associated genes and mutations

Table 3.11. Summary of different measures of diagnostic accuracy for genes and mutations related to gentamycin resistance

Molecular Marker	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Likelihood Ratio positive (LR+)**	**Likelihood Ratio negative (LR-)**	**Likelihood Ratio (LR)**	**Diagnostic odds ratio (DOR)**	**Youden index**	**Diagnostic Accuracy/effectiveness**	**Frequency of gene/mutation (%)**
*nal*C	**100**	1.5	35.7	**100**	1.015	0	2.632	ND	0.015	0.36	99
*mex*Z	**100**	2.1	35.8	**100**	1.021	0	**3.517**	ND	0.021	0.37	98.7
*amp*R	**100**	1.5	35.7	**100**	1.015	0	2.632	ND	0.015	0.36	99
*mex*RG97L	**99.5**	0.9	64.8	50	1.004	0.56	0.181	1.84	0.004	**0.65**	99.3
*mex*RL29D	**99.5**	1.9	65	66.7	1.014	0.26	1.227	**3.71**	0.014	**0.65**	99
*mex*RE70R	**97.4**	4.7	65.2	50	1.022	0.55	0.932	1.87	0.021	**0.65**	96.7
*mex*RL130T	**97.4**	3.8	64.9	44.4	1.012	0.68	0.327	1.48	0.012	**0.64**	97
*mex*RR79N	**96.9**	5.7	65.3	50	1.028	0.54	1.127	1.88	0.026	**0.65**	96
*nal*D	**95.3**	4.1	35.2	61.5	0.993	1.15	0.058	0.87	-0.006	0.36	95.7
*nal*CG71E	90.2	8.5	64.3	32.1	0.986	1.15	0.139	0.85	-0.013	**0.61**	90.7
*nal*CS209R	72.7	27.4	64.7	35.4	1.0014	0.99	0	1.00	0.001	0.57	72.7
*arm*R	59.4	41.8	35.8	65.3	1.021	0.97	0.04	1.01	0.012	0.47	58.7
*amp*RG283E	46.2	59.8	38.1	67.4	1.149	0.89	0.978	1.27	0.06	0.55	42.3
*nal*CS46A	43.4	64.4	40	67.6	1.219	0.88	1.767	1.39	0.078	0.57	38.3
*amp*RM288R	34	68.4	36.5	66	1.08	0.96	0.172	1.11	0.024	0.56	32.4
*rpl*YAla123Ser	33	66.5	35	64.5	0.99	1.01	0.007	0.98	-0.005	0.55	33.3
*lon*A499S	30.9	81.1	75	39.1	**1.63**	0.85	**5.29**	**1.93**	0.12	0.49	26.7
*pmr*ALeu71Arg	28.3	74.7	38	65.6	1.12	0.96	0.325	1.17	0.03	0.58	26.3
*amp*RE114A	23.2	78.3	66.2	35.8	1.07	0.98	0.088	1.09	0.015	0.43	22.7
*pmr*BLeu323His	19.8	84	40.4	65.7	1.24	0.95	0.692	1.29	0.038	0.61	17.3
*pmr*BSer420Arg	19.8	84	40.4	65.7	1.24	0.95	0.692	1.29	0.038	0.61	17.3
*pmr*BGly423Cys	19.8	84.5	41.2	65.9	1.28	0.95	0.902	1.35	0.043	0.61	17
*pmr*BL243Q	19.8	84.5	41.2	65.9	1.28	0.95	0.902	1.35	0.043	0.61	17
*pmr*BA248V	19.8	84	40.4	65.7	1.24	0.95	0.692	1.29	0.038	0.61	17.3
*nuo*GS468A	19.8	84	40.4	65.7	1.24	0.95	0.692	1.29	0.038	0.61	17.3
*pmr*BALA4Thr	12.3	85.6	31.7	64.1	0.85	1.02	0.277	0.83	-0.021	0.59	13.7
*pho*QY85F	12.3	92.3	46.4	65.8	1.59	0.95	1.609	1.67	0.046	**0.64**	9.3

*nal*CE153Q	11.3	**95.9**	60	66.4	**2.76**	0.92	**5.406**	**2.97**	0.072	**0.66**	6.7
*nal*CA186T	10.4	93.3	45.8	65.6	1.55	0.96	1.218	1.61	0.037	**0.64**	8
*gid*BE97Q	10.4	**97.4**	68.8	66.5	**4**	0.92	**7.816**	**4.38**	0.078	**0.67**	5.3
*gid*BE186A	10.4	**97.4**	68.8	66.5	**4**	0.92	**7.816**	**4.38**	0.078	**0.67**	5.3
*lpt*AT55A	8.8	94.3	73.9	36.1	1.54	0.97	0.974	1.60	0.031	0.39	7.7
*arn*AA170T	7.5	**99**	80	66.2	**7.5**	0.93	**8.684**	**7.84**	0.065	**0.67**	3.3
*arn*DG206C	6.6	**99**	77.8	66	**6.6**	0.94	**6.995**	**6.79**	0.056	**0.66**	3
*lpt*AR62S	6.2	**98.1**	85.7	36.4	**3.26**	0.96	**3.273**	**3.43**	0.043	0.39	4.7
*pst*BE89Q	5.7	**99**	75	65.8	**5.7**	0.95	**5.382**	**5.76**	0.047	**0.66**	2.7
*fao*AT385A	4.1	**99.1**	88.9	36.1	**4.56**	0.97	2.848	**4.52**	0.032	0.38	3
*nuo*GA574T	3.8	94.3	26.7	64.2	0.67	1.02	0.542	0.65	-0.019	**0.62**	5
*nal*Dser32Asn	2.8	**100**	**100**	65.3	ND	0.972	**6.298**	ND	0.028	**0.66**	1
*nal*Dl153Q	2.8	**99.5**	75	65.2	**5.6**	0.98	2.653	**5.62**	0.023	**0.65**	1.3
*fus*A1D588G	2.8	**100**	**100**	65.3	ND	0.972	**6.298**	ND	0.028	**0.66**	1
*nuo*GA890T	2.6	**98.1**	71.4	35.5	1.37	0.99	0.148	1.38	0.007	0.36	2.3
*mex*RR79S	1.5	**96.2**	42.9	34.8	0.39	1.02	1.412	0.40	-0.023	0.35	2.3
*amg*SE108Q	1.5	**95.3**	37.5	34.6	0.32	1.034	2.505	0.32	-0.032	0.35	2.7
*gid*BE126G	1.5	**99.1**	75	35.5	**1.67**	0.99	0.2	**1.65**	0.006	0.36	1.3
*pst*BR87C	1.5	**100**	**100**	35.7	ND	0.985	2.632	ND	0.015	0.36	1
*mva*TR80A	1	**96.2**	33.3	34.9	0.26	1.029	2.442	0.27	-0.028	0.35	2
*amp*RA51T	1	95.3	28.6	34.7	0.21	1.04	**3.81**	0.21	-0.037	0.35	2.3
*rpl*YQ41L	1	**99.1**	66.7	35.4	1.11	0.99	0.005	1.09	0.001	0.36	1
*gid*BQ28K	1	**96.2**	33.3	34.7	0.26	1.03	2.48	0.27	-0.028	0.35	2

Numbers in bold show markers with best performance parameters. These were used as input to be tested through the clustering model in an iterative multi-step process. Multiple combinations were tested based on the results shown in this table.

Table 3.12. Summary of different measures of diagnostic accuracy for genes and mutations in relation to amikacin resistance

Molecular Marker	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Likelihood Ratio positive (LR+)**	**Likelihood Ratio negative (LR-)**	**Likelihood Ratio (LR)**	**Diagnostic odds ratio (DOR)**	**Youden index**	**Diagnostic Accuracy/effectiveness**	**Frequency of gene/mutation (%)**
*mex*R	**99.3**	0.7	21.4	80	1	1	0.006	1.091	0	0.219	99.3
*nal*C	90.5	3.1	20.3	54.8	0.934	3.065	**9.151**	0.309	0.064	0.219	95.5
*amp*R	89.9	2.8	20.1	50	0.925	3.607	**12.541**	0.252	0.073	0.214	95.7
*mex*Z	89.2	3.7	20.2	55.6	0.926	2.919	**10.122**	0.315	0.071	0.219	94.8
*nal*D	87.8	4.6	20.1	58.1	0.920	2.652	**9.779**	0.349	0.076	0.224	93.8
*arm*R	86.5	25	23.9	87.2	1.153	0.54	**9.658**	**2.139**	0.115	0.382	77.4
*nal*CS46A	66.2	45.1	24.7	83.1	1.206	0.749	**6.218**	1.611	0.113	0.496	57.3
*nal*CG71E	52.3	58.8	82.3	25.1	1.269	0.811	**5.741**	1.564	0.111	0.537	49.9
*mex*RG97L	50.5	67.6	85.1	27.1	1.559	0.732	**15.513**	**2.122**	0.181	0.541	46.6
*mex*RL29D	50.3	68.9	85.6	27.4	**1.617**	0.721	**17.674**	**2.242**	0.192	0.543	46.2
*mex*RE70R	48.4	69.6	85.4	26.9	1.592	0.741	**15.726**	**2.149**	0.18	0.529	44.6
*mex*RL130T	48.4	69.6	85.4	26.9	1.592	0.741	**15.726**	**2.149**	0.18	0.529	44.6
*mex*RR79N	47.1	68.2	84.5	26	1.481	0.776	**11.453**	1.917	0.153	0.517	43.8
*nal*CS209R	41.1	72.3	84.5	25.1	1.484	0.815	**9.101**	1.8187	0.134	0.478	38.2
*rpl*YAla123Ser	32.4	68.9	22.1	78.9	1.042	0.981	0.092	1.062	0.013	**0.611**	31.4
*lon*A499S	29.8	86.5	89	25.1	**2.207**	0.812	**17.727**	**2.721**	0.163	0.419	26.3
*amp*RG283E	27	68.4	19	77.4	0.854	1.067	1.168	0.801	0.046	0.595	30.6
*pmr*ALeu71Arg	25	74	20.8	78.4	0.962	1.014	0.057	0.950	-0.01	**0.635**	25.8
*nuo*GS468A	23.6	84.7	29.7	80.3	1.542	0.902	**5.397**	1.717	0.083	**0.716**	17.1
*amp*RM288R	22.3	77	21	78.3	0.969	1.009	0.037	0.958	0.007	**0.652**	22.9
*pmr*BLeu323His	21.6	85.4	28.8	80	1.479	0.918	**4.035**	1.617	0.07	**0.717**	16.1
*pmr*BSer420Arg	21.6	85.3	28.6	80	1.469	0.919	**3.841**	1.597	0.069	**0.716**	16.2
*pmr*BGly423Cys	21.6	85.5	28.8	80	1.489	0.917	**4.071**	1.620	0.071	**0.718**	16.1
*pmr*BL243Q	21.6	85.5	28.8	80	1.489	0.9169	**4.071**	1.620	0.071	**0.718**	16.1
*pmr*BA248V	21.6	85.3	28.6	80	1.469	0.919	**3.841**	1.59	0.069	**0.72**	16.2
*pmr*BALA4Thr	18.2	83.2	22.9	78.9	1.101	0.979	0.179	1.108	0.017	**0.69**	17.1
*pho*QY85F	18.2	92.3	39.1	80.5	**2.364**	0.886	**12.55**	**2.661**	0.105	**0.76**	10
*gid*BE186A	15.5	94.3	42.6	80.4	**2.719**	0.896	**13.402**	**3.038**	0.098	**0.77**	7.8
*amp*RE114A	15.3	93.9	90.2	23.2	**2.508**	0.902	**9.931**	**2.79**	0.092	0.32	13.3
*gid*BE97Q	14.9	93.9	40	80.2	**2.443**	0.906	**10.672**	**2.698**	0.088	**0.77**	8
*arn*AA170T	10.1	**98.3**	62.5	80.1	**5.941**	0.915	**19.692**	**6.691**	0.084	**0.79**	3.5
*pst*BE89Q	9.5	**98.9**	70	80	**8.636**	0.915	**22.454**	**9.35**	0.084	**0.79**	2.9
*arn*DG206C	9.5	**98.2**	58.3	79.9	**5.278**	0.922	**16.075**	**5.57**	0.077	**0.79**	3.5
*lpt*AT55A	8.7	**95.9**	88.7	22.3	**2.122**	0.952	**3.976**	**2.2426**	0.046	0.27	7.7

*nal*CA186T	8.1	**95.2**	31.6	79.1	1.688	0.965	2.24	1.751	0.033	**0.77**	5.5
*lpt*AR62S	4.8	**98.6**	**92.9**	22	**3.429**	0.966	**4.43**	**3.671**	0.034	0.25	4.1
*nal*CE153Q	4.1	**97.2**	28.6	78.8	1.464	0.987	0.615	1.487	0.013	**0.77**	3
*nal*Dser32Asn	4.1	**99.4**	66.7	79.2	**6.833**	0.965	**8.626**	**7.606**	0.035	**0.79**	1.3
*fus*A1D588G	3.4	**100**	100	79.2	ND	0.966	**15.544**	ND	0.034	**0.79**	0.7
*nuo*GA890T	3.3	**99.3**	**94.7**	21.9	**4.714**	0.974	**4.007**	**5.04**	0.026	0.24	2.7
*fao*AT385A	2.8	**99.3**	**93.8**	21.8	**4**	0.979	2.883	**4.176**	0.021	0.23	2.3
*mex*RR79S	2	**100**	**100**	21.8	ND	0.98	**5.351**	ND	0.02	0.23	1.6
*amg*SE108Q	2	**100**	**100**	21.8	ND	0.98	**5.351**	ND	0.02	0.23	1.6
*amp*RA51T	1.8	**100**	**100**	21.8	ND	0.982	**4.877**	ND	0.018	0.23	1.5
*mva*TR80A	1.7	**100**	**100**	21.8	ND	0.983	**4.385**	ND	0.017	0.23	1.3
*gid*BQ28K	1.7	**100**	**100**	21.7	ND	0.983	**4.371**	ND	0.017	0.23	1.3
*gid*BE126G	1.5	**100**	**100**	21.7	ND	0.985	**3.882**	ND	0.015	0.23	1.2
*nal*Dl153Q	1.4	**98.5**	20	78.6	0.933	1.001	0.012	0.916	0.001	**0.78**	1.4
*nuo*GA574T	1.4	93.7	5.6	77.7	0.222	1.052	**7.398**	0.205	0.049	**0.74**	5.2
*rpl*YQ41L	1.1	**100**	**100**	21.6	ND	0.989	2.907	ND	0.011	0.22	0.9
*pst*BR87C	0.7	**100**	**100**	21.5	ND	0.993	1.935	ND	0.007	0.22	0.6

Numbers in bold show markers with best performance parameters. These were used as input to be tested through the clustering model in an iterative multi-step process. Multiple combinations were tested based on the results shown in this table.

3.7.3. Finding the best possible predictor combinations

Applying cluster analysis using 6 of previously tested gentamycin molecular markers showing the best performance among the whole group of 50 tested markers as an input to the predictive model revealed 5 clusters (combinations of molecular markers) with improved diagnostic performance. Combination 1 showed a tendency to cluster towards higher gentamycin MICs while combination 5 showed a tendency to cluster towards lower gentamycin MICs. Predictor markers forming the new combinations included *gid*BE97Q, *gid*B E186A, *arn*AA170T, *arn*DG206C, *nal*CE153Q, and *pst*BE89Q. The new clusters showed very high likelihood ratio in differentiating sensitivity from resistance phenotype (LR=345.5, p<0.0005) with strong effect sizes (Phi=0.943, p<0.0005). A summary of all combination or markers in observed clusters are shown in Table 3.13.

Table 3.13. combination of molecular markers forming the new clusters in relation to gentamycin susceptibility and resistance

Cluster	**5**	**1**	**4**	**3**	**2**
Size	**183 (61%)**	**81 (27%)**	**16 (5.3%)**	**11 (3.7%)**	**9 (3%)**
Gentamycin sensitivity (breakpoint)	**100% susceptible**	**100% resistant**	**68.8% resistant**	**72.7% resistant**	**66.7% resistant**
Average gentamycin MIC	2.9	12.94	10.91	11.73	26.89
*gid*B E186A	Marker absent 69%	Marker absent 28%	Marker present 100%	Marker absent 5%	Marker absent 3%
*gid*BE97Q	Marker absent 63%	Marker absent 29%	Marker present 100%	Marker absent 5%	Marker absent 4%
*arn*AA170T	Marker absent 62%	Marker absent 28%	Marker absent 5%	Marker present 90.9%	Marker absent 3%
*arn*DG206C	Marker absent 62%	Marker absent 28%	Marker absent 5%	Marker present 81.8%	Marker absent 3%
*nal*CE153Q	Marker absent 65%	Marker absent 29%	Marker present 87.5%	Marker absent 3%	Marker present 30%
*pst*BE89Q	Marker absent 62%	Marker absent 29%	Marker absent 6%	Marker present 72.7%	Marker absent 4%

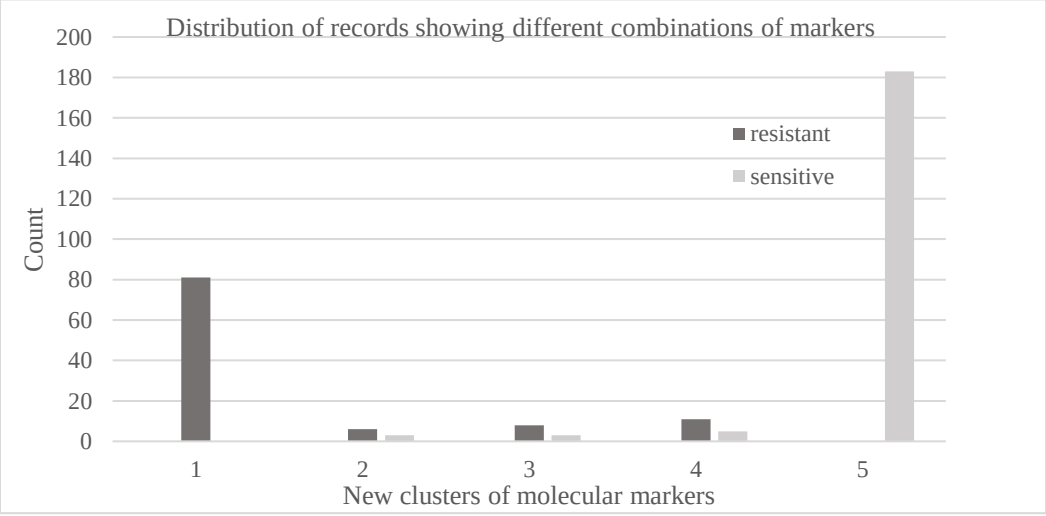


Figure 3.5. Clusters of molecular markers in relation to gentamycin susceptibility and resistance according to clinical breakpoints

Table 3.14. Distribution of observed gentamycin clusters in different categories of behavior among the studied isolates

MIC	.25	.50	1.00	2.00	4.00	8.00	16.00	32.00	64.00	Total
Cluster 1	0	0	0	0	0	35	44	2	0	81
Cluster 2	0	0	0	1	2	1	2	0	3	9
Cluster 3	0	0	1	0	2	1	7	0	0	11
Cluster 4	0	1	0	1	3	2	9	0	0	16
Cluster 5	5	6	24	45	103	0	0	0	0	183
Total	5	7	25	47	110	39	62	2	3	300

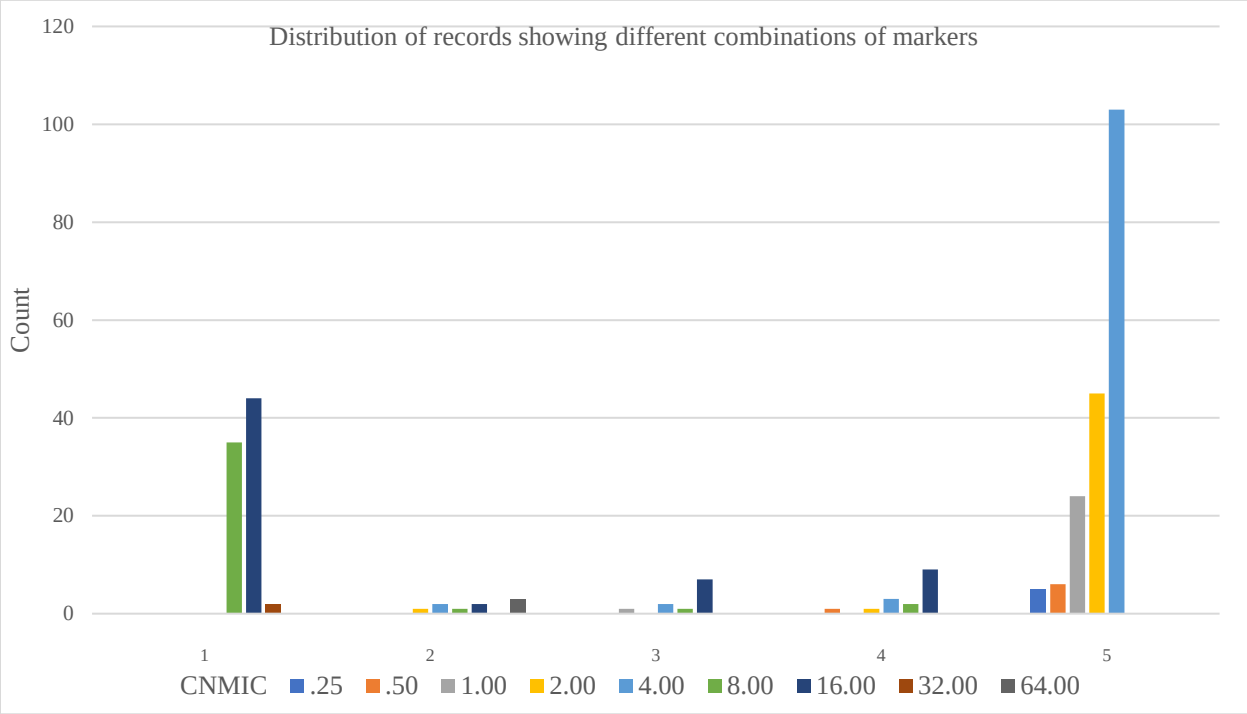


Figure 3.6. Distribution of different gentamycin levels of susceptibility within the new clusters of molecular markers

Applying cluster analysis using 8 of the previously tested amikacin molecular markers showing the best performance among the whole group of 51 tested markers as an input to the predictive model revealed 2 clusters (combination of molecular markers) with improved diagnostic performance. Combination 1 showed a tendency to cluster towards lower amikacin MICs while combination 2 showed tendency to cluster towards higher amikacin MICs. Predictor markers forming the new combination included *arn*AA170T, *arn*DG206C, *pst*BE89Q, *nal*DSer32Asn,

*lpt*AR62S, *nuo*GA890T, *fao*AT385A, and *lpt*AT55A. The new clusters showed very high likelihood ratio in differentiating sensitivity from resistance phenotype (LR=617.727, p<0.0005) with strong effect sizes (Phi=0.94, p<0.0005).

Table 3.15. combination of molecular markers forming the new clusters in relation to amikacin susceptibility and resistance

Cluster	**1**	**2**
Size	**528 (76.4%)**	**163 (23.6%)**
Amikacin sensitivity (breakpoint)	**100% susceptible**	**90.8% resistant**
Average amikacin MIC	3.69	54.47
*arn*AA170T	Mutation absent 100%	Mutation absent 85.3%
*arn*DG206C	Mutation absent 100%	Mutation absent 85.3%
*pst*BE89Q	Mutation absent 100%	Mutation absent 87.7%
*nal*DSer32Asn	Mutation absent 100%	Mutation absent 94.5%
*lpt*AR62S	Marker absent 95.1%	Marker absent 98.8%
*nuo*GA890T	Marker absent 96.6%	Marker absent 99.4%
*fao*AT385A	Marker absent 97.2%	Marker absent 99.4%
*lpt*AT55A	Marker absent 91.5%	Marker absent 95.1%

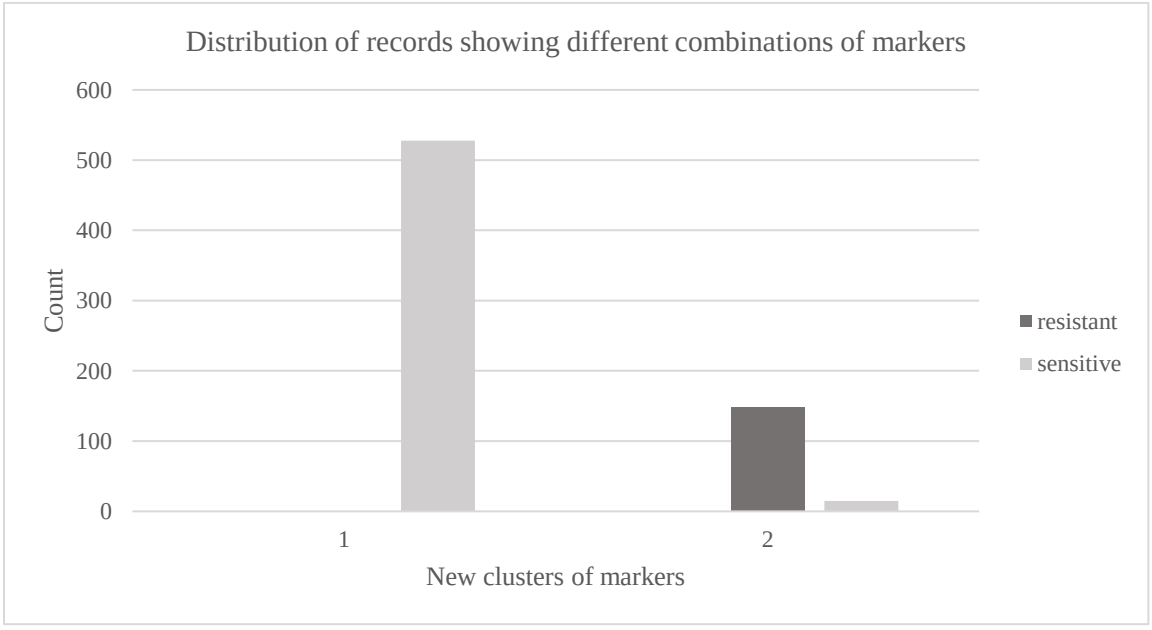


Figure 3.7. Clusters of molecular markers in relation to amikacin susceptibility and resistance according to clinical breakpoints

Table 3.16. Distribution of observed amikacin clusters in different categories of behavior among the studied isolates

MIC	.12	.25	.50	1.00	2.00	4.00	8.00	16.00	32.00	64.00	Total
Cluster 1	3	7	22	80	164	122	130	0	0	0	0
Cluster 2	0	0	0	1	5	5	4	35	40	37	36
Total	3	7	22	81	169	127	134	35	40	37	36

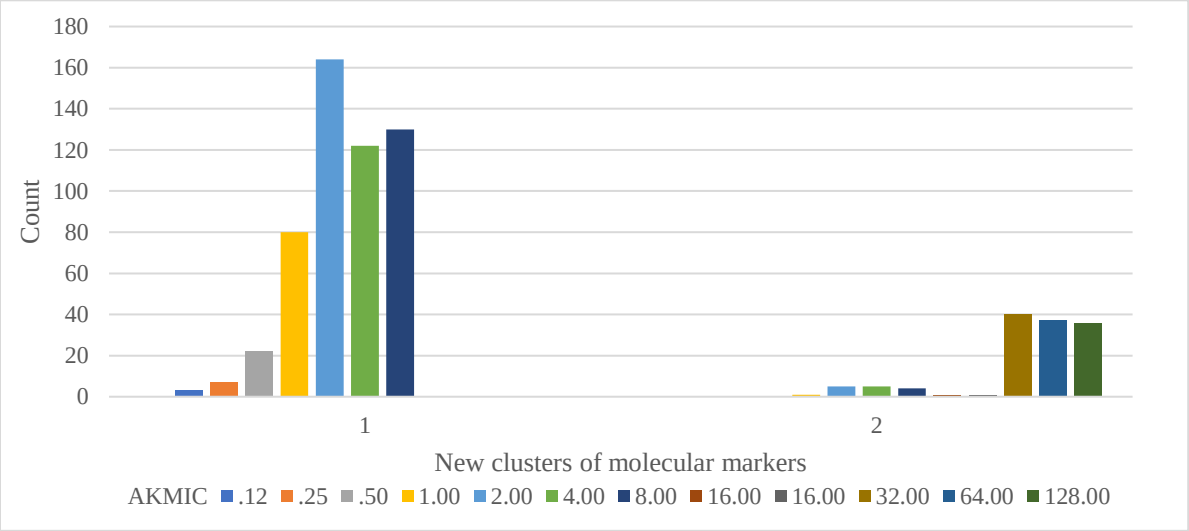


Figure 3.8. Distribution of different amikacin levels of susceptibility within the new clusters of molecular markers

3.7.4. Statistical and practical significance of individual molecular markers and new combinations

Table 3.17. Statistical significance and effect sizes for different studied gentamycin molecular marker

	chi square	**significance**	**Phi coefficient**	**significance**	**Cramer' s V**	**significance**
*nal*C	1.656	0.198	0.074	0.198	0.074	0.198
*mex*Z	2.215	0.137	0.086	0.137	0.086	0.137
*amp*R	1.656	0.198	0.074	0.198	0.074	0.198
*mex*RG97L	0.19	0.663	-0.025	0.663	0.025	0.663
*mex*RL29D	1.302	0.254	-0.066	0.254	0.066	0.254
*mex*RE70R	0.974	0.324	-0.057	0.324	0.057	0.324
*mex*RL130T	0.337	0.562	-0.034	0.562	0.034	0.562
*mex*RR79N	1.177	0.278	-0.063	0.278	0.063	0.278

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*nal*D	0.058	0.809	-0.014	0.809	0.014	0.809
*nal*CG71E	0.138	0.711	0.021	0.711	0.021	0.711
*nal*CS209R	<0.005	0.994	<0.005	0.994	<0.005	0.994
*arm*R	0.04	0.842	0.012	0.842	0.012	0.842
*amp*RG283E	0.981	0.322	0.057	0.322	0.057	0.322
*nal*CS46A	1.777	0.182	0.077	0.182	0.077	0.182
*amp*RM288R	0.173	0.678	0.024	0.678	0.024	0.678
*rpl*YAla123Ser	0.007	0.932	-0.005	0.932	0.005	0.932
*lon*A499S	5.098	0.024	-0.13	0.024	0.13	0.024
*pmr*ALeu71Arg	0.327	0.567	0.033	0.567	0.033	0.567
*amp*RE114A	0.088	0.767	-0.017	0.767	0.017	0.767
*pmr*BLeu323His	0.702	0.402	0.048	0.402	0.048	0.402
*pmr*BSer420Arg	0.702	0.402	0.048	0.402	0.048	0.402
*pmr*BGly423Cys	0.918	0.338	0.055	0.388	0.055	0.388
*pmr*BL243Q	0.918	0.338	0.055	0.388	0.055	0.388
*pmr*BA248V	0.702	0.402	0.048	0.402	0.048	0.402
*nuo*GS468A	0.702	0.402	0.048	0.402	0.048	0.402
*pmr*BALA4Thr	0.273	0.601	-0.03	0.601	0.03	0.601
*pho*QY85F	1.664	0.197	0.074	0.197	0.074	0.197
*nal*CE153Q	5.706	0.017	0.138	0.017	0.138	0.017
*nal*CA186T	1.259	0.262	0.065	0.262	0.065	0.262
*gid*BE97Q	8.26	0.004	0.166	0.004	0.166	0.004
*gid*BE186A	8.26	0.004	0.166	0.004	0.166	0.004
*lpt*AT55A	0.932	0.334	-0.056	0.334	0.056	0.334
*arn*AA170T	9.033	0.003	**0.174**	0.003	0.174	0.003
*arn*DG206C	7.316	0.007	**0.156**	0.007	0.156	0.007
*lpt*AR62S	2.847	0.092	-0.097	0.092	0.097	0.092
*pst*BE89Q	5.66	0.017	0.137	0.017	0.137	0.017
*fao*AT385A	2.383	0.123	-0.089	0.123	0.089	0.123
*nuo*GA574T	0.519	0.471	-0.042	0.471	0.042	0.471
*nal*Dser32Asn	5.546	0.019	**0.136**	0.019	0.136	0.019
*nal*Dl153Q	2.792	0.095	0.096	0.095	0.096	0.095
*fus*A1D588G	5.546	0.019	**0.136**	0.019	0.136	0.019
*nuo*GA890T	0.143	0.705	-0.022	0.705	0.022	0.705
*mex*RR79S	1.492	0.222	0.071	0.222	0.071	0.222
*amg*SE108Q	2.655	0.103	0.094	0.103	0.094	0.103
*gid*BE126G	0.189	0.663	-0.025	0.663	0.025	0.663
*pst*BR87C	1.656	0.198	-0.074	0.198	0.074	0.198
*mva*TR80A	2.584	0.108	0.093	0.108	0.093	0.108
*amp*RA51T	4.022	0.045	**0.116**	0.045	0.116	0.045

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*rpl*YQ41L	0.005	0.942	-0.004	0.942	0.004	0.942
*gid*BQ28K	2.631	0.105	0.094	0.105	0.094	0.105
New clusters	266.653	<0.0005	**0.943**	<0.0005	**0.943**	<0.0005

Table 3.18. Statistical significance and effect sizes for different studied amikacin molecular marker

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*mex*R	0.006	0.938	0.006	0.938	0.006	0.938
*nal*C	10.871	0.001	**-0.125**	0.001	0.125	0.001
*amp*R	15.222	<0.005	**-0.148**	<0.005	0.148	<0.005
*mex*Z	11.964	0.001	**-0.132**	0.001	0.132	0.001
*nal*D	11.385	0.001	**-0.128**	0.001	0.128	0.001
*arm*R	8.849	0.003	**0.113**	0.003	0.113	0.003
*nal*CS46A	6.109	0.013	0.094	0.013	0.094	0.013
*nal*CG71E	5.717	0.017	-0.091	0.017	0.091	0.017
*mex*RG97L	15.19	<0.005	**-0.148**	<0.005	0.148	<0.005
*mex*RL29D	17.242	<0.005	**-0.158**	<0.005	0.158	<0.005
*mex*RE70R	15.302	<0.005	**-0.149**	<0.005	0.149	<0.005
*mex*RL130T	15.302	<0.005	**-0.149**	<0.005	0.149	<0.005
*mex*RR79N	11.186	0.001	**-0.127**	0.001	0.127	0.001
*nal*CS209R	8.8	0.003	-0.113	0.003	0.113	0.003
*rpl*YAla123Ser	0.093	0.761	0.012	0.761	0.012	0.761
*lon*A499S	15.967	<0.005	-0.152	<0.005	0.152	<0.005
*amp*RG283E	1.148	0.284	-0.041	0.284	0.041	0.284
*pmr*ALeu71Arg	0.057	0.812	0.009	0.812	0.009	0.812
*nuo*GS468A	5.744	0.017	0.091	0.017	0.091	0.017

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*amp*RM288R	0.037	0.847	-0.007	0.847	0.007	0.847
*pmr*BLeu323His	4.276	0.039	0.079	0.039	0.079	0.039
*pmr*BSer420Arg	4.064	0.044	0.077	0.044	0.077	0.044
*pmr*BGly423Cys	4.315	0.038	0.079	0.038	0.079	0.038
*pmr*BL243Q	4.315	0.038	0.079	0.038	0.079	0.038
*pmr*BA248V	4.064	0.044	0.077	0.044	0.077	0.044
*pmr*BALA4Thr	0.181	0.671	0.016	0.671	0.016	0.671
*pho*QY85F	14.288	<0.005	**0.144**	<0.005	0.144	<0.005
*gid*BE186A	15.604	<0.005	**0.15**	<0.005	0.15	<0.005
*amp*RE114A	8.537	0.003	-0.111	0.003	0.111	0.003
*gid*BE97Q	12.259	<0.005	0.133	<0.005	0.133	<0.005
*arn*AA170T	24.932	<0.005	**0.19**	<0.005	0.19	<0.005
*pst*BE89Q	28.882	<0.005	**0.204**	<0.005	0.204	<0.005
*arn*DG206C	20.131	<0.005	**0.171**	<0.005	0.171	<0.005
*lpt*AT55A	3.477	0.062	-0.071	0.062	0.071	0.062
*nal*CA186T	2.449	0.118	0.06	0.118	0.06	0.118
*lpt*AR62S	3.533	0.06	-0.072	0.06	0.072	0.06
*nal*CE153Q	0.658	0.417	0.031	0.417	0.031	0.417
*nal*Dser32Asn	11.093	0.001	0.127	0.001	0.127	0.001
*fus*A1D588G	18.478	<0.005	**0.164**	<0.005	0.164	<0.005
*nuo*GA890T	3.03	0.082	-0.066	0.082	0.066	0.082
*fao*AT385A	2.239	0.135	-0.057	0.135	0.057	0.135
*mex*RR79S	3.047	0.081	-0.066	0.081	0.006	0.081
*amg*SE108Q	3.047	0.081	-0.066	0.081	0.066	0.081
*amp*RA51T	2.776	0.096	-0.063	0.096	0.063	0.096
*mva*TR80A	2.495	0.114	-0.06	0.114	0.06	0.114
*gid*BQ28K	2.485	0.115	-0.06	0.115	0.06	0.115

	chi square	significance	Phi coefficient	significance	Cramer' s V	significance
*gid*BE126G	2.206	0.137	-0.057	0.137	0.057	0.137
*nal*Dl153Q	0.012	0.912	0.004	0.912	0.004	0.912
*nuo*GA574T	5.678	0.017	-0.091	0.017	0.091	0.017
*rpl*YQ41L	1.65	0.199	-0.049	0.199	0.049	0.199
*pst*BR87C	1.097	0.295	-0.04	0.295	0.04	0.295
New clusters	610.079	<0.0005	**-0.94**	0.0005	**0.94**	0.0005

3.7.5. Multiple Regression Analysis for best molecular predictors

Stepwise multiple regression analysis was used to assess the ability of variants showing higher correlations with phenotype to predict the level of gentamycin and amikacin MIC. For Amikacin, after excluding variants with high multicollinearity and variants with smaller correlations with phenotype, the most important predictors in the model included *pst*BE89Q, *mex*RL29D, *nal*Dser32Asn, *amp*R, *pmr*BGly423Cys, *fus*A1D588G, and *lon*A499S. In the proposed model, only 15% of variance in the dependent variable (Amikacin MIC) is explained by the above 7 predictors in the model (p=0.016). When the statistical effect of overlapping variables was excluded, variables making significant unique contribution to the prediction of the dependent variable (Amikacin MIC) include in order of importance: *pst*BE89Q (*beta*= 0.245, p<0.0005), *mex*RL29D (*beta*= 0.177, p<0.0005), *nal*Dser32Asn (*beta*= 0.104, p<0.004), *amp*R (*beta*= 0.107, p<0.003), *pmr*BGly423Cys (*beta*= 0.110, p<0.002), *fus*A1D588G (*beta*= 0.094, p<0.009), and *lon*A499S (*beta*= 0.088, p<0.016).

When other variants were re-included as predictors for a better model, only 17.9% of variance in Amikacin MIC (p<0.0005) can be explained by 32 markers including *arn*DG206C, *amp*R, *rpl*YQ41L, *gid*BQ28K, *nal*Dser32Asn, *nuo*GA890T, *gid*BE186A, *fus*A1D588G, *mex*RG97L, *pmr*BA248V, *pmr*ALeu71Arg, *lon*A499S, *Amp*RE114A, *pmr*BALA4Thr, *nal*CS209R, *nuo*GS468A, *rpl*YAla123Ser, *nuo*GA574T, *pst*BE89Q, *mex*Z, *nal*CG71E, *gid*BE126G, *nal*C

gene, *pho*QY85F, *amg*SE108Q, *mex*RE70R, *arn*AA170T, *gid*BE97Q, *mex*RL130T, *mex*RL29D, *mex*RR79N, and *pmr*BL243Q.

For gentamycin, after excluding variants with high multicollinearity and variants with smaller correlations with the phenotypes, the most important predictors in the model included *fus*A1D588G, *lon*A499S, *pmr*ALeu71Arg, *gid*BQ28K, and *gid*BE97Q. In the proposed model, only 16.2% of variance in the dependent variable (Gentamycin MIC) is explained by the above 5 predictors in the model (p=0.02). When the statistical effect of overlapping variables was excluded, variables making significant unique contributions to the prediction of the dependent variable (Gentamycin MIC) include in order of importance: *fus*A1D588G (*beta*= 0.301, p<0.0005), *lon*A499S (*beta*= 0.182, p<0.001), *pmr*ALeu71Arg (*beta*= 0.160, p<0.004), *gid*BQ28K (*beta*= 0.134, p<0.015), and *gid*BE97Q (*beta*= 0.127, p<0.020).

When other variants were re-included as predictors for a better model, only 21.6% of variance in gentamycin MIC (p<0.0005) can be explained by 31 markers including *amg*SE108Q, *pmr*BAla4Thr, *fus*A1D588G, *nal*C gene, *mex*RL130T, *amp*R, *nuo*GA890T, *pmr*ALeu71Arg, *gid*BE186A, *pst*BE89Q, *rpl*YQ41L, *mex*Z, *nal*CG71E, *nal*Dser32Asn, *Amp*RE114A, *mex*RE70R, *lon*A499S, *pmr*BL243Q, *mex*RG97L, *nuo*GS468A, *nal*CS209R, *rpl*YAla123Ser, *nuo*GA574T, *pho*QY85F, *arn*AA170T, *mex*RL29D, *gid*BE126G, *gid*BQ28K, *mex*RR79N, *arn*DG206C, and *pmr*BA248V.

3.8. Discussion

Accurate diagnosis is an essential step for the successful management of any health problem. Diagnostic tests can be used at different time points through the management pathway. Among the essential purposes of using a diagnostic test is to select for the most appropriate next step in the decision-making process. Although higher test accuracy is often used as an important indicator for the usefulness of the test, this does not necessary indicate that tests with higher accuracy often lead to improved health outcomes and are consequently the tests of choice in clinical practice (Mustafa *et al.*, 2017).

An essential requirement for the evaluation of any medical test is to know whether the test provides more information than what is already available to doctors (Moons, Biesheuvel and Grobbee, 2004). Diagnostic accuracy is often used as an evaluation indicator (Knottnerus, van

Weel and Muris, 2002), however, a series of questions usually needs to be considered when a new test is being introduced (Bossuyt *et al.*, 2006). Knowledge of other features of the new test such as availability, invasiveness, simplicity, and cost-effectiveness can help define how the test is likely to be used, what the role of the test is and where it can be placed in the diagnostic algorithm.

On introducing a new diagnostic strategy, both accuracy of the new diagnostic test and other test features should be compared with the existing diagnostic pathway. This helps in identifying the potential contribution or, the added value of a new diagnostic strategy compared to an existing one. It may not be easy in all situations to define what the existing pathway of diagnosis is. Although this is usually provided in practice guidelines, a consensus on the optimal sequence or types of diagnostic procedures in different settings may not be easily found, and this adds to the complexity of having a comparator gold standard.

For the diagnosis of antibiotic resistance, the common standard of practice in most settings is conventional culture and sensitivity testing (Belkum and Dunne, 2013). Objective standardized methods for classifying resistance or sensitivity depend on the clinical breakpoints used to guide treatment choices in clinical practice. However, these breakpoints are continuously updated with new guidelines coming into practice. For the purpose of assessing markers in this study (Chapter), breakpoints from EUCAST (Rules, 2018) were used as the cut-off to classify the studied isolates into susceptible and resistant groups. Conventional culture has an inherent drawback of longer specimen to answer time. In addition, other limitations of conventional culture include the need for relatively large numbers of viable organisms, several steps of pre-analytic processing, and limited organism spectrum. These together with some other practical drawbacks makes the use of Sequencing-based diagnosis and treatment a better future alternative in some settings.

There has been a growing interest in developing new diagnostics for bacterial resistance with the recent advances in diagnostics technologies. Many newer technologies have been used including MALDI-TOF MS, fluorescent live/dead staining, infrared spectroscopy, microbial cell weighing by vibrating cantilevers, magnetic bead spin, and microdroplets among others (Belkum and Dunne, 2013). Sequencing-based resistance prediction is now expected to offer a better practical

diagnostic alternative for antimicrobial resistance. Mining for the best optimized set of molecular markers among what has been previously reported in the literature is the aim of this chapter.

Different factors need to be considered for the selection and application of any diagnostic test. The most important is to define the test purpose. It is important to decide whether the test is needed for screening or for confirmatory diagnosis. It is important also to consider whether it is used to rule in or to rule out the diagnosis and hence to judge subsequent treatment recommendations. It is also essential to define whether a single test, a combination of tests or, a diagnostic algorithm is needed to make the final decision (Kosack, Page and Klatser, 2017). Acceptable standards in positive and negative predictive values that can translate into changes in patient management are still lacking (Caliendo *et al.*, 2013). This makes it important to consider all parameters used to assess the accuracy of the diagnostic test in a setting-relevant basis and according to the condition being investigated.

The decisive aim of any diagnostic test is to accurately discriminate subjects with a specific disorder from subjects without it. For the evaluation of a diagnostic test, researchers need to choose the most informative indices of performance. In this study, sources of bias originating from possible deviation in disease spectrum was decreased by including the whole spectrum of the condition under evaluation (resistance and susceptibility to antibiotics). Another important point to consider when interpreting predictive values is the sample size. This is represented in the current analysis by the frequency of the molecular marker to be used as a diagnostic test. The more frequent the marker is found in a randomly selected population, the more there is confidence about the predictive values obtained because this reflects evaluation of a larger sample size.

In order to make the best selection of the diagnostic molecular markers, the purpose of the diagnostic test needs to be considered. Considering how much discriminative power is needed depends on the aim of the diagnostic test, its use, and its application. For predicting antibiotic resistance, the aim of diagnostic testing using molecular markers can be either to rule-in or to rule-out the diagnosis of resistance or the diagnosis of susceptibility.

In case of antibiotic resistance diagnosis, the cost of false positive is high because the ultimate goal is to avoid over prescription resulting from incorrect overdiagnosis of resistance. This means that high specificity of the diagnostic test is an essential priority. In addition, LR+ is

considered the best indicator for ruling- in the diagnosis of resistance. Higher values of positive results on this test (LR+) has significant contribution to diagnosis.

In addition to high specificity and LR+, high PPV is considered another additional indicator to rule-in diagnosis and to avoid the cost of false positives. For confidently considering PPV, the marker under evaluation should occur in high frequency and the population under evaluation should be randomly selected and should represent an adequate number of resistant isolates.

The other type of information that can be provided by considering a combination of parameters is ruling-out the diagnosis of resistance. When sensitivity of a particular marker is high, specificity and LR+ are poor, the presence of this diagnostic marker cannot confirm the diagnosis of resistance, but its absence practically rules it out. Having high NPV with high marker prevalence in the population is an additional way for ruling out resistance.

The group of genes and mutations previously identified in the literature as associated with resistance to quinolone and aminoglycoside group of antibiotics were reviewed and assessed for their predictive potential in this chapter.

Mutations in QRDR were not frequently encountered in the population studied. Although most show high specificity, the low frequency of these mutations reduce their practical use as molecular diagnostic markers. A more informative group of markers were those related to efflux pump regulatory genes. These genes and gene variants showed better performance in addition to higher frequency of occurrence in the population.

3.8.1. Potential markers to be applied using the rule-in algorithm

High specificity is usually considered an important parameter to rule in the diagnosis. A group of markers showing high specificity and high LR+ to the phenotype under assessment (antibiotic resistance or susceptibility) can be used as molecular markers to rule in the diagnosis under consideration (Fischer, Bachmann and Jaeschke, 2003). The diagnostic value of this type of markers is higher when the marker combines high specificity and high LR+ with high DOR, high accuracy and higher PPV towards the same phenotype.

A group of variants have met these criteria and hence, may be used to rule in quinolone resistance. These variants include: *gyr*A T83I, *nfx*B A124T, *nal*C E153Q, *nal*C Thr50Pro, *par*E

V460G, *par*C E91K, and *amp*R D135N. The corresponding measures of performance of these variants for both ciprofloxacin and levofloxacin with their corresponding statistical values are shown in Table 3.3, Table 3.6, Table 3.7, and Table 3.10. These findings can be supported by findings from other recent research observations that showed *gyr*A and *par*C QRDR mutations among the determinants of levofloxacin resistance in *Ps. aeruginosa* using four comparative methods, however, these research findings have not shown the relative importance of each identified mechanism (Jaillard *et al.*, 2017) (Jaillard *et al.*, 2018), the point that was further studied using the current analysis. Earle *et al.*, (2016) have also shown *gyr*A mutations in QRDR AA67-106, *par*C mutations in QRDR AA47-133, and *par*E mutation in QRDR AA420-458 conferring resistance to *Escherichia coli* and *Klebsiella pneumonia* which support findings from the current analysis identifying mutations within the same region showing high predictive performance.

Analysis of molecular markers associated with aminoglycoside resistance showed that variants in genes related to cell membrane permeability changes were the most important in accounting for resistance with the highest predictive values and highest performance as shown using different other measures of performance. These markers can similarly be used to rule in aminoglycoside resistance. These markers include *arn*AA170T, *arn*DG206C, *fus*A1D588G, and *nal*Dser32asn. The predictive values and other measures of performance associated with this group of markers for both gentamycin and amikacin are shown in Table 3.11, Table 3.12, Table 3.17, and Table 3.18. The Arn (PA3552-PA3559) LPS modification genes have also been linked to aminoglycosides resistance because the expression of the *arn*BCADTEF operon is recognized as a contributing factor which decreases the interaction and uptake of polycationic antibiotics (Schniederjans, Koska and Häussler, 2017). Results of the current analysis support that because the two variants identified i.e. *arn*AA170T, *arn*DG206C show higher predictive values to aminoglycoside resistance and were also predicted as deleterious using PROVEAN.

An important finding is the observation of a group of molecular markers showing high specificity and high positive predictive values to amikacin susceptibility and these can be considered as molecular susceptibility markers to rule in the diagnosis of amikacin susceptibility. These include *fao*AT385A, nuoGA890T, *lpt*AT55A, and *lpt*AR62S. Screening transposon insertion mutant library has previously shown that disruption of *fao*AB and *lpt*A are associated

with increased aminoglycosides susceptibility (Krahn *et al.*, 2012) and these findings support findings from the current analysis. *fao*AB encodes a multienzyme complex which is involved in degradative fatty acid [FA]-oxidation, and *lpt*A encodes a lysophosphatidic acid acyltransferase (LPA), responsible for adding the second FA to glycerol-3 phosphate in the synthesis of phospholipids (PLs) (Krahn *et al.*, 2012). Both *fao*AT385A and *lpt*AR62S were predicted as deleterious using PROVEAN, so it is proposed that these variants are good candidates to be used as aminoglycoside susceptibility markers. This is based on the frequent observation of these variants in aminoglycoside susceptible isolates, their higher predictive values towards susceptibility and their possible associated functional effect on aminoglycoside binding and uptake across the outer LPS membrane of *Ps. aeruginosa.*

*nuo*G operon codes for proton-translocating type I NADH oxidoreductase which is an enzymatic complex that significantly contributes to the proton electrochemical gradient. Inactivation of NADH dehydrogenase has shown to impair membrane energetics and thereby the uptake of aminoglycosides (El'Garch *et al.*, 2007). El'Garch *et al.*, (2007) has previously shown that combined simultaneous mutations in *gal*U, *nuo*G, *mex*Z, and *rpl*Y can increase survival rates in *Ps. aeruginosa* treated with tobramycin up to 16-fold while single gene mutation has a much lower effect. Pelegrin *et al.*, (2019) have recently identified a panel of *nuo*G variants including *nuo*G S468A in association with aminoglycoside resistance. A recent study investigating resistance in experimentally evolved *Ps. aeruginosa* has identified a total of 24 mutated genes in association with aminoglycoside resistance with ten mutants in genes directly involved in oxidative phosphorylation and proton motive force including *nuo*G mutants (Wardell *et al.*, 2019). Some of the identified mutants have also showed clinical relevance when re-tested. The same study has also identified the relation of *fus*A1, *mex*R, *nal*D, and *amg*S mutants to tobramycin resistance. Similar findings from transcriptional profile analysis identifying the *nuo*G among other genes encoding NADH dehydrogenases also support that. The genes were downregulated in adaptation to *Ps. aeruginosa* chronically infected lung (Kordes *et al.*, 2019). However, it is also important to consider the phenomenon of collateral sensitivity where evolved resistance to one antibiotic concomitantly cause hypersensitivity to another drug. This is sometimes explained by a fitness cost where mutations that promote resistance affect the ability of bacteria to replicate and survive in normal conditions (Barbosa *et al.*, 2019). Barbosa *et al.*, (2019) show that *Ps. aeruginosa* is one of the bacteria known to evolve collateral sensitivity to

certain drug treatments. This phenomenon may explain the observation of some *nuo*G variants identified from the current analysis in relation to amikacin susceptibility including *nuo*G S468A and *nuo*G A890T which were previously linked to resistance to other aminoglycoside agents. This may be attributable to the effect of other secondary mutations or epistatic interactions that may reduce overall fitness. Although mutations affecting the activity of *nuo*G have been associated with gradual increase in MIC, this occurred when co-operatively combined with other chromosomal mutations in other genes (Islam *et al.*, 2009) and this may explain the reason for the different observations reported in the current study. Another explanation for the different findings observed with *nuo*G mutants is that disruption of the gene may affect survival of the organism being essential for a functional respiratory complex I. Exploiting such phenomenon in mutants showing collateral sensitivity may prove useful in the setting of developing a diagnostic because it carries the advantage of directing the choice of combination therapy that could slow down the evolution of resistance.

It is interesting here to find that three of the novel aminoglycoside resistance markers identified showed high predictive performance using the multiple regression model and also showed higher diagnostic predictive values. These include *gid*BE97Q, *fus*A1D588G, and *pst*BE89Q. These findings are supported by the findings of Jaillard *et al.*, (2018) who identified *gid*B as one of genes associated with resistance in *M. tuberculosis* using four comparative methods. *gid*B has been identified earlier among the genes associated with streptomycin resistance in *M. tuberculosis* strains exhibiting no *rps*L or *rrs* mutations. Although several reports of both streptomycin resistant and susceptible isolates have shown association with *gid*B mutations, *gid*B appears to have complex substrate activity. Several mutants affecting the active site pocket or altering secondary-structure motifs and thus showing an overall effect on protein structure have been previously reported in relation to streptomycin resistance in *M. tuberculosis* (Verma *et al.*, 2014). *gidB* deletion has also shown to confer high-level aminoglycoside resistance including streptomycin and neomycin in *Salmonella* (Mikheil *et al.*, 2012).

*gid*B (glucose-inhibited division gene) is found among the gene cassettes harboring the OriC regions in some bacteria including *Ps. aeruginosa* (Rybenkov, 2014). *gid*B is known to be involved in posttranslational modification, methylation, of 16S RNA. *gid*B mutant has also shown a compromised overall bacterial fitness in *Salmonella* (Mikheil *et al.*, 2012). This may

reflect the physiologic cost of methylation deficiency. The effect of *gid*B mutations on antimicrobial susceptibility is thought to occur through mechanisms involving post-transcriptional modification which explains its relation to aminoglycoside resistance. *gid*B is considered highly conserved in both Gram-positive and Gram-negative bacteria with *gid*B protein known among the proteins involved in cell cycle control of DNA replication. Consequently, its disruption may lead to inhibition of cell division which may also explain the association observed with amikacin susceptibility from the current analysis. Three deleterious amino acid substitution in *gid*B including E126G, E97Q, and Q28K showed association with amikacin susceptibility phenotype which was not observed with gentamycin. Previously studied *gid*B mutants have shown lack of 16S rRNA methylation activity (Mikheil *et al.*, 2012) which explains the role of this mutant in conferring resistance to some of previously studied aminoglycoside agents. However, this may not prove true for amikacin. The fact that not all aminoglycosides bind to identical sites of the 16SrRNA (Ramirez and Tolmasky, 2017) may explain the different observed effects of *gid*B mutants on susceptibility to different agents from the aminoglycoside group. While amikacin binds the A site of the 16S RNA similar to other agents from the aminoglycoside group, specific interactions have been observed between its modified side group and RNA at the GC pairs C_{1404}–G_{1497} and G_{1405}–G_{1496} (Kondo *et al.*, 2006). This interaction may be different from other interaction sites that showed to be affected by *gid*B induced N7 methylation of G527 of the 16S rRNA leading to aminoglycoside resistance in *Salmonella* (Mikheil *et al.*, 2012) or low-level streptomycin resistance in *M. tuberculosis* (Verma *et al.*, 2014).

Mutations in the gene that encodes for elongation factor G, *fus*A, has been previously linked to fusidic acid resistance in *Salmonella Typhimurium*. This is thought to occur through alteration of levels of the transcriptional regulator guanosine tetraphosphate (ppGpp). Some mutations are thought to have pleiotropic effects on gene expression which may lead to fitness differences in different environments (Andersson and Hughes, 2010). The pleiotropic phenotypes of the *fus*A mutant also showed to include hypersensitivity to other unrelated antibiotics (Macvanin and Hughes, 2005). Earle *et al.*, (2016) have also reported variety of *fus*A mutations in association with *Staphylococcus aureus* resistance to fusidic acid. Similarly, another *fus*A1 variant *fus*A P413L has been identified as a factor reducing the exponential growth rate by 67% in *S. Typhimurium* and in *E. coli* (Knopp and Andersson, 2018). The study shows that the effect of

*fus*A mutants on exponential growth rate was consistent in all genetic backgrounds. Additionally, the variant appears to have remarkable effects on resistance levels. *fus*A class mutants generally appear to reduce fitness both *in vivo* and *in vitro* (Nagaev *et al.*, 2001). The role of *fus*A gene as a determinant of antimicrobial resistance has also been recently demonstrated. It has been studied as a target for argyrin B which act as a natural product with antibacterial activity. It has been shown that *fus*A sequence variation can affect the antibacterial activity of argyrin B against *Ps. aeruginosa* (Jones *et al.*, 2017). These findings support the high predictive values identified for variants in the same genes by applying the current analysis.

Similarly, other *fus*A1 variants including *fus*A1N178S, T671A, and I186V were identified by Pelegrin *et al.*, (2019) in association with different levels of amikacin susceptibility. However, these variants were rare when compared with other aminoglycosides mutations identified by Pelegrin *et al.*, (2019) including *fus*A2 and *nuo*G variants. In the current analysis, the same variant *fus*A T671A was identified in one isolate which was amikacin susceptible but gentamycin resistant. *fus*A1 Thr671Ala is located at domain V at the start of an alpha helix and is probably affecting the gene function. This finding supports findings by Bolard, Plésiat and Jeannot, (2018) who showed that replacement of wild-type *fus*A1 in PAO1 by the mutated allele of the gene has decreased amikacin MIC from 32 to 16 but has not affected gentamycin MIC. Although the variant *fus*A1 Thr671Ala has been previously identified in relation to the evolution of experimental resistance against tobramycin and gentamycin, experimental evolution of the gene occurred together with other variants including *fus*A1 I61M and *fus*A1 E100G (López-Causapé *et al.*, 2018). These other related variants have not been identified in the current analysis in relation to *fus*A1 Thr671Ala which may explain the reason for these different findings. This indicates the importance of considering the effect of individual mutations in addition to the effect of the dynamics resulting from the interaction of multiple mutations on the phenotypic behavior. In support of that is another study investigating the evolutionary trajectories of *Ps. aeruginosa* on experimental exposure to gentamycin (Sanz-García, Hernando-Amado and Martínez, 2018a). The study has identified Thr671Ala, *fus*A1 Ala595Pro, Gly545Asp, and Arg680Cys among the mutations observed in the evolved PA14 populations. Two of these mutations have been encountered in the studied set of clinical isolates.

Another variant in the same gene, *fus*A1Y552C, has also been previously reported in the literature (Bolard, Plésiat and Jeannot, 2018) (López-Causapé *et al.*, 2018) and identified in the current analysis in association with aminoglycoside resistance. These observations point to the importance of *fus*A1 gene which produce elongation factor G (EF-G1A) that is considered a key component of the translational machinery that modify/differentiate aminoglycoside susceptibility. In support of that are the findings of *fus*A mutations which are induced by the *in vitro* exposure of *Ps. aeruginosa* to increasing concentrations of tobramycin (Feng, Jonker, Moustakas, Brul and ter Kuile, 2016).

Phosphate has an important role in cellular physiology. Microbes have complex regulatory networks for sensing phosphate availability and utilization. It has been previously noted that genes involved in phosphate regulon becomes upregulated in relation to *Ps. aeruginosa* host pathogenicity and other virulence properties. *Pst*B is a phosphate uptake regulatory protein which showed to be upregulated among many other genes inducing cellular cytotoxicity under adverse conditions through phosphate acquisition (Bains, Fernández and Hancock, 2012). These transcriptional changes showing association with cellular cytotoxicity may explain the role of some observed *pst*B variants from the current analysis, e.g. *pst*B E89Q, in antibiotic resistance. Giving additional support, other mutations in *pts*B has been previously linked to low-level tobramycin resistance (Sanz-García, Hernando-Amado and Martínez, 2018b).

On the other hand, inactivation of *pst*B has previously been linked to aminoglycoside susceptibility especially when combined with inactivation of other genes including those associated with lipid biosynthesis or metabolism (*lpt*A, *fao*A) or other two component regulators (*amg*RS) (Krahn *et al.*, 2012). *pst*B R87C variant has been identified from the results of the current analysis in relation to aminoglycoside susceptibility. The variant is predicted as deleterious (PROVEAN score= - 6.874), so it is probably causing inactivation of *pst*B and these findings support the findings from the literature as shown above.

An important point to consider when studying the effects of different mutations in general is that the dynamics of mutations interactions should be overall considered. It has been shown that buildup of resistance cannot be attributable only to DNA mutations but may develop as a result of interactions between mutations and cellular adaptation (Feng, Jonker, Moustakas, Brul and ter Kuile, 2016). In addition, the fitness cost of an observed mutation which may lead to cross-

resistance or collateral sensitivity also need to be considered (Barbosa *et al.*, 2017) as this phenomenon has been previously reported in *Ps. aeruginosa.*

3.8.2. Potential markers to be applied using the rule-out algorithm

Although high specificity is considered of primary importance in ruling in the diagnosis of resistance, markers showing high sensitivity and high NPV towards the phenotype of interest can be applied using the rule-out algorithm. The presence of these markers cannot confirm the phenotype or behavior of interest, but its absence practically rules out the same conclusion. This can show higher predictive performance when high sensitivity and high NPV are combined with low specificity and low LR+ in addition to high LR, high DOR and high accuracy.

The following group of markers can be used to rule-out susceptibility which means that presence of these markers/genes does not guarantee or confirm susceptibility but their absence rule-out susceptibility and predict resistance with higher confidence. These markers included genes and mutations associated with efflux pump regulation including *amp*R, *nfx*B, *mex*S, *nal*C and *nal*D, *mex*R, *arm*R, and *mex*Z. Values of performance for these genes are shown for different studied antibiotic groups in Table 3.3, Table 3.7, Table 3.11, and Table 3.12. In addition, other *mex*R and *nal*C variants showed similar predictive capabilities. These markers include the following as quinolone resistance predictors; *nal*C S46A, *nal*C G71E, *nal*C S209R, *mex*R R79N, *mex*R E70R, *mex*R L130T, *mex*R G97L, and *mex*R L29D.

The results of the analyses performed in this chapter show that efflux-pump regulation appears to be the greatest contributor of resistance to quinolone group of antibiotics as evidenced by the regression model and the best performing predictive values. This role also applies to determining aminoglycoside resistance as evidenced by good predictive values for *mex*Z, *arm*R, *nal*D, and *nal*C and some of their related variants. This would be supported by some other recent findings that have identified the *mex*Z gene among the essential genes accounting for resistance in *Ps. aeruginosa* using PLINK association analysis and support vector machine model (Noah, 2019). In addition, Pelegrin *et al.*, (2019) have identified *nal*C G71E, *nal*C S209R, and multiple other *arm*R and *nal*D variants among those contributing to ciprofloxacin, amikacin, and carbapenem resistant *Ps. aeruginosa* isolates. An important finding here is the high predictive value for *nfxB* regulator absence and of its related variant *nfx*B A124T which have been identified from the current analysis. The importance of this finding comes from the fact that mutation in the *nfx*B

gene is considered particularly important because it is the only known basis for over-expression of MexCD-OprJ operon which is normally silent in wild-type cells (Li, Plésiat and Nikaido, 2015). Higher performance of *nfx*B and its variant *nfx*B A124T support reports from the literature which shows that MIC of different antibiotic classes can increase up to 500-fold in *nfx*B mutants including the MIC of ciprofloxacin which showed 94-fold increase (Chuanchuen et al., 2001). It was also shown that both *in vitro*-selected and clinical isolates expressing this efflux system and exhibiting high levels of quinolone resistance invariably contain mutations in *nfx*B (Purssell and Poole, 2013).

*mex*T, *mex*S, *amp*R, and *mva*T related-variants showed lower performance when considered separately but showed more contribution when combined with other efflux-pump or target-enzymes mutations. This supports findings from the literature showing variants identified in these genes and contributing to modifying the expression of MexEF-OprN probably through two independent pathways (Westfall et al., 2006). Mutations in genes that encode for the global regulators MvaT and AmpR have also been reported in other studies as activators for MexEF-OprN operon in the *in vitro* mutants. The MexEF-OprN system is known to be quiescent in wild-type cells. To produce the pump MexEF-OprN, an active MexT transcriptional regulator is required. MexS is an upper negative regulator that inactivates MexT and prevents pump expression in wild type strains. Activation and expression of the pump is known to occur through two independent pathways (Uwate et al., 2013). *nfx*C-type mutants of *Ps. aeruginosa* that produce increased amounts of MexEF-OprN and that display multidrug resistance to multiple antibiotic agents including quinolones have been repeatedly isolated *in vitro* and from the clinic (Fukuda et al., 1995) (Jalal et al., 2000) (Richardot et al., 2016). An important point to consider here is that different *mex*T, *mex*S, *mva*T, and *amp*R alterations may be encountered in different functional backgrounds. This has been observed through the analysis applied in this chapter by finding *mex*S substitutions with no associated *mex*T, the observation that would be expected by studying the regulatory function of these two regulators (Uwate et al., 2013).

*amp*R D135N also showed higher predictive performance in the current analysis which support findings from the literature about the important role of AmpR which is defined as a major global regulator. AmpR is known to play an important role in acute infections through regulation of several pathways in *Ps. aeruginosa* including virulence, biofilm formation, general metabolism,

and quorum sensing (Balasubramanian et al., 2012). It has been shown that *amp*R can regulate transcription of hundreds of genes from diverse pathways (Balasubramanian, Kumari and Mathee, 2014). In addition, gene transcriptions and other phenotypic assays have shown that *amp*R negatively regulates the transcription and function of MexEF-OprN efflux system through modulating the expression of MexT which encodes a positive regulator (Balasubramanian, Kumari and Mathee, 2014) (Cabot et al., 2012).

Cell membrane permeability changes appear to be the main contributor mechanism among the elements assessed in the current analysis for the aminoglycoside group. This is evidenced by the best performing markers, *pst*BE89Q, *arn*AA170T, *arn*DG206C, *lpt*AT55A, *lpt*AR62S, *fao*AT385A, and *nuo*GA890T in both predictive values and the regression model. The role of cell membrane permeability changes has been recognized in the literature as affecting aminoglycoside resistance. Imperfect production of membrane lipopolysaccharide (LPS) has been shown to prejudice the uptake of aminoglycosides across the outer membrane (Bryan, O'Hara and Wong, 1984). Similarly, quantitative or qualitative changes in the electron transport chain may jeopardize the active uptake process of these drugs across the cytoplasmic membrane. As a result, the intracellular concentrations of the drug may decrease to levels below that required for ribosome inhibition (Taber *et al.*, 1987). These mechanisms are probably implicated in the function of many of the variants identified using the current analysis.

For antibiotic target changes, *gyr*AT83I remains an important contributor to both ciprofloxacin and levofloxacin resistance phenotype as seen in the predictive model in section 3.5.5 and section 3.6.5. Other quinolone target site mutations including *par*CS87L, *par*CS87W, *par*CE91K, and *gyr*AD87N appeared to have some contribution to levofloxacin resistance as seen in the predictive model in section 3.6.5. On the other hand, target site changes did not appear to be a main contributor to aminoglycoside resistance. This may seem expected as observed by the very low frequency of mutations occurring at ribosomal target sites of the aminoglycoside group of antibiotics as reported in the literature. However, the novel variant *fus*A1D588G in *fus*A1 gene coding for elongation factor EF-G1A identified in the current analysis appeared to be an important contributor to aminoglycoside resistance. This is supported by different findings from the literature. EF-G1A is known to be crucial in protein synthesis since it mediates the translocation of mRNA and tRNA through the ribosome and participates in

the ribosome recycling process. This novel mechanism has been recently confirmed through site-directed mutagenesis as associated with a 1- to 3-fold increase in the MICs of tobramycin, gentamicin, and amikacin (Bolard, Plésiat and Jeannot, 2018). It has also been suggested that multiple mutations in *fusA1* co-contribute to aminoglycosides resistance as detailed above. Mutations in *fusA1*, coding for the elongation factor EF-G1A, has also been linked to aminoglycoside resistance in both clinical isolates and *in-vitro* mutants (López-Causapé *et al.*, 2018).

In conclusion, results of the analysis performed here show that complete prediction of aminoglycosides resistance appears to be challenging. Although cell membrane permeability changes-related genes appear to be the most significant, these variants do not offer full explanation for the variability in aminoglycoside phenotype. This is evidenced by the finding that evaluated elements of quinolone resistance can explain 77.2% and 60.9% of variance in quinolone phenotype for ciprofloxacin and levofloxacin respectively (section 3.5.5 and section 3.6.5) while all evaluated chromosomal elements for aminoglycosides can only explain 17.9% of variance in amikacin MIC and only 21.6% of variance in gentamycin MIC (section 3.7.5).

This result can either be explained by the fact that the main core mechanism of aminoglycoside resistance is still not understood or that aminoglycoside inactivating enzymes are the main contributors of resistance to aminoglycoside agents. The current analysis aimed at studying the chromosomal mechanisms of resistance and has not included the assessment of modifying enzymes. These enzymes were assessed in the most diverse set of isolates (87 *Ps. aeruginosa* isolates) using the three aminoglycoside-modifying enzymes which are included in ResFinder and CARD database and did not show correlation with aminoglycoside resistance phenotype (as shown in detail in *Chapter 2).* Although aminoglycosides resistant isolates in that assessed set tend to be under-represented for amikacin, acquired resistance genes seems to be similarly identified in both susceptible and resistant isolates for gentamycin. This finding can be supported by other similar findings showing the carriage of acquired resistance genes even in clinically susceptible or wild type *Ps. aeruginosa* isolates (Pitt *et al.*, 1990). Vaziri *et al.*, (2011) have also supported these findings by showing that different inactivating enzymes studied including *aac* (6′)-I, *aac* (6′)-II, *ant* (2″)-I, and *aph* (3′)-VI have been identified at different frequencies that ranged between 7 % to 36 % of aminoglycoside resistant isolates. However, a further detailed

analysis that includes all other modifying enzymes in a larger selection of isolates may be required to answer this question.

An important point to consider here when such a correlational analysis is used that some sort of confounding can be introduced with genotypic-based prediction when correlating antibiotic resistance phenotype to genotype in *Ps. aeruginosa.* This confounding can originate from genomic islands that are known to have an essential role in the dissemination of multiple antibiotic resistance (Chowdhury *et al.,* 2016). Class 1 integrons carrying several gene cassettes linked to antibiotic resistance are known to be highly mobilized (Martinez *et al.,* 2012) through lateral gene transfer which can play a major role in the dispersal of mobile genetic elements. However, the role of chromosomally located genomic islands in transferring resistance cassettes in *Ps. aeruginosa* has been generally under appreciated. Defining the genetic context of class 1 integrons is essential since several reports have shown the possibility of encountering class 1 integron in diverse clonal lines in *Ps. aeruginosa* including lines that show no evidence of plasmids (Stokes *et al.,* 2012).

Therefore, the chromosome has been suggested as an important platform for the dispersal of complex resistance regions in *Ps. aeruginosa.* This has been evidenced by the presence of class 1 integron in some core chromosomal regions or embedded in some known genomic islands (Martinez *et al.,* 2012). Additionally, class 1 integrons possess mechanisms that may facilitate re-arrangement under stress conditions as may occur in the case of antibiotic exposure and this may result in the re-arrangement of clinically important genes and also in mobilization through lateral gene transfer. This can occur at very high rates in some situations (Hocquet *et al.,* 2012).

Hocquet *et al.,* (2012) showed that SOS response induction activated by antibiotic exposure can activate the integrase IntI1 which can consequently modulate antibiotic resistance in class 1 integron based on insertion of some genetic elements (Hocquet *et al.,* 2012). The multiplicity of genomic platform through which integrons can be mobilized and the unpredictable behavior of insertion sequences that are capable of mobilizing genes in novel ways can consequently influence resistance profiles in an unpredictable way. This is more observable when co-operatively interacting with other mutations in the cell. This means that resistance profiles are sometimes not easily predictable based on the common resistance-underlying pathways. This makes it important to realize that multiple existing and emerging mechanisms may interact to

generate diverse resistance phenotypes that can be modulated and disseminated through different chromosomal or horizontally acquired platforms (Martinez *et al.*, 2013). The next chapter will evaluate the best identified molecular predictors in relation to high-risk clonal groups.

Chapter.4: Background population genomic context of Quinolone and Aminoglycoside resistance determinants and molecular markers in *Pseudomonas aeruginosa*

Introduction and Background

Accurate identification and classification of infectious agents and subsequent choice of the best intervention strategies is the ultimate goal of clinical diagnostic methods. Microbiology diagnostics can be used either to direct individual patient choice of treatment or for public health interventions to control localized outbreaks and larger epidemics across countries and continents. Central to the entire discipline of clinical microbiology is the concept of having a laboratory tool that enables precise identification and microbial classification with accurate recognition of subtypes and biologically significant markers within any microbial species. This can help to control communicable diseases through the identification of infection source, route, and mechanism of transmission. This consequently facilitate the control of its spread in susceptible populations.

Typing schemes are based on the principle of finding measurable differences for classification. Typing methods vary widely in stability, discriminatory power, and reproducibility. In all cases, the marker used for typing and classification should reflect a significant biologic character (Ranjbar *et al.*, 2014). The following section briefly explores different typing methods used with *Ps. aeruginosa* in the context of studying the population biology of the organism.

4.1. Typing approaches for *Pseudomonas aeruginosa*

Microbial typing methods can be broadly classified into conventional typing methods and molecular typing methods. (Towner and Cockayne, 1993). Conventional methods for microbial typing assess the phenotypic gross cell behavior and base the classification upon that. Conventional typing methods used with *Ps. aeruginosa* include biotyping, phage typing, serotyping, and bacteriocin typing. On the other hand, molecular typing enables the study of microbial diversity based on the information carried by the cellular macromolecules including nucleic acids, proteins, and lipopolysaccharides (Towner and Cockayne, 1993). Despite being more specific, molecular methods are also subject to variation resulting from downstream regulation and intermolecular interactions (Foxman *et al.*, 2005).

4.1.1. Conventional typing for *Ps. aeruginosa*

Typing of *Ps. aeruginosa* was first developed for purposes of hospital epidemiology. It was specifically applied for typing the organism in cystic fibrosis patients. As with other organisms,

the earliest methods used for *Ps. aeruginosa* typing were the conventional ones including phage typing, bacteriocin typing, and serotyping. These methods were commonly used in many hospital epidemiologic studies. Serotyping has been considered a benchmark typing method commonly used for *Ps. aeruginosa* classification together with other methods including pyocin typing and phage typing. Classification of the organism using serotyping is based on the identification of group-specific heat-stable lipopolysaccharide O antigen. The earliest serotyping scheme for *Ps. aeruginosa* was proposed in 1961 by Verder and Evans, (1961). Another scheme was then developed in China by J. Yuan and Z. Zhao of the Shanghai Biological Products Institute (BPI) in 1963 (Yuan, J., & Zhao, 1963). Many other modifications and schemes were also developed but the final one universally agreed was the international antigenic typing scheme (Liu and Wang, 1990) that classified the species *Ps. aeruginosa* into 20 serotypes (O1-O20) adding three more serotypes to the 17 existing groups from the international antigenic typing system (LIU *et al.*, 1983). Serotyping of *Ps. aeruginosa* has been used for several types of investigations. High risk serotypes were studied for their potential link to pathogenicity and virulence factors (Hostacká and Majtán 1997; Faure et al. 2003), for their link to specific antibiotic resistance profiles (Patzer and Dzierzanowska 1991; Bert and Lambert-Zechovsky 1996; Estahbanati, Kashani, and Ghanaatpisheh 2002; Jamasbi and Proudfoot 2008), and for their association with clinical outcomes (Lu *et al.*, 2014).

Although it is possible to find changes in structural antigens (e.g. changes that are related to lysogenic conversion), the marker used in serotyping can be considered relatively more stable when compared to pyocin typing. Serotyping has some drawbacks including lower discriminatory power, problems of poly-agglutination, changing typeability, or non-typeability of some strains (Brokopp, Gomez Lus and Farmer, 1977). However, serotyping of *Ps. aeruginosa* can be beneficially applied to track and classify cases of ventilator associated nosocomial pneumonia, burn wound infections, and keratitis. It is also useful in drawing some conclusions about the prevalence of high-risk clones. This would offer great diagnostic value in some resource-limited laboratory settings with limited access to more advanced genomic techniques. On the practical side, a continuous supply of serotype-specific antisera needs to be available, however, this may be limited to reference laboratories in some settings. A genotypic version of serotyping for *Ps. aeruginosa* which can replace the conventional method to overcome some of its technical drawbacks can be alternatively applied in laboratories with sequence-based

diagnostics infrastructure. The *in-silico* serotyping version directly assesses the gene encoding for the antigen marker used for typing. A sequence - based serotyping tool has been developed for routine hospital outbreak surveillance. The tool can determine isolate serogroup and classify high-risk clones based on the sequence of the O-specific antigen (OSA) gene cluster in *Ps. aeruginosa* (Thrane, Véronique L. Taylor, *et al.*, 2016).

4.1.2. Molecular Typing for *Ps. aeruginosa*

Genotyping uses a set of molecular markers to classify isolates' relatedness. Genotyping is often used for surveillance and hospital epidemiology purposes or for research purposes to study the population structure of the organism. It has been recognized that molecular polyphasic approaches are needed to replace phenotypic, biochemical, or morphological assays in order to classify organisms more precisely and meaningfully (Prakash *et al.*, 2007).

Multi Locus Enzyme Electrophoresis (MLEE) and Random Amplified Polymorphic DNA (RAPD) typing are two common molecular typing methods used with *Ps. aeruginosa*. These methods were commonly used in some studies to make inferences about the *Ps. aeruginosa* population structure. Other molecular typing methods that have been commonly used with *Ps. aeruginosa* include Multi-Locus Variable Number Tandem Repeat Analysis (MLVA) and Pulsed Field Gel electrophoresis (PFGE).

PFGE has been used to provide high resolution macro restriction analysis at the whole genome level. It is considered highly discriminatory with standardized criteria of interpretation. The ability of the method to resolve very large DNA fragments increases its discriminatory power (Johnson *et al.*, 2007). This has led to considering the method as a gold standard to define isolate relatedness in most studies. The method is suited for long term epidemiologic studies and population studies with higher discriminatory power. The main reason for that is its ability to provide a classification based on whole genome restriction pattern and not only on single locus analysis. However, it shows low intra-laboratory and inter-laboratory reproducibility which also limits the usefulness of its discriminatory power. In theory, being based on restriction pattern analysis, it is less informative than other locus-based methods which are considered more specific (Botes *et al.*, 2003). Another technical factor that limits its practical usefulness is the long laboratory time needed for implementation in addition to being too costly and labor

intensive to be applied on a routine basis to analyze large numbers of samples on a daily basis, or for outbreak-related investigations.

MLVA is another PCR-based molecular typing method used with *Ps. aeruginosa*. The method can detect variations in short sequence repeat motifs (called microsatellites) in the genome of interest in a rapid, easy to perform, inexpensive, and reproducible way with higher resolution when compared to PFGE. When compared to RAPD and allozyme analysis, microsatellites offer a lower number of variable loci to be analyzed with higher number of alleles per locus (Jarne and Lagoda, 1996) and these characters add to its discriminatory power.

Variable Number of Tandem Repeat (VNTR) analysis or MLVA has been used to provide information for epidemiological investigations and follow up of chronic *Ps. aeruginosa* infection. PFGE and MLVA results appear to be congruent, however, MLVA can provide additional information in some cases with the possibility of VNTR variants to reveal the evolution of strains during long term infection. Tandem repeats using VNTR typing method have higher mutation rates and can be considered a fast-evolving marker. This means that they can show distinctive patterns within a short period of time making it suitable for surveillance of local outbreaks and for tracking infection sources (Sobral *et al.*, 2012). Some MLVA-typing schemes were developed for use in *Ps. aeruginosa* and these include MLVA-15 Orsay (Vu-Thien *et al.*, 2007), MLVA-16 Orsay (Sobral *et al.*, 2012), MLVA-9 London (Turton *et al.*, 2010), and MLVA-9 Utrecht (Rosa van Mansfeld *et al.*, 2010).

MLST has been developed as a molecular version of MLEE and has been applied to study the population structure based on the concept of indexing nucleotide sequence variation present in only seven house-keeping gene fragments (Ibarz Pavón and Maiden, 2009). The method is considered an improved way of addressing evolutionary biology because it uses housekeeping genes. These genes are assumed to be under neutral or nearly neutral selection. Allele designation as used in MLST gives the same weight to both recombination events and point mutations (Maiden, 2006). Although considered an advantage, it can be less discriminatory. MLST was first introduced in 1998 to index allelic variation for essential metabolic genes. Although schemes developed in different organisms are considered highly discriminatory in addition to reflecting species diversity (Maiden, 2006), allelic loci used in typing represent less than 0.2% of the genome in question. For that reason, the MLST scheme was extended to

rMLST that indexes the variation in a larger collection of loci including 53 ribosomal protein subunit genes. The typing scheme was extended to include the complete set of coding sequences in the core genome or the whole bacterial genome in cgMLST and wgMLST respectively. This results in increasing the discriminatory power in each newly developed scheme by including more typing loci (Jolley and Maiden, 2014) (Dekker and Frank, 2016).

MLST is more usefully used to study the population biology of the organism rather than for local epidemiologic investigations of hospital outbreaks. The reason for that lies in its lower discriminatory potential to detect variability among closely related strains. Another advantage of the method that has led to its widespread use lies in its ability to easily construct an electronically accessible databases with generating digital data that are highly accurate and portable. This makes it easy to compare results among laboratories on a global basis. However, high cost and limited accessibility to routine sequencing machines in some laboratories can limit its routine application (Maiden, 2006). Another observed drawback of the method lies in the fact that different sequence types are assigned based on a combination of alleles which are assigned numbers on arbitrary bases. This means that different STs do not necessarily reflect different behavioral types. In addition, unknown and newly identified sequence types that result from finding new combinations of alleles can continuously be added. Although MLST is based on indexing diversity on the basis of allele combinations into different STs, it is not providing the full discriminatory potential.

4.2. Population structure in *Pseudomonas aeruginosa*

Proper understanding of the population structure of an organism is essential to make meaningful clinical conclusions. This understanding is required when molecular typing is used in surveillance, epidemic and outbreak investigation and for the dentification of epidemic high-risk clones (Foxman *et al.*, 2005).

Conclusions drawn about the population structure of the organism are largely based on the type of markers used in the analysis, however, this may show different sources of bias. This consequently necessitates using a combination of markers instead of using a single marker to measure different evolutionary forces (Van Belkum *et al.*, 2001). Most studies analyzing the population structure of *Ps. aeruginosa* have concluded that the organism exhibits non-clonal

structure with occasional epidemic clones. This has been suggested in most studies using different typing methods.

Ruimy *et al.*, (2001) used random amplified polymorphic DNA to study the genetic diversity of *Ps. aeruginosa* from different types of infection and from environmental sources. The study concluded that the organism exhibits high rate of recombination in the population with no stable clonal genotypes differentiating pathogenic from non-pathogenic strains and has also shown that there is no clear phylogenic separation evident between the two groups. The study also highlighted that a highly discriminant epidemiologic marker is usually needed to perform such investigations (Ruimy *et al.*, 2001).

Pirnay *et al.*, (2009) have investigated 328 unrelated *Ps. aeruginosa* isolates by combining eight parameters to analyze the population structure. The markers used for typing included Amplified Fragment Length Polymorphism (AFLP), O serotyping and several other selected genes that are related to important traits (Pirnay *et al.*, 2009). Findings from another study (Maatallah *et al.*, 2011) investigating 110 strains from five Mediterranean countries using different molecular markers including PFGE, MLST, serotyping, and some virulence genes, have supported the same observation about the panmictic population structure for *Ps. aeruginosa* clinical and environmental isolates (Maatallah *et al.*, 2011). Another study investigating 184 isolates from four countries of west and central Africa using MLST has shown a nonclonal epidemic population with observations of high-risk international clonal complexes. Eighty distinct STs were reported with 24 STs showing no correlation with any previously known sequence types (Cholley *et al.*, 2014). In another study, 501 *Ps. aeruginosa* isolates from environment, animals, and human infection (both CF and non CF) were evaluated for genetic diversity and has also concluded a nonclonal epidemic structure (Timothy J. Kidd *et al.*, 2012). Earlier studies carried out by Pirnay *et al.*, (2002) on 73 clinical and environmental isolates have inferred a panmictic population structure of the organism. Other studies have used AFLP, a set of outer membrane genes, serotypes, and pyoverdine types in combination. The same was also suggested in two earlier studies (Denamur *et al.*, 1993) (Picard *et al.*, 1994). These two studies have compared the genetic diversity of *Ps. aeruginosa* population structure using MLEE and RFLP and showed no observed correlation between both methods used (Denamur *et al.*, 1993) (Picard *et al.*, 1994). Therefore, the study recommends that no single genetic marker should be used to reflect the

population structure. In contrast, another study in 2007 has reported a high degree of clonality in the population structure of *Ps. aeruginosa*. The study used a new genotypic approach based on 58 distributed target elements for microarray typing. These targets were chosen to represent the core and accessory genome. However, the markers chosen may not be comprehensively representative or inclusive of all loci needed to represent the whole genome or to represent the evolutionary forces needed to achieve the highest level of resolution that can correctly classify the population (Wiehlmann *et al.*, 2007). These factors may have resulted into different conclusions.

In general, the population structure of *Ps. aeruginosa* is consensually believed to be of panmictic-epidemic nature (Magalhães *et al.*, 2020). Other more recent findings from the literature have also indicated that the population structure of *Ps. aeruginosa* is predominantly non-clonal based on genotyping using PFGE (Brzozowski and Jursa-kulesza, 2020). A study from a Mexican clinical collection has also supported the same findings about the epidemic structure of *Ps. aeruginosa* which showed to be represented by some globally distributed clonal complexes (Castaneda-Montes *et al.*, 2018). Some other recent findings have also supported the same conclusions showing the same epidemic structure related to previously observed high-risk clones (Pelegrin *et al.*, 2019). Another study using both MLST and core genome MLST and comparing the allelic profiles related to epidemic high-risk clones from both schemes has identified the ST235-O11 serotype cluster to contain identical cgMLST profiles (Royer *et al.*, 2020). In another study, cgMLST allelic profiles have been used to construct MST for some studied high-risk clones and did not show much heterogeneity in the epidemic clones studied (Schaumburg *et al.*, 2017).

4.2.1. Observation of high-risk clones and their linkage to behavior

High-risk clones are specific “serotypes” or “sequence types” that have been frequently observed and linked to specific types of bacterial behavior including virulence, antibiotic resistance, site specific pathogenicity (e.g. cystic fibrosis, keratitis), or infection outcome and patient prognosis. It has been suggested that a limited number of frequently observed widespread clones are responsible for human infections (A. Oliver *et al.*, 2015) with other clones showing link to higher morbidity and mortality rates in cystic fibrosis patients (Fernández-Olmos *et al.*, 2013), (Van Mansfeld *et al.*, 2009), (Caballero *et al.*, 2014), (Garcia-Castillo *et al.*, 2011).

Specific dominant multi-drug resistant and extensively drug resistant (XDR) clones appear to be disseminated in hospitals worldwide. These clones include ST235, ST111, and ST175 which were repeatedly reported in several studies (Antonio Oliver *et al.*, 2015). Marked as high-risk clones, they play a major role in the spread of resistance worldwide (Woodford, Turton and Livermore, 2011). In many studies, these three clones accounted for the majority of XDR isolates (Pena *et al.*, 2015). ST235, ST175, and ST132 have been identified as the most prevalent multi-drug resistant clones in a study including 108 blood stream isolates from 48 hospitals in 36 cities in Czech (Nemec *et al.*, 2010). Another study including 187 isolates from different hospitals in France demonstrated the dissemination of few successful international clonal complexes including ST235, ST111, and ST175. These clones were also linked to multidrug resistance (Cholley *et al.*, 2011).

In a population of 123 *Ps. aeruginosa* isolates sampled from sink fittings of 5 wards in critical care units in a non-outbreak situation, high-risk clones were shown to be over represented including ST111, ST253, and ST235 (Varin *et al.*, 2017). This can demonstrate the underestimated importance of environmental transmission in the dissemination of these high-risk clones in clinical ward environments and this also indicate the probability of water systems to act as an essential source for the dissemination of high-risk clones. Multiple *Ps. aeruginosa* outbreaks have been previously linked to hospital water systems (Loveday *et al.*, 2014) (Walker and Moore, 2015) including the 2011/12 outbreak involving a neonatal unit in Northern Ireland causing death in four neonates (Jefferies *et al.*, 2012).

Clone ST235 is commonly reported as a successful international lineage associated with hospital infection and capable of acquiring diverse resistance mechanisms (Empel *et al.*, 2007). In addition, the clone has shown the widest international distribution across the five continents and is also known to be frequently related to serotype O11 (Antonio Oliver *et al.*, 2015). The clone is known to exhibit the highest prevalence among all observed clones in different studies. Among 27 genotypes, ST235 showed to be the most prevalent clone with no significant change in observed clones over a 10 year period (Feng *et al.*, 2019). In addition, other recent studies have also identified the frequent carriage of Metallo-b-lactamases among carbapenem resistant *Ps. aeruginosa* isolates from the clonal type ST235 (Osawa *et al.*, 2019). MLST used in more recent studies has also identified both ST235 and ST357 as the most frequently observed sequence

types (Horna *et al.*, 2019). Continuous surveillance of ST235 shows that the clone is still prevalent and worrisome and also showing significant association with antibiotic resistance carrying some known resistance determinants (Vatansever *et al.*, 2020).

ST111 is another widely distributed clone and is known to be related to serotype O12. A study including 448 *Ps. aeruginosa* clinical isolates from 16 Spanish hospitals showed that the four sequence types ST175, ST646, ST532, and ST111 were most prevalent. The successful international clone ST111 was the most frequent while ST235 was represented in only small number of isolates (García-Castillo *et al.*, 2011). In the same study, ST175 was genotypically uniform using PFGE and therefore was added to the list of significant epidemic clones reported in the organism population. Although phenotypes were not consistently linked to different STs, most of the strains belonging to ST175 were quinolone resistant and a high percentage showed resistance to gentamycin, ceftazidime, and cefepime. ST175 was formerly considered as a contaminant of the hospital environment and as a colonizer. However, it was later noticed to be widespread and associated with multidrug resistant phenotype (García-Castillo *et al.*, 2011). ST175 has also been linked to serotype O4 and has been observed in several European countries and in Japan (Antonio Oliver *et al.*, 2015). In another study, a set of randomly selected 56 clinical isolates of *Ps. aeruginosa* showing a high level of diversity and including MDR and non-drug resistant isolates was studied. The majority of MDR and XDR strains were grouped under sequence types ST175 and ST235. ST253 was the third in frequency and included non-MDR isolates. The study has concluded that non-resistant and resistant isolates can co-exist within the same sequence type and that most susceptible isolates corresponds to singletons (Gomila *et al.*, 2013).

More recent studies have also identified ST111 as the most frequent and disseminated clone showing lower clonal diversity as identified in MDR/XDR isolates (Pérez *et al.*, 2019). The same study has also demonstrated that most XDR/PDR strains belonged to the international high-risk clones ST111, ST235, and ST175. The most prevalent serotypes among *Ps. aeruginosa* populations include O11, O12, O6, and O1, with O11 and O12 being particularly common among multi-resistant isolates (Pirnay *et al.*, 2009). Although serotype O12 showed a relative genetic uniformity of isolates, serotype O11 demonstrated a greater genetic diversity (Samuelsen

et al., 2010) (Woodford, Turton and Livermore, 2011). Isolates with O11 serotype belonged to either ST227 or to ST230 (Giske *et al.*, 2006).

4.2.2. Clones linked to Cystic fibrosis

Cystic fibrosis is considered the most common autosomal recessive disorder in the Caucasian population. The disorder results from an abnormal regulation of chloride channels leading to chronic pulmonary disease with associated pancreatic exocrine insufficiency (Rommens *et al.*, 1989). *Ps. aeruginosa* is considered an important pathogen determining biggest morbidity and mortality in cystic fibrosis patients with MDR/XDR strains being highly linked to disease exacerbations (Mogayzel *et al.*, 2014).

Some sequence types have been frequently observed as persistent in the diseased cystic fibrosis population. These sequence types have been reported in several studies as epidemic high-risk clones. The most prominent of those includes the Liverpool Epidemic Strain (LES), ST146 and the Australian Epidemic Strain (AES), ST649. These were repeatedly reported in the cystic fibrosis population all over the world and were linked to high level of drug resistance showing higher morbidity and mortality (A. Oliver *et al.*, 2015). However, some other studies failed to demonstrate the same link between commonly observed clones and other types of linked behavior in cystic fibrosis patients. For example, a cross sectional study investigating the population structure of *Ps. aeruginosa* among cystic fibrosis patients from a Dutch hospital showed that the two sequence types (ST406 and ST497) were prevalent among cystic fibrosis population despite being not linked genetically to the previously described epidemic clones observed in cystic fibrosis lung infection (Van Mansfeld *et al.*, 2009).

Pirnay *et al.*, (2009) have suggested that there is little evidence about the existence of a global CF clone and that resistant CF clones isolated from different parts of the world appear to be genetically diverse. He also suggested that clones belonging to a core lineage as seen in patients with CF can be seen in the environment. This may indicate an independent acquisition in different settings rather than a globally disseminated clone. Another study comparing three genotypic methods have studied the population structure of the organism and reached a similar conclusion (R. van Mansfeld *et al.*, 2010). The study has reported that the structure of the population does not show evidence of a core lineage in which major CF, hospital, or community clones co-cluster (R. van Mansfeld *et al.*, 2010).

In agreement, another study including 42 different isolates from cystic fibrosis patients reported a high degree of diversity with ST242 being identified as the most prevalent lineage. In the same study, ST17 and ST809 were the next most prevalent sequence types followed by ST620. On the other hand, other STs detected in CF isolates including ST996, ST360, ST274 and ST980 were less commonly observed (Garcia-Castillo *et al.*, 2011). Additionally, T.J. Kidd *et al.*, (2012) have suggested that CF strains are likely to be a random sample of the wider organism population and that the increased abundance of specific types in some geographic localities can be a result of chance colonization followed by adaptation and subsequent local transmission (T.J. Kidd *et al.*, 2012). Another study in a Spanish hospital showed that international CF epidemic clones previously identified were not represented within that group of chronically colonized CF patients. A new sequence type representing a double locus variant of the Dutch ST497 CF epidemic clone not previously known was also identified in the same study (Fernández-Olmos *et al.*, 2013). Other observations from more recent literature have also identified ST17, ST155, and ST179 as the most pervasive clones identified from a set of data including 298 genomes from 5 cystic fibrosis clinics (Jeukens, Freschi, Kukavica-Ibrulj, J. G. Emond-Rheault, *et al.*, 2019). Another study applying MLST has shown wider diversity among the studied Canadian *Ps. aeruginosa* cystic fibrosis population and did not show widespread sharing of the previously known dominant clones (Middleton *et al.*, 2018).

4.2.3. Importance of high-risk clones

Based on the review above, it can be concluded that several multi-locus sequence types are distributed worldwide in relation to epidemics where multidrug resistance confounds treatment success. Therefore, it is important to understand the molecular basis of success of these epidemic high-risk clones because such understanding can greatly impact treatment selection. Although horizontally acquired resistance genes have been linked to clonal dissemination and success of some high-risk clones (A. Oliver *et al.*, 2015), acquired resistance mechanisms may not be the most essential or the sole contributor to clonal success (Roy Chowdhury, Scott and Djordjevic, 2017a).

Better prognostic markers are needed to make better evidence-based patient care decisions (Jeukens, Freschi, Kukavica-Ibrulj, J.-G. Emond-Rheault, *et al.*, 2019). These markers should predict AMR or high-risk behavior more rapidly and should show clear links to disease and

patient outcome. The advantage of using predictive values related to these markers will enable rapid intervention or therapy and will also help in molecular epidemiologic surveillance. Greater understanding of the general population structure and the predictive capacity of molecular markers linked to AMR or high-risk behavior is needed. The analysis used in this chapter aims at identifying molecular markers, signatures or combinations linked to high-risk groups. When such markers are detected, they can be used to efficiently direct efforts that can further reduce the spread of epidemic clones. For example, if a marker is linked to resistance behavior and at the same time linked to high-risk clone, it can indicate that the patient or setting carry high-risk potential and consequently necessitates additional precautionary measures or isolation practices to avoid its transmission.

Building on the results of *chapter 3*, the objective of this chapter is to describe the population structure of *Ps. aeruginosa* in a large comprehensive dataset as a primary step to explore the relation of different previously identified quinolone and aminoglycoside resistance markers to high risk clones. It also aims at exploring the specific molecular markers or combination of markers that explain the global success of epidemic high-risk clones.

Several epidemiologic and population biology studies have been performed using different molecular typing tools. Although comprehensive analysis of population structure is not the primary objective for the scope of the thesis, it was necessary to describe the population structure as a first step. Next, analysis of the correlation of identified quinolones resistance markers to background context and to high-risk clones was performed. Two commonly used and well-established typing methods including MLST and serotyping were used for that purpose. High-risk clones when described and studied in the literature are usually described using the commonly known sequence types and serotypes and that was another reason for using these two typing methods. The correlation and clustering of different quinolone and aminoglycoside molecular resistance markers were then explored and studied in relation to background population structure. For that purpose, a large collection of isolates including an in-house set of *Ps. aeruginosa* clinical isolates and another set of public genomes from the Patric database (Wattam *et al.*, 2017) were studied. The analyzed set included the whole spectrum of resistance profiles for ciprofloxacin, levofloxacin, gentamycin, and amikacin antibiotics. The panel of

molecular markers studied and identified in the previous chapter are evaluated in this chapter for their potential relation to clonal success.

4.3. Methodology

4.3.1. MLST and serotypes (O-type) analysis

MLST was performed for all isolates according to previously described typing scheme by Curran (Curran *et al.*, 2004)**.** An ST was assigned to each unique allelic profile according to the *Ps. aeruginosa* PubMLST database (http://pubmlst.org/paeruginosa/). Whole genome sequence data (WGS) for our in-house group of clinical isolates (163 genomes) in addition to the selected clinical isolates from Patric database (528 genomes) were used to identify STs using the method publicly available at www.cbs.dtu.dk/services/MLST (Larsen *et al.*, 2012). WGS data was also used to determine the serogroups of all studied isolate based on the sequence of O-specific antigen (OSA) gene cluster using the *Ps. aeruginosa* serotyper (PAst) web-tool available on the Center for Genomic Epidemiology (CGE) service platform (https://cge.cbs.dtu.dk/services/PAst-1.0/) (Thrane, Véronique L Taylor, *et al.*, 2016).

4.3.2. Population structure and diversity analysis

Strain relationships were analyzed using the geoBURST Full MST algorithm (Francisco *et al.*, 2009), as implemented in the software PHYLOVIZ (Francisco *et al.*, 2012) to construct a Minimum Spanning Tree (MST) of the total set of *Ps. aeruginosa* strains based on MLST data according to the steps shown in Phylophiz documentation release 2.0 available at http://www.phyloviz.net/goeburst/Tutorial.html

The eBURST algorithm, developed specifically for MLST data, can estimate the evolution of each clonal complex from its ancestral genotype or primary founder. An eBURST group (or clonal complex) is defined as an ancestral sequence type with all its related sequence types (STs) at SLV, DLV or, TLV with at least one other ST in the group and each ST can only be assigned to a single group.

Clonal complexes (CC) were defined in the current analysis as complexes or "groups of studied isolates" containing at least three STs sharing the same allele numbers in at least five of seven loci. Isolate-specific metadata, including serotypes, quinolone, and aminoglycoside resistance data, were then overlaid on top of the minimum spanning tree. Allelic linkage disequilibrium was

assessed with two test options of both Monte Carlo methods and Parametric with 100 resampling using LIAN version 3.7 (Haubold and Hudson, 2000) available at http://guanine.evolbio.mpg.de/cgi-bin/lian/lian.cgi.pl/query.

The standardized index of Association (I^S_A) and the Mean genetic diversity (H) are two measures used to assess linkage equilibrium and degree of association between alleles. An index of association (I_A) with a lower value approaching 0 is considered highly suggestive of a non-clonal epidemic population (Smith *et al.*, 1993). Higher mean genetic diversity index also supports linkage equilibrium and low evidence of association among the alleles analyzed.

High mean genetic diversity index together with low index of association may support the idea that recombination is playing role in randomizing allele distribution in the population (Lenski, 1993).

Discriminatory power (Simpson's index) and concordance (cluster agreement) between 2 typing methods (Adjusted Rand index and Adjusted Wallace) were evaluated according to (Carriço *et al.*, 2006) using the web source http://www.comparingpartitions.info/.

The degree of concordance between the two typing schemes used was first evaluated. Simpson's index of diversity (SID: with 95% confidence intervals) was used as described by (Hunter and Gaston, 1988). Inter-method concordance was also evaluated using the Adjusted Wallace coefficient (Severiano *et al.*, 2011). Adjusted Wallace coefficient shows the probability that two strains classified as the same type by one method will also be classified as the same one when using the other method. Adjusted Wallace coefficient allows for assessing the strength and directionality of the concordance between the various typing methods according to (Carriço *et al.*, 2006). A value of 1 indicates that clusters generated by a given typing method could have been completely predicted by another methodology or in other words one method is completely concordant with the other method. In this way, it represents the prediction probability of a pair of strains that are assigned to the same type by one method are also classified in the same type by the other method.

4.3.3. Resistance genes and markers correlations

The set of previously identified resistance and sensitivity markers described in *Chapter 3* were explored for their association with high risk STs and serotypes and their potential to explain the

success of epidemic high-risk clones. Chi-square test for independence (with yates' continuity correction) was used to compare groups. A *p*-value of < 0.05 was considered as statistically significant.

A panel of 10 previously tested ciprofloxacin resistance/sensitivity markers and of 11 tested levofloxacin resistance/sensitivity markers were analyzed. In addition, the 5 previously identified clusters of molecular markers were tested for possible correlation with high-risk clones. These markers included *mex*Z, *nal*C S46A, *nal*C S209R, *nal*C G71E, *gyr*A T83I, *nal*C E153Q, *nal*C Thr50pro, *mex*S gene, *nal*D gene, *nfx*B gene, *arm*R gene, *par*E V460G, and *amp*R D135N.

For aminoglycoside, molecular markers tested included 11 markers for gentamycin and 18 markers for amikacin. These markers included: *pho*QY85F, *nuo*GA890T, *pst*BE89Q, *lpt*AT55A, *lpt*AR62S, *fao*AT385A, *arn*AA170T, *arn*DG206C, *mex*RR79N, *mex*RE70R, *mex*RL130T, *mex*RG97L, *mex*RL29D, *mex*Z, *amp*R gene, *pmr*BGly423Cys, *pmr*ALeu71Arg, *fus*A1D588G, *gid*BE186A, *arm*R (PA3719), *nal*C gene, *nal*D, *nal*Dser32Asn, *gid*BQ28K, *gid*BE97Q, *nal*CE153Q and *amp*RA51T.

4.3.4. Phylogenetic analysis

Phylogenetic analysis and hierarchical clustering were performed to evaluate the distribution of the studied set of isolates among all known *Ps. aeruginosa* genomes. The genes encoding the following metabolic enzymes: *acs*A (acetyl coenzyme A synthetase), *aro*E (shikimate dehydrogenase), *gua*A (GMP synthetase), *mut*L (DNA repair protein), *nuo*D (NADH dehydrogenase I chain C, D), *pps*A (phosphoenolpyruvate synthase) and *trp*E (anthralite synthetase component I) are commonly used for MLST typing.

Concatenated sequences of these genes for the studied set of 691 isolates as extracted from the MLST output tool provided by https://cge.cbs.dtu.dk/services/MLST/ and concatenated sequences of all other known STs for *Ps. aeruginosa* as extracted from the *Pseudomonas aeruginosa* MLST website available at https://pubmlst.org/paeruginosa/ were aligned using the MUSCLE option (Edgar, 2004) implemented in the software MEGA 7 (Kumar, Stecher and Tamura, 2016). The phylogenetic tree of the concatenated genes was constructed using the UPGMA algorithm.

4.4. Results

4.4.1. Description of population structure in the studied set of isolates

Figure 4.1 shows the minimum spanning tree (MST) analysis of *Ps. aeruginosa* strains based on all known STs identified with *Ps. aeruginosa*. The tree shows all the sequence types observed in the studied set of isolates among all known STs for the organism. Implementing the geoBURST algorithm at the TLV level showed a total of 172 clonal complexes with 125 singletons for all known *Ps. aeruginosa* STs. This indicates high diversity and shows that the isolates under study are widely distributed among all known STs. The study isolates are colored on the graph (Figure 4.1) as shades of dark green and dark blue. The graph also shows the position of high-risk clones among all other set of isolates included in the analysis.

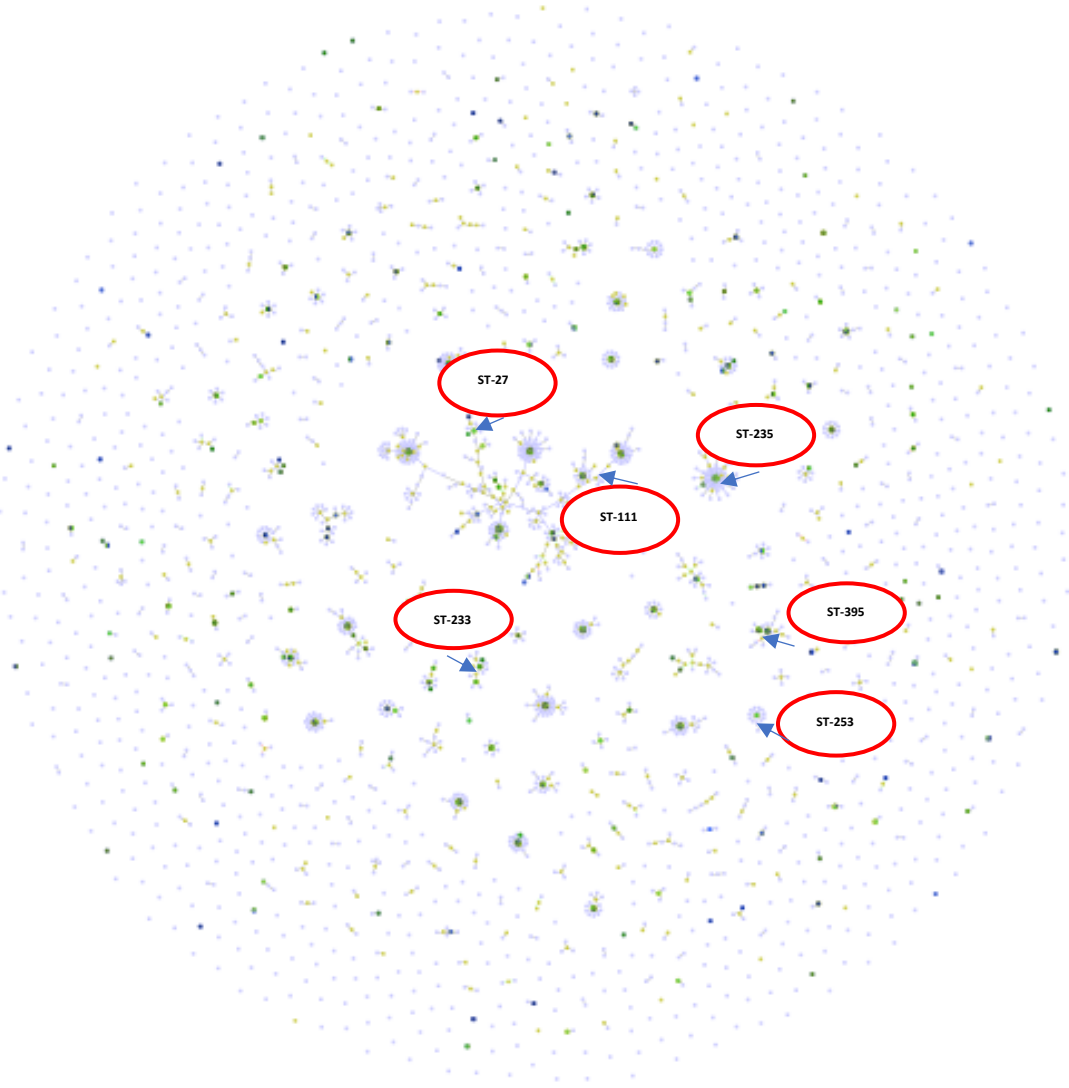


Figure 4.1. MST analysis of studied Ps. aeruginosa isolates among all known STs

Study isolates are shown in dark circles (shades of dark blue and dark green)

High-risk groups are shown in red circles and blue arrows

4.4.2. Population structure and diversity

Applying geoBURST algorithm at double locus variant level (DLV) showed 219 clonal complexes with 176 singletons. The most frequent clonal complexes observed were: CC 233 consisting of 20 STs (39 isolates), CC17 consisting of 4 STs (16 isolates), CC 532 consisting of 4 STs (17 isolates), CC 560 consisting of 7 STs (12 isolates) and CC 316 consisting of 3 STs (12 isolates). Other important high-risk clones formed clonal complexes that showed only 2 STs. Examples include: CC 446 consisting of 2 STs (15 isolates) and CC 111 consisting of 2 STs (35

isolates). The index of association (I_A) was calculated to estimate the degree of association and recombination between alleles at different loci based on MLST allelic profile data (Haubold and Hudson, 2000). When all 691 isolates were analyzed, the value of the standardized index of Association (I^S_A)=0.1302 (P<0.001) which indicates linkage equilibrium and low evidence of association among alleles analyzed. Pairwise variance (V_D =1.4063) was greater than the critical value (L=0.8007). Mean genetic diversity (H): 0.8648 +/- 0.0261. These results support that recombination plays a key role in allele distribution and support the non-clonal structure of *Ps. aeruginosa* population based on MLST classification of the studied set of isolates.

A phylogenetic analysis of the concatenated sequences of the MLST alleles was performed including the study isolates and the entire MLST database. The results showed diversity and non-clustered distribution of the study isolates among all known STs for the organism. Result shown in Figure 4.2.

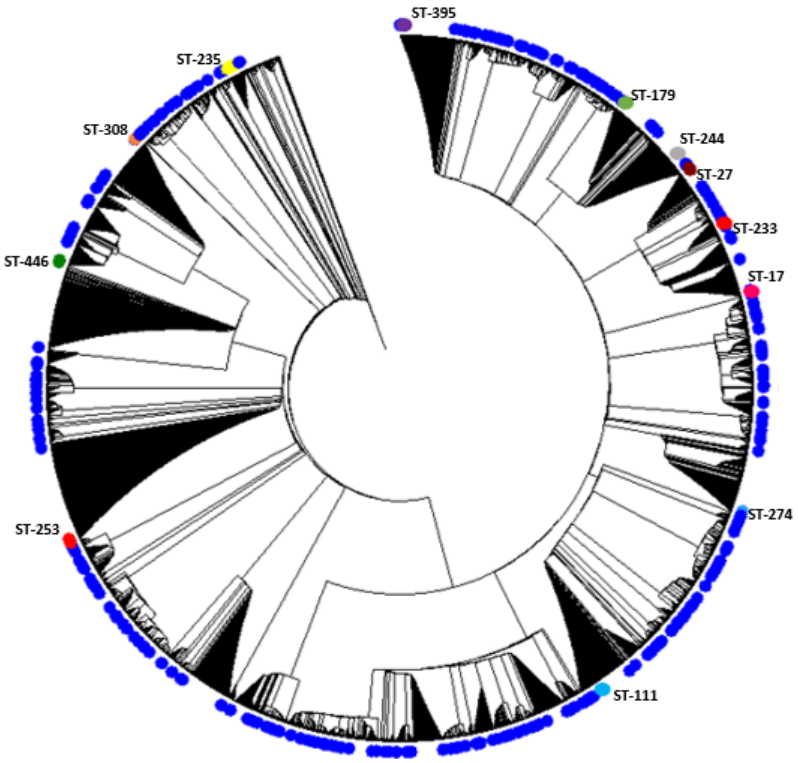


Figure 4.2. Hierarchical clustering showing the distribution of analyzed sequences among all known ST

Blue circles indicate the position of the studied set of isolates among all other Ps. aeruginosa isolates

The positions of high-risk clones are also shown on the figure using different colors

4.4.3. Analysis of MLST profile in relation to serotype

Figure 4.3 shows Minimum Spanning Tree (MST) analysis of the studied set of *Ps. aeruginosa* strains based on MLST data at SLV level: each circle corresponds to an ST identified in the studied collection of isolates. The area of each circle corresponds to the number of isolates showing certain ST. The distance relationship between the isolates is indicated by the lines connecting the isolates. Position of high-risk groups is shown in red on each graph. In this figure different ST groups are colored based on the corresponding serotype.

A population analysis of 691 *Ps. aeruginosa* isolates was performed and considerable genetic diversity was observed among the MLST results. MLST analysis identified 311 STs among all isolates including 266 known and 45 novel STs. The international clones ST235 (serotype O11[98 %]) was the most frequently identified in a total of 50 genomes, followed by ST111(serotype O12 [83.3%], serotype O4 [16.7%]) which was identified in a total of 30 genomes. ST244 (serotype O12 [35%], serotype O2 [30%], serotype O5 [30%]) was identified in 20 genomes while ST308 (serotype O11[100%]) was identified in 18 genomes. Each of sequence types ST395 (serotype O6 [100%]) and ST253 (serotype O10 [94.2%]) was identified in 17 genomes. ST348 (serotype O2 [53.3%], serotype O5 [40%]) was found in 15 isolates. ST274 was identified in 14 isolates. Each of ST179 (serotype O6 [92.3%]) and ST233 (serotype O6 [100%]) was found in 13 isolates. ST17 (serotype O1 [100%]) was identified in 12 isolates while ST27 (serotype O1 [100%]) was identified in 11 isolates and ST175 (serotype O4 [100%]) in 10 isolates. All the data above are visualized in Figure 4.3 which shows serotypes overlaid on corresponding STs.

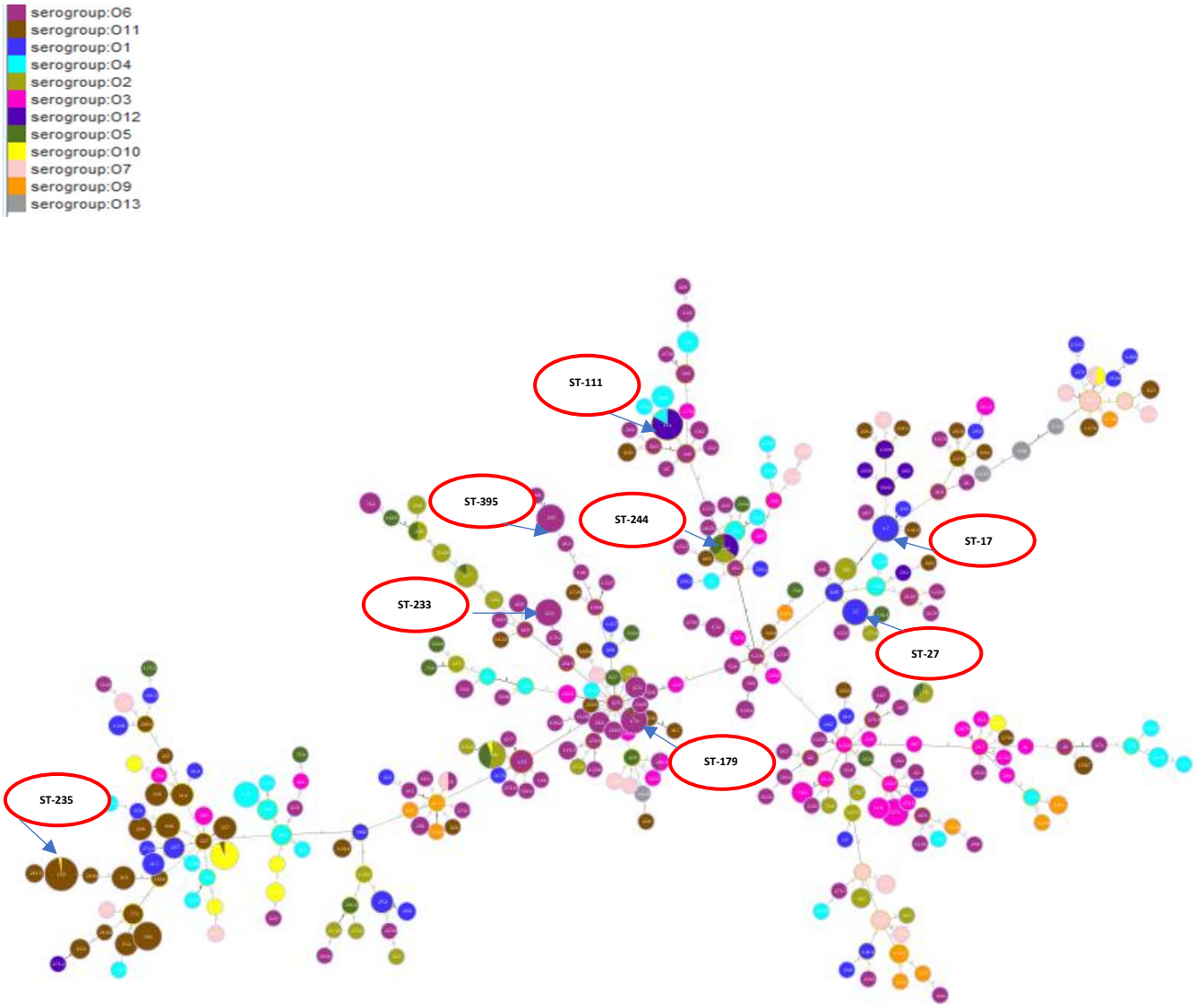


Figure 4.3. MST showing distribution of serotypes in relation to STs

Each serotype is shown in a different color as indicated in the figure ligand above

The figure also shows how different serotypes are clustering in relation to different clonal complexes.

High-risk clones are shown on the figure in red circles.

As seen in Figure 4.3, O6 serotypes showed tendency to cluster towards ST-395, ST-233, and ST-179. O12 serotype constituted the majority of isolates belonging to ST-111 and about one third of the isolates belonging to ST-244. The rest of O12 serotype group belonged to other less frequent sequence types including; ST-1006, ST-2031 and ST-3043 which are at the distances of 12, 12 and 7 locus variants respectively from ST-17 as shown in figure 4.2. O11 serotype formed nearly all the isolates belonging to ST-235 and many of its close variants. This is represented in brown color seen at the left part of figure 4.3. Serotype O1 shown in blue appeared to be more widely distributed across different distant clonal groups. It showed clustering at ST-27 and ST-17 and is also related to other sequence

types including ST-252, ST-313, and ST-207 among others. The complete set of data are shown in section 4.4.3 and are also detailed in table 4.1.

The degree of concordance between the two typing schemes used was evaluated using Simpson's index of diversity (SID: with 95% confidence intervals) as described in the methods section and showed that MLST (Simpson's ID=0.987 with 95% CI [0.984-0.990]) was more discriminatory than serotyping (Simpson's ID=0.856 with 95% CI [0.843-0.869]). Inter-method concordance was also evaluated using the Adjusted Wallace coefficient (Severiano *et al.*, 2011). Adjusted Wallace coefficient shows the probability that two strains classified as the same type by one method will also be classified as the same one when using the other method. A value of 1 indicates that clusters generated by a given typing method could have been completely predicted by another methodology which represents the prediction probability of a pair of strains that are assigned to the same type by one method are also classified in the same type by the other method.

Adjusted Wallace between ST and serotypes =0.840 with 95% CI (0.792-0.889) while that between serotypes and STs was significantly low =0.064 with 95% CI (0.048-0.080) which means that ST can predict serotype with higher confidence while the opposite is not true.

4.4.4. Quinolone resistance profile of *Ps. aeruginosa* epidemic high-risk clones

Table 4.1, Figure 4.4 and Figure 4.5 all summarize quinolone susceptibility data in relation to different STs and high-risk groups.

Table 4.1 summarize the quinolone MIC values for the high-risk clones observed in the studied collection with corresponding serotypes. Figure 4.4 is an MST tree showing levofloxacin susceptibility in relation to different STs. Figure 4.5 is an MST tree showing ciprofloxacin susceptibility in relation to different STs.

Table 4.1. Summary of isolates belonging to high-risk clones with their corresponding serotypes and quinolones susceptibility

ST	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235
Genome ID	287.1021	287.1031	287.1045	287.1058	287.1087	287.1088	287.11	287.1122	287.1124	287.1128	287.1131	287.1133	287.1134	287.1135	287.1144	287.1145	287.1147	287.1148	287.115	287.1151
serogroup	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11

levofloxacin sensitivity	R	S	R	R	R	R	R	S	R	R	R	R	R	R	R	R	R	R	R	R
ciprofloxacin sensitivity	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN
levofloxacin MIC	32	0.06	32	32	32	32	32	0.25	8	32	32	32	32	32	32	32	32	32	32	32
ST	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235
Genome ID	287.2976	287.5686	287.569	287.6327	287.6328	287.778	287.7782	287.7783	287.7785	287.7786	287.783	287.855	287.86	287.902	287.903	287.915	287.918	287.955	287.966	287.984
serogroup	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11	O10	O11	O11	O11	O11	O11	O11
levofloxacin sensitivity	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R	R	R	R	S
ciprofloxacin sensitivity	R	R	R	R	R	R	R	R	R	R	R	UN	UN	UN	UN	UN	UN	UN	UN	UN
levofloxacin MIC	8	8	8	8	8	4	4	4	4	4	4	128	0.015	4	32	4	32	32	32	0.12
ST	235	235	235	235	235	235	235	235	235	235										
Genome ID	287.1158	287.1162	287.1163	287.1184	287.12	287.1202	287.1203	287.1213	287.2973	287.2975										
serogroup	O11	O11	O11	O11	O11	O11	O11	O11	O11	O11										
levofloxacin sensitivity	R	R	R	R	R	R	R	R	R	R										
ciprofloxacin sensitivity	UN	UN	UN	UN	UN	UN	UN	UN	R	R										
levofloxacin MIC	32	32	32	32	8	32	32	32	8	8										
ST	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111
Genome ID	287.1	287.1046	287.1084	287.1093	287.1095	287.1098	287.1125	287.1126	287.1127	287.113	287.114	287.1141	287.1146	287.1169	287.1172	287.1178	287.1179	287.1185	287.1195	287.1209
serogroup	O12	O12	O12	O4	O4	O4	O12	O12	O12	O12	O12	O12	O12	O12	O12	O12	O12	O12	O12	O4
levofloxacin sensitivity	R	S	R	R	R	S	R	R	R	R	R	R	R	R	R	R	R	R	R	R
ciprofloxacin sensitivity	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN
levofloxacin MIC	32	1	8	2	8	0.25	32	32	16	32	32	32	32	8	8	32	32	16	8	32
ST	111	111	111	111	111	111	111	111	111	111										
Genome ID	287.1216	287.575	287.633	287.7772	287.7862	287.852	287.909	287.911	287.916	287.951										
serogroup	O12	O12	O12	O12	O4	O12	O12	O12	O12	O12										
levofloxacin sensitivity	R	R	R	R	R	R	R	R	R	R										
ciprofloxacin sensitivity	UN	R	R	R	S	UN	UN	UN	UN	UN										
levofloxacin MIC	32	8	8	4	2	2	16	16	32	8										
ciprofloxacin MIC	UN	8	8	2	0.5	UN	UN	UN	UN	UN										
ST	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244

Genome ID	PAE0056	PAE0064	PAE0066	PAE0069	PAE0121	PAE0143	PAE0173	287.1029	287.106	287.1065	287.1089	287.1096	287.1097	287.1208	287.7781	287.7814	287.897	287.905	287.956	287.975
serogroup	O2	O5	O5	O5	O5	O5	O5	O2	O12	O2	O2	O12	O12	O12	O12	O2	O12	O2	O12	O11
levofloxacin sensitivity	S	S	S	S	S	S	S	R	R	S	S	R	R	R	R	R	R	S	R	R
ciprofloxacin sensitivity	S	S	S	S	S	S	S	UN	UN	UN	UN	UN	UN	UN	R	R	UN	UN	UN	UN
levofloxacin MIC	0.25	1	0.5	0.5	0.12	1	0.12	32	32	0.015	0.25	32	32	32	4	4	32	0.25	16	8
ciprofloxacin MIC	0.06	0.12	0.08	0.03	0.12	0.06	0.06	UN	UN	UN	UN	UN	UN	UN	2	2	UN	UN	UN	UN
ST	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395			
Genome ID	PAE0093	PAE0015	PAE0027	PAE0054	287.1039	287.108	287.1106	287.1115	287.1116	287.5703	287.5749	287.5956	287.848	287.881	287.887	287.921	287.961			
serogroup	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6	O6			
levofloxacin sensitivity	S	S	S	S	S	S	R	S	S	R	R	R	S	S	S	R	S			
ciprofloxacin sensitivity	S	S	S	S	UN	UN	UN	UN	UN	R	R	R	UN	UN	UN	UN	UN			
levofloxacin MIC	0.25	0.25	1	0.5	0.12	0.5	2	0.5	1	2	2	8	0.5	1	0.25	4	0.25			
ciprofloxacin MIC	0.06	0.08	0.06	0.03	UN	UN	UN	UN	UN	1	1	1	UN	UN	UN	UN	UN			
ST	175	175	175	175	175	175	175	175	175	175										
Genome ID	PAE0016	PAE0025	287.1024	287.1042	287.1051	287.1052	287.1132	287.1188	287.1192	287.908										
serogroup	O4	O4	O4	O4	O4	O4	O4	O4	O4	O4										
levofloxacin sensitivity	R	S	R	R	R	R	R	R	R	R										
ciprofloxacin sensitivity	R	S	UN	UN	UN	UN	UN	UN	UN	UN										
levofloxacin MIC	32	0.5	32	32	32	32	32	8	8	32										
ciprofloxacin MIC	8	0.12	UN	UN	UN	UN	UN	UN	UN	UN										
ST	274	274	274	274	274	274	274	274	274	274	274	274	274	274						
Genome ID	PAE0045	PAE0103	PAE0159	PAE0169	287.1007	287.1019	287.1228	287.5973	287.7844	287.846	287.878	287.885	287.953	287.963						
serogroup	O3	O3	O3	O3	O3	O3	O3	O3	O3	O3	O3	O3	O3	O3						
levofloxacin sensitivity	S	S	S	S	R	R	S	R	R	S	R	S	S	S						
ciprofloxacin sensitivity	S	S	S	S	UN	UN	UN	S	R	UN	UN	UN	UN	UN						
levofloxacin MIC	0.5	0.5	1	0.25	32	32	0.25	2	4	1	8	0.03	0.5	0.12						
ciprofloxacin MIC	0.06	0.06	0.25	0.03	UN	UN	UN	0.5	2	UN	UN	UN	UN	UN						
ST	17	17	17	17	17	17	17	17	17	17	17	17								

Genome ID	PAE0067	PAE0128	287.1075	287.1129	287.124	287.7807	287.847	287.85	287.9	287.983	287.989	287.998
serogroup	O1	O1	O1	O1	O1	O1	O1	O1	O1	O1	O1	O1
levofloxacin sensitivity	R	S	S	R	R	R	R	S	R	R	S	S
ciprofloxacin sensitivity	S	S	UN	UN	UN	R	UN	UN	UN	UN	UN	UN
levofloxacin MIC	2	0.25	0.25	2	32	2	2	0.25	32	32	0.12	0.03
ciprofloxacin MIC	0.12	0.06	UN	UN	UN	1	UN	UN	UN	UN	UN	UN

Among the studied set of isolates, the international high-risk clone ST235 was the most frequently identified among all STs in the study set (50 isolates). Among those 50 isolates; 13 are ciprofloxacin resistant, 46 are levofloxacin resistant and only 4 isolates are levofloxacin susceptible. ST111 was the next most frequently observed ST in a total of 30 isolates with 28 levofloxacin resistant isolates and only 2 susceptible isolates, 3 ciprofloxacin resistant and 1 ciprofloxacin susceptible. ST244 was identified in 20 genomes; 10 are levofloxacin susceptible and 10 are levofloxacin resistant, 2 are ciprofloxacin resistant and 7 are ciprofloxacin susceptible. ST395 was identified in 17 genomes; 12 are levofloxacin susceptible and 5 are levofloxacin resistant, 4 are ciprofloxacin susceptible and 3 are ciprofloxacin resistant. ST175 was identified in 10 genomes; 9 are levofloxacin resistant and 1 is levofloxacin susceptible. The CF clone ST17 was identified in 12 isolates; 5 are levofloxacin susceptible and 7 are levofloxacin resistant, 1 is ciprofloxacin resistant and 2 are ciprofloxacin susceptible. Another CF clone ST274 was identified in 14 isolates; 9 are levofloxacin susceptible and 5 are levofloxacin resistant, 5 are ciprofloxacin susceptible and 1 is ciprofloxacin resistant.

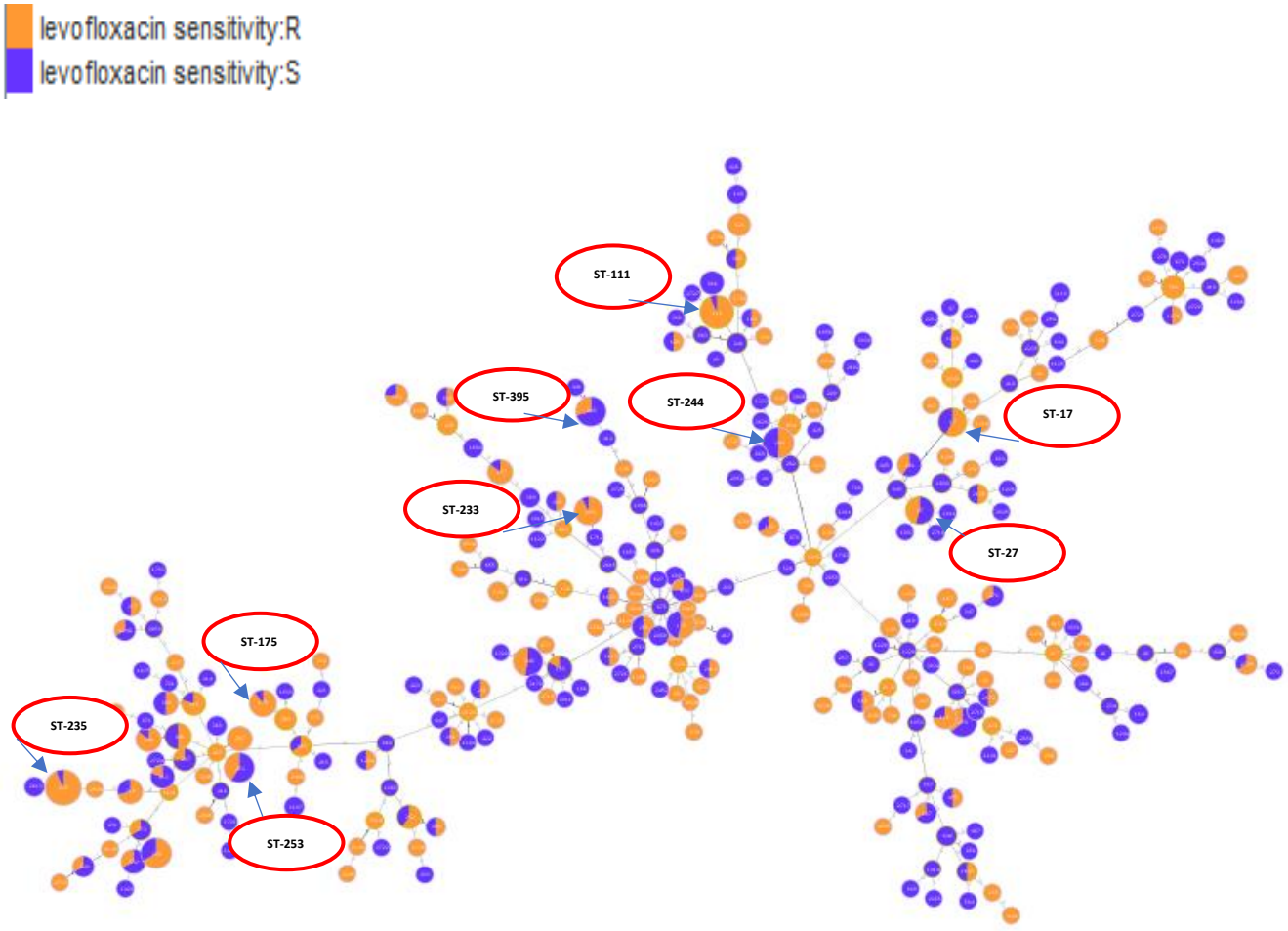


Figure 4.4. MST showing distribution of levofloxacin susceptibility in relation to STs

Orange color represents levofloxacin-resistant isolates while blue color represents levofloxacin-susceptible isolates.

The distribution of levofloxacin susceptibility and resistance in figure 4.4 does not show a specific distribution pattern for antibiotic susceptibility behavior in relation to different STs in general. This may indicate that ST classification does not overall correlate with levofloxacin susceptibility. This may be related to the inherent drawback of MLST typing method which gives random alleles designation and does not consider all variations resulting from mutations but reduces allele definition to exact matches. This may result in the assumption that MLST in general cannot be used to predict resistance or susceptibility except where a high association of specific high-risk clones with antibiotic resistance behavior is observed. Another explanation for the observed distribution is that the genomic bases differentiating resistance/susceptibility in general are global/multiple combined elements, are unpredictable or are subject to epistatic interactions and do not correlate with the MLST classification, the research question that needs more in-depth exploration.

Figure 4.4 also shows that ST-111, ST-235, ST-175, and ST-233 are majorly composed of resistant isolates. Numbers are described in detail above. This supports the literature findings about these high-risk clones. On the other hand, nearly half of the isolates classified under ST-253, ST-244, ST-17 and ST-27 were not resistant.

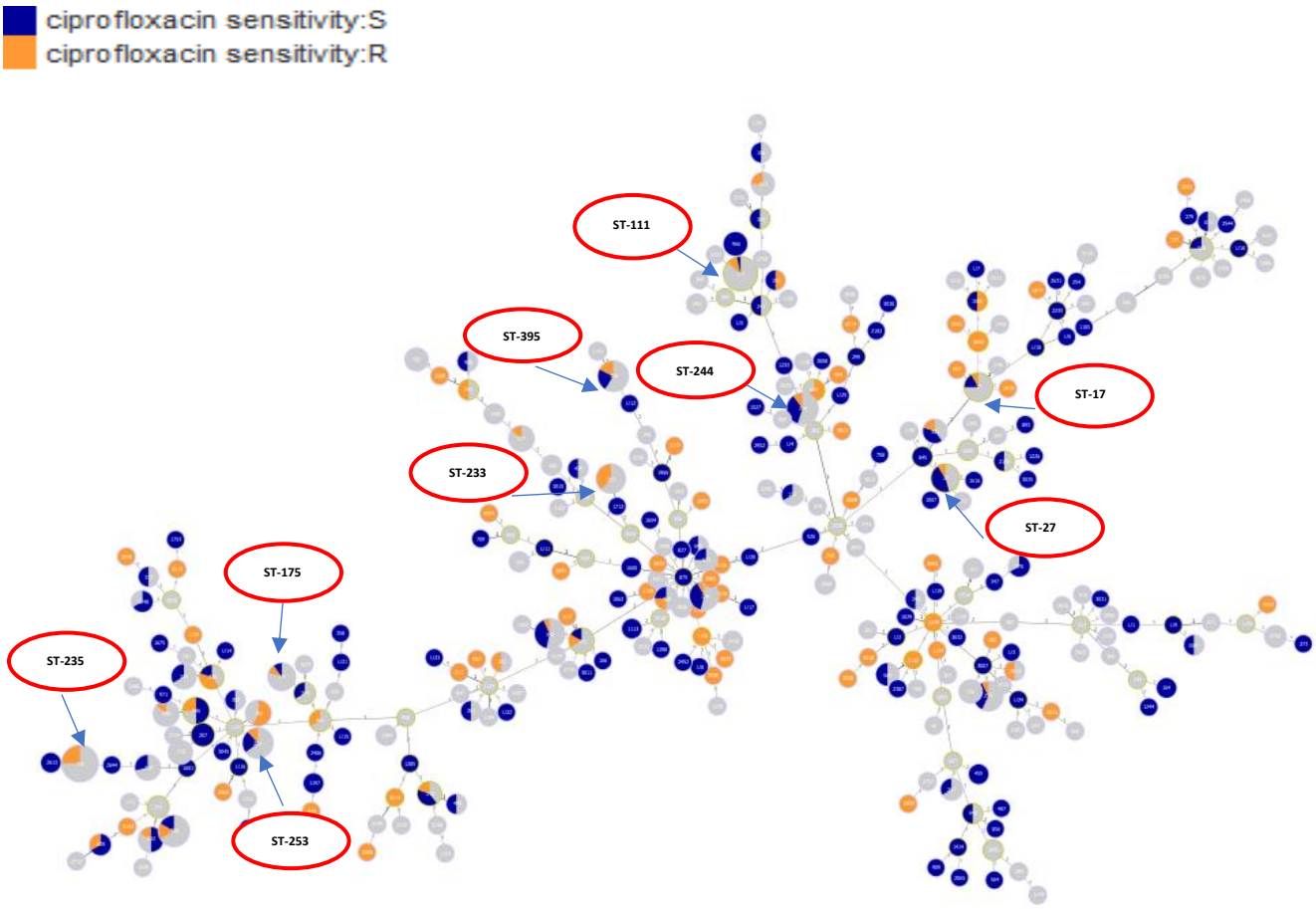


Figure 4.5. MST showing distribution of ciprofloxacin susceptibility in relation to STs

Ciprofloxacin-resistant isolates are shown in orange color while ciprofloxacin-susceptible isolates are shown in blue. Grey color represents isolates where ciprofloxacin susceptibility data are not available.

The distribution of ciprofloxacin susceptibility and resistance as seen in Figure 4.5 indicates that no specific pattern of distribution for antibiotic susceptibility behavior in relation to different STs is observed. This is similar to what was observed with levofloxacin in Figure 4.4. This may also indicate that the information drawn from MLST classification may not be sufficient to reflect a true association with an important behavior like antibiotic resistance in both levofloxacin and ciprofloxacin.

4.4.5. Aminoglycosides resistance profile of *Ps. aeruginosa* epidemic high-risk clones

Table 4.2, Figure 4.6 and Figure 4.7 all summarize aminoglycosides susceptibility data in relation to different STs and high-risk groups.

Table 4.2 summarizes the quinolone MIC values for the high-risk clones observed in the studied collection with corresponding serotypes. Figure 4.6 is an MST tree showing gentamycin

susceptibility in relation to different STs. Figure 4.7 is an MST tree showing amikacin susceptibility in relation to different STs.

Table 4.2. Summary of isolates belonging to high-risk clones with their corresponding serotypes and aminoglycosides susceptibility

ST	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235
Genome ID	287.1031	287.5686	287.1021	287.1045	287.1058	287.1087	287.1088	287.1100	287.1122	287.1124	287.1128	287.1131	287.1133	287.1134	287.1135	287.1144	287.1145	287.1147	287.1148	287.1150
serogroup	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
AK MIC	32	8	32	32	8	4	4	32	8	32	8	16	8	64	4	0.25	8	128	16	16
AK sensitivity	R	S	R	R	S	S	S	R	S	R	S	R	S	R	S	S	S	R	R	R
CN MIC	UN	16	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN
CN sensitivity	UN	R	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN	UN
ST	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235	235
Genome ID	287.1151	287.1158	287.1162	287.1163	287.1184	287.1200	287.1202	287.1203	287.1213	287.2973	287.2975	287.2976	287.5690	287.6327	287.6328	287.7780	287.7782	287.7783	287.7785	287.7786
serogroup	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
AK MIC	8	32	16	128	4	4	64	2	64	64	64	8	64	2	8	8	8	8	32	8
AK sensitivity	S	R	R	R	S	S	R	S	R	R	R	S	R	S	S	S	S	S	R	S
CN MIC	UN	UN	UN	UN	UN	UN	UN	UN	UN	16	16	16	16	0.5	4	16	16	16	16	8
CN sensitivity	UN	UN	UN	UN	UN	UN	UN	UN	UN	R	R	R	R	S	S	R	R	R	R	R
ST	235	235	235	235	235	235	235	235	235	235										
Genome ID	287.7830	287.855	287.860	287.902	287.903	287.915	287.918	287.955	287.966	287.984										
serogroup	11	11	11	10	11	11	11	11	11	11										
AK MIC	8	64	1	4	8	128	16	8	1	32										
AK sensitivity	S	R	S	S	S	R	R	S	S	R										
CN MIC	4	UN	UN	UN	UN	UN	UN	UN	UN	UN										
CN sensitivity	S	UN	UN	UN	UN	UN	UN	UN	UN	UN										
ST	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111
Genome ID	287.1000	287.1046	287.1084	287.1093	287.1095	287.1098	287.1125	287.1126	287.1127	287.1130	287.1140	287.1141	287.1146	287.1169	287.1172	287.1178	287.1179	287.1185	287.1195	287.1209

serogroup	12	12	12	4	4	4	12	12	12	12	12	12	12	12	12	12	12	12	12	12	4
AK MIC	32	4	8	2	2	1	128	8	4	4	128	128	128	8	2	128	128	8	64	128	
AK sensitivity	R	S	S	S	S	S	R	S	S	S	R	R	R	S	S	R	R	S	R	R	
ST	111	111	111	111	111	111	111	111	111	111											
Genome ID	287.1216	287.5750	287.6330	287.7772	287.7862	287.852	287.909	287.911	287.916	287.951											
serogroup	12	12	12	12	4	12	12	12	12	12											
AK MIC	128	64	32	8	8	4	1	1	64	16											
AK sensitivity	R	R	R	S	S	S	S	S	R	R											
CN MIC	UN	16	16	16	4	UN	UN	UN	UN	UN											
CN sensitivity	UN	R	R	R	S	UN	UN	UN	UN	UN											
ST	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	244	
Genome ID	PAE0056	PAE0064	PAE0066	PAE0069	PAE0121	PAE0143	PAE0173	287.1029	287.1060	287.1065	287.1089	287.1096	287.1097	287.1208	287.7781	287.7814	287.897	287.905	287.956	287.975	
serogroup	2	5	5	5	5	5	5	2	12	2	2	12	12	12	12	2	12	2	12	11	
AK MIC	0.5	2	1	1	1	4	8	2	8	0.5	1	32	64	32	32	8	64	1	4	64	
AK sensitivity	S	S	S	S	S	S	S	S	S	S	S	R	R	R	R	S	R	S	S	R	
CN MIC	1	1	2	1	2	4	UN	UN	UN	UN	UN	UN	UN	UN	16	16	UN	UN	UN	UN	
CN sensitivity	S	S	S	S	S	S	UN	UN	UN	UN	UN	UN	UN	UN	R	R	UN	UN	UN	UN	
ST	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395	395				
Genome ID	PAE0093	PAE0015	PAE0027	PAE0054	287.1039	287.1080	287.1106	287.1115	287.1116	287.5703	287.5749	287.5956	287.848	287.881	287.887	287.921	287.961				
serogroup	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6				
AK MIC	4	1	2	0.5	2	2	2	1	4	4	4	16	4	8	2	1	1				
AK sensitivity	S	S	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S				
CN MIC	4	2	4	1	UN	UN	UN	UN	UN	16	16	4	UN	UN	UN	UN	UN				
CN sensitivity	S	S	S	S	UN	UN	UN	UN	UN	R	R	S	UN	UN	UN	UN	UN				
ST	175	175	175	175	175	175	175	175	175	175											
Genome ID	PAE0025	PAE0016	287.1024	287.1042	287.1051	287.1052	287.1132	287.1188	287.1192	287.908											
serogroup	4	4	4	4	4	4	4	4	4	4											
AK MIC	2	2	4	4	4	4	2	4	4	2											

AK sensitivity	S	S	S	S	S	S	S	S	S	S				
CN MIC	4	64	UN	UN	UN	UN	UN	UN	UN	UN				
CN sensitivity	S	R	UN	UN	UN	UN	UN	UN	UN	UN				
ST	274	274	274	274	274	274	274	274	274	274	274	274	274	274
Genome ID	PAE0045	PAE0103	PAE0159	PAE0169	287.1007	287.1019	287.1228	287.5973	287.7844	287.846	287.878	287.885	287.953	287.963
serogroup	3	3	3	3	3	3	3	3	3	3	3	3	3	3
AK MIC	2	1	4	16	4	4	4	16	8	8	4	0.25	4	0.12
AK sensitivity	S	S	S	R	S	S	S	R	S	S	S	S	S	S
CN MIC	2	4	8	16	UN	UN	UN	8	4	UN	UN	UN	UN	UN
CN sensitivity	S	S	R	R	UN	UN	UN	R	S	UN	UN	UN	UN	UN
ST	17	17	17	17	17	17	17	17	17	17	17	17		
Genome ID	PAE0067	PAE0128	287.1075	287.1129	287.1240	287.7807	287.847	287.850	287.900	287.983	287.989	287.998		
serogroup	1	1	1	1	1	1	1	1	1	1	1	1		
AK MIC	2	2	2	16	32	8	4	2	2	2	0.12	4		
AK sensitivity	S	S	S	R	R	S	S	S	S	S	S	S		
CN MIC	1	4	UN	UN	UN	4	UN	UN	UN	UN	UN	UN		
CN sensitivity	S	S	UN	UN	UN	S	UN	UN	UN	UN	UN	UN		

UN= Unknown

For aminoglycosides; 23 isolates with ST235 are amikacin resistant and 27 are amikacin susceptible, 3 are gentamycin resistant and 10 are gentamycin susceptible. ST111 included a total of 14 amikacin resistant isolates and 16 amikacin susceptible isolates. Isolates belonging to ST244 included 6 amikacin resistant isolates, 14 amikacin susceptible isolates, 2 gentamycin resistant isolates and 5 gentamycin susceptible isolates. ST395 included 1 amikacin resistant, 16 amikacin susceptible, 2 gentamycin resistant and 5 gentamycin susceptible. All isolates with ST175 were amikacin susceptible, 1 isolate was gentamycin susceptible and 1 was gentamycin resistant. Isolates with ST17 included 10 amikacin susceptible isolates, 2 amikacin resistant isolates and 3 gentamycin susceptible isolates. Isolates with ST274 included 12 amikacin susceptible isolates, 2 amikacin resistant isolates, 3 gentamycin resistant isolates and 3 gentamycin susceptible isolates.

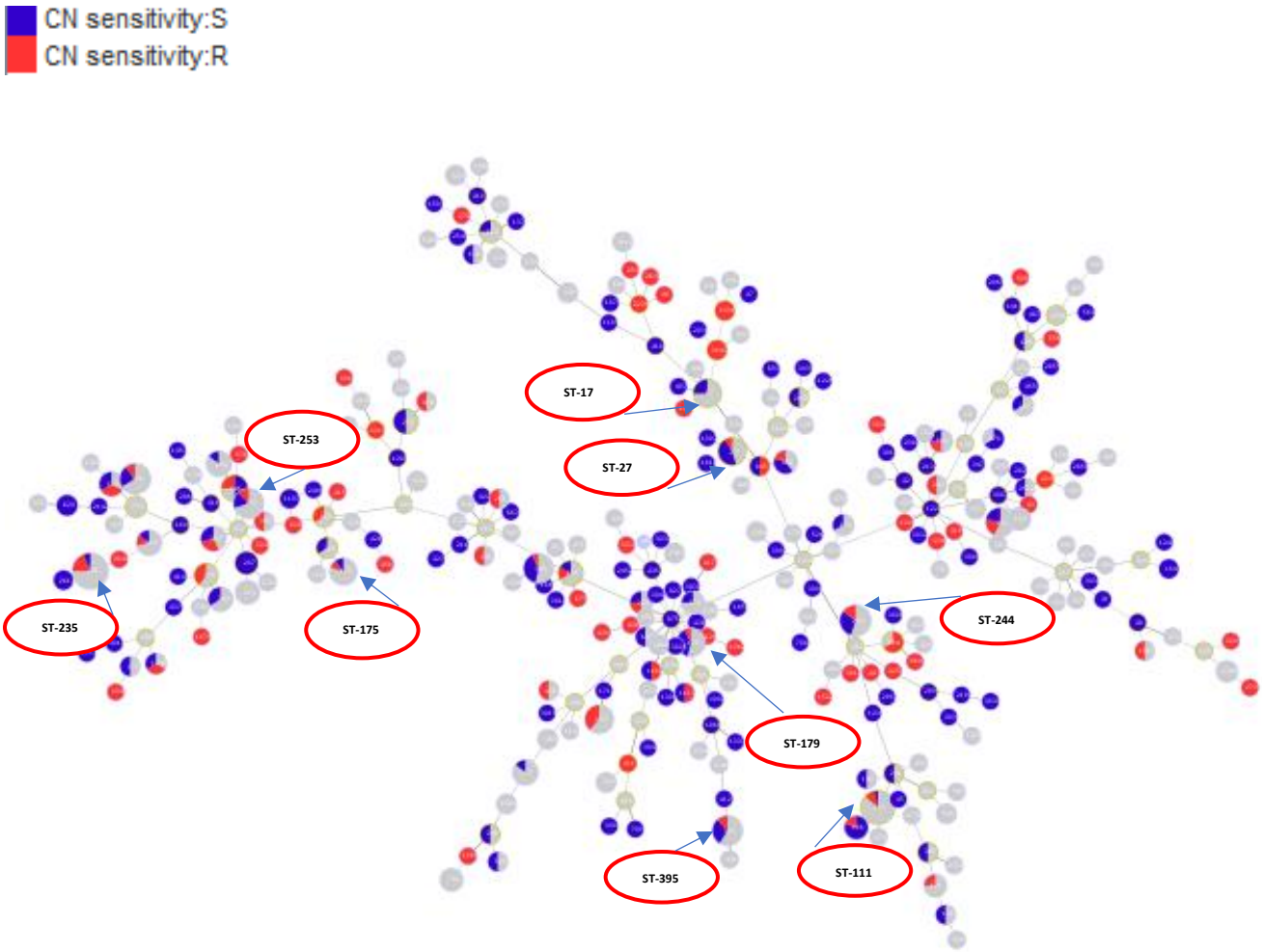


Figure 4.6. MST showing distribution of gentamycin susceptibility in relation to STs

Red color represents gentamycin-resistant isolates while blue color represents gentamycin-susceptible isolates. Grey color represents isolates where gentamycin susceptibility data are not available.

The distribution of gentamycin susceptibility and resistance in Figure 4.6 shows that both gentamycin resistant and gentamycin susceptible isolates are evenly distributed all over the MST tree. This may indicate that isolates relate to STs at different distances and are not clustering in clones or in relation to specific clonal complexes. This is also similar to what was observed with quinolones in Figure 4.4 and Figure 4.5. This may also indicate that the information drawn from MLST classification in general may not be sufficient to reflect a true association with *antibiotic resistance behavior. On the other hand, high-risk clones shown on the graph in Figure 4.6 include both resistant and susceptible isolates. This may not support the conclusion of high-risk clones' dissemination in relation to carriage of horizontally acquired resistance genes. Other markers related to increased fitness/virulence may underlie the successful dissemination of high-risk clones. This is based on the current observation of high-risk clones among both resistant and susceptible isolates.*

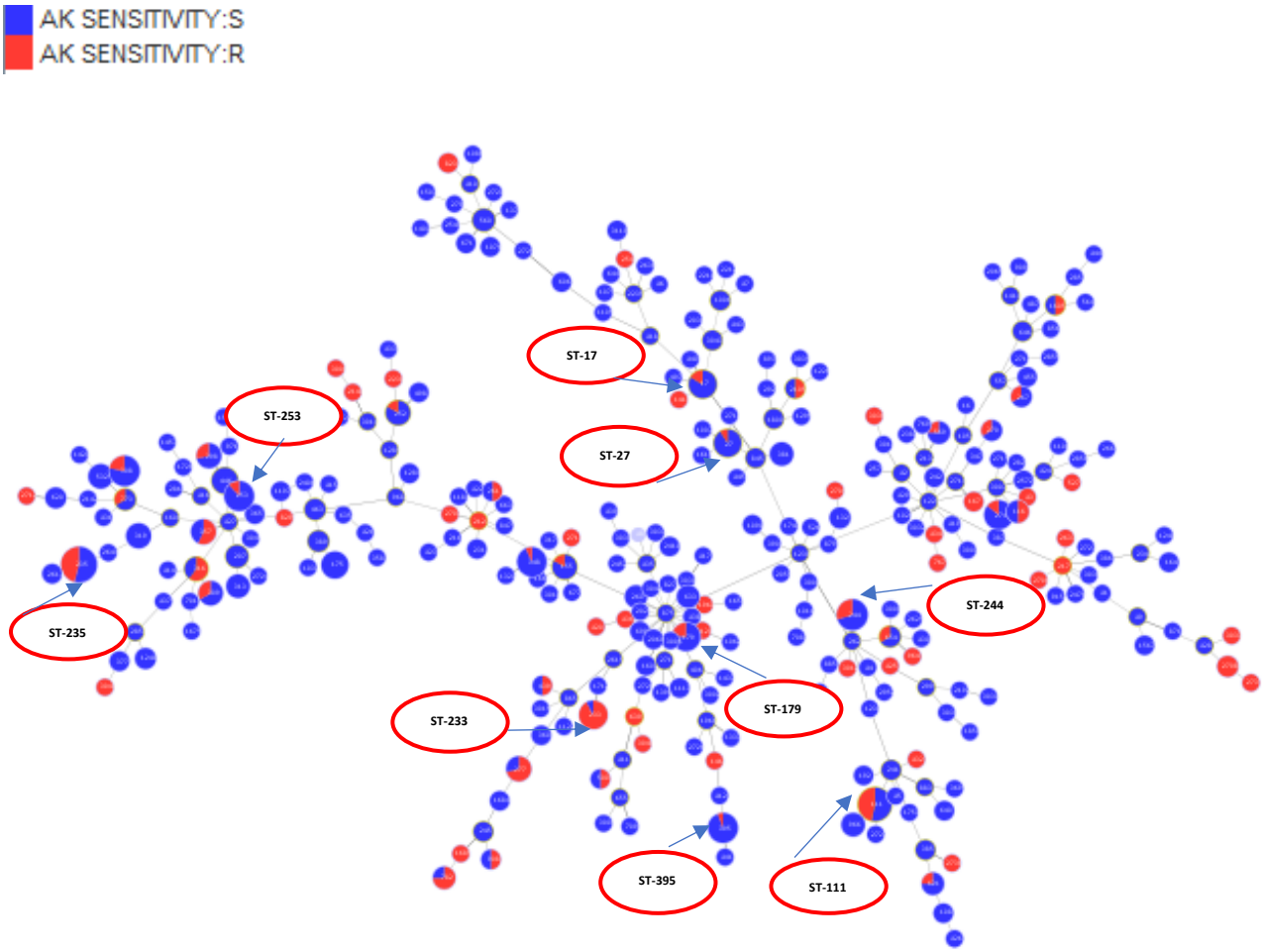


Figure 4.7. MST showing the distribution of amikacin susceptibility in relation to STs

Amikacin-resistant isolates are represented in red color while amikacin-susceptible isolates are represented in blue.

Based on the distribution of high-risk clones seen in Figure 4.7, amikacin resistance does not appear to correlate with high-risk clones. Except for ST-233, the majority of isolates forming all other high-risk STs were mostly susceptible. About half of the isolates forming each of ST-111 and ST-235 are resistant. For ST-111, 14 isolates are amikacin resistant and 16 are amikacin susceptible. For ST-235, 23 isolates are amikacin resistant and 27 are amikacin susceptible.

4.4.6. Quinolone resistance markers of *Ps. aeruginosa* epidemic high-risk clones

The distribution of previously identified quinolone resistance or susceptibility markers (in Chapter 3) and new clusters (combinations) of molecular markers is re-analyzed in this section in relation to high-risk clones. Table 4.3 and Table 4.5 summarize this distribution.

A total of 305 isolates were analyzed for ciprofloxacin sensitivity (110 resistant and 195 susceptible). Of those analyzed, a total of 82 isolates belonged to epidemic high risk clones (ST 17 [N=3], ST 27 [N=6], ST 111[N=4], ST 179 [N=6], ST 233 [N=5], ST 235 [N=13], ST 244 [N=9], ST 253 [N=6], ST 274 [N=6], ST 308 [N=6], ST 446 [N=6], ST 395 [N=7], ST 532 [N=4]). A chi-square test for independence (with yates' continuity correction) indicated significant association between ***arm*R gene** and high-risk groups, x^2 (1, n= 305) =4, *p*=0.046, *phi*= 0.114. Significant associations were also found between each of ***nal*CG71E**, x^2 (1, n= 305) =7.031, *p*=0.030, *phi*= 0.152, ***gyr*AT83I**, x^2 (1, n= 305) =8.4, *p*=0.004, *phi*= 0.166, ***nal*CE153Q**, x^2 (1, n= 305) =6.09, *p*=0.048, *phi*= 0.141 and high risk groups. Cluster 1 previously identified showed significant absence in high-risk groups, x^2 (1, n= 305) =4.066, *p*=0.044, *phi*= - 0.116. Cluster 4 also showed significant association with high-risk groups, x^2 (1, n= 305) =6.27, *p*=0.012, *phi*= - 0.144. Significant associations between individual markers for ciprofloxacin and specific high-risk clones are summarized in Table 4.4. The detailed distribution of ciprofloxacin markers in relation to high-risk clones is shown in Table 4.3.

Table 4.3. Distribution of ciprofloxacin molecular markers in relation to high-risk clones

Genome ID	ST	serogroup	Cipro MIC	Cipro sensitivity	*arm*R (PA3719)	*nal*C S46A	*nal*C G71E	*nal*C S209R	*gyr*A T83I	*nal*C E153Q	*nal*C Thr50pro	*nal*C gene	*mex*S gene	*nal*D gene	Cipro cluster 1	Cipro cluster 4
PAE0067	**17**	O1	0.12	S												
PAE0128	**17**	O1	0.06	S												
287.7807	**17**	O1	1	R												
PAE0029	**27**	O1	0.12	S												
PAE0009	**27**	O1	0.08	S												
PAE0050	**27**	O1	0.25	S												
PAE0086	**27**	O1	0.06	S												
PAE0108	**27**	O1	0.06	S												
287.7773	**27**	O1	2	R												
287.5750	**111**	O12	8	R												
287.6330	**111**	O12	8	R												
287.7772	**111**	O12	2	R												
287.7862	**111**	O4	0.5	S												
PAE0025	**175**	O4	0.12	S												
PAE0016	**175**	O4	8	R												
PAE0092	**179**	O6	0.12	S												

Genome ID	ST	serogroup	Cipro MIC	Cipro sensitivity	*armR* (PA3719)	*nalC* S46A	*nalC* G71E	*nalC* S209R	*gyrA* T83I	*nalC* E153Q	*nalC* Thr50pro	*nalC* gene	*mexS* gene	*nalD* gene	Cipro cluster 1	Cipro cluster 4
PAE0042	**179**	O6	0.03	S												
PAE0141	**179**	O6	0.12	S												
287.5748	**179**	O6	8	R												
287.5969	**179**	O6	0.25	S												
287.2980	**233**	O6	8	R												
287.5688	**233**	O6	8	R												
287.5752	**233**	O6	8	R												
287.5959	**233**	O6	8	R												
287.6492	**233**	O6	8	R												
287.2973	**235**	O11	8	R												
287.2975	**235**	O11	8	R												
287.2976	**235**	O11	8	R												
287.5690	**235**	O11	8	R												
287.6327	**235**	O11	8	R												
287.6328	**235**	O11	8	R												
287.7780	**235**	O11	2	R												
287.7782	**235**	O11	2	R												
287.7783	**235**	O11	2	R												
287.7785	**235**	O11	2	R												
287.7786	**235**	O11	2	R												
287.7830	**235**	O11	2	R												
287.5686	**235**	O11	4	R												
PAE0056	**244**	O2	0.06	S												
PAE0064	**244**	O5	0.12	S												
PAE0066	**244**	O5	0.08	S												
PAE0069	**244**	O5	0.03	S												
PAE0121	**244**	O5	0.12	S												
PAE0143	**244**	O5	0.06	S												
PAE0173	**244**	O5	0.06	S												
287.7781	**244**	O12	2	R												
287.7814	**244**	O2	2	R												
PAE0055	**253**	O10	0.25	S												
PAE0102	**253**	O10	0.06	S												
PAE0105	**253**	O10	0.12	S												
PAE0122	**253**	O10	0.12	S												
287.7843	**253**	O10	1	R												
287.7850	**253**	O10	2	R												
PAE0045	**274**	O3	0.06	S												
PAE0103	**274**	O3	0.06	S												
PAE0159	**274**	O3	0.25	S												
PAE0169	**274**	O3	0.03	S												
287.5973	**274**	O3	0.5	S												
287.7844	**274**	O3	2	R												
PAE0093	**395**	O6	0.06	S												
PAE0015	**395**	O6	0.08	S												
PAE0027	**395**	O6	0.06	S												
PAE0054	**395**	O6	0.03	S												
287.5703	**395**	O6	1	R												
287.5749	**395**	O6	1	R												
287.5956	**395**	O6	1	R												
PAE0111	**446**	O11	2	R												
PAE0127	**446**	O11	0.12	S												
PAE0139	**446**	O11	0.03	S												
PAE0150	**446**	O11	0.12	S												
PAE0166	**446**	O11	0.25	S												
287.7795	**446**	O11	2	R												

Grey shade: Marker (Gene or Gene variant) present, White: Marker (Gene or Gene variant absent), Yellow shade: Gene absent

Table 4.4. Ciprofloxacin molecular markers showing significant associations with specific high-risk clones

Molecular marker	High-risk group	Chi-square	*p*-value	Phi-coefficient
*arm*R	ST-235	6.22	0.013	0.143
***nal*CS46A**	ST-235	24.5	<0.005	0.283
***nal*CG71E**	ST-235	17.1	<0.005	0.237
***nal*CG71E**	ST-244	99.5	<0.005	0.571
***nal*CS209R**	ST-111	8.2	0.017	0.164
***nal*CS209R**	ST-233	10.5	0.005	0.185
***nal*CS209R**	ST-235	17.14	<0.005	0.237
***nal*CS209R**	ST-244	28	<0.005	0.303
***nal*CS209R**	ST-395	21.6	<0.005	0.266
***gyr*AT83I**	ST-233	30.6	<0.005	0.317
***gyr*AT83I**	ST-235	22.26	<0.005	0.27
***gyr*AT83I**	ST-308	4.28	0.039	0.118
***nal*CE153Q**	ST-235	109	<0.005	0.6
*nal*CThr50pro	ST-179	24.5	<0.005	0.3
*nal*CThr50pro	ST-235	16.6	<0.005	0.233
Cluster 1	ST-235	9	0.003	-0.172
Cluster 4	ST-233	20.5	<0.005	0.26
Cluster 4	ST-235	8	0.005	0.162
Cluster 4	ST-395	5.8	0.016	0.138

The tables and analyses presented above show that there is significant association of some of the studied ciprofloxacin resistance markers previously identified in *Chapter 3* with high-risk clones. The markers showing association with all high-risk clones evaluated included *arm*R (especially with ST-235), *nal*CG71E (especially with ST-235 and ST-244), *nal*CE153Q (especially with ST-235), and *gyr*AT83I (especially with ST-235, ST-233 and ST-308).

Ciprofloxacin resistance markers that showed association with specific high-risk groups included: *nal*CS46A (showing association with ST-235), *nal*CS209R (showing association with ST-235, ST-233, ST-395 and ST-244), *nal*CThr50Pro (showing association with ST-179 and ST-235). Ciprofloxacin Cluster 1 showing association with ciprofloxacin susceptibility (*Chapter 3*) showed significant absence in high-risk groups especially with ST-235. On the other hand, ciprofloxacin Cluster 4 showing association with ciprofloxacin resistance (*Chapter 3*) showed significant presence in high-risk clones especially with ST-235, ST-233, and ST-395. This may add additional evidence to the importance of these clusters to detect resistance behavior in

addition to being related to risky clones. This consequently means than identifying such clusters can guide specific infection control procedures.

A total of 691 isolates were analyzed for levofloxacin sensitivity (376 resistant and 315 susceptible). Of those analyzed, a total of 239 isolates belonged to epidemic high-risk clones (ST 17 [N=12], ST 27 [N=11], ST 111 [N=30], ST 235 [N=50], ST 175 [N=10], ST 179 [N=13], ST 244 [N=20], ST 233 [N=13], ST 308 [N=18], ST 395 [N=17], ST 532 [N=8], ST 446 [N=6], ST 274 [N=14], ST 253 [N=17]).

A chi-square test for independence (with yates continuity correction) indicated significant association between ***mex*Z gene**, x^2 (1, n= 691) = 5.4, *p*=0.020, *phi*= 0.088, ***arm* R**, x^2 (1, n= 691) = 19.3, *p*<0.005, *phi*= 0.167, ***nfx*B**, x^2 (1, n= 691) = 5.8, *p*=0.016, *phi*= 0.092, ***mex*S**, x^2 (1, n= 691) = 5.33, *p*=0.021, *phi*= 0.088, ***nal*C**, x^2 (1, n= 691) = 4.88, *p*=0.027, *phi*= 0.084, ***gyr*AT83I**, x^2 (1, n= 691) = 70.36, *p*<0.005, *phi*= 0.319, ***nal*CE153Q**, x^2 (1, n= 691) = 8.98, *p*=0.011, *phi*= 0.114, ***nal*CS46A,** x^2 (1, n= 691) = 14.3, *p*=0.001, *phi*= 0.144, and high-risk groups. Both previously identified clusters 1 and 5 were significantly absent in high risk groups with x^2 (1, n= 691) = 5.22, *p*=0.022, *phi*= - 0.087 and x^2 (1, n= 691) = 10.8, *p*=0.001, *phi*= - 0.125 respectively. Previously identified cluster 4 showed significant association with high risk groups, x^2 (1, n= 691) = 34.33, *p*<0.005, *phi*= 0.223. Significant associations between individual markers for levofloxacin and specific high-risk clones are summarized in Table 4.6. The detailed distribution of different markers is shown in Table 4.5.

Table 4.5. Distribution of levofloxacin molecular markers in relation to high-risk clones

Genome ID	ST	serogroup	Levo MIC	Levo sensitivity	*mex*Z	*arm*R (PA3719)	*nfx*B gene	*mex*S gene	*nal* C gene	*mex*R gene	*gyr*A T83I	*nal*C E153Q	*par*E V460G	*amp*R D135N	*nal*C S46A	Levo cluster 1	Levo cluster 4	Levo cluster 5
PAE0067	**17**	O1	2	R														
PAE0128	**17**	O1	0.25	S														
287.1075	**17**	O1	0.25	S														
287.1129	**17**	O1	2	R														
287.1240	**17**	O1	32	R														
287.7807	**17**	O1	2	R														
287.847	**17**	O1	2	R														
287.850	**17**	O1	0.25	S														
287.900	**17**	O1	32	R														
287.983	**17**	O1	32	R														
287.989	**17**	O1	0.12	S														
287.998	**17**	O1	0.03	S														
PAE0067	**17**	O1	2	R														

Genome ID	ST	serogroup	Levo MIC	Levo sensitivity	*mexZ*	*armR* (PA3719)	*nfxB* gene	*mexS* gene	*nal* C gene	*mexR* gene	*gyrA* T83I	*nalC* E153Q	*parE* V460G	*ampR* D135N	*nalC* S46A	Levo cluster 1	Levo cluster 4	Levo cluster 5
PAE0128	17	O1	0.25	S														
287.1075	17	O1	0.25	S														
287.1129	17	O1	2	R														
287.1240	17	O1	32	R														
287.7807	17	O1	2	R														
PAE0029	27	O1	0.5	S														
PAE0009	27	O1	0.25	S														
PAE0050	27	O1	4	R														
PAE0086	27	O1	0.25	S														
PAE0108	27	O1	0.25	S														
287.1086	27	O1	0.25	S														
287.1241	27	O1	4	R														
287.1301	27	O1	1	S														
287.7773	27	O1	4	R														
287.954	27	O1	32	R														
287.987	27	O1	32	R														
287.1000	111	O12	32	R														
287.1046	111	O12	1	S														
287.1084	111	O12	8	R														
287.1093	111	O4	2	R														
287.1095	111	O4	8	R														
287.1098	111	O4	0.25	S														
287.1125	111	O12	32	R														
287.1126	111	O12	32	R														
287.1127	111	O12	16	R														
287.1130	111	O12	32	R														
287.1140	111	O12	32	R														
287.1141	111	O12	32	R														
287.1146	111	O12	32	R														
287.1169	111	O12	8	R														
287.1172	111	O12	8	R														
287.1178	111	O12	32	R														
287.1179	111	O12	32	R														
287.1185	111	O12	16	R														
287.1195	111	O12	8	R														
287.1209	111	O4	32	R														
287.1216	111	O12	32	R														
287.5750	111	O12	8	R														
287.6330	111	O12	8	R														
287.7772	111	O12	4	R														
287.7862	111	O4	2	R														
287.852	111	O12	2	R														
287.909	111	O12	16	R														
287.911	111	O12	16	R														
287.916	111	O12	32	R														
287.951	111	O12	8	R														
PAE0025	175	O4	0.5	S														
PAE0016	175	O4	32	R														
287.1024	175	O4	32	R														
287.1042	175	O4	32	R														
287.1051	175	O4	32	R														
287.1052	175	O4	32	R														
287.1132	175	O4	32	R														
287.1188	175	O4	8	R														
287.1192	175	O4	8	R														
287.908	175	O4	32	R														
PAE0025	175	O4	0.5	S														
PAE0016	175	O4	32	R														

Genome ID	ST	serogroup	Levo MIC	Levo sensitivity	*mexZ*	*armR* (PA3719)	*nfxB* gene	*mexS* gene	*nal* C gene	*mexR* gene	*gyrA* T83I	*nalC* E153Q	*parE* V460G	*ampR* D135N	*nalC* S46A	Levo cluster 1	Levo cluster 4	Levo cluster 5
287.1024	**175**	O4	32	R	■	■	■	■	■	■	■				■		■	
287.1042	**175**	O4	32	R	■	■	■	■	■	■	■				■		■	
287.1051	**175**	O4	32	R	■	■	■	■	■	■	■		■					
287.1052	**175**	O4	32	R	■	■	■	■	■	■	■				■		■	
PAE0092	**179**	O6	1	S	■		■	■	■	■						■		
PAE0042	**179**	O6	0.5	S	■	■	■	■	■	■						■		
PAE0141	**179**	O6	2	R	■		■	■	■	■								
287.1103	**179**	O6	32	R	■	■	■	■	■	■	■				■		■	
287.1139	**179**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1143	**179**	O6	32	R	■	■	■	■	■	■	■				■		■	
287.1167	**179**	O6	2	R	■	■	■	■	■	■					■		■	
287.1238	**179**	O6	8	R	■	■	■	■	■	■					■		■	
287.5748	**179**	O6	8	R	■	■	■	■	■	■	■				■		■	
287.5969	**179**	O6	0.5	S	■	■	■	■	■	■					■			■
287.8029	**179**	O6	1	S	■	■	■	■	■	■					■			■
287.980	**179**	O6	0.12	S	■	■	■	■	■	■					■			■
287.988	**179**	O6	0.5	S	■	■	■	■	■	■					■			■
287.1021	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1045	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1058	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1087	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1088	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1100	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1122	**235**	O11	0.25	S	■	■	■	■	■	■					■			■
287.1124	**235**	O11	8	R	■	■	■	■	■	■	■				■		■	
287.1128	**235**	O11	32	R	■	■	■	■	■	■	■		■		■			
287.1131	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1133	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1134	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1135	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1144	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1145	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1147	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1148	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1150	**235**	O11	32	R	■	■	■	■	■	■					■		■	
287.1151	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1158	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1162	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1163	**235**	O11	32	R	■	■	■	■	■		■				■			
287.1184	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1200	**235**	O11	8	R	■	■	■	■	■	■	■				■		■	
287.1202	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1203	**235**	O11	32	R	■	■	■	■	■	■	■				■		■	
287.1213	**235**	O11	32	R		■	■		■	■					■			
287.2973	**235**	O11	8	R	■	■	■	■	■	■	■	■						
287.2975	**235**	O11	8	R	■	■	■	■	■	■	■	■						
287.2976	**235**	O11	8	R	■	■		■	■	■	■	■						
287.5690	**235**	O11	8	R	■	■	■	■	■	■	■	■			■			
287.6327	**235**	O11	8	R	■	■	■	■	■	■	■	■			■			
287.6328	**235**	O11	8	R	■	■	■	■	■	■	■	■						
287.7780	**235**	O11	4	R	■	■	■	■	■	■		■			■			
287.7782	**235**	O11	4	R	■	■	■	■	■	■		■			■			
287.7783	**235**	O11	4	R	■	■	■	■	■	■		■			■			
287.7785	**235**	O11	4	R	■	■	■	■	■	■		■			■			
287.7786	**235**	O11	4	R	■		■	■	■	■								
287.7830	**235**	O11	4	R	■	■	■	■	■	■		■			■			
287.855	**235**	O11	128	R	■	■	■	■	■	■	■				■		■	
287.860	**235**	O11	0.015	S	■	■	■	■	■	■				■	■			■

Genome ID	ST	serogroup	Levo MIC	Levo sensitivity	*mexZ*	*armR* (PA3719)	*nfxB* gene	*mexS* gene	*nal* C gene	*mexR* gene	*gyrA* T83I	*nalC* E153Q	*parE* V460G	*ampR* D135N	*nalC* S46A	Levo cluster 1	Levo cluster 4	Levo cluster 5
287.902	**235**	O10	4	R														
287.903	**235**	O11	32	R														
287.915	**235**	O11	4	R														
287.918	**235**	O11	32	R														
287.955	**235**	O11	32	R														
287.966	**235**	O11	32	R														
287.984	**235**	O11	0.12	S														
287.1031	**235**	O11	0.06	S														
287.5686	**235**	O11	8	R														
PAE0056	**244**	O2	0.25	S														
PAE0064	**244**	O5	1	S														
PAE0066	**244**	O5	0.5	S														
PAE0069	**244**	O5	0.5	S														
PAE0121	**244**	O5	0.12	S														
PAE0143	**244**	O5	1	S														
PAE0173	**244**	O5	0.12	S														
287.1029	**244**	O2	32	R														
287.1060	**244**	O12	32	R														
287.1065	**244**	O2	0.015	S														
287.1089	**244**	O2	0.25	S														
287.1096	**244**	O12	32	R														
287.1097	**244**	O12	32	R														
287.1208	**244**	O12	32	R														
287.7781	**244**	O12	4	R														
287.7814	**244**	O2	4	R														
287.897	**244**	O12	32	R														
287.905	**244**	O2	0.25	S														
287.956	**244**	O12	16	R														
287.975	**244**	O11	8	R														
PAE0055	**253**	O10	1	S														
PAE0102	**253**	O10	0.25	S														
PAE0105	**253**	O10	0.5	S														
PAE0122	**253**	O10	1	S														
287.1035	**253**	O10	0.5	S														
287.1081	**253**	O10	2	R														
287.1085	**253**	O10	1	S														
287.1154	**253**	O10	2	R														
287.1157	**253**	O10	2	R														
287.7843	**253**	O10	2	R														
287.7850	**253**	O10	4	R														
287.873	**253**	O10	1	S														
287.893	**253**	O10	0.5	S														
287.894	**253**	O10	0.25	S														
287.907	**253**	O10	2	R														
287.912	**253**	O11	32	R														
287.985	**253**	O10	1	S														
PAE0045	**274**	O3	0.5	S														
PAE0103	**274**	O3	0.5	S														
PAE0159	**274**	O3	1	S														
PAE0169	**274**	O3	0.25	S														
287.1007	**274**	O3	32	R														
287.1019	**274**	O3	32	R														
287.1228	**274**	O3	0.25	S														
287.5973	**274**	O3	2	R														
287.7844	**274**	O3	4	R														
287.846	**274**	O3	1	S														
287.878	**274**	O3	8	R														
287.885	**274**	O3	0.03	S														

Genome ID	ST	serogroup	Levo MIC	Levo sensitivity	*mexZ*	*armR* (PA3719)	*nfxB* gene	*mexS* gene	*nal* C gene	*mexR* gene	*gyrA* T83I	*nalC* E153Q	*parE* V460G	*ampR* D135N	*nalC* S46A	Levo cluster 1	Levo cluster 4	Levo cluster 5
287.953	**274**	O3	0.5	S														
287.963	**274**	O3	0.12	S														
PAE0093	**395**	O6	0.25	S														
PAE0015	**395**	O6	0.25	S														
PAE0027	**395**	O6	1	S														
PAE0054	**395**	O6	0.5	S														
287.1039	**395**	O6	0.12	S														
287.1080	**395**	O6	0.5	S														
287.1106	**395**	O6	2	R														
287.1115	**395**	O6	0.5	S														
287.1116	**395**	O6	1	S														
287.5703	**395**	O6	2	R														
287.5749	**395**	O6	2	R														
287.5956	**395**	O6	8	R														
287.848	**395**	O6	0.5	S														
287.881	**395**	O6	1	S														
287.887	**395**	O6	0.25	S														
287.921	**395**	O6	4	R														
287.961	**395**	O6	0.25	S														
PAE0111	**446**	O11	16	R														
PAE0127	**446**	O11	0.5	S														
PAE0139	**446**	O11	1	S														
PAE0150	**446**	O11	1	S														
PAE0166	**446**	O11	4	R														
287.1205	**446**	O11	2	R														
287.7795	**446**	O11	4	R														
287.875	**446**	O11	1	S														

Grey shade: Marker (Gene or Gene variant) present, White: Marker (Gene or Gene variant) absent, Yellow shade: Gene absent

Same sequence types are shown in same color

Table 4.6. Levofloxacin molecular markers showing significant associations with specific high-risk clones

Molecular marker	High-risk group	Chi-square	*p*-value	Phi co-efficient
*arm*R	ST-111	9.14	0.002	0.115
*arm*R	ST-235	8.5	0.004	0.111
*arm*R	ST-233	3.9	0.049	0.075
***gyr*AT83I**	ST-111	40.192	<0.005	0.241
***gyr*AT83I**	ST-235	56.8	<0.005	0.287
***gyr*AT83I**	ST-175	20.9	<0.005	0.174
***gyr*AT83I**	ST-233	29.3	<0.005	0.206
*gyr*AT83I	ST-308	11.25	0.001	0.128
***nal*CE153Q**	ST-235	80.4	<0.005	0.341
***Par*EV460G**	ST-175	15.7	<0.005	0.150
*nal*CS46A	ST-27	7.5	0.024	0.104
*nal*CS46A	ST-111	7	0.031	0.1
***nal*CS46A**	ST-235	19.1	<0.005	0.166
Cluster 1	ST-111	9.6	0.002	-0.118

Cluster 1	ST-235	16.5	<0.005	-0.155
Cluster 1	ST-233	4.07	0.044	-0.077
Cluster 4	ST-111	18.6	<0.005	0.164
Cluster 4	ST-235	15.7	<0.005	0.151
Cluster 4	ST-308	5.203	0.023	0.087
Cluster 5	ST-111	4.2	0.041	-0.078
Cluster 5	ST-235	6	0.014	-0.093

Data and analyses presented above in Table 4.5 and Table 4.6 in relation to levofloxacin markers shows that *mex*Z, *arm*R, *nfx*B, *mex*S and *nal*C are all significantly associated with high risk clones. In addition, *gyr*AT83I showed significant association with high-risk clones (and also individually with each of ST-111, ST-235, ST-175, ST-233, ST-308 and ST-395). Similar to markers evaluated with ciprofloxacin, each of *nal*CE153Q and *nal*CS46A also showed significant association with high-risk clones especially with ST-235 for *nal*CE153Q and with ST-111, ST-27 and ST-235 for *nal*CS46A. The variant *par*EV460G showed significant association with ST-175. Both levofloxacin Cluster 1 and levofloxacin Cluster 5 linked to levofloxacin susceptibility (*Chapter 3*) showed to be significantly absent in high-risk groups, in addition, levofloxacin cluster 4 linked to levofloxacin resistance (*Chapter 3*) showed significant association with high-risk groups.

Figure 4.8 summarizes the differential distribution of tested quinolone molecular markers in high-risk vs non high-risk clones in the whole collection of studied isolates.

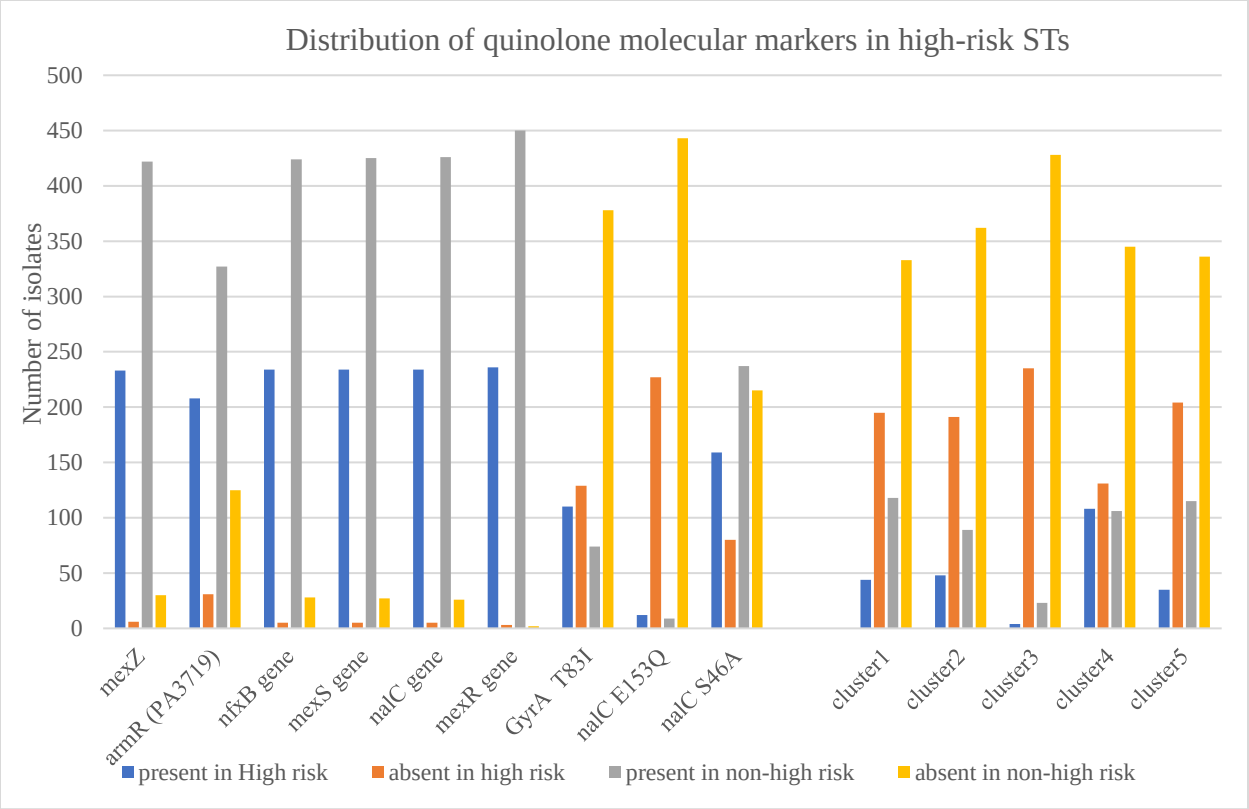


Figure 4.8. Differential distribution of tested quinolone molecular markers in high-risk vs non high-risk groups

The figure shows the distribution of variants that show significant absence/presence in high-risk groups. The distribution of variants among the four groups is represented in four colors. The difference in the distribution reflects the magnitude of significance (effect size).

gyrA T83I, nalC S46A and armR showed larger effect sizes as seen by the difference in the number of isolates belonging to each of the four groups.

Each of cluster1 and cluster5 showed significant absence in high-risk groups. These clusters were also linked to quinolone susceptibility. Cluster 4 which was linked to quinolone resistance showed significant presence in high-risk groups showing larger effect size. The difference in distribution among the four groups represents the magnitude of effect (effect size).

4.4.7. Aminoglycoside resistance markers of *Ps. aeruginosa* epidemic high-risk clones

The distribution of previously identified aminoglycoside resistance and sensitivity markers (in *Chapter 3*) is analyzed in this section in relation to high risk clones. Table 4.7 and Table 4.9 show the detailed distribution for both gentamycin and amikacin.

A total of 300 isolates were analyzed for gentamycin susceptibility phenotype (106 resistant-194 susceptible). Of those analyzed, a total of 81 isolates belonged to epidemic high risk clones (ST523 [N=4], ST446 [N=6], ST395 [N=7], ST308 [N=6], ST274 [N=6], ST253 [N=5], ST244

[N=8], ST235 [N=13], ST233 [N=5], ST179 [N=6], ST175 [N=2], ST111 [N=4], ST27 [N=6], ST17 [N=3]).

A chi-square test for independence (with yates' continuity correction) indicated significant association between ***nalC*E153Q** and high-risk groups, x^2 (1, n= 300) = 11.84, *p*=0.001, *phi*= 0.199. Significant associations were also found between each of ***pmr*ALeu71Arg**, x^2 (1, n= 300) = 9.925, *p*=0.002, *phi*= 0.182, ***gidB*E97Q**, x^2 (1, n= 300) =25.237, *p*<0.005, *phi*= 0.29, ***gidB*E186A**, x^2 (1, n= 300) =25.237, *p*<0.005, *phi*= 0.29, ***pst*BE89Q**, x^2 (1, n= 300) =5.255, *p*=0.022, *phi*= 0.132, ***arn*DG206C**, x^2 (1, n= 300) =3.839, *p*<0.05, *phi*= 0.113 and high-risk groups. Significant associations between individual markers and clusters of molecular markers for gentamycin and specific high-risk clones are summarized in Table 4.8.

Table 4.7. Distribution of gentamycin molecular markers in relation to high-risk clones

Genome ID	**ST**	**serogroup**	**Gentamycin MIC**	**Gentamycin sensitivity**	***nal*Dser32Asn**	***nal*CE153Q**	***amp*RA51T**	*pmr*ALeu71Arg	*fus*A1D588G	*gid*BQ28K	*gid*BE97Q	*gid*BE186A	*pst*BE89Q	*arn*AA170T	*arn*DG206C	GentamycinCluster1	GentamycinCluster5
PAE0044	532	11	8														
PAE0022	532	11	2														
PAE0101	532	11	2														
287.7776	532	11	16														
PAE0111	446	11	8														
PAE0127	446	11	4														
PAE0139	446	11	0.5														
PAE0150	446	11	4														
PAE0166	446	11	8														
287.7795	446	11	4														
PAE0093	395	6	4														
PAE0015	395	6	2														
PAE0027	395	6	4														
PAE0054	395	6	1														
287.5703	395	6	16														
287.5749	395	6	16														
287.5956	395	6	4														
PAE0147	308	11	4														
PAE0038	308	11	2														
PAE0052	308	11	2														
PAE0120	308	11	4														
287.5687	308	11	16														
287.6329	308	11	16														
PAE0045	274	3	2														
PAE0103	274	3	4														
PAE0159	274	3	8														
PAE0169	274	3	16														
287.5973	274	3	8														
287.7844	274	3	4														
PAE0055	253	10	8														
PAE0102	253	10	2														

Genome ID	ST	serogroup	Gentamycin MIC	Gentamycin sensitivity	*nal*Dser32Asn	*nal*CE153Q	*amp*RA51T	*pmr*ALeu71Arg	*fus*A1D588G	*gid*BQ28K	*gid*BE97Q	*gid*BE186A	*pst*BE89Q	*arn*AA170T	*arn*DG206C	GentamycinCluster1	GentamycinCluster5
PAE0105	253	10	1					■									■
PAE0122	253	10	2					■									■
287.7850	253	10	8	■				■								■	
PAE0056	244	2	1														■
PAE0064	244	5	1														■
PAE0066	244	5	2														■
PAE0069	244	5	1														■
PAE0121	244	5	2														■
PAE0143	244	5	4														■
287.7781	244	12	16	■												■	
287.7814	244	2	16	■												■	
287.5686	235	11	16	■		■					■	■					
287.2973	235	11	16	■		■					■	■					
287.2975	235	11	16	■		■					■	■					
287.2976	235	11	16	■		■					■	■					
287.5690	235	11	16	■		■					■	■					
287.6327	235	11	0.5			■					■	■					
287.6328	235	11	4			■					■	■					
287.7780	235	11	16	■		■	■				■	■					
287.7782	235	11	16	■		■					■	■					
287.7783	235	11	16	■		■					■	■					
287.7785	235	11	16	■	■	■					■	■					
287.7786	235	11	8	■							■	■					
287.7830	235	11	4			■					■	■					
287.2980	233	6	16	■				■					■	■	■		
287.5688	233	6	16	■				■					■	■	■		
287.5752	233	6	16	■				■					■	■	■		
287.5959	233	6	16	■				■					■	■	■		
287.6492	233	6	16	■				■					■	■	■		
PAE0092	179	6	2					■									■
PAE0042	179	6	2					■									■
PAE0141	179	6	4					■									■
287.5748	179	6	8	■				■								■	
287.5969	179	6	2					■									■
287.8029	179	6	4					■									■
PAE0025	175	4	4					■									■
PAE0016	175	4	64	■				■									
287.5750	111	12	16	■												■	
287.6330	111	12	16	■												■	
287.7772	111	12	16	■												■	
287.7862	111	4	4														■
PAE0029	27	1	8	■												■	
PAE0009	27	1	2														■
PAE0050	27	1	1														■
PAE0086	27	1	1														■
PAE0108	27	1	4														■
287.7773	27	1	4														■
PAE0067	17	1	1					■									■
PAE0128	17	1	4					■									■
287.7807	17	1	4					■									■

Shaded cells: variant present, White cells: variant absent

Table 4.8. Gentamycin molecular markers showing significant associations with specific high-risk clones

Molecular marker	High-risk group	Chi-square	*p*-value	Phi-coefficient
***nalC*E153Q**	ST235	160.177	<0.005	0.731
***pmrA*Leu71Arg**	ST532	11.341	0.001	0.194
***pmrA*Leu71Arg**	ST308	17.127	<0.005	0.239
***pmrA*Leu71Arg**	ST253	14.224	<0.005	0.218
***pmrA*Leu71Arg**	ST233	14.224	<0.005	0.218
***pmrA*Leu71Arg**	ST235	4.858	0.028	-0.127
***pmrA*Leu71Arg**	ST179	17.127	<0.005	0.239
***pmrA*Leu71Arg**	ST17	8.477	0.004	0.168
***gid*BE97Q**	ST235	241.202	<0.005	**0.897**
***gid*BE186A**	ST235	241.202	<0.005	**0.897**
***pst*BE89Q**	ST233	185.593	<0.005	**0.787**
***arn*AA170T**	ST233	147.458	<0.005	**0.701**
***arn*DG206C**	ST233	164.407	<0.005	**0.74**
Gentamycin cluster 1	ST111	4.739	0.029	0.126
Gentamycin cluster 5	ST235	21.254	<0.005	-0.266
Gentamycin cluster 5	ST233	7.953	0.005	-0.163

The findings and tables presented above aimed at finding the association of gentamycin markers (shown in *Chapter 3*) in relation to high-risk clones. Markers showing significant association with high-risk groups included: *nal*CE153Q (especially with ST-235, high effect size:0.73), *gid*BE97Q (especially with ST-235, high effect size: 0.897), *gid*BE186A (especially with ST-235, high effect size: 0.897), *pst*BE89Q (especially with ST-233, high effect size: 0.787) and *arn*DG206C (especially with ST-233, high effect size: 0.74). In addition, *arn*AA170T showed significant association with ST-233 with high effect size of 0.7.

Gentamycin Cluster 1 showing association with resistance (*Chapter 3*) showed to be significantly associated with the high-risk clone ST-111 while gentamycin Cluster 5 showing association with gentamycin susceptibility (*Chapter 3*) showed significant absence in high-risk clones ST-235 and ST-233.

A total of 691 isolates were analyzed for Amikacin susceptibility (148 are amikacin resistant). Of those analyzed, a total of 239 isolates belonged to epidemic high-risk clones (ST 17 [N=12], ST 27 [N=11], ST 111 [N=30], ST 235 [N=50], ST 175 [N=10], ST 179 [N=13], ST 244 [N=20], ST 233 [N=13], ST 308 [N=18], ST 395 [N=17], ST 532 [N=6], ST 446 [N=8], ST 274 [N=14], ST 253 [N=17]).

A chi-square test for independence (with yates' continuity correction) indicated significant association between ***armR*** and high-risk groups, x^2 (1, n= 691) = 19.289, p<0.005, *phi*= 0.167. Significant association was also found between each of ***nalC***, x^2 (1, n= 691) = 4.888, p=0.027, *phi*= 0.084, ***mexZ***, x^2 (1, n= 691) = 5.391, p=0.02, *phi*= 0.088, ***ampR***, x^2 (1, n= 691) = 4.452, p=0.035, *phi*= 0.08, ***pmrBGly423Cys***, x^2 (1, n= 691) = 18.239, p<0.005, *phi*= 0.162, ***pmrALeu71Arg,*** x^2 (1, n= 691) = 21.637, p<0.005, *phi*= 0.177 , ***nuo*GA890T**, x^2 (1, n= 691) = 26.014, p<0.005, *phi*= 0.194, ***pstBE89Q***, x^2 (1, n= 691) = 8.42, p=0.004, *phi*= 0.11, ***fao*AT385A**, x^2 (1, n= 691) = 11.823, p=0.001, *phi*= 0.131, ***arn*AA170T**, x^2 (1, n= 691) = 4.213, p=0.04, *phi*= 0.078, ***arn*DG206C**, x^2 (1, n= 691) = 4.213, p=0.04, *phi*= 0.078, ***phoQY85F***, x^2 (1, n= 691) = 44.957, p<0.005, *phi*= 0.255 and high-risk groups.

Each of ***mex*RR70N**, x^2 (1, n= 691) = 12.347, p<0.005, *phi*= -0.134, ***mex*RE70R**, x^2 (1, n= 691) = 18.223, p<0.005, *phi*= -0.162, ***mex*RL130T**, x^2 (1, n= 691) = 15.579, p<0.005, *phi*= -0.150, ***mex*RG97L**, x^2 (1, n= 691) = 19.259, p<0.005, *phi*= -0.167, ***mex*RL29D**, x^2 (1, n= 691) = 19.23, p<0.005, *phi*= -0.167 showed significant absence in high-risk groups.

Significant associations between individual markers or clusters of markers for amikacin and specific high-risk clones are summarized in Table 4.10.

Table 4.9. Distribution of amikacin molecular markers in relation to high-risk clones

Genome ID	ST	serogroup	Amikacin MIC	Amikacin sensitivity	*pho*QY85F	*nuo*GA890T	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C	*mex*RR79N	*mex*RE70R	*mex*RL130T	*mex*RG97L	*mex*RL29D	*pmr*BGly423Cys	*pmr*ALeu71Arg	*fus*A1D588G	*gid*BE186A	*nal*Dser32Asn	Amikacincluster1	Amikacin cluster 2
PAE0044	532	11	4										■	■	■	■	■		■				■	
PAE0022	532	11	1										■	■	■	■	■	■	■				■	
PAE0101	532	11	2										■	■	■	■	■		■				■	
287.1022	532	11	2																■				■	
287.7776	532	11	8										■	■	■	■	■		■				■	
287.971	532	11	2															■	■				■	
PAE0111	446	11	4						■				■	■	■	■	■						■	
PAE0127	446	11	2						■				■	■	■	■	■						■	
PAE0139	446	11	1						■				■	■	■	■	■						■	
PAE0150	446	11	2						■				■	■	■	■	■						■	
PAE0166	446	11	4						■				■	■	■	■	■						■	
287.1205	446	11	0.1						■														■	
287.7795	446	11	8						■				■	■	■	■	■						■	
287.875	446	11	4						■														■	
PAE0093	395	6	4			■							■	■	■	■	■						■	
PAE0015	395	6	1			■							■	■	■	■	■						■	
PAE0027	395	6	2			■							■	■	■	■	■						■	
PAE0054	395	6	1			■							■	■	■	■	■						■	

287.1039	395	6	2																					
287.1080	395	6	2																					
287.1106	395	6	2																					
287.1115	395	6	1																					
287.1116	395	6	4																					
287.5703	395	6	4																					
287.5749	395	6	4																					
287.5956	395	6	16																					
287.848	395	6	4																					
287.881	395	6	8																					
287.887	395	6	2																					
287.921	395	6	1																					
287.961	395	6	1																					
PAE0147	308	11	4																					
PAE0038	308	11	1																					
PAE0052	308	11	1																					
PAE0120	308	11	2																					
287.1038	308	11	0.5																					
287.1072	308	11	0.5																					
287.1101	308	11	0.3																					
287.1104	308	11	4																					
287.1118	308	11	0.5																					
287.1138	308	11	128																					
287.1142	308	11	64																					
287.1211	308	11	64																					
287.5687	308	11	64																					
287.6329	308	11	8																					
287.913	308	11	4																					
287.964	308	11	4																					
287.965	308	11	4																					
287.995	308	11	4																					
PAE0045	274	3	2																					
PAE0103	274	3	1																					
PAE0159	274	3	4																					
PAE0169	274	3	16																					
287.1007	274	3	4																					
287.1019	274	3	4																					
287.1228	274	3	4																					
287.5973	274	3	16																					
287.7844	274	3	8																					
287.846	274	3	8																					
287.878	274	3	4																					
287.885	274	3	0.3																					
287.953	274	3	4																					
287.963	274	3	0.1																					
PAE0055	253	10	4																					
PAE0102	253	10	2																					
PAE0105	253	10	1																					
PAE0122	253	10	1																					
287.1035	253	10	2																					
287.1081	253	10	1																					
287.1085	253	10	1																					
287.1154	253	10	1																					
287.1157	253	10	2																					
287.7843	253	10	8																					
287.7850	253	10	16																					
287.873	253	10	2																					
287.893	253	10	2																					
287.894	253	10	2																					
287.907	253	10	1																					
287.912	253	11	4																					
287.985	253	10	16																					
PAE0056	244	2	1																					
PAE0064	244	5	2																					
PAE0066	244	5	1																					
PAE0069	244	5	1																					
PAE0121	244	5	1																					
PAE0143	244	5	4																					
PAE0173	244	5	8																					
287.1029	244	2	2																					
287.1060	244	12	8																					
287.1065	244	2	0.5																					
287.1089	244	2	1																					
287.1096	244	12	32																					
287.1097	244	12	64																					

287.1208	244	12	32																					
287.7781	244	12	32																					
287.7814	244	2	8																					
287.897	244	12	64																					
287.905	244	2	1																					
287.956	244	12	4																					
287.975	244	11	64																					
287.1031	235	11	32																					
287.5686	235	11	8																					
287.1021	235	11	32																					
287.1045	235	11	32																					
287.1058	235	11	8																					
287.1087	235	11	4																					
287.1088	235	11	4																					
287.1100	235	11	32																					
287.1122	235	11	8																					
287.1124	235	11	32																					
287.1128	235	11	8																					
287.1131	235	11	16																					
287.1133	235	11	8																					
287.1134	235	11	64																					
287.1135	235	11	4																					
287.1144	235	11	0.3																					
287.1145	235	11	8																					
287.1147	235	11	128																					
287.1148	235	11	16																					
287.1150	235	11	16																					
287.1151	235	11	8																					
287.1158	235	11	32																					
287.1162	235	11	16																					
287.1163	235	11	128																					
287.1184	235	11	4																					
287.1200	235	11	4																					
287.1202	235	11	64																					
287.1203	235	11	2																					
287.1213	235	11	64																					
287.2973	235	11	64																					
287.2975	235	11	64																					
287.2976	235	11	8																					
287.5690	235	11	64																					
287.6327	235	11	2																					
287.6328	235	11	8																					
287.7780	235	11	8																					
287.7782	235	11	8																					
287.7783	235	11	8																					
287.7785	235	11	32																					
287.7786	235	11	8																					
287.7830	235	11	8																					
287.855	235	11	64																					
287.860	235	11	1																					
287.902	235	10	4																					
287.903	235	11	8																					
287.915	235	11	128																					
287.918	235	11	16																					
287.955	235	11	8																					
287.966	235	11	1																					
287.984	235	11	32																					
287.1004	233	6	128																					
287.1011	233	6	128																					
287.1119	233	6	128																					
287.1136	233	6	64																					
287.1204	233	6	64																					
287.2980	233	6	64																					
287.5688	233	6	64																					
287.5752	233	6	64																					
287.5959	233	6	64																					
287.6492	233	6	64																					
287.889	233	6	128																					
287.934	233	6	128																					
287.991	233	6	1																					
PAE0092	179	6	2																					
PAE0042	179	6	1																					
PAE0141	179	6	4																					
287.1103	179	6	4																					
287.1139	179	11	16																					

287.1143	179	6	16																					
287.1167	179	6	2																					
287.1238	179	6	8																					
287.5748	179	6	64																					
287.5969	179	6	8																					
287.8029	179	6	8																					
287.980	179	6	2																					
287.988	179	6	4																					
PAE0025	175	4	2																					
PAE0016	175	4	2																					
287.1024	175	4	4																					
287.1042	175	4	4																					
287.1051	175	4	4																					
287.1052	175	4	4																					
287.1132	175	4	2																					
287.1188	175	4	4																					
287.1192	175	4	4																					
287.908	175	4	2																					
287.1000	111	12	32																					
287.1046	111	12	4																					
287.1084	111	12	8																					
287.1093	111	4	2																					
287.1095	111	4	2																					
287.1098	111	4	1																					
287.1125	111	12	128																					
287.1126	111	12	8																					
287.1127	111	12	4																					
287.1130	111	12	4																					
287.1140	111	12	128																					
287.1141	111	12	128																					
287.1146	111	12	128																					
287.1169	111	12	8																					
287.1172	111	12	2																					
287.1178	111	12	128																					
287.1179	111	12	128																					
287.1185	111	12	8																					
287.1195	111	12	64																					
287.1209	111	4	128																					
287.1216	111	12	128																					
287.5750	111	12	64																					
287.6330	111	12	32																					
287.7772	111	12	8																					
287.7862	111	4	8																					
287.852	111	12	4																					
287.909	111	12	1																					
287.911	111	12	1																					
287.916	111	12	64																					
287.951	111	12	16																					
PAE0029	27	1	4																					
PAE0009	27	1	1																					
PAE0050	27	1	1																					
PAE0086	27	1	2																					
PAE0108	27	1	1																					
287.1086	27	1	2																					
287.1241	27	1	64																					
287.1301	27	1	8																					
287.7773	27	1	8																					
287.954	27	1	4																					
287.987	27	1	4																					
PAE0067	17	1	2																					
PAE0128	17	1	2																					
287.1075	17	1	2																					
287.1129	17	1	16																					
287.1240	17	1	32																					
287.7807	17	1	8																					
287.847	17	1	4																					
287.850	17	1	2																					
287.900	17	1	2																					
287.983	17	1	2																					
287.989	17	1	0.1																					
287.998	17	1	4																					

Grey cells show marker presence, white cells show marker absence, Same STs are highlighted the same color

Table 4.10. Amikacin molecular markers showing significant associations with specific high-risk clones

Molecular marker	High-risk group	Chi-square	*p*-value	Phi-coefficient
***arm*R**	ST235	8.473	0.004	0.111
***arm*R**	ST233	3.863	0.049	0.075
***arm*R**	ST111	9.145	0.002	0.115
***mex*RR79N**	ST235	8.625	0.003	-0.112
***mex*RR79N**	ST111	11.861	0.001	-0.131
***mex*RE70R**	ST235	11.117	0.001	-0.127
***mex*RE70R**	ST111	12.389	<0.005	-0.134
***mex*RL130T**	ST235	9.234	0.002	-0.116
***mex*RL130T**	ST111	12.389	<0.005	-0.134
***mex*RG97L**	ST235	9.191	0.002	-0.115
***mex*RG97L**	ST111	13.947	<0.005	-0.142
***mex*RL29D**	ST235	10.655	0.001	-0.124
***mex*RL29D**	ST111	13.602	<0.005	-0.14
***pmr*BGly423Cys**	ST308	21.377	<0.005	0.176
***pmr*BGly423Cys**	ST253	30.584	<0.005	0.21
***pmr*BGly423Cys**	ST244	3.942	0.047	-0.076
***pmr*BGly423Cys**	ST235	26.891	<0.005	0.197
***pmr*BGly423Cys**	ST111	21.784	<0.005	0.178
***pmr*ALeu71Arg**	ST532	17.444	<0.005	0.159
***pmr*ALeu71Arg**	ST395	6.047	0.014	-0.094
***pmr*ALeu71Arg**	ST308	38.514	<0.005	0.236
***pmr*ALeu71Arg**	ST253	35.572	<0.005	0.227
***pmr*ALeu71Arg**	ST244	4.641	0.031	-0.082
***pmr*ALeu71Arg**	ST233	30.682	<0.005	0.211
***pmr*ALeu71Arg**	ST179	38.185	<0.005	0.235
***pmr*ALeu71Arg**	ST175	29.243	<0.005	0.206
***pmr*ALeu71Arg**	ST111	10.882	0.001	-0.125
Amikacin cluster 1	ST233	42.918	<0.005	-0.245
Amikacin cluster 1	ST111	9.267	0.002	-0.116
Amikacin cluster 2	ST233	42.918	<0.005	0.245
Amikacin cluster 2	ST111	9.267	0.002	0.116

Amikacin resistance markers identified in *Chapter 3* were evaluated in relation to high-risk groups and showed association of each of *arm*R, *nal*C, *mex*Z and *amp*R with high-risk groups. Markers showing significant association with specific high-risk clones included: *gid*BE186A with ST-235, *pho*QY85F with ST-233, *nuo*GA890T with ST-395 and *lpt*AR62S with ST-446.

It is also noticed from assessment of aminoglycoside markers for both gentamycin and amikacin that *pmr*Aleu71Arg was specifically highly conserved among isolates with ST-532, ST-253, ST-

308, ST-233, ST-179, and ST-175. At the same time the marker showed significant absence in isolates with ST-446, ST-395, ST-111, ST-235 and ST-244.

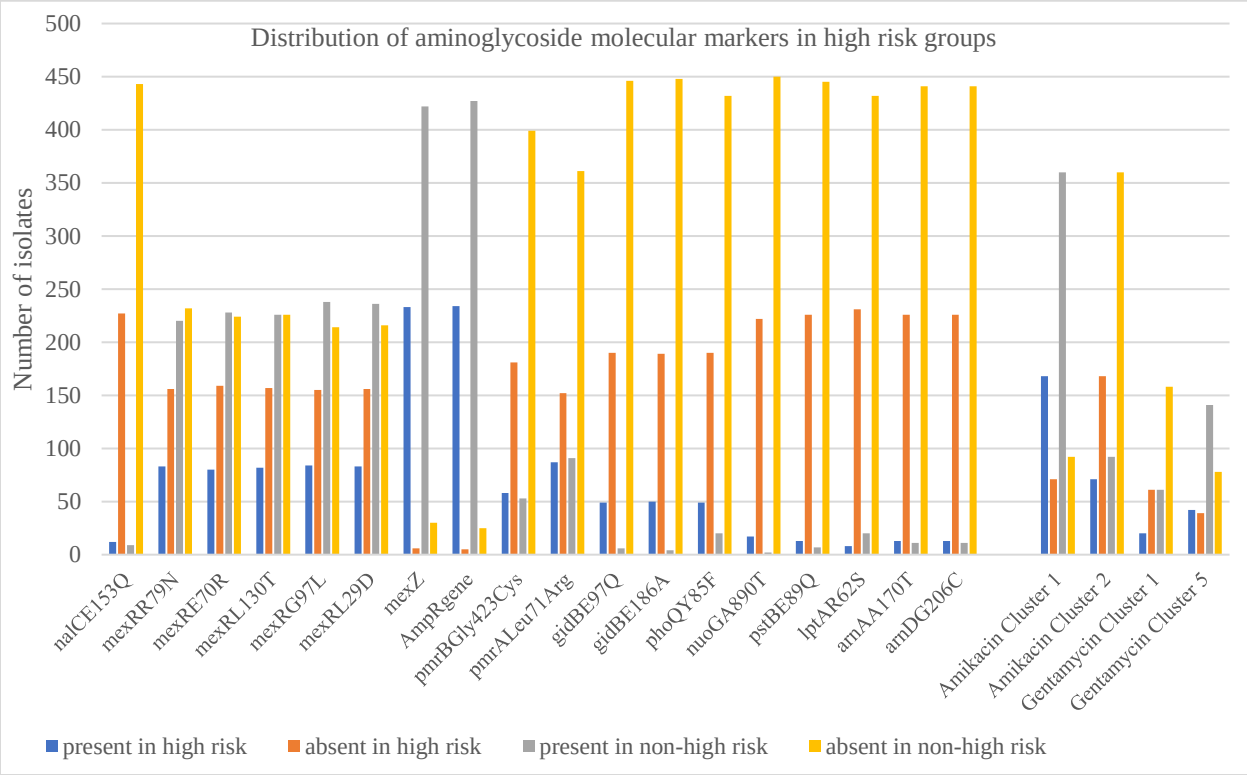


Figure 4.9. Differential distribution of tested aminoglycoside markers in high-risk versus non high-risk clones in the whole collection of studied isolates

The figure shows the distribution of variants showing significant absence/presence in relation to high-risk clones. The four groups composing each variant are shown in four different colors. The difference in the distribution among the four groups composing each variant reflects the magnitude of significance (effect size).

Markers that were significantly absent in high-risk groups include mexR R79N, mexR E70R, mexR L130T, mexR G97L, and mexR L29D. All with smaller effect sizes for the difference between the groups. This is represented by the difference in the number of isolates among the four groups.

Each of pmrA Leu71Arg, pmrBGly423Cys, gidBE97Q, gidBE186A, and phoQY85F showed significant association with high-risk groups with gidB and phoQ variants showing higher effect sizes as seen by the difference of distribution of isolates among the groups.

Gentamycin cluster 5 which clusters towards more gentamycin susceptible isolates showed significant absence in high-risk groups with moderate effect size. Gentamycin cluster 1 showed significant presence in high-risk groups in general with smaller effect size.

Amikacin cluster 2 which clusters towards more amikacin resistant isolates showed significant presence in high-risk groups. Amikacin cluster 1 linked to amikacin susceptibility showed significant absence in high-risk groups. The difference in distribution as seen in different column heights represents the magnitude of difference or the effect size.

4.5. Discussion

The chapter has attempted to provide a brief summary on the population structure of *Ps. aeruginosa* giving an overview on some conventional and molecular typing methods used with the organism. Although not the primary objective of the thesis, it was necessary to first describe the population structure of the studied set of 691 *Ps. aeruginosa* isolates. This was based on two commonly used typing methods; MLST and serotyping. The association of tested quinolone and aminoglycoside resistance/susceptibility markers was then investigated in relation to background genomic context and especially to high-risk clones.

Analysis of the population structure revealed a low value for association index (I_A) approaching zero. This is considered highly suggestive of a non-clonal epidemic population (Smith *et al.*, 1993). The standardized index of Association for the studied set of isolates was (I^S_A)=0.1302 (P<0.001). This indicates linkage equilibrium and low evidence of association among the alleles analyzed. In support of that is the finding of a relatively high Mean genetic diversity (H): 0.8648 +/- 0.0261. These results together support the idea that frequent recombination plays a key role in allele distribution in the population (Lenski, 1993).

The results of the current study generally support the description of the population structure of *Ps. aeruginosa* found in most previous studies. Several studies using MLST have consistently calculated low index of association of 0.29 (Curran *et al.*, 2004), 0.31 (Kiewitz and Tümmler, 2000), and 0.35 (Maatallah *et al.*, 2011) indicating non-clonal type of population structure. Most studies have consistently reported a panmictic structure (Morales-Espinosa *et al.*, 2012) (Griffith *et al.*, 1989) (Lomholt, Poulsen and Kilian, 2001), a finding that is supported by the current study. This also supports conclusions drawn about the population structure of *Ps. aeruginosa* which is consensually believed to be of panmictic-epidemic nature (Magalhães *et al.*, 2020).

Within the panmictic population structure, several multi-locus STs are known to successfully disseminate across diverse geographic locations and patient populations worldwide. These are commonly reported as related to multidrug resistance (MDR) and are therefore known as high-risk clones. These STs include ST111, ST175, ST235, ST244, and ST395. ST235 is the most prevalent of these international high risk clones which are widespread worldwide showing association with poor clinical outcome and are usually linked to high level of antibiotic resistance (Pirnay *et al.*, 2002)(Antonio Oliver *et al.*, 2015). Although nearly 100 different

horizontally acquired resistance elements have been reported in these clones especially for the most studied clone (ST235 isolates), high level resistance can be readily rendered by other chromosomal changes or core genome determinants (Antonio Oliver *et al.*, 2015) (Roy Chowdhury, Scott and Djordjevic, 2017). Despite their clinical importance, the molecular basis underlying the spread and success of high risk clones including the international clone ST235 (Treepong *et al.*, 2017) is not completely understood.

Different molecular epidemiology and population structure studies on Gram-negative bacteria have identified MDR strains that successfully disseminate across diverse geographic locations and different patient populations and therefore known as high-risk clones. Metallo beta-lactamases (MBLs) and Extended-spectrum Beta-lactamases (ESBLs) have both been classically linked to the successful spread and clonal dissemination of MDR *Ps. aeruginosa* in several studies (Giske *et al.*, 2006) (Empel *et al.*, 2007) (Viedma *et al.*, 2009). However, this finding was not consistent. A population analysis of 103 MDR *Ps. aeruginosa* isolates with 42 isolates belonging to ST235 and 15 isolates belonging to ST111 showed that only few ST235 and ST111 isolates are producing VIM-1, VIM-2 MBLs or other ESBLs including PER-1 or GES-7 (Guzvinec *et al.*, 2014). In the same study, isolates with multiple amino acid polymorphisms in OprD were observed. Only 4 of these isolates belonged to ST235 while 4 other isolates belonged to ST111. *gyr*AT83I was detected only in one isolate from the ST111 sequence type while the combination of *gyr*AT83I and *par*CS87L were identified only in two isolates belonging to ST235. This makes it necessary to understand the genetic environment related to mobilization of MBL genes which are frequently observed in high-risk clones.

Correa *et al.*, (2015) showed that the dissemination of extensively drug resistant *Ps. aeruginosa* has been repeatedly linked to the presence of mobile genetic elements that would facilitate their successful spread and clonal dissemination. The study has linked the clonal success of ST235 and ST111 to harboring *bla* KPC-2 and *bla* VIM-2 respectively and showed that XDR high-risk clones of *Ps. aeruginosa* rely on class 1 integrons and on the transposable element Tn4401 as principal structures for gene mobilization in ST111 and ST235.

In addition, Chowdhury *et al.*, (2016) have linked three MDR isolates representative of *Ps. aeruginosa* ST235 to genomic islands 1 (GI1) carrying Tn6162 and genomic islands 2 (GI2) carrying Tn6163. These are thought to be probably responsible for capturing and mobilizing

antibiotic resistance gene cassettes carried by class 1 integrons (Chowdhury *et al.*, 2016). The study shows that the presence of these genomic islands can be considered as a characteristic of ST235 clonal lineage. However, there is no sufficient evidence to show whether this association is linked to MDR behavior or whether it is an inherent character of the ST235 lineage. Among the resistance genes evaluated in the same study were the *nfx*B gene, the *pho*Q variant F76Y, and the *pmr*B variant V15I which were all identified in the three ST235 isolates from the same study. These findings support findings from the current analysis about the possible importance of the associations observed in *nfx*B, *pho*Q, and *pmr*AB variants to the ST235 clone.

Another study has previously evaluated the genetic characters of ST235 versus non ST235 (Cho *et al.*, 2013). Out of 68 *Ps. aeruginosa* clinical isolates collected over the time period of 4 years, 34 isolates belonged to ST235. The study has specifically evaluated aminoglycoside modifying enzymes (AME) and class 1 integrons and found that all isolates belonging to ST235 contained AME whereas 23.5 % of non-ST235 isolates also contained AME genes. The same study has also detected class 1 integrase gene in 17 isolates out of 68 isolates; 16 of which belonged to ST235 and 1 isolate belonged to ST357 (Cho *et al.*, 2013). This would again question the nature of the association observed between the genetic elements observed, MDR behavior and the high-risk clonal lineage.

Similar observations have also supported these findings and have reported the frequent carriage of GI 1 and GI 2 among the commonly observed international high-risk clones including ST235, ST111, and ST175. Chowdhury et al., (2017) showed that all *Ps. aeruginosa* strains from their study carry one or more class 1 integrons suggesting that acquisition or loss of the observed genomic islands and the transposons they harbor (including Tn6060 and Tn6249) are likely influencing the resistome carried by ST235 and consequently their global dissemination (Roy Chowdhury, Scott and Djordjevic, 2017).

Although it has been proposed that *Ps. aeruginosa* "high-risk clones" represent distinct lineages which are highly capable of acquiring and maintaining resistance genes and/or the mobile genetic elements containing these genes as compared with the general *Ps. aeruginosa* population (L. L. Wright *et al.*, 2015), this assumption needs to be further tested. Clonal success in susceptible *Ps. aeruginosa* high-risk clones tends to be under investigated in most studies which consequently necessitates the assessment of the genetic markers underlying clonal success in a

range of both resistant and susceptible high-risk clones, the point that the current study set has tried to cover. The primary aim of this chapter was to make an additional investigation of the molecular basis of success of some identified high-risk clones and to study the possible link that may exist between different antibiotic susceptibility and resistance determinants and high-risk clones. To achieve that, I have specifically searched for specific signatures in previously investigated resistance/susceptibility markers including QRDR, efflux pumps operons, cell membrane related proteins and others.

To ensure the diversity and comprehensive representation of the studied set of isolates, phylogenetic analysis and hierarchical clustering of the studied isolates were performed in relation to all known STs for *Ps. aeruginosa*. The studied set of isolates proved to be diversely distributed as shown in Figure 4.1 and Figure 4.2.

Understanding the reason for success of epidemic high-risk clones is essential for designing treatment and infection control strategies (Baquero, Coque and de la Cruz, 2011). The specific genetic resistance markers of these high-risk clones were described in detail for the first time by (Cabot *et al.*, 2012). These may include multiple combinations of chromosomal mutations and/or horizontally acquired resistance elements. The mosaic nature observed for either chromosomal or acquired resistance elements in relation to high-risk clones should draw much caution before making conclusions about the molecular bases of success for these clones. It is also important to carefully consider the fitness cost of any of these underlying mutations in relation to the genetic background of the strains (Kugelberg *et al.*, 2005).

Wastewater networks have been recently reported as trafficking sources between hospital wash basins for pathogenic bacteria including *Ps. aeruginosa.* In a recent study, 25 wash basin U-bends were investigated in five locations in a dental hospital and showed highly related *Ps. aeruginosa* strains identified in several locations with some risky clones including ST179 and ST560 (Moloney *et al.*, 2019). In addition, contaminated bronchoscopes rinsing water and connecting tubes have been also identified as a reservoir for spreading the organism (Zhang *et al.*, 2020). This situation makes it important to understand whether the clonal success of epidemic high-risk clones is essentially related to mutational resistance or to horizontally acquired resistance elements, i.e. antibiotic inactivating enzymes or integrons carrying specific

gene cassettes. Such an understanding can greatly impact the choice of the best approach required to tackle these high-risk clones and to control their dissemination.

Understanding the biologic nature of the genetic markers associated with the poor clinical outcome helps to determine whether the poor clinical outcome is related to the MDR behavior, to other virulence characters, or to other underlying biologic characters that make the organism more pathogenic or aggressive in host tissues. It is important to investigate and discriminate the origin on high-risk clones. High-risk behavior may be encountered as a conserved lineage-related character which is consistently detected in different hospitals , consequently indicating the higher probability of inter-hospital dissemination of XDR/MDR high risk clones. The other probability is that originally susceptible clones may undergo independent parallel evolution into high-risk clones in different settings by acquiring these markers.

Although noted that particular sequence types are more infectious and virulent as measured by their global spread in different clinical settings, strains carrying ST235 sequence type cannot be considered identical. It has been shown that resistance gene content may vary greatly because the plasticity of *Ps. aeruginosa* has allowed the capture and rearrangement of resistance regions. This has led to considering strains of ST235 as a potent vector for the spread and evolution of complex resistance loci (Martinez *et al.*, 2014). This has been supported in multiple studies (Liakopoulos *et al.*, 2013) (Edelstein *et al.*, 2013) including a recent work which has identified a new genomic island (PAGI-17) carrying two Tn4401b transposons and double *bla*KPC-2 chromosomal insertions and belonging to ST235 (Abril *et al.*, 2019). The study has shown higher genetic variability of the accessory genome in ST235 and has suggested the possibility of acquiring a specific genetic portion in *Ps. aeruginosa* long time ago by lateral gene transfer which has favored the incorporation and permanence of the two Tn4401b-blaKPC-2 transposons (Abril *et al.*, 2019).

The analysis performed in this chapter shows that 8 of previously identified quinolones resistance markers are exhibiting significant association with high-risk clones. The identified markers include *mex*Z, *arm*R, *nfx*B, *mex*S, *nal*C, *gyr*AT83I, *nal*CE153Q, and *nal*CS46A. Interestingly observed is that 41 isolates out of 50 isolates within the high-risk group ST235 showed the following cluster of molecular markers (*mex*Z, *arm*R, *nfx*B, *mex*S, *mex*R and *nal*C). *gyr*AT83I was identified in 36/50 isolates and *nal*CS46A in 43/50 isolates. Both *gyr*AT83I and

*nal*CS46A showed significant association with ST235 exhibiting the highest effect sizes for individual mutations among all those tested. The third molecular marker that showed the highest effect size among all tested was *nal*CE153Q. Part of these findings may be supported by findings from the literature which showed that *mex*Z G195E, leading to MexXY overexpression, and *gyr*AT83I are among the documented variants in relation to the spread of ST175 XDR phenotype. These were also identified among the mutational resistance mechanisms showing frequent occurrence among ST111 and ST235 high risk clones (Kos *et al.*, 2015). Other findings from the study performed by Treepong *et al.*, (2017) also show association between ORDR *gyr*A T83I and ST235. This finding similarly supports findings from the current analysis. However, *par*C S87(80)I variant showing association with ST235 did not show significance in the current study set. A possible explanation is that phenotypic data in relation to the high-risk clones reported by Treepong *et al.*, (2017) were not studied which may introduce a source of bias in their observations. A recent study has also identified a wide range of mutations in all efflux pump regulators in relation to high-risk clones including *nal*CE153Q which showed the highest effect size in the current analysis (Pelegrin, Saharman, Griffon, Palmieri, Mirande, Karuniawati, Sedono, Aditianingsih, Wil H.F. Goessens, *et al.*, 2019). These findings are also supported by other recent findings which demonstrate the importance of QRDR-related mutations in high-risk clones (Horna *et al.*, 2019). In the same study, Horna *et al.*, (2019) identified QRDR mutations in all isolates belonging to ST235 and ST357.

Similarly observed for the other frequent high-risk clone ST111, 27 isolates out of 30 isolates showed the same cluster of molecular markers (*mex*Z, *arm*R, *nfx*B, *mex*S, *mex*R and *nal*C). *gyr*AT83I was identified in 23/30 isolates and *nal*CS46A identified in 24/30 isolates. Both showed statistically significant association.

Some of the aminoglycoside markers identified in *Chapter 3* also showed significant association with high risk groups. These included *pmr*ALeu71Arg, *nuo*GA890T, *pst*BE89Q, *pho*QY85F, *lpt*AT55A, and *gid*BE186A. The pattern of their distribution seemed to be highly conserved for specific markers in relation to specific high-risk clones. These findings are supported by other recent findings shown by Pelegrin *et al.*, (2019) who identified the same variant *pho*QY85F as highly conserved in ST235. The same study has also shown that *pmr*ALeu71Arg was very frequently identified in both ST235 and ST446. Multiple other variants have also been identified

in *nuo*G in association with high-risk clones from the same study (Pelegrin, Saharman, Griffon, Palmieri, Mirande, Karuniawati, Sedono, Aditianingsih, Wil H.F. Goessens, *et al.*, 2019).

On the other hand, both *arn*AA170T and *arn*DG206C previously shown in *chapter 3* as resistance markers to aminoglycosides has also shown significant association with high-risk behavior especially with ST233 showing very high effect size.

*pmr*ALeu71Arg is an interesting example. It showed significant association with high-risk clones. The marker was not identified in ST111, ST235, or ST395. On the other hand, it appeared to be highly conserved in other high-risk clones including; ST532 (6/6 isolates), ST395 (15/17 isolates), ST308 (16/18 isolates), ST233 (12/13 isolates), ST179 (13/13 isolates), and ST175 (10/10 isolates). These findings are supported by findings from another recent study which showed that *pmr*AB and *pho*PQ are upregulated in relation to high-risk clones and colistin resistance (Vatansever *et al.*, 2020). The study has reported the presence of multiple insertions and deletions in the sequence of these genes but have not shown what specific SNPs were identified.

phoP (which encodes a response regulator) and *phoQ* (which encodes a sensor kinase) (Macfarlane *et al.*, 1999) are part of a three gene-operon. PhoP-PhoQ activity has been implicated in resistance to polycationic antimicrobials (e.g., polymyxins) and cationic antimicrobial peptides. Their role involves promoting an amino arabinose modification of the lipid A portion of lipopolysaccharide (LPS). It is possible that *Pho*PQ-dependent aminoglycoside resistance in *Ps. aeruginosa* involves a similar modification of LPS (McPhee, Lewenza and Hancock, 2003) (Moskowitz, Ernst and Miller, 2004b).

In addition, some mutations in the two-component sensor-regulator system *pmr*AB have been linked to a changed aminoglycoside resistance phenotype. These include *pmr*A (Leu71Arg) located within the signal receiver domain of the response regulator, *pmr*B (Thr4Ala) containing the phosphorylation site of the protein located in the protein secretion signal, and *pmr*B (Leu323His) located in close proximity to the histidine kinase A and ATP binding domains. While activation of each of the two systems separately only showed slight increases in MIC, combined activation of AmgS and the PmrAB system led to a 4-fold increase in the tobramycin MIC (Schniederjans, Koska and Häussler, 2017).

The findings shown above collectively support findings from the literature which show that mutations in several two-component regulatory systems including *pmr*AB, *pho*PQ and the associated overexpression of the *arn*BCADTEF-*pmr*E operon can lead to lipid A modification with L-Ara4N. The modification decreases the net negative charge of lipid A and consequently repels the binding to positively charged polymyxins. This effect proved to be associated with polymyxin resistance in *Ps. aeruginosa* (Han *et al.*, 2019). Similar effect may also lead to gentamycin resistance which can explain the current findings.

Another interesting example was *lpt*AR62S which appeared to be conserved to ST446. It was only identified in all 8 isolates with ST446 and not identified at all in any other isolate belonging to other high-risk groups. Similarly, *nuo*GA890T also appeared to be highly conserved to ST395 (identified in 17/17 isolates) and was not identified at all in any other high-risk group.

*gid*BE186A which showed the highest effect size for significant association with individual high-risk groups (x^2 (1,691) = 241.202, p<0.005, *phi*=0.897) appeared to be a highly conserved marker for ST235. It was identified in 49/50 isolates with ST235 sequence type and in an additional one isolate with ST253 sequence type and not at all in any other high-risk isolates. Similarly, *pho*QY85F (x^2 (1,691) = 44.957, p<0.005, *phi*=0.255) appeared to be another conserved marker that was exclusively identified in ST233 (13/13 isolates) and not identified at all in any other high-risk clone showing very high effect size for its association with ST233 (phi=0.787).

Although high-risk clonal lineages have been previously linked to some horizontally acquired resistance genes and also to genomic islands 1 (GI1) and genomic islands 2 (GI2), there is still insufficient evidence to show whether markers linked to high-risk clones are conserved markers across the lineage for each high-risk clone or that these markers occur as a change in originally susceptible and frequently occurring clones that undergo independent parallel evolution into high-risk clones in different settings. Findings from the current analysis could support the second trajectory because high-risk clone related markers analyzed here show association with quinolone resistance, but some do not show association with aminoglycoside resistance.

This observation may support the assumption that the success of these clones is related to biological functional changes caused by the associated variants. These functional modifications may give the clone its success rather than being merely related to resistance or virulence

determinants. Although inactivating enzymes were not assessed here in relation to high-risk clones, a new group of mutational variants in chromosomal genes related to efflux pumps, efflux-pump regulators, and membrane proteins showing strong association in a large diverse set of isolates can support the assumption that horizontally acquired elements ,whether through plasmids or integrons, are not the sole underlying molecular elements for the success and spread of epidemic high-risk clones. Variants identified from the current analysis can represent biologic markers showing increased fitness and leading to the acquisition of specific adaptive or beneficial traits. These variants may also represent an adaptation to chronic infections. However, this needs to be further investigated by studying the variants' biologic effect which is beyond the objectives of this work. This assumption has been previously supported by observations of *Ps. aeruginosa* sub-lineages with independent signatures of adaptation within the larger population which may result into distinct biologic activities (O'Brien *et al.*, 2017). This assumption can be supported by findings from other studies that link high-risk clones to some virulence and adaptive biologic characters including defective motility, defective pigment production and increased biofilm formation (Mulet *et al.*, 2013). Isolates with increased mutation rate (strong mutators) are known to be frequent in chronic *Ps. aeruginosa* infection while isolates with very low mutation rate are often linked to acute nosocomial infection (Mena *et al.*, 2008) (Oliver *et al.*, 2000). Although not much information is available on mutation rate of *Ps. aeruginosa* high-risk clones, it has been suggested that the mutator phenotypes observed in high-risk clones may play a role in the adaptability required for the global success and dissemination of high-risk clones showing markers similar to those observed with chronic infections (Mulet *et al.*, 2013), an assumption that could be supported by the findings from the current analysis. Another evidence suggesting the independent acquisition of adaptive characters are the findings of mutational resistance arising independently across distinct phylogenetic lineages and contributing to mutation-driven evolution of *Ps. aeruginosa* population structure (Zamudio *et al.*, 2019).

Although reported that global success of bacterial pathogens is determined by the interplay of regulatory networks interconnecting resistance and virulence (Martínez and Baquero, 2002), for the purpose and scope of this chapter, a large collection of diverse completely sequenced genomes with different resistance profiles was used to analyze a panel of 11 quinolone molecular markers, 23 aminoglycoside molecular markers in addition to the combinations of markers that have been investigated in the previous chapter. These could show potential link to clonal success

and consequently to clinical significance. Some of the previously identified molecular markers and clusters of markers showed significant association with some high-risk clones. This can add to their potential value to be used as molecular diagnostic markers and can also be useful for future design of specific treatment and infection control strategies.

In summary, the results of this chapter indicate that *Ps. aeruginosa* high-risk clones are significantly associated with a defined set of molecular markers which include; *mex*Z, *arm*R, *nfx*B, *mex*S, *mex*R, *nal*C, *gyr*AT83I, *nal*CS46A, *nal*CG71E, and *nal*CE153Q. These markers showed significant association with epidemic high-risk clones markedly observed with ST235 and ST111. It is also important to note here that only ST235, ST111, and ST175 show significant association with quinolone resistance with no other high-risk clones showing association with aminoglycoside resistance. On the other hand, some aminoglycoside markers have shown specific linkage and conservation to specific high-risk clones. These include *gid*BE186A in relation to ST235, *pho*QY85F in relation to ST233, *nuo*GA890T in relation to ST395, and *lpt*AR62S in relation to ST446. This may result in the assumption that these mutations probably offer some sort of adaptive fitness that underlie increased pathogenicity which is not necessarily related to antibiotic resistance. This may draw attention into considering the fitness characters of high-risk clones as a probable co-shared underlying element for both MDR/XDR phenotype and high-risk phenotype. This supports the importance of investigating the molecular basis of fitness in high-risk clones independent of their associated resistance or virulence characters although some overlap may exist in the genetic determinants underlying all these types of behavior. In support of that are findings reporting the lack of correlation between MDR phenotype and virulence phenotype among some common high-risk clones (Gómez-Zorrilla *et al.*, 2016). In addition, studying the characteristics underlying persistent *Ps. aeruginosa* infection in cystic fibrosis lung has concluded that cytotoxicity was a persistent character among persister versus eradicated isolates (Tramper-Stranders *et al.*, 2012)

Although susceptible or non-MDR isolates from high-risk clones tend to be under investigated in the literature (Antonio Oliver *et al.*, 2015), some recent findings have shown that pathogenicity characters and virulence genotype is highly correlated with high-risk clones independent of their associated MDR phenotype including a significant association of *exo*U-positive genotype with the high-risk clone ST235. It has also been shown in another study that *exo*U-positive isolates are

frequently detected among ST235 and O11 serotype and displaying a poor prognosis (Recio *et al.*, 2018). In support of these findings, Horna *et al.*, (2019) also showed that *exo*-U is significantly associated with MDR/XDR phenotype and with quinolone resistance. This should usually be considered within the genetic context of the strains and the associated antibiotic resistance fitness cost (Abdelraouf *et al.*, 2011). It has been shown that susceptible high-risk clones are encountered in different environmental niches (Bel Hadj Ahmed *et al.*, 2019), the fact that draws attention into the importance of investigating underlying genetic markers of success in both resistant and susceptible high-risk clones. Results of studies performed by Recio *et al.*, (2019) have shown that although the associated MDR behavior offers a fitness cost, it appears that the pathogenic or poor prognostic characters of high-risk clones are more related to its virulence genotype.

These findings when combined with the high predictive values identified in the previous chapter support the fact that efflux pump systems make essential contribution to the problematic nature of this species. This conclusion can also be supported by the distribution of susceptibility and resistance behavior of the organism population seen in MST diagram. MST diagrams showed that there is no clustering or clonality in antibiotic resistance behavior for any of the four studied antibiotic agents (amikacin, gentamycin, ciprofloxacin, levofloxacin), the observation that may indicate an unpredictable underlying genomic markers' association with behavior or the presence of confounding epistatic interactions. This leaves us in a situation where there is probably no clear definition of what represents a high-risk clone. These are clones with MDR or pathogenic phenotypes that may also share some common phenotypic features. The clones are also known to carry some MDR-related gene structures. These structures include some mobile genetic elements including integrons, transposons or plasmids in addition to specific pathogenicity islands which are frequently observed with clonality.

Horna *et al.*, (2019) showed that the spread of a successful local clone is unlikely, however, the high-risk behavior can be attributed to the dynamic nature of the species accessory genome that leads to acquisition/loss of *exo*U-encoding genomic islands which co-harbor other integrons carrying many other resistance determinant gene cassettes. The same study has also shown alteration patterns in efflux pump regulators including *mex*Z, *nfx*B, *mex*S, *nal*C, and *nal*D in association with quinolone resistance and high-risk behavior (Horna *et al.*, 2019). The study

suggests that the efflux pump system acts as an important underlying genetic determinant in strains with *exo*U- background while alterations related to efflux-pump regulation may be irrelevant in *exo*U+ background. The analysis performed in this chapter shows that efflux-pump regulators appear to be an important contributor to both quinolone resistance and high-risk behavior which support the findings of Horna *et al.*, (2019), however, the analysis have not tested for *exo*-U.

Although different conclusions can be drawn about the population structure when different typing markers are used especially with polyphasic chromosomal markers that cover the core genome, the results of the current assessment agree with what is currently known about the species. However, the work done in this chapter can be extended to re-assess the population structure using other available typing approaches especially core genome-based typing.

While a panel of molecular resistance markers were investigated in this chapter for their potential association with clonal success which greatly adds to the practical value of using these markers, the presence of specific virulence traits or markers in high-risk clones still needs to be further explored. In the previous chapter, I have investigated the predictive values for the most informative group of molecular diagnostic markers. This chapter has explored the relationship of these markers and the newly identified clusters of molecular markers in relation to high-risk clones. The chapter that follows moves on to explore for additional molecular determinants that can add more to the information and findings at this stage by applying genome wide comparative behavioral genomics. Thus far, the approach used aims at finding molecular markers that can be used as diagnostic to both resistance and/or high-risk behavior. These markers are studied with the aim of being used when enough research findings and reliable workflows become available at the practical level. Such markers can be used to guide or direct infection control procedures.

Chapter.5. Investigation of Whole System determinants of antibiotic resistance in *Ps. aeruginosa* using comparative behavioral genomics

5.6. Results. Section 3. Post-CBG identification of candidate susceptibility/resistance markers for gentamycin in *Ps. aeruginosa* (2nd round of annotation)

5.6.1. Description of significance and distribution of candidate loci across different phenotypes

5.6.2. Diagnostic Performance of candidate loci as potential molecular diagnostic markers

5.6.2.1. Candidate gentamycin markers showing high specificity and high PPV

5.6.2.2. Candidate gentamycin markers showing high NPV

5.6.3. Identification of new combinations of potential molecular markers with improved diagnostic performance

5.6.4. Distribution of candidate loci in relation to genomic background

5.7. Results. Section 4. Post-CBG identification of candidate susceptibility/resistance markers for ciprofloxacin in *Ps. aeruginosa* (2nd round of annotation)

5.7.1. Description of significance and distribution of candidate loci across different phenotypes

5.7.2. Diagnostic Performance of candidate loci as potential molecular diagnostic markers

5.7.2.1. Candidate ciprofloxacin markers showing high specificity and high PPV

5.7.2.2. Candidate ciprofloxacin markers showing higher NPV

5.7.3. Identification of new combinations of potential molecular markers with improved diagnostic performance

5.7.4. Distribution of candidate loci in relation to genomic background

5.8. Results Section 5. Predicted Functional effect of amino acid changes (nSSNPs) variants

5.9. Discussion

5.9.1. Novel insights into system-level functions associated with quinolone susceptibility phenotype

5.9.2. Exploiting GWAS for identifying predictive genomic markers

5.1. Introduction

Bacterial genome wide association studies are considered a promising research approach for the discovery of novel genetic markers and for the detailed assessment of markers' effect. The approach helps to explore the genetic basis of phenotypic variation in the population in a hypothesis free manner. Bacterial genome wide association studies (GWAS) are considered useful for predicting genetic variants associated with the phenotypes of interest using a small set of measures. The approach is considered superior to targeted and traditional genetic screens because it allows thorough testing of all possible elements at the whole system level rather than using a hypothesis-based testing which targets specific sets of genes or specific genomic regions. GWAS test genotype-phenotype associations based on either gene presence/absence pattern or based on SNP associations. The testing aims at exploring significant correlation at different levels of significance (Chen and Shapiro, 2015).

A crucial aim of microbial genomics is to understand the genetic factors underlying different phenotypic properties. In GWAS, genomes are compared to find the genetic markers systematically correlated with the characters of interest. A true signal of association should be differentiated from background noise which may lead to spurious associations. This can usually result from the confounding effect of population structure or from recombination (Power, Parkhill and de Oliveira, 2017). Microbial association studies have some limitations that need to be considered during analysis design. These include four main points. First is the strong population structure that results from microbial clonal reproduction. Second is linkage disequilibrium that can be unpredictably interrupted by homologous recombination. Third is the diversity in gene content and last, but not least, is the variability in the probability of phenotypic distribution for a given genotype (Collins and Didelot, 2018).

In human genetics, rare diseases are usually linked to highly penetrant variants while common traits are linked to multiple less-penetrant variants. There is currently some evidence that shows that the same may apply to bacterial genetics. Lack of strong associations in variants identified using genome-wide analysis is highly suggestive that common traits are affected by multiple alleles with smaller individual effect sizes. It has been shown that many bacterial variants essential for cell survival have very large effect size similar to Mendelian traits in eukaryotes. However, many bacterial traits are influenced by multiple variants with smaller effect sizes.

Their genetic basis can be determined by applying GWAS (Read and Massey, 2014). Effect size and Cluster analysis performed in the workflow applied in the thesis helps to find the best predictor combinations; an approach that can offer a great value to understand complex trait genetics.

Smaller scale genomic studies including methods based on PCR, limited sequence data and comparative genome hybridization were considered among the main approaches used for investigating genotypic-phenotypic relatedness in bacteria till 2010. The increasing availability of bacterial large-scale genomic sequence data opened the door for the increasing use of bacterial GWAS (Read and Massey, 2014).

The first bacterial GWAS was applied to study *Neisseria meningitidis.* It was chosen as a valuable approach to investigate this organism which cause exclusively human diseases with no animal model available. Bille *et al.,* (2005) have developed a gene-based microarray to investigate for the differences between normal nasopharyngeal carriage and invasive isolates causing meningococcal disease. This has brought the light to a chromosomally integrated bacteriophage which can promote the development of epidemic clones and consequently break the commensal relationship of *N. meningitidis* with humans causing invasive disease (Bille *et al.*, 2005). In a later study, Bille *et al.*, (2008) have examined a temperate bacteriophage in 1288 meningococci which were able to confirm a phage associated with hypervirulence. This was over-represented in disease isolates from young adults.

GWAS was also applied to investigate for the genetic factors underlying host adaptation in *Campylobacter jejuni* and *C. coli* and found that vitamin B5 biosynthesis is playing an important role (Sheppard *et al.,* 2013). This was followed by GWAS studies applied to methicillin-resistant *Staphylococcus aureus* clinical isolates used to investigate for vancomycin resistance and also to investigate for the ability of bacteria to lyse human cells in other studies. These studies have identified 121 novel virulence-associated loci (Laabei *et al.*, 2014) as well as novel single nucleotide polymorphism in *rpo*B gene (Alam *et al.,* 2014).

GWAS was also applied to identify polymorphism related to beta-lactam resistance in *Streptococcus pneumoniae* and has also identified multiple novel antibiotic resistance - associated loci (Chewapreecha *et al.*, 2014). Mobegi *et al.*, (2017) have used GWAS to analyze the genome sequence of 1680 *Streptococcus pneumoniae* isolates from four independent

populations and have identified hotspots of genetic variations that correspond to antibiotic resistance phenotype in different antibiotic classes (Mobegi *et al.*, 2017).

GWAS was also applied for the genomic analysis of extraintestinal pathogenic *E. coli* and identified 17 novel factors as resistance determinants for further functional testing. The study showed differential profiles of virulence factor content and antibiotic resistance phenotype among the lineages affecting different body sites (Salipante *et al.*, 2015). Genetic architecture for ecologic interaction of *E. coli* and *Staphylococcus aureus* co-culture was also investigated employing a GWAS approach called Q-ROADTRIPS and reported 66 and 111 SNPs as associated with interactions in *E. coli* and *S. aureus* respectively (He *et al.*, 2017). GWAS was also used to investigate the neural and placental tropism in the foodborne pathogen *Listeria monocytogenes* and identified novel virulent genes in relation to hypervirulent clones causing central nervous system and maternal neonatal listeriosis (Maury *et al.*, 2016).

Resistance to anti-tuberculous drugs was also investigated in *Mycobacterium tuberculosis* using GWAS in different studies. Farhat *et al.*, (2013) have identified 39 novel resistance-related loci. Another study has also identified 75 SNPs in 50 genes as potential new drivers of drug resistance in *Mycobacterium tuberculosis* (Chen and Shapiro, 2015). Large-scale genomic analysis has also been used to investigate the genetic variation underlying meningeal or pulmonary tuberculosis and has identified variations in three separate genes as correlated with the TB phenotype, one of which plays a role in host-pathogen interaction (Ruesen *et al.*, 2018).

Oppong *et al.*, (2019) has also used GWAS to study lineage specific correlations with drug resistance and showed 17 potential novel correlations between antimicrobial resistance phenotype and *Mycobacterial tuberculosis* genomic variants (Oppong *et al.*, 2019). Another recent study has also identified some other novel mutations in bacterial mono-oxygenase-Rv0565c function that contribute to Ethionamide and Prothionamide Susceptibility in *Mycobacterium tuberculosis* (Hicks *et al.*, 2019). GWAS was also applied in *Mycobacterium tuberculosis* to investigate markers related to disease transmissibility and identified genetic markers that confer increased transmissibility *in vivo.* These are related to altered immune response *in vitro* (Nebenzahl-Guimaraes *et al.*, 2017). Genome-wide studies were also used to find correlation between specific pathogen lineages and host genetic risk factors for the disease in TB (Omae *et al.*, 2017) (Zheng *et al.*, 2018). They have also been used to identify specific

correlations between host genetic factors and human gut microbiota composition and showed that host genetic composition can greatly affect bacterial abundance in gut (Davenport *et al.*, 2015).

Earle *et al.*, (2016) have used a new GWAS approach to control for population structure and to recover lineage-level genes and gene variants underlying resistance to 17 antimicrobials in over 3000 isolates from diverse clonal and recombining clinically important bacteria including *Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Escherichia coli*, and *Klebsiella pneumoniae*. Interestingly, GWAS was also used to understand how bacterial genetic variation can influence the risk of developing gastric cancer in *Helicobacter pylori* (Berthenet *et al.*, 2018). The study has been performed on 173 *H. pylori* isolates from European populations including 49 isolates with gastric cancer. The study has identified a relationship between gastric cancer phenotype and some risky genotypes including the presence of *bab*A gene and the cag pathogenicity island. Gori *et al.*, (2019) have recently used pan-GWAS to highlight lineage specific genes related to virulence and niche adaptation in *Streptococcus agalactiae.*

In *Ps. aeruginosa*, Redfern *et al.*, (2019) show the usefulness of different GWAS approaches to improve our understanding about the diversity of different important phenotypic traits for the species including biofilm formation and AMR. A new extended Kmer-based GWAS method has previously suggested novel genotype-phenotype correlations in *Mycobacterium tuberculosis*, *Staphylococcus aureus* and in *Pseudomonas aeruginosa* (Jaillard *et al.*, 2018). *Ps. aeruginosa* has also been previously investigated as an example of a highly plastic genome using a modified alignment free GWAS methodology in another study (Jaillard *et al.*, 2017). Jaillard *et al.*, (2017) have shown the power of GWAS approach to explore for the genetic basis of several antibiotic resistance phenotypes and have identified 46 novel putative resistance-related polymorphisms (Jaillard *et al.*, 2017). In another study, the genetic bases of biofilm production in MDR *Ps. aeruginosa* have been investigated. The study has identified several SNPs in relation to biofilm forming phenotype including the arsenic reduction genes and *lpx*O that encodes a lipid hydroxylase (Redfern *et al.*, 2019).

Beyond investigating pathogenesis, Sutton *et al.*, (2019) showed that GWAS can also be applied to explore the predatory behavior in myxobacteria. By studying 29 myxobacterial genomes including 10 prey organisms including *Ps. aeruginosa*, Sutton *et al.*, (2019) have identified 139

predation genes. *Ps. aeruginosa* showed that 42 genes are associated with its predatory activity (Sutton *et al.*, 2019). The most notable finding in the same study was the demonstration of formaldehyde secretion as a potential predation resistant trait in *Ps. aeruginosa*.

Based on the examples presented above, GWAS has shown promising results and can be considered a high-throughput tool for genetic marker discovery in a wide range of organisms showing different ranges of genome plasticity. As an approach, GWAS allows the identification of significant correlations. This helps to achieve a better understanding of disease etiology by exploring the genetic basis of microbial traits. Understanding the different biologic mechanisms underlying infectious diseases can guide new drugs and vaccine development. It can also help in designing predictive rapid diagnostic tools which guides surveillance for the sake of public health interventions.

5.2. Objectives

This chapter aims at:

- Identification of system-level functions associated with susceptibility phenotype to quinolone group of antibiotics (1st round of CBG annotation).
- Identification of gene variants showing significant association with susceptibility and resistance to ciprofloxacin and gentamycin (2nd round of CBG annotation).
- Testing for the potential practical application of the identified variants for use as molecular diagnostics by evaluating the predictive diagnostic accuracy of candidate markers.
- Examining the functional effect of prioritized gene variants.
- Finding the best combination of variants showing improved diagnostic performance.
- Examining the distribution of candidate alleles in relation to background genomic structure.

5.3. Methods

5.3.1. Phenotype testing

Strains were tested phenotypically using three experimental methods. Modified Stoke's method was chosen for the classification of studied isolates during the first exploratory stage of comparative behavioral genomics (CBG). MIC values were then used to classify isolates into

susceptible and resistant during the next stage of testing predictive values according to clinical breakpoints.

5.3.1.1. Kirby-Bauer Method

Bacterial inoculum of approximately $1.5*10^8$ CFU/mL (corresponding to 0.5 McFarland standard) was applied to the surface of a Mueller-Hinton agar plate.

Commercially prepared, fixed concentration, paper antibiotic disks were then placed on the inoculated agar surface.

Plates were incubated for 18–20 h at 35 ℃ prior to determination of results. The zones of growth inhibition around each of the antibiotic disks were measured to the nearest millimeter.

Detailed experimental protocol was implemented as shown in the following steps according to (Andrews and Howe, 2011)

- Enough molten agar is poured into sterile Petri dishes to give a depth of 4 ± 0.5 mm (25 mL in a 90 mm Petri dish).
- The surface of the agar is dried to remove excess moisture without over-drying.
- Four to five colonies (from a fresh overnight culture) are taken directly from the plate and then suspended in 3 mL broth or saline. in 100*12 mm glass tube to give turbidity that is just visible. The suspension is matched to the density of the 0.5 McFarland standard.
- To inoculate the plate, a sterile cotton-wool swab is dipped into the suspension and excess liquid removed by turning the swab against the side of the tube.
- The inoculum is spread evenly over the entire surface of the plate by swabbing in three directions.
- The plate is then allowed to dry before applying discs. Discs are then applied to the surface of the agar within 15 min of inoculation.
- A disc dispenser is used to apply discs to inoculated agar surface which should be firmly placed to the dried agar surface with even contact.
- Plates are then incubated immediately or within 15 min of disc application.
- For *Pseudomonas* spp., incubation should be at 35–37.8 ℃ temperature in air for 18–20 hours.

- Plates should not be stacked more than six high in the incubator.
- The diameters of zones of inhibition are measured to the nearest millimeter (zone edge is taken as the point of inhibition as judged by the naked eye) with a ruler or a digital caliper.

5.3.1.2. Modified Stokes Method

Stokes' method was known and used for a long time as a routine method of testing antimicrobial sensitivity in most laboratories in the UK. It was listed by BSAC together with MIC as the main methods used for sensitivity testing ('A guide to sensitivity testing. Report of the Working Party on Antibiotic Sensitivity Testing of the British Society for Antimicrobial Chemotherapy.', 1991). The comparative disc diffusion techniques based on Stokes' method for sensitivity testing was replaced by standardized Kirby Baeur disc diffusion protocol from 2001 by the BSAC Working Party Report (Andrews and BSAC Working Party on Susceptibility Testing, 2001) because of technical validity-related drawbacks and also due to some criticism related to the non-plausibility of using the same interpretation criteria to classify sensitivity and resistance to all classes of antibiotics.

Although the method is not currently used in routine clinical practice, it was used in this work to classify strains phenotype for comparative purposes. It was chosen to control for sources of variability that are sometimes encountered during routine susceptibility testing.

Stokes' method is usually done by comparing disc zone diameter of the test strains to a normally susceptible control strain. The original Stokes' Method had interpretive criteria in the clinical laboratory (Potz *et al.*, 2004), however these were not used for the purpose of this research because of many technical drawbacks and lack of validity of such interpretations. For the purpose of this research, the method is modified to be used for the purpose of group comparison without related cutoff interpretations. In the current analysis, zone difference values were used as a scale measurement for sensitivity or resistance behavior for comparative analysis. This would help each individual test isolate to be compared to a control strain under the same technical conditions of the medium composition, incubation time, atmosphere, temperature, and disc content or quality, so both the test and control strain can be measured directly on the same plate.

5.3.1.3. Broth micro-dilution for MIC measurement

The detailed experimental protocol for MIC testing using broth microdilution method was implemented as shown in the following steps according to Wiegand, Hilpert and Hancock, (2008).

- Bacterial isolates to be tested were streaked onto nutrient-rich agar plates without inhibitors and incubated for 18-24 h at 37 ℃ to obtain single colonies.
- For each isolate, 3 to 5 similar colonies were selected from the fresh agar plate and transferred into sterile cation adjusted Muller-Hinton broth using sterile loop or cotton swab and vortexed well to prepare the initial inoculum.
- Inoculum suspension was adjusted to McFarland Standard of 0.5 (OD_{625}=0.08-0.13).
- Preparation and storage of antibiotic stock and working solutions were done according to Andrews, (2001). Details are shown in Appendix I
- Antibiotic solutions were prepared at 10 different concentrations according to recommended test concentrations shown in Andrews, (2001).
- Sterile 96-well microtiter plate was labelled so that 10 different concentrations are added to 10 columns of the microtiter plate. Column 11 was used as growth control well and column 12 as sterility control well.
- 100 µl of sterile broth solution was pipetted into sterility control well (column 12) and 50 µl of sterile broth solution into growth control well (column 11).
- For each tested isolate, 50 µl of each antibiotic dilution was added into the respective well.
- Bacterial suspension adjusted to $1.5*10^8$ CFU/ml was mixed and vortexed well and then diluted 1:100.
- Each well containing the different antibiotic solutions and the growth control well was inoculated with 50 µl of bacterial suspension. This results into the final desired inoculum of $5*10^5$ CFU/ml.
- 10 µl aliquot from growth control well was immediately removed after inoculating the plate and pipetted into a sterile Eppendorf tube holding 990 µl of sterile broth and mixed well by vortexing.

- Further (1:10) dilution of the suspension was made by pipetting 100 µl into 900 µl sterile broth and mixed well.
- 100 µl of each of the two dilutions was plated onto two different antibiotic-free nutrient rich agar plates using sterile cell spreader to spread the liquid.
- Microtiter plates were sealed with gas permeable membrane and then incubated at 37 ℃ for 18-20 hours.
- Agar plates were incubated together with the microtiter plate.
- All experiments were done in triplicate.

Reading Results and MIC determination

For the test to be valid, cell counts plated on agar were checked to verify that the right number of CFU were used.

The presence of around 50 colonies on the lower of the two dilutions (1:1000) of the initial bacterial suspension was expected when the correct bacterial inoculum density of $5*10^5$ CFU/ml was used.

If the cell number was within the desired range, the test could be analyzed to determine the MIC.

Antibiotic-free growth control wells were checked. It was necessary to observe sufficient growth for the test to be valid (a definite turbidity or a sediments-button size > 2mm-in microtiter plates).

Test results cannot be read if the sterility control well (no bacterial inoculum) is turbid. The MIC was defined as the lowest concentration of the antimicrobial agent that inhibits visible growth of the tested isolate as observed with the unaided eye.

5.3.2. Comparative behavioral genomics

A set of genomes of *Ps. aeruginosa* species including 167 lab isolates in addition to the reference *Ps. aeruginosa* isolates PA14 and PAO1 were used as an initial set for annotation. Each genome was assembled using de novo sequence assembly program SPAdes (Bankevich *et al.*, 2012). Assembly quality and downstream sequence analyses were carried out using MUMer (Delcher, 2002), BLAST (Altschul *et al.*, 1990), and in-house perl scripts. The first round of genome annotation was performed using known genes with NCBI annotation and in-house pipeline of

annotation to find consensus gene sequences which are then used to resolve a search element for each gene. This was used to compose a library of genes for the species. During round 2 of genome annotation, the gene search elements were used to search back the genomes in order to resolve the alleles for each gene. During this step, allele consolidation was revised manually. Consolidated alleles were used to revise the gene library generated and to re-annotate all consistent core SNPs in all samples. This was then used to define the core genome and the diversity of the collection was defined based on the core genome. During round 3, the most diverse set of genomes based on core genome annotation was chosen as an optimized set to apply CBG analysis. The dataset included in this study showed a total of 877,218 SNPs in the directly comparable pan and core genomes. The isolate distance measure used to ascertain diversity was pairwise SNP distance of an isolate from its nearest neighbor in the core genome. The isolates had an average nearest neighbor distance of 18,213 SNPs, and the isolates with the minimum and maximum distance being 13,731 and 63,011 SNPs, respectively. Based on the identified diverse set of isolates, a matrix of gene content (gene presence/absence) and a SNP ID matrix was generated. CBG is then performed to find statistically significant association of different genes and SNPs using Fischer exact at 0.05 level of significance. Output from CBG shows lists of significant correlations for each of the genes and SNPs at different levels of significance. Venn diagrams were used to visualize co-shared variants among different pair-wise comparisons (Bardou *et al.*, 2014).

5.3.3. Post-CBG statistical and functional analysis

5.3.3.1. Correspondence analysis, Effect sizes and predictive values

This stage of analysis aimed at testing practical and applied significance of prioritized molecular markers.

Statistical analysis: In order to filter the output and to find the most important SNPs candidates, statistical techniques used to investigate for practical significance including phi-coefficient, effect sizes, and confidence interval risk estimation were calculated using SPSS. To find the deviation of different identified genomic elements from independence, principle component analysis was used. In order to suit the type of study variables which are qualitative binomial data, correspondence analysis was applied using SPSS.

Functional analysis: MEGA software was used to map nucleotide changes in individual genes to PA-14 as a reference nucleotide changes. After that, functional analysis was applied to predict the functional effect of observed amino acid substitutions resulting from non-synonymous mutations on gene function. To do so, PROVEAN (Protein Variation Effect Analyzer) was applied for that purpose as described previously in Methods section 3.4.3

Analysis of both effect sizes and predictive values was applied to a larger population of isolates. The population is more inclusive of different phenotypes. At the second stage of CBG annotation, studied strains were classified into susceptible and resistant groups. This classification was then used for diagnostic accuracy testing. This type of case/control binning was necessary for predictive values' calculations which was used in the 2nd round of CBG annotation. The binary phenotypes were determined using latest EUCAST guidelines (EUCAST, 2018). Strains with a gentamycin MIC > 4 were classified as resistant. Using these cutoffs across the studied group of strains classified 112 strains as susceptible and 50 strains as resistant. Strains with ciprofloxacin MIC > 0.5 were classified as resistant. Using these cutoffs across the studied group of strains classified 157 strains as susceptible and only 5 strains as resistant.

Candidate gene variants were prioritized using primary CBG output sheets based on correspondence analysis and were re-identified in 162 isolates using NCBI BLAST (Cock *et al.*, 2015) . Blast output was visualized and analyzed for identifying SNP distribution using MEGA software (Kumar, Stecher and Tamura, 2016). Statistics including principle component analysis, predictive values, significance ,and effect sizes were implemented using IBM SPSS statistics (SPSS.V21).

5.3.3.2. Testing for the functional effect of identified gene variants

Protein sequence information of prioritized genes showing the evaluated variants were retrieved from *Pseudomonas* genome database (Winsor *et al.*, 2009). Possible functional effect of amino acid changes identified in the variants of interest was evaluated using PROVEAN (Protein Variation Effect Analyzer) available at http://provean.jcvi.org./index.php (Choi and Chan, 2015) and I-Mutant v2.0 (Predictor of Protein Stability Changes upon Mutations) available at http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi (Capriotti, Fariselli and Casadio, 2005).

5.3.3.3. Cluster analysis used to find best possible combination of molecular markers

Two-step cluster analysis algorithm developed by (Chiu *et al.*, 2001) for distance-based analysis of large datasets was implemented using IBM SPSS statistics (SPSS.V21) . Different gene variants of interest were used as input to the clustering function implemented in SPSS to find clusters predictive of a good model in relation to antibiotic resistance/susceptibility phenotype. Mixed type attributes can be handled using this function and the number of clusters is automatically determined (Bacher, Wenzig and Vogler, 2004).

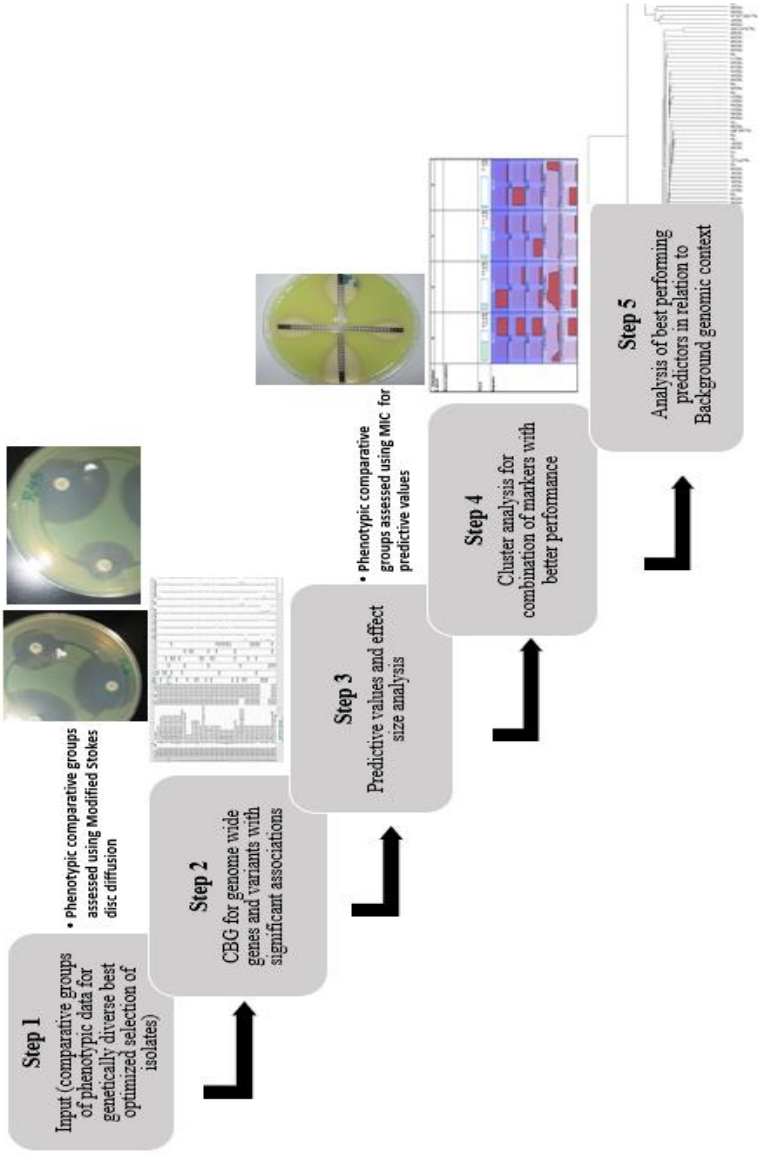
5.3.3.4. Background genomic context

The CSI phylogeny tool by the DTU available at https://cge.cbs.dtu.dk/services/CSIPhylogeny/ was used to construct whole genome SNP-based dendogram (Kaas *et al.*, 2014). SAMtools (H. Li *et al.*, 2009) were used for sequence alignment and mapping and MuMmer (Delcher *et al.*, 2002) was used for large scale genome alignment and comparison. Concatenated alignment of high-quality SNPs was then used to infer phylogeny. The minimum depth at any SNP position used was 10x and the minimum distance between SNPs was 10bp. Approximate maximum-likelihood for large alignments as implemented in Fast tree (Price, Dehal and Arkin, 2010) was used to infer phylogeny.

To test for the effect of recombination on phylogeny, RDP was used to re-test for clonal relationship in studied isolates while accounting for recombinational events (Martin *et al.*, 2010). RDP4 recombination detection program was used for detection and visualization of recombination in whole genome sequence alignments of 162 of tested *Ps. aeruginosa* isolates. Whole genome sequence alignment of 162 isolates was generated using Burrows-Wheeler Transform-Based Mapping Algorithms (Li and Durbin, 2009) and (Heng Li *et al.*, 2009). The aligned sequences were used as an input to the software. The default exploratory method of the software was used to scan for recombination signal in genomes alignment including the primary scanning algorithms; RDP (Martin and Rybicki, 2000), MaxCHI (Smith, 1992) and GENECONV (Padidam, Sawyer and Fauquet, 1999) which were used to characterize patterns of recombination without any prior information. This was followed by BOOTSCAN (Martin *et al.*, 2005) and SISSCAN (Gibbs, Armstrong and Gibbs, 2000) algorithms to check for the recombination signal detected using other exploratory methods.

Analysis of recombination patterns in whole genome alignments was performed to minimize the possible disruptive impact recombinant segments may introduce on inferring phylogeny. Therefore, it was important to identify recombination breakpoints and to re-analyze phylogeny based on genome regions unbroken by recombination breakpoints. To do so, Maximum likelihood tree algorithm (Price, Dehal and Arkin, 2010) was used to re-draw the phylogenetic tree after excluding regions of recombination.

Phandango interactive web application available at www.phandango.net was used to explore and to visualize markers' distribution in relation to background genomic context (Hadfield *et al.*, 2018). The stages and method workflow are summarized in the following diagram.



5.4. Results. Section 1. Underlying resistance/susceptibility determinants in quinolones (individual and co-shared resistance determinants for ciprofloxacin/norfloxacin/ofloxacin)

5.4.1. Phenotypic correlations of optimized CBG test group for the quinolone group

This part of the analysis aimed at finding significant correlation between quinolone susceptibility phenotype and the underlying genes/gene variants. The phenotype was expressed using the difference in zone diameter. This was expressed in millimeter using a susceptible control strain. This is shown in method section 5.3.1.2 using modified stokes method. Binning of phenotypic groups was used for comparison. This is based on the distribution of phenotypic data as shown in Figure 5.1, Figure 5.2, and Figure 5.3. Modified stokes' method measures the difference in zone diameter in mm by subtracting reading of the test isolate out of the control isolate. Overlapping MIC and/or zone diameter data were additionally used to resolve the BINs. The BINs were based on phenotypic readings using the three methods. These included modified stokes, MIC, and zone diameter. Strains were primarily selected based on genotypic diversity which were then phenotypically measured. The following graphs show isolates included in different BINs and were used as input to CBG analysis. Readings shown in the graph are based on the phenotypic data measured using modified stokes method.

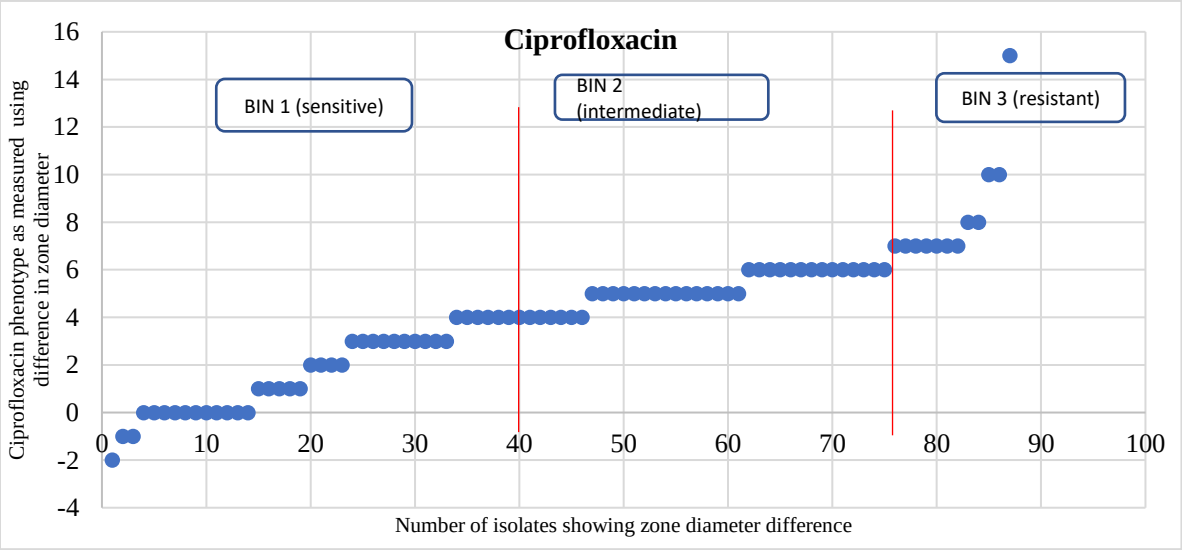


Figure 5.1. Distribution of ciprofloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

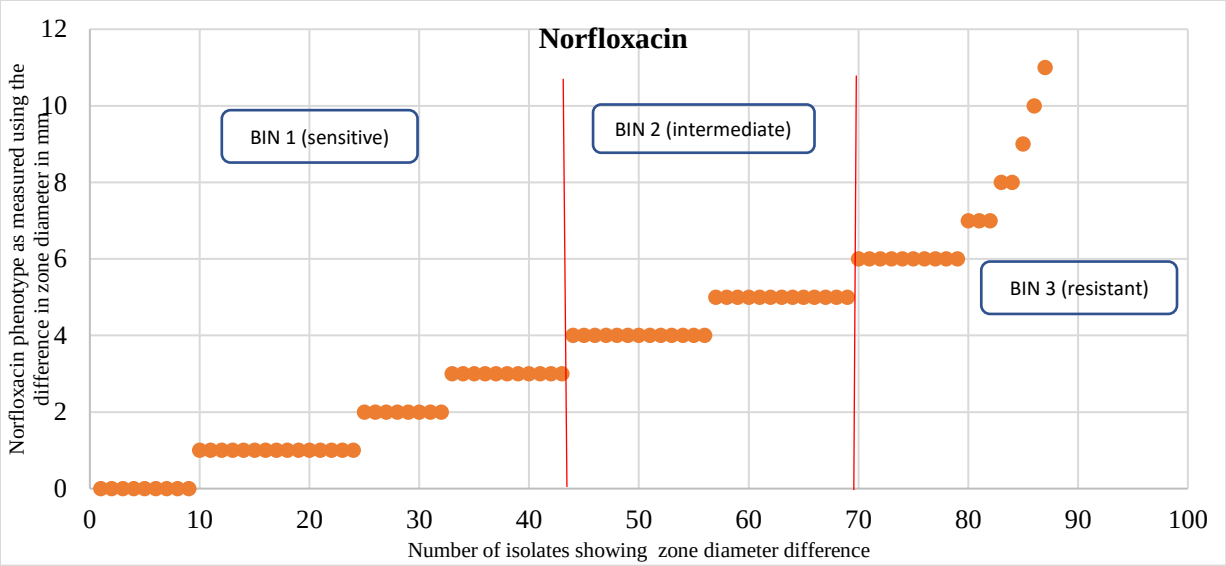


Figure 5.2. Distribution of norfloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

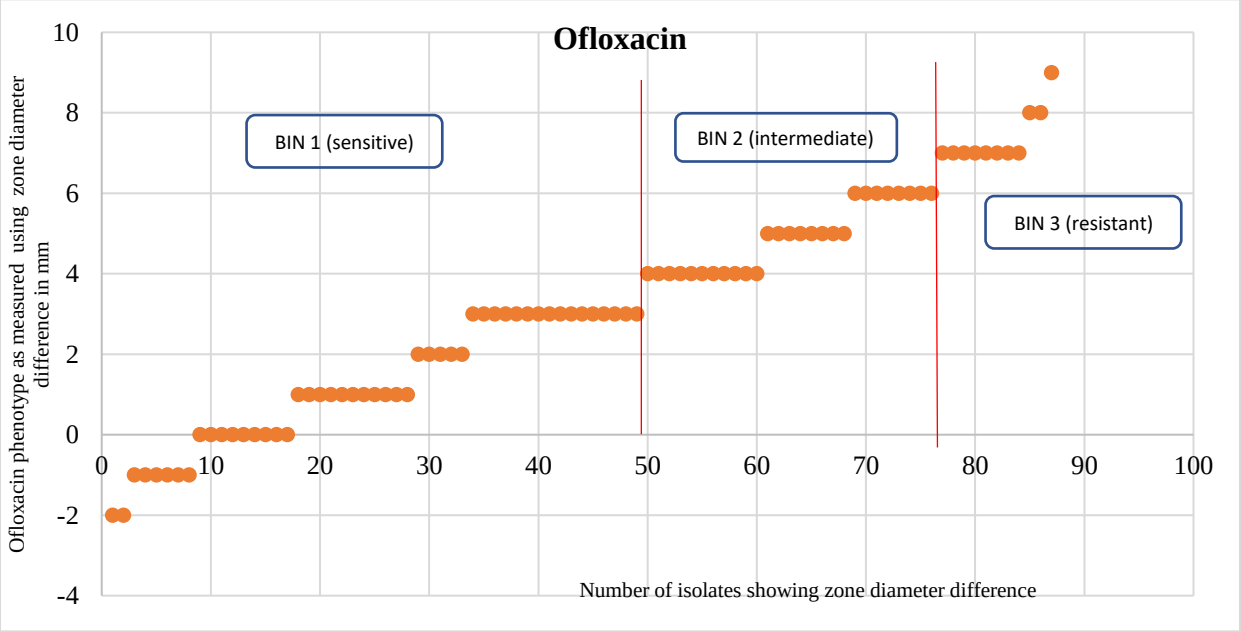


Figure 5.3. Distribution of ofloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

Another analysis was performed to find correlations for the three quinolone agents analyzed using CBG. This was performed in order to find the co-shared and the unique determinants for each agent within the same antibiotic class. This is shown in Figure 5.4, Figure 5.5 and Figure 5.6

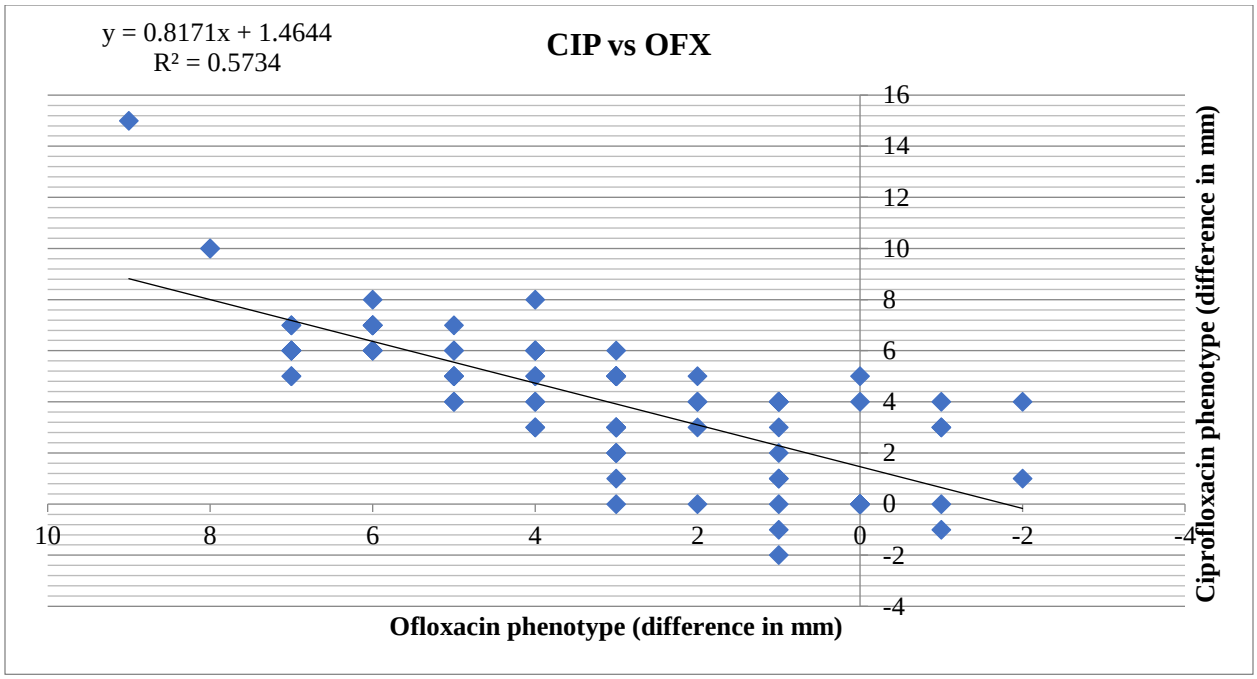


Figure 5.4. Correlation between Ciprofloxacin and Ofloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

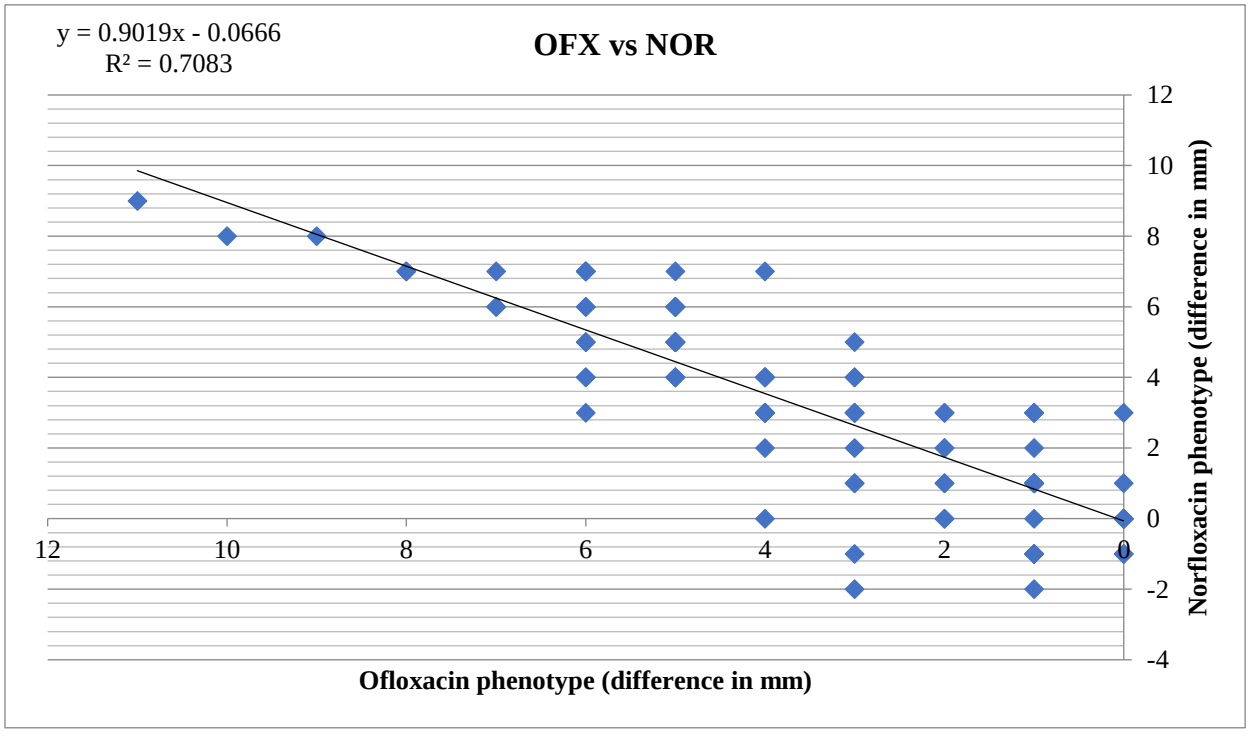


Figure 5.5. Correlation between Norfloxacin and Ofloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

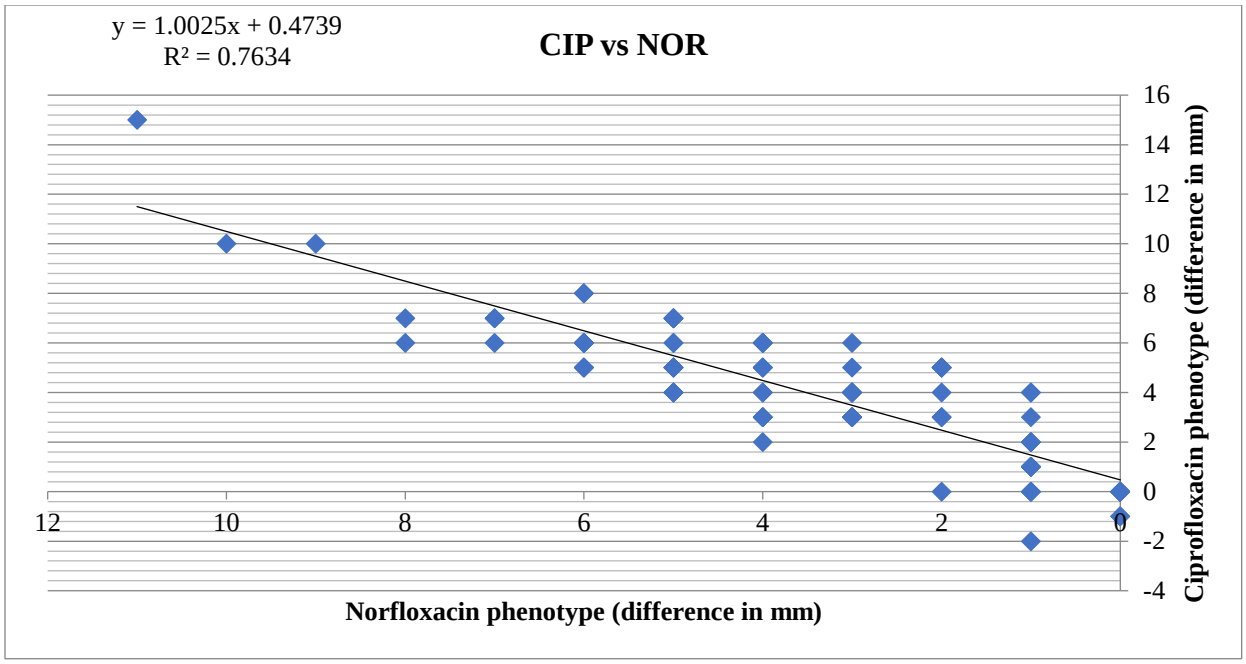


Figure 5.6. Correlation between Norfloxacin and Ciprofloxacin phenotype in 87 optimized CBG set of Ps. aeruginosa isolates

5.4.2. Co-shared resistance determinants for the quinolone group

Figure 5.4, Figure 5.5, and Figure 5.6 show that norfloxacin/ciprofloxacin correlation and norfloxacin/ofloxacin correlation are higher than ciprofloxacin/ofloxacin correlation. To identify co-shared variants/ genes for each of the three quinolone agents, primary CBG sheets were manually and visually filtered to remove significant correlations for intermediate BINs and to include extremes of behavior. This has included the most sensitive and the most resistant at the first stage. Phi coefficient was then calculated for the remaining variants. Genes showing the highest effect size and those co-occurring among the three quinolones agents are shown in Table 5.1, Table 5.2, and Table 5.3 with their related-statistical values.

Table 5.1. Prioritized genes showing best performance for ciprofloxacin resistance and susceptibility phenotype

Ciprofloxacin	Fisher exact sig	Chi square sig	Phi coefficient value	Odds ratio value for gene (presence-absence)	95% CI	Odds ratio value for cohort behavior susceptible	95% CI	Odds ratio value for cohort behavior resistant	95%CI
Alcohol dehydrogenase	0.046	0.035	-0.301	0.283	(0.086-0.937)	0.531	(0.292-0.963)	1.872	(0.974-3.597)
Gluconate-5-dehydrogenase	0.013	0.006	0.392	6	(1.566-22.989)	2.250	(1.286-3.936)	0.375	(0.155-0.906)
Glutaredoxin arsenate reductase	0.072	0.052	0.280	3.333	(0.967-11.487)	1.824	(1.010-3.291)	0.547	(0.273-1.096)
Glutathione-S-transferase	0.082	0.069	0.259	3.056	(0.897-10.406)	1.725	(0.977-3.047)	0.565	(0.281-1.134)
***p*-benzoquinone reductase**	0.082	0.069	0.259	3.056	(0.897-10.406)	1.725	(0.977-3.047)	0.565	(0.281-1.134)
Zinc type alcohol dehydrogenase like	0.082	0.069	0.259	3.056	(0.897-10.406)	1.725	(0.977-3.047)	0.565	(0.281-1.134)
Lipoprotein signal peptidase	0.057	0.036	0.302	4.038	(1.047-15.581)	1.868	(1.088-3.2080)	0.463	(0.194-1.104)
Transcriptional regulator *Znt*r	0.119	0.066	0.263	3.231	(0.901-11.586)	1.744	(0.999-3.044)	0.540	(0.252-1.157)
***Mbt*B**	0.052	0.036	0.303	4.444	(1.039-19.015)	2.476	(0.881-6.958)	0.557	(0.343-0.905)
Chromosome partitioning ATPase *soj*	0.032	0.019	0.335	6.400	(1.195-34.285)	2.080	(1.257-3.442)	0.325	(0.092-1.151)
NAD(P)H azoreductase	0.054	0.034	0.313	4.167	(1.067-16.277)	1.905	(1.092-3.323)	0.457	(0.191-1.092)
Gramicidin-S-synthase	0.096	0.057	0.272			2.300	(1.654-3.198)		
Vit B12 transporter	0.205	0.124	0.220	2.700	(0.747-9.764)	1.607	(0.915-2.822)	0.595	(0.280-1.264)
Ethanolamine ammonia lyase heavy chain	0.041	0.026	-0.317	0.113	(0.012-1.028)	0.472	(0.294-0.760)	4.167	(0.667-26.017)
Ubiquinone -O-methyl transferase	0.355	0.252	0.167	2.036	(0.599-6.922)	1.453	(0.783-2.698)	0.714	(0.383-1.330)

P-values highlighted in red shows the analysis method chosen. This was based on the suitability of BINs for the analysis.

Odds ratio for cohort behavior susceptible > 1 indicates that the gene tends to occur more frequently in susceptible isolates

Odds ratio for cohort behavior resistant >1 shows that the gene tends to occur more frequently in resistant isolates.

Confidence interval for any parameter shows the range of values which is likely to contain the parameter of interest in the population.

Table 5.2. Prioritized genes showing best performance for norfloxacin resistance and susceptibility phenotype

Norfloxacin	Fisher exact sig	Chi square sig	Phi coefficient value	Odds ratio value for gene (presence-absence)	95% CI	Odds ratio value for cohort behavior susceptible	95% CI	Odds ratio value for cohort behavior resistant	95%CI
Alcohol dehydrogenase	0.009	0.004	-0.362	0.211	(0.070-0.635)	0.488	(0.296-0.806)	2.317	(1.183-4.535)
Gluconate-5-dehydrogenase	0.003	0.002	0.388	5.893	(1.807-19.225)	2.112	(1.330-3.254)	0.358	(0.160-0.801)
Glutaredoxin arsenate reductase	0.056	0.030	0.276	3.431	(1.100-10.704)	1.729	(1.086-2.755)	0.504	(0.247-1.030)

NAD(P)H dehydrogenase (quinone)	0.017	0.011	0.321	4.167	(1.348-12.882)	1.864	(1.176-2.953)	0.447	(0.217-0.923)
***Mbt*B**	0.002	0.001	0.406	7	(1.968-24.895)	3.093	(1.260-7.591)	0.442	(0.276-0.707)
Chromosome partitioning ATPase *soj*	0.129	0.080	0.221	3.068	(0.845-11.134)	1.591	(1.011-2.505)	0.519	(0.218-1.233)
NAD(P)H azoreductase	0.015	0.011	0.323	4.412	(1.349-14.428)	1.853	(1.187-2.893)	0.420	(0.189-0.934)
Gramicidin-S-synthase	0.196	0.094	0.211	5.556	(0.610-50.597)	1.759	(1.121-2.760)	0.317	(0.052-1.927)
Ethanolamine ammonia lyase heavy chain	0.026	0.014	-0.311	0.1	(0.012-0.856)	0.500	(0.343-0.729)	5	(0.776-32.222)
Ubiquinone -O-methyl transferase	0.041	0.032	-0.270	0.260	(0.072-0.934)	0.457	(0.191-1.094)	1.76	(1.119-2.768)

P-values highlighted in red shows the analysis method chosen. This was based on the suitability of BINs for the analysis.

Odds ratio for cohort behavior susceptible > 1 indicates that the gene tends to occur more frequently in susceptible isolates

Odds ratio for cohort behavior resistant >1 shows that the gene tends to occur more frequently in resistant isolates.

Confidence interval for any parameter shows the range of values which is likely to contain the parameter of interest in the population.

Table 5.3. Prioritized genes showing best performance for ofloxacin resistance and susceptibility phenotype

Ofloxacin	**Fisher exact sig**	**Chi square sig**	**Phi coefficient value**	**Odds ratio value for gene (presence-absence)**	**95% CI**	**Odds ratio value for cohort behavior susceptible**	**95% CI**	**Odds ratio value for cohort behavior resistant**	**95%CI**
Alcohol dehydrogenase	0.118	0.077	-0.228	0.393	(0.138-1.117)	0.653	(0.398-1.072)	1.662	(0.934-2.958)
Gluconate-5-dehydrogenase	0.299	0.194	-0.168	0.507	(0.180-1.422)	0.737	(0.461-1.178)	1.455	(0.817-2.591)
Glutaredoxin arsenate reductase	0.032	0.028	0.284	3.667	(1.117-12.034)	1.667	(1.089-2.552)	0.455	(0.202-1.021)
Glutathione-S-transferase	0.032	0.028	0.284	3.667	(1.117-12.034)	1.667	(1.089-2.552)	0.455	(0.202-1.021)
NAD(P)H dehydrogenase (quinone)	0.116	0.082	0.224	5.778	(0.650-51.339)	1.683	(1.126-2.514)	0.291	(0.046-1.825)
Zinc type alcohol dehydrogenase like	0.069	0.053	0.25	2.850	(0.974-8.336)	1.569	(0.994-2.478)	0.551	(0.288-1.054)
Lipoprotein signal peptidase	0.081	0.047	-0.257	0.138	(0.015-1.259)	0.281	(0.046-1.706)	2.045	(1.264-3.309)
Transcriptional regulator *Zntr*	0.028	0.017	0.316	3.818	(1.241-11.752)	1.886	(1.038-3.424)	0.494	(0.276-0.884)
Chromosome partitioning ATPase *soj*	0.768	0.653	0.058	1.313	(0.400-4.303)	1.125	(0.685-1.846)	0.857	(0.428-1.716)
NAD(P)H azoreductase	0.028	0.015	0.313	4.141	(1.264-13.572)	1.748	(1.137-2.686)	0.422	(0.187-0.952)
Gramicidin-S-synthase									
Vit B12 transporter	0.168	0.099	0.213	2.579	(0.825-8.064)	1.474	(0.955-2.275)	0.571	(0.275-1.188)
***Mnm*C**	0.754	0.550	0.078	1.467	(0.417-5.161)	1.179	(0.707-1.966)	0.804	(0.380-1.704)
Ubiquinone -O-methyl transferase	0.063	0.047	0.256	7	(0.804-60.981)	1.750	(1.200-2.552)	0.250	(0.039-1.595)

P-values highlighted in red shows the analysis method chosen. This was based on the suitability of BINs for the analysis.

Odds ratio for cohort behavior susceptible > 1 indicates that the gene tends to occur more frequently in susceptible isolates

Odds ratio for cohort behavior resistant >1 shows that the gene tends to occur more frequently in resistant isolates.

Confidence interval for any parameter shows the range of values which is likely to contain the parameter of interest in the population.

Results of statistical analysis including effect sizes, Fischer exact/Chi square, and risk estimate measures are shown in Table 5.1, Table 5.2, and Table 5.3. These were used to extract the genes shared among the three quinolone agents. These are illustrated in Figure 5.7.

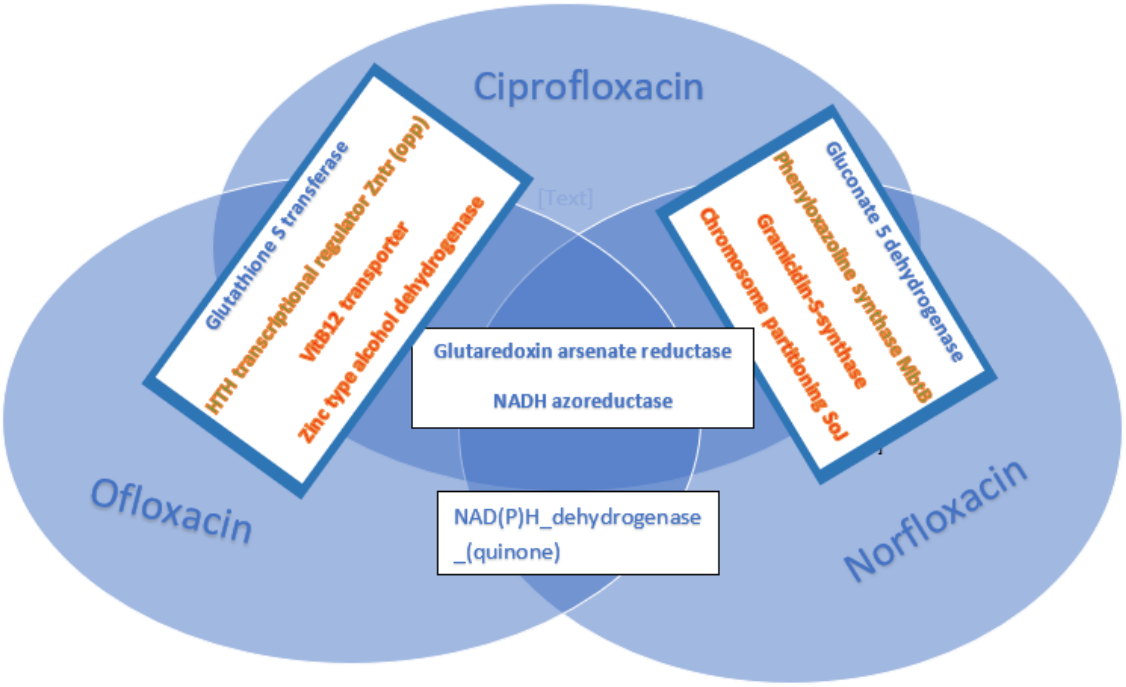


Figure 5.7. Shared genes showing best performance among the three studied quinolones agents; ciprofloxacin, norfloxacin and ofloxacin

These shared quinolones genes showed significant associations at extremes of phenotypic behavior and lacked significance at intermediate groups. Risk estimates as shown in Table 5.1, 5.2, and 5.3 also support the role of these genes. Applying correspondence analysis has also shown that some of these genes were also linked to some behavioral BINs as shown in Figure 5.8 highlighted in the red circles.

Correspondence analysis was performed in parallel. At this stage, finer BINs were used to re-divide the phenotypic behavior into four groups including sensitive, intermediate 1, intermediate 2, and resistant. This was performed in order to observe the signal of deviation. The first analysis was implemented using three-BIN classification and then the second analysis was performed using the finer four-BIN classification. This aimed at finding non-observable patterns using larger BINs.

Results of Correspondence analysis representing deviation from independence for the genes showing significant association within different quinolones' phenotypic categories are shown in Figure 5.8. Correspondence analysis summarizes row and column profiles into relative frequencies. It calculates co-ordinates representing the distance or the differential proportions between row entries that represent the gene/gene variant and column entries that represent presence/absence in different phenotypic groups. This can be used as an indication of association.

The distance between lines shown in Figure 5.8 represents the distance between individual row and column profiles and the line length represent the distance to the average row and column profiles which is shown as zero point. The shorter the line and the closer to zero, the less important the difference in the relative frequency is. More important differences show longer lines which represents the largest deviation from the expected average in proportion. This allows the visualization of how groups are related in addition to visualizing significant correlations resulting from the first stage of analysis.

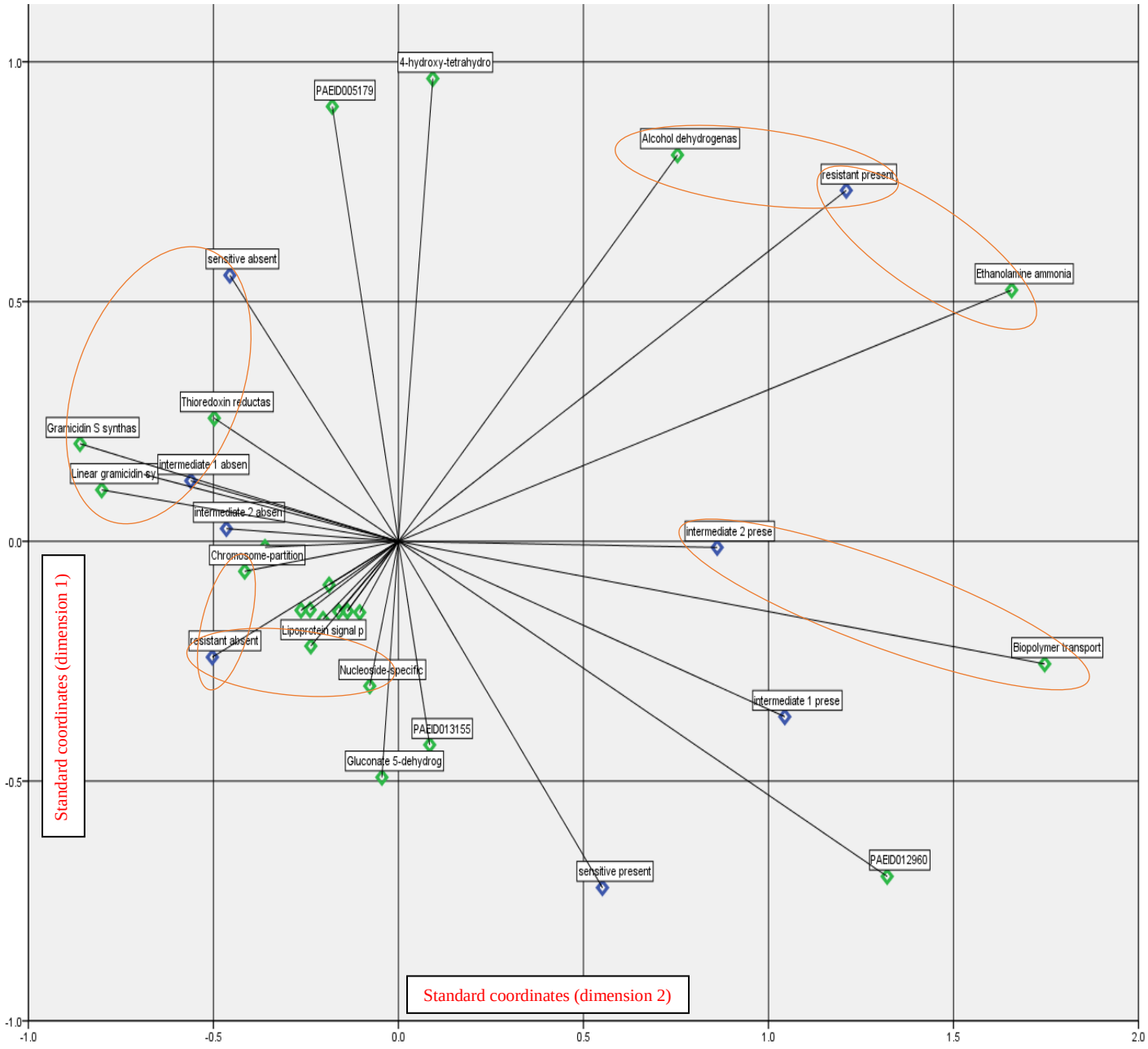


Figure 5.8. Correspondence analysis showing 2-dimension deviation from independence for the best performing quinolone genes

Green diamonds show the genes included in the analysis

Blue diamonds show different phenotypic categories.

Red circles show possible links between genes and phenotypic groups representing deviation from independence.

Based on the 2nd stage Binning of phenotypes into four BINs/groups which has been identified from results of correspondence analysis shown in Figure 5.8, the distribution of isolates among the-four BINs for each gene is shown in detail in Figure 5.9.

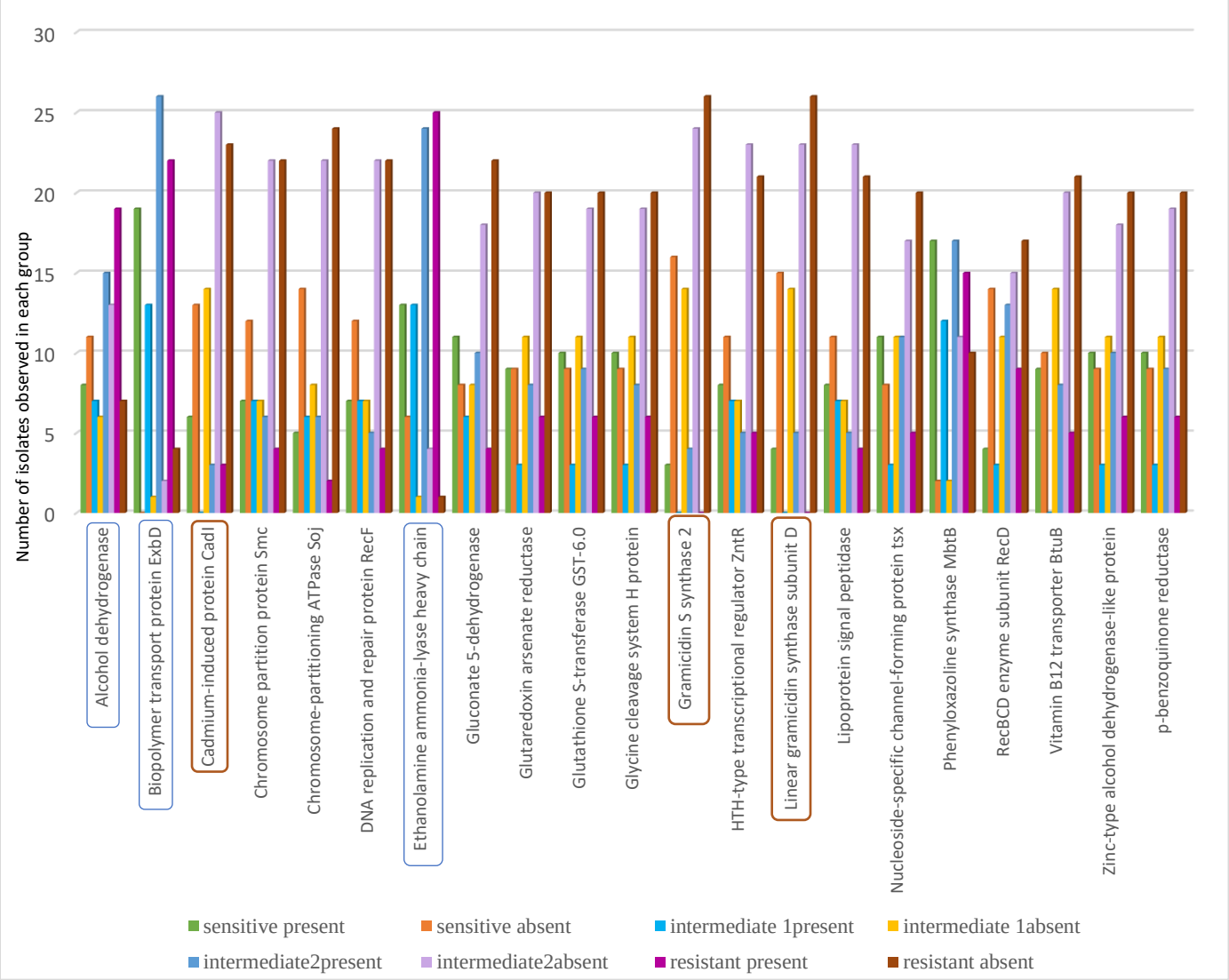


Figure 5.9. Distribution of genes showing best performance among different categories of susceptibility behavior

Genes in blue rectangles show tendency to present at more resistant groups (resistant and intermediate 2 groups). The number of isolates occurring at each of these groups is represented in columns.

Genes in brown rectangles show tendency to be absent at more susceptible groups (susceptible and intermediate 1 groups). The number of isolates occurring at each of these groups is represented in columns.

In conclusion, genes/variants shared among different quinolones agents were identified using the analyses presented in Table 5.1, Table 5.2, Table 5.3, Figure 5.8, and Figure 5.9. Based on the combined findings of all analyses performed, chromosome partitioning *soj*, Gramicidin-S-synthase, glutathione-S-transferase, azoreductases, and *p*-benzoquinone reductases all tend to correlate with susceptibility behavior while both alcohol dehydrogenase and ethanolamine ammonia lyase tend to correlate with resistance phenotype. In addition, the three genes related to susceptibility showed exactly the same distribution. These include NAD(P)H azoreductase, glutathione-S-transferase, and p-benzoquinone reductases. The possible role of these genes will be shown in detail in discussion section.

5.5. Results. Section 2. System-level functional determinants of antibiotic resistance to ciprofloxacin identified using CBG (1st round of annotation)

The same approach and steps shown in methodology section were applied to identify a group of genes/gene variants belonging to some novel functional pathways. Gene variants appear to be linked to resistance. These gene variants and functional pathways are illustrated in this section.

5.5.1. Susceptibility/resistance determinant mutations identified by comparative behavioral genomics

Table 5.4 shows single nucleotide changes and statistical values expressing their practical significance. Table 5.5 shows the predicted functional effect of nucleotide changes using PROVEAN. Sequences of genes from PA-14 was used as a reference.

Table 5.4. Variants showing significance in relation to susceptibility and resistance identified using comparative behavioral genomics

SNP position	Gene	Reference PA14 SNP	Isolates SNP change	PA14 amino acid	Isolates amino acid change	P-value (chi square or Fischer exact)	Phi-coefficient	Risk estimate (confidence Interval)	Lower Limit	Upper Limit
snp_118	*trp*E (Anthranilate synthase component 1)	A	G	Ala 40	Thr	0.003	0.383	3.096	1.234	7.767

snp_10	*phn*A (Anthranilate synthase component 1_ pyocyanine specific)	G	A	Gly 4	Arg	0.088	-0.255	1.8	0.937	3.459
snp_118	*phn*A (Anthranilate synthase component 1_ pyocyanine specific)	G	A	Met 40	Val	0.002	0.407	3.035	1.228	7.499
snp_11	*ant*A (Anthranilate 1_2-dioxygenase large subunit)	C	A	Thr 4	Asn	0.003	0.383	3.096	1.234	7.767
snp_800	*pqs*A (Anthranilate--CoA ligase)	C	A	Ala 267	Glu	0.211	0.188	1.816	0.569	5.794
snp_366	*glt*B (Glutamate synthase [NADPH] large chain)	A	C	Lys 122	Asn	0.013	0.32	2.085	1.038	4.189
snp_1638	*gdh*B (NAD-specific glutamate dehydrogenase)	C	G	Leu 546	Phe	0.039	0.269	1.837	0.844	3.997
snp_1201	*gsh*A (Glutamate--cysteine ligase)	C	A	Lys 401	Gln	0.000	0.497	2.400	1.568	3.674
snp_1130	*msu*D (Methanesulfonate monooxygenase)	C	T	Ala 377	Val	0.001	0.426	4	1.367	11.703
snp_936	*msu*D (Methanesulfonate monooxygenase	G	C	Glutamic acid 312	Aspartic acid	0.257	-0.169	1.482	0.759	2.895
snp_749	*ssu*D (Alkanesulfonate monooxygenase)	T	G	Leu 250	Arg	0.001	0.502	0.316	0.198	0.504

Table 5.5. Predicted functional effect of observed nucleotide changes using PROVEAN

Protein variant	**PROVEAN score**	**Prediction (cutoff)**
*gdh*B, Leu 546 Phe	-3.752	Deleterious (-1.3)
*glt*B, Lys 122 Asn	-0.479	Neutral (-1.3)
*gsh*A, Lys 401 Gln	-0.168	Neutral (-1.3)
*pqs*A, Ala 267 Glu	0.568	Neutral (-1.3)
*ant*A, Thr 4 Asn	-0.946	Neutral (-1.3)
*phn*A, Gly 4 Arg	0.429	Neutral (-1.3)
*phn*A, Met 40 Val	-1.324	Deleterious (-1.3)

*trp*E, Ala 40 Thr	-1.984	Deleterious (-1.3)
*msu*D Ala 377 Val	-1.194	Neutral (-1.3)

PROVEAN score lower than the prediction cutoff used (-1.3) is predicted as deleterious

PROVEAN score higher than the prediction cutoff used (-1.3) is predicted as neutral

Table 5.6 shows the predictive values of the identified genes/gene variants in relation to ciprofloxacin susceptibility.

Tables 5.6(A-B). Ciprofloxacin susceptibility and resistance markers identified from 1st round of CBG annotation

(A) Susceptibility markers				
	Ciprofloxacin			
	Specificity		Negative predictive value	
	Zone data	MIC data	zone data	MIC data
Azoreductases	**71.20**	**69.6**	**92.90**	**97.4**
Glutathione_S-transferase_GST-6.0	**71.20**	**69.6**	**92.90**	**97.4**
p-benzoquinone reductase	**72.60**	**70.2**	**93.00**	**97.4**
***trp*E (Ala40Thr)**	**66.20**	**82.6**	**84.70**	**97.1**
***ant*A (Thr4Asn)**	**82.60**	**82.6**	**84.40**	**98.5**
***msu*D (Ala377Val)**	**90.60**	**90.1**	**83.30**	**96.7**
***gdh*B (Leu546Phe)**	62.30	57.1	90.50	96.8

(B) Resistance markers				
	Ciprofloxacin			
	Sensitivity		Negative predictive value	
	Zone data	MIC data	Zone data	MIC data
Glutarate-semialdehyde dehydrogenase DavD	**71.4**	**80**	**78.9**	**97.4**
Alcohol dehydrogenase	**60.70**	**82.30**	**40**	**95.2**
***phn*A (Gly4Arg)**	**75.00**	**87.00**	**60**	**96.3**
***pqs*A (Ala267Glu)**	**85.70**	**84.60**	**100**	**100**
***mur*I (Arg25Arg)**	**83.30**	**91.70**	**100**	**100**
***gdh*B (Leu546Phe)**	**86.40**	**40.00**	90.50	96.8

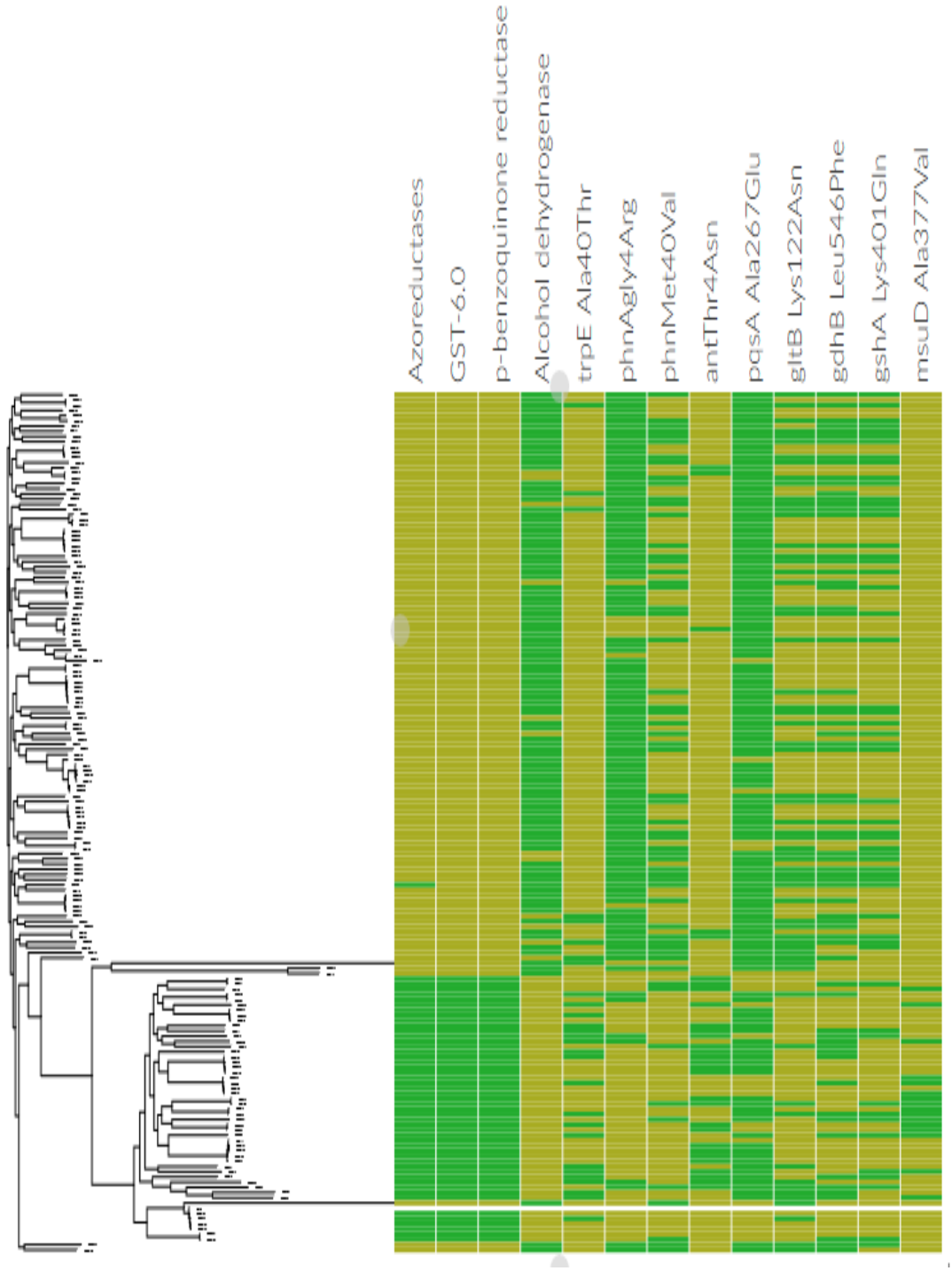


Figure 5.10. Distribution of CBG markers in relation to background context

Green color represents marker/gene presence, yellow color represents marker/gene absence

The results shown in this section hypothesize a probable role of these functional pathways in cell survival to stress. Different functional pathways appear to offer complementary roles in determining resistance phenotypes. The distribution of different genetic markers in relation to the background genomic context may point to that. Three of the genes identified as probable susceptibility determinants were co-absent across the upper section of the tree and co-present across the lower section. Absence of these genes is assumed to make the cell susceptible to xenobiotic agents including antibiotics. This will be shown in detail in discussion section. However, it appears from the distribution of markers as shown on the phylogenetic tree (Figure 5.10) that alternative pathways may be acting in that case to protect the cell. This may include the *Pseudomonas* quinolone signal pathway as will be discussed later. It can also involve a possible role of alcohol dehydrogenase. Predictive values shown in Table 5.6 give additional support to the observed distribution of markers in relation to phylogeny. This is represented by the higher specificity and higher NPV for susceptibility-related markers which cluster across the upper section of the phylogenetic tree. These include azoreductases, GST-6.0, p-benzoquinone reductase, *trp*E Ala40Thr and *ant*A Thr4Asn. In addition, resistance markers summarized in Table 5.6 showed an opposite clustering to susceptibility markers. This indicates that predictive values are complementary to other statistical measures and to phylogeny to understand the biologic behavior in relation to background genomic context.

5.5.2. Functional pathways with compounds and gene changes linked to resistance (Secondary intermediary metabolism)

5.5.2.1. Anthranilate biosynthesis and *Pseudomonas* quinolone signal

Based on the polymorphisms identified in the genes from the anthranilate biosynthetic pathway, changes in anthranilate biosynthesis appear to be significantly correlated with resistance behavior. This assumption was based on variants showing significant correlations with quinolone resistance phenotype and these are shown in Table 5.4. Anthranilate biosynthesis occurs from two different sources including chorismite conversion and tryptophan degradation. Parts of these two pathways are illustrated here as extracted from KEGG database (Kanehisa and Goto, 2000) in Figure 5.11. The figure illustrates two probable pathways showing the identified variants in relation to ciprofloxacin susceptibility and their related statistical assessment.

Anthranilate can act as a primary metabolite precursor for the well-known PQS (2-heptyl-3,4-dihydroxyquinoline biosynthesis). *Pseudomonas* quinolone signal is involved in multiple cellular functions and pathways including the regulation of several virulence factors, intercellular signaling in biofilms and iron chelating activity (Bredenbruch *et al.*, 2006).

Results of the 1st round of annotation from comparative behavioral genomics suggest the involvement of this signaling system in quinolone resistance. This may occur as a result of upregulation of oxidative stress response genes triggered by PQS including; superoxide dismutases, catalases, and alkyl hydroperoxide reductases. Figure 5.11 (A) illustrates the two sources of anthranilate synthesis through tryptophan degradation or corismite conversion. Both are considered precursors to PQS signal.

This hypothesis is based on the observation of several polymorphisms in genes encoding for some enzymes involved in several key steps along this pathway. Identified polymorphisms showed significant correlation with ciprofloxacin resistance behavior. The identified SNPs are summarized in Figure 5.11(B). Two of the identified nucleotide changes were predicted as deleterious. These include: *phn*A Met40Val and *trp*E Ala40Thr. By observing the distribution of the five variants related to that pathway, it appears that both *phn*A Met40Val and *phn*A Gly4Arg tend to be more frequent at the upper segment of the phylogenetic tree while both *trp*E Ala40Thr and *ant*A Thr4Asn tend to be more frequent at the lower segment of the phylogenetic tree (Figure 5.10).

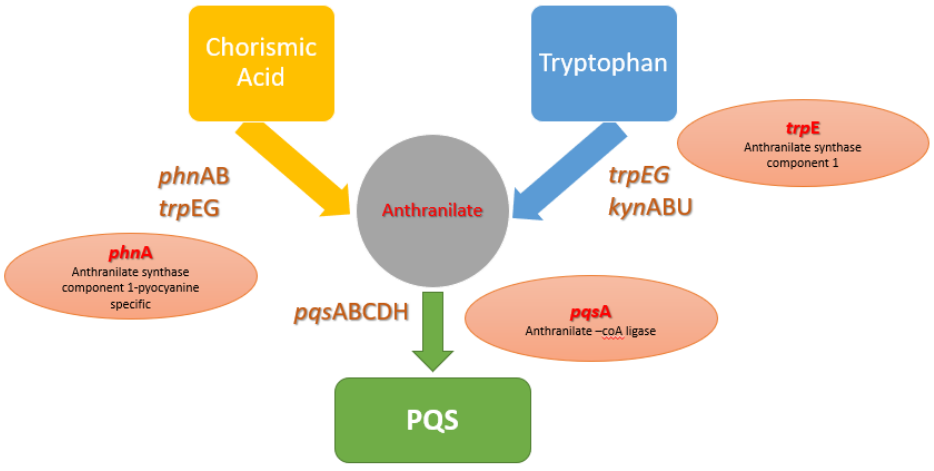
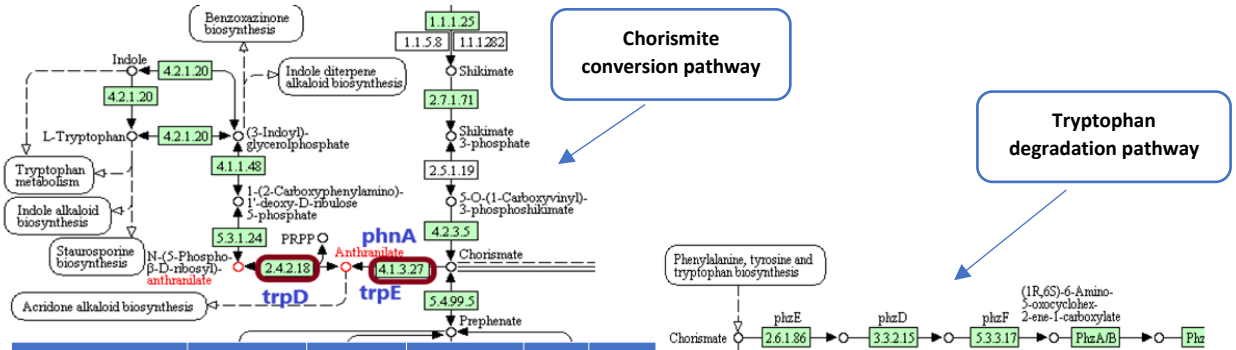


Figure 5.11 (A). The figure represents the two sources of anthranilate synthesis. The variants shown in red exhibited significance when CBG was applied



Gene	P-value (chi square or fischer exact)	Phi-coeffecient	Risk estimate (Confidence interval)	Upper limit	Lower limit
Anthranilate phosphoribosyltransferase (TrpD)	0.001	0.42	4.942	1.312	18.616
Anthranilate synthase component 1 (TrpE)	0.003	0.383	3.096	1.234	7.767
Anthranilate synthase component 1_ pyocyanine specific (phnA)	0.036	0.271	2.273	0.923	5.598
Anthranilate--CoA ligase (pqsA)	0.211	0.188	1.816	0.569	5.794
Anthranilate 1_2-dioxygenase large subunit (antA)	0.003	0.383	3.096	1.234	7.767

Figure 5.11(B). The figure shows gene variants identified using CBG in the pathway of Anthranilate biosynthesis and Pseudomonas quinolone signal

Genes in the pathway of anthranilate biosynthesis are labelled in red

5.5.2.2. glutamate and glutathione metabolism

Changes in genes involved in glutamate metabolism also showed significant correlation with ciprofloxacin resistance. Observed variants and their statistical values are shown in Table 5.4. The genes identified can be responsible for maintaining the cellular pool of glutamate needed by the cell on exposure to oxidative stress as in the case of exposure to quinolones. Increased glutamate production can be linked to bacterial cell protection and survival on exposure to oxidative stress. In that case, glutamate can be converted in the bacterial cytoplasm into a number of compounds such as glutamine, glutathione, GABA or the TCA cycle intermediate oxoglutarate or succinate and fumarate to a lesser extent. Figure 5.12 (A) shows glutathione/glutamate metabolism. Glutathione is considered one of the antioxidant molecules that protect the bacterial cell against reactive oxygen radicals when the bacterial cell is exposed to oxidative stress. The level of glutathione has also been noticed to increase during oxidative stress in relation to increased glutamate level. On the other hand, oxoglutarate is known as a potent antioxidant molecule which can be converted, in absence of any enzymatic reaction, into succinate in the presence of H_2O_2. In addition, conversion of glutamate to oxoglutarate by the glutamate dehydrogenase *gdh*A may increase the production of NADPH, which might also contribute to the anti-oxidant effect of glutamate acquisition (Ramond *et al.*, 2014). The deactivating nucleotide changes (markers) from CBG output are shown in Table 6.2. These were tested using PROVEAN and showed that the gene *gdh*B is non-functional which can consequently affect glutamate utilization pathways. Figure 5.12 (B) shows parts of the pathways of glutathione metabolism and glutamate metabolism extracted from KEGG database (Kanehisa and Goto, 2000). The graph shows the gene variants identified through CBG and their statistical values of performance.

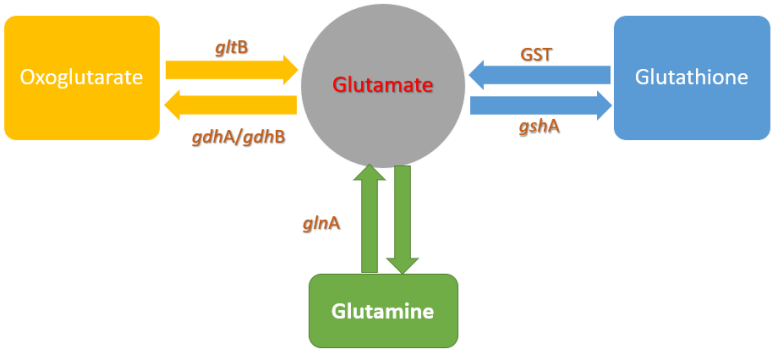
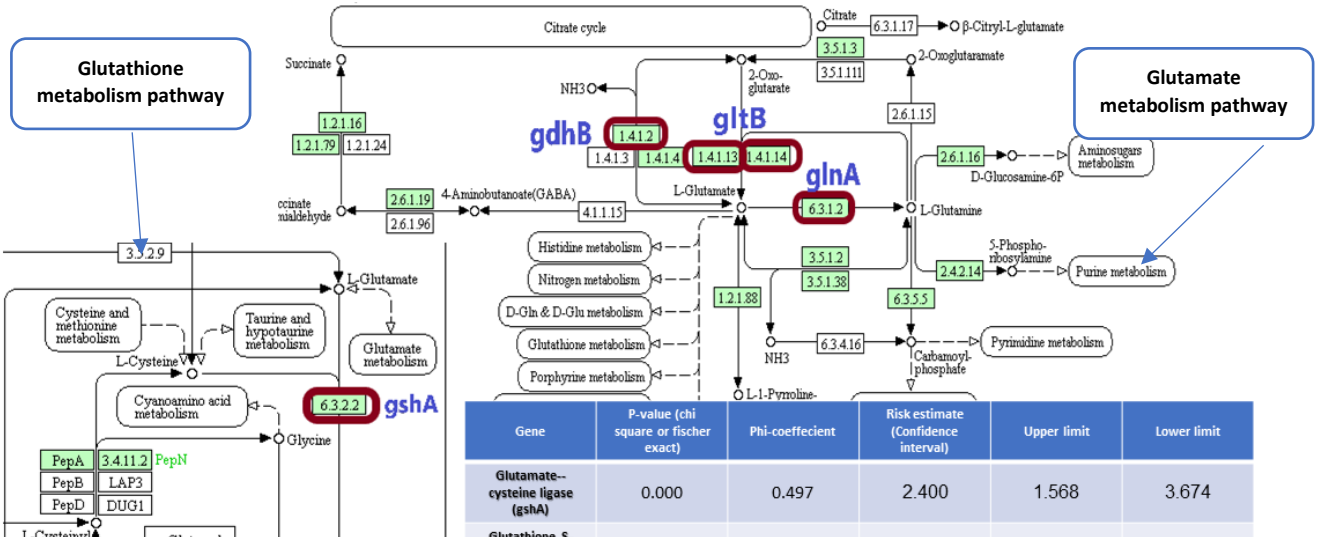


Figure 5.12 (A). The diagram illustrates the central nitrogen cycle and the pathways involved in glutamate metabolism



Gene	P-value (chi square or fischer exact)	Phi-coeffecient	Risk estimate (Confidence interval)	Upper limit	Lower limit
Glutamate--cysteine ligase (gshA)	0.000	0.497	2.400	1.568	3.674
Glutathione S-transferase GST-6.0	0.041	0.305	2.014	1.039	3.905
Glutamate synthase [NADPH] large chain (gltB)	0.013	0.32	2.085	1.038	4.189
NAD-specific glutamate dehydrogenase (gdhB)	0.039	0.269	1.837	0.844	3.997
Glutamine synthetase (glnA)	0.017	0.311	2.88	1.559	5.322

Figure 5.12 (B). The diagram illustrates gene variants identified in the pathway of glutamate and glutathione metabolism by using CBG

5.6. Results. Section 3. Post-CBG identification of candidate susceptibility/resistance markers for gentamycin in *Ps. aeruginosa* (2nd round of annotation)

At this stage, CBG sheets were filtered for finding significant variants at extremes of behavior. Genes were re-checked against gene annotations available from public databases. The variants defined below and used for calculation of predictive values are based on that annotation.

5.6.1. Description of the significance and distribution of candidate loci across different phenotypes

The BINs used for CBG comparison consisted of four phenotypic comparison groups. These four groups were analyzed in six pairwise comparisons. Each BIN /group was compared to all other groups. In order to find the most important differences, the intermediate pairwise comparisons were removed and those analyzed included: PWC 2_3, PWC3_4 and PWC1_2.

The venn diagram shown below represents the number of single nucleotide polymorphisms (SNPs) related to gentamycin susceptibility/resistance phenotype for the three analyzed pairwise comparison.

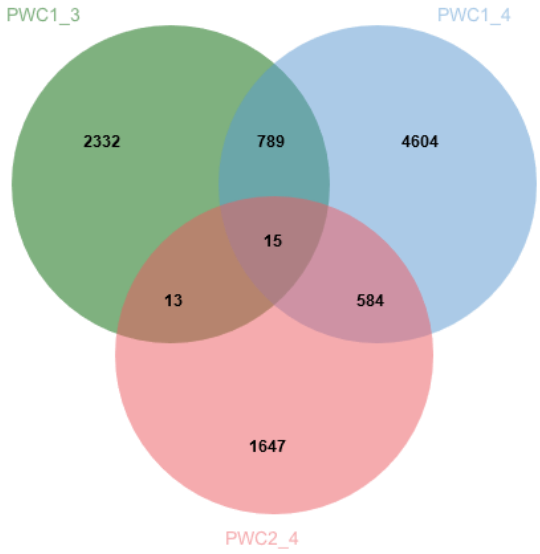


Figure 5.13. The venn diagram illustrates the number of SNPs showing significance at the cutoff of 0.05. Three pairwise comparison groups were prioritized, and these are illustrated at the three circles shown. The intersections represent those SNPs co-detected among the 3 groups

PWC1_3; Green circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 1 and BIN 3

PWC1_4; Blue circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 1 and BIN 4

PWC2_4; Pink circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 2 and BIN 4

The venn diagram shown in Figure 5.13 has helped to filter the variants. Variants at the intersections of the three groups were studied.

5.6.2. Diagnostic Performance of candidate loci as potential molecular diagnostic markers

Gene variants at the intersections of venn diagram shown above were visually and manually scanned in order to select those variants showing significant correlation at the extremes of gentamycin comparative groups and at the same time lacking significant correlations among middle groups.

As a result, markers shown below were prioritized and were tested for their diagnostic accuracy using predictive values. Primary distribution of CBG-based markers in the studied group of lab strains (162 isolates) is shown in Supplementary Table 6. These markers were classified into susceptible and resistant groups based on the latest clinical breakpoints (Rules, 2018). Predictive values and effect sizes were consequently analyzed to reflect practical significance. Measures of diagnostic accuracy for gentamycin CBG-based markers are shown in Figure 5.14. This figure shows how different measures of performance are distributed for different variants.

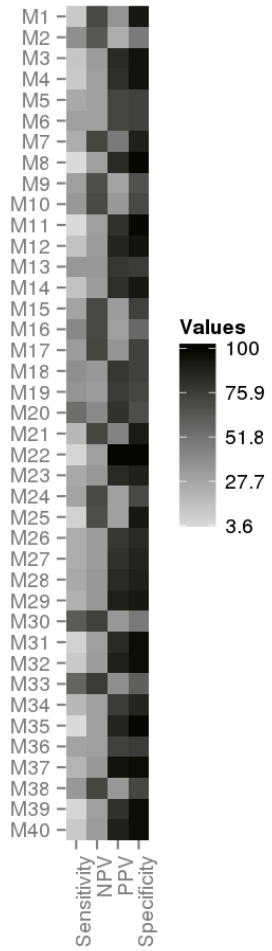


Figure 5.14. Heatmap showing diagnostic performance for CBG-based gentamycin molecular diagnostic markers

The analysis showed some variants linked to gentamycin susceptibility and those variants showed tendency to be absent at higher MICs. These variants were assessed in two ways. The first group showed high specificity and high PPV to susceptibility phenotype. This can consequently rule-in susceptibility. The second group showed high NPV to resistance phenotype which can consequently rule-out resistance. These variants are shown in detail in Table 5.7 and in Table 5.8.

5.6.2.1. Candidate gentamycin markers showing high specificity and high PPV

Variants showing higher specificity and higher PPV to gentamycin susceptibility phenotype are shown in Table 5.7. These markers have the potential to be used as rule- in susceptibility diagnostic.

Table 5.7. CBG-based Gentamycin Molecular Markers showing higher specificity/PPV

Gene ID	Gene Identifier	SNP position	SNP change	AA position	AA change	CBG Marker ID	sensitivity	NPV	PPV	specificity	Frequency of SNP change %
PA4744	Translation initiation factor IF-2	261	C>T	87	R	M22	7.1	32.5	100	100	4.93
PA1041	Outer membrane porin F	543	A>G	181	G	M8	4.5	31.4	83.3	98	3.7
PA1207	Glutathione-regulated potassium-efflux system protein KefC	393	C>T	131	A	M11	3.6	31.2	80	98	3.08
PA3860	Long-chain-fatty-acid--CoA ligase	1623	C>G	541	R	M35	5.4	31.6	85.7	98	4.3
PA3431	Inner membrane protein YohK	502	C>T	168	L	M31	8.9	32	83.3	96	7.4
PA3431	Inner membrane protein YohK	517	T>C	173	L	M32	12.5	32.9	87.5	96	9.9
PA5112	Esterase EstA	1188	T>C	396	N	M37	21.4	35.3	92.3	96	16.04
PA2194	Hydrogen cyanide synthase subunit HcnB	714	A>G	238	V	M39	7.1	31.6	80	96	6.2
PA4055	Riboflavin synthase	372	T>G	124	R	M40	12.5	32.9	87.5	96	9.9
PA2836	**putative multidrug resistance protein EmrK**	**993**	T>C	**331**	**R**	M3	13.4	32.6	83.3	94	11.11
PA2836	**putative multidrug resistance protein EmrK**	**1014**	A>G	**338**	A	M4	11.6	32.2	81.3	94	9.9
PA1207	Glutathione-regulated potassium-efflux system protein KefC	543	C>T	181	G	M12	15.2	33.1	85	94	12.34
PA3614	Ribonuclease	1287	A>C	429	P	M14	15.2	32.6	81	92	12.96
PA5201	30S ribosomal protein S1	386	C>A	129	A129E	M29	23.2	34.8	86.7	92	18.52
PA4744	Translation initiation factor IF-2	129	G>A	43	K	M23	26.8	34.9	83.3	88	22.22
PA5197	Ribosomal protein S6--L-glutamate ligase	744	C>T	248	D	M28	26.8	34.9	83.3	88	22.22
PA5181	putative oxidoreductase	1128	G>A	376	Q	M27	25.9	34.1	80.6	86	22.22
PA0009	Glycine--tRNA ligase alpha subunit	540	C>T	180	D	M34	19.6	32.3	75.9	86	17.9
PA5181	putative oxidoreductase	1062	G>A	354	E	M26	25.9	33.6	78.4	84	22.84
PA3614	Ribonuclease	1268	G>C	423	G423A	M13	36.3	34.9	77.4	76	32.72
PA0108	Cytochrome c oxidase subunit 3	846	C>T	282	D	M36	29.5	32.5	73.3	76	27.8
PA2870	Response regulator PleD	1206	C>T	402	H	M5	28.6	31.6	71.1	74	27.8
PA2910	putative manganese efflux pump MntP	534	C>T	178	G	M6	31.3	32.5	72.9	74	29.63
PA1529	DNA ligase	1698	C>G	566	L	M18	40.2	35.6	77.6	74	35.8
PA1630	HTH-type transcriptional regulator TsaQ1/TsaQ2	174	G>A	58	T	M19	37.5	34	75	72	34.6
PA1754	HTH-type transcriptional regulator CysB	174	T>G	58	R	M20	55.4	40.5	79.5	68	48.15

Markers with highest significance and highest effect sizes among all 26 markers included: M20 (HTH-type transcriptional regulator *Cys*B T 174 G), (*phi*= 0.216, *p*=0.006), M37 (Esterase EstA T 1188 C), (*phi*= 0.219, *p*=0.005), M22 (Translation initiation factor IF-2 C 261 T), (*phi*= 0.152, *p*=0.053), M29 (30S ribosomal protein S1 C 386 A), (*phi*= 0.181, *p*=0.021), M28 (Ribosomal protein S6--L-glutamate ligase C 744 T), (*phi*= 0.164, *p*=0.037) and M23 (Translation initiation factor IF-2 G 129 A), (*phi*= 0.164, *p*=0.037).

The distribution of the most important molecular predictors across different MICs are shown in Figure 5.15 (A-F).

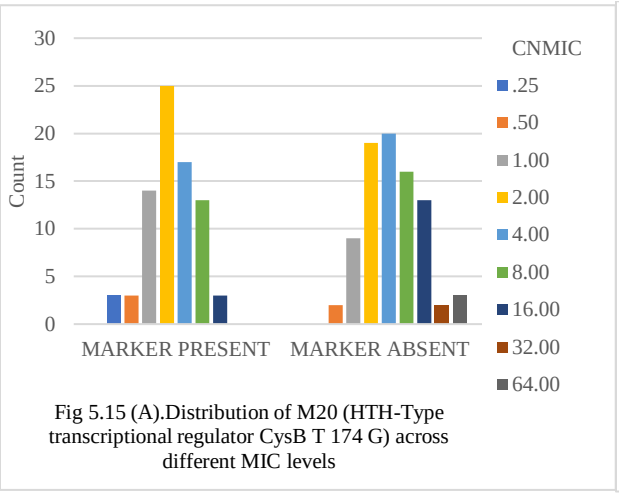


Fig 5.15 (A).Distribution of M20 (HTH-Type transcriptional regulator CysB T 174 G) across different MIC levels

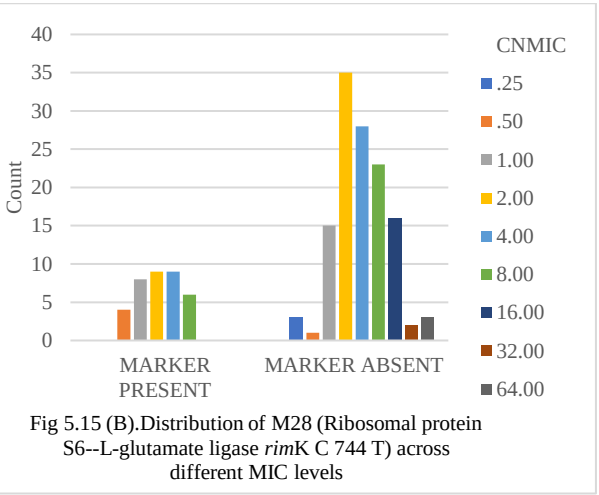


Fig 5.15 (B).Distribution of M28 (Ribosomal protein S6--L-glutamate ligase *rim*K C 744 T) across different MIC levels

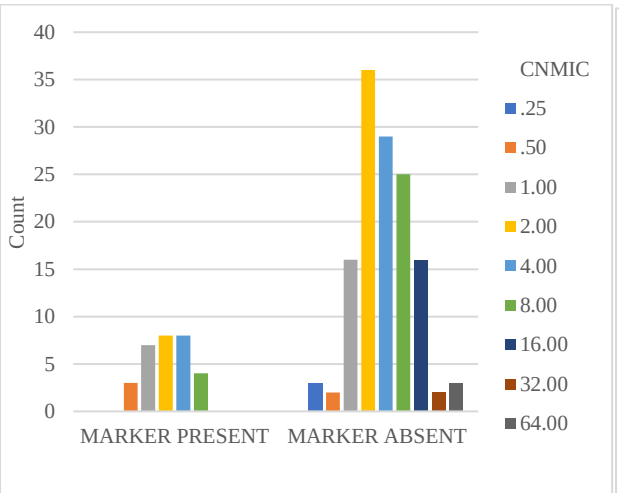


Fig 5.15 (C).Distribution of M29 (30S ribosomal protein S1 C 386 A) across different MIC levels

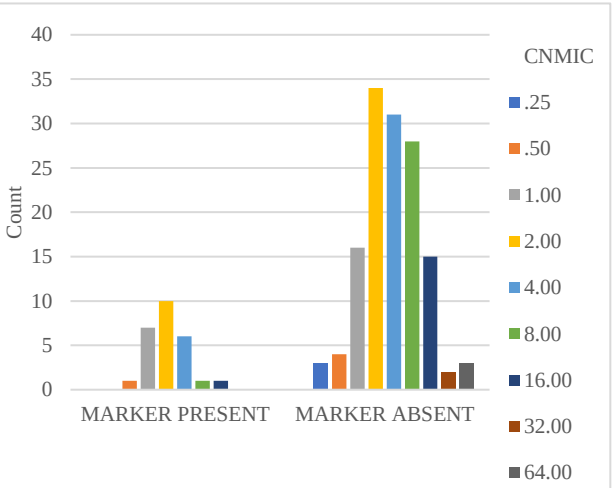


Fig 5.15 (D).Distribution of M37 (Esterase *Est*A T 1188 C) across different MIC levels

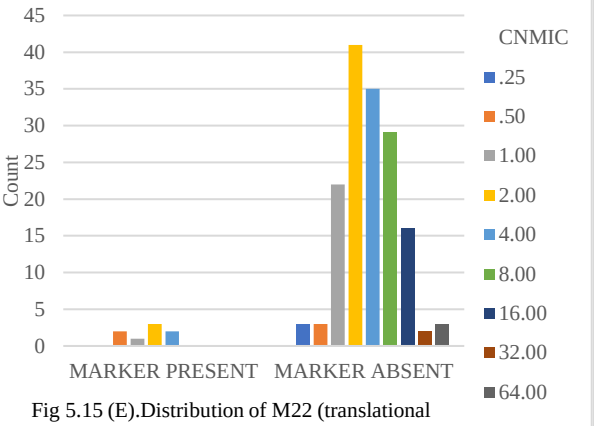


Fig 5.15 (E).Distribution of M22 (translational initiation Factor IF-2 *inf*B C 261 T) across different MIC levels

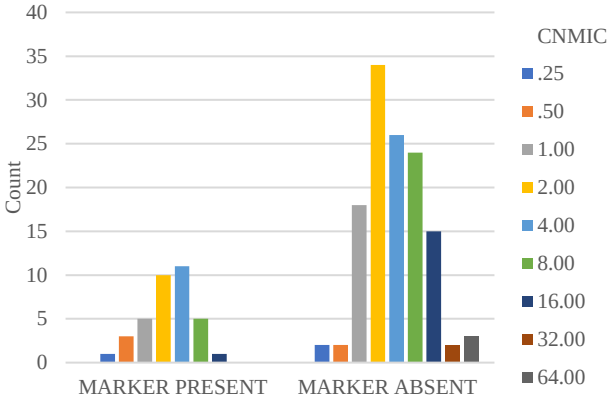


Fig 5.15 (F).Distribution of M23 (translational initiation Factor IF-2 *inf*B G 129 A) across different MIC levels

Figure 5.15 shows six gentamycin-related molecular markers with higher specificity and higher PPV. These are listed in Table 5.7.

The figure shows that the markers tend to be absent at higher gentamycin MICs.

5.6.2.2. Candidate gentamycin markers showing higher NPV

Variants showing high NPV to gentamycin resistance phenotype are shown in Table 5.8. These markers have the potential to be used as rule-out resistance diagnostic.

The two markers showing the highest significance and highest effect sizes among those evaluated included; M7 (Cell division inhibitor *Sul*A C111T), (*phi*= 0.167, *p*=0.033) and M33 (Ribosomal RNA small subunit methyltransferase B C423A), (*phi*= 0.174, *p*=0.027). The distribution of these markers across different MICs is shown in Figure 5.16 (A-B).

Table 5.8. CBG-based Gentamycin Molecular Markers showing higher NPV

Gene ID	Gene Identifier	SNP position	nucleotide change	AA position	AA	CBG Marker ID	sensitivity	NPV	PPV	specificity	Frequency of SNP change %
PA0017	Ribosomal RNA small subunit methyltransferase B	423	C>A	141	G	M33	58	76.4	40	60.7	45.06
PA3431	Inner membrane protein YohK	132	T>C	44	Y	M30	62	74	35	48.2	54.93
PA3008	Cell division inhibitor SulA	111	C>T	37	S	M7	26	72.6	48	87.5	16.7
PA3587	HTH-type transcriptional regulator CynR	543	C>T	181	Y	M17	34	71.6	37	74.1	28.4
PA0473	putative GST-like protein YibF	678	C>T	226	R	M21	20	71.4	46	89.3	13.6
PA5112	Esterase EstA	1779	C>T	593	F	M38	36	71.4	36	71.4	30.9

PA1147	Inner membrane protein YjeH	1257	A>G	419	G	M10	36	70.9	35	69.6	32.1
PA3587	HTH-type transcriptional regulator CynR	142	T>C	47.33	L	M15	30	70	33	73.2	27.8
PA2716	NADH oxidase	681	T>C	227	D	M1	12	69.9	38	91.1	9.9
PA3587	HTH-type transcriptional regulator CynR	225	G>A	75	Q	M16	44	69.6	31	57.1	43.2
PA0815	HTH-type transcriptional regulator YjiE	426	G>A	142	V	M24	30	69.3	31	70.5	29.6
PA0782	**Bifunctional protein PutA**	**2958**	T>C	**986**	R	M25	8	69.1	31	92	8.02
PA1060	Inner membrane protein YtfF	123	T>C	41	G	M9	32	68.8	30	67	32.7
PA2691	NADH dehydrogenase-like protein	1032	C>T	344	D	M2	40	64.7	26	49.1	47.53

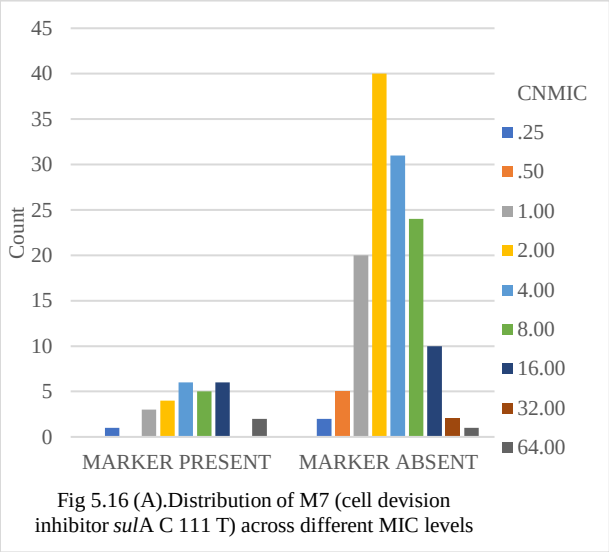


Fig 5.16 (A).Distribution of M7 (cell devision inhibitor *sul*A C 111 T) across different MIC levels

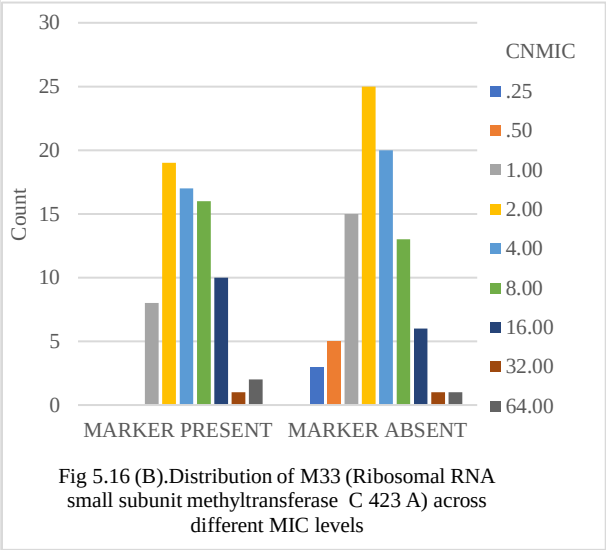


Fig 5.16 (B).Distribution of M33 (Ribosomal RNA small subunit methyltransferase C 423 A) across different MIC levels

Figure 5.16 shows two gentamycin-related molecular markers with the highest NPV as listed in Table 5.8 which also shows higher effect size.

The figure shows that the markers tend to be absent at higher gentamycin MICs.

5.6.3. Identification of new combinations of potential molecular markers with improved diagnostic performance

Cluster analysis was applied to explore for the possibility of finding better combination of markers showing better performance. This was achieved by using the same approach as explained in detail in *Chapter* 3. The following 2 combinations of markers were identified.

Combination1: A combination of 5 molecular markers including M20, M23, M28, M37 and *nal*C A186 T showed higher effect size in relation to gentamycin susceptibility phenotype. This was higher than any of the above markers separately, (*chi*=15.308, *p*=0.005, *phi*=0.307).

The detailed composition of these markers and their distribution are shown in Figure 5.17 and in Table 5.8.

Red Column to the right-hand side of each cluster shows “marker absence”

Red Column to the left-hand side of each cluster shows “marker presence”.

Distribution of CN MIC in each cluster is shown as continuous scale with higher MICs at the right-hand side and lower MICs at the left-hand side.

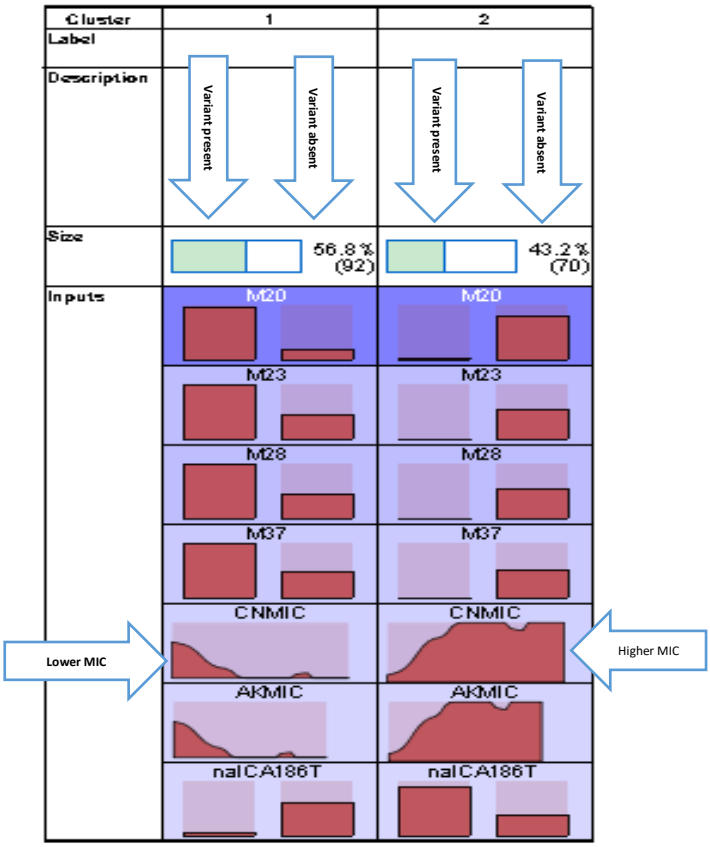


Figure 5.17. The figure illustrates the relative distribution of gentamycin molecular markers that form combination 1 across the 2 identified clusters

This figure shows how markers are distributed (present/absent) in relation to different MICs.

Table 5.9. Distribution of the components of combination 1 into two clusters in relation to gentamycin susceptibility

Cluster	Cluster.1	Cluster.2
Size	56.8 % (92)	43.2 % (70)
M20	Marker Present (81.5%)	Marker Absent (95.7%)
M23	Marker Absent (60.9%)	Marker Absent (100%)
M28	Marker Absent (60.9%)	Marker Absent (100%)
M37	Marker Absent (71.7%)	Marker Absent (100%)
*nal*C A186T	Mutation Absent (98.9%)	Mutation Absent (84.3%)
Average gentamycin MIC	3.55	9.71

The table shows that M20, M23, M28, and M37 tend to be absent at higher gentamycin MICs.

These 5 markers have divided the studied isolates into two clusters; Cluster 1 with lower gentamycin MICs (56.8%, 92 isolates) and Cluster 2 with higher gentamycin MICs (43.2%, 70 isolates). Figure 5.18 shows the distribution of the two new clusters of molecular markers among different MICs.

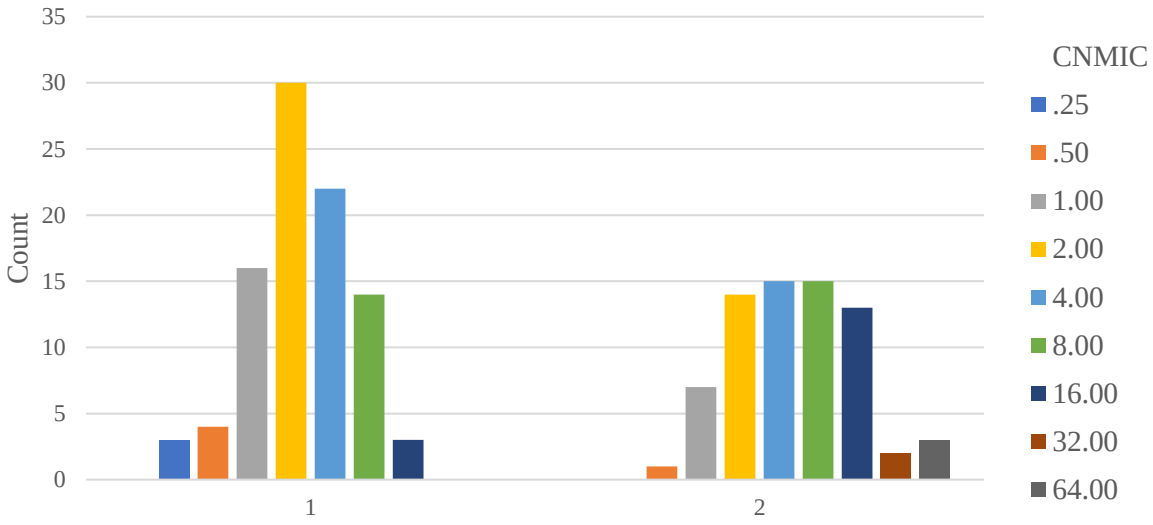


Fig 5.18. The distribution of different gentamycin MICs into two new clusters of molecular markers that form combination 1

The figure shows that cluster 2 tends to occur at higher gentamycin MICs while cluster 1 tends to occur at lower gentamycin MICs.

Cluster 2 was tested for its predictive performance for resistance phenotype and showed sensitivity of 66%, PPV of 47.1%, specificity of 67% and NPV of 81.5% with (*chi*=15.308, *p*<0.005, *Phi*=0.307). The likelihood ratio of this cluster to differentiate different MICs is 27.697 with (*phi*=0.387, *P*=0.002).

Cluster 1 was also tested for its predictive performance for sensitivity phenotype and showed sensitivity of 37.5%, PPV of 77.8%, specificity of 76% and NPV of 35.2% with (*chi*=2.835, *p*=0.092, *phi*=0.132). The likelihood ratio of this cluster to differentiate different MICs =14.028 with (*phi*=0.273, *p*=0.081).

Combination 2: Applying cluster analysis showed a second model with improved performance. This has divided the studied isolates into 3 clusters of markers. The markers ordered starting with the most important predictors included; M29, M33, M23, *Amp*RE114A, M28, M20, M7, M22

and *nal*C E153Q. The combination of 9 molecular markers showed higher effect size in relation to gentamycin susceptibility phenotype when compared to any of the individual markers separately, (*chi*=25.74, *p*=0.058, *phi*=0.399). The differential distribution of markers in relation to MIC is shown in Table 5.10,

Table 5.10. Distribution of the components of combination 2 into three clusters in relation to gentamycin susceptibility

	Cluster.1	Cluster.2	Cluster.3
	33.3% (54)	50% (81)	16.7% (27)
M29	Marker Absent (94.4%)	Marker Absent (100%)	Marker Present (100%)
M33	Marker Absent (100%)	Marker Present (71.6%)	Marker Present (55.6%)
M23	Marker Present (59.3%)	Marker Absent (95.1%)	Marker Absent (100%)
*Amp*R E114A	Mutation Present (59.3%)	Mutation Absent (95.1%)	Mutation Absent (77.8%)
M28	Marker Absent (59.3%)	Marker Absent (100%)	Marker present (51.9%)
M20	Marker Present (81.5%)	Marker Absent (72.8%)	Marker Absent (55.6%)
M7	Marker Absent (100%)	Marker Absent (67.9%)	Marker Absent (96.3%)
M22	Marker Absent (85.2%)	Marker Absent (100%)	Marker Absent (100%)
*nal*C E153Q	Mutation Absent (92.6%)	Mutation Absent (100%)	Mutation Absent (96.3%)
Average gentamycin MIC	3.87	8.86	2.94

The table shows that each of M20, M28, and *amp*RE114A tends to be absent at higher MICs while M33 tends to be present at higher MICs.

Cluster 2 showed tendency to occur towards higher gentamycin MICs and included 50% of isolates (81) while cluster 3 showed tendency to occur towards lower gentamycin MICs and included 16.7 % (27) isolates.

Predictive performance of cluster 2 was tested to resistance phenotype and showed sensitivity of 70%, PPV of 43.2, specificity of 58.9% and NPV of 81.5% with (*chi*=11.571, *p*=0.001, *phi*=0.267). Likelihood ratio of the cluster to differentiate different MICs= 22.648 with (*p*=0.004, *phi*=0.346).

Cluster 3 was also tested for its predictive performance to susceptibility phenotype and showed a sensitivity of 67%, PPV of 81.5%, specificity of 66% and NPV of 47.1% with (*chi*=15.308,

p<0.005, *phi*=0.307). Likelihood ratio of cluster to differentiate different MICs= 27.697 with (*p*=0.002, *phi*=0.387).

Red Column to the right-hand side of each cluster shows "marker absence"

Red Column to the left-hand side of each cluster shows "marker presence".

Distribution of CN MIC in each cluster is shown as continuous scale with higher MICs to the right-hand side and lower MICs to the left-hand side.

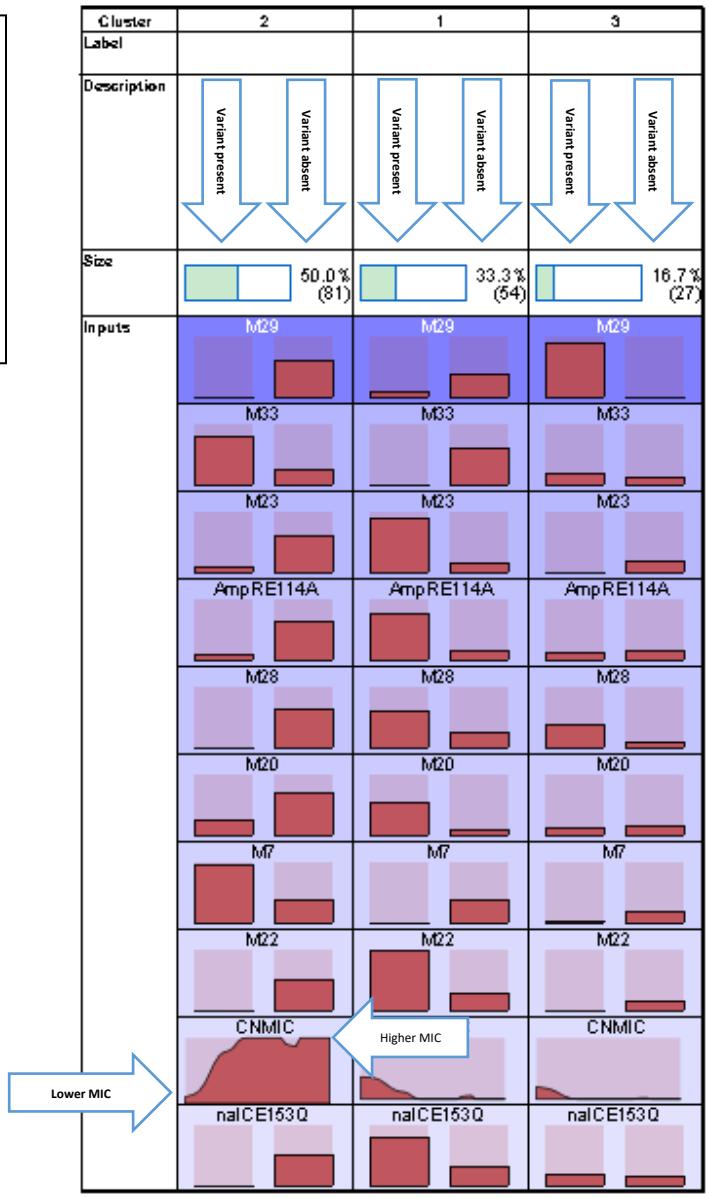


Figure 5.19. The figure illustrates the relative distribution of gentamycin molecular markers that form combination 2 across the 3 identified clusters

This figure shows how markers are distributed (present/absent) in relation to different MICs.

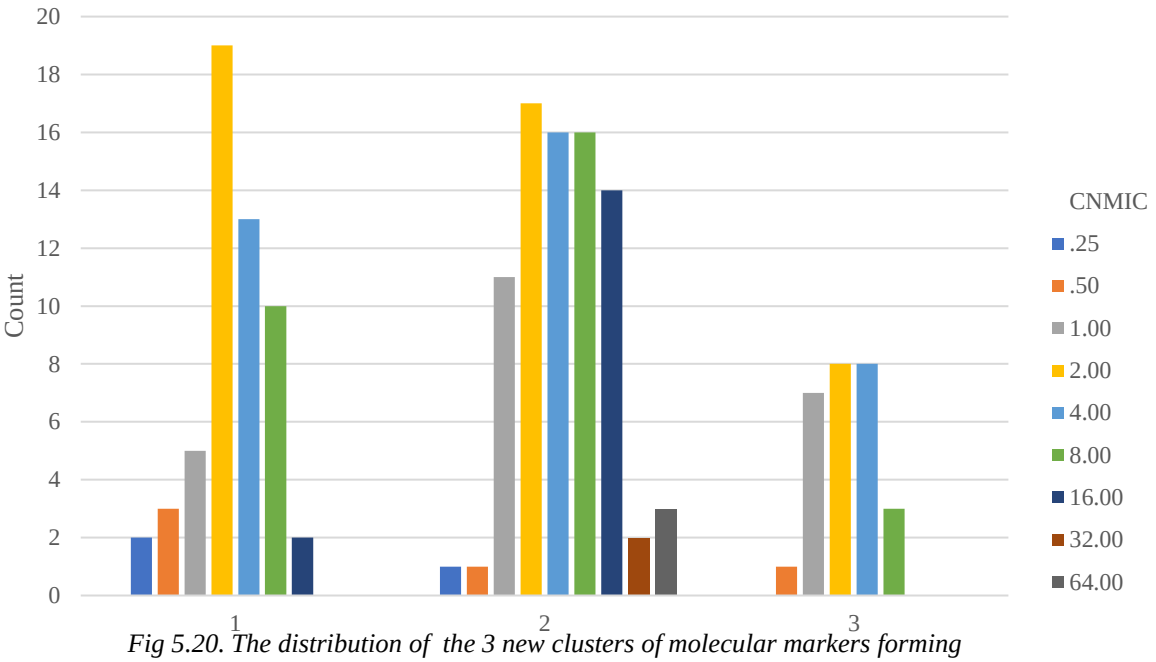


Fig 5.20. The distribution of the 3 new clusters of molecular markers forming combination 2 across different gentamycin MICs

The figure shows the tendency of cluster 2 to present at higher gentamycin MICs and the tendency of cluster 1 and cluster 3 to present at lower gentamycin MICs.

5.6.4. Distribution of candidate loci in relation to genomic background

Microbial genomes are subject to different evolutionary forces that shapes its populations. These forces include; mutation, gene gain and loss, genetic re-arrangements and recombination (Didelot and Maiden, 2010). Considering that, correct inference of evolutionary tree needs accurate identification of its population genetic structure and mosaicism which may have been introduced by recombination.

Recombination is defined as the exchange of genetic information between two lineages. It is very common in nature and can show remarkable impact on the evolutionary history especially in bacterial populations. Recombination is considered a process that takes genetic material from one genetic background and insert it into another (Hanage, 2016). The prevalence of recombination suggests its high importance for microbial evolution. Bacterial recombination has proved to provide a powerful adaptive benefit (Polz, Alm and Hanage, 2013). A good understanding of the impact of recombination on bacterial genomes is therefore crucial for the correct interpretation of genome analyses. In order to use bacterial genomic data for

epidemiologic purposes and to understand the spread of bacterial pathogens, a good understanding of how genetic exchange affects conclusions drawn from microbial genomic information is crucial. RDP4 (Martin *et al.*, 2010) was used to generate an alignment cleaned from variations which are assumed to have arisen through recombination. The remaining part of the genome was then considered as "clonal frame" in which variations are assumed to have accumulated only through mutations. This was used to infer true phylogeny and to infer the genetic relatedness using maximum likelihood algorithm (Price, Dehal and Arkin, 2010).

Figure 5.21 shows the distribution of all candidate gentamycin loci analyzed in this section in relation to background phylogeny. Candidate loci distributions for the studied set of 162 *Ps. aeruginosa* isolates were then re-mapped to the dendogram after removing recombinant regions to study the effect of recombination on the distribution of loci as shown in Figure 5.22.

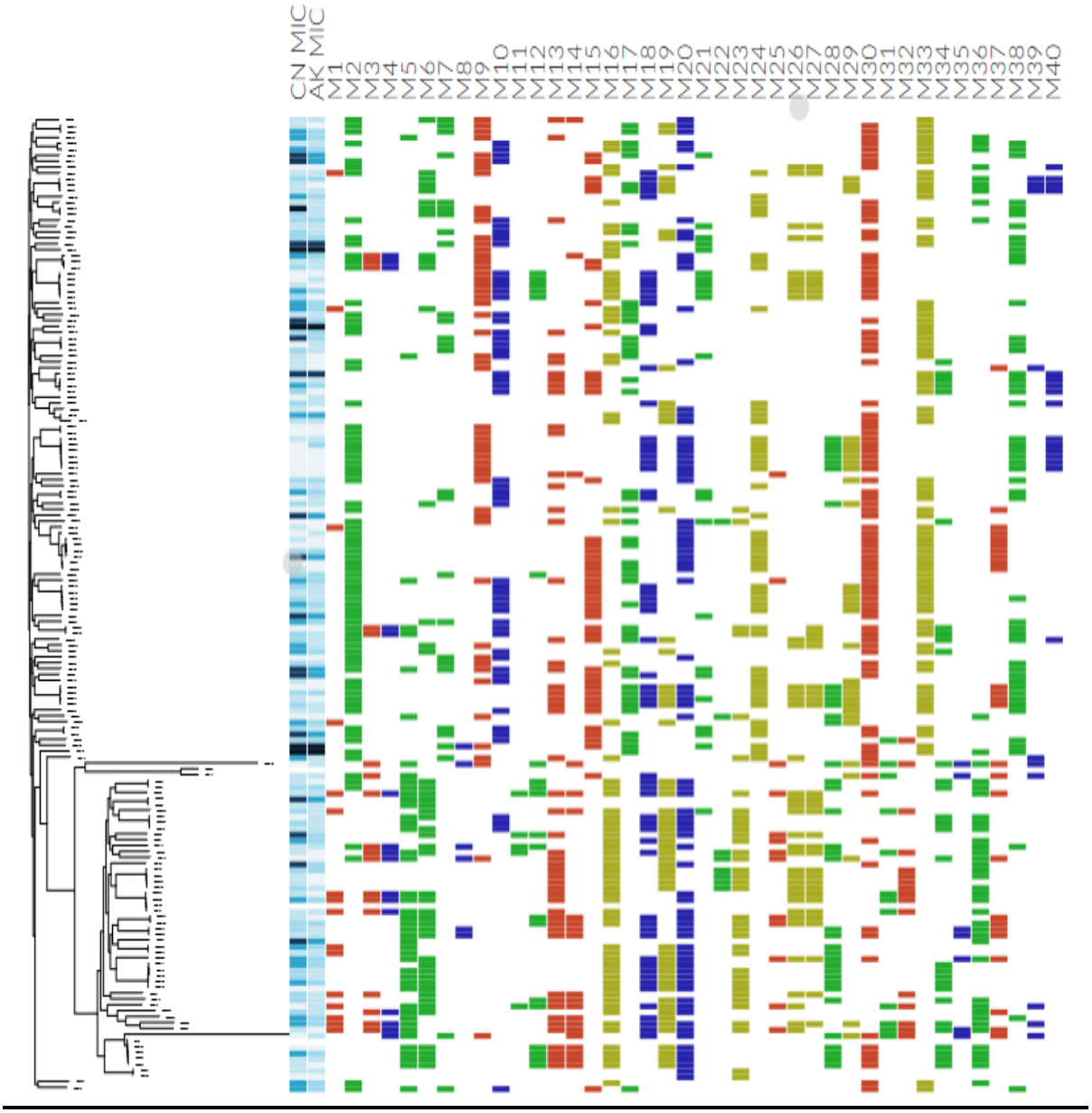


Figure 5.21. Distribution of candidate gentamycin markers in relation to background phylogeny

Gentamycin and amikacin susceptibility columns resistance/susceptibility as shades of blue where darker shades show higher resistance and lighter shades show lower resistance

Nucleotide changes evaluated at each marker position are shown in different colors where Red color refers to “C” nucleotide substitution, Green color refers to “T” nucleotide substitution, Blue color refers to “G” nucleotide substitution, and Yellow color refers to “A” nucleotide substitution

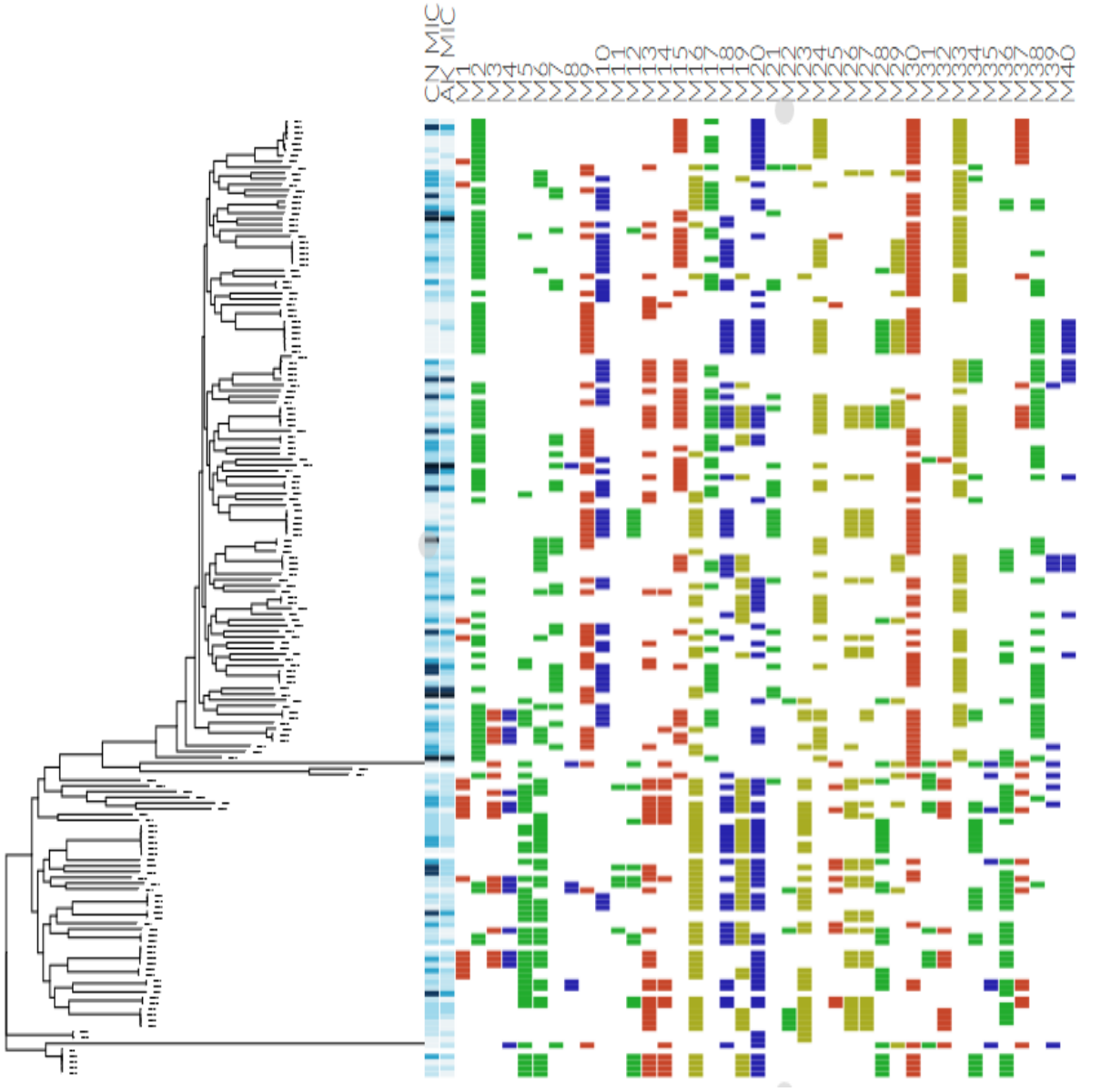


Figure 5.22. Dendogram showing re-constructed clonal genealogy mapped to candidate gentamycin markers

Gentamycin and amikacin susceptibility columns show different resistance/ susceptibility as shades of blue where darker shades show higher resistance and lighter shades show lower resistance

Nucleotide changes evaluated at each marker position are shown in different colors where Red color refers to “C” nucleotide substitution, Green color refers to “T” nucleotide substitution, Blue color refers to “G” nucleotide substitution, and Yellow color refers to “A” nucleotide substitution

The resulting corrected phylogenetic tree showed change in inter-isolate genetic distances and different branch lengths as a result of removing recombinant regions. However, the overall structure of the population remains separated into two large distinct groups. In addition, the differential distribution of candidate markers across the two large phylogenetic clusters remains

the same. This may indicate that these markers are probably more informative. The most important observation seen here is that the smaller subpopulation at the lower half of the dendogram in Figure 5.21 and Figure 5.22 appears to be more affected by recombination. This is seen as change in genetic distance/position of the whole cluster relative to the other subpopulation on the upper half of the dendogram when recombination was accounted for. This observation may indicate that markers clustering through this subpopulation offer real adaptive advantage as opposed to markers clustering through the upper half of the tree showing decreased fitness advantage.

Combining all the analyses of gentamycin markers shown in this section show that M7 and M33 have higher negative predictive values towards resistance and are distributed across the upper section of the phylogenetic tree. This may probably indicate lower fitness and the possibility to be used as rule-out resistance diagnostics markers. On the other hand, M20, M28, M29, and M37 which form a newly identified cluster showing better combined predictive values and higher effect size was distributed across different branches of the phylogenetic tree i.e. not clonally or evolutionary clustering. This probably indicates a genuine link to important biologic traits.

5.7. Results. Section 4. Post-CBG identification of candidate susceptibility/resistance markers for ciprofloxacin in *Ps. aeruginosa* (2nd round of annotation)

The approach applied here was the same approach used in gentamycin results section 5.7. Ciprofloxacin CBG sheets were filtered for finding significant variants at extremes of behavior. Genes were re-checked against gene annotations available from public databases and the variants defined and used for calculation of predictive values were based on that annotation.

5.7.1. Description of the significance and distribution of candidate loci across different phenotypic groups

The BINs used for CBG comparison consisted of four phenotypic comparison groups. These four groups were analyzed in six pairwise comparisons. Each BIN /group is compared to all other groups. In order to find important differences, intermediate pairwise comparisons were removed and those analyzed included PWC 2_3, PWC3_4 and PWC1_2.

The venn diagram shown below represents the number of single nucleotide polymorphisms (SNPs) associated with ciprofloxacin susceptibility/resistance phenotype for the three analyzed pairwise comparison.

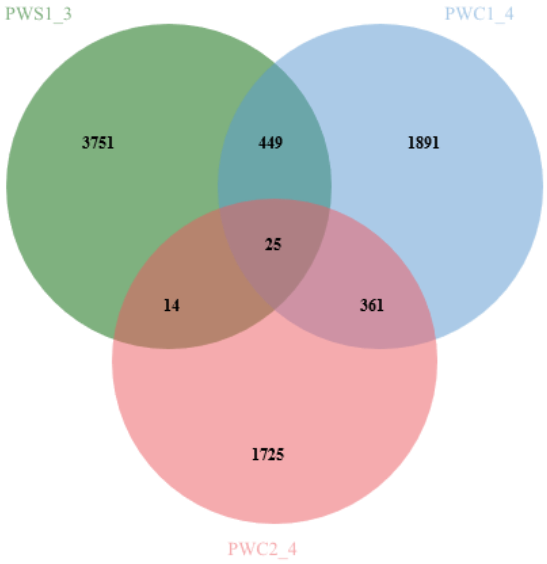


Figure 5.23. The venn diagram summarizes the number of SNPs with significance cutoff of 0.05 in each of the 3 prioritized pairwise comparison groups and also those co-detected among the 3 groups

PWC1_3; Green circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 1 and BIN 3

PWC1_4; Blue circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 1 and BIN 4

PWC2_4; Pink circle refers to number of gene variants identified in pair-wise comparison of phenotypic BIN 2 and BIN 4

The venn diagram shown in Figure 5.23 has helped to filter the variants, variants at the intersections of the three groups were prioritized to study.

5.7.2. Diagnostic Performance of candidate loci as potential molecular diagnostic markers

Gene variants at the intersections of venn diagram shown above were visually and manually scanned to select those variants showing significant correlations at extremes of gentamycin comparative groups and at the same time lacking significant correlations among middle groups.

As a result, markers were prioritized and were tested for their diagnostic accuracy using predictive values. A list of these markers is shown in Table 5.11 and Tables 5.12. Primary distribution of CBG-based markers in studied group of lab strains (162 isolates) is shown in Supplementary Table 7. These markers were classified into susceptible and resistant based on latest clinical breakpoints (Rules, 2018) and then predictive values and effect sizes were analyzed to reflect practical significance.

Measures of diagnostic accuracy for ciprofloxacin CBG-based markers are shown in Figure 5.24. The figure shows how different measures of performance are distributed for different variants.

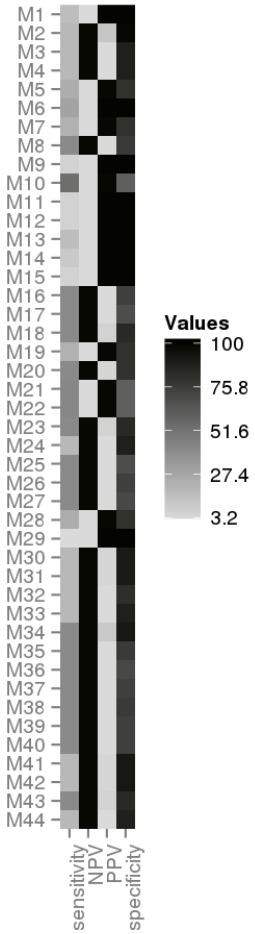


Figure 5.24. Heatmap showing diagnostic performance for CBG-based ciprofloxacin molecular diagnostic markers

The analysis showed that some variants were linked to ciprofloxacin susceptibility and those variants showed to be absent at higher MICs. These variants were assessed in two ways.

The first group showed high specificity and high PPV to susceptibility phenotype and can consequently rule-in susceptibility. The second group showed high NPV to resistance phenotype and can consequently rule-out resistance. These variants are shown in detail in Table 5.11 and Table 5.12.

5.7.2.1. Candidate ciprofloxacin markers showing high specificity and high PPV

Variants showing higher specificity and higher PPV to ciprofloxacin susceptibility phenotype are shown in Table 5.11. These markers have the potential to be used to rule- in susceptibility.

Table 5.11. CBG-based ciprofloxacin Molecular Markers showing higher specificity/PPV

Gene ID	Gene Identifier	SNP position	SNP change	AA position	AA change	CBG Marker ID	sensitivity	NPV	PPV	specificity	Frequency of SNP change %
PA2788	Methyl-accepting chemotaxis protein PctC	828	C>T	276	G	M1	18.5	3.8	100	100	17.9
PA3078	putative sensor histidine kinase TcrY	1032	A>G	344	A	M5	25.5	3.3	97.6	80	25.3
PA1273	c-diamide synthase	156	T>C	52	I	M6	28.7	4.3	100	100	27.77
PA1273	c-diamide synthase	498	C>T	166	P	M7	24.2	3.3	97.4	80	24.07
PA1374	putative HTH-type transcriptional regulator	99	A>T	33	L	M9	7.6	3.3	100	100	7.4
PA1374	putative HTH-type transcriptional regulator	222	T>C	74	R	M10	54.1	4	97.7	60	53.7
PA1374	putative HTH-type transcriptional regulator	271	G>C	91	V>L	M11	7.6	3.3	100	100	7.4
PA1374	putative HTH-type transcriptional regulator	300	C>T	100	G	M12	7.6	3.3	100	100	7.4
PA1374	putative HTH-type transcriptional regulator	376	A>C	126	M>L	M13	18.5	3.8	100	100	17.9
PA1408	Mechanosusceptible channel MscK	1150	G>A	384	A>T	M14	11.5	3.5	100	100	11.11
PA1408	Mechanosusceptible channel MscK	1329	G>A	443	V	M15	8.3	3.4	100	100	8.02
PA1694	Yop proteins translocation protein Q	668	C>T	223	T>I	M19	23.6	3.2	97.4	80	23.45
PA1805	Peptidyl-prolyl cis-trans isomerase D	1149	A>G	383	K	M21	42.7	3.2	97.1	60	42.59
PA1805	Peptidyl-prolyl cis-trans isomerase D	1167	C>T	389	R	M22	43.3	3.3	97.1	60	43.2
PA0454	Inner membrane protein YccS	2184	G>A	728	P	M28	26.1	3.3	97.6	80	25.92
PA0386	Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR	13	A>G	5	T>A	M29	5.1	3.2	100	100	4.93

M14 showed highest effect sizes among all 16 markers; Mechanosusceptible channel *Msc*K G 1150 A, (*phi*= 0.376, *p*=0.003). Other markers with lower significance included; M5 (putative sensor histidine kinase *Tcr*Y A 1032 G), (*phi*= 0.259, *p*=0.208) and M6 (c-diamide synthase T 156 C), (*phi*= 0.111, *p*=0.159). The distribution of these markers across different MICs is shown in Figure 5.25 (A-C).

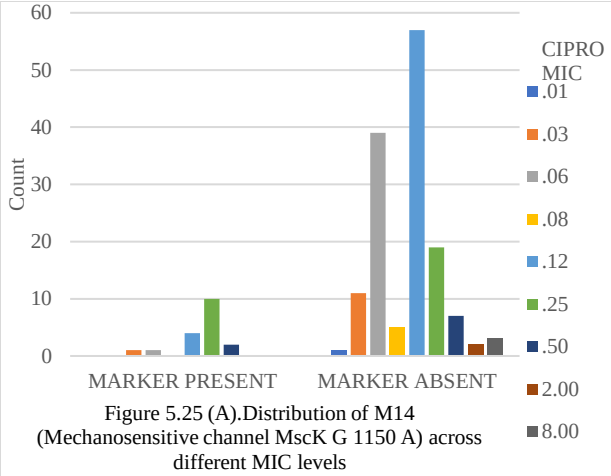


Figure 5.25 (A).Distribution of M14 (Mechanosensitive channel MscK G 1150 A) across different MIC levels

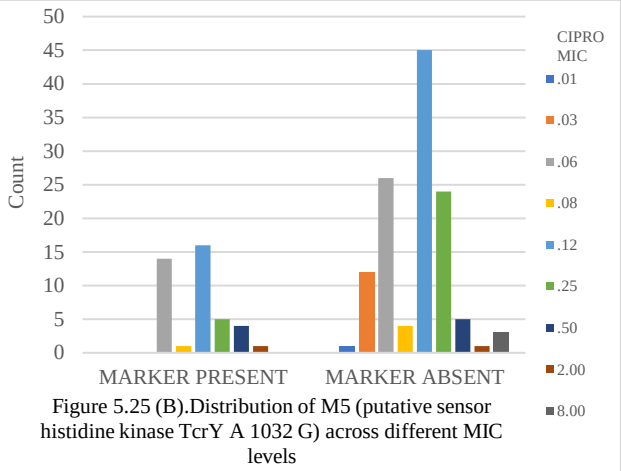


Figure 5.25 (B).Distribution of M5 (putative sensor histidine kinase TcrY A 1032 G) across different MIC levels

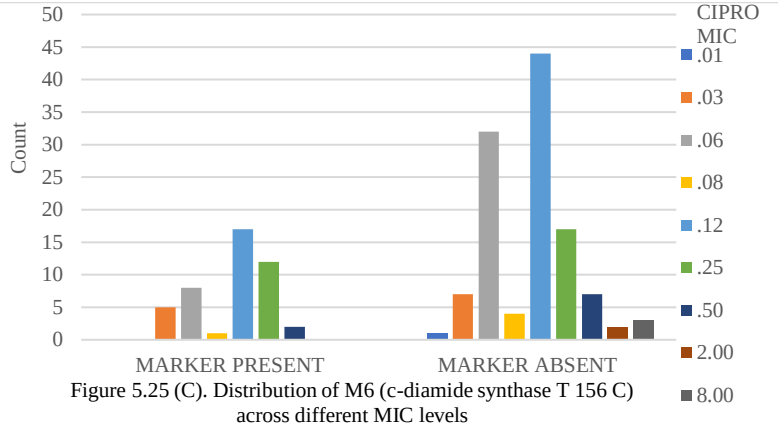


Figure 5.25 (C). Distribution of M6 (c-diamide synthase T 156 C) across different MIC levels

Figure 5.25 (A-C) shows three ciprofloxacin-associated molecular markers showing the higher specificity and higher PPV as listed in Table 5.11.

The figure shows that markers tend to be absent at higher gentamycin MICs.

5.7.2.2. Candidate ciprofloxacin markers showing higher NPV

Variants showing higher NPV to ciprofloxacin resistance phenotype are shown in Table 5.12. These markers have the potential to be used as rule-out resistance diagnostic.

The markers showing the highest significance and highest effect sizes among those evaluated included; M2 (Methyl-accepting chemotaxis protein PctC C1482 T), (*phi*= 0.138, *p*=0.08), M18 (*Yop* proteins translocation protein Q C 417 G), (*phi*= 0.329, *p*=0.025), M20 (Cysteine--tRNA ligase C 738 T), (*phi*= 0.306, *p*=0.056), M23 (Anthranilate phosphoribosyl transferase G 96 A), (*phi*= 0.112, *p*=0.155), M34 (Amino-acid permease *Roc*C C 1071 A), (*phi*= 0.388, *p*=0.002) and

M44 (HTH-type transcriptional regulator *Yof*A G 565 A), (*phi*= 0.369, *p*=0.005).The distribution of these markers across different MIC levels is shown in Figures 5.38- 5.43.

Table 5.12. CBG-based ciprofloxacin Molecular Markers showing higher NPV

Gene ID	Gene Identifier	SNP position	SNP change	AA position	AA change	CBG Marker ID	sensitivity	NPV	PPV	specificity	Frequency of SNP change %
PA2788	Methyl-accepting chemotaxis protein PctC	1482	C>T	494	V	M2	20	97.4	14.3	96.2	4.32
PA2811	Inner membrane transport permease YadH	444	G>A	148	V	M3	20	97.2	4.8	87.3	12.96
PA2811	Inner membrane transport permease YadH	450	G>A	150	V	M4	20	97.2	4.8	87.3	12.96
PA1273	c-diamide synthase	795	A>G	265	E	M8	40	97.5	4.9	75.2	25.3
PA3650	1-deoxy-D-xylulose 5-phosphate reductoisomerase	111	T>C	37	T	M16	40	97.5	4.7	73.9	26.54
PA1607	putative HTH-type transcriptional regulator	243	C>T	81	R	M17	40	97.3	3.8	68.2	32.09
PA1694	Yop proteins translocation protein Q	417	C>G	139	P	M18	40	97.8	7.1	83.4	17.28
PA1795	Cysteine--tRNA ligase	738	C>T	246	T	M20	40	97.7	5.9	79.6	20.98
PA0650	Anthranilate phosphoribosyl transferase	96	G>A	32	Q	M23	40	97.8	7.4	84.1	16.66
PA0650	Anthranilate phosphoribosyl transferase	594	C>T	198	T	M24	20	97.2	4.8	87.3	12.96
PA0650	Anthranilate phosphoribosyl transferase	957	G>C	319	L	M25	40	97.2	3.8	67.5	32.71
PA0484	Formyltetrahydrofolate deformylase	495	G>A	165	E	M26	40	97.5	4.8	74.5	25.92
PA0454	Inner membrane protein YccS	1803	G>A	601	E	M27	40	97.3	4.1	70.1	30.24
PA0386	Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR	51	C>T	17	H	M30	20	97.2	5.6	89.2	11.11
PA0386	Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR	57	G>C	19	P	M31	20	97.2	5.6	89.2	11.11
PA0306	HTH-type transcriptional regulator GadX	834	G>C	278	G	M32	20	97	3.4	82.2	17.9
PA4981	Amino-acid permease RocC	516	G>A	172	G	M33	20	97.2	5	87.9	12.34
PA4981	Amino-acid permease RocC	1071	C>A	357	A	M34	40	97.9	12.5	91.1	9.87
PA4979	Acyl-CoA dehydrogenase	603	G>A	201	G	M35	40	97.5	4.9	75.2	25.3
PA4830	Proofreading thioesterase EntH	24	T>C	8	D	M36	40	97.3	4	69.4	30.86
PA4830	Proofreading thioesterase EntH	151	C>G	51	Q>E	M37	40	97.5	4.8	74.5	25.92
PA4830	Proofreading thioesterase EntH	153	A>G	51	Q>E	M38	40	97.5	4.9	75.2	25.3
PA4830	Proofreading thioesterase EntH	162	T>C	54	C	M39	40	97.5	4.7	73.9	26.54
PA4830	Proofreading thioesterase EntH	300	C>T	100	D	M40	40	97.5	4.8	74.5	25.92
PA4749	Phosphoglucosamine mutase	213	T>C	71	D	M41	20	97.3	6.3	90.4	9.87
PA4749	Phosphoglucosamine mutase	228	A>G	76	G	M42	20	97.3	6.7	91.1	9.25
PA0816	HTH-type transcriptional regulator YofA	511	G>A	171	A>T	M43	40	97.8	7.1	83.4	17.28
PA0816	HTH-type transcriptional regulator YofA	565	G>A	189	V>M	M44	20	97.2	5	87.9	12.34

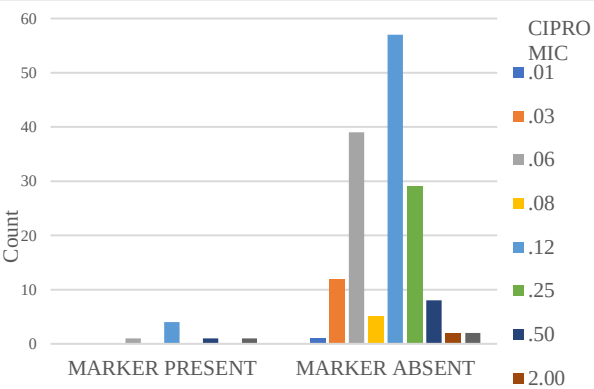


Figure 5.26 (A). Distribution of M2 (Methyl-accepting chemotaxis protein PctC C1482 T) across different MIC levels

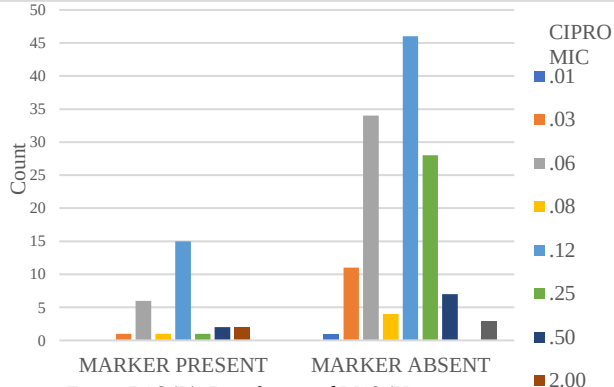


Figure 5.26 (B). Distribution of M18 (Yop proteins translocation protein Q C 417 G) across different MIC levels

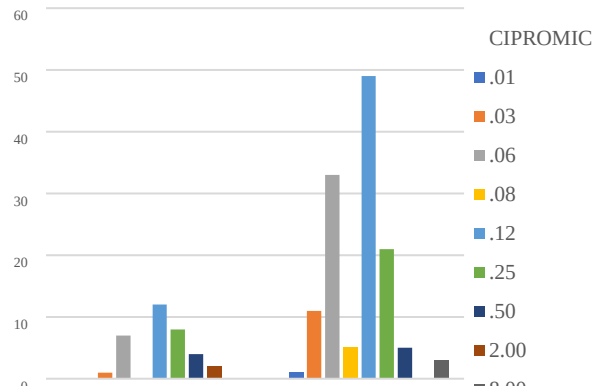


Figure 5.26 (C). Distribution of M20 (Cysteine--tRNA ligase C 738 T) across different MIC levels

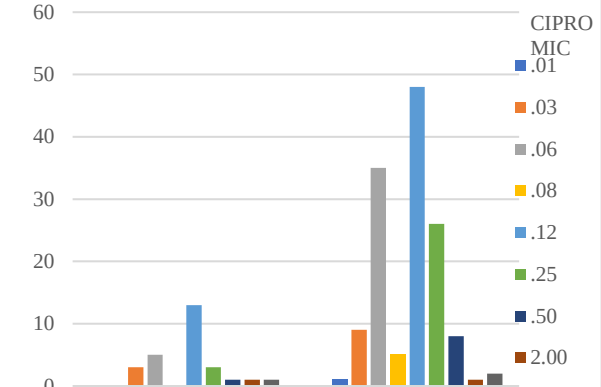


Figure 5.26 (D). Distribution of M23 (Anthranilate phosphoribosyltransferase G 96 A) across different MIC levels

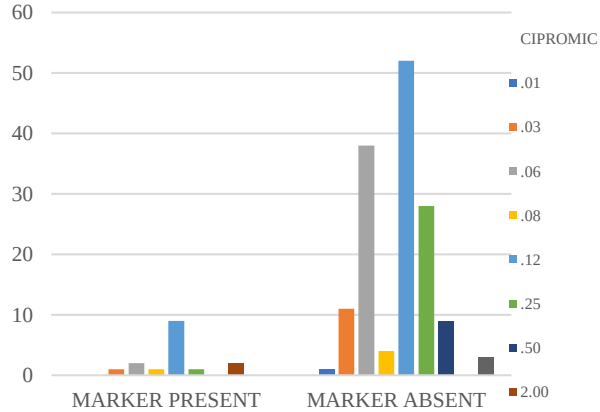


Figure 5.26 (E). Distribution of M34 (Amino-acid permease RocC C 1071 A) across different MIC levels

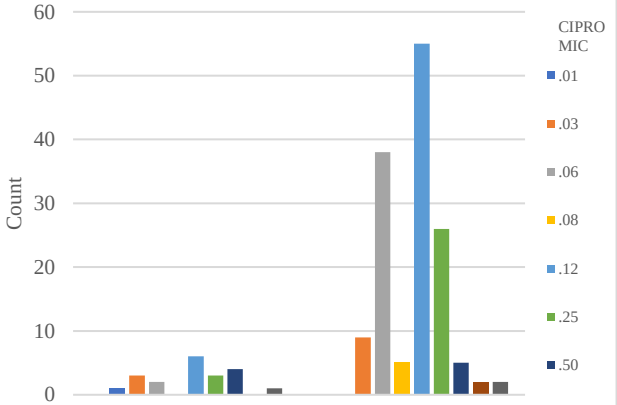


Figure 5.26 (F). Distribution of M44 (HTH-type transcriptional regulator YofA G 565 A) across different MIC levels

Figure 5.26 (A-F) shows six ciprofloxacin-associated molecular markers showing higher NPV as listed in Table 5.12.

The figure shows that the markers tend to be absent at higher gentamycin MICs.

5.7.3. Identification of new combinations of potential molecular markers showing improved diagnostic performance

Cluster analysis was applied to explore for the possibility of finding better combination of markers with better performance and the following 2 combinations of markers were identified.

Combination1: A combination of 3 molecular markers M20, M23 and M2 showed higher significance and effect size in relation to ciprofloxacin susceptibility phenotype than any of the markers above separately, (*chi*=75.997, *p*<0.0005, *phi*=0.685).

These 3 markers have divided the studied isolates into four clusters; Both Cluster 1 (11.1%, 18 isolates) and Cluster 4 (63.6%, 103 isolates) showed tendency to cluster towards lower ciprofloxacin MIC while cluster 2 (5.6%, 9 isolates) showed tendency to cluster towards higher ciprofloxacin MIC. This distribution is shown in Figure 5.27 (distribution of the four new clusters of molecular markers among different levels of MIC).

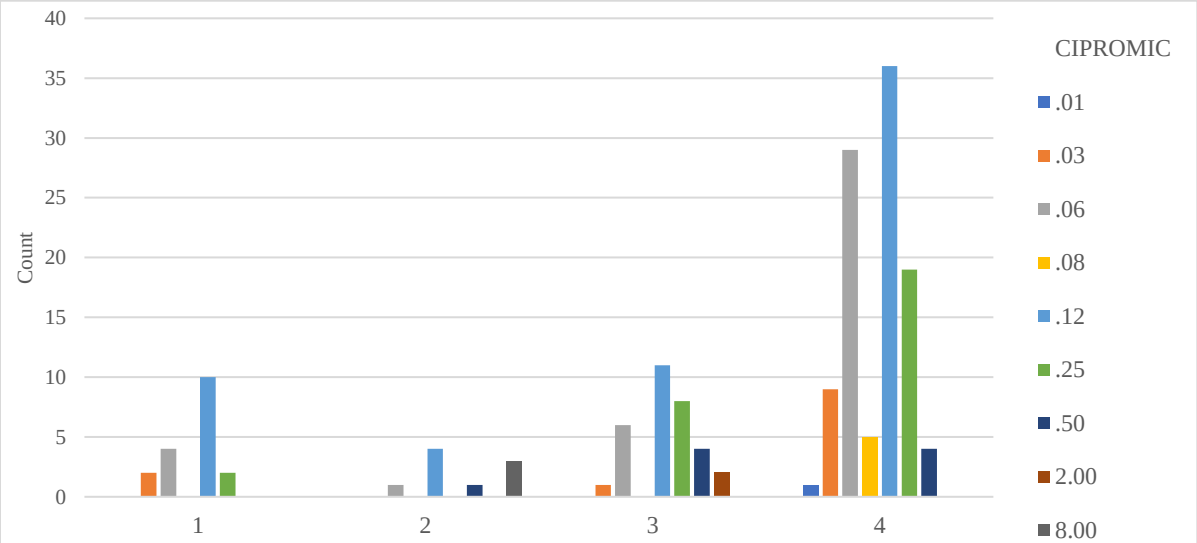


Figure 5.27. Distribution of the 4 clusters of molecular markers forming combination-1 across different levels of ciprofloxacin resistance

Predictive performance of cluster 2 was tested for resistance phenotype and showed sensitivity of 60%, PPV of 33.3%, specificity of 96.2% and NPV of 98.7% with (*chi*=29.147, *p*<0.005, *phi*=-0.424). Likelihood ratio of the cluster to differentiate different MICs= 24.357 with (*p*=0.002, *phi*=0.584).

Cluster 4 was also tested for its predictive performance for susceptibility phenotype which showed a sensitivity of 65.6%, PPV of 100%, specificity of 100% and NPV of 8.5% with

(*chi*=9.007, *p*=0.003, *phi*=0.236). Likelihood ratio of cluster to differentiate different MICs= 19.631 with (*p*=0.039, *phi*=0.317).

Table 5.13. Distribution of the components of combination 1 into four clusters in relation to ciprofloxacin susceptibility

	Cluster.4	Cluster.3	Cluster.1	Cluster.2
Size	63.6% (103)	19.8% (32)	11.1% (18)	5.6 % (9)
M20	Marker Absent (100%)	Marker Present (100%)	Marker Absent (100%)	Marker Absent (77.8%)
M2	Marker Absent (100%)	Marker Absent (100%)	Marker Absent (100%)	Marker Present (77.8%)
M23	Marker Absent (100%)	Marker Absent (78.1%)	Marker Present (100%)	Marker Absent (77.8%)
Average Cipro MIC	0.13	0.3	0.11	2.78
Cipro sensitivity	100% susceptible	93.8% susceptible	100% susceptible	66.7% susceptible

The table shows that the three markers (M20, M2, M23) tend to be absent at lower ciprofloxacin MICs and present at higher ciprofloxacin MICs.

Red Column to the right-hand side of each cluster shows “marker absence”

Red Column to the left-hand side of each cluster shows “marker presence”.

Distribution of CIPRO MIC in each cluster is shown as continuous scale with higher MICs shown to the right-hand side and lower MICs shown to the left-hand side.

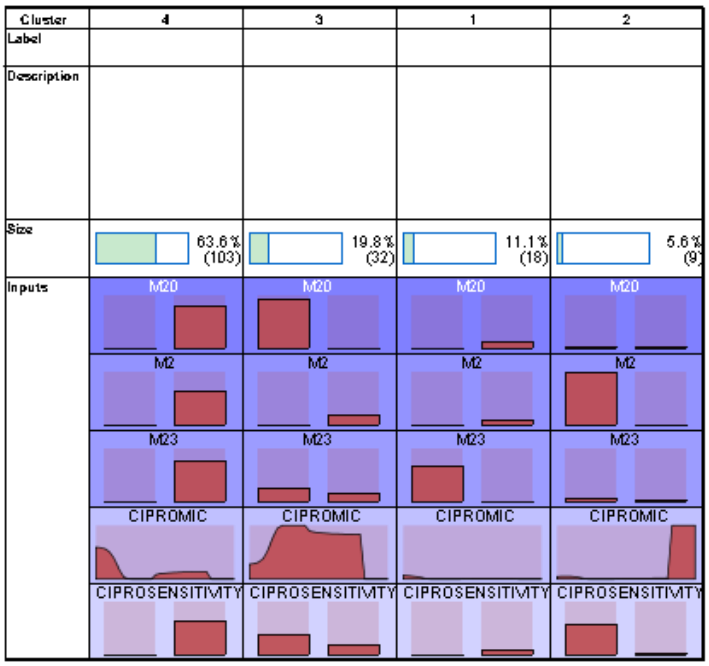


Figure 5.28. The figure illustrates the relative distribution of ciprofloxacin molecular markers composing combination 1 across the 4 identified clusters

This figure shows how markers are distributed (present/absent) in relation to ciprofloxacin MIC

Combination2: A combination of 5 molecular markers M20, M23, M2, *nal*C S209R and *nal*C A186T showed higher significance and effect size in relation to ciprofloxacin susceptibility phenotype than any of the above markers separately, (*chi*=88.068, $p<0.0005$, *phi*=0.737). The distribution of these markers is shown in Table 5.14.

Table 5.14. Distribution of the components of combination 2 into five clusters in relation to ciprofloxacin susceptibility

	Cluster.2	Cluster.1	Cluster.3	Cluster.5	Cluster.4
Size	44.4% (72)	21.6% (35)	17.3% (28)	11.7 % (19)	4.9% (8)
*nal*C S209R	Present 100%	Absent 97.1%	Present 100%	Present 94.7%	Present 50%
M20	Marker Absent 100%	Marker Absent 88.6%	Marker Present 100%	Marker Absent 100%	Marker Absent 75%
M23	Marker Absent 100%	Marker Absent 100%	Marker Absent 75%	Marker Present 100%	Marker Absent 87.5%
M2	Marker Absent 100%	Marker Absent 100%	Marker Absent 100%	Marker Absent 94.7%	Marker Present 75%

*nal*C A186T	Absent 100%	Absent 68.6%	Absent 100%	Absent 100%	Absent 87.5%
Average Cipro MIC	0.14	0.12	0.32	0.11	3.12
Cipro sensitivity	100% susceptible	100% susceptible	92.9% susceptible	100% susceptible	62.5% susceptible

Red Column to the right-hand side of each cluster shows "marker absence"

Red Column to the left-hand side of each cluster shows "marker presence".

Distribution of CIPRO MIC in each cluster is shown as continuous scale with higher MICs shown to the right-hand side and lower MICs shown towards left-hand side.

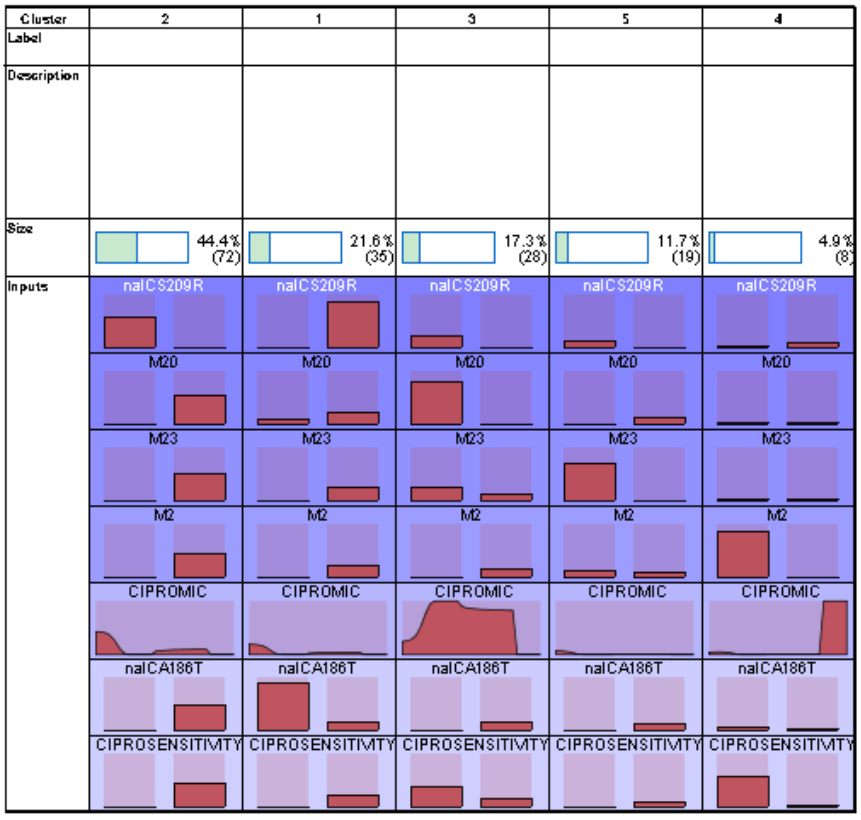


Figure 5.29. The figure illustrates the relative distribution of ciprofloxacin molecular markers forming combination 2 across the 5 identified clusters

The figure shows how markers are distributed (present/absent) in relation to ciprofloxacin MIC.

These 5 markers have divided the studied isolates into five clusters; Cluster 4 (4.9%, 9 isolates) showed tendency to cluster towards higher ciprofloxacin MIC while cluster 2 (44.4%, 72 isolates) showed tendency to cluster towards lower ciprofloxacin MIC. The distribution of the five new clusters of molecular markers among different MICs is shown in Figure 5.30.

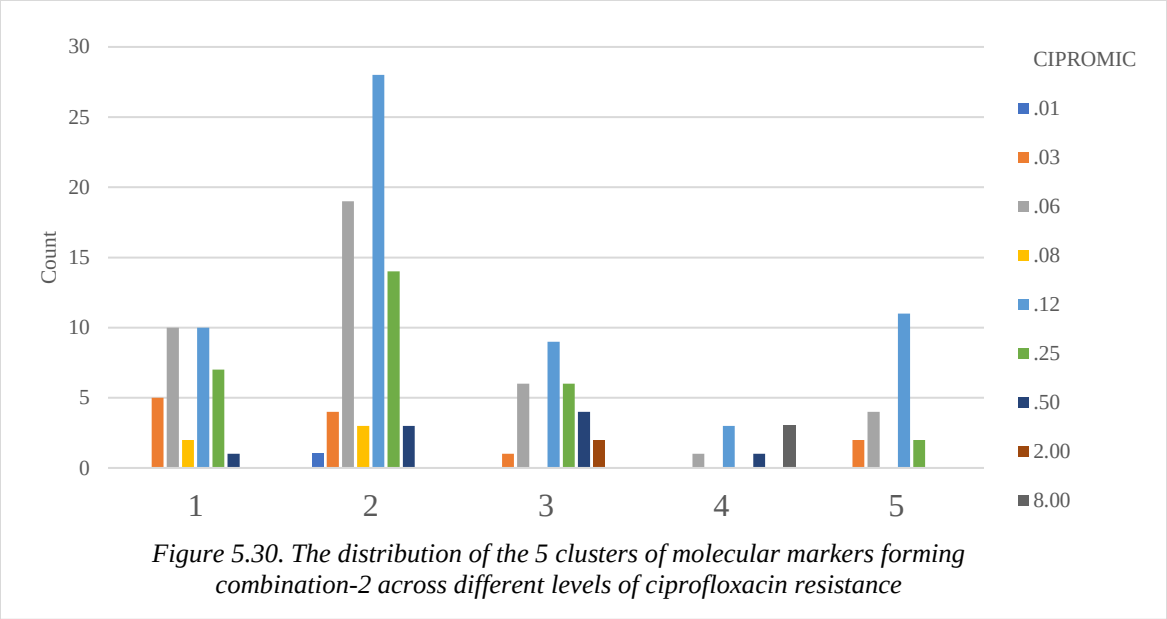


Figure 5.30. The distribution of the 5 clusters of molecular markers forming combination-2 across different levels of ciprofloxacin resistance

The figure shows that cluster 4 tends to occur at higher ciprofloxacin MICs while cluster 1 and cluster 2 tend to occur at lower ciprofloxacin MICs.

Predictive performance of cluster 4 was tested towards resistance phenotype and showed sensitivity of 60%, PPV of 37.5%, specificity of 96.8% and NPV of 98.7% with (*chi*=33.32, $p<0.0005$, *phi*=-0.454). Likelihood ratio of the cluster to differentiate different MIC levels= 24.174 with ($p=0.002$, *phi*=0.616).

Cluster 2 was also tested for its predictive performance towards sensitivity phenotype and showed a sensitivity of 45.9%, PPV of 100%, specificity of 100% and NPV of 5.6% with (*chi*=4.127, $p=0.042$, *phi*=0.16). Likelihood ratio of cluster to differentiate different levels of MIC= 9.438 with (*phi*=0.21)

5.7.4. Distribution of candidate loci in relation to genomic background

Figure 5.31 shows the distribution of all candidate ciprofloxacin loci analyzed in this section in relation to background phylogeny. Candidate loci distributions for studied 162 isolates of *Ps. aeruginosa* were then remapped to dendogram after removing recombinant regions to study the effect of recombination on loci distribution as shown in Figure 5.32.

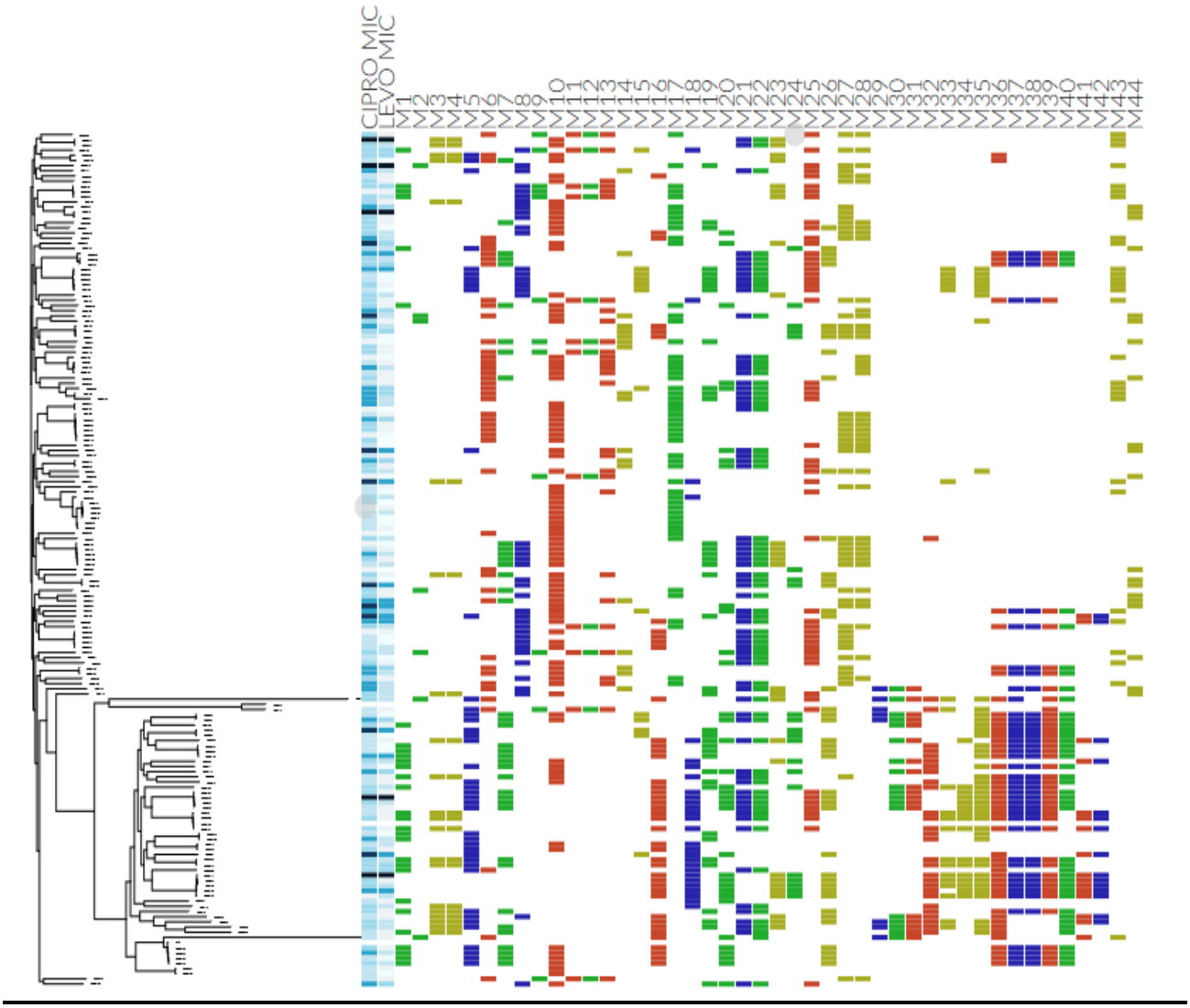


Figure 5.31: Distribution of candidate ciprofloxacin markers in relation to background phylogeny

Ciprofloxacin and levofloxacin susceptibility columns show resistance/ susceptibility as shades of blue where darker shades show higher resistance and lighter shades show lower resistance

Nucleotide changes evaluated at each marker position is shown in different color where Red color refers to “C” nucleotide substitution, Green color refers to “T” nucleotide substitution, Blue color refers to “G” nucleotide substitution, and Yellow color refers to “A” nucleotide substitution

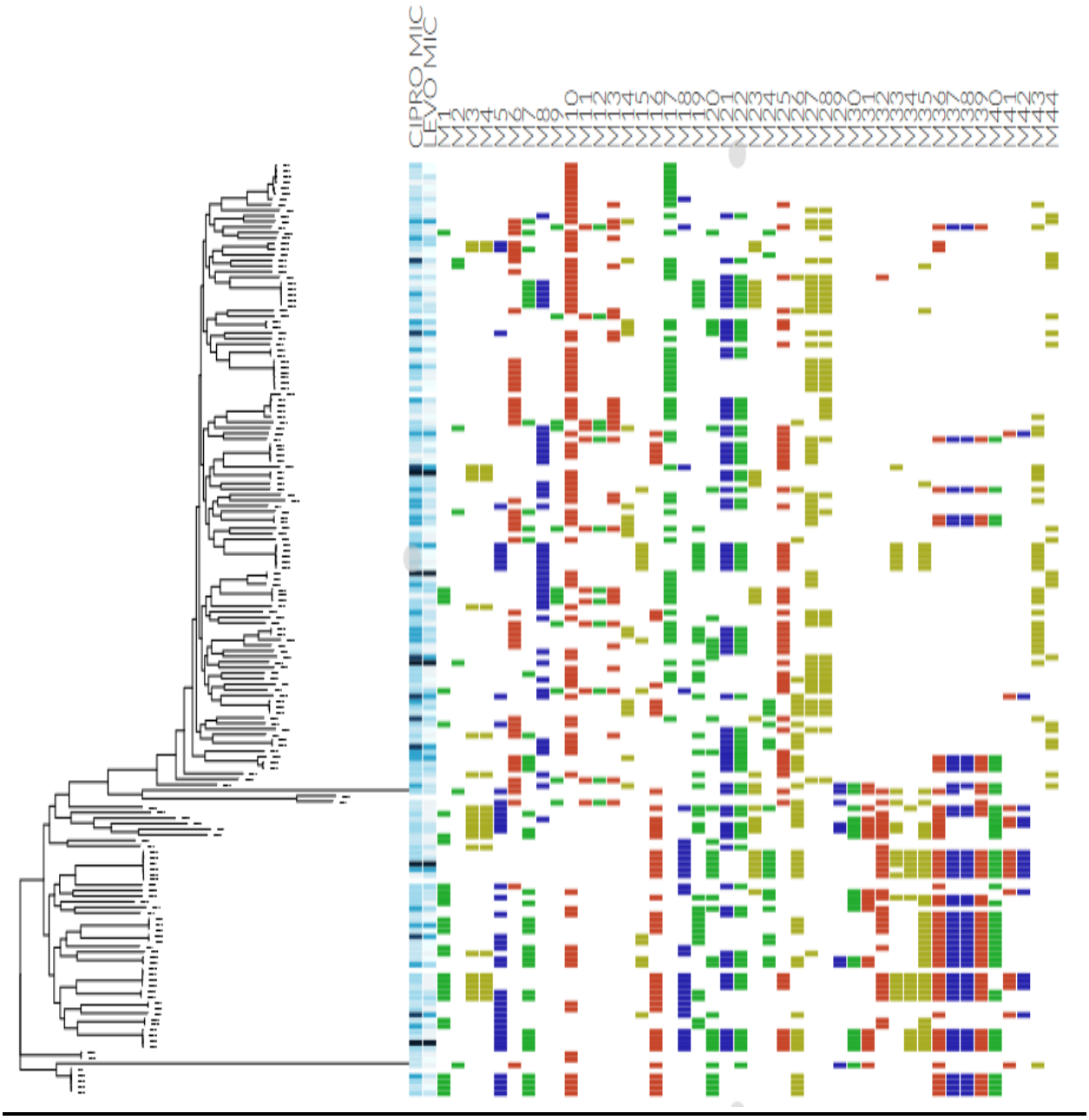


Figure 5.32. Dendogram showing re-constructed clonal genealogy mapped to candidate ciprofloxacin markers

Ciprofloxacin and levofloxacin susceptibility columns show resistance/susceptibility as shades of blue where darker shades show higher resistance and lighter shades show lower resistance

Nucleotide changes evaluated at each marker position is shown in different color where Red color refers to “C” nucleotide substitution, Green color refers to “T” nucleotide substitution, Blue color refers to “G” nucleotide substitution, and Yellow color refers to “A” nucleotide substitution

5.8. Results Section 5. Predicted Functional effect of amino acid changes (nSSNPs) variants

For all variants identified in results sections, functional effect of nSSNPs changes were tested to understand their possible role. These results would support some of the observed predictive values in addition to variants' distribution in relation to background phylogeny.

Table 5.15. Summary of predicted functional effect of non-synonymous SNPs identified as candidate markers

Marker	Gene	Amino acid change	PROVEAN Prediction		I-mutant Prediction	
			Prediction (cutoff -1.3)	PROVEAN score	Protein Stability	Reliability Index
M29 (gentamycin)	30S ribosomal protein S1	A129E	Neutral	0.210	Decrease	6
M13 (gentamycin)	Ribonuclease	G423A	Neutral	0.378	Decrease	1
M11 (ciprofloxacin)	Putative HTH-transcriptional regulator	V91L	Neutral	-0.765	Decrease	6
M13 (ciprofloxacin)	Putative HTH-transcriptional regulator	M126L	Neutral	0.502	Decrease	6
M14 (ciprofloxacin)	Mechanosusceptible channel MscK	A384T	Neutral	-0.751	Decrease	7
M19 (ciprofloxacin)	Yop proteins translocation protein Q	T223I	Deleterious	-2.040	Decrease	4
M29 (ciprofloxacin)	Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR	T5A	Neutral	0.048	Decrease	3
M37 (ciprofloxacin)	Proofreading thioesterase EntH	Q51E	Neutral	1.739	Decrease	2
M43 (ciprofloxacin)	HTH-type transcriptional regulator YofA	A171T	Deleterious	-1.385	Decrease	6
M44 (ciprofloxacin)	HTH-type transcriptional regulator YofA	V189M	Neutral	-0.161	Decrease	8

PROVEAN SCORE < -1.3 cutoff is predicted as deleterious

Reliability index shows the degree of reliability of the predicted effect on protein function stability, the higher the number, the more reliable is the prediction.

Different analyses including different ciprofloxacin variants analyzed in this section show that M14, M11 ,and M13 all showed very high specificity and PPV towards susceptibility phenotype and were all distributed across the upper section of the phylogenetic tree. These variants were also predicted to affect protein stability and may lead to loss of function. This may indicate the possibility of being used as susceptibility markers. On the other hand, ciprofloxacin markers M43 and M44 showed high NPV and were also distributed across the upper section of the tree carrying the possibility of being used as rule-out resistance markers.

5.9. Discussion

5.9.1. Novel insights into system-level functions associated with quinolone susceptibility phenotype

Results of comparative behavioral genomics has identified multiple co-shared genomic elements underlying quinolone susceptibility phenotype as shown in Results section 1. These genes and variants are identified as candidates for future research. Although the identified genes and variants were not all exhaustively studied, the potential importance of some of these elements and their link to the studied phenotypes in relation to findings from the literature will be discussed in this section. Chromosome partitioning *soj* and Gramicidin-S-synthase (non – ribosomal peptide synthase) are two genes identified as co-shared susceptibility markers between ciprofloxacin and norfloxacin.

These two genes can be linked to the lethal action of quinolones through their role in the pathway of chromosomal fragmentation and DNA damage that follows replication fork arrest (Drlica *et al.*, 2008). Two pathways are known to be involved in releasing DNA ends from QL-gyrase-DNA complexes. One pathway was linked to an identified protein factor, and is called the protein synthesis dependent mode, while the other pathway is known as the chloramphenicol-insusceptible mode or the protein synthesis independent mode. The lethal effect of quinolones is known to arise from the release of DNA ends from QL-gyrase-DNA complexes (Drlica *et al.*, 2009). Based on that, non-ribosomal type of peptides may be among the proteins responsible for releasing these complexes and consequently the final bactericidal effect of quinolones through the protein synthesis independent mode. In the current analysis, Gramicidin-S-synthase and linear gramicidin synthase can be considered as probable candidates for performing this role. In support of this finding are findings from a recent GWAS study investigating *Ps. aeruginosa* in CF which has identified non-ribosomal peptide synthase among the gene presence/absence genetic markers set. The analysis was based on using two different approaches including PLINK and random forest (Noah, 2019).

Results of the current CBG analysis has also shed the light on the possible role of Ethanolamine utilization in relation to quinolone phenotype. Both ethanolamine ammonia lyase and vit B12 transporter showed significant association at extremes of phenotypic behavioral groups and were prioritized from the current CBG analysis as shown in results section 1.

Ethanolamine ammonia lyase catalyzes the AdoCbl dependent conversion of ethanolamine (EA) to acetaldehyde and ammonia (Shibata *et al.*, 2010). The enzyme is considered essential for the growth of many bacteria on ethanolamine in the presence of exogenous vitamin B12 (Chang and Chang, 1975). This is also part of the pathway of glycerophospholipid metabolism (Kanehisa and Goto, 2000). It is the first enzyme in the degradative pathway of bacteria carrying the ethanolamine utilization (*eut*) operon (Kofoid *et al.*, 1999). Inhibition of this pathway showed an important link in organisms that carry ethanolamine utilization genes and that cause food poisoning (Tsoy, Ravcheev and Mushegian, 2009).

Alcohol dehydrogenase is another gene showing a co-shared resistance behavior to ciprofloxacin and ofloxacin. The gene showed tendency to present at more resistant groups as seen by its distribution and odds ratio for resistance behavior. This is illustrated in results section 1. Alcohol dehydrogenases are ubiquitous enzymes that present in nearly all life forms. These are oxidoreductases that catalyze the reversible oxidation of alcoholic compounds into corresponding aldehydes or ketones (Levin *et al.*, 2004). Evidence from the literature may support the role of alcohol dehydrogenase because many bacterial species possess alcohol dehydrogenase with its ability to produce the reactive and toxic acetaldehyde (SALASPURO, 1997). In addition, acetaldehyde production by microbial alcohol dehydrogenase has been linked to some pathogenic, toxic, and carcinogenic effect (Pavlova *et al.*, 2013). The current analysis shows that alcohol dehydrogenase is more distributed across resistant isolates showing a possible co-interaction in resistance and virulence mechanisms. This finding can be related to the predatory behavior of *Ps. aeruginosa* represented in its acetaldehyde production by alcohol dehydrogenase. This assumption may be supported by similar findings from other GWAS studies (Sutton *et al.*, 2019). The findings showed that formaldehyde secreted by *Ps. aeruginosa* together with its possession of other formaldehyde detoxifying enzymes are linked to its predatory behavior.

It has been suggested that mucin degradation can act as a carbon source contributing to the pathophysiology of the organism in CF lung. Tnseq analysis has shown that some genes related to amino acid biosynthesis and cofactor biosynthesis are linked to increased growth fitness (Flynn, Phan and Hunter, 2017). Both *trp*E ,linked to tryptophan biosynthesis, and *gsh*A ,linked to glutathione biosynthesis, were among the genes required for growth on mucin as identified

using TnSeq fitness screen (Flynn, Phan and Hunter, 2017). The same study has hypothesized that amino acids were being liberated from mucin and can act as primary source of carbon and energy. In the current analysis, a possible deleterious *trp*E variant and a *gsh*A variant were identified among the genetic markers showing higher effect sizes in relation to ciprofloxacin resistance. Although the link to CF phenotype was not studied here, this finding may indicate a possible overlap between quinolone resistance phenotype and CF pathogenicity.

It was also interesting to observe that some of the genes identified in the current analysis in relation to ciprofloxacin resistance have been previously identified as fitness determinants in *Pseudomonas fluorescence*. These determinants showed the ability to evade the host immune system in plant (Liu *et al.*, 2018). Cole *et al.*, (2017) have used transposon mutagenesis coupled with high-throughput transposon sequencing (Tnseq) and have identified hundreds of genes that increase fitness in wild type and in immunocompromised plant. Some of these genes were amino acid biosynthetic genes which have been similarly identified in previous other studies (Cole *et al.*, 2017).

Variants in *trp*E, *glt*B, *ant*A, and *ssu*D were linked to quinolone resistance as identified using the current CBG analysis. Research findings reported by Liu *et al.*, (2018) may support that through identifying the same genes, *trp*E, *glt*B, *ant*A, and *ssu*D as fitness determinants. These observations may indicate that general fitness determinants in the species are contributing to antibiotic resistance phenotypes and that *Ps. aeruginosa* clinical success especially in immunocompromised patients may originate from the general fitness characters the species possess. These determinants can be contributing to its success in wide variety of hosts including human and plant. The role of shared fitness determinants and its contribution to antibiotic resistance behavior may be underestimated in *Pseudomonas* species.

A very important finding from the output of the analysis performed in the chapter is observing a probable group of non-specific susceptibility determinant genes co-shared among different agents from the quinolone group. These genes include glutathione-S-transferase, azoreductases, and *p*-benzoquinone reductases. The genes have shown correlation with susceptibility behavior and also showed the same distribution in relation to background genomic context as illustrated in results section 1. A possible interpretation is that the three genes share similar functions acting as non-specific detoxification enzymes for any type of toxin or xenobiotic agents (non-nutritional

foreign chemical species). They may also act as general oxidative-stress related defense agents. The ability of any cell to survive any threat including both endogenously produced and exogenous xenobiotic agents is fundamental to survival and it is proposed that absence of these genes are related to the inability of the bacterial cell to detoxify toxic agents especially reactive oxygen species making the cell more susceptible to killing. The absence of these enzymes can probably make the cell easily killed and damaged by ROS and consequently make the bacterial cell susceptible to the action of antibiotics. However, the presence of these enzymes alone does not necessarily lead to resistance in all situations.

The first gene is one of the glutathione transferases (GSTs). These are major phase II detoxification enzymes found in the cytosol of living organisms. This family of enzymes can bind non-catalytically to a wide range of endogenous and exogenous ligands. They have wide range of functions including catalyzing substrates conjugation to glutathione, isomerases and peroxidase activities. Novel classes are continuously being identified in non-mammalian species carrying out functions that show similarities to other non-GST stress related proteins. These enzymes can provide protective function against H_2O_2-induced cell death by removal of reactive oxygen species and re-generation of S-thiolated proteins which are consequences of oxidative stress (Sheehan *et al.*, 2001). Although not widely recognized in relation to antibiotic resistance, there is some research evidence showing that glutathione-S-transferase can mediate Fosfomycin resistance in bacteria (Arca, Hardisson and Suárez, 1990). This mechanism is not known as a common mechanism in relation to quinolone resistance and is not commonly reported in recent literature among the main quinolone resistance mechanisms. However, an important finding from a recent GWAS investigating predation behavior in myxobacteria has shown that the same gene identified from CBG (Glutathione S-transferase GST-6.0) is linked to predation behavior in some studied organisms (Sutton *et al.*, 2019). In addition, the gene was reported among the fitness determinants genes in *Ps. fluorescence* using transposon mutagenesis coupled with Tnseq (Liu *et al.*, 2018).

Pseudomonas aeruginosa possesses three types of azoreductase enzymes. These enzymes have wide substrate activities and can act as NADH quinone oxidoreductases (Ryan *et al.*, 2010). Although it has been suggested that these enzymes play a role in quinone detoxification, the exact physiologic role of these enzymes has not been clear. It has been shown that these enzymes

have a wider role in the metabolism of a wide range of drug classes (Ryan *et al.*, 2011). Although azoreductases were previously considered a distinct group of NAD(P)H dependent flavoenzymes that have been identified and constitutively expressed in a wide range of bacteria (Ryan *et al.*, 2014) , the diverse nature of these enzymes along with their broad substrate specificity and their ability to reduce wide variety of both endogenous and exogenous compounds shows that they can play a general role in cell survival under adverse or stressful conditions. Results of comparative behavioral genomics analysis showed that three genes encoding for enzymes with co-shared function were significantly related to quinolone susceptibility behavior. These include Glutaredoxin arsenate reductase (*ars* H), NADH azoreductase, and NAD(P)H Dehydrogenase_(quinone). In support of that are findings from a recent GWAS study on *Ps. aeruginosa* showing several arsenate reductase genes including *ars*H. These findings show that *ars*H is related to different biofilm forming phenotype (Redfern *et al.*, 2019).

All these three enzymes share the common flavodoxin like functional fold and are Flavin mononucleotide (FMN) dependent enzymes that, despite having low sequence similarity, are recently grouped into the same family due to sharing the common function of NAD(P)H dependent quinone reductase. All three enzymes have a range of substrate specificity and one of their shared substrates is quinone. Quinones are toxic to *Ps. aeruginosa* and are produced as part of innate defense mechanism by many organisms to which *Ps. aeruginosa* is pathogenic. This includes plants, fungi, invertebrates, and human. *Ps. aeruginosa* has a number of enzymes with azoreductase-similar function and new enzymes are discovered and proving to belong to the same family of enzymes. These enzymes are usually related to the organism host range. This is a reason for their larger number in *Ps. aeruginosa* than in other *Pseudomonads* which may be attributed to the species wider host range. Quinones production is known as a defense mechanism produced by many hosts against *Ps. aeruginosa* and their toxicity is probably linked to their production of hydrogen peroxide. In the current CBG analysis, the genes were significantly absent in susceptible isolates. This means that their role is possibly linked to getting rid of the free radicals produced by quinones. It is thus assumed that absence of these genes makes bacteria more susceptible to antibiotic-mediated killing in general.

p-benzoquinone reductase was the third gene showing the same distribution as GST and azoreductases. Looking the literature up for benzoquinone reductases in *Pseudomonas aeruginosa* returned the *p*-benzoquinone reductase (*pnp*B) gene with an alternative name: NAD(P)H dehydrogenase (quinone). The gene was identified in *Pseudomonas sp.* (strain WBC-3)(Zhang *et al.*, 2009) as involved in the degradation of para-nitrophenol (PNP) by catalyzing the reduction of p-benzoquinone to hydroquinone. This protein is involved in the pathway of 4-nitrophenol degradation, which is part of Xenobiotic degradation. Search results also returned the gene PA1225 which has been previously defined as "probable NAD(P)H dehydrogenase" and was also recently re-annotated as "NADPH: quinone reductase" acting on 1,4-benzoquinone and 2,6-dimethoxy-1,4-benzoquinone as substrates rather than on quinones which are the substrates for azoreductases (Flores and Gadda, 2018). This may indicate that this gene also shows the co-shared non-specific function of getting rid of other classes of xenobiotic agents.

5.9.2. Exploiting GWAS for the identification of predictive genomic markers

The new era of high throughput sequencing technologies can greatly inform the way scientists use genomic information to impact decision making for public health interventions. Applying WGS-based comparative genome analysis to 162 isolates of *Ps. aeruginosa* has identified some candidate genomic markers as potential molecular diagnostics. Identified candidates were tested for their informative value as a diagnostic on the basis of their predictive values and their potential effect on gene function in the frame of genetic background.

The phylogenetic approach is the most popular method used to describe the genetic relatedness and the population structure. Phylogenetic analysis allows the detailed identification of genetic relatedness for sub-populations and also for individuals. Genome comparison aims at identifying and inferring the biologic significance of similarities and differences. To explore whether identified genetic markers are truly related to the studied property of interest, prioritized markers were examined in relation to phylogenetic background to exclude spurious results that originate from the confounding effects of population structure or recombination.

Mutations are considered an important source of variation in any organism and can lead either to decreased fitness or to beneficial adaptation. Therefore, determining the distribution of mutations in different genetic backgrounds and their related fitness effect is important for a proper understanding of important traits. An additional point to consider when interpreting genetic

differences from association studies is their effect on protein function. The effect of genetic differences on protein function can vary widely making it difficult to interpret phenotype-genotype relationship. Single point mutations can affect protein function in several ways including; stability changes, conformational changes, physical and chemical structural changes, changes in electrostatic properties, and intermolecular interactions. In addition, changes in subcellular localization which may also affect normal cell function through changing protein concentration. This also needs to be considered. Although the ability to predict the functional effect of observed variants is complicated, it can offer an additional line of evidence that helps in differentiating the fitness effect of mutations. In the current analysis, I-mutant was used to identify free energy change upon mutation in order to predict mutation effect on protein stability. PROVEAN was used to identify the possible variant effect on protein function and activity.

Although exploring the functional importance of the newly identified antibiotic susceptibility/resistance related mutations from GWAS may be complicated by other epistatic interactions influencing these correlations, the identified variants were studied by exploring their functional effect, their distribution in relation to background genomic context, and by reviewing the literature for the possibility of finding a functional correlation.

The problem of population stratification encountered in bacterial GWAS can be especially problematic with highly clonal species (Read and Massey, 2014) and this is not the case in the current analysis of *Ps. aeruginosa* which is known to be poorly clonal. This consequently means that the distribution of variants observed here in relation to phylogenetic background may represent a strong association signal rather than a spurious effect of population stratification especially with a highly recombining species like *Ps. aeruginosa*. In the current analysis and as shown in the results section 2 and section 3, the corrected phylogenetic tree showed changed inter-isolate genetic distances exhibited as different branch lengths as a result of removing recombinant regions, however, the overall structure of the population remains separated into two large distinct groups. In addition, the differential distribution of candidate markers across the two large phylogenetic clusters remains the same which may indicate that the markers shown are probably more informative. These observations may indicate that markers clustering through the lower subpopulation which is affected by recombination may offer real adaptive advantage as

opposed to markers clustering through the upper half of the tree which may be associated with decreased fitness advantage.

Analysis of the distribution pattern of candidate gentamycin markers in relation to their background genomic structure showed that some variants showing high NPV towards gentamycin resistance phenotype are more frequently occurring in isolates on the upper section of the tree; i.e. occur earlier in evolutionary history. These markers have the potential to be used to rule-out resistance. These include; M7, M9, M15, M17, M24 and M33.

On the other hand, it appears that M31, M32 and M36 are beneficial mutations. The mutations can provide incremental fitness advantage to resistant isolates. These variants showed to be distributed at the lower part of the phylogenetic tree which is probably more affected by recombination. The possible higher fitness combined with higher PPV and higher specificity of these variants to resistance behavior may indicate the possibility of using these markers to rule-in resistance. Findings from the literature can support the current CBG findings for M36 (cytochrome c oxidase subunit 3). Results of GWAS identifying clusters of predation genes in myxobacteria showed that cytochrome oxidase subunit 1 and subunit 2 were among the genes identified as predation genes at 100% specificity and 44-50% sensitivity among the studied group of organisms (Sutton *et al.*, 2019). PA3431 is a hypothetical cytoplasmic membrane in *Ps. aeruginosa* showing a functional prediction from interproscan with functional domains similar to *cid*B/*Lrg*B family (Winsor *et al.*, 2016). This membrane proteins have been previously linked to penicillin tolerance in *Staphlococcus aureus* (Rice *et al.*, 2003) (Groicher *et al.*, 2000). Findings from the literature also show increased expression of PA3431 in response to low oxygen conditions (Alvarez-Ortega and Harwood, 2007), (data available at https://www.ncbi.nlm.nih.gov/geoprofiles/43497634). These findings from the literature can support the assumption from CBG about the possible role of M31 and M32 (variants of PA3431) in increasing fitness which can consequently contribute to increased antibiotic resistance as shown above.

Some other gentamycin markers forming best predictor combination 1 (including M20, M28, M37) and combination 2 (including M20, M28, M29) with better combined predictive values and higher effect sizes did not show clonal or evolutionary clustering and were distributed across different branches of the evolutionary tree. This may increase the informative value of these

markers because such a distribution may indicate a genuine link to important biologic traits. Sutton *et al.*, (2019) has identified the same protein (30S ribosomal protein S1) as one of the predation genes in myxobacteria using GWAS. This finding can support the CBG finding shown here with M29 (a 30S ribosomal protein S1 variant). Another variant identified with better combined predictive performance is another ribosomal modification protein variant (M28). The Ribosomal protein S6--L-glutamate ligase (PA5197), with a functional domain similar to Escherichia coli *rim*K (Winsor *et al.*, 2016) is known to add additional Glu residues to the native Glu-Glu C-terminus of ribosomal protein S6. Mutation of the Glu-Glu terminus to Lys-Glu has shown to block this addition (Zhao *et al.*, 2013). This fact can explain the role of *rim*K variant (M28) identified from CBG analysis in affecting gentamycin resistance.

Esterase *est*A (PA5112) variant identified from CBG (M37) in relation to gentamycin resistance appears to be linked to many other phenotypic behaviors in *Ps. aeruginosa. est*A function was probably first identified as a lipolytic enzyme which is related to cellular membrane by Wilhelm, Tommassen and Jaeger, (1999). The study has shown that this esterase was unique as it was the first identified in the outer membrane and is the first example of type IV secretion system (Wilhelm, Tommassen and Jaeger, 1999). Later, reports have shown the possible role of membrane bound esterase activity in the production of rhamnolipids which are known to affect outer membrane composition, biofilm formation, and cell motility (Wilhelm *et al.*, 2007). More recently, the relation of *est*A expression modification to the overproduction of rhamnolipids has been confirmed (Dobler *et al.*, 2017). Rhamnolipids have multiple roles and their presence is known to alter cell surface polarity and to promote the uptake of hydrophobic substrates (Al-Tahhan *et al.*, 2000), a fact that would interpret its possible relation to gentamycin resistance as identified in the current CBG analysis. Another recent research finding also supports that role. Radlinski *et al.*, (2019) have recently found that rhamnolipids produced by *Ps. aeruginosa* can induce proton motive force-independent aminoglycoside uptake in *Staph. aureus* and showed that rhamnolipids can potentiate aminoglycoside activity by restoring sensitivity to tolerant, persister, or the biofilm small colony variant of *Staph. aureus* (Radlinski *et al.*, 2019).

Combined testing for ciprofloxacin showed that M14 is an important candidate. M14 showed 100% specificity and 100% PPV towards sensitivity phenotype. In addition, it showed to be exclusively distributed across the upper half of the tree and lost in the lower half of the tree; i.e.

more recently, probably because it is disadvantageous. M14 (Mechanosusceptible channel MscK A384T) showed the highest significance and the highest effect size (*phi*=0.376, *p*=0.003) among identified variants. When tested for its functional effect, I-mutant showed that it is possible for the variant to exhibit decreased protein stability with high reliability index (RI=7). Although PROVEAN did not predict the mutation as deleterious, PROVEAN score was -0.751 which is near to the cutoff point. These pieces of information taken together can indicate the higher informative value of the marker as susceptibility- related marker. Findings from the literature can support the possible role of PA1408 variant (M14) identified from CBG. Expression profiling studies showed that the gene is overexpressed in small colony variant (Wei *et al.*, 2011), data available at https://www.ncbi.nlm.nih.gov/geoprofiles/91464421. This would support the assumption shown in the current analysis about the identified variant which may lead to loss of gene function or protein instability in relation to antibiotic susceptibility.

Similarly, ciprofloxacin markers M11 (Putative HTH-transcriptional regulator V91L) and M13 (Putative HTH-transcriptional regulator M126L) showed 100% PPV and 100% specificity and were exclusively distributed across the upper half of the phylogenetic tree. When tested for functional effect, I-mutant showed that variant may show decreased protein stability (reliability index, RI=6). Although not predicted as deleterious by PROVEAN, the SNP change needs to be tested for its possible effect on protein subcellular localization. This can affect its concentration and subsequently its function.

Ciprofloxacin marker M44 (HTH-type transcriptional regulator YofA V189M) showed NPV of 97.2 % to resistance phenotype. This means it has the potential to be used to rule out resistance. The marker showed high effect size (*phi*= 0.369, *p*= 0.005) and was similarly exclusively distributed across the upper half of the phylogenetic tree. Functional effect testing showed decreased stability with reliability index=8. Although M43 showed lower effect size (*phi*=0.107, *p*=0.172), it showed high NPV (97.8%) and was also distributed across the upper half of the tree. Moreover, it was predicted as deleterious using PROVEAN at -1.3 cutoff. I-mutant predicted decreased stability with reliability index=6 in relation to the same variant.

Results of the analysis performed in this chapter have collectively identified new functional elements in addition to markers with good predictive values. However, some may occur in lower frequency in the population. The current results need to be re-tested in a larger population in

order to re-test for the frequency of encountering these variants and their possible association with the studied phenotype in a larger set of population. The analysis performed in this chapter has also shown that multiple lines of evidence are usually required to assess any candidate marker for its practical applications. Statistical significance tests cannot be considered as the only guide about the usefulness of observed markers because they can often be associated with many flaws that commonly result from choosing cutoffs for comparing groups. Using predictive values, effect sizes, and information about variant frequency in the population in addition to testing the functional effect of variants under study should all be considered collectively.

A possible limitation of the current analysis is the lower number of resistant strains included especially with ciprofloxacin. This may falsely underestimate the significance of tested correlations. Another possible limitation is the need to use breakpoints for the calculation of predictive values. Breakpoints were used because they are considered the standard practice agreed and used in diagnostic laboratories and in clinical practice to prescribe treatment. However, using breakpoints to classify groups suffers from the limitation of being not completely informative; carrying the possibility of misclassification. This can occur at both borders of the cutoff points or at was previously considered as intermediate category. This category can carry some clinical and practical uncertainty about clinical treatment success.

Declaration of contribution:

CBG analysis output as described in section 5.3.2 in methods section was provided through Brunel Systems and Synthetic Biology lab. I would like to thank Prof. Nigel Saunders and Mr. Arshad Khan for providing and running the CBG informatics work analysis and providing CBG sheets for further analysis and interpretation.

Chapter 6. General discussion and conclusions

6.1. Role of rapid diagnostics in addressing antimicrobial resistance problem

6.2. Opportunities and limitations for technologies used in rapid bacterial identification and antibiotic resistance profiling

6.3. Challenges for developing rapid point of care diagnostics for antibiotic resistance detection

6.3.1. Lack of complete understanding of genotype-phenotype relatedness

6.3.2. Diversity of resistance mechanisms

6.3.3. Geographic and personal variation

6.3.4. Guidelines and benchmarking for transforming research findings into applied diagnostic tools

6.4. Thesis Conclusions and future directions

6.5. Example: Rule-In algorithm

6.6. Example: Rule-Out algorithm

6.7. General Summary

6.8. Applied/practical example

6.1. Role of rapid diagnostics in addressing antimicrobial resistance problem

Although great advances have been achieved in diagnostic technologies, empiric antimicrobial prescription is still widely used to deal with suspected critical infections. This consequently results in overuse of our small inventory of effective and last line antimicrobial agents. This leads to the aggravation of antibiotic resistance problem by driving the emergence and spread of multi-drug resistant organisms. The gap that currently exists between the traditional microbiology workflow and the need for more rapid results, especially in some critical conditions, has led to the current problem of overtreatment. Rapid diagnostic tests for infectious diseases need to distinguish bacterial from viral infections, to identify the type of infectious agent, and to provide reliable information on susceptibility to antimicrobial agents. The test needs to be simple and rapid and also to provide results within hours. Rapid diagnostics should be able to clarify disease etiology, to influence treatment choice, and to enable public health surveillance and interventions. In addition to avoiding the cost of over-prescription and reducing time-to-results, the application of rapid diagnostics helps to avoid the complications and patient mortality that may arise from antimicrobial therapy and prolonged hospitalization.

Rapid diagnostics used to identify antimicrobial susceptibility (AST) profile within hours can make significant advancement in the management of infectious diseases by improving antimicrobial use and hence, clinical outcomes. The timely guided administration of the right antibiotic within few hours can be lifesaving in some critical cases. Although standard phenotypic testing requires 48-72 hours to provide the final results, rapid diagnostics are capable of identifying the correct answers in a shorter time. Although several approaches have been proposed as possible solutions to antibiotic resistance problem as discussed in *Chapter 1*, a practical and helpful way for defeating the emergence of antibiotic resistance is to implement targeted antibiotic prescription practices through using point of care (POC) testing by general practitioners (GPs).

Although some point of care diagnostics currently exists to differentiate bacterial from viral infection and to detect specific types of target organisms and resistance markers, the feasibility and the practicality of developing antibiotic resistance detection panel for use as rapid point of care diagnostic (POC) is considered challenging and more demanding than appreciated. Some of the currently available multiplex diagnostics panels include antibiotic resistance markers in

addition to species identification function, however, it is usually considered a real challenge to choose the most informative resistance targets to be reliably included into a rapid diagnostic.

6.2. Opportunities and limitations for technologies used in rapid bacterial identification and antibiotic resistance profiling

As an alternative to growth-based methods, molecular diagnostic tests aim at the rapid detection of antibiotic resistance genes from bacterial cultures or from primary specimens within hours of collection from patients. Although the main advantage of rapid molecular-based AST diagnostics lies in its capability to significantly reduce turn-around time to impact patient management, available tests often lack the representative set of antibiotic resistance genes that enables comprehensive prediction of resistance phenotype. On the technical side, the primary stages of sample processing needed before molecular-based systems can be used reduce the chance of full automation. Primary sample processing may require pre-culture steps due to the low numbers of target pathogens that may exist in some complex biologic samples

Different molecular platforms and accelerated phenotypic systems are currently available for rapid bacterial ID/AST testing as discussed in *Chapter 1*. Accelerated phenotypic systems for rapid bacterial ID/AST testing depend on confining bacteria in small volumes at single-cell levels in microfluidic devices to accelerate bacterial growth and biochemical reactions. Marker concentrations in these isolated environments can then reach detectable levels much more quickly to be measured. Microfluidic systems can be used to determine resistance profiles based on electrochemical reduction of a redox-active molecule resazurin, and can detect growth at very low concentrations of 1 CFU/ml (Besant, Sargent and Kelley, 2015). On the other hand, molecular platforms used for detection of antibiotic resistance markers use nucleic acid-based markers indicative of the presence of bacteria and/or antibiotic resistance for bacterial ID/AST as discussed in *Chapter 1*. The main challenge here is related to the limited spectrum of the genes included in the detection panel. Detection panels often include specific types of acquired resistance rather than intrinsic resistance elements. This includes most of currently available rapid “Multiples Syndromic Panels”. Most available panels detect carbapenemases-related genes or Extended-spectrum B-lactamases (ESBL). For example, Xpert CarbaR® detects 5 gene families; KPC, NDM, VIM, IMP and OXA-48; Verigene® system detects the most common carbapenemases and CTX-M ESBL and FilmArray® BCID detect KPC (Tuite *et al.*, 2014). The

scope being investigated here is the potential of sequence-based resistance prediction and markers' interpretation can offer within the background genomic context. The purpose is to offer evidence about the use of sequencing technologies as an alternative automated platform for infectious disease management using *Ps. aeruginosa* as an example.

6.3. Challenges for developing rapid point of care diagnostics for antibiotic resistance detection

Although different platforms exist for rapid bacterial identification and antibiotic resistance profiling, the main challenge remains in choosing the best set of molecular markers for inclusion in the identification platform. These challenges are summarized in the following subsections.

6.3.1. Lack of complete understanding of genotype-phenotype relatedness

The comprehensive understanding of the phenotype-genotype relationship is a difficult to answer biologic question. Although analysis of microbial phenotypic networks inter-relatedness suggests common genetic underpinning, this can be jeopardized by different factors including different gene repertoires, codon usage bias and proteome compositions (Brbić *et al.*, 2016). This basic biologic concept makes the prediction of any phenotypic trait a challenging task.

6.3.2. Diversity of resistance mechanisms related to different antibiotic targets, different classes, and different bacterial species

It is not feasible to develop POC diagnostics capable of detecting all different antibiotic resistance targets for all species using one tool due to the wide array and multiplicity of antimicrobial resistance mechanisms (van Hoek *et al.*, 2011). Instead, prioritizing individual markers and combination of markers based on their diagnostic predictive values and their frequency of occurrence in different bacterial populations can offer the first valuable step to antibiotic resistance diagnostics innovation. Choosing marker combinations that act as resistance targets needs to be optimized to include the best performing markers. This is important because the inclusion of less informative markers can negatively affect the tool "value: cost" ratio. Inclusion of a large number of antibiotic resistance targets in the rapid POC diagnostic panel can consequently increase the cost without adding much diagnostic information.

Novel POC rapid diagnostics can be developed to include an informative combination of resistance targets. This can be used to build a flexible adaptable format which allows the inclusion of new types of emerging resistance targets (Mitsakakis *et al.*, 2018). Optimized choice of best-performing combinations needs much investigation before being translated into clinically applied tools. First, because different bacterial species carry different types of intrinsic resistance and second, due to different types of transferable resistance mechanisms carried by different species.

6.3.3. Geographic and personal variation

Having an up-to-date comprehensive knowledge about the incidence and the changing trends of specific resistance mechanisms in different geographic localities is not easily achievable. The ongoing genomic epidemiologic analysis of different resistance mechanisms and monitoring of the emergence of new variants' is necessary to get a guided grasp about the characters of the bacterial populations and the resistance genes prevalent within target locations. Diagnostic innovators need to have access to updated epidemiologic data on antibiotic resistance for targeted health care setting. This can be readily achieved when sequencing becomes integrated into routine laboratory diagnostic workflow. This is considered an indirect but a beneficial outcome of introducing different sequencing technologies in diagnostic clinical settings. Other challenging points include the changing geographic-related epidemiology due to travel or ethnic background-related carriage and the confounding effect of normal human non-pathogenic microbiota which might carry some transferable resistance genes leading to false positive results.

6.3.4. Guidelines and benchmarking for transforming research findings into applied diagnostic tools

Implementing innovative strategies to prioritize the best-performing markers is considered a valuable step for in-depth exploration for the elements underpinning antibiotic resistance phenotypes. Nevertheless, there remains the challenge of absence of guidelines needed to transform research findings into applied tools. Finding cutoffs for predictive values that shows which resistance-related markers are practically used is a highly challenging point.

The clinical integration of improved diagnostics is considered a demanding process. It requires the fulfillment of a wide landscape of practical, technical, strategic, and financial needs. This

includes the stimulation of research and development diagnostics to translate new technologies and research findings into practical tests. The development of the guidelines needed to conduct the diagnostic tests and to communicate the results in addition to ensuring that the necessary healthcare infrastructure, including personnel and information technology are met, are all needed for practical application.

To practically adopt a new test, consensus guidelines for diagnostic test interpretation need to be accepted by professional societies for incorporating the into clinical practice. This depends on many factors including the availability of clinical outcome data and cost-effectiveness of the test, especially in resource-limited settings. All these factors can form major challenges for adopting a new diagnostic test. Another challenging operational point is the requirement for assay validation and verification of performance.

6.4. Thesis Conclusions and future directions

It has been discussed in *Chapter 2* that the special nature of *Ps. aeruginosa* makes AMR prediction in the species specifically challenging (Jeukens, Freschi, Kukavica-Ibrulj, J.-G. Emond-Rheault, *et al.*, 2019). This is in part attributable to the large pan-genome of the species in addition to the species-specific complicated regulation of resistance. In their work, J. Jeukens *et al.*, (2017) highlight the complexity of correlating phenotype and AMR genotype in *Ps. aeruginosa*. Their results showed that the identified resistome using the CARD database does not correlate with phenotype and that some resistance genes can be found in some susceptible isolates, the fact that was supported by the analysis performed in *Chapter 2.* That is the reason why it is important to consider that even antibiotic susceptible strains of the species can carry some resistance mechanisms due to intrinsic AMR, the fact that makes it necessary to differentiate and understand the relative importance of intrinsic resistance as compared to clinically meaningful resistance. Another possible source of bias in predicting clinically important resistance determinants in *Ps. aeruginosa* is the current lack of clarity of understanding about the difference of gene content for virulence and AMR between clinical and environmental isolates (Freschi *et al.*, 2015). *Chapter 2* concludes that available NGS-based pipelines used for identifying genomic bases of resistance do not perform well in predicting quinolones and aminoglycosides resistance in *Ps. aeruginosa* and shed light on some practical steps needed to adopt the use of NGS in diagnostic settings.

Based on these results, the known gene panels used to predict resistance seem to be insufficient to predict a specific phenotype with sufficient certainty. The tools used showed high false positive results because it does not consider all known gene variants but instead, only a specific subset of genes. This necessitates more comprehensive evaluation and investigation of system-level resistance determinants in order to provide more comprehensive understanding of all possible resistance mechanisms to be included into a more inclusive predictor panel of markers.

This objective was achieved in the thesis in two steps. The first step was to evaluate all the resistance elements known from the literature and to find possible clusters or combinations of molecular markers that can provide a better prediction of resistance. The second step was to apply genome wide association approach to mine for unknown elements underlying resistance. Out of 57 tested molecular markers identified from the literature, 13 markers were prioritized to have the most significant association with quinolone resistance phenotype and 5 clusters of combined markers were identified as well. For Aminoglycosides, out of 51 molecular markers, 8 were prioritized as the best cluster combination of predictors.

Different methods of typing are usually used to describe the population structure of different bacterial species and to provide the framework that helps to characterize lineages associated with resistance, virulence, or other high-risk characters. These markers can then be linked to the risk or to the complications of health care related infections. Studying the population structure helps to identify the biologic significance of candidate markers in relation to background genomic context. Understanding the population structure combined with the predictive capacity of some identified molecular markers can help to efficiently reduce the spread of high-risk clones by designing better diagnostic markers and better targets for improved infection control and other intervention strategies. In *Chapter 4*, epidemic high-risk clones were studied in relation to the previously identified molecular markers. Different markers as well as combinations of markers have shown significant association with some high-risk clones. These pieces of information can add to the practical value of using these markers in molecular diagnostics.

Although phenotypic AST testing is currently considered the gold standard, it has some drawbacks. Chief is the “time to result” and also is its sensitivity when compared to nucleic acid amplification- based methods (Maurer *et al.*, 2017). The majority of the currently available POC diagnostics for infectious diseases rely on nucleic acid-based amplification including PCR or

isothermal amplification (Craw and Balachandran, 2012). Although there is still an argument about PCR being superior to genome sequencing in terms of cost and sensitivity of detection, antibiotic resistance occurring as a result of mutations or gene variants are less easily detected using nucleic acid amplification.

OpGen is a newly developed high-throughput multiplex PCR that can test for a large panel of genes (Walker *et al.,* 2019). The majority of genes included in the test are acquired resistance or plasmids-associated genes. These were chosen based on surveys of resistance genes databases. Although the genomes included in the assessment for the purpose of developing this Multiplex PCR tool were chosen from different continents and they included 1484 *Ps. aeruginosa* isolates, no information about the genetic diversity of the isolates were provided, the point that may question the informative value of the panel of markers included.

Although it has been shown that resistance genes included in the multiplex panel have included the resistance mechanisms for the isolates assessed in the study, it is important to consider more inclusive core-genome based markers. This can be considered a more cost-effective approach. Including a very large number of horizontally acquired non-specific elements does not achieve high cost-value advantage. Acquired and enzymatic resistance cannot be reliably used to judge overall resistance in all situations. Acquired resistance is not species-specific and is only part of the resistance baggage carried by the bacterial cell. The same theoretical drawback may also apply to other accelerated phenotypic methods used for the purpose of rapid AST testing. These usually depends on detection of antibiotic-hydrolyzing enzymes very early in the growth cycle using colorimetric methods. It becomes obvious here that choosing the best potential set of molecular markers for predicting antibiotic resistance phenotype at the practical level represents the current gap in knowledge addressed in the thesis as shown in *Chapter 1*. The thesis also suggests a diagnostic and decision algorithm for antibiotic prescription based on the predictive values studied for the markers. For example, if we have some sequencing information available for *Ps. aeruginosa*, the studied markers shown can be used as a list to check for markers fulfilling certain criteria of performance/predictive values in the sequence and steps shown in Figure 6.1, Figure 6.2, and Figure 6.3. This is also illustrated in the examples shown in sections 6.5, section 6.6, and section 6.8. These markers' lists can be incorporated into a diagnostic panel or can be identified using metagenomic or NGS diagnostic platform.

The relative importance of each resistance-related mechanism and the quantitative contribution of each type of resistance remains an unanswered question. The notion of different resistance and virulence elements showing combinatorial effect and building towards determining the phenotype has been shown in the literature. However, no information is available about the relative importance or the quantitative contribution of different types of resistance-related mechanisms on overall resistance and this was one of the objectives of the analysis performed in *Chapter 3*. Having such information about the variants showing higher quantitative contribution indicates that these variants can be prioritized for inclusion into diagnostic panels especially when combined with higher predictive values and this is also considered a new way to use existing knowledge.

For antibiotic target changes, *gyr*AT83I remains an important contributor to both ciprofloxacin and levofloxacin resistance phenotype as seen in the predictive model in section 3.5.5 and section 3.6.5. Other quinolone target site mutations including *par*CS87L, *par*CS87W, *par*CE91K, and *gyr*AD87N appeared to have some contribution to levofloxacin resistance as seen in the predictive model in section 3.6.5. On the other hand, target site changes did not appear to be a main contributor to aminoglycoside resistance. This may seem obvious by the very low frequency of mutations occurring at ribosomal target sites of the aminoglycoside group of antibiotics as reported in the literature. However, the novel variant *fus*A1D588G in *fus*A1 gene coding for elongation factor EF-G1A which has been identified from the current analysis appeared to be an important contributor to aminoglycoside resistance. Although cell membrane permeability change-related genes appear to be the most significant contributor to aminoglycosides resistance, these variants do not offer a full explanation for the variability in aminoglycoside phenotype as shown by the predictive model. Evaluated elements of quinolone resistance can explain 77.2% and 60.9% of variance in quinolone phenotype for ciprofloxacin and levofloxacin respectively (section 3.5.5 and section 3.6.5) while all evaluated chromosomal elements for aminoglycosides can only explain 17.9% of variance in amikacin MIC and only 21.6% of variance in gentamycin MIC (section 3.7.5).

To consider another important aspect of the diagnostic landscape of antibiotic resistance, an additional complementary analysis which shows background population structure and the relation of high-risk clones to the analyzed set of molecular markers was applied in *Chapter 4*.

The chapter has described the population structure of the organism using the analyzed set of *Ps. aeruginosa* isolates and has identified some quinolones and aminoglycosides resistance-related markers to be significantly related to high-risk clones including: *nalC*E153Q, *pmr*ALeu71Arg, *gidB*E97Q, *gidB*E186A, *pst*BE89Q, *arn*DG206C, *arn*AA170T, *pmrBGly423Cys*, *nuo*GA890T, *fao*AT385A, *phoQY85F*, *lptAT55A*, *mex*RR70N, *mex*RE70R, *mex*RL130T, *mex*RG97L, *mex*RL29D, *arm*R gene, *nal*CG71E, *nal*CS46A, *nal*D, *nfx*B, *mex*S, and *mex*Z. Some other combinations of predictors assessed in *Chapter 3* has also shown significant associations with high risk clones including, amikacin *cluster 1*, amikacin *cluster 2* and ciprofloxacin *cluster 4*. These markers showing high effect size in relation to the risk clones studied, can be used as signals to risky clones or behavior. This can consequently be used to direct infection control measures.

Based on the results described above, it becomes necessary to investigate for core genome - based markers rather than using markers related to antibiotic target changes or horizontally acquired elements. Comparative genomics was performed in *Chapter 5* to gain more understanding and to mine for previously unexplored novel resistance-related elements at the whole system level. This can build up to add to the existing body of knowledge.

Using this approach has additively identified two essential functional systems linked to quinolone resistance including anthranilate metabolism and glutamate metabolism in addition to several diagnostic markers for ciprofloxacin and gentamycin. These can be used in combination with some other established markers to provide several alternative combinations of predictors showing improved performance.

The thesis conceptually concludes that choosing a diagnostic panel of molecular markers for inclusion into a rapid diagnostic platform is not an easy task. It requires evidence-based research in order to choose the most informative and consequently the most cost-effective combination. Most of the currently available rapid diagnostic platforms do not consider this aspect, but instead includes a set of horizontally acquired elements, the topic that has been discussed above in much detail.

Based on the results of the analyses performed in *Chapter 3, Chapter 4* and *Chapter 5*, it is concluded that there is no single marker, gene or gene variant, or combination to be used as a diagnostic option, but instead there are several alternative combinations with different

improved criteria of performance. These markers occur at different frequencies. These individual markers or combination of markers can be alternatively used in a diagnostic setting. The composition of these alternative predictors and combinations of predictors are shown in detail with their respective criteria of performance in Tables 6.1 - 6.4. There is a lack of standard criteria that can be used to interpret predictive values. Also, the difference in the frequency of encountering these variants in the population necessitates using this list. Table 6.3 and Table 6.4 show the best performing combinations. These include C3, C7, C8, C9 and C10. The detailed composition of these clusters is shown in Table 7.2. Cluster C3 showed 94.6% sensitivity, 100% specificity, 100% PPV, 98.4% NPV, *phi*=0.965 and Youden Index=0.95. Cluster C7 showed 76.4% sensitivity, 100% specificity, 100%PPV, 88.6%NPV, *phi*=0.823 and Youden Index=0.76. Cluster C8 showed 94.3% sensitivity, 100% specificity, 100% PPV, 90.6%NPV, *phi*= 0.924 and Youden Index=0.94. Cluster C9 showed 100% sensitivity, 97.2% specificity, 90.8% PPV, 100% NPV, *phi*=0.94 and Youden Index=0.97. Cluster C10 showed 97.2% sensitivity, 100% specificity, 100% PPV, 90.8% NPV, *phi*=0.94 and Youden Index=0.97.

To translate these findings into practical steps, a suggested diagnostic algorithm can be used to guide antibiotic selection. This is illustrated in Assessment algorithms and Decision algorithms shown in Figure 6.1, Figure 6.2 , and Figure 6.3. These workflows can be implemented using NGS-based platforms or metagenomic sequencing. The combination of markers analyzed throughout the thesis and summarized in this Chapter can be plugged into these algorithms. The suggested marker can also be incorporated into rapid diagnostic panels. The performance criteria of the combinations of markers shown in Table 6.1 and Table 6.4 can be applied using the sequence of steps illustrated in the diagrams below.

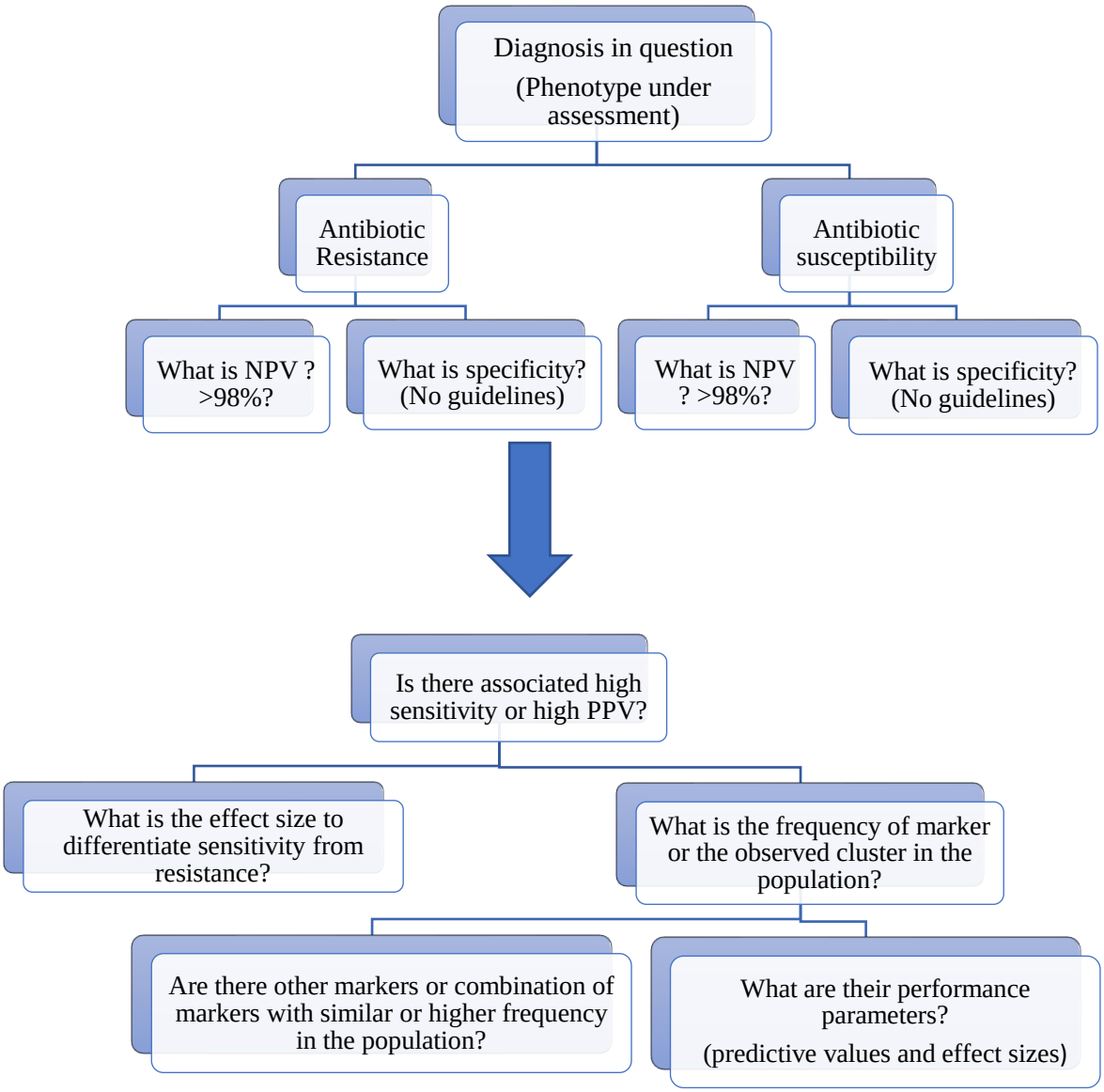


Figure 6.1. Diagram showing suggested diagnostic /Assessment algorithm

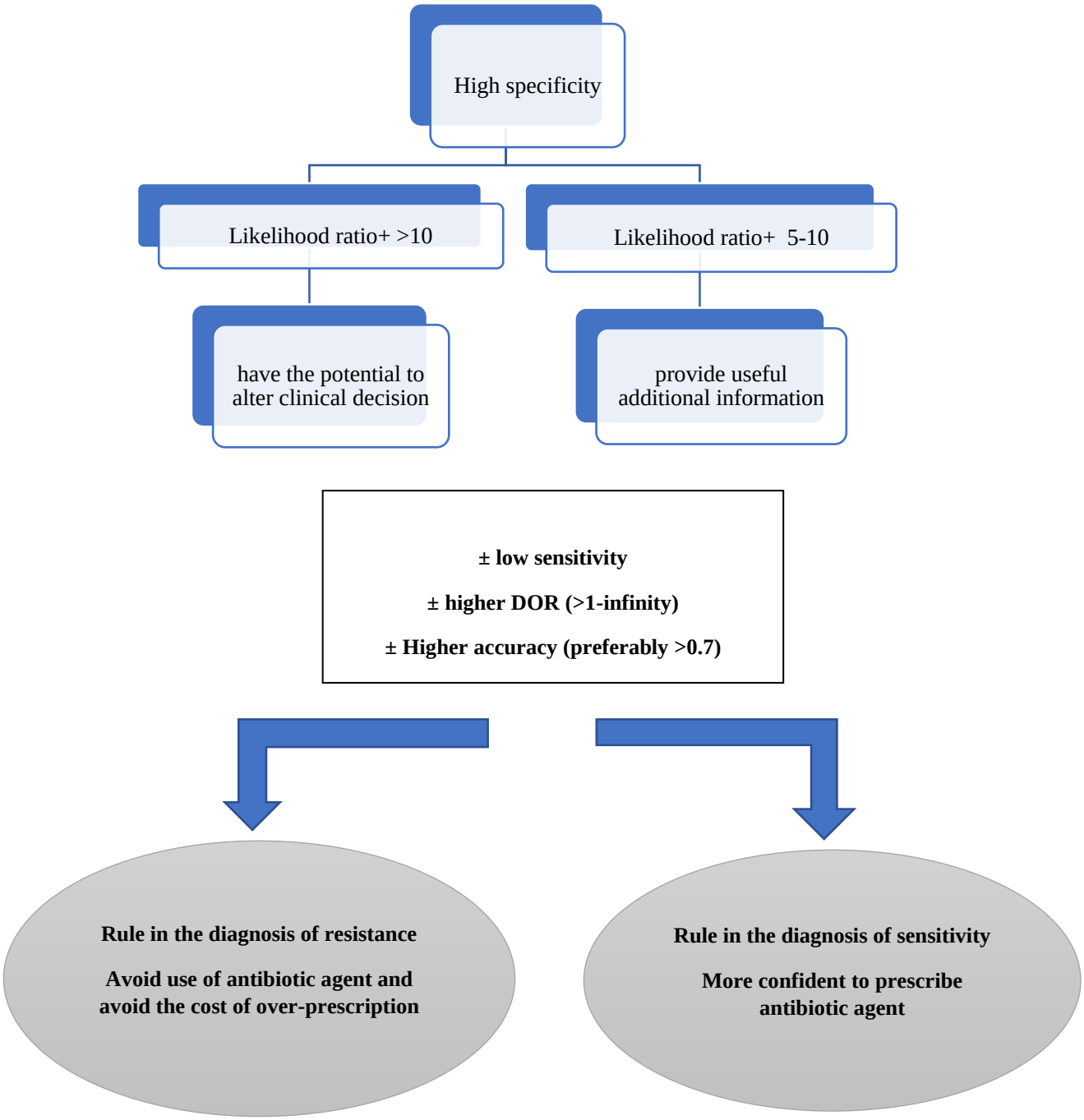


Figure 6.2. Diagram showing suggested decision algorithm 1

6.5. Example.1. Rule-in algorithm

This example illustrates the steps shown in Figure 6.1 and Figure 6.2.

By following the steps shown in Figure 6.2 to rule in the diagnosis, we would first look-up markers showing higher specificity ± higher NPV. We would then look up if higher specificity is also combined with a high LR+ for the same marker. In case higher specificity is also combined with a LR+ value which is greater than 10, then the marker has the potential to alter clinical decision. If the higher specificity is combined with a LR+ value which ranges between 5 and 10, then the marker can provide useful additional information. When markers fulfil one of the above two criteria and additionally show high DOR with value > (1-infinity) ± higher accuracy (>0.7), this would then give more confidence to rule in the diagnosis.

In case the criteria of high specificity, high LR+, high accuracy, and high DOR are achieved for resistance behavior, the marker can then be used to rule in resistance. This helps to avoid the use of antibiotic agent and consequently to avoid the cost of over-prescription.

To use these markers, we would start by looking up markers with higher frequencies. These have a higher chance of being encountered in a random sample. In case the marker is not identified, we would then look up markers with lower frequencies.

Examples for algorithm 1, Rule in Quinolone resistance:

- *gyr*AT83I: Specificity 99.5%, PPV 96.6%, LR 54.08, DOR 65.67, Accuracy 0.73
- *nfx*BA124T: Specificity 99%, PPV 80%, LR 10.07, DOR 7.61, Accuracy 0.66
- *nal*CE153Q: Specificity 97.5-98.4%, PPV 68.8-81%, LR 70.5, DOR 4-5.74, Accuracy 0.72
- *nal*CThr50Pro: Specificity 99.5%, PPV 50%, LR 60, DOR 2.35, Accuracy 0.69
- *amp*RD135N: Specificity 99.5%, PPV 80%, LR 12.6, DOR 7.65, Accuracy 0.66
- *par*CE91k: Specificity 99.7-100%, PPV 80-100%, LR 2-3.36, DOR 3.35, Accuracy 0.46-0.64
- *par*EV460G: Specificity 99.7-100%, PPV 75-100%, LR 2-2.68, DOR 2.51, Accuracy 0.46-0.64

Examples for algorithm 1, Rule in Aminoglycosides resistance:

- *arn*AA170T: Specificity 98.3-99%, PPV 62.5-80%, LR 8.7-19.7, DOR 6.7-7.84, Accuracy 0.67-0.79
- *arn*DG206C: Specificity 98.2-99%, PPV 58.3-77.8%, LR 7-16, DOR 5.57-6.79, Accuracy 0.66-0.79
- *fus*A1D588G: Specificity 100%, PPV 100%, LR 6.3-15.5, Accuracy 0.66-0.79
- *nal*DSer32Asn: Specificity 99.4-100%, PPV 66.7-100%, LR 6.3-8.6, DOR 7.6, Accuracy 0.66-0.79

6.6. Example.2. Rule-out algorithm

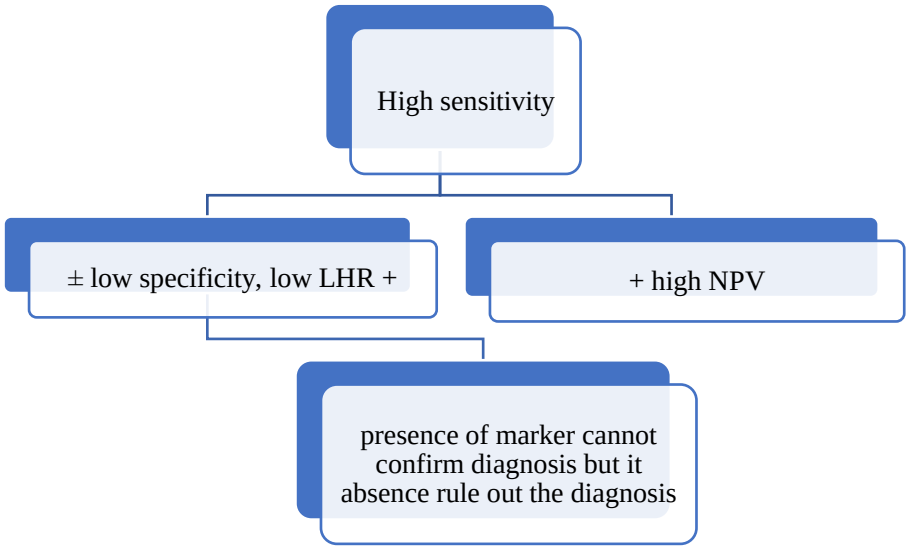


Figure 6.3. Diagram showing suggested decision algorithm 2

Using the algorithm shown in Figure 6.3, we would start by looking up markers with high sensitivity. For those showing high sensitivity, we can then look up if the same marker shows high NPV. This performance can be additionally combined with low specificity ± low LR. Also, high LR and high accuracy values would offer useful additional diagnostic information.

Examples for algorithm 2, Rule out Quinolone susceptibility:

Presence does not confirm susceptibility, but absence predict resistance

- *nfx*B: Sensitivity 99.7%, NPV 83.3-97.3%, LR 5.8-34.5, Accuracy 0.5-0.67
- *mex*S: Sensitivity 99.7%, NPV 91.7-97.1%, LR 17-31.67, Accuracy 0.5-0.67
- *amp*R: Sensitivity 100%, NPV 100%, LR 8.23, Accuracy 0.65
- *nal*C: Sensitivity 99.4-100%, NPV 94.3-100%, LR 15.79-59.7, Accuracy 0.5-0.72
- *nal*D: Sensitivity 100%, NPV 100%, LR 38.6, Accuracy 0.5-0.69
- *mex*R: Sensitivity 98.3-99%, NPV 62.5-80%, LR 8.7-19.7, DOR 6.7-7.84, Accuracy 0.67
- *arm*R: Sensitivity 82.4-91.1%, NPV 58.4-92.2%, LR 1.98-85.8, Accuracy 0.58-0.71
- *mex*Z: Sensitivity 97.3-99.1-%, NPV 50.92.5%, LR 11.82, Accuracy 0.5
- *mex*R R79N: Sensitivity 50.8-98.5%, NPV 52.6-75%, LR 10, Accuracy 0.56-0.66
- *mex*R E70R: Sensitivity 52.4-99.5%, NPV 53.4-83.3%, LR 7.9, Accuracy 0.57-0.65
- *mex*R L130T: Sensitivity 53-100%, NPV 54-100%, LR:16.7, Accuracy 0.58-0.66
- *mex*R G97L: Sensitivity 53-100%, NPV 51.7-100%, LR 8.23, Accuracy 0.56-0.65
- *mex*R L29D: Sensitivity 53-100%, NPV 52.2-100%, LR 12.43, Accuracy 0.56-0.65

I here also summarize the final results of the best performing molecular predictors analyzed throughout the thesis. Individual predictors and clusters of predictors can be used as "flexible format alternative options" for use as diagnostic markers as illustrated in the examples shown above. These may be incorporated into a rapid diagnostic panel or alternatively identified using NGS platforms or other metagenomic platforms. The thesis provides proof of principle approach for some potential diagnostic markers for 2 agents from the aminoglycoside group and 2 agents from the quinolone group of antibiotics in *Ps. aeruginosa*. This can be re-implemented with other clinically important organisms and other antibiotic classes and can also be adapted and modified based on new research findings.

Summary of the performance of best analyzed alternative predictor markers and combination of predictors to be applied using the suggested assessment and diagnostic algorithms shown above are shown in Tables 6.1-6.4.

Individual predictors

Table 6.1. Measures of performance for best predictor quinolone and aminoglycoside molecular markers

Molecular Markers	Antibiotic agent assessed	Sensitivity	Specificity	PPV	NPV	LR+	LR	DOR	Youden Index	Accuracy	Frequency	Phi coefficient	Significant association with phenotype (P-value)	High-risk ST association *
*arm*R	ciprofloxacin	**91.1**	59.6	56	**92.2**	**2.25**	**85.8**	**15.05**	**0.507**	**0.71**	**58.7**	0.494	<0.0005	*
*nal*CS46A		**71.4**	81.8	69	**97**	**3.92**	**217**	**72**	**0.532**	**0.86**	**37.4**	0.765	<0.0005	*****
*nal*D gene		**100**	16.1	67.8	**100**	1.19	38.6	ND	0.161	**0.7**	**94.2**	0.33	<0.0005	*
*nal*C gene		**100**	24.1	70	**100**	1.32	**59.7**	ND	0.241	**0.73**	**91.3**	0.411	<0.0005	
*mex*S gene		**99.5**	9.8	66.1	**91.7**	1.1	17	21.46	0.093	0.67	**96.1**	0.232	<0.0005	
*nfx*B gene		**99.5**	4.5	64.8	**83.3**	1.04	5.8	9.2	0.04	0.65	**98.1**	0.138	0.015	
*amp*R gene		**100**	3.6	64.7	**100**	1.03	8.23	ND	0.036	0.65	**98.7**	0.152	0.007	
*gyr*A T83I		25	**99.5**	96.6	70.1	**50**	**54.08**	**65.67**	0.245	**0.73**	9.4	0.404	<0.0005	*
*nfx*B gene	Levofloxacin	99.7	9.4	47.8	97.3	1.1	34.48	**32.97**	0.091	0.504	**94.7**	0.203	<0.0005	*
*mex*R gene		100	1.8	45.9	100	1.02	ND	ND	0.018	0.46	**99**	0.092	0.015	
*mex*S gene		99.7	8.7	47.6	97.1	1.1	31.67	**29.96**	0.084	0.5	**95.1**	0.0193	<0.0005	*
*nal*C gene		99.4	8.7	47.5	94.3	1.1	15.79	**14.94**	0.081	0.49	**95**	0.183	<0.0005	*
*nal*D gene		100	12.3	48.7	100	1.14	ND	ND	0.123	0.52	**93.3**	0.245	<0.0005	
*amp*R gene		100	8.7	47.7	100	1.1	ND	ND	0.087	0.5	**95.3**	0.203	<0.0005	
*mex*Z		99.1	9.7	47.7	92.5	1.1	11.8	**11.26**	0.088	0.5	**94.3**	0.188	<0.0005	*
*gyr*AT83I		48.3	99.4	98.9	61.5	**80.5**	154.8	**147.1**	0.477	0.71	26.6	0.537	<0.0005	*
*nal*C	Gentamycin	**100**	1.5	35.7	**100**	1.02	2.632	ND	0.015	0.36	99	0.074	0.198	*
*mex*Z		**100**	2.1	35.8	**100**	1.02	**3.517**	ND	0.021	0.37	98.7	0.086	0.137	*
*amp*R		**100**	1.5	35.7	**100**	1.02	2.632	ND	0.015	0.36	99	0.074	0.198	*
*nal*Dser32Asn		2.8	**100**	**100**	65.3	ND	**6.3**	ND	0.03	**0.66**	**6.3**	**0.136**	0.019	
*nal*Dl153Q		2.8	**99.5**	75	65.2	**5.6**	2.65	**5.62**	0.03	**0.66**	2.65	0.096	0.095	
*fus*A1D588G		2.8	**100**	**100**	65.3	ND	**6.3**	ND	0.03	**0.66**	**6.3**	**0.136**	0.019	
*arn*AA170T		7.5	**99**	80	66.2	**7.5**	**8.68**	**7.84**	0.07	**0.67**	3.3	**0.174**	0.003	

*arn*DG206C		6.6	**99**	77.8	66	**6.6**	**6.99**	**6.79**	0.06	**0.66**	3	**0.156**	0.007	
*arn*AA170T	**Amikacin**	10.1	**98.3**	62.5	80.1	**5.94**	**19.69**	**6.69**	0.084	**0.79**	3.5	**0.19**	0	*
*pst*BE89Q		9.5	**98.9**	70	80	**8.64**	**22.45**	**9.35**	0.084	**0.79**	2.9	**0.204**	0	*
*arn*DG206C		9.5	**98.2**	58.3	79.9	**5.28**	**16.08**	**5.57**	0.077	**0.79**	3.5	**0.171**	0	*
*fus*A1D588G		3.4	**100**	100	79.2	ND	**15.54**	ND	0.034	**0.79**	0.7	**0.164**	0	
*mex*RR79S		2	**100**	**100**	21.8	ND	**5.351**	ND	0.02	0.23	1.6	-0.07	0.081	
*amg*SE108Q		2	**100**	**100**	21.8	ND	**5.351**	ND	0.02	0.23	1.6	-0.07	0.081	
*nuo*GA574T		1.4	93.7	5.6	77.7	0.22	**7.398**	0.21	0.049	**0.74**	5.2	-0.09	0.017	

Clusters of predictors

Table 6.2. Components of predictor clusters showing the distribution of best performing predictor combination of molecular markers

Predictor cluster	**C 1**	**C 2**
Markers composition		
*nal*CS209R	present96%	absent3-100%
*pmr*BGly423Cys	absent100%	present0-70%
*nal*CG71E	present100%	absent 3-45%
M7	present100%	absent 74.3-100%
M21	present51.9%	absent 82.9-97.4%
*nal*CA186T	absent 100%	present 3-33%
M25	absent 100%	present 3-31%
M1	absent 100%	present 6-33%
*fao*AT385A	absent 100%	present 0-23%
*lpt*AR62S	absent 100%	present 0-17%
*lpt*At55A	absent100%	present 3-23%
M11	absent 100%	present 0-13%
*nuo*GA890T	absent 100%	present 0-12%
*gid*BE186A	absent 100%	present 0-8%
*gid*BE97Q	absent 100%	present 0-8%
M22	absent100%	present 5-13%
*fus*AD588G	present 4%	absent 100%
*nal*CE153Q	absent100%	present 2-8%
*amp*RA51T	absent100%	present 2.5%

*nal*Dser32Asn	absent100%	present 2.5%
M35	absent100%	present 3-8%
*arn*AA170T	absent100%	present 1.5%
*arn*DG206C	absent 100%	present 1.5%
*mex*Z	present 100%	absent 1.5-3%
Markers composition	**C 3 and C 5**	**C 4 and C 6**
M11	absent 100%	present 100%
M12	absent 100%	present 100%
M13	absent 91.4%	present 100%
M20	present 37%	present 8%
M43	present 25%	present 42%
*nal*CA186T	present 12%	present 25%
*amp*RD135N	present 3%	absent 100%
Markers composition	**C 7**	**C 8**
*gid*B E186A	Marker absent 28%	Marker absent 69%
*gid*BE97Q	Marker absent 29%	Marker absent 63%
*arn*AA170T	Marker absent 28%	Marker absent 62%
*arn*DG206C	Marker absent 28%	Marker absent 62%
*nal*CE153Q	Marker absent 29%	Marker absent 65%
*pst*BE89Q	Marker absent 29%	Marker absent 62%
Markers composition	**C 9**	**C 10**
*arn*AA170T	Mutation absent 85.3%	Mutation absent 100%
*arn*DG206C	Mutation absent 85.3%	Mutation absent 100%
*pst*BE89Q	Mutation absent 87.7%	Mutation absent 100%
*nal*DSer32Asn	Mutation absent 94.5%	Mutation absent 100%
*lpt*AR62S	Marker absent 98.8%	Marker absent 95.1%
*nuo*GA890T	Marker absent 99.4%	Marker absent 96.6%
*fao*AT385A	Marker absent 99.4%	Marker absent 97.2%
Markers composition	**C 11**	**C 12**
*nal*CE153Q	Mutation absent 100%	Mutation absent 82.1%
*nal*CThr50pro	Mutation absent 100%	Mutation absent 97.4%
*arm*R	Gene present 100%	Gene present 97.4%
*nal*D	Gene present 100%	Gene present 97.4%
*gyr*A T83I	Mutation absent 100%	Mutation absent 56.4%
*nfx*B A124T	Mutation absent 100%	Mutation absent 97.4%
Markers composition	**C 13**	**C 14**

*nal*CS209R		present 50%
M20	present 22%	present25%
M23	present 22%	present 12%
M2	present 77.8	absent 25%
*nal*CA186T		present 12%

Table 6.3. Best performing quinolone and aminoglycoside predictor combinations

Predictor cluster	Phenotype	Average MIC	Cluster frequency	Phi coefficient	P-value
C1	Aminoglycoside resistant	AK=8.59 CN=13.96	16.7%	0.479	<0.0005
C2	Aminoglycoside susceptible	AK=2.55 CN=4	83.3%	0.155	0.13
C3	Quinolone resistant	Levo=6.57	21.6%	0.965	<0.0005
C4	Quinolone susceptible	Levo=0.83	7.4%	0.042	0.597
C5	Quinolone resistant	Cipro=1.05	21.6%	0.34	<0.0005
C6	Quinolone susceptible	Cipro=0.1	7.4%	0.05	0.521
C7	Aminoglycoside resistant	CN=12.94	27%	0.823	<0.0005
C8	Aminoglycoside susceptible	CN=2.9	61%	0.924	<0.0005
C9	Aminoglycoside resistant	AK=54.47	23.6%	0.94	<0.0005
C10	Aminoglycoside susceptible	AK=3.69	76.4%	0.94	<0.0005
C11	Quinolone susceptible	Cipro=0.22	19.1	0.366	<0.0005
C12	Quinolone resistant	Cipro=4.82	12.8	0.509	<0.0005
C13	Quinolone susceptible	Cipro=2.78	5.6	0.424	<0.0005
C14	Quinolone resistant	Cipro=3.12	4.9	0.454	<0.0005

Table 6.4. Measures of diagnostic accuracy for best performing quinolone and aminoglycoside predictor combinations

Performance Parameters	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**
Sensitivity	**100**	28.8	**94.6**	8	**100**	7.6	76.4	**94.3**	**100**	**97.2**	29.9	35.5	60	60
Specificity	86.5	**100**	**100**	**94.6**	80.9	**100**	**100**	**100**	**97.2**	**100**	**100**	**100**	**96.2**	**96.8**
PPV	22.2	**100**	**100**	83.3	14.3	**100**	**100**	**100**	**90.8**	**100**	**100**	**100**	33.3	37.5
NPV	**100**	51.6	**98.4**	23.3	**100**	3.3	88.6	90.6	**100**	**90.8**	44.7	73.2	**98.7**	**98.7**
LR+	**7.4**	-	-	1.5	**5.2**	-	-	-	**35.7**	-	-	-	**15.8**	**18.8**
LR	-	-	-	1.5	-	-	-	-	-	-	-	-	**38**	**45.4**
DOR	-	-	-	1.5	-	-	-	-	-	-	-	-	**37.8**	**45.4**
Youden Index	0.87	0.29	**0.95**	0.03	0.81	0.08	0.76	**0.94**	**0.97**	**0.97**	0.3	0.36	0.56	0.57
Accuracy	**0.87**	**0.87**	**0.99**	0.28	**0.81**	0.1	**0.92**	**0.96**	**0.98**	**0.98**	0.55	**0.77**	**0.95**	**0.96**

High-risk ST association *	UN	UN	UN	UN	UN	UN	UN	UN	***0.006**	***0.006**	UN	***0.012**	UN	UN

UN: unknown -not evaluated, *: significant association (p-value<0.05)

6.7. General summary

As discussed in *Chapter 2,* automated tools and bioinformatics platforms that predict for antibiotic resistance may suffer the theoretic limitation of restricted spectrum of genes/resistance markers identified. This may originate from the fact that our understanding about key resistance conferring genes and mutations is still growing. Especially for *Ps. aeruginosa*, the special nature of the species makes AMR prediction in *Ps. aeruginosa* specifically challenging. This highlights the complexity of associating antibiotic resistance phenotype to genotype in the organism.

When the tools were used to assess genotype-phenotype correlations in *Ps. aeruginosa* as studied in *Chapter 2*, the results as applied to 87 selected highly diverse isolates showed that there is tendency to overpredict resistance with false positive rate of 2-25% for ciprofloxacin susceptibility, 24-45% for levofloxacin susceptibility, and 5-21% for amikacin susceptibility for the MICRA platform. This is a situation that needs to be minimized to avoid the cost of antibiotic over-prescription.

Using CARD and ResFinder, different sets of predictors were identified in the tested collection of isolates. Those were tested for their correlational pattern and predictive capacity to identify antibiotic resistance phenotype. The ranges of predictive values for these sets of predictors are summarized in Figure 6.4 to Figure 6.7.

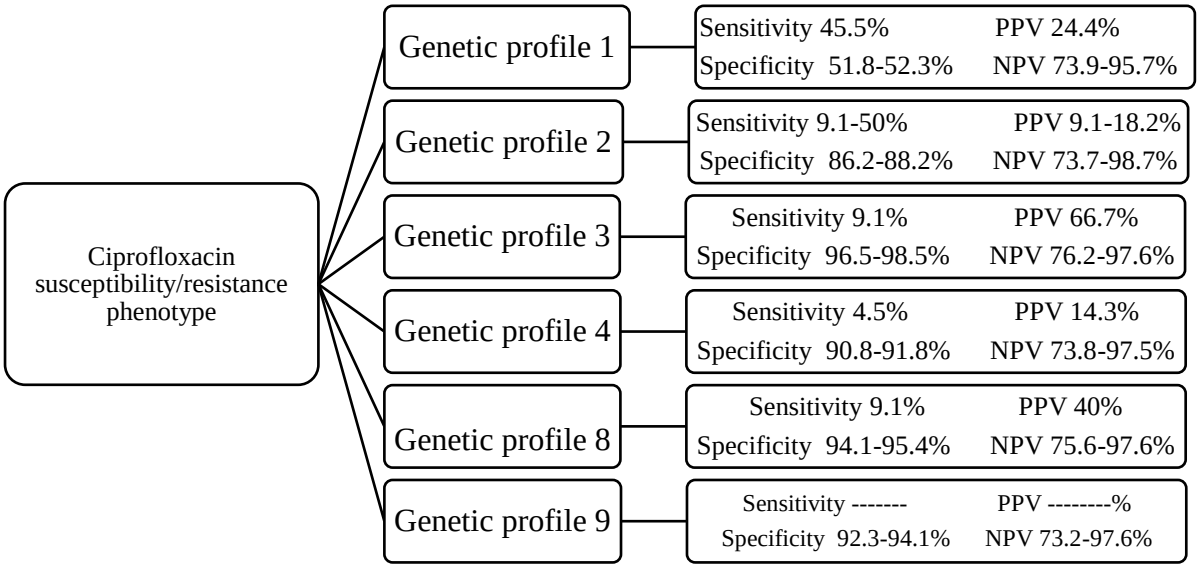


Figure 6.4. Predictive capacity of identified predictor sets using CARD/ResFinder for differentiating ciprofloxacin susceptibility/resistance phenotype in the 87 selected diverse Ps. aeruginosa isolates

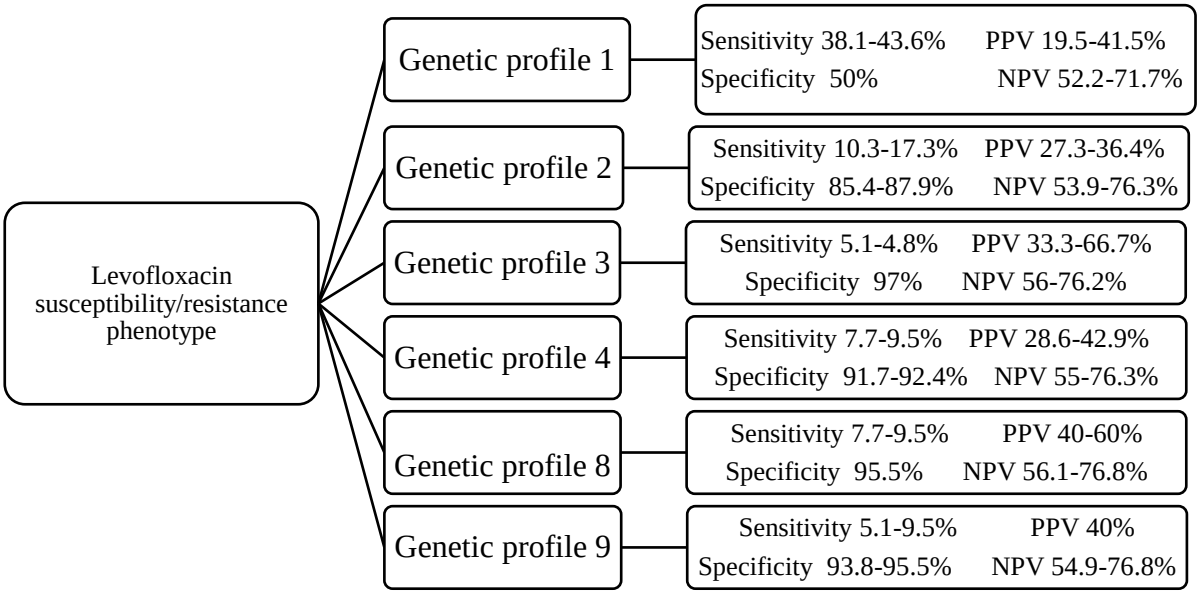


Figure 6.5. Predictive capacity of identified predictor sets using CARD/ResFinder for differentiating levofloxacin susceptibility/resistance phenotype in the 87 selected diverse Ps. aeruginosa isolates

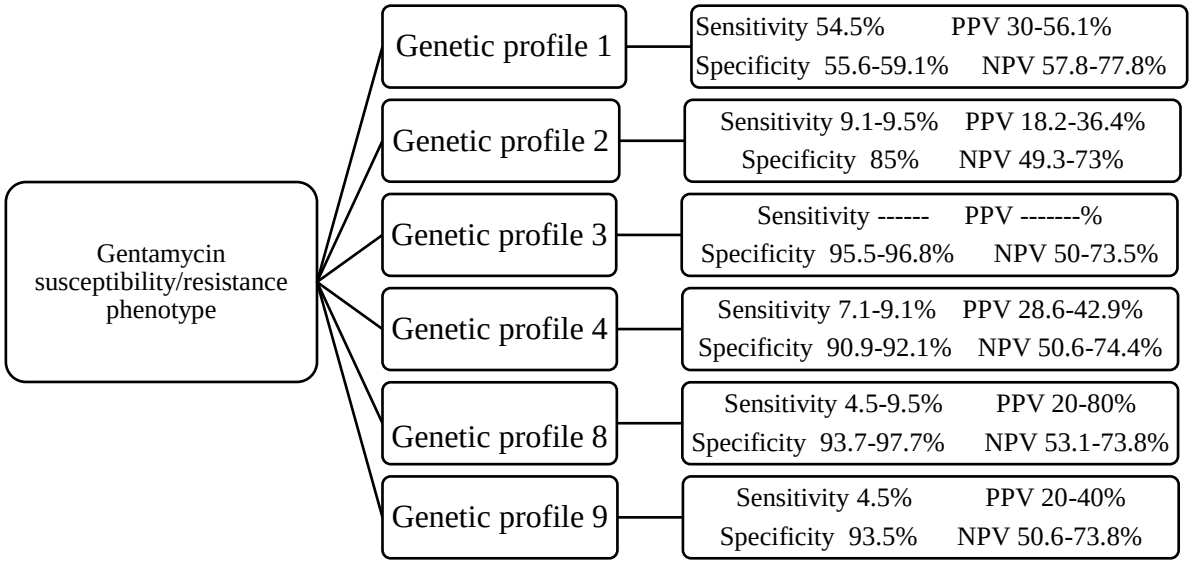


Figure 6.6. Predictive capacity of identified predictor sets using CARD/ResFinder for differentiating gentamycin susceptibility/resistance phenotype in the 87 selected diverse Ps. aeruginosa isolates

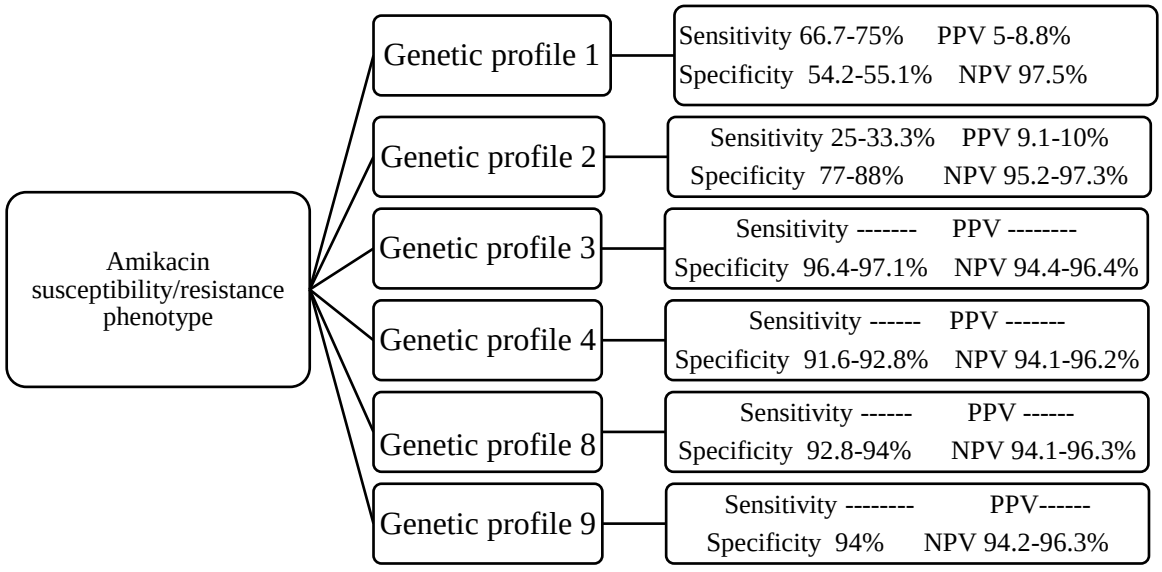


Figure 6.7. Predictive capacity of identified predictor sets using CARD/ResFinder for differentiating amikacin susceptibility/resistance phenotype in the 87 selected diverse Ps. aeruginosa isolates

That is the reason why a larger more genetically and phenotypically inclusive collection of isolates have been tested to find elements/ combination of elements that can improve antibiotic resistance prediction. Based on observing a group of markers with different tendencies to cluster towards higher/lower susceptibility phenotypes, new combinations of markers appeared to offer better prediction/differentiation for the studied phenotypes. The predictive capacity of the sets of markers shown in *Chapter 3* for differentiating antibiotic susceptibility from resistance are summarized in Figure 6.8 to Figure 6.11.

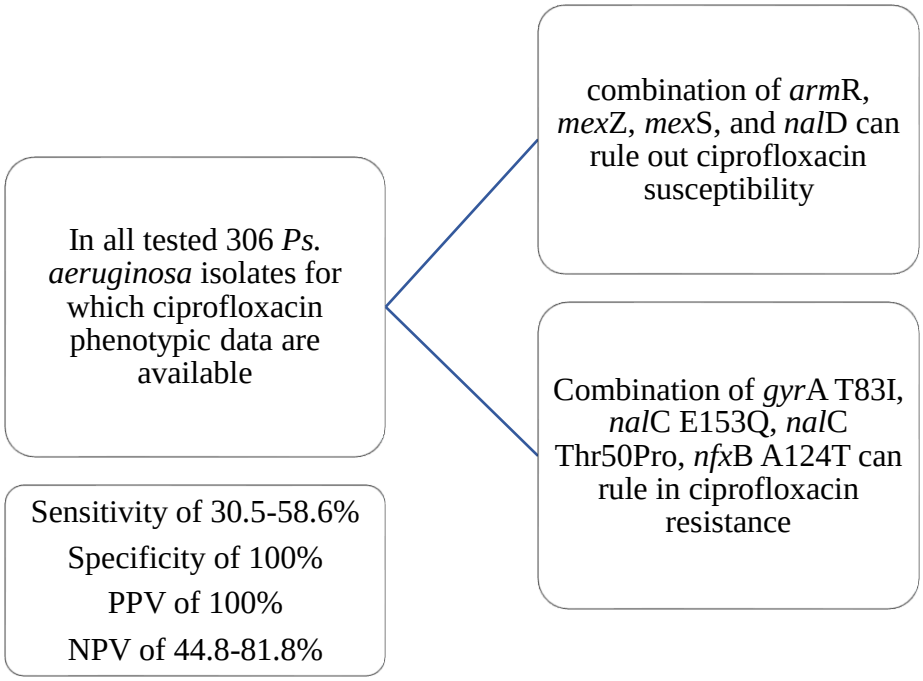


Figure 6.8. Predictive capacity of combinations in differentiating ciprofloxacin susceptibility/resistance phenotype in all 306 Ps. aeruginosa isolates for which ciprofloxacin susceptibility data are available

The diagram above is based on the results as extracted in *Chapter 3* from Table 3.4 and Figure 3.1.

Cluster 4 composed of ***39 ciprofloxacin resistant*** *Ps. aeruginosa* isolates with average MIC of 4.82 showed presence of *gyr*A T83I, *nal*C E153Q, *nal*C Thr50Pro, *nfx*B A124T in 43.6%, 17.9%, 2.6% and 2.6% of isolates respectively and absence of *arm*R, *nal*D in 2.6% of isolates.

Cluster 3 composed of ***26 ciprofloxacin resistant*** *Ps. aeruginosa* isolates with average MIC of 4.62 showed presence of *gyr*A T83I and *nfx*B A124T in 42.3% and 7.7 % of isolates respectively

absence of *nal*D, *mex*S, *arm*R and *mex*Z in 42.3 %, 38.5 %, 7.7 % and 7.7% of isolates respectively.

Cluster 2 composed of ***60 ciprofloxacin susceptible*** *Ps. aeruginosa* isolates with average MIC of 0.22 showed absence of *gyr*A T83I, *nal*C E153Q, *nal*C Thr50Pro, *nfx*B A124T in 100 % of isolates and presence of *arm*R, *nal*D, *mex*S and *mex*Z in 100% of isolates.

Based on the observed differential distribution among clusters, in all tested 306 *Ps. aeruginosa* isolates for which ciprofloxacin susceptibility phenotype is available, the combination of *arm*R + *mex*Z + *mex*S + *nal*D + *gyr*A T83I + *nal*C E153Q + *nal*C Thr50Pro + *nfx*B A124T can differentiate ciprofloxacin resistance from ciprofloxacin susceptibility with overall ***sensitivity of 30.5 % - 58.6%***, ***specificity of 100%***, ***PPV of 100%*** and ***NPV of 44.8%-81.1%***.

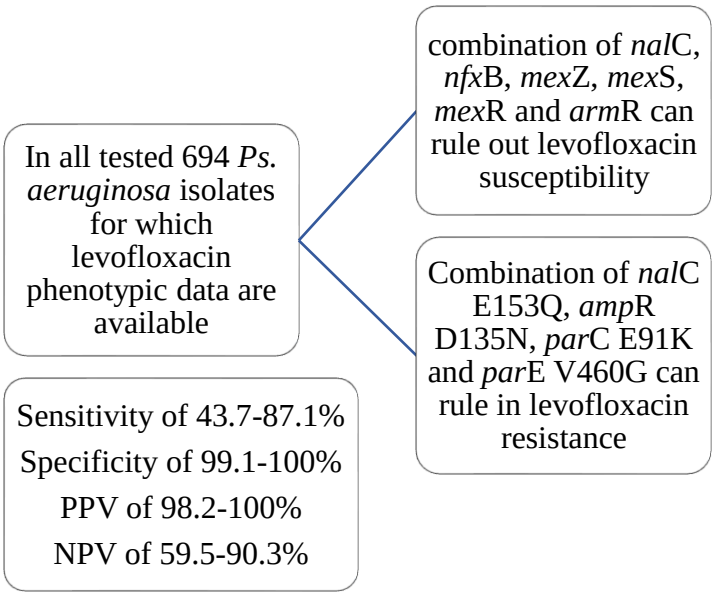


Figure 6.9. Predictive capacity of combinations in differentiating levofloxacin susceptibility/resistance phenotype in all 694 Ps. aeruginosa isolates for which levofloxacin susceptibility data are available

The diagram above is based on the results as extracted in *Chapter 3* from Table 3.8 and Figure 3.3,

Cluster 1 composed of ***164 levofloxacin susceptible*** *Ps. aeruginosa* isolates with average MIC of 0.56 showed presence of *nal*C, *nfx*B, *mex*Z, *mex*S, *mex*R and *arm*R in 99.4%, 100%, 100%, 100%, 100%, 43.9% and absence of *nal*C E153Q, *amp*R D135N, *par*C E91K and *par*E V460G in 97.6%, 100%, 100% and 100% isolates respectively.

Cluster 5 composed of ***150 levofloxacin susceptible*** *Ps. aeruginosa* isolates with average MIC of 0.54 showed presence of *nal*C, *nfx*B, *mex*Z, *mex*S, *mex*R and *arm*R in 100%, 100%, 99.3%, 99.3%, 100%, 99.3% and absence of *nal*C E153Q, *amp*R D135N, *par*C E91K and *par*E V460G in 100%, 99.3%, 100% and 99.3% isolates respectively.

Cluster 3 composed of ***31 levofloxacin resistant*** *Ps. aeruginosa* isolates with average MIC of 4.13 showed absence of *nal*C, *nfx*B, *mex*Z, *mex*S and *arm*R in 100%, 96.8%, 100%, 64.5%, 96.8% and presence of *amp*R D135N, *par*C E91K in 6.5% and 3.2% of isolates respectively.

Cluster 2 composed of ***134 levofloxacin resistant*** *Ps. aeruginosa* isolates with average MIC of 11.05 showed absence of *mex*Z, *mex*S and *arm*R in 5.8%, 8.7%, 27.5% and presence of *nal*C E153Q, *amp*R D135N, *par*C E91K and *par*E V460G in 14.5%, 5.8%, 2.9% and 2.2% of isolates respectively.

Based on the observed differential distribution among clusters, in all tested 694 *Ps. aeruginosa* isolates for which levofloxacin susceptibility phenotype is available, combination of *nal*C + *nfx*B + *mex*Z + *mex*S + *mex*R + *arm*R + *nal*C E153Q + *amp*R D135N + *par*C E91K + *par*E V460G can differentiate levofloxacin resistance from levofloxacin susceptibility with overall ***sensitivity of 43.7- 87.1%***, ***specificity of 99.1-100%***, ***PPV of 98.2-100%*** and ***NPV of 59.5-90.3%.***

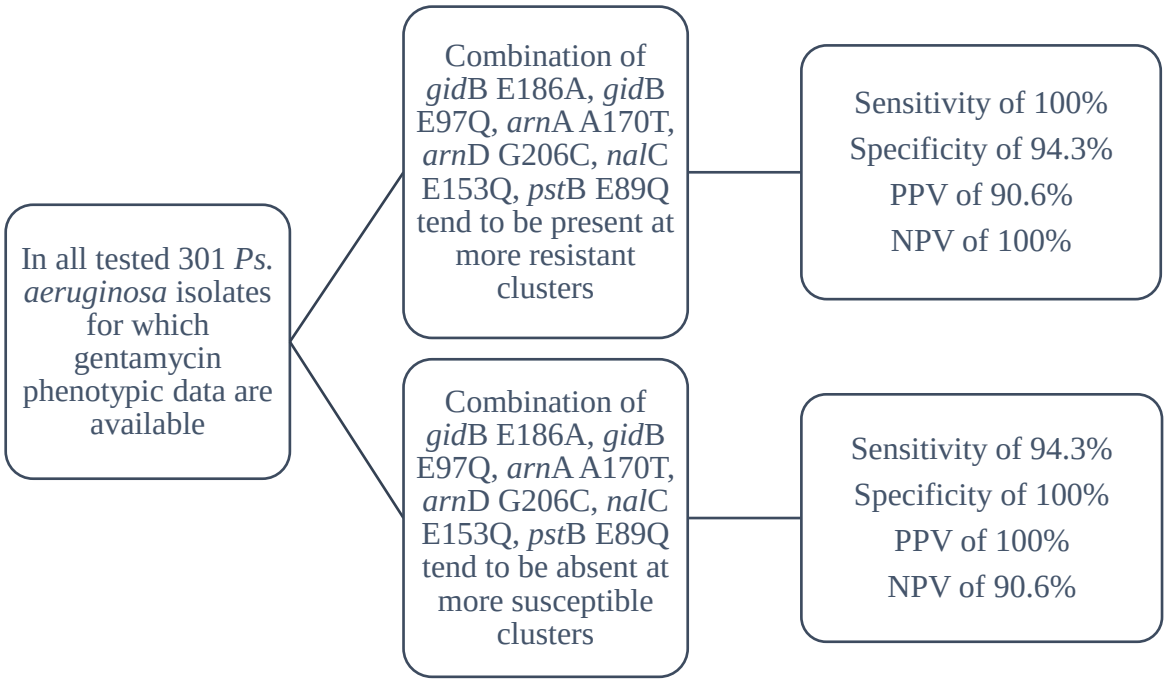


Figure 6.10. Predictive capacity of combinations in differentiating gentamycin susceptibility/resistance phenotype in all 301 Ps. aeruginosa isolates for which gentamycin susceptibility data are available

The diagram above is based on the results as extracted in *Chapter 3* from Table 3.13 and Figure 3.5,

Cluster 1 collectively composed of ***81 gentamycin resistant*** *Ps. aeruginosa* isolates with average MIC of 12.94 showed presence of *gid*B E186A, *gid*B E97Q, *arn*A A170T, *arn*D G206C, *nal*C E153Q, *pst*B E89Q in 72%, 71%, 72%, 72%, 71%, 71% of isolates respectively.

Cluster 2, 3, 4 collectively composed of ***25 gentamycin resistant*** *Ps. aeruginosa* isolates with average MIC ranging from 10.91-26.89 showed presence of *gid*B E186A, *gid*B E97Q, *arn*A A170T, *arn*D G206C, *nal*C E153Q, *pst*B E89Q in (95-100%), (95-100%), (90.9-97%), (81.8-97%), (30-87.5%), (72.7-96%) of isolates respectively.

Cluster 5 collectively composed of ***183 gentamycin susceptible*** *Ps. aeruginosa* isolates with average MIC of 2.9 showed absence of *gid*B E186A, *gid*B E97Q, *arn*A A170T, *arn*D G206C, *nal*C E153Q, *pst*B E89Q in 69%, 63%, 62%, 62%, 65%, 62% of isolates respectively.

Based on the observed differential distribution among clusters, in all tested 301 *Ps. aeruginosa* isolates for which gentamycin susceptibility phenotype is available, combination of *gid*B E186A,

*gid*B E97Q, *arn*A A170T, *arn*D G206C, *nal*C E153Q, *pst*B E89Q can differentiate gentamycin resistance from gentamycin susceptibility with overall ***sensitivity of 94.3-100%***, ***specificity 94.3-100%***, ***PPV of 90.6-100%*** and ***NPV of 90.6-100%.***

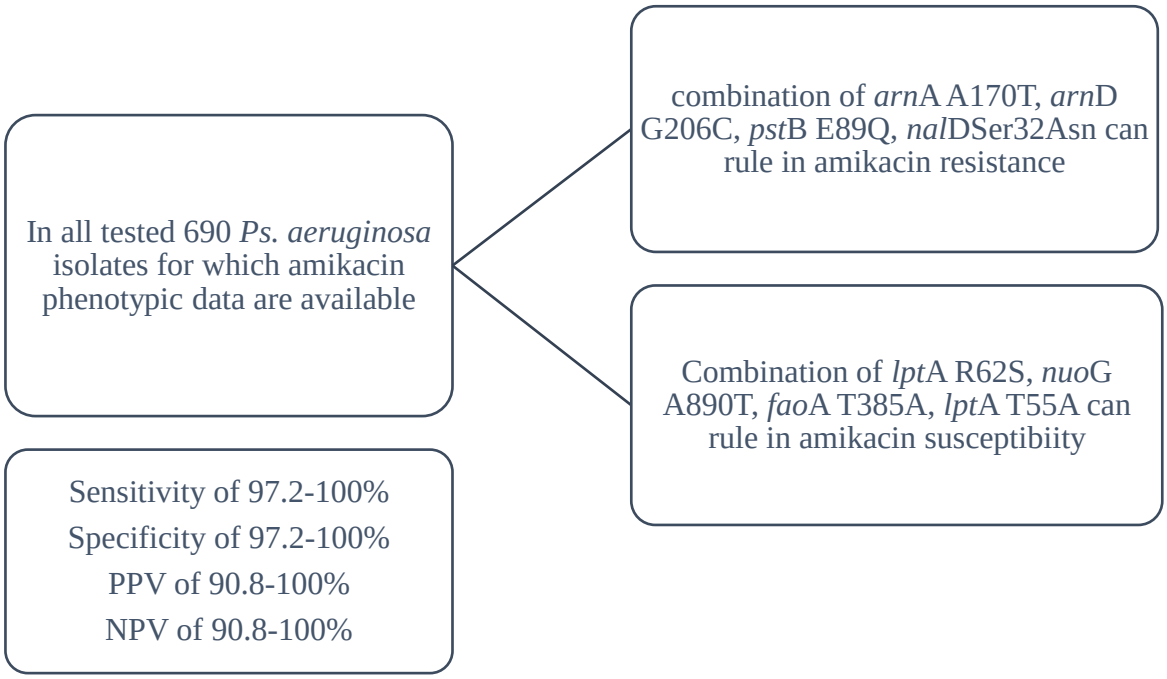


Figure 6.11. Predictive capacity of combinations in differentiating amikacin susceptibility/resistance phenotype in all 690 Ps. aeruginosa isolates for which amikacin susceptibility data are available

The diagram above is based on the results as extracted in *Chapter 3* from Table 3.15 and Figure 3.7.

Cluster 1 composed of ***528 amikacin susceptible*** *Ps. aeruginosa* isolates with average MIC of 3.69 showed absence of *arn*A A170T, *arn*D G206C, *pst*B E89Q, *nal*D Ser32Asn in 100% of isolates and presence of *lpt*A R62S, *nuo*G A890T, *fao*A T385A, *lpt*A T55A in 4.9%, 3.4%, 2.8%, 8.5% of isolates respectively.

Cluster 2 composed of ***148 amikacin resistant*** *Ps. aeruginosa* isolates with average MIC of 54.47 showed presence of *arn*A A170T, *arn*D G206C, *pst*B E89Q, *nal*D Ser32Asn in 14.7%, 14.7%, 12.3%, 5.5% of isolates respectively and absence of *lpt*A R62S, *nuo*G A890T, *fao*A T385A, *lpt*A T55A in 98.8%, 99.4%, 99.4%, 95.1% of isolates respectively.

Based on the observed differential distribution among clusters, in all tested 690 *Ps. aeruginosa* isolates for which amikacin susceptibility phenotype is available, combination of *arn*A A170T, *arn*D G206C, *pst*B E89Q, *nal*D Ser32Asn, *lpt*A R62S, *nuo*G A890T, *fao*A T385A, *lpt*A T55A can differentiate amikacin resistance from amikacin susceptibility with overall ***sensitivity of 97.2-100%***, ***specificity 97.2-100%***, ***PPV of 90.8-100%*** and ***NPV of 90.8-100%***

Based on summary diagrams shown, the set of predictors identified through clustering as summarized in Fig 6.8, Fig 6.9, Fig 6.10, and Fig 6.11 appears to offer better differentiation/ prediction towards quinolone and aminoglycoside susceptibility when compared to genomic profiles/predictor sets identified using CARD and ResFinder. Genomic profiles are shown in Table 2.1 and Table 2.5. These are also summarized in Figure 6.4, Figure 6.5, Figure 6.6, Figure 6.7.

The difference in predictive capabilities for different sets assessed can be summarized as follows:

- From the results shown in Fig 6.4 and Fig 6.8, the set of markers identified for ciprofloxacin appears to offer better prediction/differentiation for ciprofloxacin susceptibility phenotype with improved sensitivity of 8.6-26%, improved specificity of 48-98%, improved PPV of 33-90%.

- From the results shown in Fig 6.5 and Fig 6.9, the set of markers identified for levofloxacin appears to offer better prediction/differentiation for levofloxacin susceptibility phenotype with improved sensitivity of 43-82%, improved specificity of 12-50%, improved PPV of 31-80% and improved NPV of 13-38%.

- From the results shown in Fig 6.6 and Fig 6.10, the set of markers identified for gentamycin appears to offer better prediction/differentiation for gentamycin susceptibility phenotype with improved sensitivity of 45-90%, improved specificity of 30-45%, improved PPV of 20-80% and improved NPV of 25-50%

- From the results shown in Fig 6.7 and Fig 6.11, the set of markers identified for amikacin appears to offer better prediction/differentiation for amikacin susceptibility phenotype with improved sensitivity of 25-75%, improved specificity of 20-45%, improved PPV of 80-90%.

In addition, when using comparative behavioral genomics as shown in *Chapter 5*, the following group of markers showed improved prediction.

- M28 (Inner membrane protein YccS G2184A) showed 30% improved specificity as compared to genomic profile 1 and 31-87% improved PPV as a ciprofloxacin molecular marker.
- M29 (Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR A13G) showed 6-15% improved specificity and 34-90% improved PPV as a ciprofloxacin molecular marker.
- M20 (Cysteine--tRNA ligase C738T) showed 27% improved specificity and 22% improved NPV as a ciprofloxacin molecular marker.
- M31 (Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR G57C) showed 37% improved specificity and 22% improved NPV as a ciprofloxacin molecular marker.
- M32 (HTH-type transcriptional regulator GadX G834C) showed 30% improved specificity and 22% improved NPV as a ciprofloxacin molecular marker.
- M43 (HTH-type transcriptional regulator YofA G511A) showed 31% improved specificity and 22% improved NPV as a ciprofloxacin molecular marker.
- The combination of M11(putative HTH-type transcriptional regulator V 91 L), M12 (putative HTH-type transcriptional regulator C 300 T), M13 (putative HTH-type transcriptional regulator M 126 L), M20 (Cysteine--tRNA ligase C 738 T), M43 (HTH-type transcriptional regulator YofA A 171 T), *nal*CA186T, and *amp*RD135N can differentiate ciprofloxacin resistance from susceptibility at sensitivity of (94.6-100%), specificity of (80.9-100%), PPV of (14.3-100%), and NPV of (98.4-100%). This shows improved performance when compared to the set of predictors identified using the evaluated databases at improved sensitivity of 50%, improved specificity of 10-30%, improved PPV of 4-40%.
- M31(Inner membrane protein YohK C502T) showed 13-36% improved specificity and 28-63% improved PPV as a gentamycin molecular marker.

- M32 (Inner membrane protein YohK T517C) showed 13-36% improved specificity and 32-60% improved PPV as a gentamycin molecular marker.
- M29 (30S ribosomal protein S1 C386A) showed 7-32% improved specificity and 32-60% improved PPV as a gentamycin molecular marker.
- M28 (Ribosomal protein S6--L-glutamate ligase C744T) showed 3-28% improved specificity and 28-63% improved PPV as a gentamycin molecular marker.
- M36 (Cytochrome c oxidase subunit 3 C846A) showed 16% improved specificity and 18-53% improved PPV as a gentamycin molecular marker.
- The combination of M7 (Cell division inhibitor *Sul*A C 111 T), M21 (putative GST-like protein *Yib*F C 678 T), M25 (Bifunctional protein *Put*A T 2958 C), M1 (NADH oxidase T 681 C), M11 (Glutathione-regulated potassium-efflux system protein KefC C 393 T), M22 (Translation initiation factor IF-2 C 261 T), M35 (Long-chain-fatty-acid--CoA ligase C1623 G) in addition to combinations shown above with gentamycin (C1 and C2, Table 7.2) can differentiate gentamycin resistance from susceptibility at sensitivity of (28.8-100%), specificity of (86.5-100%), PPV of (22.2-100%), and NPV of (51.6-100%). This shows improved performance when compared to the set of predictors identified using the evaluated databases at improved sensitivity of 24-46%, improved specificity of 15-30%, improved PPV of 20% and improved NPV of 22%.

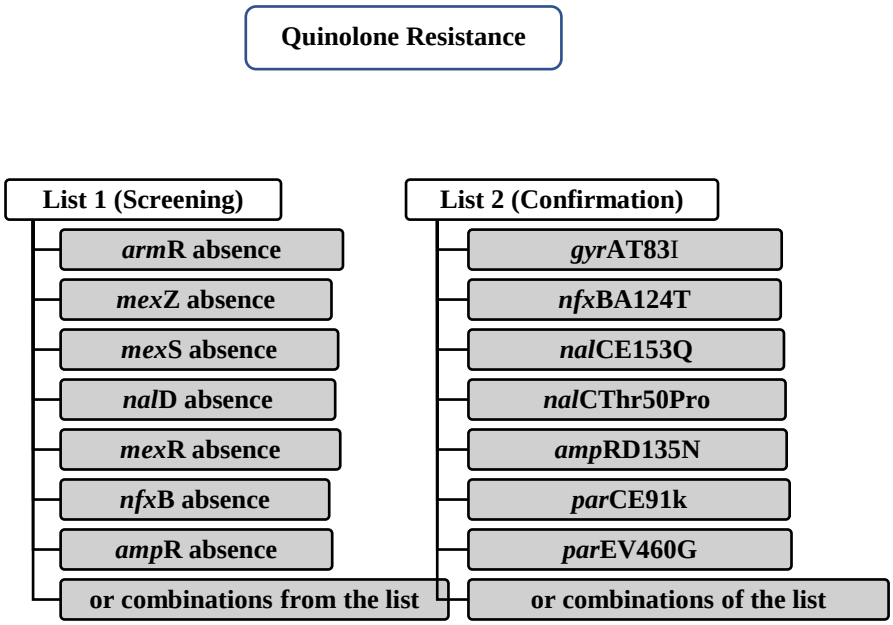
6.8. Applied Example/Summary

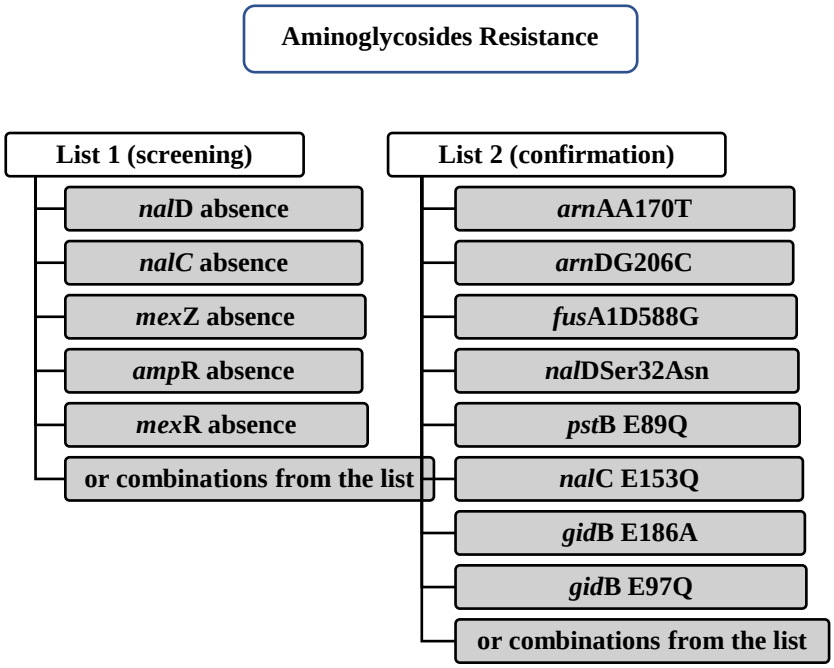
- In a diagnostic setting, the combinations of tested markers shown can be used as a list to choose from. Each marker shows different form of improved predictive capabilities.
- Each of the markers/combinations tested have a different frequency of occurrence in the population. This means that in a randomly selected sample, there are different chances of encountering the marker when a test is performed.
- That is the case where we can use the list of markers showing higher predictive capabilities. Some markers can be infrequent or uncommon and this consequently results in a lower chance of being detected when the diagnostic test is performed.
- In case a marker is not detected in a sample, we need to test for another marker/combination from the list of markers showing best performance.

Example: If we have an unknown clinical sample which is positive for *Ps. aeruginosa* with some available sequencing information, we would start by testing for combinations/markers showing

higher sensitivities as a first step screening, then for those positive for any of the listed markers/combinations showing higher sensitivity, we would then test for markers/combinations showing higher specificities as a second step.

The first step would be used as a screening diagnostic. This step aims at detecting all possible positive cases showing antibiotic resistance. By doing that, we increase the possibility of detecting all cases of resistance and thus ensuring the best possible clinical outcome and better patients' health. This consequently results in reducing mortality and complications related to antibiotic resistance. However, tests with high sensitivity also have the drawback of including/detecting some false negative cases. That is the reason why sensitive tests should be followed by more specific tests. By doing so, we increase the chance of correctly identifying all true negative cases using a specific test together with identifying all true positive cases using a sensitive test. At the same time, this helps to reduce false positive cases which ultimately results in avoiding the cost of antibiotic over-prescription.





List 1 should be first used to look up individual or/combinations of markers as a 1st step screening. Markers included in list 1 show high sensitivity ± high NPV.

List 2 should then be used to look up individual/combinations of markers as a second step confirmation. Markers included in list 2 show higher specificity ± high PPV.

Predictive performance of markers and combinations are detailed in Table 6.1 and Table 6.4.

Values of improved performance including improved Sensitivity, Specificity, PPV, and NPV are illustrated in diagrams 6.8 – 6.11.

Concluding statement

Review of available rapid diagnostic platforms showed that some POC diagnostics can detect mutations and can also detect some resistance-related gene variants in combination through gene amplification followed by microarray detection or gene sequencing. However, there is a lack of research evidence that informs about the best potential set of molecular markers with the possibility to act as best diagnostic predictors. All available research and diagnostic platforms build on a very small subset of established resistance elements and even do not expand to consider the breadth of information available in the literature in addition to including results from comprehensive research approaches such as GWAS. The thesis has contributed to fill this gap in knowledge by creating a new understanding about the combinatorial quantitative

contribution of different resistance mechanisms and by suggesting a way of using this approach to be applied in a diagnostic setting. In that way, it has combined the existing body of knowledge about resistance-related variants with results of comparative behavioral genomics into a new way to understand resistance-underlying elements and applying it at the practical level. This can be achieved by translating this new understanding into suggested assessment and decision algorithms for antibiotic prescription in addition to the potential of using this knowledge to develop new antibiotic-resistance diagnostic panels using different sequencing platforms including both NGS and metagenomics.

Results of the analysis performed here suggest that genome sequencing, once introduced into clinical practice with the additional availability of all related infrastructure, would offer the ultimate resolution and broadest coverage for detecting all resistance variants. This would also offer the added advantage of adapting the panels used for detection. It also carries the additional value of attaching all other genomic information which can offer the best optimized and personalized infectious diseases management options. Using genome sequencing as a clinical diagnostic tool to detect the best combination of markers showing higher predictive performance from clinical specimens would consequently impact treatment and offer better choices. Adapting that approach would also provide up-to-date epidemiologic information about the incidence and trends of new resistance mechanisms that would inform the future development of this type of "flexible adaptable format "diagnostic. To achieve that and to bring it closer to practice, rapid new portable or benchtop sequencing technologies need to be used together with rapid DNA extraction, rapid direct colony sequencing or metagenomic sequencing directly from clinical samples. This would greatly help to shorten the time to answer cycle.

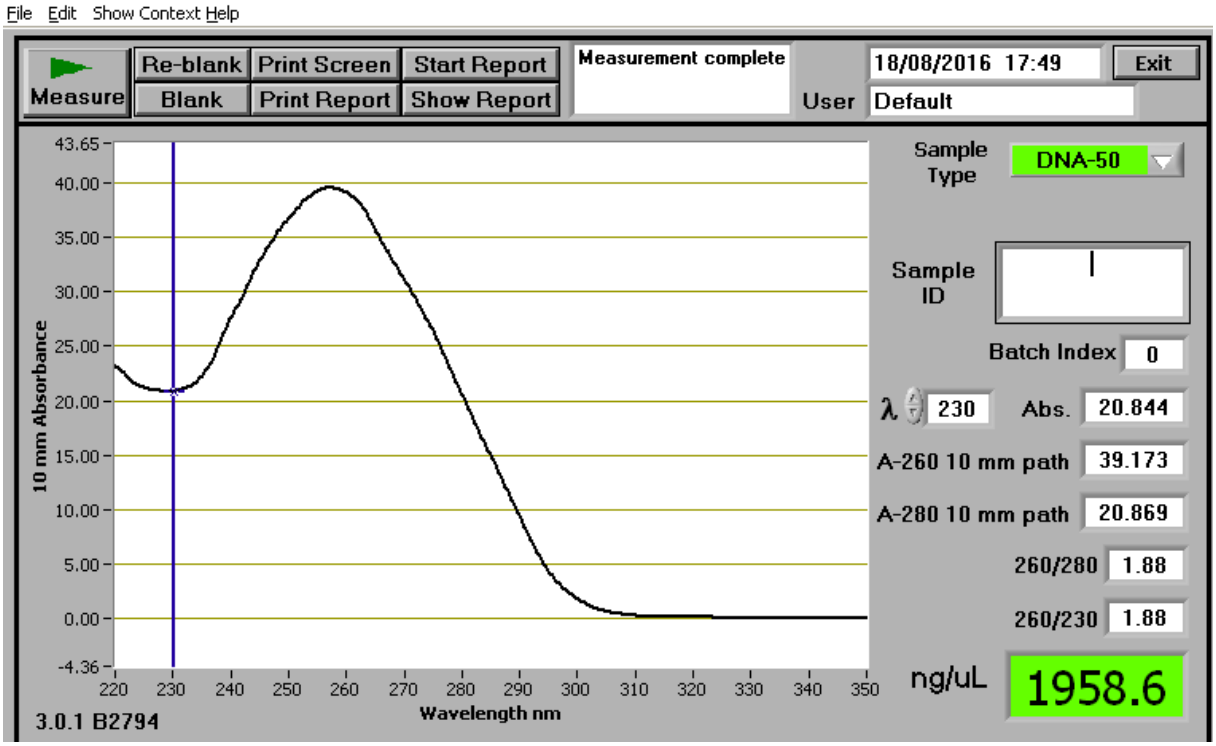
The thesis has attempted to follow a novel approach that explores for the best informative panel of markers which can be used for rapid antimicrobial resistance detection using different sequencing platforms. Development of high throughput research approaches that improve our understanding about the genomic bases of phenotype help to provide personalized sequence-based diagnostic algorithms which can greatly improve infectious disease management. The highest informative value genome sequencing can provide together with the exploitation of the advancement in DNA rapid extraction and new sequencing technologies can bring this new research-based knowledge closer to the clinic.

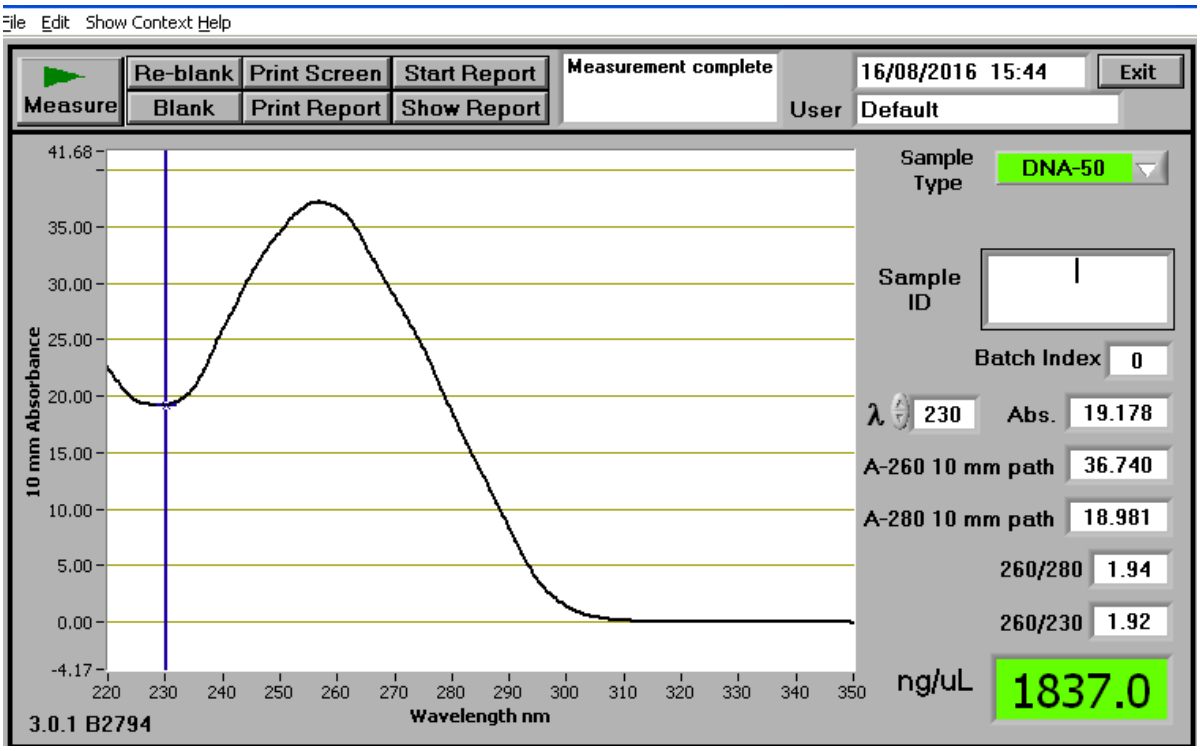
Appendix I

McFarland Standard

A $BaSO_4$ 0.5 McFarland standard used to standardize the inoculum density for a susceptibility test was prepared by adding 0.5-ml aliquot of 0.048 mol/L $BaCl_2$ to 99.5 ml of 0.18 mol/L H_2SO_4 with constant stirring to maintain a suspension. The correct density of the turbidity standard should be verified using a spectrophotometer. The absorbance at 625 nm should be 0.008 to 0.10 for the 0.5 McFarland standard.

Nanodrop -DNA quality check curves





Qubit DNA quantification

Qubit® dsDNA HS Assay Kits and the Qubit® Fluorometer instrument were used to quantify dsDNA according to the following steps:

Clear thin walled 0.5 ml PCR tubes are set up and labelled according to the number of samples and standards. The Qubit® dsDNA HS Assay requires 2 standards.

The Qubit® working solution is prepared by diluting the Qubit® dsDNA HS Reagent 1:200 in Qubit® dsDNA HS Buffer in a clean plastic tube.

Sufficient Qubit® working solution to accommodate all standards and samples so that the final volume in each tube is 200 µl.

To prepare standards, 190 µL of Qubit® working solution is added to each of the tubes used for standards with 10 µL of each Qubit® standard then mixed by vortexing for 2–3 seconds.

To prepare assay tubes, Qubit® working solution is added at the volume anywhere from 180–199 µL to each assay tube based on the sample volume which can be anywhere from 1–20 µL. The assay tubes with final volume of 200 µl are then mixed by vortexing for 2–3 seconds.

All tubes are then allowed to incubate at room temperature for 2 minutes.

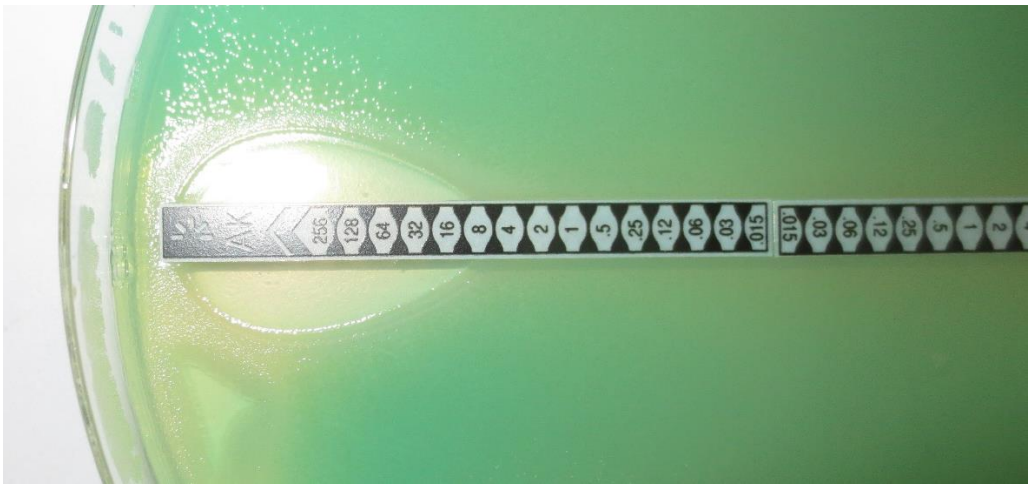
Both standards are then used to measure samples using the Qubit® 3.0 Fluorometer according to instructions shown in the Qubit® 3.0 Fluorometer User Guide.

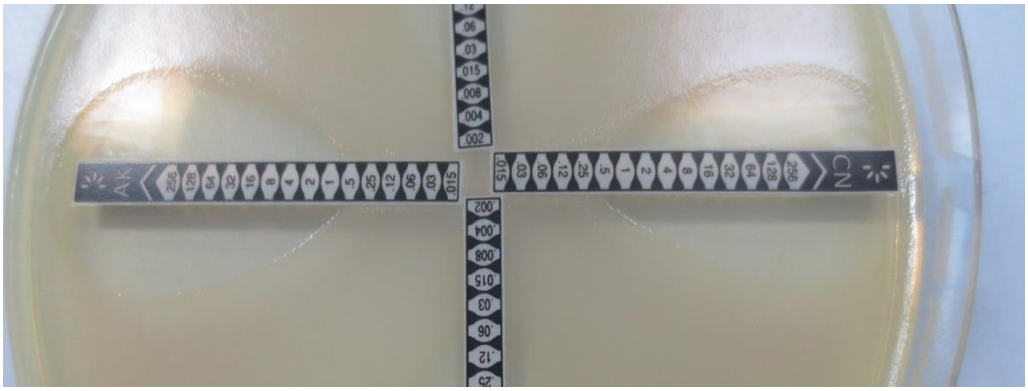
The Qubit® 2.0 Fluorometer gives values for the Qubit® dsDNA HS Assay in ng/mL which corresponds to the concentration after the sample was diluted into the assay tube.

The concentration in the sample is calculated according to the following equation:

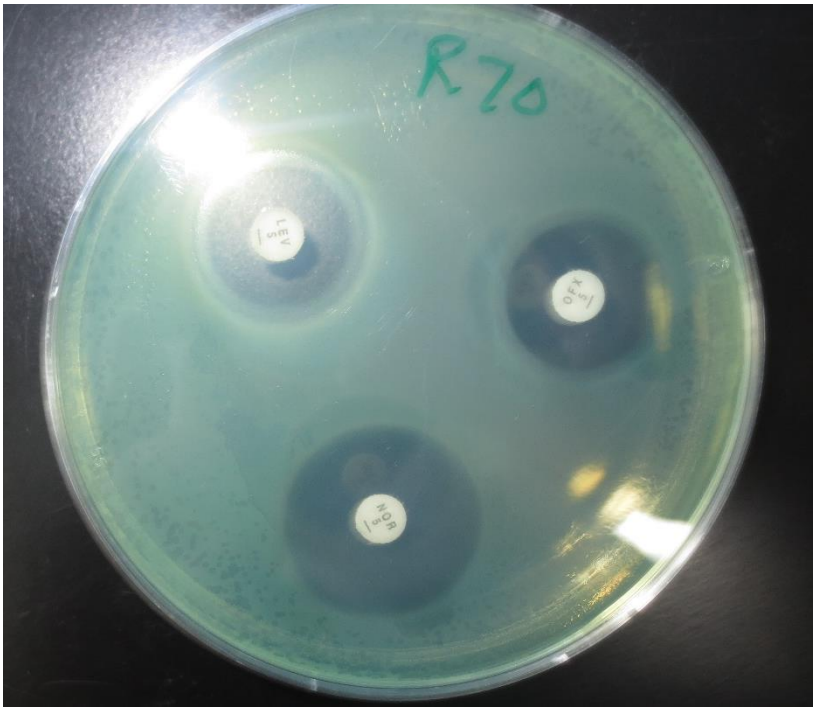
Concentration of the sample = QF value ×(200/X) where QF value = the value given by the Qubit® 2.0 Fluorometer, x = the number of microliters of sample added to the assay tube

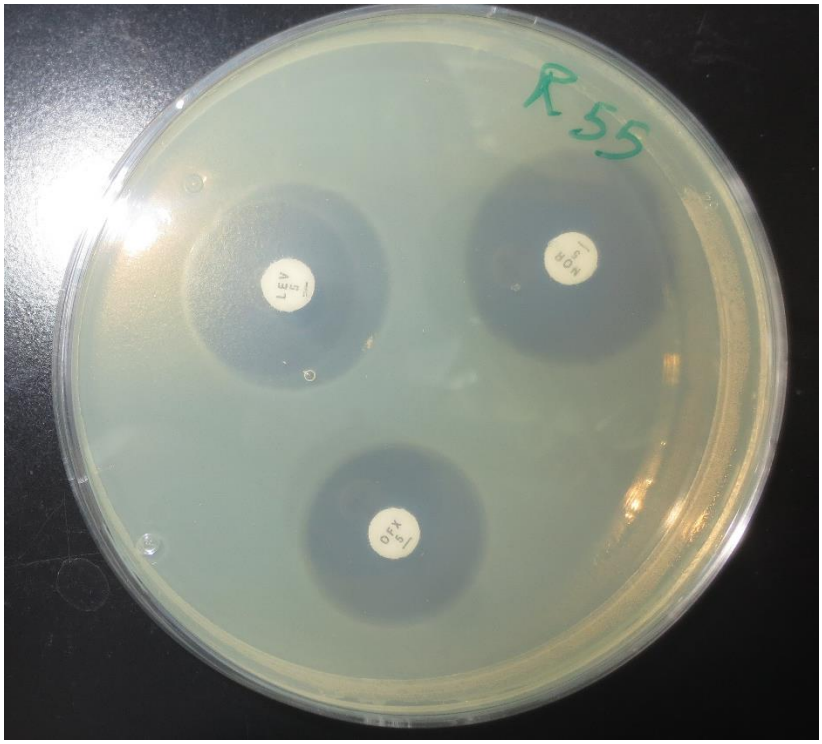
MIC evaluator strip



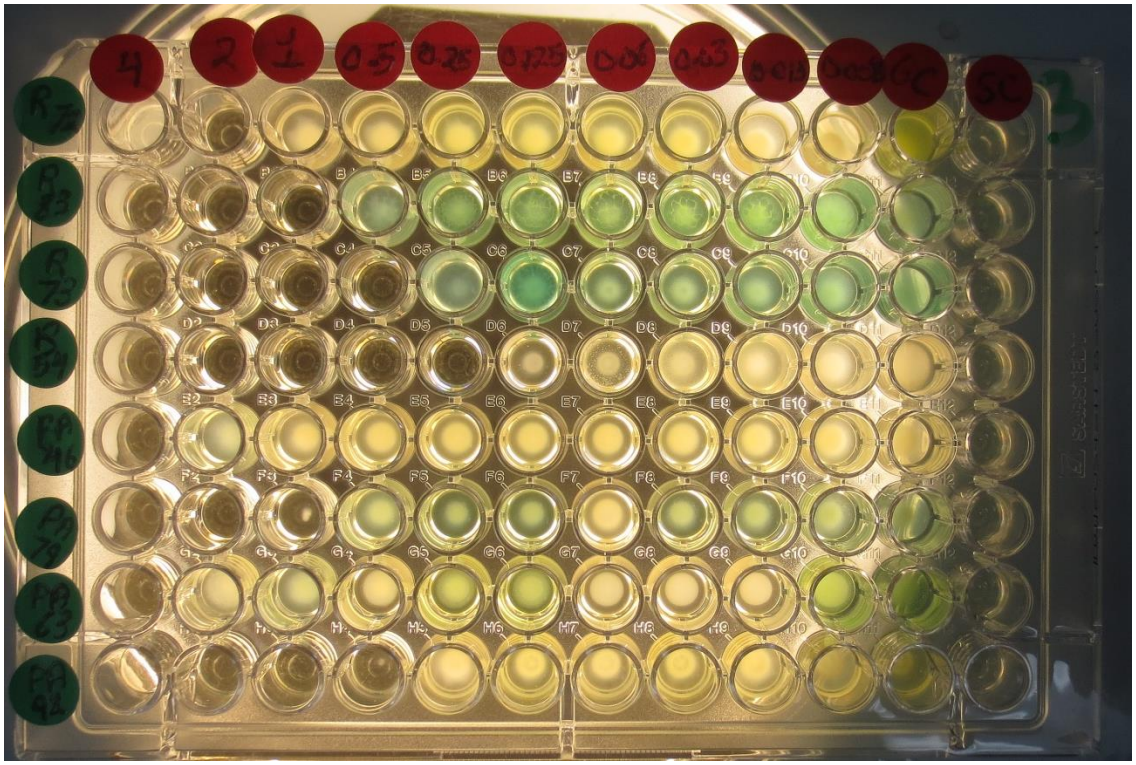


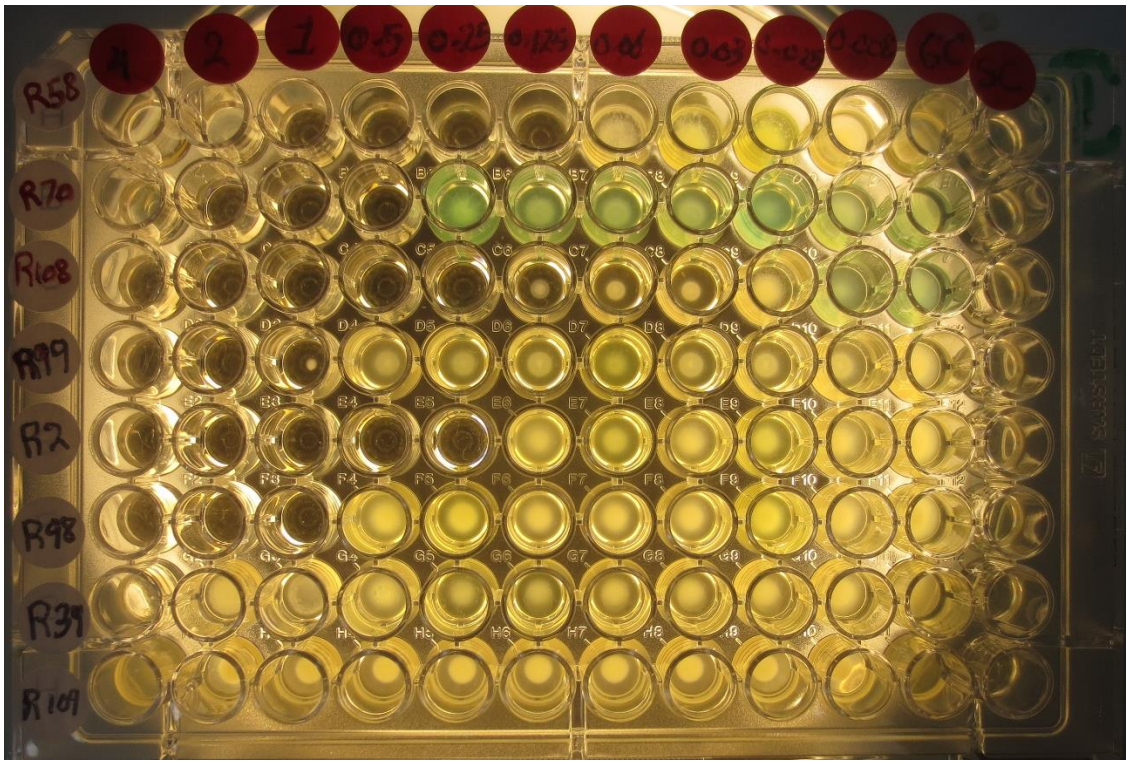
Disc-diffusion-Kirby-Bauer method



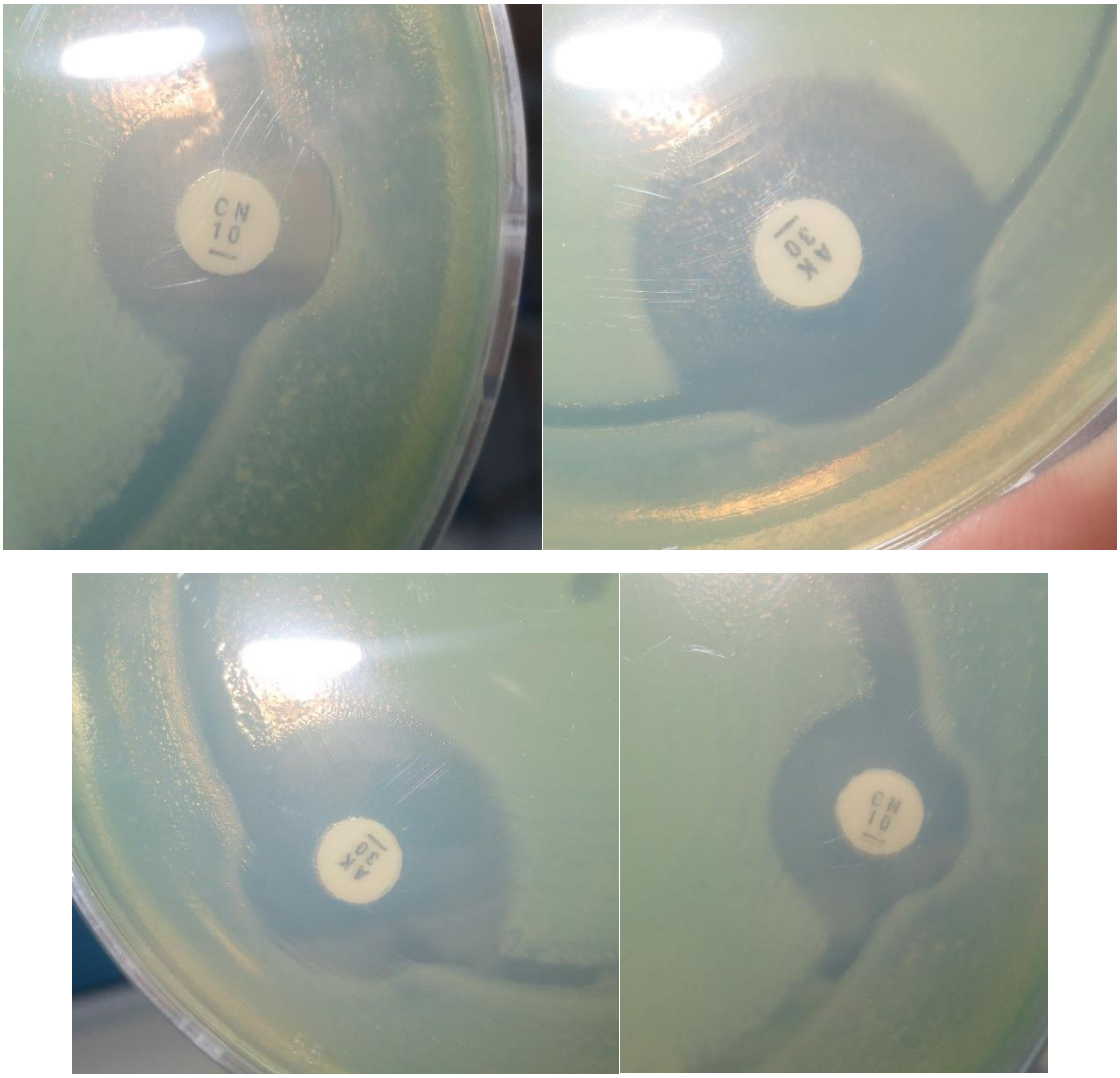


MIC by broth Microdilution





Disc-diffusion-Modified Stokes'



Preparation and storage of antibiotic stock and working solutions

Preparation of antibiotic stock solutions

Antibiotic stock solutions are prepared at high concentration (e.g. 1000 µg/ml) from commercially available antimicrobial powders (with given potency). The amount needed and the diluents can be calculated by using either of the following formulas to determine the amount of antimicrobial powder (1) or diluent (2) needed for a standard solution.

$$(1)\text{Weight (mg)} = \frac{\text{Volume (mL)} \cdot \text{Concentration (µg/mL)}}{\text{Potency (µg/mg)}}$$

or

$$(2)\text{Volume (mL)} = \frac{\text{Weight (mg)} \cdot \text{Potency (µg/mg)}}{\text{Concentration (µg/mL)}}$$

Preparation of antibiotic dilution ranges according to the following table:

Stock	Volume from stock	CAMH broth	Concentration In working stock	Final concentration in test plate after adding bacterial suspension
10.000	256ul	20ml	128	64
10.000	128ul	20ml	64	32
10.000	64ul	20ml	32	16
1000	320ul	20ml	16	8
1000	160ul	20ml	8	4
1000	80ul	20ml	4	2
1000	40ul	20ml	2	1
100	200ul	20ml	1	0.5
100	100ul	20ml	0.5	0.25
100	50ul	20ml	0.25	0.125
10	250ul	20ml	0.125	0.06
10	125ul	20ml	0.06	0.03
1	625ul	20ml	0.03	0.015
1	313ul	20ml	0.015	0.008
1	156ul	20ml	0.008	0.004

Storage of antibiotic Stock solutions

Antibiotic	Solvent	Diluent	4℃	-20℃	-70℃
Amikacin	water	water	7days	1 month	------
gentamycin			6 months	NR	NR
levofloxacin			---------	---------	---------
Ciprofloxacin	DMSO	water	2 weeks	3 months	3 months

NR: Not Recommended

DMSO: Dimethyl Sulphoxide

APPENDIX II

Supplementary Table 1A: List of Ps. aeruginosa genomes from Patric database analyzed for ciprofloxacin and levofloxacin resistance prediction

Genome ID	Strain	Completion Date	BioProject Accession	GenBank Accessions	Genome Length	GC Content
1611770.3	MRSN12121	2015-02-23	PRJNA273956	CP010892,CP010893,CP010894	6986763	63.79
287.1477	MRSN 20176	2015-05-05	PRJNA273956	JYGC02000000	6730817	65.93
287.1482	MRSN18971	2014-06-20	PRJNA240151	JFJU01000000	6391538	66.45
287.2972	AR_0103	2016-11-10	PRJNA292904	MPBP00000000	6900175	65.99
287.2973	AR_0092	2016-11-10	PRJNA292904	MPBS00000000	6963676	66.02
287.2975	AR_0100	2016-11-10	PRJNA292904	MPBQ00000000	6929574	65.98
287.2976	AR_0105	2016-11-10	PRJNA292904	MPBO00000000	6425652	66.22
287.2977	AR_0054	2016-11-10	PRJNA292904	MPBV00000000	7267705	65.81
287.2978	AR_0064	2016-11-10	PRJNA292904	MPBU00000000	6827144	66.05
287.2979	AR_0094	2016-11-10	PRJNA292904	MPBR00000000	6882581	66.05
287.2980	AR_0108	2016-11-10	PRJNA292904	MPBN00000000	6971125	65.98
287.5685	AR_0360	2018-04-25	PRJNA316321	CP027165	6463575	66.44145
287.5686	AR_0354	2018-04-25	PRJNA316321	CP027171	6747010	66.09722
287.5687	AR_0353	2018-04-25	PRJNA316321	CP027172,CP027173	7282236	65.70483
287.5688	AR_0230	2018-04-25	PRJNA316321	CP027174,CP027175,CP027176	7086054	65.87829
287.5689	AR_0356	2018-04-25	PRJNA316321	CP027169,CP027168,CP027170,CP027167	7247865	65.59437
287.5690	AR_0357	2018-04-25	PRJNA316321	CP027166	7162784	65.79707
287.5701	AR_0355	2018-03-14	PRJNA316321	PSQQ01000000	7172470	65.84778
287.5702	AR_0359	2018-03-14	PRJNA316321	PSQS01000000	6940345	65.91524
287.5703	AR_0358	2018-03-14	PRJNA316321	PSQR01000000	7284124	65.71017
287.5704	AR_0351	2018-03-14	PRJNA316321	PSQP01000000	6611054	66.24956
287.5746	AR_0443	2018-04-30	PRJNA316321	CP029147	6776714	65.83632
287.5747	AR_0440	2018-04-30	PRJNA316321	CP029148	7167215	65.84894
287.5748	AR439	2018-04-30	PRJNA316321	CP029097,CP029095,CP029096	7578039	65.41879
287.5749	AR442	2018-04-30	PRJNA316321	CP029090	7267567	65.75234
287.5750	AR445	2018-04-30	PRJNA316321	CP029088	7125975	65.78442
287.5751	AR441	2018-04-30	PRJNA316321	CP029093,CP029091,CP029092,CP029094	7245771	65.59486
287.5752	AR444	2018-04-30	PRJNA316321	CP029089	6853499	66.05279
287.5778	AR_0095	2018-03-14	PRJNA292904	CP027538	6822666	66.10762
287.5955	AR_0446	2018-06-06	PRJNA316321	CP029660	6475581	66.33031
287.5956	CCUG 70744	2018-06-06	PRJNA401330	CP023255	6859232	66.04034
287.5959	AR_0110	2018-06-10	PRJNA292904	CP029745	6799785	66.01369
287.5969	AR_0447	2018-06-22	PRJNA316321	QHCK01000000	7178792	65.86686
287.5971	AR_460	2018-07-05	PRJNA316321	CP030351	6303875	66.55481

287.5972	AR_455	2018-07-05	PRJNA316321	CP030328	6540996	65.8842
287.5973	AR_458	2018-07-05	PRJNA316321	CP030327	6685102	66.23354
287.6326	AR_0459	2018-07-16	PRJNA316321	QMGQ01000000	6752712	66.24611
287.6327	AR_0457	2018-07-16	PRJNA316321	QMGP01000000	7381251	65.63855
287.6328	AR_0456	2018-07-16	PRJNA316321	QMGO01000000	7079305	65.85722
287.6329	AR_0449	2018-07-16	PRJNA316321	QMGL01000000	6792240	65.96245
287.6330	AR_0241	2018-07-16	PRJNA316321	QMGI01000000	7226502	65.64898
287.6331	AR_0352	2018-07-16	PRJNA316321	QMGJ01000000	6418505	66.43134
287.6492	AR_0111	2018-09-17	PRJNA292904	CP032257,CP032256	7075653	65.83051
287.7771	MRSN8914	2018-12-24	PRJNA446057	RXTB01000000	7356800	65.33966
287.7772	MRSN8915	2018-12-24	PRJNA446057	RXTA01000000	7064713	65.6805
287.7773	MRSN994	2018-12-24	PRJNA446057	RXSX01000000	6734108	66.12464
287.7774	MRSN8139	2018-12-24	PRJNA446057	RXTE01000000	6575059	66.36821
287.7775	MRSN7014	2018-12-24	PRJNA446057	RXTH01000000	6420453	66.4322
287.7776	MRSN8912	2018-12-24	PRJNA446057	RXTC01000000	6986301	65.71874
287.7777	MRSN8136	2018-12-24	PRJNA446057	RXTF01000000	6907900	65.71381
287.7778	MRSN6695	2018-12-24	PRJNA446057	RXTJ01000000	6235816	66.57632
287.7779	MRSN6241	2018-12-24	PRJNA446057	RXTL01000000	7161217	65.6252
287.7780	MRSN6678	2018-12-24	PRJNA446057	RXTK01000000	6782774	66.11963
287.7781	MRSN6220	2018-12-24	PRJNA446057	RXTM01000000	6910198	66.00948
287.7782	MRSN5524	2018-12-24	PRJNA446057	RXTO01000000	6725073	65.99564
287.7783	MRSN5539	2018-12-24	PRJNA446057	RXTN01000000	6769662	66.07058
287.7784	MRSN552	2018-12-24	PRJNA446057	RXTP01000000	6444919	66.35499
287.7785	MRSN5519	2018-12-24	PRJNA446057	RXTQ01000000	6751302	66.13514
287.7786	MRSN4841	2018-12-24	PRJNA446057	RXTT01000000	6475204	66.25277
287.7787	MRSN5498	2018-12-24	PRJNA446057	RXTS01000000	6821230	66.00017
287.7788	MRSN443463	2018-12-24	PRJNA446057	RXTU01000000	6393241	66.4404
287.7789	MRSN315	2018-12-24	PRJNA446057	RXUI01000000	6696919	66.20516
287.7790	MRSN19711	2018-12-24	PRJNA446057	RXUX01000000	6265327	66.5761
287.7791	MRSN20190	2018-12-24	PRJNA446057	RXUV01000000	6191883	66.59317
287.7792	MRSN9873	2018-12-24	PRJNA446057	RXSY01000000	7146707	65.48819
287.7793	MRSN9718	2018-12-24	PRJNA446057	RXSZ01000000	6861353	66.01952
287.7794	MRSN8141	2018-12-24	PRJNA446057	RXTD01000000	7161718	65.6272
287.7795	MRSN8130	2018-12-24	PRJNA446057	RXTG01000000	6828028	66.09563
287.7796	MRSN6739	2018-12-24	PRJNA446057	RXTI01000000	6298064	66.47807
287.7797	MRSN5508	2018-12-24	PRJNA446057	RXTR01000000	6439859	66.34678
287.7798	MRSN435288	2018-12-24	PRJNA446057	RXTW01000000	6363888	66.22477
287.7799	MRSN321	2018-12-24	PRJNA446057	RXUG01000000	6260875	66.54782
287.7800	MRSN25678	2018-12-24	PRJNA446057	RXUN01000000	6509751	66.21528

287.7801	MRSN18754	2018-12-24	PRJNA446057	RXVH01000000	6270751	66.54153
287.7802	MRSN390231	2018-12-24	PRJNA446057	RXTZ01000000	6149846	66.37745
287.7803	MRSN436311	2018-12-24	PRJNA446057	RXTV01000000	6473360	66.41605
287.7804	MRSN373401	2018-12-24	PRJNA446057	RXUA01000000	6985982	65.97148
287.7805	MRSN3705	2018-12-24	PRJNA446057	RXUB01000000	6419743	66.57632
287.7806	MRSN29192	2018-12-24	PRJNA446057	RXUK01000000	6375161	66.43941
287.7807	MRSN25762	2018-12-24	PRJNA446057	RXUM01000000	6845305	66.2181
287.7808	MRSN2444	2018-12-24	PRJNA446057	RXUP01000000	6843047	66.06668
287.7809	MRSN3587	2018-12-24	PRJNA446057	RXUU01000000	6494873	66.36721
287.7810	MRSN2101	2018-12-24	PRJNA446057	RXUT01000000	6346725	66.37656
287.7811	MRSN23861	2018-12-24	PRJNA446057	RXUQ01000000	7025487	65.94264
287.7812	MRSN20176	2018-12-24	PRJNA446057	RXUW01000000	6717057	65.99654
287.7813	MRSN1948	2018-12-24	PRJNA446057	RXUY01000000	6674435	66.19558
287.7814	MRSN1906	2018-12-24	PRJNA446057	RXVB01000000	7049219	65.7835
287.7815	MRSN1938	2018-12-24	PRJNA446057	RXUZ01000000	7052906	66.11615
287.7816	MRSN1899	2018-12-24	PRJNA446057	RXVD01000000	6573139	66.22041
287.7817	MRSN1902	2018-12-24	PRJNA446057	RXVC01000000	6344431	66.45721
287.7818	MRSN18970	2018-12-24	PRJNA446057	RXVE01000000	6370604	66.51189
287.7819	MRSN17849	2018-12-24	PRJNA446057	RXVK01000000	6738933	66.12466
287.7820	MRSN16744	2018-12-24	PRJNA446057	RXVO01000000	6776946	66.11751
287.7821	MRSN16740	2018-12-24	PRJNA446057	RXVP01000000	6670940	66.08019
287.7822	MRSN1688	2018-12-24	PRJNA446057	RXVM01000000	6606996	66.14838
287.7823	MRSN16383	2018-12-24	PRJNA446057	RXVQ01000000	6302224	66.53333
287.7824	MRSN16345	2018-12-24	PRJNA446057	RXVR01000000	6333273	66.51163
287.7825	MRSN1613	2018-12-24	PRJNA446057	RXVU01000000	6276349	66.51006
287.7826	MRSN1601	2018-12-24	PRJNA446057	RXVW01000000	6411494	66.28397
287.7827	MRSN1583	2018-12-24	PRJNA446057	RXVX01000000	6463376	66.37041
287.7828	MRSN1617	2018-12-24	PRJNA446057	RXVT01000000	6780137	66.19642
287.7829	MRSN1388	2018-12-24	PRJNA446057	RXWC01000000	7171428	65.86237
287.7830	MRSN15678	2018-12-24	PRJNA446057	RXVZ01000000	6780257	66.09457
287.7831	MRSN14981	2018-12-24	PRJNA446057	RXWB01000000	6508561	66.21667
287.7832	MRSN409937	2018-12-24	PRJNA446057	RXTX01000000	6487691	66.33823
287.7833	MRSN351791	2018-12-24	PRJNA446057	RXUE01000000	6812488	65.93882
287.7834	MRSN346179	2018-12-24	PRJNA446057	RXUF01000000	6293739	66.51709
287.7835	MRSN30858	2018-12-24	PRJNA446057	RXUJ01000000	6382502	66.4238
287.7836	MRSN26263	2018-12-24	PRJNA446057	RXUL01000000	6430277	66.3206
287.7837	MRSN1925	2018-12-24	PRJNA446057	RXVA01000000	7130564	65.83213
287.7838	MRSN25623	2018-12-24	PRJNA446057	RXUO01000000	6635285	66.08652
287.7839	MRSN18562	2018-12-24	PRJNA446057	RXVI01000000	6494310	66.3681

287.7840	MRSN16344	2018-12-24	PRJNA446057	RXVS01000000	6318735	66.45731
287.7841	MRSN1612	2018-12-24	PRJNA446057	RXVV01000000	6506983	66.235
287.7842	MRSN1739	2018-12-24	PRJNA446057	RXVL01000000	7031489	65.82428
287.7843	MRSN16847	2018-12-24	PRJNA446057	RXVN01000000	6594082	65.97868
287.7844	MRSN15753	2018-12-24	PRJNA446057	RXVY01000000	6547438	66.18349
287.7845	MRSN1380	2018-12-24	PRJNA446057	RXWD01000000	6339603	66.48981
287.7846	MRSN15566	2018-12-24	PRJNA446057	RXWA01000000	6408848	66.38998
287.7847	MRSN1356	2018-12-24	PRJNA446057	RXWE01000000	6555630	66.19885
287.7848	MRSN12914	2018-12-24	PRJNA446057	RXWH01000000	6756810	65.92381
287.7849	MRSN12283	2018-12-24	PRJNA446057	RXWK01000000	6301640	66.49963
287.7850	MRSN12365	2018-12-24	PRJNA446057	RXWJ01000000	6869499	65.87287
287.7851	MRSN358800	2018-12-24	PRJNA446057	RXUD01000000	6501460	66.35765
287.7852	MRSN11538	2018-12-24	PRJNA446057	RXWN01000000	6528585	66.2616
287.7853	MRSN11278	2018-12-24	PRJNA446057	RXWS01000000	6893110	66.04308
287.7854	MRSN11281	2018-12-24	PRJNA446057	RXWR01000000	6323478	66.49251
287.7855	MRSN13488	2018-12-24	PRJNA446057	RXWF01000000	5912421	66.50286
287.7856	MRSN1344	2018-12-24	PRJNA446057	RXWG01000000	6232278	66.49448
287.7857	MRSN12368	2018-12-24	PRJNA446057	RXWI01000000	6340159	66.43633
287.7858	MRSN11285	2018-12-24	PRJNA446057	RXWQ01000000	6650291	66.16359
287.7859	MRSN12282	2018-12-24	PRJNA446057	RXWL01000000	6876949	65.96845
287.7860	MRSN11536	2018-12-24	PRJNA446057	RXWO01000000	6935556	65.91659
287.7861	MRSN11286	2018-12-24	PRJNA446057	RXWP01000000	6711313	65.91158
287.7862	MRSN11976	2018-12-24	PRJNA446057	RXWM01000000	6991711	65.90075
287.8027	MRSN401528	2018-12-24	PRJNA446057	RXTY01000000	6486368	66.14036
287.8028	MRSN369569	2018-12-24	PRJNA446057	RXUC01000000	6316238	66.51856
287.8029	MRSN2144	2018-12-24	PRJNA446057	RXUR01000000	7032738	65.79075
287.8030	MRSN18855	2018-12-24	PRJNA446057	RXVF01000000	6296953	66.48956
287.8031	MRSN2108	2018-12-24	PRJNA446057	RXUS01000000	6394452	66.39729
287.8032	MRSN18803	2018-12-24	PRJNA446057	RXVG01000000	6191380	66.5096
287.8033	MRSN317	2018-12-24	PRJNA446057	RXUH01000000	6349621	66.47357
287.8034	MRSN18560	2018-12-24	PRJNA446057	RXVJ01000000	6268481	66.5069

Supplementary Table 1B: List of Ps. aeruginosa genomes from Patric database analyzed for levofloxacin resistance prediction:

Genome ID	Strain	Completion Date	BioProject Accession	GenBank Accessions	Genome Length	GC Content
1611770.3	MRSN12121	2015-02-23	PRJNA273956	CP010892,CP010893,CP010894	6986763	63.79
287.1000	AZPAE14933	2014-12-04	PRJNA264310	JTRI00000000	7092428	65.74
287.1001	AZPAE14932	2014-12-04	PRJNA264310	JTRJ00000000	6638040	66.21
287.1002	AZPAE14931	2014-12-04	PRJNA264310	JTRK00000000	6359186	66.43
287.1003	AZPAE14930	2014-12-04	PRJNA264310	JTRL00000000	6302622	66.48
287.1004	AZPAE14929	2014-12-04	PRJNA264310	JTRM00000000	6750744	66.09
287.1005	AZPAE14928	2014-12-04	PRJNA264310	JTRN00000000	6257878	66.55
287.1006	AZPAE14927	2014-12-04	PRJNA264310	JTRO00000000	6885375	65.93
287.1007	AZPAE14926	2014-12-04	PRJNA264310	JTRP00000000	6886320	65.94
287.1008	AZPAE14925	2014-12-04	PRJNA264310	JTRQ00000000	6462755	66.29
287.1009	AZPAE14924	2014-12-04	PRJNA264310	JTRR00000000	7042302	65.8
287.1010	AZPAE14923	2014-12-04	PRJNA264310	JTRS00000000	6724711	66.11
287.1011	AZPAE14922	2014-12-04	PRJNA264310	JTRT00000000	6839077	65.98
287.1012	AZPAE14921	2014-12-04	PRJNA264310	JTRU00000000	6815406	65.99
287.1013	AZPAE14920	2014-12-04	PRJNA264310	JTRV00000000	6882761	66.06
287.1014	AZPAE14919	2014-12-04	PRJNA264310	JTRW00000000	6284486	66.28
287.1015	AZPAE14918	2014-12-04	PRJNA264310	JTRX00000000	6318316	66.47
287.1016	AZPAE14917	2014-12-04	PRJNA264310	JTRY00000000	6308307	66.52
287.1017	AZPAE14916	2014-12-04	PRJNA264310	JTRZ00000000	6242056	66.5
287.1018	AZPAE14915	2014-12-04	PRJNA264310	JTSA00000000	6880572	65.9
287.1019	AZPAE14914	2014-12-04	PRJNA264310	JTSB00000000	6858678	65.91
287.1020	AZPAE14913	2014-12-04	PRJNA264310	JTSC00000000	6454858	66.27
287.1021	AZPAE14912	2014-12-04	PRJNA264310	JTSD00000000	6835920	66.08
287.1022	AZPAE14911	2014-12-04	PRJNA264310	JTSE00000000	6808509	65.93
287.1023	AZPAE14910	2014-12-04	PRJNA264310	JTSF00000000	6253654	66.5
287.1024	AZPAE14909	2014-12-04	PRJNA264310	JTSG00000000	7056926	65.94
287.1025	AZPAE14908	2014-12-04	PRJNA264310	JTSH00000000	6815570	66.14
287.1026	AZPAE14907	2014-12-04	PRJNA264310	JTSI00000000	6392112	66.38
287.1027	AZPAE14906	2014-12-04	PRJNA264310	JTSJ00000000	6690011	66.19
287.1028	AZPAE14905	2014-12-04	PRJNA264310	JTSK00000000	6692731	66.2
287.1029	AZPAE14904	2014-12-04	PRJNA264310	JTSL00000000	6937853	65.87
287.1030	AZPAE14903	2014-12-04	PRJNA264310	JTSM00000000	6846030	65.86
287.1031	AZPAE14902	2014-12-04	PRJNA264310	JTSN00000000	6886932	66.05
287.1032	AZPAE14901	2014-12-04	PRJNA264310	JTSO00000000	6881448	66.13
287.1033	AZPAE14900	2014-12-04	PRJNA264310	JTSP00000000	6851898	65.79
287.1034	AZPAE14899	2014-12-04	PRJNA264310	JTSQ00000000	6470623	66.16

287.1035	AZPAE14898	2014-12-04	PRJNA264310	JTSR00000000	6577690	66.12
287.1036	AZPAE14897	2014-12-04	PRJNA264310	JTSS00000000	6448155	66.24
287.1037	AZPAE14895	2014-12-04	PRJNA264310	JTST00000000	6330823	66.42
287.1038	AZPAE14894	2014-12-04	PRJNA264310	JTSU00000000	7003752	65.96
287.1039	AZPAE14893	2014-12-04	PRJNA264310	JTSV00000000	6769601	65.96
287.1040	AZPAE14892	2014-12-04	PRJNA264310	JTSW00000000	6609151	66.05
287.1041	AZPAE14891	2014-12-04	PRJNA264310	JTSX00000000	6447447	66.38
287.1042	AZPAE14890	2014-12-04	PRJNA264310	JTSY00000000	7016044	65.97
287.1043	AZPAE14889	2014-12-04	PRJNA264310	JTSZ00000000	6479466	66.26
287.1044	AZPAE14888	2014-12-04	PRJNA264310	JTTA00000000	6680559	66.15
287.1045	AZPAE14887	2014-12-04	PRJNA264310	JTTB00000000	6903570	66.01
287.1046	AZPAE14886	2014-12-04	PRJNA264310	JTTC00000000	7057811	65.77
287.1047	AZPAE14885	2014-12-04	PRJNA264310	JTTD00000000	6406036	66.22
287.1048	AZPAE14884	2014-12-04	PRJNA264310	JTTE00000000	6434371	66.45
287.1049	AZPAE14883	2014-12-04	PRJNA264310	JTTF00000000	6309393	66.42
287.1050	AZPAE14882	2014-12-04	PRJNA264310	JTTG00000000	6455645	66.42
287.1051	AZPAE14881	2014-12-04	PRJNA264310	JTTH00000000	6703242	66.06
287.1052	AZPAE14880	2014-12-04	PRJNA264310	JTTI00000000	6864526	66.11
287.1053	AZPAE14879	2014-12-04	PRJNA264310	JTTJ00000000	6226729	66.54
287.1054	AZPAE14878	2014-12-04	PRJNA264310	JTTK00000000	6953921	65.92
287.1055	AZPAE14877	2014-12-04	PRJNA264310	JTTL00000000	6300065	66.37
287.1056	AZPAE14876	2014-12-04	PRJNA264310	JTTM00000000	6386042	66.4
287.1057	AZPAE14875	2014-12-04	PRJNA264310	JTTN00000000	6980738	65.87
287.1058	AZPAE14874	2014-12-04	PRJNA264310	JTTO00000000	6850351	66.05
287.1059	AZPAE14873	2014-12-04	PRJNA264310	JTTP00000000	6259215	66.47
287.1060	AZPAE14872	2014-12-04	PRJNA264310	JTTQ00000000	7250921	65.47
287.1061	AZPAE14871	2014-12-04	PRJNA264310	JTTR00000000	7054233	65.55
287.1062	AZPAE14870	2014-12-04	PRJNA264310	JTTS00000000	6899104	66.03
287.1063	AZPAE14869	2014-12-04	PRJNA264310	JTTT00000000	6796086	66.04
287.1064	AZPAE14868	2014-12-04	PRJNA264310	JTTU00000000	6390696	66.38
287.1065	AZPAE14867	2014-12-04	PRJNA264310	JTTV00000000	6785052	65.93
287.1066	AZPAE14866	2014-12-04	PRJNA264310	JTTW00000000	6452948	66.3
287.1067	AZPAE14865	2014-12-04	PRJNA264310	JTTX00000000	6894960	65.79
287.1068	AZPAE14864	2014-12-04	PRJNA264310	JTTY00000000	6160523	66.57
287.1069	AZPAE14863	2014-12-04	PRJNA264310	JTTZ00000000	7070053	65.55
287.1070	AZPAE14862	2014-12-04	PRJNA264310	JTUA00000000	7067572	65.64
287.1071	AZPAE14861	2014-12-04	PRJNA264310	JTUB00000000	6713012	66.14
287.1072	AZPAE14860	2014-12-04	PRJNA264310	JTUC00000000	6768731	66.15
287.1073	AZPAE14859	2014-12-04	PRJNA264310	JTUD00000000	6527863	66.22

287.1074	AZPAE14858	2014-12-04	PRJNA264310	JTUE00000000	6660551	65.95
287.1075	AZPAE14857	2014-12-04	PRJNA264310	JTUF00000000	6869309	66.04
287.1076	AZPAE14856	2014-12-04	PRJNA264310	JTUG00000000	6335571	66.42
287.1077	AZPAE14855	2014-12-04	PRJNA264310	JTUH00000000	6342011	66.15
287.1078	AZPAE14853	2014-12-04	PRJNA264310	JTUI00000000	6829167	66.09
287.1079	AZPAE14852	2014-12-04	PRJNA264310	JTUJ00000000	6860603	65.91
287.1080	AZPAE14851	2014-12-04	PRJNA264310	JTUK00000000	6617149	66.24
287.1081	AZPAE14850	2014-12-04	PRJNA264310	JTUL00000000	6729517	66.06
287.1082	AZPAE14848	2014-12-04	PRJNA264310	JTUM00000000	6252044	66.43
287.1083	AZPAE14847	2014-12-04	PRJNA264310	JTUN00000000	6345563	66.43
287.1084	AZPAE14846	2014-12-04	PRJNA264310	JTUO00000000	6819636	65.85
287.1085	AZPAE14845	2014-12-04	PRJNA264310	JTUP00000000	6621008	66
287.1086	AZPAE14844	2014-12-04	PRJNA264310	JTUQ00000000	6732772	66.12
287.1087	AZPAE14843	2014-12-04	PRJNA264310	JTUR00000000	6737478	65.98
287.1088	AZPAE14842	2014-12-04	PRJNA264310	JTUS00000000	6889407	65.93
287.1089	AZPAE14841	2014-12-04	PRJNA264310	JTUT00000000	6845205	66.01
287.1090	AZPAE14840	2014-12-04	PRJNA264310	JTUU00000000	7049244	65.61
287.1091	AZPAE14839	2014-12-04	PRJNA264310	JTUV00000000	6615080	66.18
287.1092	AZPAE14838	2014-12-04	PRJNA264310	JTUW00000000	7113432	65.42
287.1093	AZPAE14837	2014-12-04	PRJNA264310	JTUX00000000	6449289	66.22
287.1094	AZPAE14836	2014-12-04	PRJNA264310	JTUY00000000	6438844	66.27
287.1095	AZPAE14835	2014-12-04	PRJNA264310	JTUZ00000000	6500090	66.12
287.1096	AZPAE14834	2014-12-04	PRJNA264310	JTVA00000000	6792091	65.97
287.1097	AZPAE14833	2014-12-04	PRJNA264310	JTVB00000000	6876687	65.96
287.1098	AZPAE14832	2014-12-04	PRJNA264310	JTVC00000000	6543022	66.02
287.1099	AZPAE14831	2014-12-04	PRJNA264310	JTVD00000000	7184792	65.56
287.1100	AZPAE14830	2014-12-04	PRJNA264310	JTVE00000000	6816534	66
287.1101	AZPAE14829	2014-12-04	PRJNA264310	JTVF00000000	6995537	66.05
287.1102	AZPAE14828	2014-12-04	PRJNA264310	JTVG00000000	6872882	66.04
287.1103	AZPAE14827	2014-12-04	PRJNA264310	JTVH00000000	7354112	65.59
287.1104	AZPAE14826	2014-12-04	PRJNA264310	JTVI00000000	7108825	65.87
287.1105	AZPAE14825	2014-12-04	PRJNA264310	JTVJ00000000	6426191	66.4
287.1106	AZPAE14824	2014-12-04	PRJNA264310	JTVK00000000	7034546	65.84
287.1107	AZPAE14823	2014-12-04	PRJNA264310	JTVL00000000	6849147	66.04
287.1108	AZPAE14822	2014-12-04	PRJNA264310	JTVM00000000	6668034	66.1
287.1109	AZPAE14821	2014-12-04	PRJNA264310	JTVN00000000	6811510	66.04
287.1110	AZPAE14820	2014-12-04	PRJNA264310	JTVO00000000	6364162	66.41
287.1111	AZPAE14819	2014-12-04	PRJNA264310	JTVP00000000	6660543	66.16
287.1112	AZPAE14818	2014-12-04	PRJNA264310	JTVQ00000000	6303594	66.39

287.1113	AZPAE14817	2014-12-04	PRJNA264310	JTVR00000000	6295916	66.49
287.1114	AZPAE14816	2014-12-04	PRJNA264310	JTVS00000000	6738242	66.04
287.1115	AZPAE14815	2014-12-04	PRJNA264310	JTVT00000000	7030097	65.89
287.1116	AZPAE14814	2014-12-04	PRJNA264310	JTVU00000000	6843476	65.98
287.1117	AZPAE14813	2014-12-04	PRJNA264310	JTVV00000000	6470062	66.2
287.1118	AZPAE14812	2014-12-04	PRJNA264310	JTVW00000000	7093737	65.99
287.1119	AZPAE14811	2014-12-04	PRJNA264310	JTVX00000000	6981966	65.94
287.1120	AZPAE14810	2014-12-04	PRJNA264310	JTVY00000000	6316519	66.4
287.1121	AZPAE14809	2014-12-04	PRJNA264310	JTVZ00000000	6210274	66.51
287.1122	AZPAE14732	2014-12-04	PRJNA264310	JTWA00000000	6793558	66.06
287.1123	AZPAE14731	2014-12-04	PRJNA264310	JTWB00000000	6231966	66.5
287.1124	AZPAE14730	2014-12-04	PRJNA264310	JTWC00000000	6865760	66.02
287.1125	AZPAE14729	2014-12-04	PRJNA264310	JTWD00000000	7015575	65.78
287.1126	AZPAE14728	2014-12-04	PRJNA264310	JTWE00000000	7356424	65.59
287.1127	AZPAE14727	2014-12-04	PRJNA264310	JTWF00000000	7362494	65.57
287.1128	AZPAE14726	2014-12-04	PRJNA264310	JTWG00000000	6885273	65.84
287.1129	AZPAE14725	2014-12-04	PRJNA264310	JTWH00000000	6802459	66.14
287.1130	AZPAE14724	2014-12-04	PRJNA264310	JTWI00000000	7091089	65.79
287.1131	AZPAE14723	2014-12-04	PRJNA264310	JTWJ00000000	6639036	66.13
287.1132	AZPAE14722	2014-12-04	PRJNA264310	JTWK00000000	6781659	66.07
287.1133	AZPAE14721	2014-12-04	PRJNA264310	JTWL00000000	6867543	66.01
287.1134	AZPAE14720	2014-12-04	PRJNA264310	JTWM00000000	6813733	66.02
287.1135	AZPAE14719	2014-12-04	PRJNA264310	JTWN00000000	6855346	66
287.1136	AZPAE14718	2014-12-04	PRJNA264310	JTWO00000000	6910404	65.98
287.1137	AZPAE14717	2014-12-04	PRJNA264310	JTWP00000000	7039359	65.87
287.1138	AZPAE14716	2014-12-04	PRJNA264310	JTWQ00000000	7038646	65.93
287.1139	AZPAE14715	2014-12-04	PRJNA264310	JTWR00000000	7119339	65.79
287.1140	AZPAE14714	2014-12-04	PRJNA264310	JTWS00000000	7022276	65.78
287.1141	AZPAE14713	2014-12-04	PRJNA264310	JTWT00000000	7238306	65.6
287.1142	AZPAE14712	2014-12-04	PRJNA264310	JTWU00000000	7010520	65.95
287.1143	AZPAE14711	2014-12-04	PRJNA264310	JTWV00000000	7121620	65.79
287.1144	AZPAE14710	2014-12-04	PRJNA264310	JTWW00000000	6748304	66.14
287.1145	AZPAE14708	2014-12-04	PRJNA264310	JTWX00000000	6707536	66.05
287.1146	AZPAE14707	2014-12-04	PRJNA264310	JTWY00000000	7081209	65.78
287.1147	AZPAE14706	2014-12-04	PRJNA264310	JTWZ00000000	6921381	66.06
287.1148	AZPAE14705	2014-12-04	PRJNA264310	JTXA00000000	6888268	65.97
287.1149	AZPAE14704	2014-12-04	PRJNA264310	JTXB00000000	6282527	66.51
287.1150	AZPAE14703	2014-12-04	PRJNA264310	JTXC00000000	6892377	65.93
287.1151	AZPAE14702	2014-12-04	PRJNA264310	JTXD00000000	7131510	65.84

287.1152	AZPAE14701	2014-12-04	PRJNA264310	JTXE00000000	6990663	65.93
287.1153	AZPAE14700	2014-12-04	PRJNA264310	JTXF00000000	6754658	66.09
287.1154	AZPAE14699	2014-12-04	PRJNA264310	JTXG00000000	5743255	65.28
287.1155	AZPAE14698	2014-12-04	PRJNA264310	JTXH00000000	6467367	65.44
287.1156	AZPAE14697	2014-12-04	PRJNA264310	JTXI00000000	6531914	66.18
287.1157	AZPAE14695	2014-12-04	PRJNA264310	JTXJ00000000	6757024	66.06
287.1158	AZPAE14694	2014-12-04	PRJNA264310	JTXK00000000	6964286	65.86
287.1159	AZPAE14693	2014-12-04	PRJNA264310	JTXL00000000	6380431	66.44
287.1160	AZPAE14692	2014-12-04	PRJNA264310	JTXM00000000	7011166	65.9
287.1161	AZPAE14691	2014-12-04	PRJNA264310	JTXN00000000	6301926	66.57
287.1162	AZPAE14690	2014-12-04	PRJNA264310	JTXO00000000	6730564	66.1
287.1163	AZPAE14689	2014-12-04	PRJNA264310	JTXP00000000	7252889	65.55
287.1164	AZPAE14688	2014-12-04	PRJNA264310	JTXQ00000000	7007711	66.04
287.1165	AZPAE14687	2014-12-04	PRJNA264310	JTXR00000000	6995963	65.89
287.1166	AZPAE14570	2014-12-04	PRJNA264310	JTXS00000000	7094694	65.94
287.1167	AZPAE14566	2014-12-04	PRJNA264310	JTXT00000000	6870259	65.99
287.1168	AZPAE14557	2014-12-04	PRJNA264310	JTXU00000000	6772788	66.1
287.1169	AZPAE14554	2014-12-04	PRJNA264310	JTXV00000000	7101679	65.75
287.1170	AZPAE14550	2014-12-04	PRJNA264310	JTXW00000000	6317165	66.47
287.1171	AZPAE14538	2014-12-04	PRJNA264310	JTXX00000000	6267671	66.54
287.1172	AZPAE14535	2014-12-04	PRJNA264310	JTXY00000000	6944986	65.88
287.1173	AZPAE14533	2014-12-04	PRJNA264310	JTXZ00000000	6661879	66.27
287.1174	AZPAE14526	2014-12-04	PRJNA264310	JTYA00000000	6289098	66.55
287.1175	AZPAE14509	2014-12-04	PRJNA264310	JTYB00000000	6519283	66.29
287.1176	AZPAE14505	2014-12-04	PRJNA264310	JTYC00000000	6823044	66.04
287.1177	AZPAE14499	2014-12-04	PRJNA264310	JTYD00000000	6698732	66.12
287.1178	AZPAE14463	2014-12-04	PRJNA264310	JTYE00000000	7034157	65.85
287.1179	AZPAE14453	2014-12-04	PRJNA264310	JTYF00000000	7052856	65.76
287.1180	AZPAE14443	2014-12-04	PRJNA264310	JTYG00000000	6268752	66.59
287.1181	AZPAE14442	2014-12-04	PRJNA264310	JTYH00000000	7007955	65.97
287.1182	AZPAE14441	2014-12-04	PRJNA264310	JTYI00000000	6810460	65.93
287.1183	AZPAE14437	2014-12-04	PRJNA264310	JTYJ00000000	6714753	66.04
287.1184	AZPAE14422	2014-12-04	PRJNA264310	JTYK00000000	6732223	66.05
287.1185	AZPAE14415	2014-12-04	PRJNA264310	JTYL00000000	6982687	65.83
287.1186	AZPAE14410	2014-12-04	PRJNA264310	JTYM00000000	6295059	66.5
287.1187	AZPAE14404	2014-12-04	PRJNA264310	JTYN00000000	6575587	66.3
287.1188	AZPAE14403	2014-12-04	PRJNA264310	JTYO00000000	6907387	66.08
287.1189	AZPAE14402	2014-12-04	PRJNA264310	JTYP00000000	6466306	66.33
287.1190	AZPAE14398	2014-12-04	PRJNA264310	JTYQ00000000	6693786	66.09

287.1191	AZPAE14395	2014-12-04	PRJNA264310	JTYR00000000	6325978	66.49
287.1192	AZPAE14394	2014-12-04	PRJNA264310	JTYS00000000	6977785	66.01
287.1193	AZPAE14393	2014-12-04	PRJNA264310	JTYT00000000	6723039	66.23
287.1194	AZPAE14390	2014-12-04	PRJNA264310	JTYU00000000	6884000	65.97
287.1195	AZPAE14381	2014-12-04	PRJNA264310	JTYV00000000	6791911	65.83
287.1196	AZPAE14379	2014-12-04	PRJNA264310	JTYW00000000	6253630	66.51
287.1197	AZPAE14373	2014-12-04	PRJNA264310	JTYX00000000	6337010	66.43
287.1198	AZPAE14372	2014-12-04	PRJNA264310	JTYY00000000	6794294	66.13
287.1199	AZPAE14359	2014-12-04	PRJNA264310	JTYZ00000000	6642619	66.29
287.1200	AZPAE14353	2014-12-04	PRJNA264310	JTZA00000000	6835254	66.01
287.1201	AZPAE14352	2014-12-04	PRJNA264310	JTZB00000000	6561378	66
287.1202	AZPAE13880	2014-12-04	PRJNA264310	JTZC00000000	6855952	66
287.1203	AZPAE13879	2014-12-04	PRJNA264310	JTZD00000000	6948193	65.92
287.1204	AZPAE13877	2014-12-04	PRJNA264310	JTZE00000000	6788523	66.1
287.1205	AZPAE13876	2014-12-04	PRJNA264310	JTZF00000000	6795623	65.85
287.1206	AZPAE13872	2014-12-04	PRJNA264310	JTZG00000000	6958295	65.9
287.1207	AZPAE13866	2014-12-04	PRJNA264310	JTZH00000000	7258328	65.73
287.1208	AZPAE13864	2014-12-04	PRJNA264310	JTZI00000000	6909435	65.95
287.1209	AZPAE13860	2014-12-04	PRJNA264310	JTZJ00000000	6888735	65.91
287.1210	AZPAE13858	2014-12-04	PRJNA264310	JTZK00000000	7017824	66.02
287.1211	AZPAE13856	2014-12-04	PRJNA264310	JTZL00000000	7373661	65.7
287.1212	AZPAE13853	2014-12-04	PRJNA264310	JTZM00000000	7040099	65.99
287.1213	AZPAE13850	2014-12-04	PRJNA264310	JTZN00000000	5501656	65.73
287.1214	AZPAE13848	2014-12-04	PRJNA264310	JTZO00000000	6958460	66.07
287.1215	AZPAE13757	2014-12-04	PRJNA264310	JTZP00000000	6469470	66.38
287.1216	AZPAE13756	2014-12-04	PRJNA264310	JTZQ00000000	7335469	65.54
287.1217	AZPAE12423	2014-12-04	PRJNA264310	JTZR00000000	6535360	66.21
287.1218	AZPAE12422	2014-12-04	PRJNA264310	JTZS00000000	6331130	66.45
287.1219	AZPAE12421	2014-12-04	PRJNA264310	JTZT00000000	6243887	66.4
287.1220	AZPAE12420	2014-12-04	PRJNA264310	JTZU00000000	6283659	66.37
287.1221	AZPAE12419	2014-12-04	PRJNA264310	JTZV00000000	6290772	66.46
287.1222	AZPAE12418	2014-12-04	PRJNA264310	JTZW00000000	6737495	66.21
287.1223	AZPAE12417	2014-12-04	PRJNA264310	JTZX00000000	6360200	66.3
287.1224	AZPAE12416	2014-12-04	PRJNA264310	JTZY00000000	6527924	66.15
287.1225	AZPAE12415	2014-12-04	PRJNA264310	JTZZ00000000	6740389	66.26
287.1226	AZPAE12414	2014-12-04	PRJNA264310	JUAA00000000	6736819	66.28
287.1227	AZPAE12413	2014-12-04	PRJNA264310	JUAB00000000	6742361	66.26
287.1228	AZPAE12412	2014-12-04	PRJNA264310	JUAC00000000	6445104	66.39
287.1229	AZPAE12411	2014-12-04	PRJNA264310	JUAD00000000	6341911	66.42

287.1230	AZPAE12410	2014-12-04	PRJNA264310	JUAE00000000	6457280	66.37
287.1231	AZPAE12409	2014-12-04	PRJNA264310	JUAF00000000	6758579	65.98
287.1232	AZPAE12156	2014-12-04	PRJNA264310	JUAG00000000	6571733	66.05
287.1233	AZPAE12155	2014-12-04	PRJNA264310	JUAH00000000	6732273	66.12
287.1234	AZPAE12154	2014-12-04	PRJNA264310	JUAI00000000	6299347	66.38
287.1235	AZPAE12153	2014-12-04	PRJNA264310	JUAJ00000000	6193644	66.55
287.1237	AZPAE12151	2014-12-04	PRJNA264310	JUAL00000000	6332026	66.47
287.1238	AZPAE12150	2014-12-04	PRJNA264310	JUAM00000000	6594453	65.98
287.1239	AZPAE12149	2014-12-04	PRJNA264310	JUAN00000000	6355203	66.36
287.1240	AZPAE12148	2014-12-04	PRJNA264310	JUAO00000000	6650984	66.28
287.1241	AZPAE12147	2014-12-04	PRJNA264310	JUAP00000000	6522044	66.25
287.1242	AZPAE12146	2014-12-04	PRJNA264310	JUAQ00000000	6313502	66.44
287.1243	AZPAE12145	2014-12-04	PRJNA264310	JUAR00000000	6452923	66.4
287.1245	AZPAE12143	2014-12-04	PRJNA264310	JUAT00000000	6473536	66.34
287.1246	AZPAE12142	2014-12-04	PRJNA264310	JUAU00000000	6323807	66.42
287.1247	AZPAE12140	2014-12-04	PRJNA264310	JUAV00000000	6250380	66.55
287.1248	AZPAE12138	2014-12-04	PRJNA264310	JUAW00000000	6576673	66.14
287.1249	AZPAE12137	2014-12-04	PRJNA264310	JUAX00000000	6304247	66.52
287.1250	AZPAE12136	2014-12-04	PRJNA264310	JUAY00000000	6303512	66.43
287.1301	AZPAE12135	2014-12-04	PRJNA264310	JUAZ00000000	6826162	66.07
287.1477	MRSN 20176	2015-05-05	PRJNA273956	JYGC02000000	6730817	65.93
287.1482	MRSN18971	2014-06-20	PRJNA240151	JFJU01000000	6391538	66.45
287.2972	AR_0103	2016-11-10	PRJNA292904	MPBP00000000	6900175	65.99
287.2973	AR_0092	2016-11-10	PRJNA292904	MPBS00000000	6963676	66.02
287.2975	AR_0100	2016-11-10	PRJNA292904	MPBQ00000000	6929574	65.98
287.2976	AR_0105	2016-11-10	PRJNA292904	MPBO00000000	6425652	66.22
287.2977	AR_0054	2016-11-10	PRJNA292904	MPBV00000000	7267705	65.81
287.2978	AR_0064	2016-11-10	PRJNA292904	MPBU00000000	6827144	66.05
287.2979	AR_0094	2016-11-10	PRJNA292904	MPBR00000000	6882581	66.05
287.2980	AR_0108	2016-11-10	PRJNA292904	MPBN00000000	6971125	65.98
287.5685	AR_0360	2018-04-25	PRJNA316321	CP027165	6463575	66.44145
287.5686	AR_0354	2018-04-25	PRJNA316321	CP027171	6747010	66.09722
287.5687	AR_0353	2018-04-25	PRJNA316321	CP027172,CP027173	7282236	65.704834
287.5688	AR_0230	2018-04-25	PRJNA316321	CP027174,CP027175,CP027176	7086054	65.87829
287.5689	AR_0356	2018-04-25	PRJNA316321	CP027169,CP027168,CP027170,CP027167	7247865	65.59437
287.5690	AR_0357	2018-04-25	PRJNA316321	CP027166	7162784	65.79707
287.5701	AR_0355	2018-03-14	PRJNA316321	PSQQ01000000	7172470	65.84778
287.5702	AR_0359	2018-03-14	PRJNA316321	PSQS01000000	6940345	65.91524
287.5703	AR_0358	2018-03-14	PRJNA316321	PSQR01000000	7284124	65.71017

287.5704	AR_0351	2018-03-14	PRJNA316321	PSQP01000000	6611054	66.24956
287.5746	AR_0443	2018-04-30	PRJNA316321	CP029147	6776714	65.83632
287.5747	AR_0440	2018-04-30	PRJNA316321	CP029148	7167215	65.84894
287.5748	AR439	2018-04-30	PRJNA316321	CP029097,CP029095,CP029096	7578039	65.418785
287.5749	AR442	2018-04-30	PRJNA316321	CP029090	7267567	65.752335
287.5750	AR445	2018-04-30	PRJNA316321	CP029088	7125975	65.78442
287.5751	AR441	2018-04-30	PRJNA316321	CP029093,CP029091,CP029092,CP029094	7245771	65.59486
287.5752	AR444	2018-04-30	PRJNA316321	CP029089	6853499	66.05279
287.5778	AR_0095	2018-03-14	PRJNA292904	CP027538	6822666	66.10762
287.5955	AR_0446	2018-06-06	PRJNA316321	CP029660	6475581	66.33031
287.5956	CCUG 70744	2018-06-06	PRJNA401330	CP023255	6859232	66.04034
287.5959	AR_0110	2018-06-10	PRJNA292904	CP029745	6799785	66.01369
287.5969	AR_0447	2018-06-22	PRJNA316321	QHCK01000000	7178792	65.86686
287.5971	AR_460	2018-07-05	PRJNA316321	CP030351	6303875	66.55481
287.5972	AR_455	2018-07-05	PRJNA316321	CP030328	6540996	65.8842
287.5973	AR_458	2018-07-05	PRJNA316321	CP030327	6685102	66.23354
287.6326	AR_0459	2018-07-16	PRJNA316321	QMGQ01000000	6752712	66.24611
287.6327	AR_0457	2018-07-16	PRJNA316321	QMGP01000000	7381251	65.63855
287.6328	AR_0456	2018-07-16	PRJNA316321	QMGO01000000	7079305	65.857216
287.6329	AR_0449	2018-07-16	PRJNA316321	QMGL01000000	6792240	65.96245
287.6330	AR_0241	2018-07-16	PRJNA316321	QMGI01000000	7226502	65.64898
287.6331	AR_0352	2018-07-16	PRJNA316321	QMGJ01000000	6418505	66.43134
287.6492	AR_0111	2018-09-17	PRJNA292904	CP032257,CP032256	7075653	65.830505
287.7771	MRSN8914	2018-12-24	PRJNA446057	RXTB01000000	7356800	65.33966
287.7772	MRSN8915	2018-12-24	PRJNA446057	RXTA01000000	7064713	65.680504
287.7773	MRSN994	2018-12-24	PRJNA446057	RXSX01000000	6734108	66.12464
287.7774	MRSN8139	2018-12-24	PRJNA446057	RXTE01000000	6575059	66.36821
287.7775	MRSN7014	2018-12-24	PRJNA446057	RXTH01000000	6420453	66.4322
287.7776	MRSN8912	2018-12-24	PRJNA446057	RXTC01000000	6986301	65.71874
287.7777	MRSN8136	2018-12-24	PRJNA446057	RXTF01000000	6907900	65.713806
287.7778	MRSN6695	2018-12-24	PRJNA446057	RXTJ01000000	6235816	66.576324
287.7779	MRSN6241	2018-12-24	PRJNA446057	RXTL01000000	7161217	65.6252
287.7780	MRSN6678	2018-12-24	PRJNA446057	RXTK01000000	6782774	66.11963
287.7781	MRSN6220	2018-12-24	PRJNA446057	RXTM01000000	6910198	66.00948
287.7782	MRSN5524	2018-12-24	PRJNA446057	RXTO01000000	6725073	65.99564
287.7783	MRSN5539	2018-12-24	PRJNA446057	RXTN01000000	6769662	66.07058
287.7784	MRSN552	2018-12-24	PRJNA446057	RXTP01000000	6444919	66.35499
287.7785	MRSN5519	2018-12-24	PRJNA446057	RXTQ01000000	6751302	66.13514
287.7786	MRSN4841	2018-12-24	PRJNA446057	RXTT01000000	6475204	66.25277

287.7787	MRSN5498	2018-12-24	PRJNA446057	RXTS01000000	6821230	66.00017
287.7788	MRSN443463	2018-12-24	PRJNA446057	RXTU01000000	6393241	66.4404
287.7789	MRSN315	2018-12-24	PRJNA446057	RXUI01000000	6696919	66.20516
287.7790	MRSN19711	2018-12-24	PRJNA446057	RXUX01000000	6265327	66.576096
287.7791	MRSN20190	2018-12-24	PRJNA446057	RXUV01000000	6191883	66.59317
287.7792	MRSN9873	2018-12-24	PRJNA446057	RXSY01000000	7146707	65.48819
287.7793	MRSN9718	2018-12-24	PRJNA446057	RXSZ01000000	6861353	66.019516
287.7794	MRSN8141	2018-12-24	PRJNA446057	RXTD01000000	7161718	65.6272
287.7795	MRSN8130	2018-12-24	PRJNA446057	RXTG01000000	6828028	66.09563
287.7796	MRSN6739	2018-12-24	PRJNA446057	RXTI01000000	6298064	66.478065
287.7797	MRSN5508	2018-12-24	PRJNA446057	RXTR01000000	6439859	66.34678
287.7798	MRSN435288	2018-12-24	PRJNA446057	RXTW01000000	6363888	66.22477
287.7799	MRSN321	2018-12-24	PRJNA446057	RXUG01000000	6260875	66.54782
287.7800	MRSN25678	2018-12-24	PRJNA446057	RXUN01000000	6509751	66.21528
287.7801	MRSN18754	2018-12-24	PRJNA446057	RXVH01000000	6270751	66.541534
287.7802	MRSN390231	2018-12-24	PRJNA446057	RXTZ01000000	6149846	66.37745
287.7803	MRSN436311	2018-12-24	PRJNA446057	RXTV01000000	6473360	66.416046
287.7804	MRSN373401	2018-12-24	PRJNA446057	RXUA01000000	6985982	65.97148
287.7805	MRSN3705	2018-12-24	PRJNA446057	RXUB01000000	6419743	66.576324
287.7806	MRSN29192	2018-12-24	PRJNA446057	RXUK01000000	6375161	66.43941
287.7807	MRSN25762	2018-12-24	PRJNA446057	RXUM01000000	6845305	66.2181
287.7808	MRSN2444	2018-12-24	PRJNA446057	RXUP01000000	6843047	66.06668
287.7809	MRSN3587	2018-12-24	PRJNA446057	RXUU01000000	6494873	66.36721
287.7810	MRSN2101	2018-12-24	PRJNA446057	RXUT01000000	6346725	66.376564
287.7811	MRSN23861	2018-12-24	PRJNA446057	RXUQ01000000	7025487	65.94264
287.7812	MRSN20176	2018-12-24	PRJNA446057	RXUW01000000	6717057	65.99654
287.7813	MRSN1948	2018-12-24	PRJNA446057	RXUY01000000	6674435	66.19558
287.7814	MRSN1906	2018-12-24	PRJNA446057	RXVB01000000	7049219	65.7835
287.7815	MRSN1938	2018-12-24	PRJNA446057	RXUZ01000000	7052906	66.11615
287.7816	MRSN1899	2018-12-24	PRJNA446057	RXVD01000000	6573139	66.22041
287.7817	MRSN1902	2018-12-24	PRJNA446057	RXVC01000000	6344431	66.457214
287.7818	MRSN18970	2018-12-24	PRJNA446057	RXVE01000000	6370604	66.51189
287.7819	MRSN17849	2018-12-24	PRJNA446057	RXVK01000000	6738933	66.124664
287.7820	MRSN16744	2018-12-24	PRJNA446057	RXVO01000000	6776946	66.11751
287.7821	MRSN16740	2018-12-24	PRJNA446057	RXVP01000000	6670940	66.08019
287.7822	MRSN1688	2018-12-24	PRJNA446057	RXVM01000000	6606996	66.148384
287.7823	MRSN16383	2018-12-24	PRJNA446057	RXVQ01000000	6302224	66.533325
287.7824	MRSN16345	2018-12-24	PRJNA446057	RXVR01000000	6333273	66.51163
287.7825	MRSN1613	2018-12-24	PRJNA446057	RXVU01000000	6276349	66.510056

287.7826	MRSN1601	2018-12-24	PRJNA446057	RXVW01000000	6411494	66.28397
287.7827	MRSN1583	2018-12-24	PRJNA446057	RXVX01000000	6463376	66.37041
287.7828	MRSN1617	2018-12-24	PRJNA446057	RXVT01000000	6780137	66.19642
287.7829	MRSN1388	2018-12-24	PRJNA446057	RXWC01000000	7171428	65.862366
287.7830	MRSN15678	2018-12-24	PRJNA446057	RXVZ01000000	6780257	66.094574
287.7831	MRSN14981	2018-12-24	PRJNA446057	RXWB01000000	6508561	66.21667
287.7832	MRSN409937	2018-12-24	PRJNA446057	RXTX01000000	6487691	66.33823
287.7833	MRSN351791	2018-12-24	PRJNA446057	RXUE01000000	6812488	65.93882
287.7834	MRSN346179	2018-12-24	PRJNA446057	RXUF01000000	6293739	66.51709
287.7835	MRSN30858	2018-12-24	PRJNA446057	RXUJ01000000	6382502	66.4238
287.7836	MRSN26263	2018-12-24	PRJNA446057	RXUL01000000	6430277	66.320595
287.7837	MRSN1925	2018-12-24	PRJNA446057	RXVA01000000	7130564	65.83213
287.7838	MRSN25623	2018-12-24	PRJNA446057	RXUO01000000	6635285	66.08652
287.7839	MRSN18562	2018-12-24	PRJNA446057	RXVI01000000	6494310	66.368095
287.7840	MRSN16344	2018-12-24	PRJNA446057	RXVS01000000	6318735	66.457306
287.7841	MRSN1612	2018-12-24	PRJNA446057	RXVV01000000	6506983	66.235
287.7842	MRSN1739	2018-12-24	PRJNA446057	RXVL01000000	7031489	65.82428
287.7843	MRSN16847	2018-12-24	PRJNA446057	RXVN01000000	6594082	65.978676
287.7844	MRSN15753	2018-12-24	PRJNA446057	RXVY01000000	6547438	66.18349
287.7845	MRSN1380	2018-12-24	PRJNA446057	RXWD01000000	6339603	66.48981
287.7846	MRSN15566	2018-12-24	PRJNA446057	RXWA01000000	6408848	66.389984
287.7847	MRSN1356	2018-12-24	PRJNA446057	RXWE01000000	6555630	66.19885
287.7848	MRSN12914	2018-12-24	PRJNA446057	RXWH01000000	6756810	65.92381
287.7849	MRSN12283	2018-12-24	PRJNA446057	RXWK01000000	6301640	66.499626
287.7850	MRSN12365	2018-12-24	PRJNA446057	RXWJ01000000	6869499	65.87287
287.7851	MRSN358800	2018-12-24	PRJNA446057	RXUD01000000	6501460	66.35765
287.7852	MRSN11538	2018-12-24	PRJNA446057	RXWN01000000	6528585	66.261604
287.7853	MRSN11278	2018-12-24	PRJNA446057	RXWS01000000	6893110	66.043076
287.7854	MRSN11281	2018-12-24	PRJNA446057	RXWR01000000	6323478	66.49251
287.7855	MRSN13488	2018-12-24	PRJNA446057	RXWF01000000	5912421	66.50286
287.7856	MRSN1344	2018-12-24	PRJNA446057	RXWG01000000	6232278	66.494484
287.7857	MRSN12368	2018-12-24	PRJNA446057	RXWI01000000	6340159	66.43633
287.7858	MRSN11285	2018-12-24	PRJNA446057	RXWQ01000000	6650291	66.16359
287.7859	MRSN12282	2018-12-24	PRJNA446057	RXWL01000000	6876949	65.96845
287.7860	MRSN11536	2018-12-24	PRJNA446057	RXWO01000000	6935556	65.91659
287.7861	MRSN11286	2018-12-24	PRJNA446057	RXWP01000000	6711313	65.911575
287.7862	MRSN11976	2018-12-24	PRJNA446057	RXWM01000000	6991711	65.90075
287.8027	MRSN401528	2018-12-24	PRJNA446057	RXTY01000000	6486368	66.14036
287.8028	MRSN369569	2018-12-24	PRJNA446057	RXUC01000000	6316238	66.518555

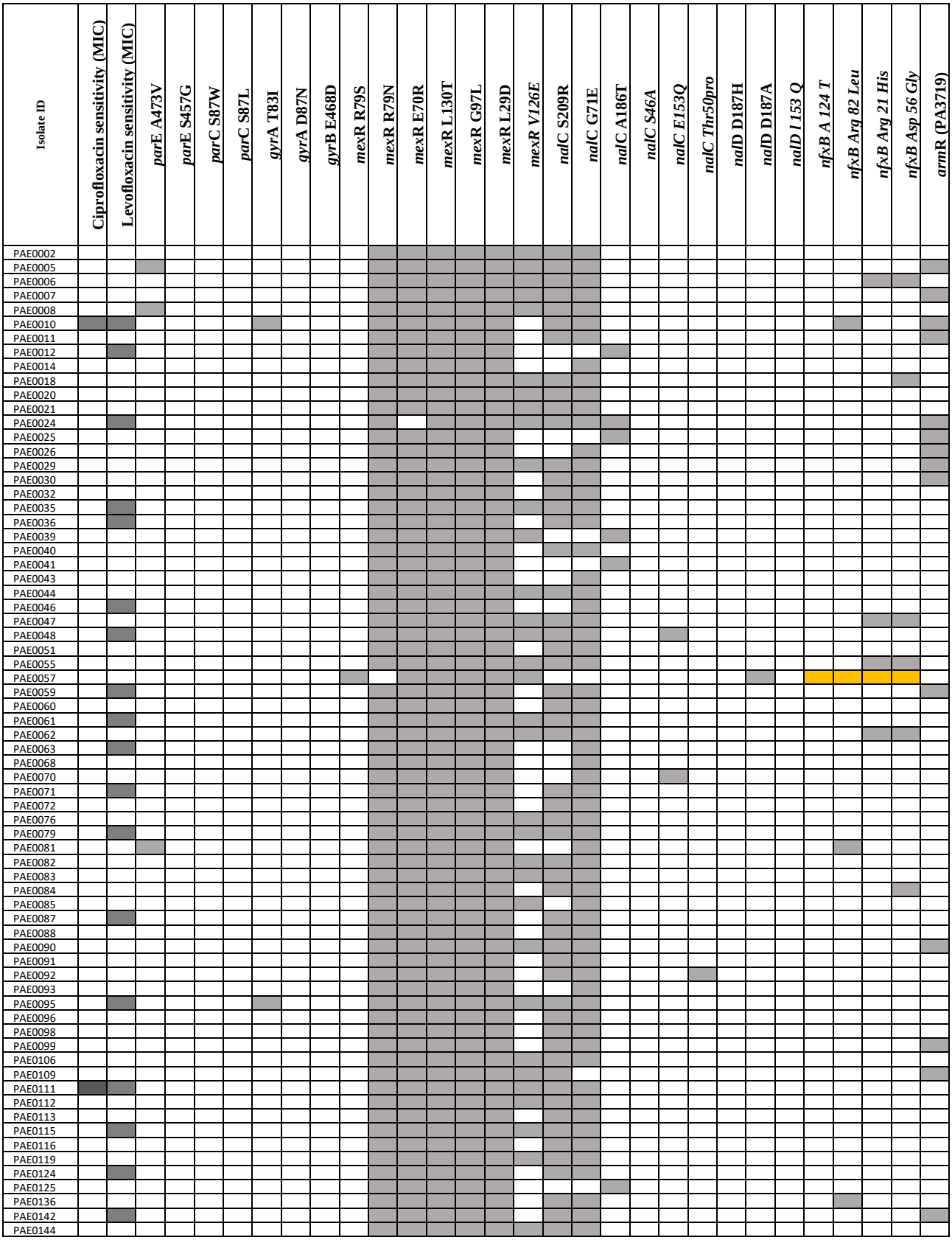
287.8029	MRSN2144	2018-12-24	PRJNA446057	RXUR01000000	7032738	65.79075
287.8030	MRSN18855	2018-12-24	PRJNA446057	RXVF01000000	6296953	66.489555
287.8031	MRSN2108	2018-12-24	PRJNA446057	RXUS01000000	6394452	66.39729
287.8032	MRSN18803	2018-12-24	PRJNA446057	RXVG01000000	6191380	66.5096
287.8033	MRSN317	2018-12-24	PRJNA446057	RXUH01000000	6349621	66.47357
287.8034	MRSN18560	2018-12-24	PRJNA446057	RXVJ01000000	6268481	66.5069
287.846	AZPAE15072	2014-12-04	PRJNA264310	JTMA00000000	6437786	66.3
287.847	AZPAE15071	2014-12-04	PRJNA264310	JTMB00000000	7403124	65.72
287.848	AZPAE15070	2014-12-04	PRJNA264310	JTMC00000000	7164330	65.83
287.849	AZPAE15069	2014-12-04	PRJNA264310	JTMD00000000	6980745	65.93
287.850	AZPAE15068	2014-12-04	PRJNA264310	JTME00000000	6856096	66.14
287.851	AZPAE15067	2014-12-04	PRJNA264310	JTMF00000000	6651822	66.15
287.852	AZPAE15066	2014-12-04	PRJNA264310	JTMG00000000	6959180	65.82
287.853	AZPAE15065	2014-12-04	PRJNA264310	JTMH00000000	6184461	66.47
287.854	AZPAE15064	2014-12-04	PRJNA264310	JTMI00000000	6687065	66.19
287.855	AZPAE15063	2014-12-04	PRJNA264310	JTMJ00000000	6768150	66.03
287.856	AZPAE15062	2014-12-04	PRJNA264310	JTMK00000000	6481258	66.31
287.857	AZPAE15061	2014-12-04	PRJNA264310	JTML00000000	6832661	65.96
287.858	AZPAE15060	2014-12-04	PRJNA264310	JTMM00000000	6716256	66.11
287.859	AZPAE15059	2014-12-04	PRJNA264310	JTMN00000000	6657813	66.25
287.860	AZPAE15058	2014-12-04	PRJNA264310	JTMO00000000	6567883	66.22
287.861	AZPAE15057	2014-12-04	PRJNA264310	JTMP00000000	6451298	66.43
287.862	AZPAE15056	2014-12-04	PRJNA264310	JTMQ00000000	6287511	66.49
287.863	AZPAE15055	2014-12-04	PRJNA264310	JTMR00000000	6360589	66.37
287.864	AZPAE15054	2014-12-04	PRJNA264310	JTMS00000000	6647129	65.96
287.865	AZPAE15053	2014-12-04	PRJNA264310	JTMT00000000	6522359	66.25
287.866	AZPAE15052	2014-12-04	PRJNA264310	JTMU00000000	6339616	66.38
287.867	AZPAE15051	2014-12-04	PRJNA264310	JTMV00000000	6814130	66.09
287.868	AZPAE15050	2014-12-04	PRJNA264310	JTMW00000000	6901290	65.78
287.869	AZPAE15049	2014-12-04	PRJNA264310	JTMX00000000	6259893	66.46
287.870	AZPAE15048	2014-12-04	PRJNA264310	JTMY00000000	6924072	65.73
287.871	AZPAE15047	2014-12-04	PRJNA264310	JTMZ00000000	6898282	66.03
287.872	AZPAE15046	2014-12-04	PRJNA264310	JTNA00000000	6345609	66.45
287.873	AZPAE15045	2014-12-04	PRJNA264310	JTNB00000000	6703034	66
287.874	AZPAE15044	2014-12-04	PRJNA264310	JTNC00000000	6555614	66.21
287.875	AZPAE15043	2014-12-04	PRJNA264310	JTND00000000	6978967	65.98
287.876	AZPAE15042	2014-12-04	PRJNA264310	JTNE00000000	6640508	66.27
287.877	AZPAE15041	2014-12-04	PRJNA264310	JTNF00000000	6739630	66.02
287.878	AZPAE15040	2014-12-04	PRJNA264310	JTNG00000000	6601216	66.23

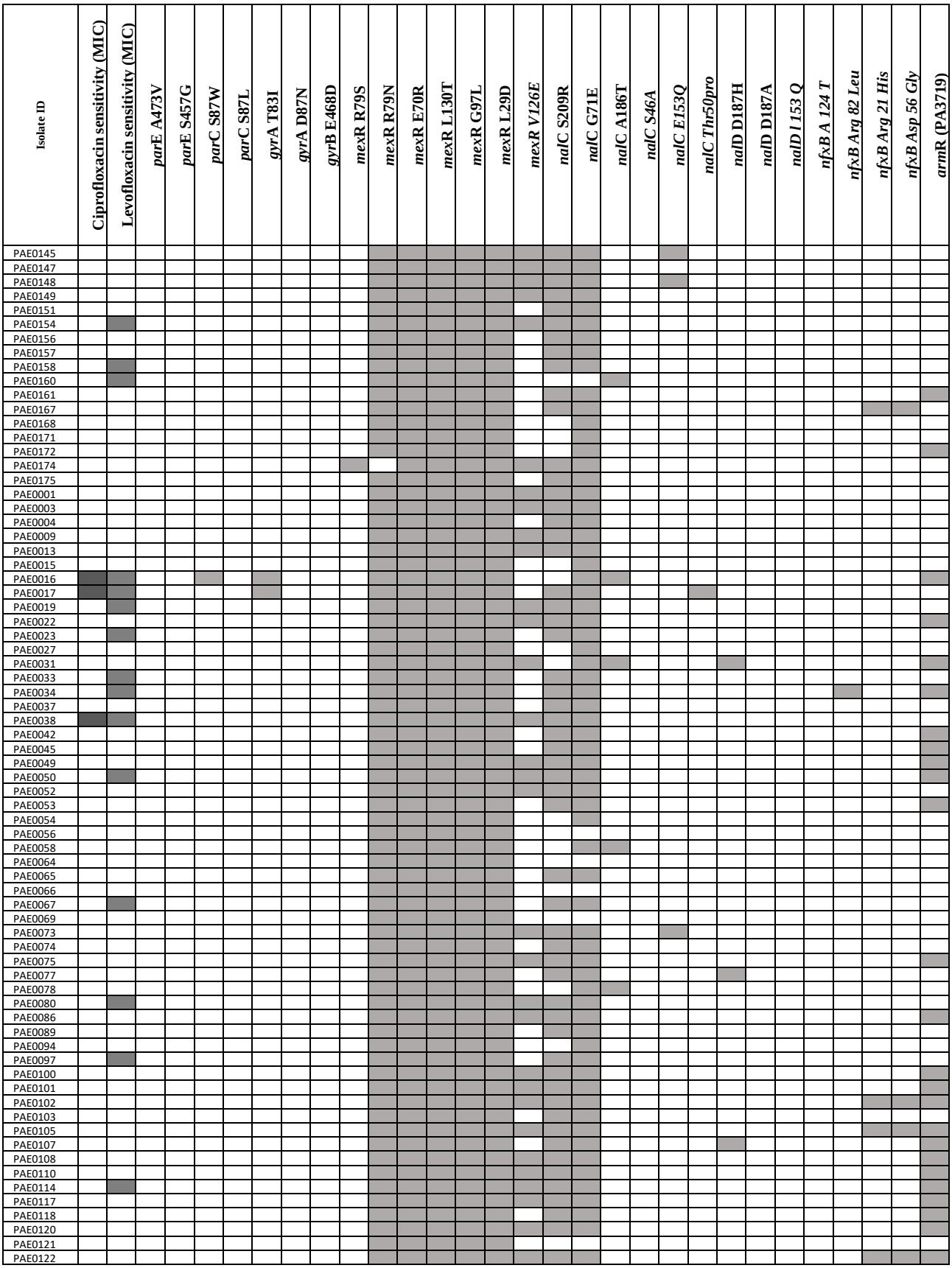
287.879	AZPAE15039	2014-12-04	PRJNA264310	JTNH00000000	6697373	66.07
287.880	AZPAE15038	2014-12-04	PRJNA264310	JTNI00000000	6261442	66.49
287.881	AZPAE15037	2014-12-04	PRJNA264310	JTNJ00000000	7200442	65.83
287.882	AZPAE15036	2014-12-04	PRJNA264310	JTNK00000000	6412424	66.31
287.883	AZPAE15035	2014-12-04	PRJNA264310	JTNL00000000	6673047	66.09
287.884	AZPAE15034	2014-12-04	PRJNA264310	JTNM00000000	6794620	66.14
287.885	AZPAE15033	2014-12-04	PRJNA264310	JTNN00000000	6087226	66.4
287.886	AZPAE15032	2014-12-04	PRJNA264310	JTNO00000000	6484573	66.2
287.887	AZPAE15031	2014-12-04	PRJNA264310	JTNP00000000	6894718	65.98
287.888	AZPAE15030	2014-12-04	PRJNA264310	JTNQ00000000	6608631	66.18
287.889	AZPAE15029	2014-12-04	PRJNA264310	JTNR00000000	6860916	66.03
287.890	AZPAE15028	2014-12-04	PRJNA264310	JTNS00000000	6834546	65.7
287.891	AZPAE15027	2014-12-04	PRJNA264310	JTNT00000000	6438293	66.36
287.892	AZPAE15026	2014-12-04	PRJNA264310	JTNU00000000	6358197	66.41
287.893	AZPAE15025	2014-12-04	PRJNA264310	JTNV00000000	6726954	66
287.894	AZPAE15024	2014-12-04	PRJNA264310	JTNW00000000	6776478	66
287.895	AZPAE15023	2014-12-04	PRJNA264310	JTNX00000000	6351355	66.4
287.896	AZPAE15022	2014-12-04	PRJNA264310	JTNY00000000	6219362	66.56
287.897	AZPAE15021	2014-12-04	PRJNA264310	JTNZ00000000	6686424	66.05
287.898	AZPAE15020	2014-12-04	PRJNA264310	JTOA00000000	7163187	65.8
287.899	AZPAE15019	2014-12-04	PRJNA264310	JTOB00000000	6717129	66.06
287.900	AZPAE15018	2014-12-04	PRJNA264310	JTOC00000000	6933997	66.06
287.901	AZPAE15017	2014-12-04	PRJNA264310	JTOD00000000	6594600	66.09
287.902	AZPAE15016	2014-12-04	PRJNA264310	JTOE00000000	6756824	65.84
287.903	AZPAE15015	2014-12-04	PRJNA264310	JTOF00000000	6804229	66.08
287.904	AZPAE15014	2014-12-04	PRJNA264310	JTOG00000000	6400436	66.42
287.905	AZPAE15013	2014-12-04	PRJNA264310	JTOH00000000	6358832	66.41
287.906	AZPAE15012	2014-12-04	PRJNA264310	JTOI00000000	6462525	66.32
287.907	AZPAE15011	2014-12-04	PRJNA264310	JTOJ00000000	6590340	65.85
287.908	AZPAE15010	2014-12-04	PRJNA264310	JTOK00000000	7014236	65.9
287.909	AZPAE15009	2014-12-04	PRJNA264310	JTOL00000000	6958019	65.85
287.910	AZPAE15008	2014-12-04	PRJNA264310	JTOM00000000	6345734	66.18
287.911	AZPAE15007	2014-12-04	PRJNA264310	JTON00000000	6814810	65.79
287.912	AZPAE15006	2014-12-04	PRJNA264310	JTOO00000000	6832413	65.95
287.913	AZPAE15005	2014-12-04	PRJNA264310	JTOP00000000	6889490	66.07
287.914	AZPAE15004	2014-12-04	PRJNA264310	JTOQ00000000	6345076	66.39
287.915	AZPAE15003	2014-12-04	PRJNA264310	JTOR00000000	6882469	66.11
287.916	AZPAE15002	2014-12-04	PRJNA264310	JTOS00000000	6802969	65.79
287.917	AZPAE15001	2014-12-04	PRJNA264310	JTOT00000000	6582180	66.3

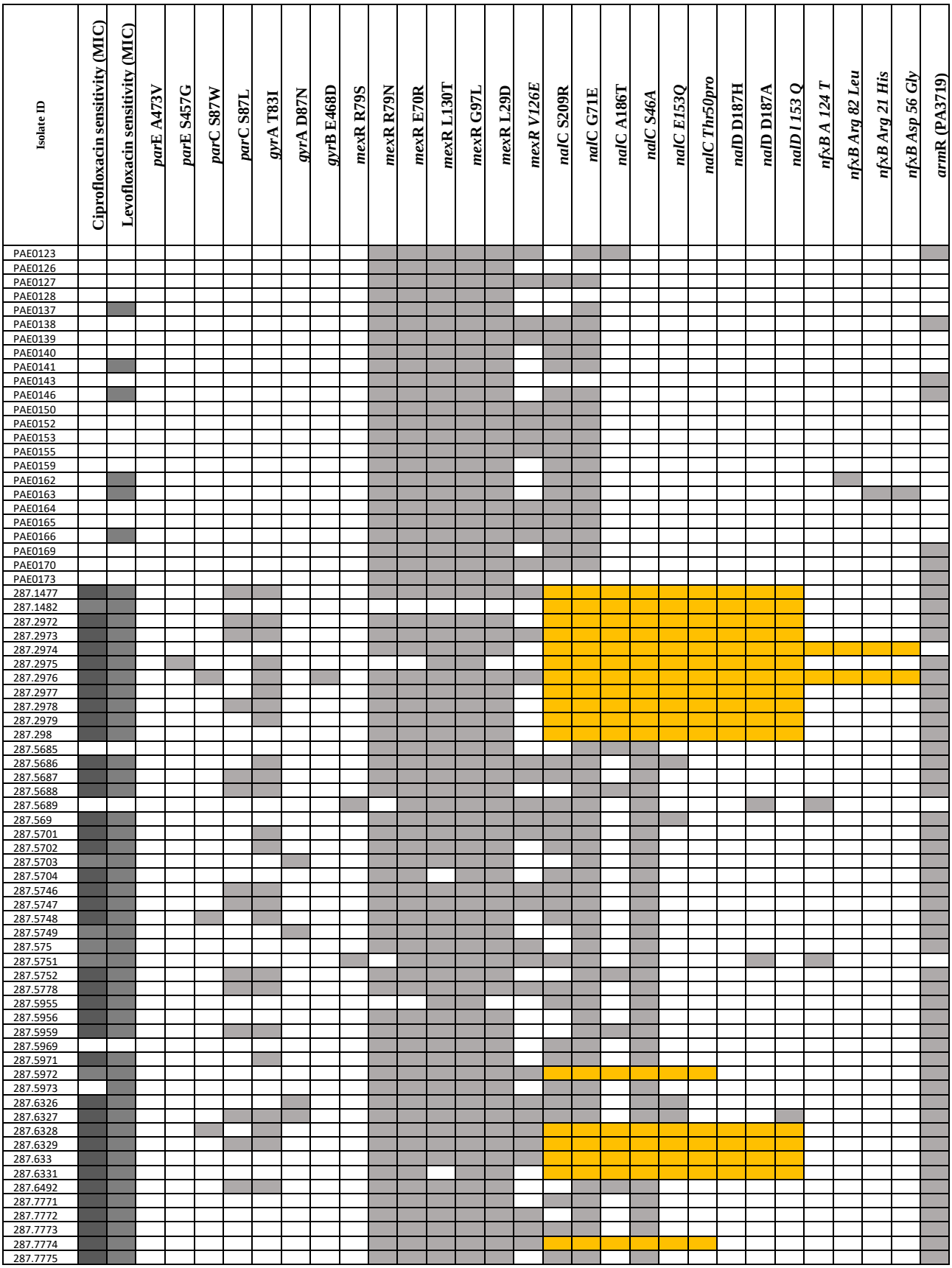
287.918	AZPAE15000	2014-12-04	PRJNA264310	JTOU00000000	6858205	66
287.919	AZPAE14999	2014-12-04	PRJNA264310	JTOV00000000	6427887	66.3
287.920	AZPAE14998	2014-12-04	PRJNA264310	JTOW00000000	6712745	66.17
287.921	AZPAE14997	2014-12-04	PRJNA264310	JTOX00000000	6715214	65.98
287.922	AZPAE14996	2014-12-04	PRJNA264310	JTOY00000000	6459808	66.34
287.923	AZPAE14995	2014-12-04	PRJNA264310	JTOZ00000000	6389451	66.32
287.924	AZPAE14994	2014-12-04	PRJNA264310	JTPA00000000	6476787	66.07
287.925	AZPAE14993	2014-12-04	PRJNA264310	JTPB00000000	6914857	65.97
287.926	AZPAE14992	2014-12-04	PRJNA264310	JTPC00000000	6510002	66.3
287.927	AZPAE14991	2014-12-04	PRJNA264310	JTPD00000000	6467607	66.31
287.928	AZPAE14990	2014-12-04	PRJNA264310	JTPE00000000	6809137	66.09
287.929	AZPAE14989	2014-12-04	PRJNA264310	JTPF00000000	6549370	66.19
287.930	AZPAE14988	2014-12-04	PRJNA264310	JTPG00000000	6387737	66.41
287.931	AZPAE14987	2014-12-04	PRJNA264310	JTPH00000000	6855155	66.02
287.932	AZPAE14986	2014-12-04	PRJNA264310	JTPI00000000	6471510	66.35
287.933	AZPAE14985	2014-12-04	PRJNA264310	JTPJ00000000	6781336	66.05
287.934	AZPAE14984	2014-12-04	PRJNA264310	JTPK00000000	6809111	66.07
287.951	AZPAE14983	2014-12-04	PRJNA264310	JTPL00000000	7082650	65.77
287.952	AZPAE14982	2014-12-04	PRJNA264310	JTPM00000000	6455689	66.31
287.953	AZPAE14981	2014-12-04	PRJNA264310	JTPN00000000	6276540	66.48
287.954	AZPAE14980	2014-12-04	PRJNA264310	JTPO00000000	6768315	66.16
287.955	AZPAE14979	2014-12-04	PRJNA264310	JTPP00000000	6751299	66.03
287.956	AZPAE14978	2014-12-04	PRJNA264310	JTPQ00000000	7130480	65.74
287.957	AZPAE14977	2014-12-04	PRJNA264310	JTPR00000000	6336142	66.39
287.958	AZPAE14976	2014-12-04	PRJNA264310	JTPS00000000	6631621	66.06
287.959	AZPAE14975	2014-12-04	PRJNA264310	JTPT00000000	6369755	66.42
287.960	AZPAE14974	2014-12-04	PRJNA264310	JTPU00000000	7085926	65.75
287.961	AZPAE14973	2014-12-04	PRJNA264310	JTPV00000000	7183565	65.78
287.962	AZPAE14972	2014-12-04	PRJNA264310	JTPW00000000	6399923	66.38
287.963	AZPAE14971	2014-12-04	PRJNA264310	JTPX00000000	6203904	66.44
287.964	AZPAE14970	2014-12-04	PRJNA264310	JTPY00000000	6920457	66.05
287.965	AZPAE14969	2014-12-04	PRJNA264310	JTPZ00000000	6832241	66
287.966	AZPAE14968	2014-12-04	PRJNA264310	JTQA00000000	6897223	66
287.967	AZPAE14967	2014-12-04	PRJNA264310	JTQB00000000	6567690	66.17
287.968	AZPAE14965	2014-12-04	PRJNA264310	JTQC00000000	7042523	65.91
287.969	AZPAE14964	2014-12-04	PRJNA264310	JTQD00000000	7039582	65.91
287.970	AZPAE14963	2014-12-04	PRJNA264310	JTQE00000000	6380431	66.44
287.971	AZPAE14962	2014-12-04	PRJNA264310	JTQF00000000	6963110	65.83
287.972	AZPAE14961	2014-12-04	PRJNA264310	JTQG00000000	6441305	66.34

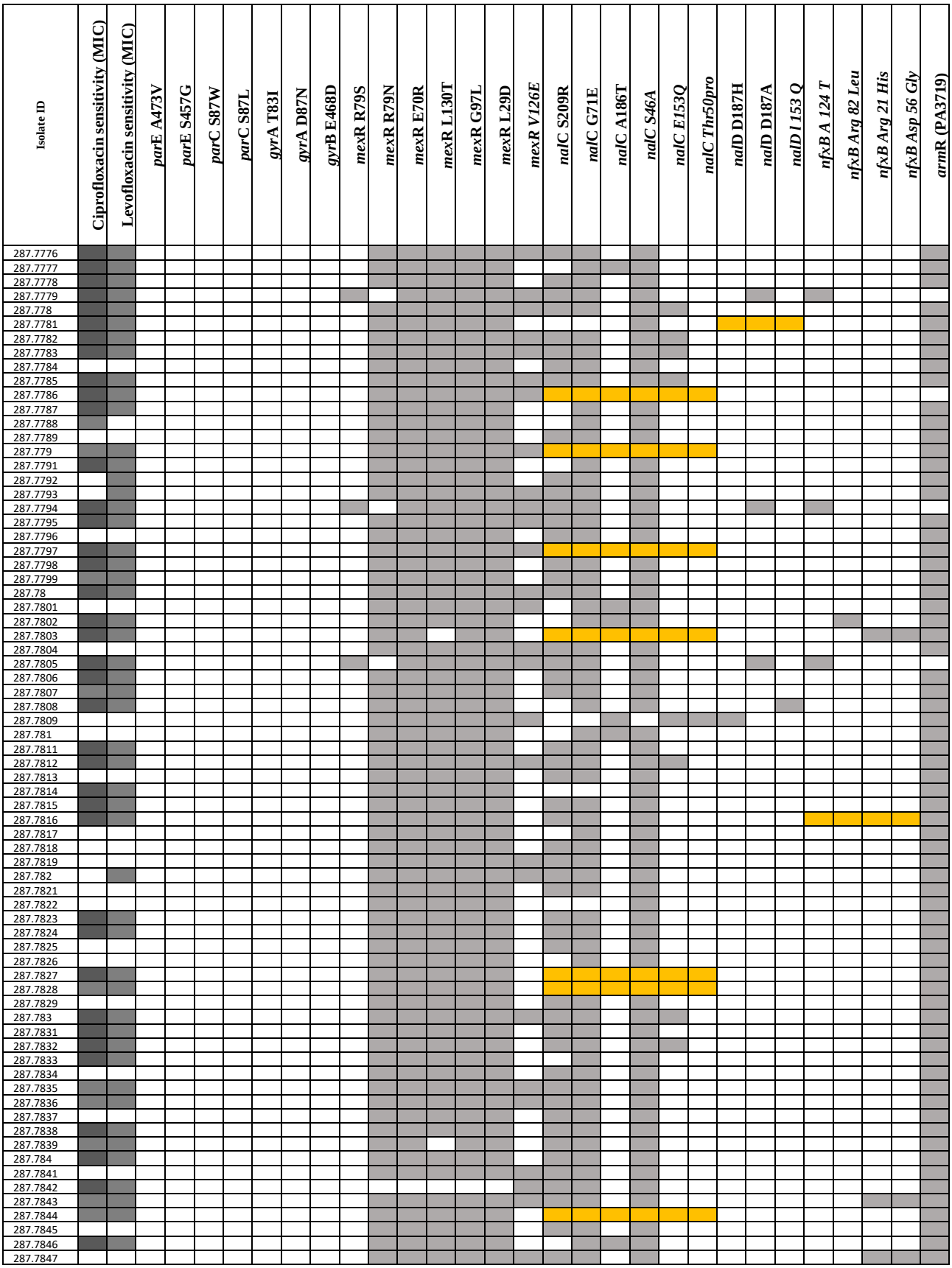
287.973	AZPAE14960	2014-12-04	PRJNA264310	JTQH00000000	6370668	66.45
287.974	AZPAE14959	2014-12-04	PRJNA264310	JTQI00000000	6743682	66.15
287.975	AZPAE14958	2014-12-04	PRJNA264310	JTQJ00000000	7130390	65.74
287.976	AZPAE14957	2014-12-04	PRJNA264310	JTQK00000000	6621214	66.11
287.977	AZPAE14956	2014-12-04	PRJNA264310	JTQL00000000	7110106	65.36
287.978	AZPAE14955	2014-12-04	PRJNA264310	JTQM00000000	6375587	66.42
287.979	AZPAE14954	2014-12-04	PRJNA264310	JTQN00000000	6368200	66.43
287.980	AZPAE14953	2014-12-04	PRJNA264310	JTQO00000000	6809063	66.15
287.981	AZPAE14952	2014-12-04	PRJNA264310	JTQP00000000	6365190	66.39
287.982	AZPAE14951	2014-12-04	PRJNA264310	JTQQ00000000	6876017	65.85
287.983	AZPAE14950	2014-12-04	PRJNA264310	JTQR00000000	6332797	66.24
287.984	AZPAE14949	2014-12-04	PRJNA264310	JTQS00000000	6899528	66.05
287.985	AZPAE14948	2014-12-04	PRJNA264310	JTQT00000000	6828755	65.86
287.986	AZPAE14947	2014-12-04	PRJNA264310	JTQU00000000	6482138	66.33
287.987	AZPAE14946	2014-12-04	PRJNA264310	JTQV00000000	6592636	66.29
287.988	AZPAE14945	2014-12-04	PRJNA264310	JTQW00000000	6866953	66.05
287.989	AZPAE14944	2014-12-04	PRJNA264310	JTQX00000000	6824257	66.11
287.990	AZPAE14943	2014-12-04	PRJNA264310	JTQY00000000	6250531	66.53
287.991	AZPAE14942	2014-12-04	PRJNA264310	JTQZ00000000	6787557	66.05
287.992	AZPAE14941	2014-12-04	PRJNA264310	JTRA00000000	6881480	65.98
287.993	AZPAE14940	2014-12-04	PRJNA264310	JTRB00000000	6929735	65.97
287.994	AZPAE14939	2014-12-04	PRJNA264310	JTRC00000000	6351078	66.39
287.995	AZPAE14938	2014-12-04	PRJNA264310	JTRD00000000	6900154	66.12
287.996	AZPAE14937	2014-12-04	PRJNA264310	JTRE00000000	6892469	65.86
287.997	AZPAE14936	2014-12-04	PRJNA264310	JTRF00000000	6265859	66.33
287.998	AZPAE14935	2014-12-04	PRJNA264310	JTRG00000000	6764634	66.22
287.999	AZPAE14934	2014-12-04	PRJNA264310	JTRH00000000	6346157	66.37

Supplementary Table 2A: Distribution of known quinolone resistance mutations (ciprofloxacin and levofloxacin)




Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
parE A473V
parE S457G
parC S87W
parC S87L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nalC S209R
nalC G71E
nalC A186T
nalC S46A
nalC E153Q
nalC Thr50pro
nalD D187H
nalD D187A
nalD I 153 Q
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
armR (PA3719)
PAE0145
PAE0147
PAE0148
PAE0149
PAE0151
PAE0154
PAE0156
PAE0157
PAE0158
PAE0160
PAE0161
PAE0167
PAE0168
PAE0171
PAE0172
PAE0174
PAE0175
PAE0001
PAE0003
PAE0004
PAE0009
PAE0013
PAE0015
PAE0016
PAE0017
PAE0019
PAE0022
PAE0023
PAE0027
PAE0031
PAE0033
PAE0034
PAE0037
PAE0038
PAE0042
PAE0045
PAE0049
PAE0050
PAE0052
PAE0053
PAE0054
PAE0056
PAE0058
PAE0064
PAE0065
PAE0066
PAE0067
PAE0069
PAE0073
PAE0074
PAE0075
PAE0077
PAE0078
PAE0080
PAE0086
PAE0089
PAE0094
PAE0097
PAE0100
PAE0101
PAE0102
PAE0103
PAE0105
PAE0107
PAE0108
PAE0110
PAE0114
PAE0117
PAE0118
PAE0120
PAE0121
PAE0122


Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
parE A473V
parE S457G
parC S87W
parC S87L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nalC S209R
nalC G71E
nalC A186T
nalC S46A
nalC E153Q
nalC Thr50pro
nalD D187H
nalD D187A
nalD I 153 Q
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
armR (PA3719)
PAE0123
PAE0126
PAE0127
PAE0128
PAE0137
PAE0138
PAE0139
PAE0140
PAE0141
PAE0143
PAE0146
PAE0150
PAE0152
PAE0153
PAE0155
PAE0159
PAE0162
PAE0163
PAE0164
PAE0165
PAE0166
PAE0169
PAE0170
PAE0173
287.1477
287.1482
287.2972
287.2973
287.2974
287.2975
287.2976
287.2977
287.2978
287.2979
287.298
287.5685
287.5686
287.5687
287.5688
287.5689
287.569
287.5701
287.5702
287.5703
287.5704
287.5746
287.5747
287.5748
287.5749
287.575
287.5751
287.5752
287.5778
287.5955
287.5956
287.5959
287.5969
287.5971
287.5972
287.5973
287.6326
287.6327
287.6328
287.6329
287.633
287.6331
287.6492
287.7771
287.7772
287.7773
287.7774
287.7775

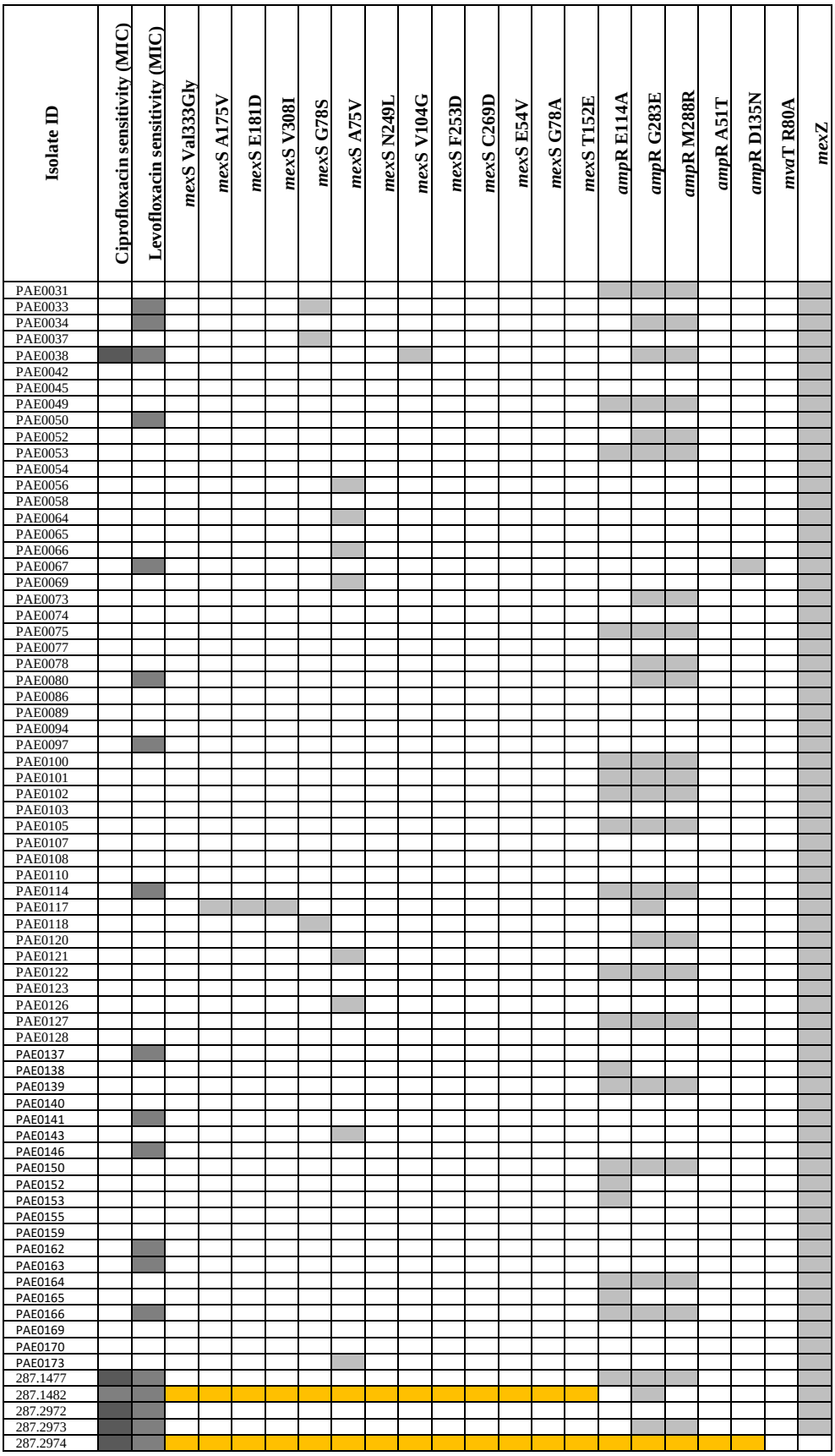

Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
parE A473V
parE S457G
parC S87W
parC S87L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nalC S209R
nalC G71E
nalC A186T
nalC S46A
nalC E153Q
nalC Thr50pro
nalD D187H
nalD D187A
nalD I 153 Q
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
armR (PA3719)
287.7776
287.7777
287.7778
287.7779
287.778
287.7781
287.7782
287.7783
287.7784
287.7785
287.7786
287.7787
287.7788
287.7789
287.779
287.7791
287.7792
287.7793
287.7794
287.7795
287.7796
287.7797
287.7798
287.7799
287.78
287.7801
287.7802
287.7803
287.7804
287.7805
287.7806
287.7807
287.7808
287.7809
287.781
287.7811
287.7812
287.7813
287.7814
287.7815
287.7816
287.7817
287.7818
287.7819
287.782
287.7821
287.7822
287.7823
287.7824
287.7825
287.7826
287.7827
287.7828
287.7829
287.783
287.7831
287.7832
287.7833
287.7834
287.7835
287.7836
287.7837
287.7838
287.7839
287.784
287.7841
287.7842
287.7843
287.7844
287.7845
287.7846
287.7847

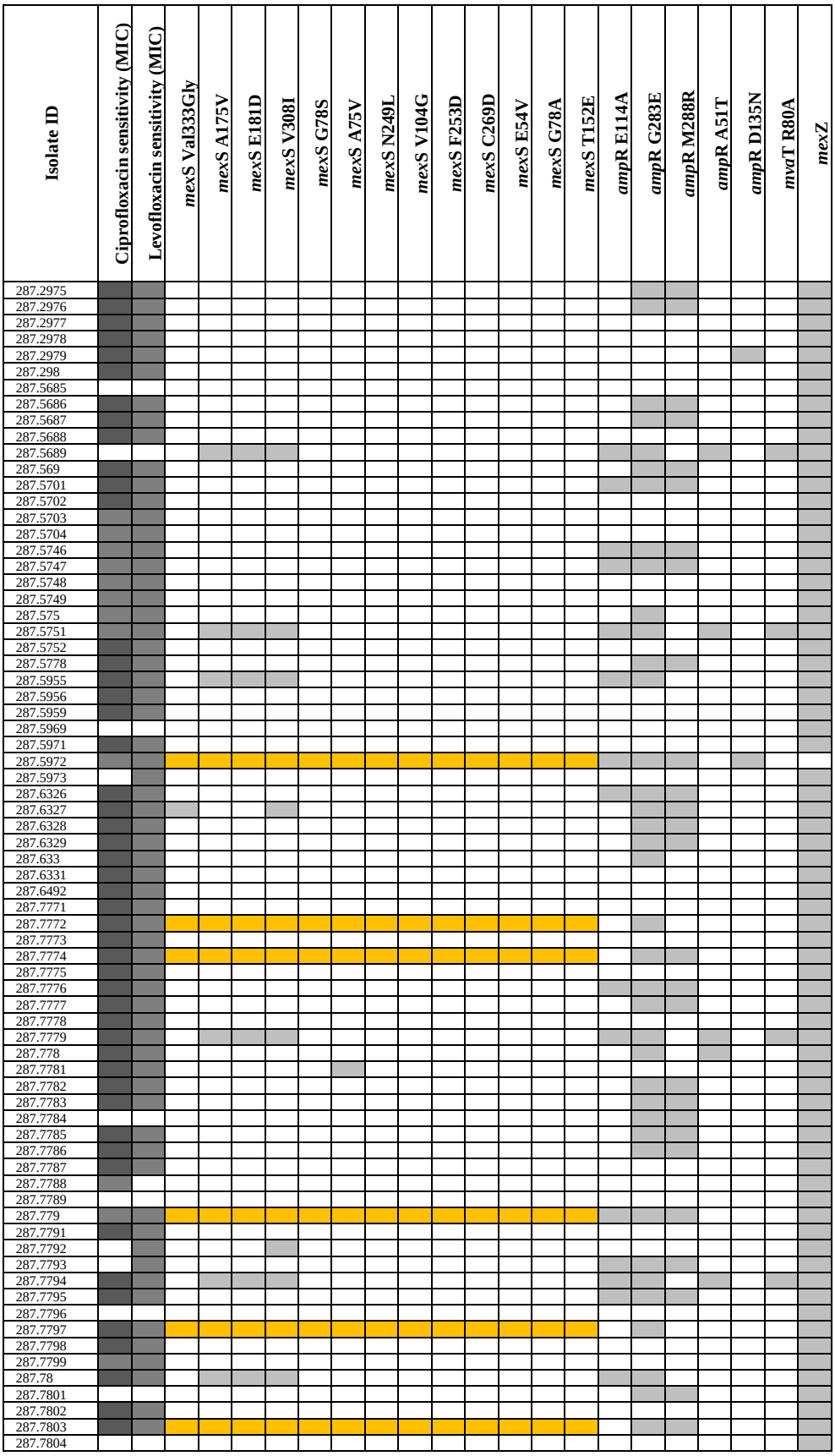
Isolate ID	Ciprofloxacin sensitivity (MIC)	Levofloxacin sensitivity (MIC)	*par*E A473V	*par*E S457G	*par*C S87W	*par*C S87L	*gyr*A T83I	*gyr*A D87N	*gyr*B E468D	*mex*R R79S	*mex*R R79N	*mex*R E70R	*mex*R L130T	*mex*R G97L	*mex*R L29D	*mexR V126E*	*nal*C S209R	*nal*C G71E	*nal*C A186T	*nalC S46A*	*nalC E153Q*	*nalC Thr50pro*	*nal*D D187H	*nal*D D187A	*nalD I 153 Q*	*nfxB A 124 T*	*nfxB Arg 82 Leu*	*nfxB Arg 21 His*	*nfxB Asp 56 Gly*	*arm*R (PA3719)
287.7848																														
287.7849																														
287.785																														
287.7851																														
287.7852																														
287.7853																														
287.7854																														
287.7855																														
287.7856																														
287.7857																														
287.7858																														
287.7859																														
287.786																														
287.7861																														
287.7862																														
287.8027																														
287.8028																														
287.8029																														
287.803																														
287.8031																														
287.8032																														
287.8033																														
287.8034																														

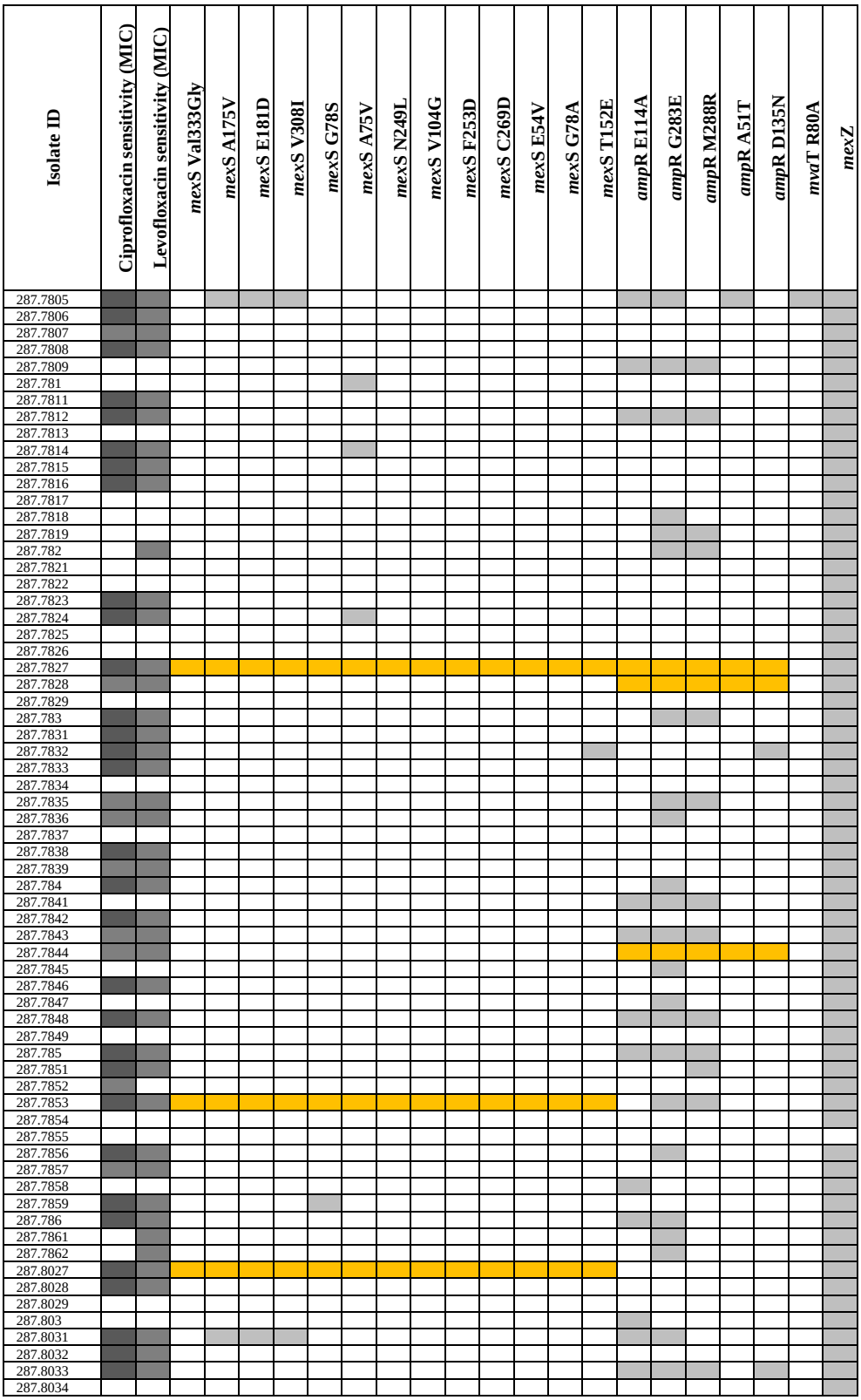
Supplementary Table 2B: Distribution of known quinolone resistance mutations (ciprofloxacin and levofloxacin)

Isolate ID	Ciprofloxacin sensitivity (MIC)	Levofloxacin sensitivity (MIC)	*mex*S Val333Gly	*mex*S A175V	*mex*S E181D	*mex*S V308I	*mex*S G78S	*mex*S A75V	*mex*S N249L	*mex*S V104G	*mex*S F253D	*mex*S C269D	*mex*S E54V	*mex*S G78A	*mex*S T152E	*ampR* E114A	*ampR* G283E	*ampR* M288R	*ampR* A51T	*ampR* D135N	*mvaT* R80A	*mexZ*
PAE0002																						
PAE0005																						
PAE0006																						
PAE0007																						
PAE0008																						
PAE0010																						
PAE0011																						
PAE0012																						
PAE0014																						
PAE0018																						
PAE0020																						
PAE0021																						
PAE0024																						
PAE0025																						
PAE0026																						
PAE0029																						
PAE0030																						
PAE0032																						
PAE0035																						
PAE0036																						
PAE0039																						
PAE0040																						
PAE0041																						
PAE0043																						
PAE0044																						
PAE0046																						
PAE0047																						

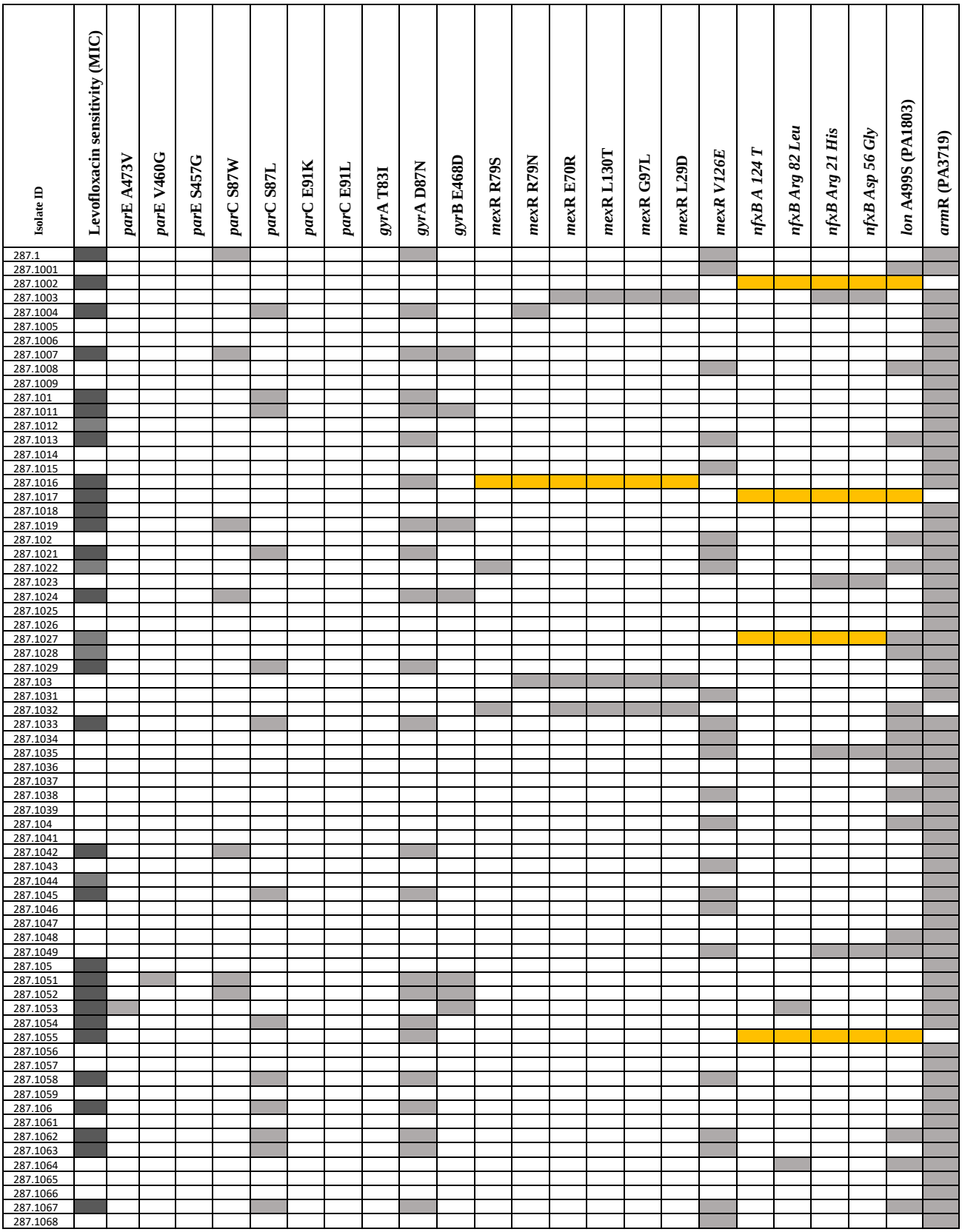
Isolate ID	Ciprofloxacin sensitivity (MIC)	Levofloxacin sensitivity (MIC)	*mex*S Val333Gly	*mex*S A175V	*mex*S E181D	*mex*S V308I	*mex*S G78S	*mex*S A75V	*mex*S N249L	*mex*S V104G	*mex*S F253D	*mex*S C269D	*mex*S E54V	*mex*S G78A	*mex*S T152E	*ampR* E114A	*ampR* G283E	*ampR* M288R	*ampR* A51T	*ampR* D135N	*mvaT* R80A	*mexZ*
PAE0048																						
PAE0051																						
PAE0055																						
PAE0057																						
PAE0059																						
PAE0060																						
PAE0061																						
PAE0062																						
PAE0063																						
PAE0068																						
PAE0070																						
PAE0071																						
PAE0072																						
PAE0076																						
PAE0079																						
PAE0081																						
PAE0082																						
PAE0083																						
PAE0084																						
PAE0085																						
PAE0087																						
PAE0088																						
PAE0090																						
PAE0091																						
PAE0092																						
PAE0093																						
PAE0095																						
PAE0096																						
PAE0098																						
PAE0099																						
PAE0106																						
PAE0109																						
PAE0111																						
PAE0112																						
PAE0113																						
PAE0115																						
PAE0116																						
PAE0119																						
PAE0124																						
PAE0125																						
PAE0136																						
PAE0142																						
PAE0144																						
PAE0145																						
PAE0147																						
PAE0148																						
PAE0149																						
PAE0151																						
PAE0154																						
PAE0156																						
PAE0157																						
PAE0158																						
PAE0160																						
PAE0161																						
PAE0167																						
PAE0168																						
PAE0171																						
PAE0172																						
PAE0174																						
PAE0175																						
PAE0001																						
PAE0003																						
PAE0004																						
PAE0009																						
PAE0013																						
PAE0015																						
PAE0016																						
PAE0017																						
PAE0019																						
PAE0022																						
PAE0023																						
PAE0027																						

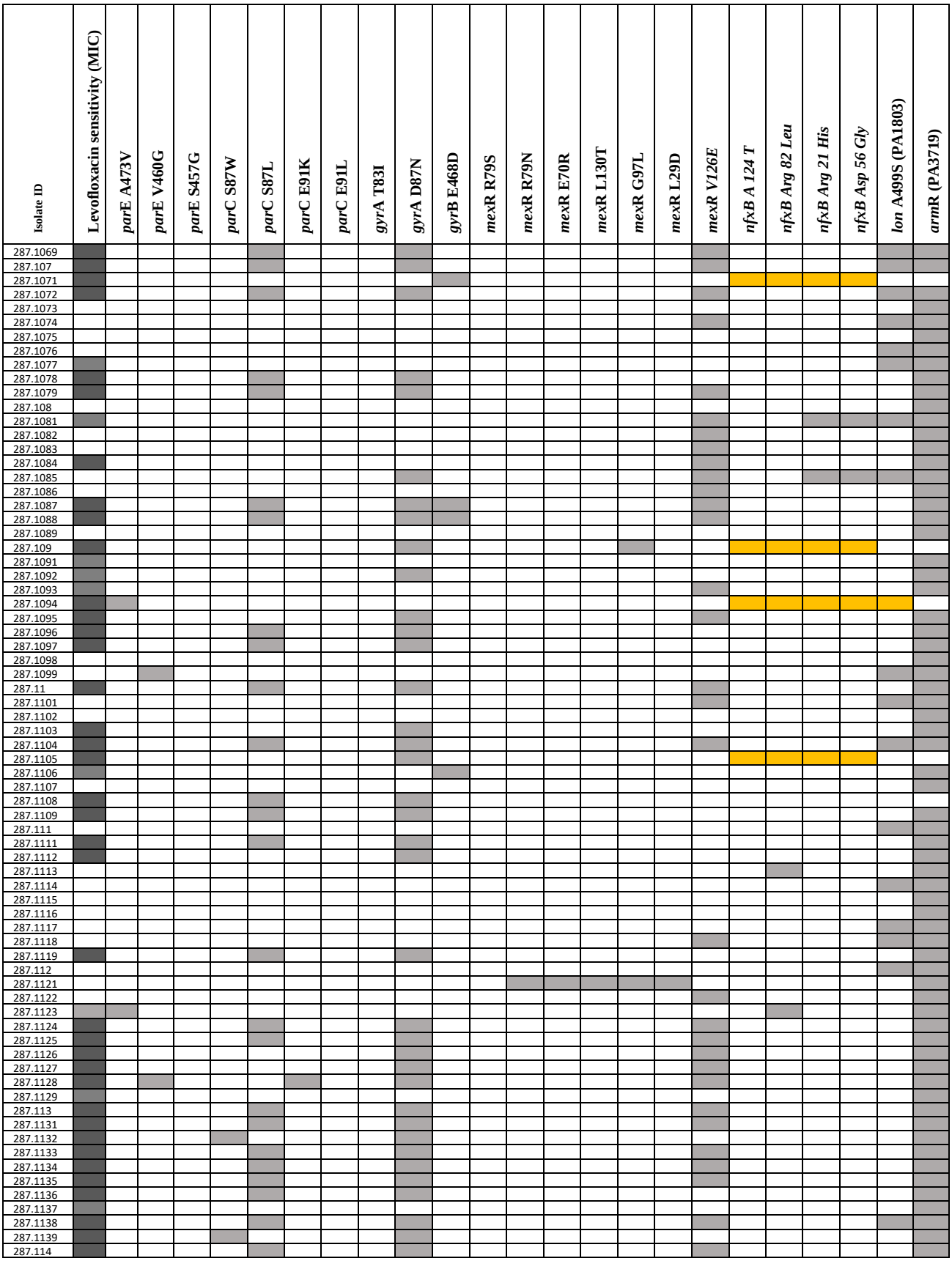

Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
PAE0031
PAE0033
PAE0034
PAE0037
PAE0038
PAE0042
PAE0045
PAE0049
PAE0050
PAE0052
PAE0053
PAE0054
PAE0056
PAE0058
PAE0064
PAE0065
PAE0066
PAE0067
PAE0069
PAE0073
PAE0074
PAE0075
PAE0077
PAE0078
PAE0080
PAE0086
PAE0089
PAE0094
PAE0097
PAE0100
PAE0101
PAE0102
PAE0103
PAE0105
PAE0107
PAE0108
PAE0110
PAE0114
PAE0117
PAE0118
PAE0120
PAE0121
PAE0122
PAE0123
PAE0126
PAE0127
PAE0128
PAE0137
PAE0138
PAE0139
PAE0140
PAE0141
PAE0143
PAE0146
PAE0150
PAE0152
PAE0153
PAE0155
PAE0159
PAE0162
PAE0163
PAE0164
PAE0165
PAE0166
PAE0169
PAE0170
PAE0173
287.1477
287.1482
287.2972
287.2973
287.2974

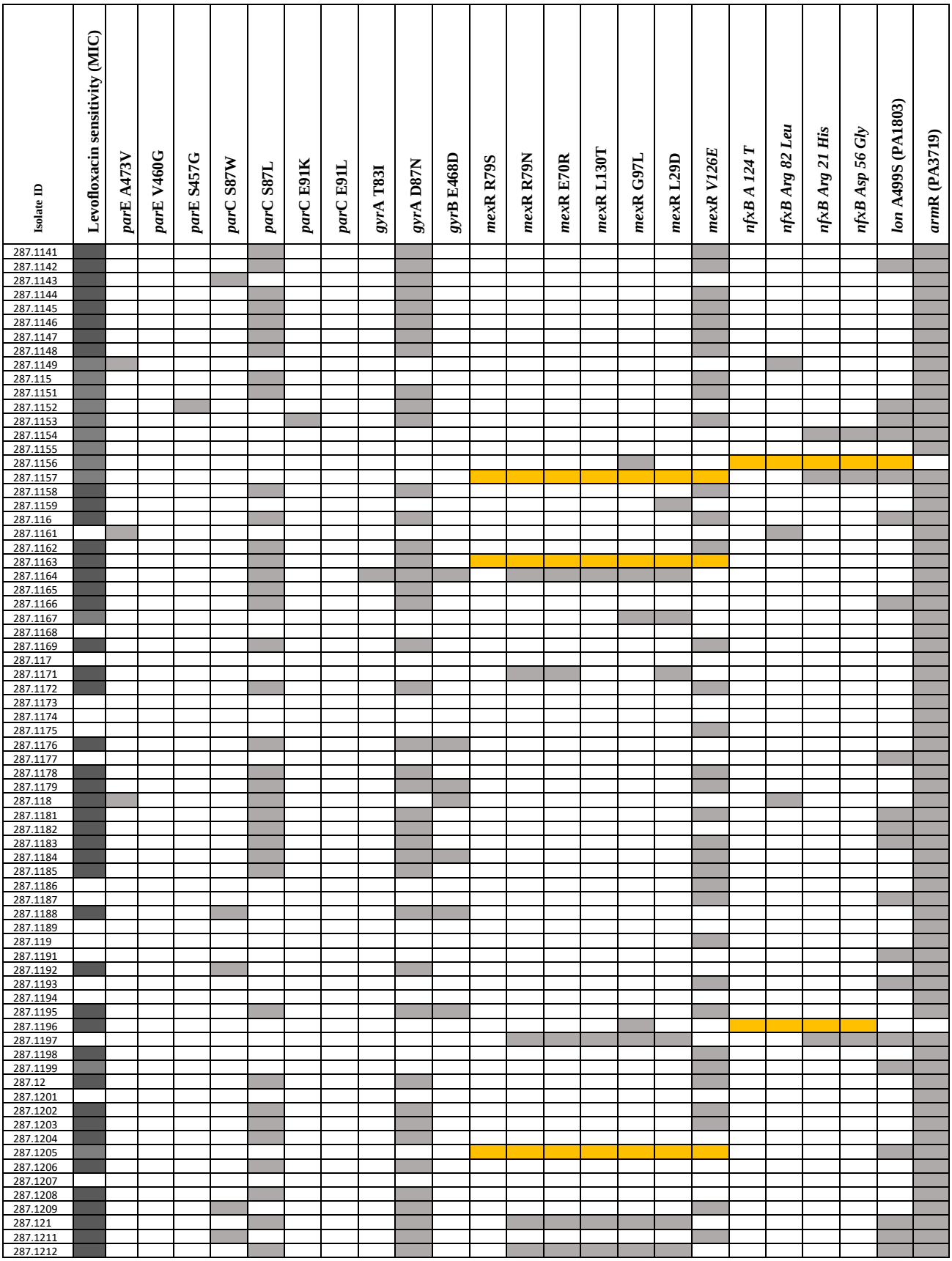

Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.2975
287.2976
287.2977
287.2978
287.2979
287.298
287.5685
287.5686
287.5687
287.5688
287.5689
287.569
287.5701
287.5702
287.5703
287.5704
287.5746
287.5747
287.5748
287.5749
287.575
287.5751
287.5752
287.5778
287.5955
287.5956
287.5959
287.5969
287.5971
287.5972
287.5973
287.6326
287.6327
287.6328
287.6329
287.633
287.6331
287.6492
287.7771
287.7772
287.7773
287.7774
287.7775
287.7776
287.7777
287.7778
287.7779
287.778
287.7781
287.7782
287.7783
287.7784
287.7785
287.7786
287.7787
287.7788
287.7789
287.779
287.7791
287.7792
287.7793
287.7794
287.7795
287.7796
287.7797
287.7798
287.7799
287.78
287.7801
287.7802
287.7803
287.7804

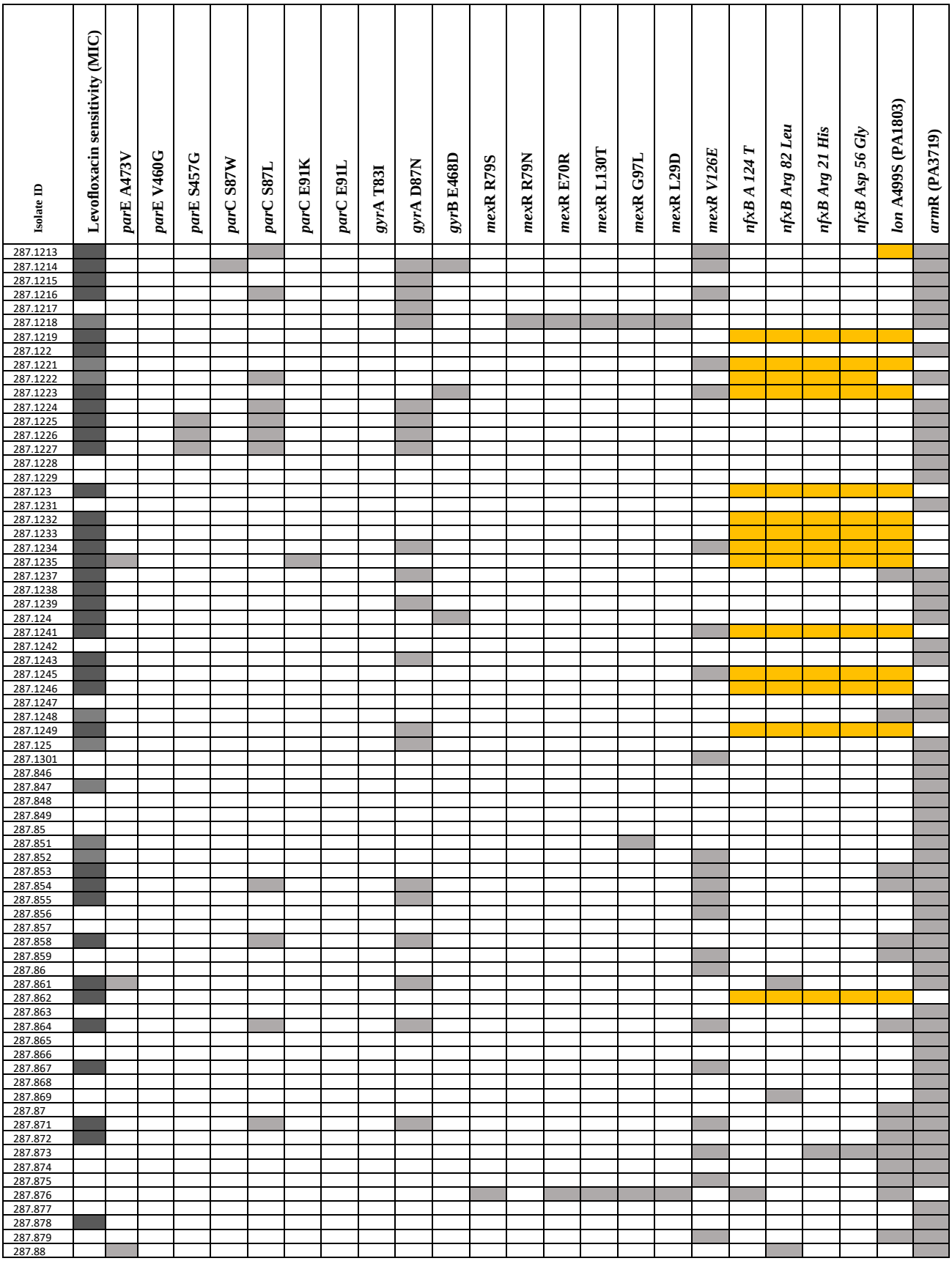

Isolate ID
Ciprofloxacin sensitivity (MIC)
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.7805
287.7806
287.7807
287.7808
287.7809
287.781
287.7811
287.7812
287.7813
287.7814
287.7815
287.7816
287.7817
287.7818
287.7819
287.782
287.7821
287.7822
287.7823
287.7824
287.7825
287.7826
287.7827
287.7828
287.7829
287.783
287.7831
287.7832
287.7833
287.7834
287.7835
287.7836
287.7837
287.7838
287.7839
287.784
287.7841
287.7842
287.7843
287.7844
287.7845
287.7846
287.7847
287.7848
287.7849
287.785
287.7851
287.7852
287.7853
287.7854
287.7855
287.7856
287.7857
287.7858
287.7859
287.786
287.7861
287.7862
287.8027
287.8028
287.8029
287.803
287.8031
287.8032
287.8033
287.8034

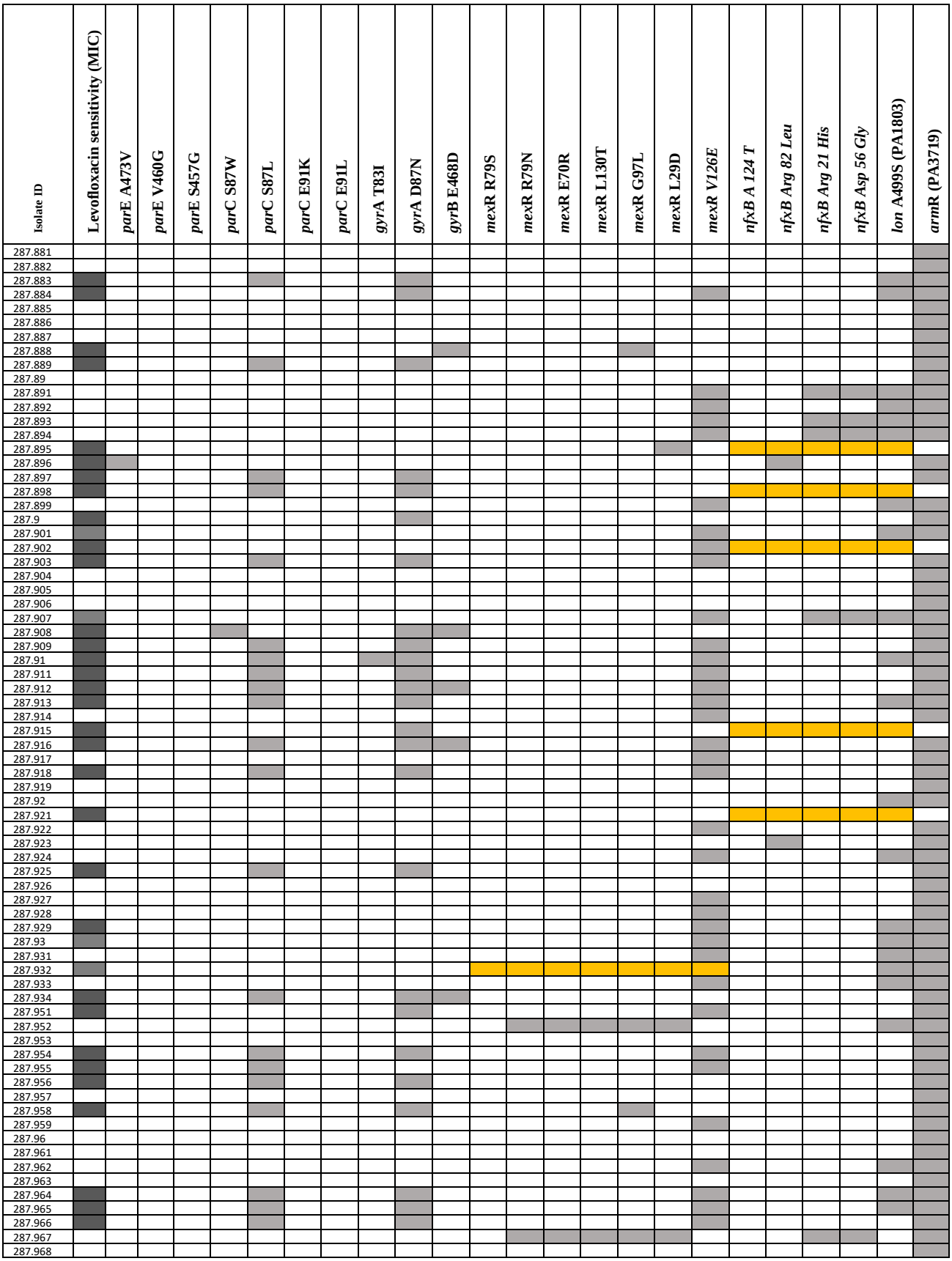
Supplementary Table 3A: Distribution of known quinolone resistance mutations (Levofloxacin data only)




Isolate ID
Levofloxacin sensitivity (MIC)
parE A473V
parE V460G
parE S457G
parC S87W
parC S87L
parC E91K
parC E91L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
lon A499S (PA1803)
armR (PA3719)
287.1069
287.107
287.1071
287.1072
287.1073
287.1074
287.1075
287.1076
287.1077
287.1078
287.1079
287.108
287.1081
287.1082
287.1083
287.1084
287.1085
287.1086
287.1087
287.1088
287.1089
287.109
287.1091
287.1092
287.1093
287.1094
287.1095
287.1096
287.1097
287.1098
287.1099
287.11
287.1101
287.1102
287.1103
287.1104
287.1105
287.1106
287.1107
287.1108
287.1109
287.111
287.1111
287.1112
287.1113
287.1114
287.1115
287.1116
287.1117
287.1118
287.1119
287.112
287.1121
287.1122
287.1123
287.1124
287.1125
287.1126
287.1127
287.1128
287.1129
287.113
287.1131
287.1132
287.1133
287.1134
287.1135
287.1136
287.1137
287.1138
287.1139
287.114


Isolate ID
Levofloxacin sensitivity (MIC)
parE A473V
parE V460G
parE S457G
parC S87W
parC S87L
parC E91K
parC E91L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
lon A499S (PA1803)
armR (PA3719)


Isolate ID
Levofloxacin sensitivity (MIC)
parE A473V
parE V460G
parE S457G
parC S87W
parC S87L
parC E91K
parC E91L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
lon A499S (PA1803)
armR (PA3719)
287.1213
287.1214
287.1215
287.1216
287.1217
287.1218
287.1219
287.122
287.1221
287.1222
287.1223
287.1224
287.1225
287.1226
287.1227
287.1228
287.1229
287.123
287.1231
287.1232
287.1233
287.1234
287.1235
287.1237
287.1238
287.1239
287.124
287.1241
287.1242
287.1243
287.1245
287.1246
287.1247
287.1248
287.1249
287.125
287.1301
287.846
287.847
287.848
287.849
287.85
287.851
287.852
287.853
287.854
287.855
287.856
287.857
287.858
287.859
287.86
287.861
287.862
287.863
287.864
287.865
287.866
287.867
287.868
287.869
287.87
287.871
287.872
287.873
287.874
287.875
287.876
287.877
287.878
287.879
287.88

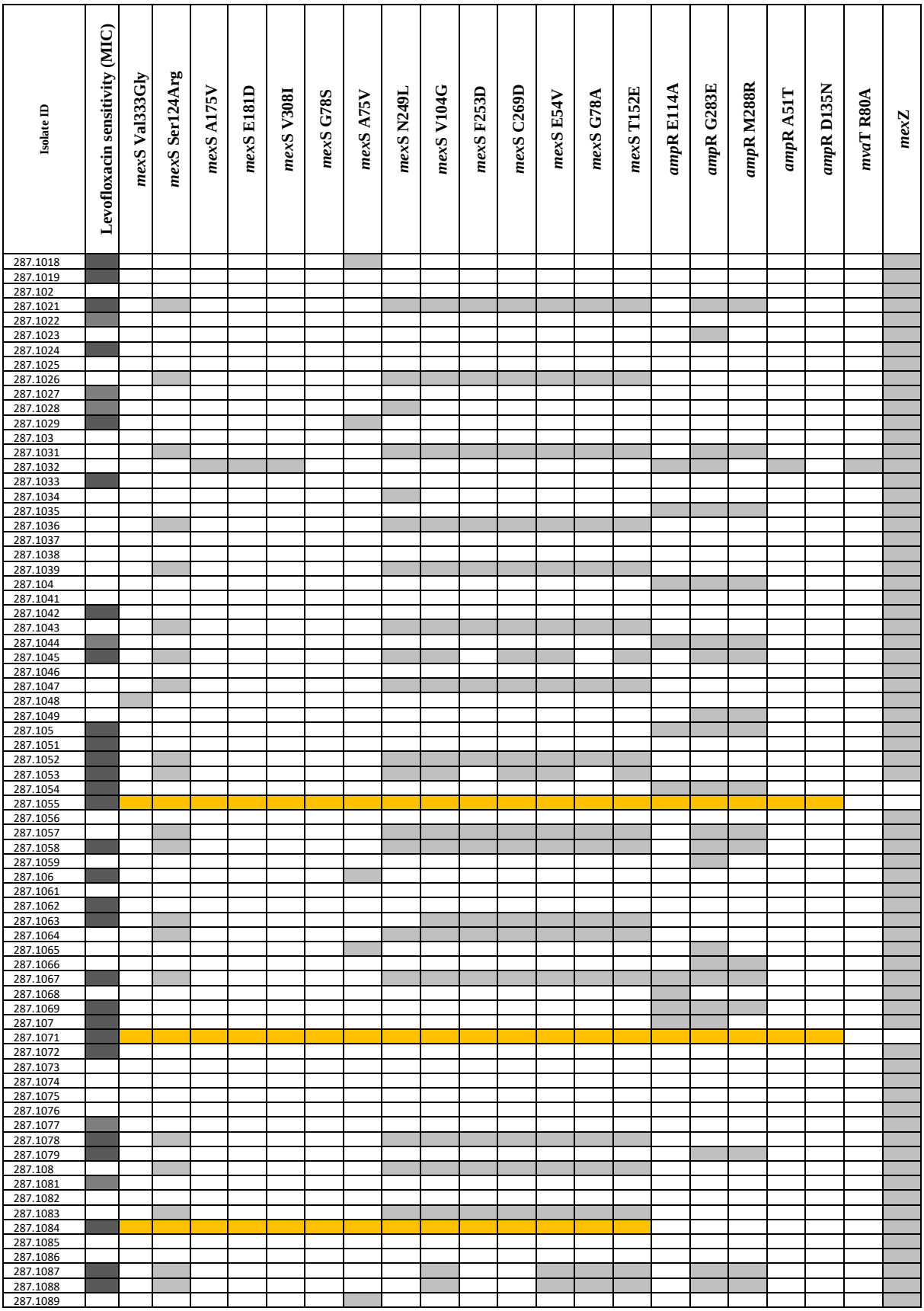

Isolate ID
Levofloxacin sensitivity (MIC)
parE A473V
parE V460G
parE S457G
parC S87W
parC S87L
parC E91K
parC E91L
gyrA T83I
gyrA D87N
gyrB E468D
mexR R79S
mexR R79N
mexR E70R
mexR L130T
mexR G97L
mexR L29D
mexR V126E
nfxB A 124 T
nfxB Arg 82 Leu
nfxB Arg 21 His
nfxB Asp 56 Gly
lon A499S (PA1803)
armR (PA3719)
287.881
287.882
287.883
287.884
287.885
287.886
287.887
287.888
287.889
287.89
287.891
287.892
287.893
287.894
287.895
287.896
287.897
287.898
287.899
287.9
287.901
287.902
287.903
287.904
287.905
287.906
287.907
287.908
287.909
287.91
287.911
287.912
287.913
287.914
287.915
287.916
287.917
287.918
287.919
287.92
287.921
287.922
287.923
287.924
287.925
287.926
287.927
287.928
287.929
287.93
287.931
287.932
287.933
287.934
287.951
287.952
287.953
287.954
287.955
287.956
287.957
287.958
287.959
287.96
287.961
287.962
287.963
287.964
287.965
287.966
287.967
287.968

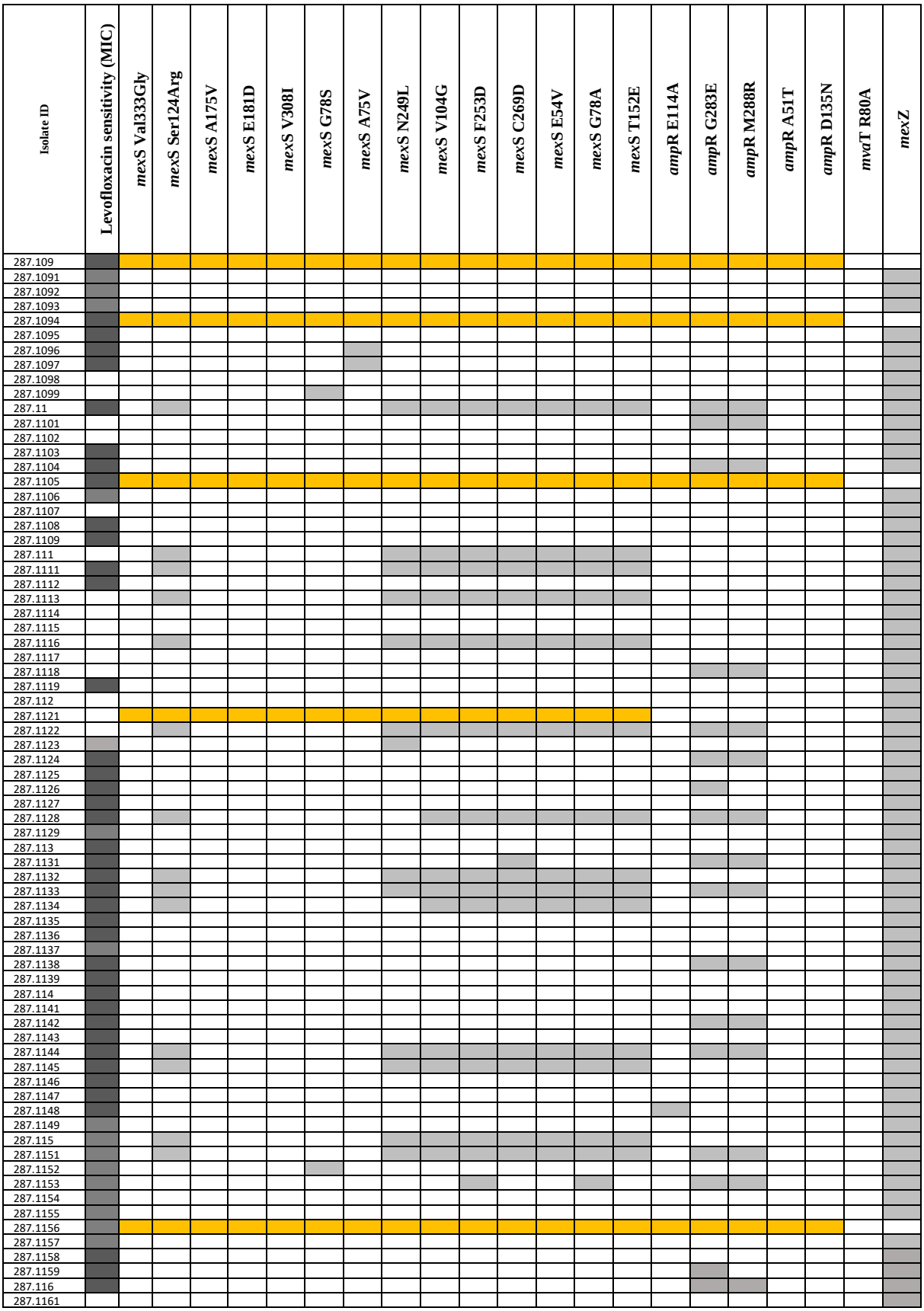
Isolate ID	Levofloxacin sensitivity (MIC)	*par*E A473V	*par*E V460G	*par*E S457G	*par*C S87W	*par*C S87L	*par*C E91K	*par*C E91L	*gyr*A T83I	*gyr*A D87N	*gyr*B E468D	*mex*R R79S	*mex*R R79N	*mex*R E70R	*mex*R L130T	*mex*R G97L	*mex*R L29D	*mex*R *V126E*	*nfx*B *A 124 T*	*nfx*B *Arg 82 Leu*	*nfx*B *Arg 21 His*	*nfx*B *Asp 56 Gly*	*lon* A499S (PA1803)	*arm*R (PA3719)
287.969																								
287.97																								
287.971																								
287.972																								
287.973																								
287.974																								
287.975																								
287.976																								
287.977																								
287.978																								
287.979																								
287.98																								
287.981																								
287.982																								
287.983																								
287.984																								
287.985																								
287.986																								
287.987																								
287.988																								
287.989																								
287.99																								
287.991																								
287.992																								
287.993																								
287.994																								
287.995																								
287.996																								
287.997																								
287.998																								
287.999																								
1611770																								

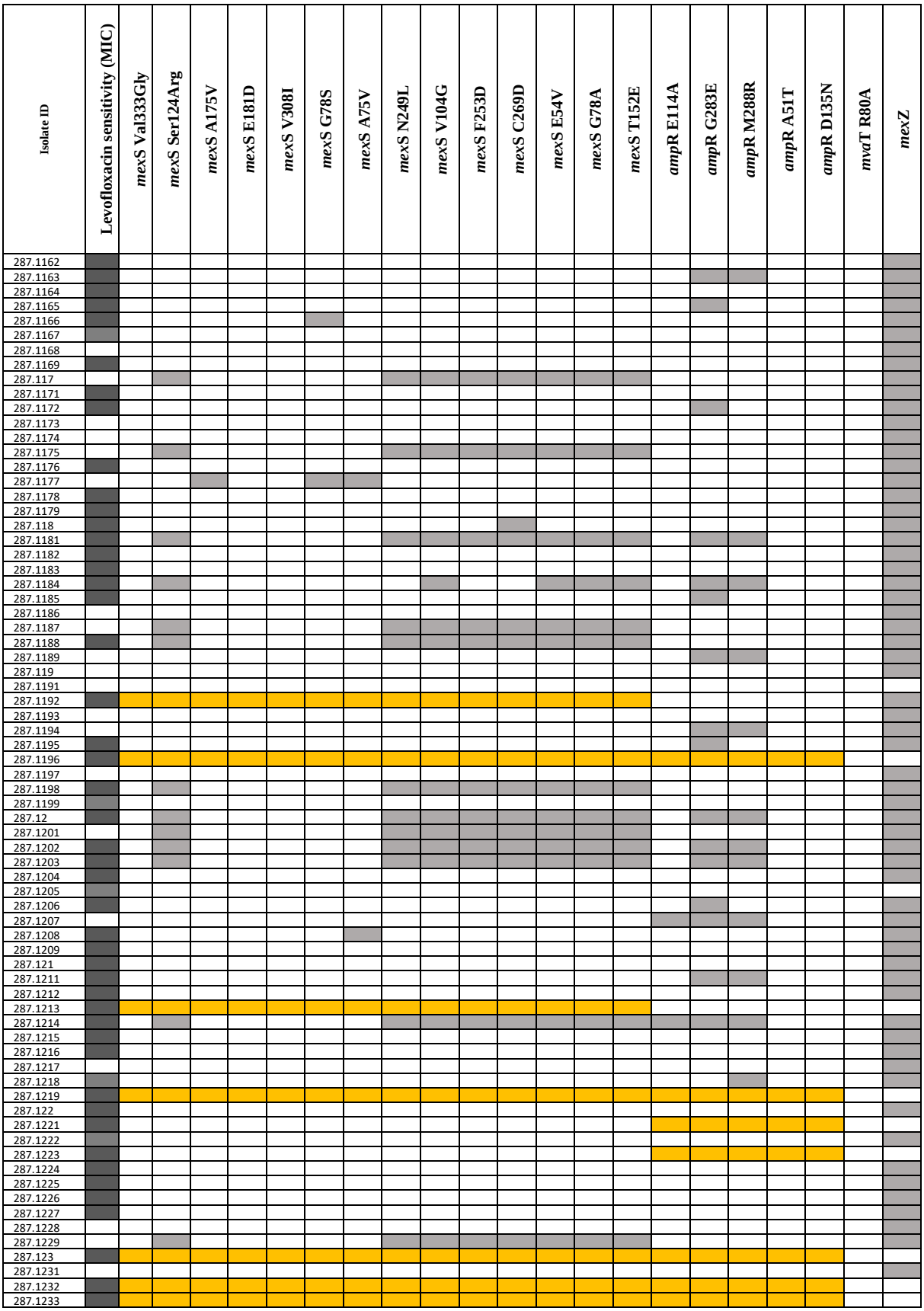
Grey shade: variant present, Grey shade: Antibiotic resistance, Yellow shade: gene absent

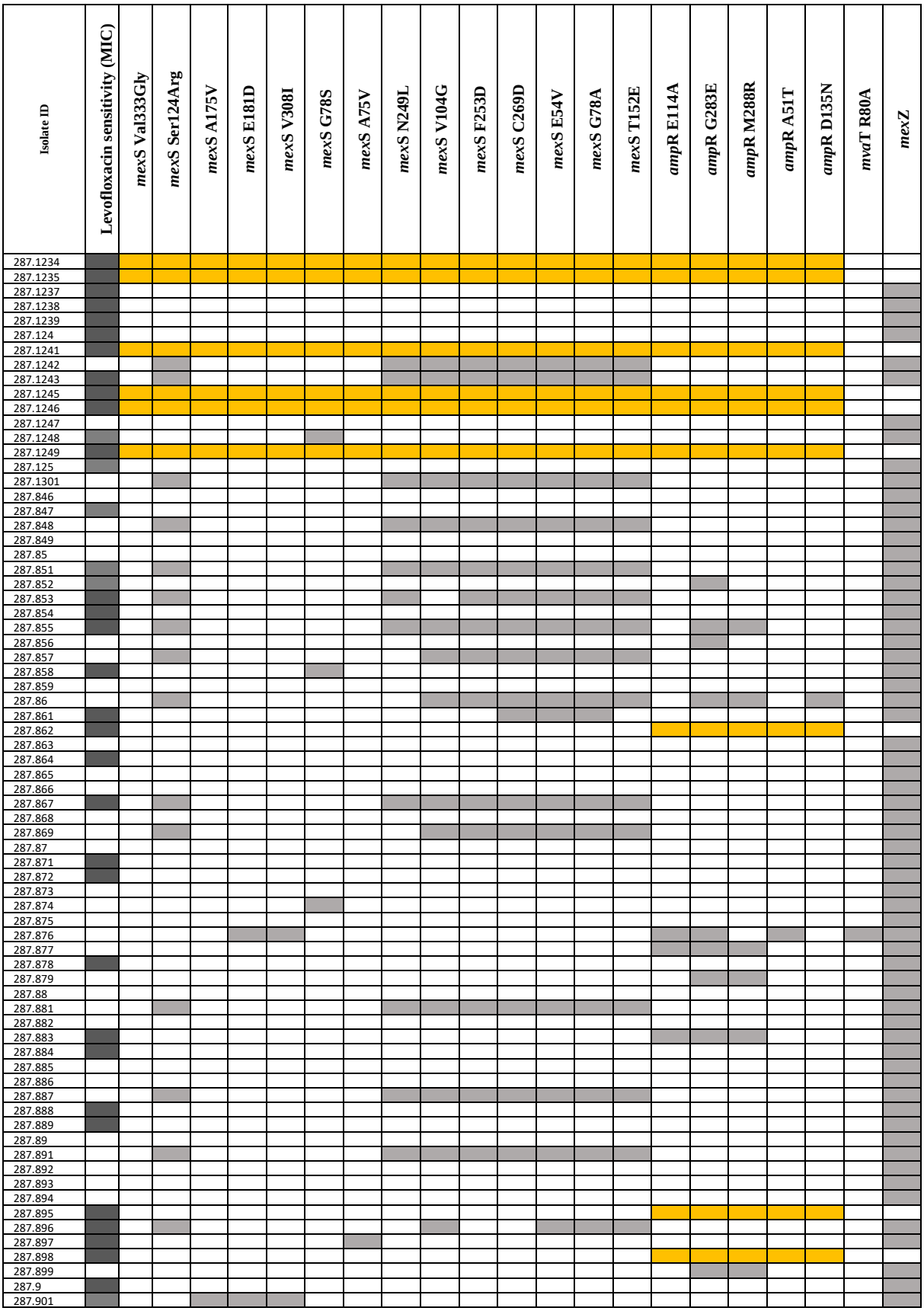
Supplementary Table 3B: Distribution of known quinolone resistance mutations (Levofloxacin data only):

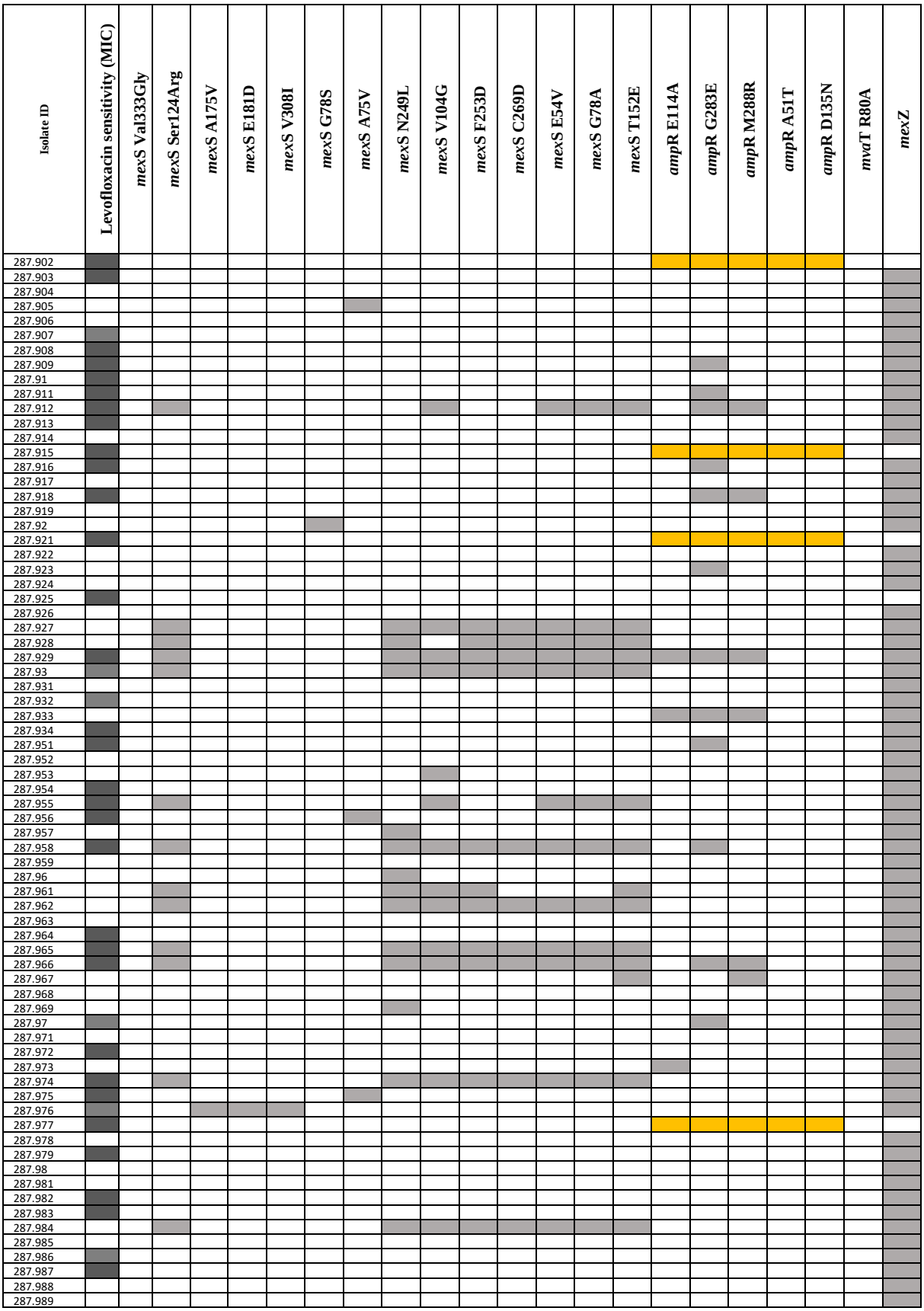
Isolate ID	Levofloxacin sensitivity (MIC)	*mex*S Val333Gly	*mex*S Ser124Arg	*mex*S A175V	*mex*S E181D	*mex*S V308I	*mex*S G78S	*mex*S A75V	*mex*S N249L	*mex*S V104G	*mex*S F253D	*mex*S C269D	*mex*S E54V	*mex*S G78A	*mex*S T152E	*amp*R E114A	*amp*R G283E	*amp*R M288R	*amp*R A51T	*amp*R D135N	*mva*T R80A	*mex*Z
287.1																						
287.1001																						
287.1002																						
287.1003																						
287.1004																						
287.1005																						
287.1006																						
287.1007																						
287.1008																						
287.1009																						
287.101																						
287.1011																						
287.1012																						
287.1013																						
287.1014																						
287.1015																						
287.1016																						
287.1017																						


Isolate ID
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS Ser124Arg
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.1018
287.1019
287.102
287.1021
287.1022
287.1023
287.1024
287.1025
287.1026
287.1027
287.1028
287.1029
287.103
287.1031
287.1032
287.1033
287.1034
287.1035
287.1036
287.1037
287.1038
287.1039
287.104
287.1041
287.1042
287.1043
287.1044
287.1045
287.1046
287.1047
287.1048
287.1049
287.105
287.1051
287.1052
287.1053
287.1054
287.1055
287.1056
287.1057
287.1058
287.1059
287.106
287.1061
287.1062
287.1063
287.1064
287.1065
287.1066
287.1067
287.1068
287.1069
287.107
287.1071
287.1072
287.1073
287.1074
287.1075
287.1076
287.1077
287.1078
287.1079
287.108
287.1081
287.1082
287.1083
287.1084
287.1085
287.1086
287.1087
287.1088
287.1089


Isolate ID
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS Ser124Arg
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
nvaT R80A
mexZ
287.109
287.1091
287.1092
287.1093
287.1094
287.1095
287.1096
287.1097
287.1098
287.1099
287.11
287.1101
287.1102
287.1103
287.1104
287.1105
287.1106
287.1107
287.1108
287.1109
287.111
287.1111
287.1112
287.1113
287.1114
287.1115
287.1116
287.1117
287.1118
287.1119
287.112
287.1121
287.1122
287.1123
287.1124
287.1125
287.1126
287.1127
287.1128
287.1129
287.113
287.1131
287.1132
287.1133
287.1134
287.1135
287.1136
287.1137
287.1138
287.1139
287.114
287.1141
287.1142
287.1143
287.1144
287.1145
287.1146
287.1147
287.1148
287.1149
287.115
287.1151
287.1152
287.1153
287.1154
287.1155
287.1156
287.1157
287.1158
287.1159
287.116
287.1161


Isolate ID
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS Ser124Arg
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.1162
287.1163
287.1164
287.1165
287.1166
287.1167
287.1168
287.1169
287.117
287.1171
287.1172
287.1173
287.1174
287.1175
287.1176
287.1177
287.1178
287.1179
287.118
287.1181
287.1182
287.1183
287.1184
287.1185
287.1186
287.1187
287.1188
287.1189
287.119
287.1191
287.1192
287.1193
287.1194
287.1195
287.1196
287.1197
287.1198
287.1199
287.12
287.1201
287.1202
287.1203
287.1204
287.1205
287.1206
287.1207
287.1208
287.1209
287.121
287.1211
287.1212
287.1213
287.1214
287.1215
287.1216
287.1217
287.1218
287.1219
287.122
287.1221
287.1222
287.1223
287.1224
287.1225
287.1226
287.1227
287.1228
287.1229
287.123
287.1231
287.1232
287.1233


Isolate ID
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS Ser124Arg
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.1234
287.1235
287.1237
287.1238
287.1239
287.124
287.1241
287.1242
287.1243
287.1245
287.1246
287.1247
287.1248
287.1249
287.125
287.1301
287.846
287.847
287.848
287.849
287.85
287.851
287.852
287.853
287.854
287.855
287.856
287.857
287.858
287.859
287.86
287.861
287.862
287.863
287.864
287.865
287.866
287.867
287.868
287.869
287.87
287.871
287.872
287.873
287.874
287.875
287.876
287.877
287.878
287.879
287.88
287.881
287.882
287.883
287.884
287.885
287.886
287.887
287.888
287.889
287.89
287.891
287.892
287.893
287.894
287.895
287.896
287.897
287.898
287.899
287.9
287.901


Isolate ID
Levofloxacin sensitivity (MIC)
mexS Val333Gly
mexS Ser124Arg
mexS A175V
mexS E181D
mexS V308I
mexS G78S
mexS A75V
mexS N249L
mexS V104G
mexS F253D
mexS C269D
mexS E54V
mexS G78A
mexS T152E
ampR E114A
ampR G283E
ampR M288R
ampR A51T
ampR D135N
mvaT R80A
mexZ
287.902
287.903
287.904
287.905
287.906
287.907
287.908
287.909
287.91
287.911
287.912
287.913
287.914
287.915
287.916
287.917
287.918
287.919
287.92
287.921
287.922
287.923
287.924
287.925
287.926
287.927
287.928
287.929
287.93
287.931
287.932
287.933
287.934
287.951
287.952
287.953
287.954
287.955
287.956
287.957
287.958
287.959
287.96
287.961
287.962
287.963
287.964
287.965
287.966
287.967
287.968
287.969
287.97
287.971
287.972
287.973
287.974
287.975
287.976
287.977
287.978
287.979
287.98
287.981
287.982
287.983
287.984
287.985
287.986
287.987
287.988
287.989

Isolate ID	Levofloxacin sensitivity (MIC)	*mex*S Val333Gly	*mex*S Ser124Arg	*mex*S A175V	*mex*S E181D	*mex*S V308I	*mex*S G78S	*mex*S A75V	*mex*S N249L	*mex*S V104G	*mex*S F253D	*mex*S C269D	*mex*S E54V	*mex*S G78A	*mex*S T152E	*ampR* E114A	*ampR* G283E	*ampR* M288R	*ampR* A51T	*ampR* D135N	*mvaT* R80A	*mexZ*
287.99																						
287.991																						
287.992																						
287.993																						
287.994																						
287.995																						
287.996																						
287.997																						
287.998																						
287.999																						
1611770																						

Grey shade: Resistant, No shade: susceptible

Grey shade: Mutant variant or gene present, No shade: variant or gene absent, Yellow shade: gene absent

Supplementary Table 4A: Distribution of aminoglycoside resistance variants (amikacin and gentamycin)

Isolate ID	AK sensitivity	CN sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
PAE0002																												
PAE0005																												
PAE0006																												
PAE0007																												
PAE0008																												
PAE0010																												
PAE0011																												
PAE0012																												
PAE0014																												
PAE0018																												
PAE0020																												
PAE0021																												
PAE0024																												
PAE0025																												
PAE0026																												
PAE0029																												
PAE0030																												
PAE0032																												
PAE0035																												
PAE0036																												
PAE0039																												
PAE0040																												
PAE0041																												
PAE0043																												
PAE0044																												
PAE0046																												
PAE0047																												
PAE0048																												
PAE0051																												
PAE0055																												

Isolate ID	AK sensitivity	CN sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
PAE0057											■			■														
PAE0059		■																										
PAE0060																												
PAE0061		■		■	■	■	■	■	■																			
PAE0062				■	■	■	■	■				■																
PAE0063		■										■																
PAE0068																						■						
PAE0070																												
PAE0071				■	■	■	■	■	■			■																
PAE0072		■																										
PAE0076												■																
PAE0079			■										■		■													
PAE0081																			■					■				
PAE0082			■									■						■	■						■			
PAE0083												■						■										
PAE0084			■						■			■						■										
PAE0085		■									■													■			■	■
PAE0087																												
PAE0088		■																										
PAE0090			■						■			■							■									
PAE0091		■																										
PAE0092									■																			
PAE0093																				■	■							
PAE0095																								■				
PAE0096																												
PAE0098																												
PAE0099												■							■					■				
PAE0106				■	■	■	■	■	■			■							■									
PAE0109		■																										
PAE0111		■										■													■			
PAE0112			■									■						■				■			■			
PAE0113		■																										
PAE0115				■	■	■	■	■	■			■							■									
PAE0116																												
PAE0119				■	■	■	■	■	■										■									
PAE0124		■																										
PAE0125		■																										
PAE0136		■																										
PAE0142		■																	■									
PAE0144		■	■																									
PAE0145		■							■			■							■									
PAE0147				■	■			■	■			■																
PAE0148			■									■				■	■	■	■									
PAE0149												■							■									
PAE0151			■									■																
PAE0154		■																										
PAE0156		■							■										■									
PAE0157	■	■																										
PAE0158												■																
PAE0160		■																										
PAE0161																				■				■				
PAE0167		■							■																			
PAE0168	■	■								■																		
PAE0171		■		■	■	■	■	■																■				
PAE0172																												
PAE0174			■								■																	
PAE0175	■	■							■																			
PAE0001												■							■									
PAE0003																												
PAE0004																				■				■				
PAE0009												■																
PAE0013			■						■			■							■						■			
PAE0015																				■	■							
PAE0016		■							■																			
PAE0017																												
PAE0019				■	■	■	■	■				■																
PAE0022				■	■	■	■	■	■			■																
PAE0023									■																			
PAE0027																				■	■							
PAE0031																												
PAE0033																				■				■				
PAE0034																												

Isolate ID	AK sensitivity	CN sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
PAE0037																												
PAE0038																												
PAE0042																												
PAE0045																												
PAE0049																												
PAE0050																												
PAE0052																												
PAE0053																												
PAE0054																												
PAE0056																												
PAE0058																												
PAE0064																												
PAE0065																												
PAE0066																												
PAE0067																												
PAE0069																												
PAE0073																												
PAE0074																												
PAE0075																												
PAE0077																												
PAE0078																												
PAE0080																												
PAE0086																												
PAE0089																												
PAE0094																												
PAE0097																												
PAE0101																												
PAE0102																												
PAE0103																												
PAE0105																												
PAE0107																												
PAE0108																												
PAE0110																												
PAE0114																												
PAE0117																												
PAE0118																												
PAE0120																												
PAE0121																												
PAE0122																												
PAE0123																												
PAE0126																												
PAE0127																												
PAE0128																												
PAE0137																												
PAE0138																												
PAE0139																												
PAE0140																												
PAE0141																												
PAE0143																												
PAE0150																												
PAE0152																												
PAE0153																												
PAE0155																												
PAE0159																												
PAE0162																												
PAE0163																												
PAE0165																												
PAE0166																												
PAE0169																												
PAE0170																												
287.2972																												
287.2973																												
287.2975																												
287.2976																												
287.2977																												
287.2978																												
287.2979																												
287.2980																												
287.568																												
287.5686																												
287.5687																												
287.5688																												

Isolate ID	AK sensitivity	CN sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.5689																												
287.5690																												
287.5701																												
287.5702																												
287.5703																												
287.5704																												
287.5746																												
287.5747																												
287.5748																												
287.5749																												
287.5750																												
287.5751																												
287.575																												
287.5778																												
287.5955																												
287.5956																												
287.5959																												
287.5969																												
287.5971																												
287.5972																												
287.5973																												
287.6326																												
287.6327																												
287.6328																												
287.6329																												
287.6330																												
287.6331																												
287.6492																												
287.7771																												
287.777																												
287.7773																												
287.7774																												
287.7775																												
287.7776																												
287.7777																												
287.7778																												
287.7779																												
287.7780																												
287.7781																												
287.7782																												
287.7783																												
287.7784																												
287.7785																												
287.7786																												
287.7787																												
287.7788																												
287.7789																												
287.7790																												
287.7791																												
287.7792																												
287.7793																												
287.7794																												
287.7795																												
287.7796																												
287.7797																												
287.7798																												
287.7799																												
287.7800																												
287.7801																												
287.7802																												
287.7803																												
287.7804																												
287.7805																												
287.7806																												
287.7807																												
287.7808																												
287.7809																												
287.7810																												
287.7811																												
287.7812																												
287.7813																												
287.7814																												

Isolate ID	AK sensitivity	CN sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.7815																												
287.7816																												
287.7817																												
287.7818																												
287.7819																												
287.7820																												
287.7821																												
287.7822																												
287.7823																												
287.7824																												
287.7825																												
287.7826																												
287.7827																												
287.7828																												
287.7829																												
287.7830																												
287.7831																												
287.7832																												
287.7833																												
287.7834																												
287.7835																												
287.7836																												
287.7837																												
287.783																												
287.7839																												
287.7840																												
287.7841																												
287.7842																												
287.7844																												
287.7845																												
287.7846																												
287.7847																												
287.7848																												
287.7849																												
287.7850																												
287.7851																												
287.7852																												
287.7853																												
287.7854																												
287.7855																												
287.7856																												
287.7857																												
287.7858																												
287.7859																												
287.7860																												
287.7861																												
287.7862																												
287.8027																												
287.8028																												
287.8029																												
287.8030																												
287.8031																												
287.8032																												
287.8033																												
287.8034																												

Supplementary Table 4B: Distribution of aminoglycoside resistance variants (amikacin only)

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rplY*Ala123Ser	*rplY*O41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.1000																											
287.1001																											
287.1002																											
287.1003																											
287.1004																											
287.1005																											
287.1006																											
287.1007																											
287.1008																											
287.1009																											
287.1010																											
287.1011																											
287.1012																											
287.1013																											
287.1014																											
287.1015																											
287.1016																											
287.1017																											
287.1018																											
287.1019																											
287.1020																											
287.1021																											
287.1022																											
287.1023																											
287.1024																											
287.1025																											
287.1026																											
287.1027																											
287.1028																											
287.1029																											
287.1030																											
287.1031																											
287.1032																											
287.1033																											
287.1034																											
287.1035																											
287.1036																											
287.1037																											
287.1038																											
287.1039																											
287.1040																											
287.1041																											
287.1042																											
287.1043																											
287.1044																											
287.1045																											
287.1046																											
287.1047																											
287.1048																											
287.1049																											
287.1050																											
287.1051																											
287.1052																											
287.1053																											
287.1054																											
287.1055																											
287.1056																											
287.1057																											
287.1058																											
287.1059																											
287.1060																											
287.1061																											
287.1062																											
287.1063																											
287.1064																											
287.1065																											
287.1066																											
287.1067																											
287.1068																											
287.1069																											

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*nmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*qid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.1070																											
287.1071																											
287.1072																											
287.1073																											
287.1074																											
287.1075																											
287.1076																											
287.1077																											
287.1078																											
287.1079																											
287.1080																											
287.1081																											
287.1082																											
287.1083																											
287.1084																											
287.1085																											
287.1086																											
287.1087																											
287.1088																											
287.1089																											
287.1090																											
287.1091																											
287.1092																											
287.1093																											
287.1095																											
287.1096																											
287.1097																											
287.1098																											
287.1099																											
287.1100																											
287.1101																											
287.1102																											
287.1103																											
287.1104																											
287.1105																											
287.1106																											
287.1107																											
287.1108																											
287.1109																											
287.1110																											
287.1111																											
287.1112																											
287.1113																											
287.1114																											
287.1115																											
287.1116																											
287.1117																											
287.1118																											
287.1119																											
287.1120																											
287.1121																											
287.1122																											
287.1123																											
287.1124																											
287.1125																											
287.1126																											
287.1127																											
287.1128																											
287.1129																											
287.1130																											
287.1131																											
287.1132																											
287.1133																											
287.1134																											
287.1135																											
287.1136																											
287.1137																											
287.1138																											
287.1139																											
287.1140																											
287.1141																											
287.1142																											

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*nmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*qid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.1143																											
287.1144																											
287.1145																											
287.1146																											
287.1147																											
287.1148																											
287.1149																											
287.1150																											
287.1151																											
287.1152																											
287.1153																											
287.1154																											
287.1155																											
287.1156																											
287.1157																											
287.1158																											
287.1159																											
287.1160																											
287.1161																											
287.1162																											
287.1163																											
287.1164																											
287.1165																											
287.1166																											
287.1167																											
287.1168																											
287.1169																											
287.1170																											
287.1171																											
287.1172																											
287.1173																											
287.1174																											
287.1175																											
287.1176																											
287.1177																											
287.1178																											
287.1179																											
287.1180																											
287.1181																											
287.1182																											
287.1183																											
287.1184																											
287.1185																											
287.1186																											
287.1187																											
287.1188																											
287.1189																											
287.1190																											
287.1191																											
287.1192																											
287.1193																											
287.1194																											
287.1195																											
287.1196																											
287.1197																											
287.1198																											
287.1199																											
287.1200																											
287.1201																											
287.1202																											
287.1203																											
287.1204																											
287.1205																											
287.1206																											
287.1207																											
287.1208																											
287.1209																											
287.1210																											
287.1211																											
287.1212																											
287.1213																											
287.1214																											

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*qid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.1215																											
287.1216																											
287.1217																											
287.1218																											
287.1219																											
287.1220																											
287.1221																											
287.1222																											
287.1223																											
287.1224																											
287.1225																											
287.1226																											
287.1227																											
287.1228																											
287.1229																											
287.123																											
287.1231																											
287.1232																											
287.1233																											
287.1234																											
287.1235																											
287.1237																											
287.1238																											
287.1239																											
287.1240																											
287.1241																											
287.1242																											
287.1243																											
287.1245																											
287.1246																											
287.1247																											
287.1248																											
287.1249																											
287.1250																											
287.1301																											
287.846																											
287.847																											
287.848																											
287.849																											
287.850																											
287.851																											
287.852																											
287.853																											
287.854																											
287.855																											
287.856																											
287.857																											
287.858																											
287.859																											
287.860																											
287.861																											
287.862																											
287.863																											
287.864																											
287.865																											
287.866																											
287.867																											
287.868																											
287.869																											
287.870																											
287.871																											
287.872																											
287.873																											
287.874																											
287.875																											
287.876																											
287.877																											
287.878																											
287.879																											
287.880																											
287.881																											
287.882																											

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*nmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*qid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.883			■	■	■	■	■	■																			
287.884											■													■			
287.885																											
287.886																											
287.887																			■	■							
287.888																											
287.889	■							■														■				■	■
287.890																											
287.891											■																
287.892		■																						■			
287.893			■	■	■	■	■	■			■																
287.894		■						■			■																
287.895																											
287.896																											
287.897	■																										
287.899		■									■																
287.900								■										■									
287.901		■												■													
287.902			■	■	■	■	■	■			■																
287.903			■	■	■	■	■								■	■	■	■									
287.904																											
287.905																									■		
287.906																											
287.907			■	■	■	■	■	■			■																
287.908								■																			
287.909			■	■	■	■	■																				
287.910											■																
287.911		■																									
287.912		■									■				■	■	■	■									
287.913		■						■			■																
287.914		■						■			■							■									
287.915	■		■	■	■	■	■				■				■	■	■	■									
287.916	■	■																									
287.917											■							■					■				
287.918	■	■									■				■	■	■	■									
287.919											■																
287.920																			■				■				
287.921																			■	■							
287.922																											
287.923								■																			
287.924		■										■		■													
287.925								■										■									
287.926																											
287.927		■						■			■						■										
287.928			■	■	■	■	■																				
287.929			■	■	■	■	■	■			■																
287.930			■	■	■	■	■	■			■																
287.931											■													■			
287.932																							■				
287.933			■	■	■	■	■	■			■																
287.934	■							■														■				■	■
287.951	■		■	■	■	■	■																				
287.952		■									■																
287.953																											
287.954											■																
287.955			■	■	■	■	■				■				■	■	■	■									
287.956								■																			
287.957																											
287.958																											
287.959																							■				
287.960								■																			
287.961																			■	■							
287.962								■			■							■									
287.963																											
287.964			■	■	■	■	■	■			■																
287.965		■						■			■																
287.966			■	■	■	■	■				■				■	■	■	■									
287.967																											
287.968																			■	■							
287.969																			■	■							
287.970																											
287.971			■	■	■	■	■	■			■																

Isolate ID	AK sensitivity	*pmr*BALA4Thr	*pmr*BLeu323His	*pmr*BSer420Arg	*pmr*BGly423Cys	*pmr*BL243Q	*pmr*BA248V	*pmr*ALeu71Arg	*fus*A1D588G	*amg*SE108Q	*rpl*YAla123Ser	*rpl*YQ41L	*gid*BQ28K	*gid*BE126G	*gid*BE97Q	*gid*BE186A	*pho*QY85F	*nuo*GS468A	*nuo*GA574T	*nuo*GA890T	*pst*BR87C	*pst*BE89Q	*lpt*AT55A	*lpt*AR62S	*fao*AT385A	*arn*AA170T	*arn*DG206C
287.972																											
287.973																											
287.974																											
287.975																											
287.976																											
287.977																											
287.978																											
287.979																											
287.980																											
287.981																											
287.982																											
287.983																											
287.984																											
287.985																											
287.986																											
287.987																											
287.988																											
287.989																											
287.990																											
287.991																											
287.992																											
287.993																											
287.994																											
287.995																											
287.996																											
287.997																											
287.998																											
287.999																											

Supplementary Table 5: Summary of QRDR mutations in relation to quinolones resistance

Study	no. of isolates studied	Type of mutation	Number of isolates	Mutation pattern	Resistance rate
1	114 clinical isolates 22 isolates were resistant to Levofloxacin (19.3%)	**six types of mutations**		combination of *gyr*A and *par*C mutations were distributed in 10 different patterns	
		Thr83Ile (*gyr*A)	48 of 114 isolates (42.1%)	Thr83 to Ile in *gyr*A alone was most frequently found (25 strains; 21.9%)	3/25 are resistant
		Asp87Tyr (*gyr*A)	3 of 114 (2.6%)	combination of Thr83 to Ile in *gyr*A with Asp87 to Asn in *gyr*A (4 strains; 3.5%)	2/4 are resistant

		Asp87Asn (*gyr*A)	9 of 114 isolates (7.9%)	combination of Thr83 to Ile in *gyr*A with Asp87 to Tyr in *gyr*A (3 strains; 2.6%)	2/3 are resistant
		Glu91Arg (*par*C)	6 of 114 (5.3%)	combination of Thr83 to Ile in *gyr*A with Ser87 to Leu in *par*C (5 strains; 4.4%)	4/5 are resistant
		Ser87Leu (*par*C)	12 of 114(10.5%)	combination of Thr83 to Ile in *gyr*A with Ser87 to Trp in *par*C (2 strains;1.8 %)	2/2 are resistant
		Ser87Trp (*par*C)	2 of 114 (1.8%)	combination of Thr83 to Ile in *gyr*A with Glu91 to Arg in *par*C (2 strains; 1.8%)	2/2 are resistant
				combination of Thr83 to Ile in *gyr*A with Asp87 to Asn in *gyr*A with Ser87 to Leu in *par*C (3 strains; 2.6%)	3/3 are resistant
				combination of Thr83 to Ile in *gyr*A with Ser87 to Leu in *par*C with Glu91 to Arg in par C (2 strains; 1.8%)	2/2 are resistant
				combination of Thr83 to Ile in *gyr*A with Asp87 to Asn in *gyr*A, Ser87 to Leu in *par*C with Glu91 to Arg in *par*C (2 strains;1.8%)	2/2 are resistant
2	38 clinical isolates were non susceptible to at least one of the three fluoroquinolones tested	**five types of mutations**			
		Thr83Ile (*gyr*A mutation)	19 of 38 isolates	mutation in *gyr*A	10 out of 22 resistant to ciprofloxacin, ofloxacin, pefloxacin
		His80Arg (*gyr*A mutation)	1 isolate	mutations in both *gyr*A and *par*C	8 out of 22 resistant to ciprofloxacin, ofloxacin, pefloxacin
				no mutations in *gyr*A or *par*C	4 out of 22 resistant to ciprofloxacin, ofloxacin, pefloxacin
		Ser80Leu (*par*C mutation)	8 of 38 isolates	no mutations in all QRDR	11 out of 38 susceptible to ciprofloxacin

		Gln84Asp (*par*C mutation)	1 isolate	mutation in *gyr*A	1 out of 5 resistant to ofloxacin, pefloxacin/ intermediate to ciprofloxacin
		Ala85Gly (*par*C mutation)	1 isolate	mutations in both *gyr*A and *par*C	4 out of 5 resistant to ofloxacin, pefloxacin/ intermediate to ciprofloxacin
				no mutations in all QRDR	7 out of 7 resistant to ofloxacin, pefloxacin/ susceptible to ciprofloxacin
		four types of mutations			
3	100 clinical isolates 64 resistants to ciprofloxacin 63 resistants to levofloxacin	Thr83Ile (*gyr*A mutation)		Thr83Ile (*gyr*A mutation) alone	20 out of 64 resistant to ciprofloxacin and levofloxacin
		Asp87Asn (*gyr*A mutation)		combination of Thr83Ile in *gyr*A with Ala 88 Pro in *par*C	8 out of 64 resistant to ciprofloxacin
		Ala88Pro (*par*C mutation)		combination of Thr83Ile in *gyr*A withSer87 to Leu in *par*C	31 out of 64 resistant to ciprofloxacin and levofloxacin
		Ser87Leu (*par*C mutation)		combination of Thr83Ile in *gyr*A with Asp87 to Asn in *gyr*A with Ser87 to Leu in *par*C	5 out of 64 resistant to ciprofloxacin
		twenty types of mutations			
4	262 clinical isolates 232 resistants to ciprofloxacin 184 resistants to cipro + other drugs 48 resistants to cipro alone	Thr83Ile (*gyr*A mutation)		combination of Thr83Ile (*gyr*A mutation) with any listed *par*C mutation	99 out of 232
		Asp87Asn (*gyr*A mutation)		combination of Asp87Tyr (gyrA mutation) with Ser87Leu *par*C mutation	1 out of 232
		Asp87Tyr (*gyr*A mutation)		No mutations in other QRDR	42 out of 232
		Asp87His (*gyr*A mutation)		*gyr*A mutation alone (any of five types listed)	58 out of 232
		Asp87Gly (*gyr*A mutation)		combination of Thr83Ile (*gyr*A mutation) with different listed *par*E mutations	15 out of 232

		Ser87Leu (*par*C mutation)	combination of Asp87Asn (gyrA mutation) with Ala473Thr (*par*E mutation)	1 out of 232
		Gly85Cys (*par*C mutation)	Ser466Phe (*gyr*B mutation) only	10 out of 232
		Ser87Trp (*par*C mutation)	combination of Thr83Ile in *gyr*A withSer87 to Leu in *par*C with additional *par*E mutation	4 out of 232
		Glu91Lys (*par*C mutation)	Isolated Ser87Leu in *par*C	2 out of 232
		Ser466Phe (*gyr*B mutation)		
		Leu501Phe (*par*E mutation)		
		Ser457Gly (*par*E mutation)		
		Glu453Asp (*par*E mutation)		
		Asp419Asn (*par*E mutation)		
		Lys388Glu (*par*E mutation)		
		Glu459Val (*par*E mutation)		
		Val460Phe (*par*E mutation)		
		Ala473Val (*par*E mutation)		
		Ala473Thr (*par*E mutation)		
		Ser457Arg (*par*E mutation)		
5	210 clinical isolates 18 selected resistant isolates	Thr83Ile (*gyr*A mutation)	mutations were not studied in all QRDR	single Thr83Ile mutation (*gyr*A) 10 strains out of 15
		Asp87Asn (*gyr*A mutation)	only *gyr*A mutations studied	single Asp87Asn (*gyr*A) 1 strain out of 15
		Asp87Gly (*gyr*A mutation)		combination of Thr83Ile and Asp87Gly (*gyr*A) 1 strain out of 15

		Asp87His (*gyr*A mutation)		combination of Thr83Ile and Asp87Asn (*gyr*A) 2 strains out of 15
				combination of Thr83Ile and Asp87His (*gyr*A) 1 strain out of 15
6	20 cystic fibrosis isolates from 6 patients	Thr83Ile (*gyr*A mutation)	11 of 20 isolates had mutations in *gyr*A	Thr83Ile (*gyr*A mutation) 7 strains out of 20
		Asp87Asn (*gyr*A mutation)	19 isolates have efflux system mutations	Asp87Asn (*gyr*A mutation) 3 strains out of 20
		Asp87Tyr (*gyr*A mutation)	no mutations in parC or *mex*R in any isolate	Asp87Tyr (*gyr*A mutation) 1 strain out of 20
7	513 clinical isolates 150 levofloxacin intermediate and resistant isolates 127 ciprofloxacin resistant isolates	**23 types of mutations**		
		112 out of 150 isolates levofloxacin non susceptible		
		Thr83Ala (*gyr*A mutation) 1isolate		
		Asp87Asn (*gyr*A mutation) 16 isolates		
		Asp87Tyr (*gyr*A mutation) 2 isolates		
		Asp87Gly (*gyr*A mutation) 2 isolates		
		Ser87Leu (parC mutation) 75 isolates		
		Ser87Trp (*par*C mutation) 9isolates		
		Glu91Lys (parC mutation)1isolate		
		Ala88Pro (*par*C mutation) 1isolate		
		Leu95Gln (*par*C mutation)1isolate		
		Glu470Asp (*gyr*B mutation) 13 isolates		
		Gln469Val (*gyr*B mutation) 1isolate		
		Ser468Tyr (*gyr*B mutation) 1isolate		
		Ser468Phe (*gyr*B mutation) 3 isolates		
		Gln459Arg (*gyr*B mutation) 1isolate		
		Thr-473 Met (gyrB mutation) 1 isolate		
		Ala477 Val (*gyr*B mutation) 7 isolates		

<table>
<tr><td></td><td></td><td colspan="3">Ala473Val (parE mutation) 1 isolate</td></tr>
<tr><td></td><td></td><td colspan="3">Glu459Lys (parE mutation) 1isolate</td></tr>
<tr><td></td><td></td><td colspan="3">Ala473Val (parE mutation) 1isolate</td></tr>
<tr><td></td><td></td><td colspan="3">Glu459Val (parE mutation) 3 isolates</td></tr>
<tr><td></td><td></td><td colspan="3">Asp419Asn (parE mutation) 7 isolates</td></tr>
<tr><td></td><td></td><td colspan="3">Ala425Val (parE mutation) 3 isolates</td></tr>
<tr><td rowspan="8">8</td><td></td><td colspan="3">five types of mutations</td></tr>
<tr><td rowspan="7">22 clinical isolates
14 isolates with decreased quinolones susceptibility</td><td>Thr83Ile (gyrA mutation)
16 isolate out of 22</td><td>combination of Thr83Ile (gyrA mutation) with Ser80Leu (parC mutation)</td><td>5 isolates</td></tr>
<tr><td>Asp87Gly (gyrA mutation)
3 isolates out of 22</td><td>combination of Thr83Ile (gyrA mutation) with Asp87Gly (gyrA mutation) with Ser80Leu (parC mutation)</td><td>2 isolates</td></tr>
<tr><td>Asp87Asn (gyrA mutation)
1 isolate out of 22</td><td>combination of Thr83Ile (gyrA mutation) with Asp87Asn (gyrA mutation) with Ser80Leu (parC mutation)</td><td>1 isolate</td></tr>
<tr><td></td><td>combination of Thr83Ile (gyrA mutation) with Glu84Lys (parC mutation)</td><td>2 isolates</td></tr>
<tr><td>Ser80Leu (parC mutation)
8 isolates out of 22</td><td>single Thr83Ile (gyrA mutation)</td><td>6 isolates</td></tr>
<tr><td>Glu84Lys (parC mutation)
2 isolates out of 22</td><td>single Asp87Gly (gyrA mutation)</td><td>1 isolate</td></tr>
<tr><td rowspan="6">9</td><td rowspan="6">30 clinical strains</td><td colspan="3">Thr83Ile (gyrA mutation) 27 isolates out of 30 isolates</td></tr>
<tr><td colspan="3">Asp87Gly (gyrA mutation) 1 isolate out of 30 isolates</td></tr>
<tr><td colspan="3">Ser464Phe (gyrB mutation) 2 isolates out of 30 isolates</td></tr>
<tr><td colspan="3">Ser80Leu (parC mutation) 7 isolates out of 30 isolates</td></tr>
<tr><td colspan="3">Ser80Trp (parC mutation) 2 isolates out of 30 isolates</td></tr>
<tr><td colspan="3">Glu84Lys (parC mutation) 1 isolate</td></tr>
</table>

Supplementary Table 6: Distribution of SNP changes in 40 gentamycin CBG markers among study isolates

ID	CN MIC	AK MIC	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40
PAE0095	**0.25**	**0.5**		T					T		C	G										G																				
PAE0149	**0.25**	**0.5**																				G			A																	
PAE0151	**0.25**	**0.5**		T							C				C	C						G					C															
PAE0060	**0.5**	**2**																																								
PAE0081	**0.5**	**1**					T				C											G		T	A					T	A					T					
PAE0119	**0.5**	**0.5**		T	C	G	T			G	C				C			A			A			T	A		C			T	A				T		T	C			
PAE0004	**0.5**	**0.5**		T							C									G		G				A				T	A	C							T		G
PAE0139	**0.5**	**0.5**					T	T										A		G	A	G			A					T					T						
PAE0018	**1**	**2**	C				T	T					T	T	C	C		A		G	A	G			A		C	A					T	C			T			G	
PAE0041	**1**	**0.5**		T							C							A			A	G						A	A			C					T				G
PAE0057	**1**	**2**				G	T		T		C					C				G		G								T	A	C	T	C			G		C		G	
PAE0099	**1**	**1**		T							C				C			A	T			G	T	T	A										T						
PAE0161	**1**	**1**		T							C									G		G				A				T	A	C							T		G
PAE0174	**1**	**2**			C		T			G	C			T		C											C			T	A	C	T	C		T	G	T	C		G	
PAE0033	**1**	**4**		T							C									G		G				A				T	A	C							T		G
PAE0037	**1**	**1**		T							C									G		G				A				T	A	C							T		G
PAE0050	**1**	**1**									C	G		T				A		G			T					A	A			C								
PAE0054	**1**	**0.5**		T											C		C		T	G	A	G				A		A	A	T	A				A				C	T		
PAE0056	**1**	**0.5**		T													C		T			G				A						C			A				C			
PAE0058	**1**	**1**		T							C				C			A	T		A				A						C			A				C			
PAE0064	**1**	**2**		T																		G				A						C			A				C			
PAE0065	**1**	**1**		T	C	G	T					G					C		T						A	A			A			C			A	T				T			
PAE0067	**1**	**2**							T			G							T													C			A				T			
PAE0069	**1**	**1**		T													C		T			G				A						C			A				C			
PAE0075	**1**	**2**					T	T				G						A		G	A	G			A										T		T					
PAE0077	**1**	**1**		T							C				C																	C										
PAE0078	**1**	**2**							T			G							T	G			T									C			A				T			
PAE0086	**1**	**2**									C	G		T				A		G			T					A	A			C										
PAE0105	**1**	**0.5**					T	T						T	C	C		A			A	G								T		C				T		T				
PAE0107	**1**	**0.5**		T							C				C																	C										
PAE0118	**1**	**0.5**		T							C									G		G				A				T	A	C							T		G
PAE0043	**2**	**2**		T				T	T		C				C	C						G													A						
PAE0062	**2**	**1**	C		C	G	T								C	C		A		G	A	G			A		C	A					T	C		T	G	T	C			
PAE0070	**2**	**2**		T								G			C		C		T												A				A				T			
PAE0076	**2**	**2**	C	T				T			C							A		G						A		A	A						A							

PAE0084	2	1			C	G	T	T					T		C			A		G	A			T	A		C	A	A	T			T	C				T				
PAE0090	2	2					T	T		G					C	C				G		G			A					T		C					G	T	C			
PAE0092	2	2						T									C		T	G	A										A				A			T			G	G
PAE0098	2	1		T							C				C			A				G										C				T						
PAE0106	2	2	C		C	G	T	T					T	T	C	C		A		G	A	G			A		C	A	A			C						T	C			
PAE0115	2	2		T	C	G		T		G			T	T				A			A	G						A	A	T								T		T		
PAE0116	2	1		T							C									G	A																		C		G	
PAE0001	2	2																				G			A																	
PAE0009	2	1									C	G		T				A		G			T					A	A			C										
PAE0015	2	1		T											C		C		T	G	A	G	T			A		A	A	T	A				A				C	T		
PAE0017	2	1		T					T		C								T		A	G										C			A							
PAE0019	2	2	C				T											A			A	G			A					T												
PAE0022	2	1	C		C	G	T	T							C			A				G						A	A				T	C				T				
PAE0023	2	2						T	T									A								A						C						T		T		
PAE0031	2	1					T				C	G			C			A	T				T												A							
PAE0034	2	2		T								G						A	T			G										C			A			T		T		
PAE0038	2	1													C			A			A			T	A			A	A					C				T				
PAE0042	2	0.5						T									C		T	G	A										A				A			T			G	G
PAE0045	2	2										G			C		C		T																A	T				T		G
PAE0049	2	2					T	T																				A	A													
PAE0052	2	1													C			A			A			T	A			A	A					C								
PAE0053	2	1		T			T	T						T				A		G	A	G								T						T		T				
PAE0066	2	1		T													C		T			G				A						C			A				C			
PAE0073	2	1					T	T						T	C	C		A		G		G			A		C	A	A										C			
PAE0074	2	2		T					T		C	G						A	T																A							
PAE0080	2	1																			A	G				A									A							
PAE0089	2	2										G						A	T				T					A	A						A					T		
PAE0094	2	2		T	C	G		T			C						C					G				A						C								T		
PAE0101	2	2	C		C	G	T	T							C			A				G						A	A				T	C				T				
PAE0102	2	2					T	T						T	C	C		A			A	G								T		C				T		T				
PAE0110	2	1		T							C	G					C									A					A									T		
PAE0114	2	2					T	T				G						A		G	A	G			A											T		T				
PAE0117	2	2		T	C		T										C			G											A	C	T				G		C		G	
PAE0121	2	1		T													C		T			G				A						C			A				C			
PAE0122	2	1					T	T						T	C	C		A			A	G								T		C				T		T				
PAE0123	2	2										G			C											A									A							
PAE0126	2	1	C	T																		G										C			A				C			
PAE0140	2	2		T							C									G		G				A				T	A	C								T		G

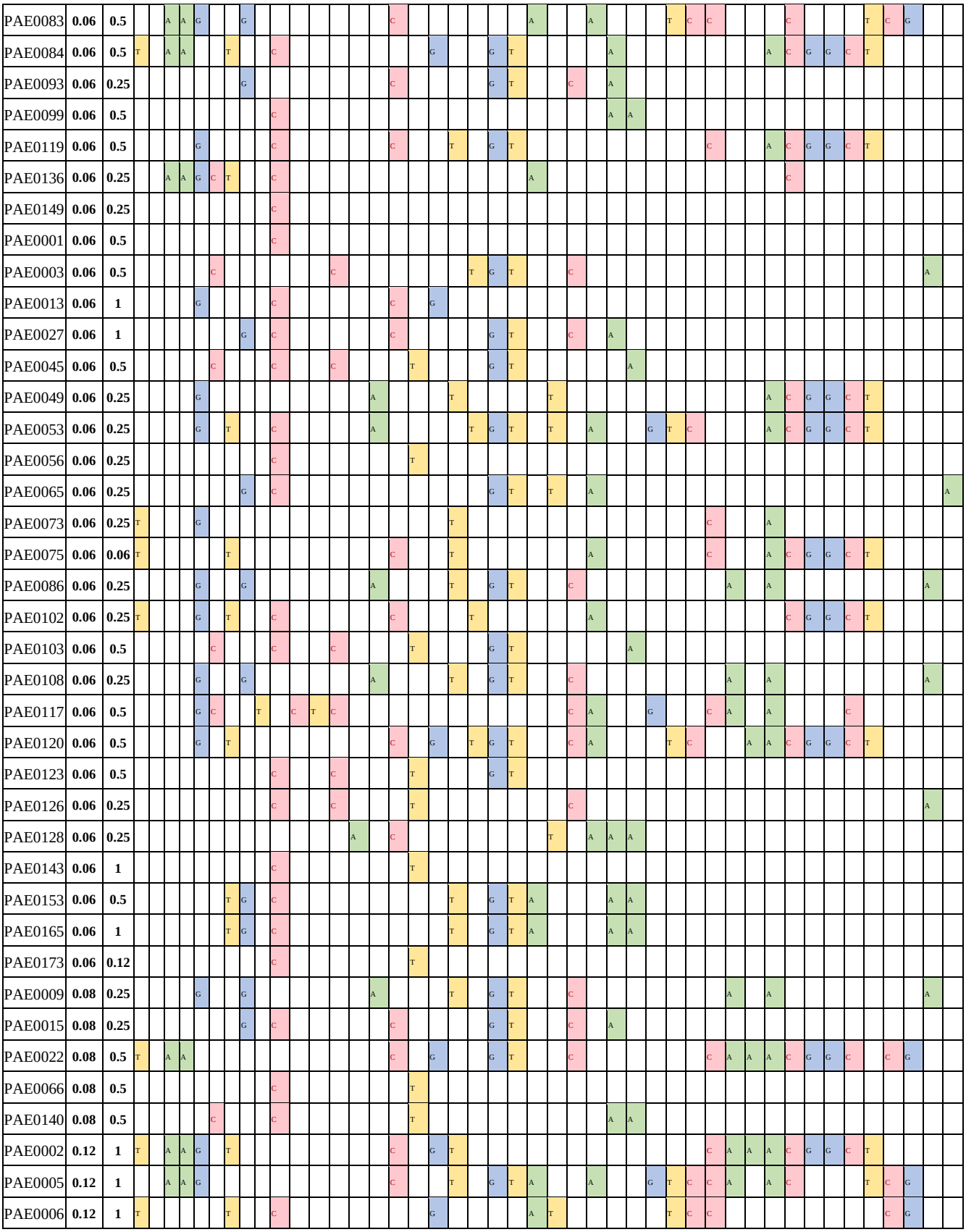
PAE0153	2	2		T								G					C			G						A					A	C			A							
PAE0155	2	2																A			A	G				A						C			A							
PAE0007	4	2		T				T				G																			T	A	C										
PAE0008	4	4		T								G					C		T														T	C						T		
PAE0020	4	2	C					T							C	C		A			A	G	T		A			A	A	T			T	C								
PAE0021	4	2	C		C			T							C	C		A							A			A	A					C								
PAE0025	4	2						T	T		C															A						C								T		
PAE0026	4	4		T							C	G									A	G	T					A	A			C			A					T		
PAE0030	4	2							T			G						A	T		A	G										C										
PAE0036	4	2		T					T		C				C			A														C			A					T		
PAE0040	4	4		T							C	G			C							G										C			A			T				
PAE0068	4	4		T					T					T			C		T													C			A							
PAE0071	4	4		T			T	T						T				A		G	A	G								T						T		T				
PAE0082	4	2					T					G						A		G	A	G			A											T		T				
PAE0083	4	2			C	G	T	T										A			A	G																T	C			
PAE0087	4	4		T							C	G					C					G									A	C			A					T		
PAE0093	4	4		T											C		C		T	G	A	G				A		A	A	T	A				A				C	T		
PAE0096	4	2		T											C		C		T	G	A							A	A			C				T				T		G
PAE0112	4	2						T						T	C	C		A			A	G			A					T						T		T				
PAE0147	4	4													C			A			A			T	A			A	A					C				T				
PAE0148	4	4					T	T						T	C	C		A		G		G			A		C	A	A									T	C			
PAE0158	4	2		T	C	G	T					G					C		T						A	A			A			C			A	T				T		
PAE0172	4	4		T							C	G			C			A		G												C			A							
PAE0003	4	2		T																G	A					A						C								T		G
PAE0013	4	2					T	T		G					C	C				G		G			A					T		C					G	T	C			
PAE0027	4	2		T											C		C		T	G	A	G				A		A	A	T	A				A				C	T		
PAE0103	4	1										G			C		C		T																A	T				T		G
PAE0108	4	1									C	G		T				A		G			T					A	A			C										
PAE0120	4	2													C			A			A			T	A			A	A					C				T				
PAE0127	4	2					T	T										A		G	A	G			A					T						T						
PAE0128	4	2							T			G							T													C			A					T		
PAE0137	4	2		T	C	G		T			C						C					G				A						C										
PAE0138	4	2		T								G					C			G						A					A	C			A					T		
PAE0141	4	4						T									C			G	A										A				A			T			G	G
PAE0143	4	4		T													C					G				A						C			A				C			
PAE0150	4	2					T	T										A		G	A	G			A					T						T						
PAE0152	4	2		T								G					C			G						A					A	C			A							
PAE0163	4	2		T					T			G					C						T			A						C			A							

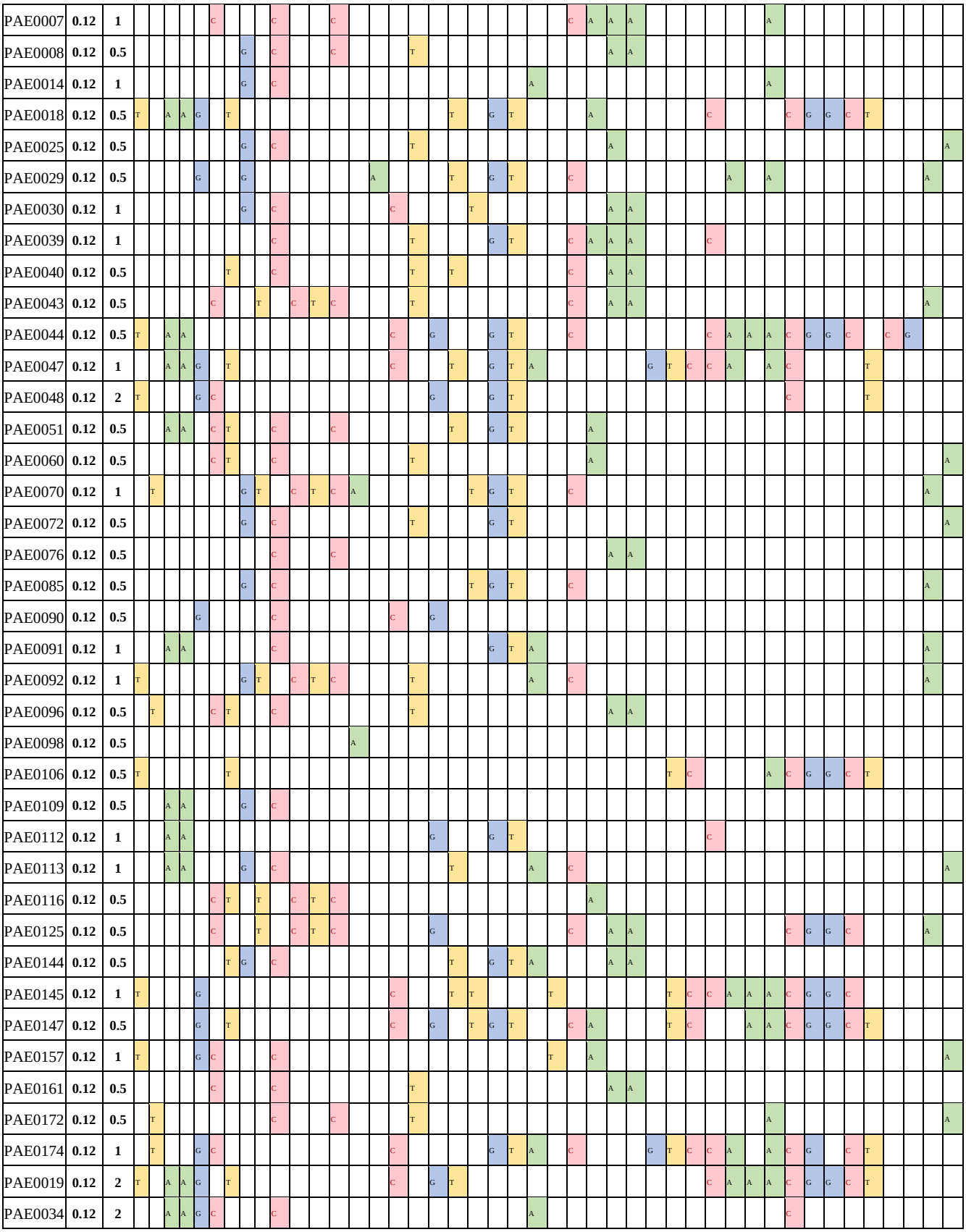
PAE0165	4	2		T								G					C			G						A					A	C			A							
PAE0002	8	4	C				T											A			A	G			A					T												
PAE0005	8	4	C				T								C	C				G	A																T		T		
PAE0011	8	4																A		G	A				A		C					C						T				
PAE0012	8	4							T			G							T	G			T									C			A					T		
PAE0014	8	4		T							C						C		T	G															A					T		
PAE0029	8	4									C	G		T				A		G			T					A	A			C										
PAE0032	8	4		T														A														C			A			T				
PAE0039	8	4		T			T				C	G			C		C		T			G					C					C			A							
PAE0044	8	4	C		C	G	T	T							C			A				G						A	A				T	C				T				
PAE0047	8	4	C		C	G	T								C	C		A		G	A	G			A			A	A		A		T	C		T		T			G	
PAE0048	8	4					T	T										A		G	A	G					C	A	A	T		C					G	T	C			
PAE0051	8	4		T				T	T			G																							A					T		
PAE0055	8	4					T	T						T	C	C		A			A	G								T		C				T		T				
PAE0063	8	4		T	C	G		T			C					C		A				G				A						C								T		
PAE0072	8	2		T				T			C																	A	A		A	C			A							
PAE0085	8	4	C															A			A					A				T	A											
PAE0091	8	4		T					T		C								T		A	G										C			A							
PAE0109	8	4																		G						A									A							
PAE0111	8	4						T										A		G	A	G								T						T						
PAE0113	8	2		T					T		C				C			A							A	A		A				C									G	
PAE0125	8	4	C					T										A	T			G				A									A							
PAE0136	8	4										G						A	T			G										C			A			T		T		
PAE0144	8	4		T								G					C		T	G						A					A	C			A							
PAE0154	8	8																A			A	G				A						C			A							
PAE0160	8	4		T			T		T			G					C		T													C			A					T		
PAE0097	8	4		T				T				G						A			A											C			A	T						
PAE0159	8	4										G			C		C																		A	T				T		G
PAE0162	8	4					T				C				C																	C			A			T				
PAE0166	8	4					T	T										A		G	A	G			A					T						T						
PAE0006	16	4						T					T	T	C			A			A	G			A		C	A	A													
PAE0010	16	8							T		C	G					C		T				T									C			A					T		
PAE0024	16	8					T	T																				A	A													
PAE0035	16	16		T					T		C	G						A	T				T												A					T		
PAE0046	16	8										G					C		T	G			T			A						C								T		
PAE0059	16	8									C															A						C			A							
PAE0061	16	8					T																															T				
PAE0088	16	8		T													C						T									C										

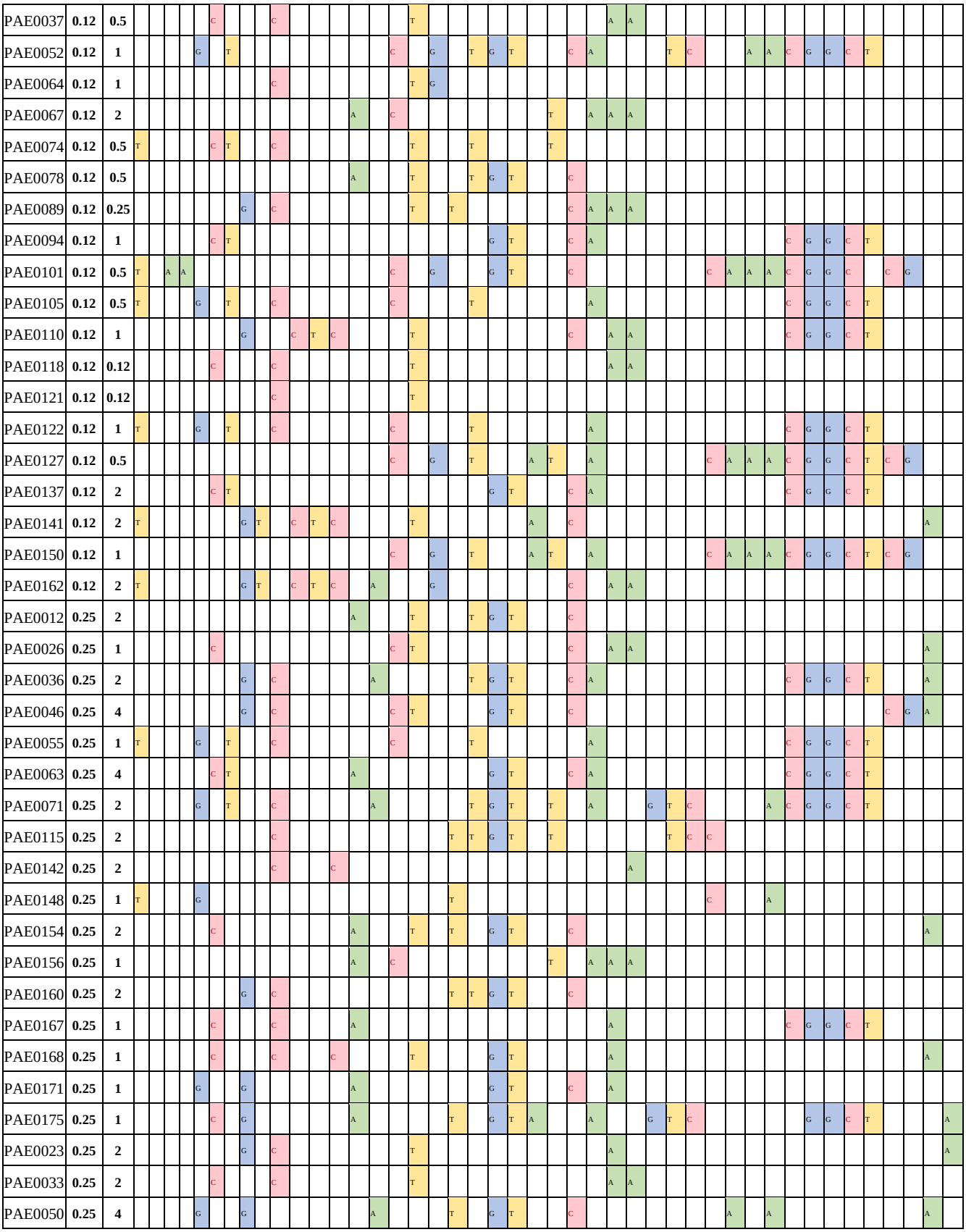
PAE0124	**16**	**8**		T			T		T		C	G			C		C		T				T			A						C			A					T		
PAE0142	**16**	**4**		T					T			G						A	T													C			A							
PAE0145	**16**	**4**													C			A			A	G										C						T				
PAE0156	**16**	**4**							T			G							T													C			A					T		
PAE0167	**16**	**8**		T					T			G					C		T				T			A						C			A							
PAE0171	**16**	**8**		T							C	G					C															C			A							
PAE0169	**16**	**16**										G			C		C																		A	T				T		G
PAE0173	**16**	**8**		T													C					G				A						C			A				C			
PAE0157	**32**	**32**									C							A					T																	T		
PAE0170	**32**	**32**		T													C			G															A							
PAE0168	**64**	**32**							T	G	C						C		T				T			A						C			A					T		
PAE0175	**64**	**32**		T															T							A						C			A			T		T		
PAE0016	**64**	**2**						T	T		C															A						C								T		

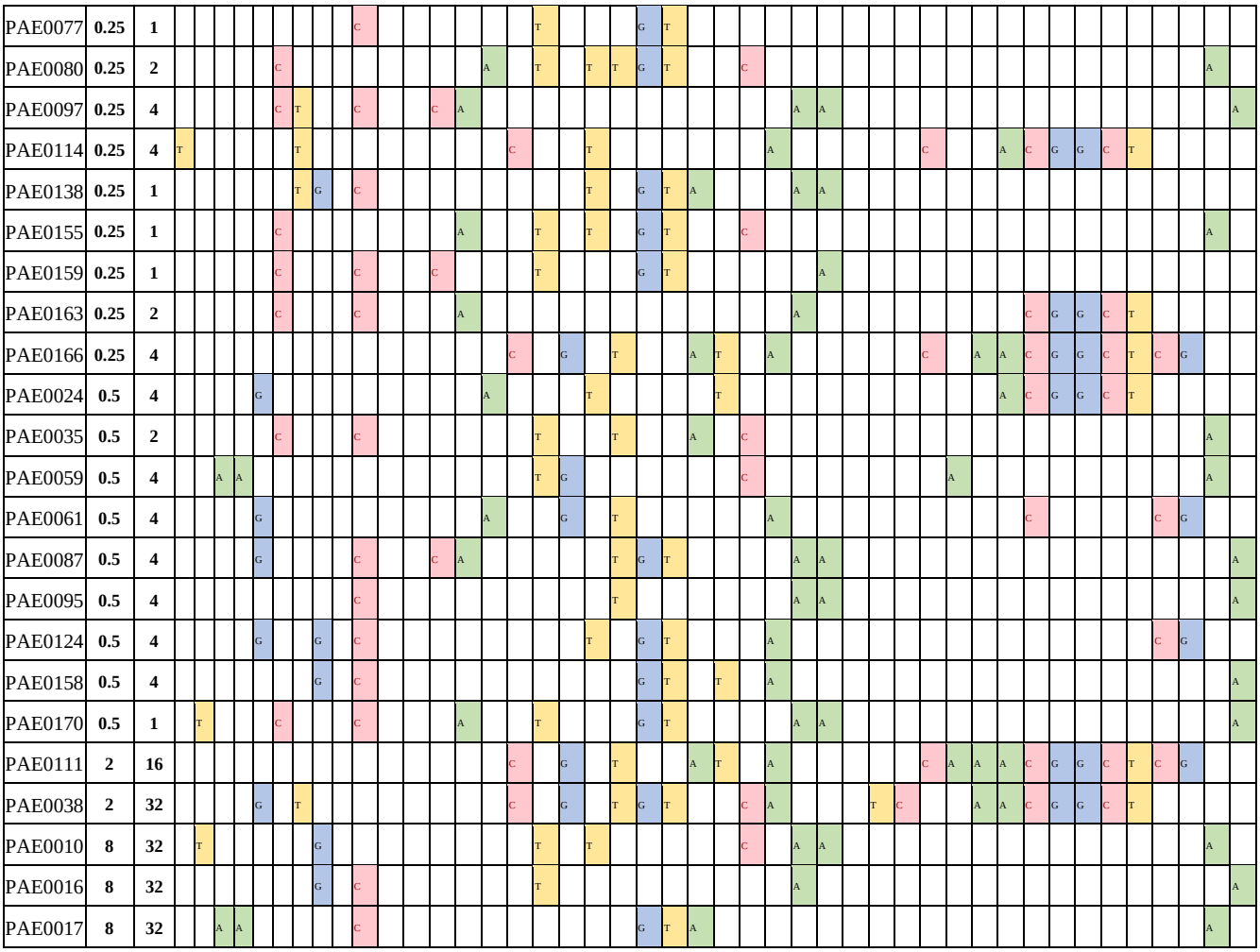
Supplementary Table 7: Distribution of SNP changes in 44 ciprofloxacin CBG markers among study isolates

	Cipro MIC	**Levo MIC**	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40	M41	M42	M43	M44
PAE0151	**0.008**	**0.06**																									C		A	A																A
PAE0068	**0.03**	**0.25**						C				C							T																											
PAE0088	**0.03**	**0.12**						C																		T																				A
PAE0004	**0.03**	**0.25**						C				C							T										A	A																
PAE0031	**0.03**	**0.25**						C	T		T		C	T	C	A			T		T																									A
PAE0042	**0.03**	**0.5**	T							G	T				C				T						A		C																		A	
PAE0054	**0.03**	**0.5**								G		C						C					G	T			C		A																	
PAE0058	**0.03**	**0.25**									T		C	T	C																															A
PAE0069	**0.03**	**0.5**										C							T																											
PAE0107	**0.03**	**0.12**										C							T				G	T																						
PAE0139	**0.03**	**1**																C		G		T			A	T		A						C	A	A	A	C	G	G	C	T	C	G		
PAE0152	**0.03**	**0.25**							T	G		C									T		G	T	A				A	A																
PAE0169	**0.03**	**0.25**						C				C			C				T				G	T						A															A	
PAE0011	**0.06**	**0.25**	T				G		T			C								G														C			A	C	G	G	C	T				
PAE0020	**0.06**	**0.5**			A	A	G											C		G	T	T	G	T	A	T		A				T	C			A		C	G	G	C	T	C	G		
PAE0021	**0.06**	**0.5**	T																	G		T																								
PAE0032	**0.06**	**0.5**						C			T		C	T	C														A	A																
PAE0041	**0.06**	**0.5**										C						C									C		A	A																
PAE0057	**0.06**	**0.5**		T				C										C				T		T							G	T	C					C					C		A	
PAE0062	**0.06**	**0.5**	T		A	A			T									C					G	T								T	C	C			A	C				T				
PAE0081	**0.06**	**0.25**						C															G	T			C		A	A																A
PAE0082	**0.06**	**0.25**	T						T									C			T							A						C			A	C	G	G	C	T				


PAE0083 0.06 0.5
PAE0084 0.06 0.5
PAE0093 0.06 0.25
PAE0099 0.06 0.5
PAE0119 0.06 0.5
PAE0136 0.06 0.25
PAE0149 0.06 0.25
PAE0001 0.06 0.5
PAE0003 0.06 0.5
PAE0013 0.06 1
PAE0027 0.06 1
PAE0045 0.06 0.5
PAE0049 0.06 0.25
PAE0053 0.06 0.25
PAE0056 0.06 0.25
PAE0065 0.06 0.25
PAE0073 0.06 0.25
PAE0075 0.06 0.06
PAE0086 0.06 0.25
PAE0102 0.06 0.25
PAE0103 0.06 0.5
PAE0108 0.06 0.25
PAE0117 0.06 0.5
PAE0120 0.06 0.5
PAE0123 0.06 0.5
PAE0126 0.06 0.25
PAE0128 0.06 0.25
PAE0143 0.06 1
PAE0153 0.06 0.5
PAE0165 0.06 1
PAE0173 0.06 0.12
PAE0009 0.08 0.25
PAE0015 0.08 0.25
PAE0022 0.08 0.5
PAE0066 0.08 0.5
PAE0140 0.08 0.5
PAE0002 0.12 1
PAE0005 0.12 1
PAE0006 0.12 1







BIBLIOGRAPHY

‘A guide to sensitivity testing. Report of the Working Party on Antibiotic Sensitivity Testing of the British Society for Antimicrobial Chemotherapy.’ (1991) *The Journal of antimicrobial chemotherapy*, 27 Suppl D, pp. 1–50.

Abdelraouf, K. *et al.* (2011) ‘Effect of multidrug resistance-conferring mutations on the fitness and virulence of *Pseudomonas aeruginosa*’, *Journal of Antimicrobial Chemotherapy*, 66(6), pp. 1311–1317. doi: 10.1093/jac/dkr105.

Abril, D. *et al.* (2019) ‘Genome plasticity favours double chromosomal Tn4401b-bla KPC-2 transposon insertion in the *Pseudomonas aeruginosa* ST235 clone’, *BMC Microbiology*. BMC Microbiology, 19(1), pp. 1–12. doi: 10.1186/s12866-019-1418-6.

Adewoye, L. *et al.* (2002) 'The *mex*R repressor of the *mex*AB-oprM multidrug efflux operon in *Pseudomonas aeruginosa*: characterization of mutations compromising activity.', *Journal of bacteriology*, 184(15), pp. 4308–12.

Adu-Oppong, B., Gasparrini, A. J. and Dantas, G. (2017) 'Genomic and functional techniques to mine the microbiome for novel antimicrobials and antimicrobial resistance genes', *Annals of the New York Academy of Sciences*, 1388(1), pp. 42–58. doi: 10.1111/nyas.13257.

Akasaka, T. *et al.* (2001) 'Type II topoisomerase mutations in fluoroquinolone-resistant clinical strains of *Pseudomonas aeruginosa* isolated in 1998 and 1999: role of target enzyme in mechanism of fluoroquinolone resistance.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology, 45(8), pp. 2263–8. doi: 10.1128/AAC.45.8.2263-2268.2001.

Al-Tahhan, R. A. *et al.* (2000) 'Rhamnolipid-induced removal of lipopolysaccharide from *Pseudomonas aeruginosa*: Effect on cell surface properties and interaction with hydrophobic substrates', *Applied and Environmental Microbiology*, 66(8), pp. 3262–3268. doi: 10.1128/AEM.66.8.3262-3268.2000.

Alam, M. T. *et al.* (2014) 'Dissecting vancomycin-intermediate resistance in *Staphylococcus aureus* using genome-wide association', *Genome Biology and Evolution*, 6(5), pp. 1174–1185. doi: 10.1093/gbe/evu092.

Aldred, K. J., Kerns, R. J. and Osheroff, N. (2014) 'Mechanism of quinolone action and resistance.', *Biochemistry*. American Chemical Society, 53(10), pp. 1565–74. doi: 10.1021/bi5000564.

Allcock, R. J. N., Jennison, A. V and Warrilow, D. (2017) 'Towards a Universal Molecular Microbiological Test', *Journal of Clinical Microbiology*. Edited by C. S. Kraft, 55(11), p. 3175 LP-3182. doi: 10.1128/JCM.01155-17.

Alonso, A., Rojo, F. and Martínez, J. L. (1999) 'Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin.', *Environmental microbiology*, 1(5), pp. 421–30.

Alonso, A., Sánchez, P. and Martínez, J. L. (2001) 'Environmental selection of antibiotic resistance genes.', *Environmental microbiology*, 3(1), pp. 1–9.

Altschul, S. F. *et al.* (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Álvarez-Lerma, F., Grau, S. and Álvarez-Beltrán, M. (2006) 'Levofloxacin in the treatment of ventilator-associated pneumonia', *Clinical Microbiology and Infection*, pp. 81–92. doi: 10.1111/j.1469-0691.2006.01399.x.

Alvarez-Ortega, C. and Harwood, C. S. (2007) 'Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration', *Molecular Microbiology*, 65(1), pp. 153–165. doi: 10.1111/j.1365-2958.2007.05772.x.

Andersson, D. I. and Hughes, D. (2010) 'Antibiotic resistance and its cost: is it possible to reverse resistance?', *Nature Reviews Microbiology*, 8(4), pp. 260–271. doi: 10.1038/nrmicro2319.

Andersson, P. *et al.* (2013) 'Sequences of multiple bacterial genomes and a *Chlamydia trachomatis* genotype from direct sequencing of DNA derived from a vaginal swab diagnostic specimen', *Clinical Microbiology and Infection*, 19(9), pp. E405–E408. doi: 10.1111/1469-0691.12237.

Andrews, J. M. (2001) 'Determination of minimum inhibitory concentrations', *Journal of Antimicrobial Chemotherapy*. Narnia, 48(suppl_1), pp. 5–16. doi: 10.1093/jac/48.suppl_1.5.

Andrews, J. M. and BSAC Working Party on Susceptibility Testing, for the (2001) 'BSAC standardized disc susceptibility testing method', *Journal of Antimicrobial Chemotherapy*, 48(suppl_1), pp. 43–57. doi: 10.1093/jac/48.suppl_1.43.

Andrews, J. M. and Howe, R. A. (2011) 'BSAC standardized disc susceptibility testing method (version 10)', *Journal of Antimicrobial Chemotherapy*. Oxford University Press, 66(12), pp. 2726–2757. doi: 10.1093/jac/dkr359.

'Antibiotic susceptibility diagnostics for the future' (2019) *Nature Microbiology*, 4(10), p. 1603. doi: 10.1038/s41564-019-0577-4.

Arca, P., Hardisson, C. and Suárez, J. E. (1990) 'Purification of a glutathione S-transferase that mediates fosfomycin resistance in bacteria.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 34(5), pp. 844–8.

Austin, A., Lietman, T. and Rose-Nussbaumer, J. (2017) 'Update on the Management of Infectious Keratitis', *Ophthalmology*. Elsevier, 124(11), pp. 1678–1689. doi: 10.1016/j.ophtha.2017.05.012.

Azucena, E. and Mobashery, S. (2001) 'Aminoglycoside-modifying enzymes: mechanisms of catalytic processes and inhibition', *Drug Resistance Updates*, 4(2), pp. 106–117. doi: 10.1054/drup.2001.0197.

Bacher, J., Wenzig, K. and Vogler, M. (2004) 'SPSS TwoStep Cluster - a first evaluation', *2004-2*. DEU, 2004–2, p. 23.

Bains, M., Fernández, L. and Hancock, R. E. W. (2012) 'Phosphate Starvation Promotes Swarming Motility and Cytotoxicity of *Pseudomonas aeruginosa*, *Applied and Environmental Microbiology*, 78(18), p. 6762 LP-6768. doi: 10.1128/AEM.01015-12.

Bankevich, A. *et al.* (2012) 'SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing', *Journal of Computational Biology*, 19(5), pp. 455–477. doi: 10.1089/cmb.2012.0021.

Baquero, F., Alvarez-Ortega, C. and Martinez, J. L. (2009) 'Ecology and evolution of antibiotic resistance', *Environmental Microbiology Reports*. John Wiley & Sons, Ltd (10.1111), 1(6), pp. 469–476. doi: 10.1111/j.1758-2229.2009.00053.x.

Baquero, F., Coque, T. M. and de la Cruz, F. (2011) 'Ecology and evolution as targets: the need for novel eco-evo drugs and strategies to fight antibiotic resistance.', *Antimicrobial agents and chemotherapy*, 55(8), pp. 3649–60. doi: 10.1128/AAC.00013-11.

Barbosa, C. *et al.* (2017) 'Alternative Evolutionary Paths to Bacterial Antibiotic Resistance Cause Distinct Collateral Effects', *Molecular Biology and Evolution*, 34(9), pp. 2229–2244. doi: 10.1093/molbev/msx158.

Barbosa, C. *et al.* (2019) 'Evolutionary stability of collateral sensitivity to antibiotics in the model pathogen *Pseudomonas aeruginosa*', *eLife*. Edited by P. J. Wittkopp et al. eLife Sciences Publications, Ltd, 8, p. e51481. doi: 10.7554/eLife.51481.

Bardou, P. *et al.* (2014) 'jvenn: an interactive Venn diagram viewer', *BMC Bioinformatics*. BioMed Central, 15(1), p. 293. doi: 10.1186/1471-2105-15-293.

Bauer, K. A. *et al.* (2014) 'Review of rapid diagnostic tests used by antimicrobial stewardship programs', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 59(Suppl 3), pp. S134–S145. doi: 10.1093/cid/ciu547.

Bayot ML, B. B. (2020) *Antimicrobial Susceptibility Testing. [Updated 2020 Mar 31].*, *StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-*. Available at: https://www.ncbi.nlm.nih.gov/books/NBK539714/.

Becker, B. and Cooper, M. A. (2013) 'Aminoglycoside Antibiotics in the 21st Century', *ACS Chemical Biology*, 8(1), pp. 105–115. doi: 10.1021/cb3005116.

van Beek, J. *et al.* (2019) 'Tracing local and regional clusters of carbapenemase-producing *Klebsiella pneumoniae* ST512 with whole genome sequencing, Finland, 2013 to 2018', *Eurosurveillance*, 24(38), pp. 1–9. doi: 10.2807/1560-7917.ES.2019.24.38.1800522.

Bej, A. K. *et al.* (1996) 'Detection of viable *Vibrio cholerae* by reverse-transcriptase polymerase chain reaction (RT-PCR)', *Molecular Biotechnology*. Humana Press, 5(1), pp. 1–10. doi: 10.1007/BF02762407.

Bel Hadj Ahmed, A. *et al.* (2019) 'Characterization of *Pseudomonas aeruginosa* isolated from various environmental niches: New STs and occurrence of antibiotic susceptible "high-risk clones"', *International Journal of Environmental Health Research*. Taylor & Francis, pp. 1–10. doi: 10.1080/09603123.2019.1616080.

Van Belkum, A. *et al.* (2001) 'Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology', *Clinical Microbiology Reviews*. doi: 10.1128/CMR.14.3.547-560.2001.

Belkum, A. van and Dunne, W. M. (2013) 'Next-Generation Antimicrobial Susceptibility Testing', *Journal of Clinical Microbiology*. American Society for Microbiology Journals, 51(7), pp. 2018–2024. doi: 10.1128/JCM.00313-13.

Bert, F. and Lambert-Zechovsky, N. (1996) 'Comparative distribution of resistance patterns and serotypes in *Pseudomonas aeruginosa* isolates from intensive care units and other wards.', *The Journal of antimicrobial chemotherapy*, 37(4), pp. 809–13.

Berthenet, E. *et al.* (2018) 'A GWAS on *Helicobacter pylori* strains points to genetic variants associated with gastric cancer risk', *BMC Biology*. BMC Biology, 16(1), pp. 1–11. doi:

10.1186/s12915-018-0550-3.

Besant, J. D., Sargent, E. H. and Kelley, S. O. (2015) 'Rapid electrochemical phenotypic profiling of antibiotic-resistant bacteria.', *Lab on a chip*, 15(13), pp. 2799–807. doi: 10.1039/c5lc00375j.

Besser, J. *et al.* (2018) 'Next-generation sequencing technologies and their application to the study and control of bacterial infections', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2017/10/23, 24(4), pp. 335–341. doi: 10.1016/j.cmi.2017.10.013.

Bhullar, K. *et al.* (2012) 'Antibiotic resistance is prevalent in an isolated cave microbiome.', *PloS one*. Edited by R. K. Aziz, 7(4), p. e34953. doi: 10.1371/journal.pone.0034953.

Bille, E. *et al.* (2005) 'A chromosomally integrated bacteriophage in invasive meningococci', *Journal of Experimental Medicine*, 201(12), pp. 1905–1913. doi: 10.1084/jem.20050112.

Bille, E. *et al.* (2008) 'Association of a bacteriophage with meningococcal disease in young adults', *PLoS ONE*, 3(12). doi: 10.1371/journal.pone.0003885.

Blair, J. M. A. *et al.* (2015) 'Molecular mechanisms of antibiotic resistance', *Nature Reviews Microbiology*. Nature Publishing Group, 13(1), pp. 42–51. doi: 10.1038/nrmicro3380.

Blázquez, J. *et al.* (2006) 'PBP3 inhibition elicits adaptive responses in *Pseudomonas aeruginosa*', *Molecular Microbiology*, 62(1), pp. 84–99. doi: 10.1111/j.1365-2958.2006.05366.x.

Bolard, A., Plésiat, P. and Jeannot, K. (2018) 'Mutations in Gene *fus*A1 as a Novel Mechanism of Aminoglycoside Resistance in Clinical Strains of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology Journals, 62(2), pp. e01835-17. doi: 10.1128/AAC.01835-17.

Bonfiglio, G. (2001) 'Is levofloxacin as active as ciprofloxacin against *Pseudomonas aeruginosa*?', *Chemotherapy*. Switzerland, 47(4), pp. 239–242. doi: 10.1159/000048529.

Borovinskaya, M. A. *et al.* (2007) 'Structural basis for aminoglycoside inhibition of bacterial ribosome recycling', *Nature Structural & Molecular Biology*, 14(8), pp. 727–732. doi:

10.1038/nsmb1271.

Bossuyt, P. M. *et al.* (2006) 'Comparative accuracy: assessing new tests against existing diagnostic pathways.', *BMJ (Clinical research ed.)*. BMJ Publishing Group, 332(7549), pp. 1089–92. doi: 10.1136/bmj.332.7549.1089.

Botes, J. *et al.* (2003) 'Genomic typing of *Pseudomonas aeruginosa* isolates by comparison of Riboprinting and PFGE: correlation of experimental results with those predicted from the complete genome sequence of isolate PAO1', *Journal of Microbiological Methods*. Elsevier, 55(1), pp. 231–240. doi: 10.1016/S0167-7012(03)00156-8.

Bourdon, N. *et al.* (2010) 'Rapid detection of vancomycin-resistant enterococci from rectal swabs by the Cepheid Xpert vanA/vanB assay', *Diagnostic Microbiology and Infectious Disease*, 67(3), pp. 291–293. doi: 10.1016/j.diagmicrobio.2010.02.009.

Boutoille, D. *et al.* (2004) 'Detection of an IS21 insertion sequence in the *mex*R gene of *Pseudomonas aeruginosa* increasing beta-lactam resistance.', *FEMS microbiology letters*, 230(1), pp. 143–6.

Bowles, E. C. *et al.* (2011) 'Xpert MTB/RIF®, a novel automated polymerase chain reaction–based tool for the diagnosis of tuberculosis [Technical note]', *The International Journal of Tuberculosis and Lung Disease*, 15(7), pp. 988–989. doi: 10.5588/ijtld.10.0574.

Boykin, L. M. *et al.* (2018) 'Real time portable genome sequencing for global food security', *bioRxiv*. Cold Spring Harbor Laboratory, p. 314526. doi: 10.1101/314526.

Bradley, P. *et al.* (2015) 'Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*', *Nature communications*. Nature Publishing Group, 6, p. 10063. doi: 10.1038/ncomms10063.

Brazas, M. D. and Hancock, R. E. W. (2005) 'Ciprofloxacin induction of a susceptibility determinant in *Pseudomonas aeruginosa.*', *Antimicrobial agents and chemotherapy*. American Society for Microbiology Journals, 49(8), pp. 3222–7. doi: 10.1128/AAC.49.8.3222-3227.2005.

Brbić, M. *et al.* (2016) 'The landscape of microbial phenotypic traits and associated genes.', *Nucleic acids research*. Oxford University Press, 44(21), pp. 10074–10090. doi: 10.1093/nar/gkw964.

Bredenbruch, F. *et al.* (2006) 'The *Pseudomonas aeruginosa* quinolone signal (PQS) has an iron-chelating activity', *Environmental Microbiology*, 8(8), pp. 1318–1329. doi: 10.1111/j.1462-2920.2006.01025.x.

Breidenstein, E. B. M. *et al.* (2008) 'Complex ciprofloxacin resistome revealed by screening a Pseudomonas aeruginosa mutant library for altered susceptibility.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 52(12), pp. 4486–91. doi: 10.1128/AAC.00222-08.

Breidenstein, E. B. M., Bains, M. and Hancock, R. E. W. (2012) 'Involvement of the lon protease in the SOS response triggered by ciprofloxacin in *Pseudomonas aeruginosa* PAO1.', *Antimicrobial agents and chemotherapy*, 56(6), pp. 2879–87. doi: 10.1128/AAC.06014-11.

Briales, A. *et al.* (2012) 'Prevalence of plasmid-mediated quinolone resistance determinants qnr and aac(6')-Ib-cr in *Escherichia coli* and *Klebsiella pneumoniae* producing extended-spectrum β-lactamases in Spain.', *International journal of antimicrobial agents*, 39(5), pp. 431–4. doi: 10.1016/j.ijantimicag.2011.12.009.

Brokopp, C. D., Gomez Lus, R. and Farmer, J. J. (1977) 'Serological typing of *Pseudomonas aeruginosa*: use of commercial antisera and live antigens', *Journal of Clinical Microbiology*, 5(6), pp. 640–649.

Bruchmann, S. *et al.* (2013) 'Quantitative Contributions of Target Alteration and Decreased Drug Accumulation to *Pseudomonas aeruginosa* Fluoroquinolone Resistance', *Antimicrobial Agents and Chemotherapy*, 57(3), pp. 1361–1368. doi: 10.1128/AAC.01581-12.

Bryan, L. E., O'Hara, K. and Wong, S. (1984) 'Lipopolysaccharide changes in impermeability-type aminoglycoside resistance in *Pseudomonas aeruginosa.*', *Antimicrobial agents and chemotherapy*, 26(2), pp. 250–5. doi: 10.1128/aac.26.2.250.

Brzozowski, M. and Jursa-kulesza, J. (2020) 'Genotypic characterisation and antimicrobial resistance of *Pseudomonas aeruginosa* strains isolated from patients of different hospitals and medical centres from Poland', *BMC Infectious Diseases*, pp. 1–19. doi: 10.21203/rs.3.rs-18693/v1.

Burnham, C.-A. D. *et al.* (2017) 'Diagnosing antimicrobial resistance', *Nature Reviews*

Microbiology. Nature Publishing Group, 15(11), pp. 697–703. doi: 10.1038/nrmicro.2017.103

Caballero, J. D. D. *et al.* (2014) '59 Population structure and antimicrobial susceptibility of *Pseudomonas aeruginosa* from a national survey involving 24 cystic fibrosis units in Spain', *Journal of Cystic Fibrosis*, 13, p. S61. doi: 10.1016/S1569-1993(14)60196-X.

Caboche, S. *et al.* (2017) 'MICRA: an automatic pipeline for fast characterization of microbial genomes from high-throughput sequencing data', *Genome Biology*. BioMed Central, 18(1), p. 233. doi: 10.1186/s13059-017-1367-z.

Cabot, G. *et al.* (2012) 'Genetic markers of widespread extensively drug-resistant *Pseudomonas aeruginosa* high-risk clones.', *Antimicrobial agents and chemotherapy*, 56(12), pp. 6349–57. doi: 10.1128/AAC.01388-12.

Cabot, G. *et al.* (2016) 'Evolution of *Pseudomonas aeruginosa* Antimicrobial Resistance and Fitness under Low and High Mutation Rates.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 60(3), pp. 1767–78. doi: 10.1128/AAC.02676-15.

Caille, O. *et al.* (2014) 'Structural and functional characterization of *Pseudomonas aeruginosa* global regulator AmpR.', *Journal of bacteriology*, 196(22), pp. 3890–902. doi: 10.1128/JB.01997-14.

Caliendo, A. M. *et al.* (2013) 'Better tests, better care: improved diagnostics for infectious diseases.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. Oxford University Press, 57 Suppl 3(Suppl 3), pp. S139-70. doi: 10.1093/cid/cit578.

Campbell, F. *et al.* (2018) 'When are pathogen genome sequences informative of transmission events?', *PLOS Pathogens*. Public Library of Science, 14(2), p. e1006885. Available at: https://doi.org/10.1371/journal.ppat.1006885.

Campbell, P. J. *et al.* (2011) 'Molecular Detection of Mutations Associated with First- and Second-Line Drug Resistance Compared with Conventional Drug Susceptibility Testing of *Mycobacterium tuberculosis*', *Antimicrobial Agents and Chemotherapy*, 55(5), pp. 2032–2041. doi: 10.1128/AAC.01550-10.

Cao, L., Srikumar, R. and Poole, K. (2004) 'MexAB-OprM hyperexpression in NalC-type multidrug-resistant *Pseudomonas aeruginosa*: identification and characterization of the *nal*C

gene encoding a repressor of PA3720-PA3719', *Molecular Microbiology*, 53(5), pp. 1423–1436. doi: 10.1111/j.1365-2958.2004.04210.x.

Capriotti, E., Fariselli, P. and Casadio, R. (2005) 'I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure.', *Nucleic acids research*. Oxford University Press, 33(Web Server issue), pp. W306-10. doi: 10.1093/nar/gki375.

Carriço, J. A. *et al.* (2006) 'Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*.', *Journal of clinical microbiology*. American Society for Microbiology Journals, 44(7), pp. 2524–32. doi: 10.1128/JCM.02536-05.

Castaneda-Montes, F. J. *et al.* (2018) 'Population structure of *Pseudomonas aeruginosa* through a MLST approach and antibiotic resistance profiling of a Mexican clinical collection.', *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. Netherlands, 65, pp. 43–54. doi: 10.1016/j.meegid.2018.06.009.

Centers for Disease Control and Prevention (2018) *Biggest Threats and Data | Antibiotic/Antimicrobial Resistance | CDC, November 26, 2018*.

Chai, J. H. *et al.* (2018) 'Cost-benefit analysis of introducing next-generation sequencing (metagenomic) pathogen testing in the setting of pyrexia of unknown origin', *PLOS ONE*. Edited by H. Tse. Public Library of Science, 13(4), p. e0194648. doi: 10.1371/journal.pone.0194648.

Chan, J. P. *et al.* (2019) 'Using Bacterial Transcriptomics to Investigate Targets of Host-Bacterial Interactions in *Caenorhabditis elegans*', *Scientific Reports*, 9(1), pp. 1–12. doi: 10.1038/s41598-019-41452-2.

Chang, G. W. and Chang, J. T. (1975) 'Evidence for the B12-dependent enzyme ethanolamine deaminase in *Salmonella*.', *Nature*, 254(5496), pp. 150–1. doi: 10.1038/254150a0.

Charalampous, T. *et al.* (2018) 'Rapid Diagnosis of Lower Respiratory Infection using Nanopore-based Clinical Metagenomics', *bioRxiv*, p. 387548. doi: 10.1101/387548.

Charalampous, T. *et al.* (2019) 'Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection', *Nature Biotechnology*. Springer US, 37(7), pp. 783–792. doi: 10.1038/s41587-019-0156-5.

Chen, P. E. and Shapiro, B. J. (2015) 'The advent of genome-wide association studies for bacteria', *Current Opinion in Microbiology*. Elsevier Current Trends, 25, pp. 17–24. doi: 10.1016/J.MIB.2015.03.002.

Chen, W. *et al.* (2016) 'Novobiocin binding to NalD induces the expression of the MexAB-OprM pump in *Pseudomonas aeruginosa*.', *Molecular microbiology*, 100(5), pp. 749–58. doi: 10.1111/mmi.13346.

Cheng, G. *et al.* (2013) 'Antibacterial action of quinolones: From target to network', *European Journal of Medicinal Chemistry*, 66, pp. 555–562. doi: 10.1016/j.ejmech.2013.01.057.

Chewapreecha, C. *et al.* (2014) 'Comprehensive Identification of Single Nucleotide Polymorphisms Associated with Beta-lactam Resistance within Pneumococcal Mosaic Genes', *PLoS Genetics*, 10(8). doi: 10.1371/journal.pgen.1004547.

Chiu, C. Y. and Miller, S. A. (2019) 'Clinical metagenomics', *Nature Reviews Genetics*, 20(6), pp. 341–355. doi: 10.1038/s41576-019-0113-7.

Chiu, T. *et al.* (2001) 'A robust and scalable clustering algorithm for mixed type attributes in large database environment', in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '01*. New York, New York, USA: ACM Press, pp. 263–268. doi: 10.1145/502512.502549.

Cho, H. H. *et al.* (2013) 'Prevalence and genetic analysis of multidrug-resistant *Pseudomonas aeruginosa* ST235 isolated from a hospital in Korea, 2008-2012.', *Annals of clinical and laboratory science*, 43(4), pp. 414–9.

Choi, Y. and Chan, A. P. (2015) 'PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels.', *Bioinformatics (Oxford, England)*. Oxford University Press, 31(16), pp. 2745–7. doi: 10.1093/bioinformatics/btv195.

Cholley, P. *et al.* (2011) 'Most multidrug-resistant *Pseudomonas aeruginosa* isolates from hospitals in eastern France belong to a few clonal types.', *Journal of clinical microbiology*. American Society for Microbiology, 49(7), pp. 2578–83. doi: 10.1128/JCM.00102-11.

Cholley, P. *et al.* (2014) 'Population structure of clinical *Pseudomonas aeruginosa* from West and Central African countries', *PLoS ONE*. doi: 10.1371/journal.pone.0107008.

Choudhury, D. *et al.* (2016) 'Premature Termination of MexR Leads to Overexpression of MexAB-OprM Efflux Pump in *Pseudomonas aeruginosa* in a Tertiary Referral Hospital in India', *PLOS ONE*. Edited by B. Lin, 11(2), p. e0149156. doi: 10.1371/journal.pone.0149156.

Chowdhury, P. R. *et al.* (2016) 'Genomic islands 1 and 2 play key roles in the evolution of extensively drug-resistant ST235 isolates of *Pseudomonas aeruginosa*', *Open Biology*. Royal Society of London, 6(3). doi: 10.1098/rsob.150175.

Chuanchuen, R. *et al.* (2001) 'Cross-resistance between triclosan and antibiotics in *Pseudomonas aeruginosa* is mediated by multidrug efflux pumps: exposure of a susceptible mutant strain to triclosan selects *nfx*B mutants overexpressing MexCD-OprJ.', *Antimicrobial agents and chemotherapy*, 45(2), pp. 428–32. doi: 10.1128/AAC.45.2.428-432.2001.

Cirz, R. T. *et al.* (2006) 'Defining the *Pseudomonas aeruginosa* SOS response and its role in the global response to the antibiotic ciprofloxacin.', *Journal of bacteriology*, 188(20), pp. 7101–10. doi: 10.1128/JB.00807-06.

Clausen, P. T. L. C. *et al.* (2016) 'Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data', *Journal of Antimicrobial Chemotherapy*, 71(9), pp. 2484–2488. doi: 10.1093/jac/dkw184.

Clerc, O. and Greub, G. (2010) 'Routine use of point-of-care tests: usefulness and application in clinical microbiology', *Clinical Microbiology and Infection*, 16(8), pp. 1054–1061. doi: 10.1111/j.1469-0691.2010.03281.x.

Cock, P. J. A. *et al.* (2015) 'NCBI BLAST+ integrated into Galaxy', *GigaScience*. Narnia, 4(1), p. 39. doi: 10.1186/s13742-015-0080-7.

Cole, B. J. *et al.* (2017) 'Genome-wide identification of bacterial plant colonization genes', *PLoS Biology*, 15(9), pp. 1–24. doi: 10.1371/journal.pbio.2002860.

Collins, C. and Didelot, X. (2018) 'A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination', *PLOS Computational Biology*. Edited by A. C. McHardy. Public Library of Science, 14(2), p. e1005958. doi: 10.1371/journal.pcbi.1005958.

Committee, J. F. (2020) 'British National Formulary (online) London', in *BMJ Group and*

Pharmaceutical Press. BMJ Group and Pharmaceutical Press. Available at: http://www.medicinescomplete.com.

Correa, A. *et al.* (2015) 'Dissemination of high-risk clones of extensively drug-resistant *Pseudomonas aeruginosa* in Colombia', *Antimicrobial Agents and Chemotherapy*. American Society for Microbiology, 59(4), pp. 2421–2425. doi: 10.1128/AAC.03926-14.

Correia, S. *et al.* (2017) 'Mechanisms of quinolone action and resistance: where do we stand?', *Journal of Medical Microbiology*, 66(5), pp. 551–559. doi: 10.1099/jmm.0.000475.

Craw, P. and Balachandran, W. (2012) 'Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review', *Lab on a Chip*, 12(14), p. 2469. doi: 10.1039/c2lc40100b.

Creecy, J. P. and Conway, T. (2015) 'Quantitative bacterial transcriptomics with RNA-seq', *Current Opinion in Microbiology*. Elsevier Ltd, 23, pp. 133–140. doi: 10.1016/j.mib.2014.11.011.

Curran, B. *et al.* (2004) 'Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa.*', *Journal of clinical microbiology*. American Society for Microbiology, 42(12), pp. 5644–9. doi: 10.1128/JCM.42.12.5644-5649.2004.

Cuzon, G. *et al.* (2012) 'Evaluation of a DNA microarray for the rapid detection of extended-spectrum -lactamases (TEM, SHV and CTX-M), plasmid-mediated cephalosporinases (CMY-2-like, DHA, FOX, ACC-1, ACT/MIR and CMY-1-like/MOX) and carbapenemases (KPC, OXA-48, VIM, IMP and NDM)', *Journal of Antimicrobial Chemotherapy*, 67(8), pp. 1865–1869. doi: 10.1093/jac/dks156.

D'Costa, V. M. *et al.* (2011) 'Antibiotic resistance is ancient.', *Nature*, 477(7365), pp. 457–61. doi: 10.1038/nature10388.

Dancer, S. J. *et al.* (2019) 'Tracking *Staphylococcus aureus* in the intensive care unit using whole-genome sequencing', *Journal of Hospital Infection*. W.B. Saunders, 103(1), pp. 13–20. doi: 10.1016/J.JHIN.2019.04.016.

Davenport, E. R. *et al.* (2015) 'Genome-wide association studies of the human gut microbiota', *PLoS ONE*, 10(11), pp. 1–22. doi: 10.1371/journal.pone.0140301.

Davey, L. and Valdivia, R. H. (2020) 'ScienceDirect Bacterial genetics and molecular pathogenesis in the age of high throughput DNA sequencing', *Current Opinion in Microbiology*. Elsevier Ltd, 54, pp. 59–66. doi: 10.1016/j.mib.2020.01.007.

Davis, B. D., Chen, L. L. and Tai, P. C. (1986) 'Misread protein creates membrane channels: an essential step in the bactericidal action of aminoglycosides.', *Proceedings of the National Academy of Sciences*, 83(16), pp. 6164–6168. doi: 10.1073/pnas.83.16.6164.

Davis, J. J. *et al.* (2016) 'Antimicrobial Resistance Prediction in PATRIC and RAST', *Scientific Reports*, 6(1), p. 27930. doi: 10.1038/srep27930.

Dean, C. R. and Goldberg, J. B. (2002) '*Pseudomonas aeruginosa galU* is required for a complete lipopolysaccharide core and repairs a secondary mutation in a PA103 (serogroup O11) *wbpM* mutant', *FEMS Microbiology Letters*, 210(2), pp. 277–283. doi: 10.1111/j.1574-6968.2002.tb11193.x.

Dekker, J. P. and Frank, K. M. (2016) 'Next-generation epidemiology: Using real-time core genome multilocus sequence typing to support infection control policy', *Journal of Clinical Microbiology*, 54(12), pp. 2850–2853. doi: 10.1128/JCM.01714-16.

Delcher, A. L. *et al.* (2002) 'Fast algorithms for large-scale genome alignment and comparison', *Nucleic Acids Research*, 30(11), pp. 2478–2483. doi: 10.1093/nar/30.11.2478.

Denamur, E. *et al.* (1993) 'The absence of correlation between allozyme and rrn RFLP analysis indicates a high gene flow rate within human clinical *Pseudomonas aeruginosa* isolates', *FEMS Microbiology Letters*, 110(3).

Desjardins, P. and Conklin, D. (2010) 'NanoDrop microvolume quantitation of nucleic acids.', *Journal of visualized experiments : JoVE*. MyJoVE Corporation, (45). doi: 10.3791/2565.

Deurenberg, R. H. *et al.* (2017) 'Application of next generation sequencing in clinical microbiology and infection prevention', *Journal of Biotechnology*. Elsevier, 243, pp. 16–24. doi: 10.1016/J.JBIOTEC.2016.12.022.

Didelot, X. and Maiden, M. C. J. (2010) 'Impact of recombination on bacterial evolution', *Trends in Microbiology*, 18(7), pp. 315–322. doi: 10.1016/j.tim.2010.04.002.

Diggle, S. P. *et al.* (2002) 'Advancing the quorum in *Pseudomonas aeruginosa*: MvaT and the regulation of N-acylhomoserine lactone production and virulence gene expression.', *Journal of bacteriology*. American Society for Microbiology (ASM), 184(10), pp. 2576–86. doi: 10.1128/JB.184.10.2576-2586.2002.

Dobler, L. *et al.* (2017) 'Enhanced rhamnolipid production by *Pseudomonas aeruginosa* overexpressing estA in a simple medium', *PLoS ONE*, 12(8), pp. 1–12. doi: 10.1371/journal.pone.0183857.

Doi, Y., Wachino, J. and Arakawa, Y. (2016) 'Aminoglycoside Resistance', *Infectious Disease Clinics of North America*, 30(2), pp. 523–537. doi: 10.1016/j.idc.2016.02.011.

Dötsch, A. *et al.* (2009) 'Genomewide identification of genetic determinants of antimicrobial drug resistance in *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 53(6), pp. 2522–31. doi: 10.1128/AAC.00035-09.

Dou, Y. *et al.* (2017) '*Pseudomonas aeruginosa* prevalence, antibiotic resistance and antimicrobial use in Chinese burn wards from 2007 to 2014', *Journal of International Medical Research*, 45(3), pp. 1124–1137. doi: 10.1177/0300060517703573.

Drlica, K. *et al.* (2008) 'Quinolone-mediated bacterial death.', *Antimicrobial agents and chemotherapy*, 52(2), pp. 385–92. doi: 10.1128/AAC.01617-06.

Drlica, K. *et al.* (2009) 'Quinolones: action and resistance updated.', *Current topics in medicinal chemistry*, 9(11), pp. 981–98.

Earle, S. G. *et al.* (2016) 'Identifying lineage effects when controlling for population structure improves power in bacterial association studies', *Nature Microbiology*, 1(5), pp. 1–21. doi: 10.1038/nmicrobiol.2016.41.

Edelstein, M. V. *et al.* (2013) 'Spread of extensively resistant VIM-2-positive ST235 *Pseudomonas aeruginosa* in Belarus, Kazakhstan, and Russia: A longitudinal epidemiological and clinical study', *The Lancet Infectious Diseases*. Elsevier Ltd, 13(10), pp. 867–876. doi: 10.1016/S1473-3099(13)70168-3.

Edgar, R. C. (2004) 'MUSCLE: multiple sequence alignment with high accuracy and high throughput.', *Nucleic acids research*. Oxford University Press, 32(5), pp. 1792–7. doi:

10.1093/nar/gkh340.

Edson, R. S. and Terrell, C. L. (1999) ‘The Aminoglycosides’, *Mayo Clinic Proceedings*, 74(5), pp. 519–528. doi: 10.4065/74.5.519.

Ehsan, Z. and Clancy, J. P. (2015) ‘Management of *Pseudomonas aeruginosa* infection in cystic fibrosis patients using inhaled antibiotics with a focus on nebulized liposomal amikacin’, *Future Microbiology*. Future Medicine, 10(12), pp. 1901–1912. doi: 10.2217/fmb.15.117.

El’Garch, F. *et al.* (2007) ‘Cumulative effects of several nonenzymatic mechanisms on the resistance of *Pseudomonas aeruginosa* to aminoglycosides.’, *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 51(3), pp. 1016–21. doi: 10.1128/AAC.00704-06.

Ellington, M. J. *et al.* (2017) ‘The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee’, *Clinical Microbiology and Infection*. doi: 10.1016/j.cmi.2016.11.012.

Empel, J. *et al.* (2007) ‘Outbreak of *Pseudomonas aeruginosa* infections with PER-1 extended-spectrum beta-lactamase in Warsaw, Poland: further evidence for an international clonal complex.’, *Journal of clinical microbiology*. American Society for Microbiology, 45(9), pp. 2829–34. doi: 10.1128/JCM.00997-07.

Engstrom, A. *et al.* (2011) ‘Comparison of clinical isolates and in vitro selected mutants reveals that *tly*A is not a sensitive genetic marker for capreomycin resistance in *Mycobacterium tuberculosis*’, *Journal of Antimicrobial Chemotherapy*, 66(6), pp. 1247–1254. doi: 10.1093/jac/dkr109.

Erill, I., Campoy, S. and Barbé, J. (2007) ‘Aeons of distress: an evolutionary perspective on the bacterial SOS response’, *FEMS Microbiology Reviews*, 31(6), pp. 637–656. doi: 10.1111/j.1574-6976.2007.00082.x.

Estahbanati, H. K., Kashani, P. P. and Ghanaatpisheh, F. (2002) ‘Frequency of *Pseudomonas aeruginosa* serotypes in burn wound infections and their resistance to antibiotics.’, *Burns : journal of the International Society for Burn Injuries*, 28(4), pp. 340–8.

EUCAST (2018) *The European Committee on Antimicrobial Susceptibility Testing. Breakpoint*

tables for interpretation of MICs and zone diameters. Available at: http://www.eucast.org.

Eusebi, P. (2013) *Diagnostic Accuracy Measures, Cerebrovascular diseases (Basel, Switzerland)*. doi: 10.1159/000353863.

Everett, M. J. *et al.* (1996) 'Contributions of individual mechanisms to fluoroquinolone resistance in 36 *Escherichia coli* strains isolated from humans and animals.', *Antimicrobial agents and chemotherapy*, 40(10), pp. 2380–6.

Fajardo, A. *et al.* (2008) 'The Neglected Intrinsic Resistome of Bacterial Pathogens', *PLoS ONE*. Edited by M. Falagas. Public Library of Science, 3(2), p. e1619. doi: 10.1371/journal.pone.0001619.

Farhat, M. R. *et al.* (2013) 'Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*', *Nature Genetics*, 45(10), pp. 1183–1189. doi: 10.1038/ng.2747.

Faria, N. R. *et al.* (2016) 'Mobile real-time surveillance of Zika virus in Brazil', *Genome Medicine*. BioMed Central, 8(1), p. 97. doi: 10.1186/s13073-016-0356-2.

Faure, K. *et al.* (2003) 'O-antigen serotypes and type III secretory toxins in clinical isolates of *Pseudomonas aeruginosa*.', *Journal of clinical microbiology*, 41(5), pp. 2158–60.

Feng, W. *et al.* (2019) 'Changes in the resistance and epidemiological characteristics of *Pseudomonas aeruginosa* during a ten-year period', *Journal of Microbiology, Immunology and Infection*. Elsevier Taiwan LLC. doi: 10.1016/j.jmii.2019.08.017.

Feng, Y., Jonker, M. J., Moustakas, I., Brul, S. and Ter Kuile, B. H. (2016) 'Dynamics of Mutations during Development of Resistance by *Pseudomonas aeruginosa* against Five Antibiotics.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology Journals, 60(7), pp. 4229–36. doi: 10.1128/AAC.00434-16.

Fernández-Olmos, A. *et al.* (2013) 'Population structure and Antimicrobial susceptibility of both nonpersistent and persistent *Pseudomonas aeruginosa* isolates recovered from cystic fibrosis patients', *Journal of Clinical Microbiology*, 51(8). doi: 10.1128/JCM.00802-13.

Ferrero, L., Cameron, B. and Crouzet, J. (1995) 'Analysis of *gyr*A and *grl*A mutations in

stepwise-selected ciprofloxacin-resistant mutants of *Staphylococcus aureus*.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 39(7), pp. 1554–8.

Feuerriegel, S. *et al.* (2009) 'Sequence Analyses of Just Four Genes To Detect Extensively Drug-Resistant *Mycobacterium tuberculosis* Strains in Multidrug-Resistant Tuberculosis Patients Undergoing Treatment', *Antimicrobial Agents and Chemotherapy*, 53(8), pp. 3353–3356. doi: 10.1128/AAC.00050-09.

Fischbach, M. A. and Walsh, C. T. (2009) 'Antibiotics for Emerging Pathogens', *Science*, 325(5944), pp. 1089–1093. doi: 10.1126/science.1176667.

Fischer, J. E., Bachmann, L. M. and Jaeschke, R. (2003) 'A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis', *Intensive Care Medicine*. Springer-Verlag, 29(7), pp. 1043–1051. doi: 10.1007/s00134-003-1761-8.

Fischer, M. A. *et al.* (2020) 'Non-Infection-Related And Non-Visit-Based Antibiotic Prescribing Is Common Among Medicaid Patients', *Health Affairs*. Health Affairs, 39(2), pp. 280–288. doi: 10.1377/hlthaff.2019.00545.

Fleming-Dutra, K. E. *et al.* (2016) 'Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011', *JAMA*, 315(17), pp. 1864–1873. doi: 10.1001/jama.2016.4151.

Flores, E. and Gadda, G. (2018) 'Kinetic Characterization of PA1225 from *Pseudomonas aeruginosa* PAO1 Reveals a New NADPH:Quinone Reductase', *Biochemistry*, 57(21), pp. 3050–3058. doi: 10.1021/acs.biochem.8b00090.

Flynn, J. M., Phan, C. and Hunter, R. C. (2017) 'Genome-Wide Surveyof *Pseudomonas aeruginosa* PA14 Reveals a Role for the Glyoxylate Pathway and Extracellular Proteases in the Utilization of Mucin', *Infection and Immunity*. Edited by V. B. Young, 85(8), pp. e00182-17. doi: 10.1128/IAI.00182-17.

Forsberg, K. J. *et al.* (2012) 'The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens', *Science*, 337(6098), pp. 1107–1111. doi: 10.1126/science.1220761.

Foxman, B. *et al.* (2005) 'Choosing an appropriate bacterial typing technique for epidemiologic studies.', *Epidemiologic perspectives & innovations : EP+I*, 2, p. 10. doi: 10.1186/1742-5573-2-

10.

Francisco, A. P. *et al.* (2009) 'Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach', *BMC Bioinformatics*. BioMed Central, 10(1), p. 152. doi: 10.1186/1471-2105-10-152.

Francisco, A. P. *et al.* (2012) 'PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods', *BMC Bioinformatics*. BioMed Central, 13(1), p. 87. doi: 10.1186/1471-2105-13-87.

Fraud, S. and Poole, K. (2011) 'Oxidative Stress Induction of the MexXY Multidrug Efflux Genes and Promotion of Aminoglycoside Resistance Development in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 55(3), pp. 1068–1074. doi: 10.1128/AAC.01495-10.

Freschi, L. *et al.* (2015) 'Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium', *Frontiers in Microbiology*, 6, p. 1036. doi: 10.3389/fmicb.2015.01036.

Freschi, L. *et al.* (2019) 'The *Pseudomonas aeruginosa* Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity', *Genome Biology and Evolution*. Oxford University Press, 11(1), pp. 109–120. doi: 10.1093/gbe/evy259.

Frimodt-Møller, J. *et al.* (2018) 'Mutations causing low level antibiotic resistance ensure bacterial survival in antibiotic-treated hosts', *Scientific Reports*, 8(1), p. 12512. doi: 10.1038/s41598-018-30972-y.

G. M. Moons, K. *et al.* (1997) *Limitations of Sensitivity, Specificity, Likelihood Ratio, and Bayes' Theorem in Assessing Diagnostic Probabilities, Epidemiology (Cambridge, Mass.)*. doi: 10.1097/00001648-199701000-00002.

Gales, A. C. *et al.* (2001) 'Characterization of *Pseudomonas aeruginosa* Isolates: Occurrence Rates, Antimicrobial Susceptibility Patterns, and Molecular Typing in the Global SENTRY Antimicrobial Surveillance Program, 1997–1999', *Clinical Infectious Diseases*, 32(s2), pp. S146–S155. doi: 10.1086/320186.

García-Álvarez, L. *et al.* (2011) 'Meticillin-resistant *Staphylococcus aureus* with a novel *mec*A homologue in human and bovine populations in the UK and Denmark: a descriptive study.', *The*

Lancet. Infectious diseases. Elsevier, 11(8), pp. 595–603. doi: 10.1016/S1473-3099(11)70126-8.

Garcia-Castillo, M. *et al.* (2011) '128 Population structure of *Pseudomonas aeruginosa* from cystic fibrosis patients', *Journal of Cystic Fibrosis*. Elsevier, 10, p. S33. doi: 10.1016/S1569-1993(11)60145-8.

García-Castillo, M. *et al.* (2011) 'Wide dispersion of ST175 clone despite high genetic diversity of carbapenem-nonsusceptible *Pseudomonas aeruginosa* clinical strains in 16 Spanish hospitals.', *Journal of clinical microbiology*. American Society for Microbiology, 49(8), pp. 2905–10. doi: 10.1128/JCM.00753-11.

García, C. *et al.* (2011) 'Knowledge, attitudes and practice survey about antimicrobial resistance and prescribing among physicians in a hospital setting in Lima, Peru', *BMC Clinical Pharmacology*. BioMed Central, 11(1), p. 18. doi: 10.1186/1472-6904-11-18.

Gardy, J. L. and Loman, N. J. (2018) 'Towards a genomics-informed, real-time, global pathogen surveillance system', *Nature Reviews Genetics*. Nature Publishing Group, 19(1), pp. 9–20. doi: 10.1038/nrg.2017.88.

Genestet, C. *et al.* (2019) 'Prospective Whole-Genome Sequencing in Tuberculosis Outbreak Investigation, France, 2017-2018', *Emerging infectious diseases*. Centers for Disease Control and Prevention, 25(3), pp. 589–592. doi: 10.3201/eid2503.181124.

Georghiou, S. B. *et al.* (2012) 'Evaluation of Genetic Mutations Associated with Mycobacterium tuberculosis Resistance to Amikacin, Kanamycin and Capreomycin: A Systematic Review', *PLoS ONE*. Edited by L. E. Via, 7(3), p. e33275. doi: 10.1371/journal.pone.0033275.

Gibbs, M. J., Armstrong, J. S. and Gibbs, A. J. (2000) 'Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences', *Bioinformatics*. Narnia, 16(7), pp. 573–582. doi: 10.1093/bioinformatics/16.7.573.

Gilbert, D. N., Eubanks, N. and Jackson, J. (1977) 'Comparison of amikacin and gentamicin in the treatment of urinary tract infections', *The American Journal of Medicine*, 62(6), pp. 924–929. doi: 10.1016/0002-9343(77)90662-3.

Giske, C. G. *et al.* (2006) 'Establishing Clonal Relationships between VIM-1-Like Metallo- -Lactamase-Producing *Pseudomonas aeruginosa* Strains from Four European Countries by

Multilocus Sequence Typing', *Journal of Clinical Microbiology*, 44(12), pp. 4309–4315. doi: 10.1128/JCM.00817-06.

Glas, A. S. *et al.* (2003) 'The diagnostic odds ratio: a single indicator of test performance', *Journal of Clinical Epidemiology*. Pergamon, 56(11), pp. 1129–1135. doi: 10.1016/S0895-4356(03)00177-X.

Gómez-Zorrilla, S. *et al.* (2016) 'Impact of multidrug resistance on the pathogenicity of *Pseudomonas aeruginosa*: In vitro and in vivo studies', *International Journal of Antimicrobial Agents*, 47(5), pp. 368–374. doi: 10.1016/j.ijantimicag.2016.02.010.

Gomila, M. *et al.* (2013) 'Genetic diversity of clinical *Pseudomonas aeruginosa* isolates in a public hospital in Spain', *BMC Microbiology*, 13(1). doi: 10.1186/1471-2180-13-138.

Gonzalez, M. D. and McElvania, E. (2018) 'New Developments in Rapid Diagnostic Testing for Children', *Infectious Disease Clinics of North America*, 32(1), pp. 19–34. doi: 10.1016/j.idc.2017.11.006.

Goodall, E. C. A. *et al.* (2018) 'The Essential Genome of Escherichia coli K-12.', *mBio*, 9(1). doi: 10.1128/mBio.02096-17.

Goossens, H. (2003) 'Susceptibility of multi-drug-resistant *Pseudomonas aeruginosa* in intensive care units: results from the European MYSTIC study group.', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 9(9), pp. 980–3.

Gordon, N. C. *et al.* (2014) 'Prediction of *Staphylococcus aureus* antimicrobial resistance by whole-genome sequencing.', *Journal of clinical microbiology*. American Society for Microbiology, 52(4), pp. 1182–91. doi: 10.1128/JCM.03117-13.

Gori, A. *et al.* (2019) 'Pan-GWAS of *Streptococcus agalactiae* highlights lineage-specific genes associated with virulence and niche adaptation', *bioRxiv*, p. 574152. doi: 10.1101/574152.

Griffith, S. J. *et al.* (1989) 'The Epidemiology of *Pseudomonas aeruginosa* in Oncology Patients in a General Hospital', *Journal of Infectious Diseases*, 160(6), pp. 1030–1036. doi: 10.1093/infdis/160.6.1030.

Grillon, A. *et al.* (2016) 'Comparative Activity of Ciprofloxacin, Levofloxacin and Moxifloxacin against *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Stenotrophomonas maltophilia* Assessed by Minimum Inhibitory Concentrations and Time-Kill Studies', *PloS one*. Public Library of Science, 11(6), pp. e0156690–e0156690. doi: 10.1371/journal.pone.0156690.

Groicher, K. H. *et al.* (2000) 'The *Staphylococcus aureus* lrgAB operon modulates murein hydrolase activity and penicillin tolerance', *Journal of Bacteriology*, 182(7), pp. 1794–1801. doi: 10.1128/JB.182.7.1794-1801.2000.

Grosso-Becerra, M.-V. *et al.* (2014) '*Pseudomonas aeruginosa* clinical and environmental isolates constitute a single population with high phenotypic diversity', *BMC Genomics*, 15(1), p. 318. doi: 10.1186/1471-2164-15-318.

Gupta, S. K. *et al.* (2014) 'ARG-ANNOT, a New Bioinformatic Tool To Discover Antibiotic Resistance Genes in Bacterial Genomes', *Antimicrobial Agents and Chemotherapy*, 58(1), pp. 212–220. doi: 10.1128/AAC.01310-13.

Guzvinec, M. *et al.* (2014) 'Sequence types 235, 111, and 132 Predominate among multidrug-resistant *Pseudomonas aeruginosa* clinical isolates in croatia', *Antimicrobial Agents and Chemotherapy*. American Society for Microbiology, 58(10), pp. 6277–6283. doi: 10.1128/AAC.03116-14.

Hadfield, J. *et al.* (2018) 'Phandango: an interactive viewer for bacterial population genomics', *Bioinformatics*. Edited by J. Kelso. Narnia, 34(2), pp. 292–293. doi: 10.1093/bioinformatics/btx610.

Han, M.-L. *et al.* (2019) ' Comparative Metabolomics and Transcriptomics Reveal Multiple Pathways Associated with Polymyxin Killing in *Pseudomonas aeruginosa* ', *mSystems*, 4(1), pp. 1–18. doi: 10.1128/msystems.00149-18.

Hanage, W. P. (2016) 'Not So Simple After All: Bacteria, Their Population Genetics, and Recombination', *Cold Spring Harbor Perspectives in Biology*, 8(7), p. a018069. doi: 10.1101/cshperspect.a018069.

Hanson, K. E. and Couturier, M. R. (2016) 'Multiplexed Molecular Diagnostics for Respiratory, Gastrointestinal, and Central Nervous System Infections.', *Clinical infectious diseases : an*

official publication of the Infectious Diseases Society of America. Oxford University Press, 63(10), pp. 1361–1367. doi: 10.1093/cid/ciw494.

Harris, S. R. *et al.* (2013) 'Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study.', *The Lancet. Infectious diseases*. Elsevier, 13(2), pp. 130–6. doi: 10.1016/S1473-3099(12)70268-2.

Hasegawa, M. *et al.* (1997) 'Drug-Resistance Patterns of Clinical Isolates of *Pseudomonas aeruginosa* with Reference to Their Lipopolysaccharide Compositions', *Chemotherapy*, 43(5), pp. 323–331. doi: 10.1159/000239585.

Hasman, H. *et al.* (2014) 'Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples.', *Journal of clinical microbiology*. American Society for Microbiology Journals, 52(1), pp. 139–46. doi: 10.1128/JCM.02452-13.

Haubold, B. and Hudson, R. R. (2000) 'LIAN 3.0: detecting linkage disequilibrium in multilocus data', *Bioinformatics*, 16(9), pp. 847–849. doi: 10.1093/bioinformatics/16.9.847.

Hay, T. *et al.* (2013) 'Antibiotic Inducibility of the *mex*XY Multidrug Efflux Operon of *Pseudomonas aeruginosa*: Involvement of the MexZ Anti-Repressor ArmZ', *PLoS ONE*. Edited by W. van Schaik, 8(2), p. e56858. doi: 10.1371/journal.pone.0056858.

He, X. *et al.* (2017) 'Bacterial genetic architecture of ecological interactions in co-culture by GWAS-taking *Escherichia coli* and *Staphylococcus aureus* as an example', *Frontiers in Microbiology*, 8(NOV), pp. 1–10. doi: 10.3389/fmicb.2017.02332.

Hellen Gelband, Molly Miller-Petrie, Suraj Pant, Sumanth Gandra, Jordan Levinson, Devra Barter, Andrea White, R. L. (2015) *The State of the World's Antibiotics, 2015 | Center for Disease Dynamics, Economics & Policy (CDDEP)*.

Henrichfreise, B. *et al.* (2007) 'Resistance mechanisms of multiresistant *Pseudomonas aeruginosa* strains from Germany and correlation with hypermutation.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology, 51(11), pp. 4062–70. doi: 10.1128/AAC.00148-07.

Hicks, N. D. *et al.* (2019) 'Bacterial genome-wide association identifies novel factors that contribute to ethionamide and prothionamide susceptibility in *Mycobacterium tuberculosis*',

mBio, 10(2). doi: 10.1128/mBio.00616-19.

Higgins, P. G. *et al.* (2003) 'Mutations in GyrA, ParC, MexR and NfxB in clinical isolates of *Pseudomonas aeruginosa.*', *International journal of antimicrobial agents*, 21(5), pp. 409–13.

Hilden, J. and Glasziou, P. (1996) *Regret Graphs, Diagnostic Uncertainty and Youden's Index*, *Statistics in medicine*. doi: 10.1002/(SICI)1097-0258(19960530)15:10<969::AID-SIM211>3.0.CO;2-9.

Hocquet, D. *et al.* (2012) 'Evidence for induction of integron-based antibiotic resistance by the SOS response in a clinical setting', *PLoS Pathogens*, 8(6). doi: 10.1371/journal.ppat.1002778.

van Hoek, A. H. A. M. *et al.* (2011) 'Acquired antibiotic resistance genes: an overview.', *Frontiers in microbiology*. Frontiers Media SA, 2, p. 203. doi: 10.3389/fmicb.2011.00203.

Høiby, N. *et al.* (2016) 'The phenotypic evolution of *Pseudomonas aeruginosa* populations changes in the presence of subinhibitory concentrations of ciprofloxacin', *Microbiology*. Microbiology Society, 162(5), pp. 865–875. doi: 10.1099/mic.0.000273.

Hooper, D. C. (1999) 'Mechanisms of fluoroquinolone resistance.', *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, 2(1), pp. 38–55. doi: 10.1054/drup.1998.0068.

Hooper, D. C. (2000) 'Mechanisms of action and resistance of older and newer fluoroquinolones.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 31 Suppl 2(Supplement_2), pp. S24-8. doi: 10.1086/314056.

Hooper, D. C. and Jacoby, G. A. (2015) 'Mechanisms of drug resistance: quinolone resistance', *Annals of the New York Academy of Sciences*. 2015/07/17, 1354(1), pp. 12–31. doi: 10.1111/nyas.12830.

Horna, G. *et al.* (2019) 'High frequency of the *exo*U+/*exo*S+ genotype associated with multidrug-resistant "high-risk clones" of *Pseudomonas aeruginosa* clinical isolates from Peruvian hospitals', *Scientific Reports*. Springer US, 9(1), pp. 1–13. doi: 10.1038/s41598-019-47303-4.

Hostacká, A. and Majtán, V. (1997) 'Serotyping and virulence factors of *Pseudomonas*

aeruginosa clinical isolates.', *Acta microbiologica et immunologica Hungarica*, 44(2), pp. 141–6.

Housseini B Issa, K., Phan, G. and Broutin, I. (2018) 'Functional Mechanism of the Efflux Pumps Transcription Regulators From *Pseudomonas aeruginosa* Based on 3D Structures', *Frontiers in Molecular Biosciences*. Frontiers, 5, p. 57. doi: 10.3389/fmolb.2018.00057.

Humphreys, H. and Coleman, D. C. (2019) 'Contribution of whole-genome sequencing to understanding of the epidemiology and control of meticillin-resistant *Staphylococcus aureus*', *Journal of Hospital Infection*. W.B. Saunders, 102(2), pp. 189–199. doi: 10.1016/J.JHIN.2019.01.025.

Hunt, M. *et al.* (2017) 'ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads.', *Microbial genomics*. Microbiology Society, 3(10), p. e000131. doi: 10.1099/mgen.0.000131.

Hunter, P. R. and Gaston, M. A. (1988) 'Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity.', *Journal of clinical microbiology*, 26(11), pp. 2465–6.

Ibarz Pavón, A. B. and Maiden, M. C. J. (2009) 'Multilocus sequence typing', *Methods in molecular biology (Clifton, N.J.)*, 551, pp. 129–140. doi: 10.1007/978-1-60327-999-4_11.

Inouye, M. *et al.* (2014) 'SRST2: Rapid genomic surveillance for public health and hospital microbiology labs', *Genome Medicine*, 6(11), p. 90. doi: 10.1186/s13073-014-0090-6.

Islam, S. *et al.* (2009) 'Chromosomal mechanisms of aminoglycoside resistance in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients', *Clinical Microbiology and Infection*, 15(1), pp. 60–66. doi: 10.1111/j.1469-0691.2008.02097.x.

Jaillard, M. *et al.* (2017) 'Representing Genetic Determinants in Bacterial GWAS with Compacted De Bruijn Graphs', *bioRxiv*, p. 113563. doi: 10.1101/113563.

Jaillard, M. *et al.* (2018) 'A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events', *PLoS Genetics*, 14(11), pp. 1–28. doi: 10.1371/journal.pgen.1007758.

Jalal, S. *et al.* (2000) 'Molecular mechanisms of fluoroquinolone resistance in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients.', *Antimicrobial agents and chemotherapy*, 44(3), pp. 710–2.

Jamasbi, R. J. and Proudfoot, E. M. (2008) 'Phenotypic and Genotypic Characteristics of Clinical Isolates of *Pseudomonas aeruginosa*: Rate of Occurrence and Distribution of Different Serotypes, Antimicrobial Susceptibility Profiles', *Lab Med*, 39(3), pp. 155–161. doi: 10.1309/1BAWW0951N7V71CE.

Jarne, P. and Lagoda, P. J. L. (1996) 'Microsatellites, from molecules to populations and back', *Trends in Ecology & Evolution*. Elsevier Current Trends, 11(10), pp. 424–429. doi: 10.1016/0169-5347(96)10049-5.

Jefferies, J. M. C. *et al.* (2012) '*Pseudomonas aeruginosa* outbreaks in the neonatal intensive care unit--a systematic review of risk factors and environmental sources.', *Journal of medical microbiology*. England, 61(Pt 8), pp. 1052–1061. doi: 10.1099/jmm.0.044818-0.

Jensen, L. J. *et al.* (2004) 'Comparative Genomics of Four *Pseudomonas* Species', in *Pseudomonas*. Boston, MA: Springer US, pp. 139–164. doi: 10.1007/978-1-4419-9086-0_5.

Jeukens, J. *et al.* (2017) 'Comparative genomics of a drug-resistant *Pseudomonas aeruginosa* panel and the challenges of antimicrobial resistance prediction from genomes', *FEMS Microbiology Letters*. Narnia, 364(18). doi: 10.1093/femsle/fnx161.

Jeukens, J. *et al.* (2017) 'Genomics of antibiotic-resistance prediction in *Pseudomonas aeruginosa*', *Annals of the New York Academy of Sciences*, 1435(1), pp. 5–17. doi: 10.1111/nyas.13358.

Jeukens, J., Freschi, L., Kukavica-Ibrulj, I., Emond-Rheault, J.-G., *et al.* (2019) 'Genomics of antibiotic-resistance prediction in *Pseudomonas aeruginosa*', *Annals of the New York Academy of Sciences*, 1435(1), pp. 5–17. doi: 10.1111/nyas.13358.

Jeukens, J., Freschi, L., Kukavica-Ibrulj, I., Emond-Rheault, J. G., *et al.* (2019) 'The *Pseudomonas aeruginosa* Population among Cystic Fibrosis Patients in Quebec, Canada: A disease hot spot without known epidemic isolates', *Journal of Clinical Microbiology*, 57(6), pp. 1–3. doi: 10.1128/JCM.02019-18.

Jia, B. *et al.* (2017) 'CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database', *Nucleic Acids Research*, 45(D1), pp. D566–D573. doi: 10.1093/nar/gkw1004.

Johnson, J. K. *et al.* (2007) 'Multilocus Sequence Typing Compared to Pulsed-Field Gel Electrophoresis for Molecular Typing of *Pseudomonas aeruginosa*, *Journal of Clinical Microbiology*, 45(11), p. 3707 LP-3712. doi: 10.1128/JCM.00560-07.

Jolley, K. A. and Maiden, M. C. J. (2014) 'Using MLST to study bacterial variation: Prospects in the genomic era', *Future Microbiology*, 9(5). doi: 10.2217/fmb.14.24.

Jones, A. K. *et al.* (2017) 'Determinants of antibacterial spectrum and resistance potential of the elongation factor G inhibitor argyrin B in key Gram-negative pathogens', *Antimicrobial Agents and Chemotherapy*, 61(4). doi: 10.1128/AAC.02400-16.

Jorth, P. *et al.* (2017) 'Evolved Aztreonam Resistance Is Multifactorial and Can Produce Hypervirulence in *Pseudomonas aeruginosa*', *mBio*. Edited by J. B. Goldberg, 8(5). doi: 10.1128/mBio.00517-17.

Juarez, P. *et al.* (2018) 'Constitutive Activation of MexT by Amino Acid Substitutions Results in MexEF-OprN Overproduction in Clinical Isolates of *Pseudomonas aeruginosa*, *Antimicrobial Agents and Chemotherapy*, 62(5), pp. e02445-17. doi: 10.1128/AAC.02445-17.

Judge, K. *et al.* (2015) 'Early insights into the potential of the Oxford Nanopore MinION for the detection of antimicrobial resistance genes', *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkv206.

Jugheli, L. *et al.* (2009) 'High Level of Cross-Resistance between Kanamycin, Amikacin, and Capreomycin among *Mycobacterium tuberculosis* Isolates from Georgia and a Close Relation with Mutations in the rrs Gene', *Antimicrobial Agents and Chemotherapy*, 53(12), pp. 5064–5068. doi: 10.1128/AAC.00851-09.

Kaas, R. S. *et al.* (2014) 'Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms', *PLoS ONE*. Edited by A. Friedrich. Public Library of Science, 9(8), p. e104984. doi: 10.1371/journal.pone.0104984.

Kaatz, G. W., Seo, S. M. and Foster, T. J. (1999) 'Introduction of a *nor*A promoter region

mutation into the chromosome of a fluoroquinolone-susceptible strain of *Staphylococcus aureus* using plasmid integration.', *Antimicrobial agents and chemotherapy*, 43(9), pp. 2222–4.

Kallen, A. J. and Srinivasan, A. (2010) 'Current Epidemiology of Multidrug-Resistant Gram-Negative Bacilli in the United States', *Infection Control & Hospital Epidemiology*, 31(S1), pp. S51–S54. doi: 10.1086/655996.

Kanehisa, M. and Goto, S. (2000) 'KEGG: kyoto encyclopedia of genes and genomes.', *Nucleic acids research*. Oxford University Press, 28(1), pp. 27–30. doi: 10.1093/nar/28.1.27.

Karlowsky, J. A. *et al.* (2003) 'Surveillance for antimicrobial susceptibility among clinical isolates of *Pseudomonas aeruginosa* and *Acinetobacter baumannii* from hospitalized patients in the United States, 1998 to 2001.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 47(5), pp. 1681–8. doi: 10.1128/AAC.47.5.1681-1688.2003.

Karlowsky, J. A. *et al.* (2005) 'Stable Antimicrobial Susceptibility Rates for Clinical Isolates of *Pseudomonas aeruginosa* from the 2001–2003 Tracking Resistance in the United States Today Surveillance Studies', *Clinical Infectious Diseases*, 40(Supplement_2), pp. S89–S98. doi: 10.1086/426188.

Kaye, K. S. *et al.* (2006) 'Differential Effects of Levofloxacin and Ciprofloxacin on the Risk for Isolation of Quinolone-Resistant *Pseudomonas aeruginosa*, *Antimicrobial Agents and Chemotherapy*, 50(6), p. 2192 LP-2196. doi: 10.1128/AAC.00060-06.

Kettner, M. *et al.* (1995) 'Incidence and mechanisms of aminoglycoside resistance in *Pseudomonas aeruginosa* serotype O11 isolates.', *Infection*, 23(6), pp. 380–3.

Khaledi, A. *et al.* (2020) 'Predicting antimicrobial resistance in *Pseudomonas aeruginosa* with machine learning-enabled molecular diagnostics', *EMBO Molecular Medicine*. John Wiley & Sons, Ltd, 12(3), p. e10264. doi: 10.15252/emmm.201910264.

Kidd, T. J. *et al.* (2012) '*Pseudomonas aeruginosa* Exhibits Frequent Recombination, but Only a Limited Association between Genotype and Ecological Setting', *PLoS ONE*, 7(9). doi: 10.1371/journal.pone.0044199.

Kiewitz, C. and Tümmler, B. (2000) 'Sequence diversity of *Pseudomonas aeruginosa*: impact on population structure and genome evolution.', *Journal of bacteriology*. American Society for

Microbiology, 182(11), pp. 3125–35. doi: 10.1128/JB.182.11.3125-3135.2000.

Kim, J.-S. *et al.* (2016) 'Evaluation of Verigene Blood Culture Test Systems for Rapid Identification of Positive Blood Cultures.', *BioMed research international*. Hindawi Limited, 2016, p. 1081536. doi: 10.1155/2016/1081536.

Kim, Y. A. *et al.* (2018) 'Correlation of Aminoglycoside Consumption and Amikacin- or Gentamicin-Resistant *Pseudomonas aeruginosa* in Long-Term Nationwide Analysis: Is Antibiotic Cycling an Effective Policy for Reducing Antimicrobial Resistance?', *Annals of laboratory medicine*. The Korean Society for Laboratory Medicine, 38(2), pp. 176–178. doi: 10.3343/alm.2018.38.2.176.

Kindrachuk, K. N. *et al.* (2011) 'Involvement of an ATP-dependent protease, PA0779/AsrA, in inducing heat shock in response to tobramycin in *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 55(5), pp. 1874–82. doi: 10.1128/AAC.00935-10.

Kiser, T. H. *et al.* (2010) 'Efflux Pump Contribution to Multidrug Resistance in Clinical Isolates of *Pseudomonas aeruginosa*', *Pharmacotherapy*, 30(7), pp. 632–638. doi: 10.1592/phco.30.7.632.

Klockgether, J. *et al.* (2011) '*Pseudomonas aeruginosa* Genomic Structure and Diversity', *Frontiers in Microbiology*, 2, p. 150. doi: 10.3389/fmicb.2011.00150.

Kłodzińska, S. N. *et al.* (2016) 'Inhalable Antimicrobials for Treatment of Bacterial Biofilm-Associated Sinusitis in Cystic Fibrosis Patients: Challenges and Drug Delivery Approaches.', *International journal of molecular sciences*, 17(10). doi: 10.3390/ijms17101688.

Kluge, R. M. *et al.* (1974) 'Comparative activity of tobramycin, amikacin, and gentamicin alone and with carbenicillin against *Pseudomonas aeruginosa*', *Antimicrobial agents and chemotherapy*, 6(4), pp. 442–446. doi: 10.1128/aac.6.4.442.

Knopp, M. and Andersson, D. I. (2018) 'Predictable phenotypes of antibiotic resistance mutations', *mBio*, 9(3), pp. 1–14. doi: 10.1128/mBio.00770-18.

Knottnerus, J. A., van Weel, C. and Muris, J. W. M. (2002) 'Evaluation of diagnostic

procedures.', *BMJ (Clinical research ed.)*. BMJ Publishing Group, 324(7335), pp. 477–80. doi: 10.1136/bmj.324.7335.477.

Kofoid, E. *et al.* (1999) 'The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins.', *Journal of bacteriology*. American Society for Microbiology (ASM), 181(17), pp. 5317–29.

Köhler, T. *et al.* (1997) 'Characterization of MexE-MexF-OprN, a positively regulated multidrug efflux system of *Pseudomonas aeruginosa*.', *Molecular microbiology*, 23(2), pp. 345–54.

Komp Lindgren, P. *et al.* (2005) 'Biological cost of single and multiple norfloxacin resistance mutations in *Escherichia coli* implicated in urinary tract infections.', *Antimicrobial agents and chemotherapy*, 49(6), pp. 2343–51. doi: 10.1128/AAC.49.6.2343-2351.2005.

Kondo, J. *et al.* (2006) 'Crystal structure of the bacterial ribosomal decoding site complexed with amikacin containing the gamma-amino-alpha-hydroxybutyryl (haba) group.', *Biochimie*. France, 88(8), pp. 1027–1031. doi: 10.1016/j.biochi.2006.05.017.

König, J. *et al.* (2012) 'Protein–RNA interactions: new genomic technologies and perspectives', *Nature Reviews Genetics*, 13(2), pp. 77–83. doi: 10.1038/nrg3141.

Kordes, A. *et al.* (2019) 'Genetically diverse *Pseudomonas aeruginosa* populations display similar transcriptomic profiles in a cystic fibrosis explanted lung', *Nature Communications*, 10(1), p. 3397. doi: 10.1038/s41467-019-11414-3.

Kos, V. N. *et al.* (2015) 'The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology, 59(1), pp. 427–36. doi: 10.1128/AAC.03954-14.

Kosack, C. S., Page, A.-L. and Klatser, P. R. (2017) 'A guide to aid the selection of diagnostic tests.', *Bulletin of the World Health Organization*. World Health Organization, 95(9), pp. 639–645. doi: 10.2471/BLT.16.187468.

Köser, C. U., Holden, M. T. G., *et al.* (2012) 'Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak', *New England Journal of Medicine*. Massachusetts Medical Society , 366(24), pp. 2267–2275. doi: 10.1056/NEJMoa1109910.

Köser, C. U., Ellington, M. J., *et al.* (2012) 'Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology', *PLoS Pathogens*. Edited by G. F. Rall, 8(8), p. e1002824. doi: 10.1371/journal.ppat.1002824.

Kowalski, R. P. *et al.* (2001) 'An in vitro resistance study of levofloxacin, ciprofloxacin, and ofloxacin using keratitis isolates of *Staphylococcus aureus* and *Pseudomonas aeruginosa*', *Ophthalmology*. Elsevier, 108(10), pp. 1826–1829. doi: 10.1016/S0161-6420(01)00724-2.

Krahn, T. *et al.* (2012) 'Determinants of Intrinsic Aminoglycoside Resistance in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 56(11), pp. 5591–5602. doi: 10.1128/AAC.01446-12.

Kugelberg, E. *et al.* (2005) 'Reduction of the fitness burden of quinolone resistance in *Pseudomonas aeruginosa', Journal of Antimicrobial Chemotherapy*, 55(1), pp. 22–30. doi: 10.1093/jac/dkh505.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular Biology and Evolution*, 33(7), pp. 1870–1874. doi: 10.1093/molbev/msw054.

Laabei, M. *et al.* (2014) 'Predicting the virulence of MRSA from its genome sequence', *Genome Research*, 24(5), pp. 839–849. doi: 10.1101/gr.165415.113.

LaBorwit, S. E. *et al.* (2001) 'Topical 0.3% ciprofloxacin vs topical 0.3% ofloxacin in early treatment of *Pseudomonas aeruginosa* keratitis in a rabbit model', *Annals of Ophthalmology*, 33(1), pp. 48–52. doi: 10.1007/s12009-001-0055-4.

Ladner, J. T. *et al.* (2019) 'Precision epidemiology for infectious disease control', *Nature Medicine*, 25(2), pp. 206–211. doi: 10.1038/s41591-019-0345-2.

Lakin, S. M. *et al.* (2017) 'MEGARes: an antimicrobial resistance database for high throughput sequencing.', *Nucleic acids research*. Oxford University Press, 45(D1), pp. D574–D580. doi: 10.1093/nar/gkw1009.

Lambert, D. *et al.* (2015) 'GeneSippr: A Rapid Whole-Genome Approach for the Identification and Characterization of Foodborne Pathogens such as Priority Shiga Toxigenic *Escherichia coli*', *PLOS ONE*. Edited by P. Fratamico. Public Library of Science, 10(4), p. e0122928. doi:

10.1371/journal.pone.0122928.

Langridge, G. C. *et al.* (2009) 'Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants.', *Genome research*, 19(12), pp. 2308–2316. doi: 10.1101/gr.097097.109.

Larsen, M. V. *et al.* (2012) 'Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria', *Journal of Clinical Microbiology*, 50(4), pp. 1355–1361. doi: 10.1128/JCM.06094-11.

Lau, C. H.-F. *et al.* (2013) 'Mutational Activation of the AmgRS Two-Component System in Aminoglycoside-Resistant *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 57(5), pp. 2243–2251. doi: 10.1128/AAC.00170-13.

Lau, C. H.-F. *et al.* (2015) 'AmgRS-mediated envelope stress-inducible expression of the *mex*XY multidrug efflux operon of *Pseudomonas aeruginosa*.', *MicrobiologyOpen*. Wiley-Blackwell, 4(1), pp. 121–35. doi: 10.1002/mbo3.226.

Lau, W. K. *et al.* (1977) 'Amikacin therapy of exacerbations of *Pseudomonas aeruginosa* infections in patients with cystic fibrosis.', *Pediatrics*. United States, 60(3), pp. 372–377.

Lee, Y.-J. *et al.* (2010) 'Fluoroquinolone resistance of *Pseudomonas aeruginosa* isolates causing nosocomial infection is correlated with levofloxacin but not ciprofloxacin use.', *International journal of antimicrobial agents*. Netherlands, 35(3), pp. 261–264. doi: 10.1016/j.ijantimicag.2009.11.007.

Leibowitz, H. M. (1991) 'Clinical evaluation of ciprofloxacin 0.3% ophthalmic solution for treatment of bacterial keratitis.', *American journal of ophthalmology*. United States, 112(4 Suppl), p. 34S–47S.

Lenski, R. E. (1993) 'Assessing the genetic structure of microbial populations.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 90(10), pp. 4334–6. doi: 10.1073/pnas.90.10.4334.

Leung, K. L. *et al.* (2010) 'Usefulness of resistant gene markers for predicting treatment outcome on second-line anti-tuberculosis drugs', *Journal of Applied Microbiology*, 109(6), pp. 2087–2094. doi: 10.1111/j.1365-2672.2010.04840.x.

Levin, I. *et al.* (2004) 'The ternary complex of *Pseudomonas aeruginosa* alcohol dehydrogenase with NADH and ethylene glycol', *Protein Science*. John Wiley & Sons, Ltd, 13(6), pp. 1547–1556. doi: 10.1110/ps.03531404.

Lewis, K. (2013) 'Platforms for antibiotic discovery', *Nature Reviews Drug Discovery*, 12(5), pp. 371–387. doi: 10.1038/nrd3975.

Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools.', *Bioinformatics (Oxford, England)*. Oxford University Press, 25(16), pp. 2078–9. doi: 10.1093/bioinformatics/btp352.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754–1760. doi: 10.1093/bioinformatics/btp324.

Liakopoulos, A. *et al.* (2013) 'Carbapenemase-producing *Pseudomonas aeruginosa* from central Greece: Molecular epidemiology and genetic analysis of class I integrons', *BMC Infectious Diseases*. BMC Infectious Diseases, 13(1), p. 1. doi: 10.1186/1471-2334-13-505.

Lila, G. *et al.* (2017) 'The prevalence and resistance patterns of *Pseudomonas aeruginosa* in a tertiary care hospital in Kosovo.', *Le infezioni in medicina : rivista periodica di eziologia, epidemiologia, diagnostica, clinica e terapia delle patologie infettive*, 25(1), pp. 21–26.

Linnet, K. (1988) *A review on the methodology for assessing diagnostic tests*, *Clinical chemistry*.

Liu, B. and Pop, M. (2009) 'ARDB--Antibiotic Resistance Genes Database', *Nucleic Acids Research*, 37(Database), pp. D443–D447. doi: 10.1093/nar/gkn656.

Liu, C. *et al.* (2019) 'Intrinsic and external determinants of antibiotic prescribing: a multi-level path analysis of primary care prescriptions in Hubei, China', *Antimicrobial Resistance & Infection Control*. BioMed Central, 8(1), p. 132. doi: 10.1186/s13756-019-0592-5.

LIU, P. V. *et al.* (1983) 'Survey of Heat-Stable, Major Somatic Antigens of *Pseudomonas aeruginosa*', *International Journal of Systematic Bacteriology*. Microbiology Society, 33(2), pp. 256–264. doi: 10.1099/00207713-33-2-256.

Liu, P. V and Wang, S. (1990) 'Three new major somatic antigens of *Pseudomonas aeruginosa*.', *Journal of clinical microbiology*. American Society for Microbiology (ASM),

28(5), pp. 922–5.

Liu, Z. *et al.* (2018) ‘A Genome-Wide Screen Identifies Genes in Rhizosphere-Associated *Pseudomonas* Required to Evade Plant Defenses’, *mBio*, 9(6). doi: 10.1128/mBio.00433-18.

Livermore, D. M., Williams, R. J. and Williams, J. D. (1981) ‘Comparison of the beta-lactamase stability and the in-vitro activity of cefoperazone, cefotaxime, cefsulodin, ceftazidime, moxalactam and ceftriaxone against *Pseudomonas aeruginosa*.’, *The Journal of antimicrobial chemotherapy*, 8(4), pp. 323–31.

Llanes, C. *et al.* (2004) ‘Clinical strains of *Pseudomonas aeruginosa* overproducing MexAB-OprM and MexXY efflux pumps simultaneously.’, *Antimicrobial agents and chemotherapy*, 48(5), pp. 1797–802.

Llanes, C. *et al.* (2011) ‘Role of the MexEF-OprN efflux system in low-level resistance of *Pseudomonas aeruginosa* to ciprofloxacin.’, *Antimicrobial agents and chemotherapy*, 55(12), pp. 5676–84. doi: 10.1128/AAC.00101-11.

Loit, K. *et al.* (2019) ‘Relative performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) third generation sequencing instruments in identification of agricultural and forest fungal pathogens’, *Applied and Environmental Microbiology*, 85(21), pp. 1–20. doi: 10.1128/AEM.01368-19.

Loman, N. J. *et al.* (2013) ‘A Culture-Independent Sequence-Based Metagenomics Approach to the Investigation of an Outbreak of Shiga-Toxigenic Escherichia coli O104:H4’, *JAMA*, 309(14), p. 1502. doi: 10.1001/jama.2013.3231.

Loman, N. J. and Watson, M. (2015) ‘Successful test launch for nanopore sequencing’, *Nature Methods*, 12(4), pp. 303–304. doi: 10.1038/nmeth.3327.

Lomholt, J. A., Poulsen, K. and Kilian, M. (2001) ‘Epidemic population structure of *Pseudomonas aeruginosa*: Evidence for a clone that is pathogenic to the eye and that has a distinct combination of virulence factors’, *Infection and Immunity*, 69(10). doi: 10.1128/IAI.69.10.6284-6295.2001.

López-Causapé, C. *et al.* (2018) ‘Evolution of the *Pseudomonas aeruginosa* Aminoglycoside Mutational Resistome In Vitro and in the Cystic Fibrosis Setting’, *Antimicrobial Agents and*

Chemotherapy. American Society for Microbiology Journals, 62(4), pp. e02583-17. doi: 10.1128/AAC.02583-17.

Loveday, H. P. *et al.* (2014) 'Association between healthcare water systems and *Pseudomonas aeruginosa* infections: a rapid systematic review.', *The Journal of hospital infection*. England, 86(1), pp. 7–15. doi: 10.1016/j.jhin.2013.09.010.

Lu, Q. *et al.* (2014) '*Pseudomonas aeruginosa* serotypes in nosocomial pneumonia: prevalence and clinical outcomes', *Critical Care*, 18(1), p. R17. doi: 10.1186/cc13697.

Maatallah, M. *et al.* (2011) 'Population structure of *Pseudomonas aeruginosa* from five Mediterranean countries: Evidence for frequent recombination and epidemic occurrence of CC235', *PLoS ONE*, 6(10). doi: 10.1371/journal.pone.0025617.

Macfarlane, E. L. A. *et al.* (1999) 'PhoP-PhoQ homologues in *Pseudomonas aeruginosa* regulate expression of the outer-membrane protein OprH and polymyxin B resistance', *Molecular Microbiology*, 34(2), pp. 305–316. doi: 10.1046/j.1365-2958.1999.01600.x.

MacGowan, A. P., Wootton, M. and Holt, H. A. (1999) 'The antibacterial efficacy of levofloxacin and ciprofloxacin against *Pseudomonas aeruginosa* assessed by combining antibiotic exposure and bacterial susceptibility', *Journal of Antimicrobial Chemotherapy*, 43(3), pp. 345–349. doi: 10.1093/jac/43.3.345.

Macvanin, M. and Hughes, D. (2005) 'Hyper-susceptibility of a fusidic acid-resistant mutant of Salmonella to different classes of antibiotics.', *FEMS microbiology letters*. England, 247(2), pp. 215–220. doi: 10.1016/j.femsle.2005.05.007.

Madigan, T. *et al.* (2018) 'Whole-genome sequencing for methicillin-resistant *Staphylococcus aureus* (MRSA) outbreak investigation in a neonatal intensive care unit', *Infection Control & Hospital Epidemiology*. Cambridge University Press, 39(12), pp. 1412–1418. doi: 10.1017/ice.2018.239.

Magalhães, B. *et al.* (2020) 'Combining Standard Molecular Typing and Whole Genome Sequencing to Investigate *Pseudomonas aeruginosa* Epidemiology in Intensive Care Units ', *Frontiers in Public Health* , p. 3. Available at: https://www.frontiersin.org/article/10.3389/fpubh.2020.00003.

Maiden, M. C. J. (2006) ‘Multilocus Sequence Typing of Bacteria’, *Annual Review of Microbiology*. Annual Reviews, 60(1), pp. 561–588. doi: 10.1146/annurev.micro.59.030804.121325.

de Man, T. J. B. and Limbago, B. M. (2016) ‘SSTAR, a Stand-Alone Easy-To-Use Antimicrobial Resistance Gene Predictor’, *mSphere*. Edited by P. Dunman, 1(1). doi: 10.1128/mSphere.00050-15.

van Mansfeld, R. *et al.* (2010) ‘The Population Genetics of *Pseudomonas aeruginosa* Isolates from Different Patient Populations Exhibits High-Level Host Specificity’, *PLoS ONE*. Edited by E. M. Braga, 5(10), p. e13482. doi: 10.1371/journal.pone.0013482.

Van Mansfeld, R. *et al.* (2009) ‘*Pseudomonas aeruginosa* genotype prevalence in Dutch Cystic Fibrosis patients and age dependency of colonization by various *Ps. aeruginosa* sequence types’, *Journal of Clinical Microbiology*, 47(12). doi: 10.1128/JCM.01462-09.

Marchetti, F. and Viale, P. (2003) ‘Current and future perspectives for levofloxacin in severe *Pseudomonas aeruginosa* infections.’, *Journal of chemotherapy (Florence, Italy)*. England, 15(4), pp. 315–322. doi: 10.1179/joc.2003.15.4.315.

Martin, D. P. *et al.* (2005) ‘A Modified Bootscan Algorithm for Automated Identification of Recombinant Sequences and Recombination Breakpoints’, *AIDS Research and Human Retroviruses*, 21(1), pp. 98–102. doi: 10.1089/aid.2005.21.98.

Martin, D. P. *et al.* (2010) ‘RDP3: a flexible and fast computer program for analyzing recombination’, *Bioinformatics*, 26(19), pp. 2462–2463. doi: 10.1093/bioinformatics/btq467.

Martin, D. and Rybicki, E. (2000) ‘RDP: detection of recombination amongst aligned sequences’, *Bioinformatics*, 16(6), pp. 562–563. doi: 10.1093/bioinformatics/16.6.562.

Martinez, E. *et al.* (2012) ‘Diverse mobilized class 1 integrons are common in the chromosomes of pathogenic *Pseudomonas aeruginosa* clinical isolates’, *Antimicrobial Agents and Chemotherapy*, 56(4), pp. 2169–2172. doi: 10.1128/AAC.06048-11.

Martinez, E. *et al.* (2013) ‘Emerging and existing mechanisms co-operate in generating diverse β-lactam resistance phenotypes in geographically dispersed and genetically disparate *Pseudomonas aeruginosa* strains’, *Journal of Global Antimicrobial Resistance*, 1(3), pp. 135–

142. doi: 10.1016/j.jgar.2013.03.013.

Martinez, E. *et al.* (2014) 'Establishment and multi drug resistance evolution of ST235 *Pseudomonas aeruginosa* strains in the intensive care unit of a Colombian hospital', *Research in Microbiology*. Elsevier Masson SAS, 165(10), pp. 852–856. doi: 10.1016/j.resmic.2014.10.011.

Martínez, J. L. (2019) 'Mechanisms of Action and of Resistance to Quinolones', *Antibiotic Drug Resistance*. (Wiley Online Books), pp. 39–55. doi: doi:10.1002/9781119282549.ch2.

Martínez, J. L. and Baquero, F. (2002) 'Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance.', *Clinical microbiology reviews*, 15(4), pp. 647–79. doi: 10.1128/cmr.15.4.647-679.2002.

Masuda, N. *et al.* (1996) 'Quantitative correlation between susceptibility and OprJ production in NfxB mutants of *Pseudomonas aeruginosa.*', *Antimicrobial agents and chemotherapy*, 40(4), pp. 909–13.

Masuda, N. *et al.* (2000) 'Contribution of the MexX-MexY-*opr*M efflux system to intrinsic resistance in *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 44(9), pp. 2242–6.

Matsumoto, M. *et al.* (2012) 'Mutations in the *gyr*A and *par*C genes and in vitro activities of fluoroquinolones in 114 clinical isolates of *Pseudomonas aeruginosa* derived from urinary tract infections and their rapid detection by denaturing high-performance liquid chromatography', *International Journal of Antimicrobial Agents*, 40(5), pp. 440–444. doi: 10.1016/j.ijantimicag.2012.06.021.

Maurer, F. P. *et al.* (2017) 'Advances in rapid identification and susceptibility testing of bacteria in the clinical microbiology laboratory: implications for patient care and antimicrobial stewardship programs', *Infectious Disease Reports*, 9(1), p. 6839. doi: 10.4081/idr.2017.6839.

Maury, M. M. *et al.* (2016) 'Europe PMC Funders Group Uncovering Listeria monocytogenes hypervirulence by harnessing its biodiversity', *Nature Genetics*, 48(3), pp. 308–313. doi: 10.1038/ng.3501.

Maus, C. E., Plikaytis, B. B. and Shinnick, T. M. (2005) 'Molecular Analysis of Cross-Resistance to Capreomycin, Kanamycin, Amikacin, and Viomycin in Mycobacterium

tuberculosis', *Antimicrobial Agents and Chemotherapy*, 49(8), pp. 3192–3197. doi: 10.1128/AAC.49.8.3192-3197.2005.

McArthur, A. G. *et al.* (2013) 'The comprehensive antibiotic resistance database.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology, 57(7), pp. 3348–57. doi: 10.1128/AAC.00419-13.

McDonald, M. J. (2019) 'Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery', *EMBO reports*. John Wiley & Sons, Ltd, 20(8), p. e46992. doi: 10.15252/embr.201846992.

McGann, P. *et al.* (2016) 'Real time application of whole genome sequencing for outbreak investigation – What is an achievable turnaround time?', *Diagnostic Microbiology and Infectious Disease*. doi: 10.1016/j.diagmicrobio.2016.04.020.

McPhee, J. B., Lewenza, S. and Hancock, R. E. W. (2003) 'Cationic antimicrobial peptides activate a two-component regulatory system, PmrA-PmrB, that regulates resistance to polymyxin B and cationic antimicrobial peptides in *Pseudomonas aeruginosa*', *Molecular Microbiology*, 50(1), pp. 205–217. doi: 10.1046/j.1365-2958.2003.03673.x.

Mena, A. *et al.* (2008) 'Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation', *Journal of Bacteriology*, 190(24), pp. 7910–7917. doi: 10.1128/JB.01147-08.

Mensa, J. *et al.* (2018) 'Antibiotic selection in the treatment of acute invasive infections by *Pseudomonas aeruginosa*: Guidelines by the Spanish Society of Chemotherapy', *Revista espanola de quimioterapia : publicacion oficial de la Sociedad Espanola de Quimioterapia*. 2018/02/23. Sociedad Española de Quimioterapia, 31(1), pp. 78–100. Available at: https://pubmed.ncbi.nlm.nih.gov/29480677.

Mezger, A. *et al.* (2015) 'A General Method for Rapid Determination of Antibiotic Susceptibility and Species in Bacterial Infections', *Journal of Clinical Microbiology*. Edited by S. S. Richter, 53(2), pp. 425–432. doi: 10.1128/JCM.02434-14.

Miao, R. *et al.* (2020) 'Inappropriate antibiotic prescriptions among pediatric inpatients in different type hospitals', *Medicine*. Wolters Kluwer Health, 99(2), pp. e18714–e18714. doi:

10.1097/MD.0000000000018714.

Middleton, M. A. *et al.* (2018) 'Epidemiology of Clonal *Pseudomonas aeruginosa* Infection in a Canadian Cystic Fibrosis Population', *Annals of the American Thoracic Society*. American Thoracic Society - AJRCCM, 15(7), pp. 827–836. doi: 10.1513/AnnalsATS.201801-007OC.

Mikheil, D. M. *et al.* (2012) 'Deletion of gene encoding methyltransferase (gidB) confers high-level antimicrobial resistance in Salmonella', *The Journal of Antibiotics*, 65(4), pp. 185–192. doi: 10.1038/ja.2012.5.

Miller, C. *et al.* (2004) 'SOS response induction by beta-lactams and bacterial defense against antibiotic lethality.', *Science (New York, N.Y.)*, 305(5690), pp. 1629–31. doi: 10.1126/science.1101630.

Mima, T. *et al.* (2005) 'Gene cloning and properties of the RND-type multidrug efflux pumps MexPQ-OpmE and MexMN-OprM from *Pseudomonas aeruginosa*.', *Microbiology and immunology*, 49(11), pp. 999–1002.

Mingeot-Leclercq, M. P., Glupczynski, Y. and Tulkens, P. M. (1999) 'Aminoglycosides: activity and resistance.', *Antimicrobial agents and chemotherapy*, 43(4), pp. 727–37.

Mitsakakis, K. *et al.* (2018) 'Challenges in identifying antibiotic resistance targets for point-of-care diagnostics in general practice.', *Future microbiology*. Future Science Group, 13(10), pp. 1157–1164. doi: 10.2217/fmb-2018-0084.

Mobegi, F. M. *et al.* (2017) 'Deciphering the distance to antibiotic resistance for the *Pneumococcus* using genome sequencing data', *Scientific Reports*. Nature Publishing Group, 7(October 2016), pp. 1–13. doi: 10.1038/srep42808.

Mogayzel, P. J. *et al.* (2014) 'Cystic Fibrosis Foundation Pulmonary Guideline. Pharmacologic Approaches to Prevention and Eradication of Initial *Pseudomonas aeruginosa* Infection', *Annals of the American Thoracic Society*. American Thoracic Society - AJRCCM, 11(10), pp. 1640–1650. doi: 10.1513/AnnalsATS.201404-166OC.

Moloney, E. M. *et al.* (2019) 'Whole-genome sequencing identifies highly related *Pseudomonas aeruginosa* strains in multiple washbasin U-bends at several locations in one hospital: evidence for trafficking of potential pathogens via wastewater pipes', *Journal of Hospital Infection*.

Elsevier Ltd, pp. 4–11. doi: 10.1016/j.jhin.2019.11.005.

Moons, K. G. M., Biesheuvel, C. J. and Grobbee, D. E. (2004) 'Test Research versus Diagnostic Research', *Clinical Chemistry*, 50(3), pp. 473–476. doi: 10.1373/clinchem.2003.024752.

Morales-Espinosa, R. *et al.* (2012) 'Genetic and phenotypic characterization of a *Pseudomonas aeruginosa* population with high frequency of genomic islands', *PLoS ONE*, 7(5). doi: 10.1371/journal.pone.0037459.

Morgan-Linnell, S. K. and Zechiedrich, L. (2007) 'Contributions of the combined effects of topoisomerase mutations toward fluoroquinolone resistance in *Escherichia coli*.', *Antimicrobial agents and chemotherapy*, 51(11), pp. 4205–8. doi: 10.1128/AAC.00647-07.

Morganti, S. *et al.* (2020) 'Role of Next-Generation Sequencing Technologies in Personalized Medicine', in Pravettoni, G. and Triberti, S. (eds) *P5 eHealth: An Agenda for the Health Technologies of the Future*. Cham: Springer International Publishing, pp. 125–154. doi: 10.1007/978-3-030-27994-3_8.

Morita, Y., Tomida, J. and Kawamura, Y. (2012) 'MexXY multidrug efflux system of *Pseudomonas aeruginosa*', *Frontiers in Microbiology*, 3, p. 408. doi: 10.3389/fmicb.2012.00408.

Morita, Y., Tomida, J. and Kawamura, Y. (2014) 'Responses of *Pseudomonas aeruginosa* to antimicrobials', *Frontiers in Microbiology*, 4, p. 422. doi: 10.3389/fmicb.2013.00422.

Mosammaparast, N., Nolte, F. S. and McAdam, A. J. (2012) 'Point-Counterpoint: Molecular Testing for Infectious Diseases Should Be Done in the Clinical Microbiology Laboratory: Table 1', *Journal of Clinical Microbiology*. Edited by P. H. Gilligan, 50(6), pp. 1836–1840. doi: 10.1128/JCM.00488-12.

Mösges, R., Nematian-Samani, M. and Eichel, A. (2011) 'Treatment of acute otitis externa with ciprofloxacin otic 0.2% antibiotic ear solution.', *Therapeutics and clinical risk management*, 7, pp. 325–336. doi: 10.2147/TCRM.S6769.

Moskowitz, S. M., Ernst, R. K. and Miller, S. I. (2004) 'PmrAB, a two-component regulatory system of *Pseudomonas aeruginosa* that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A.', *Journal of bacteriology*, 186(2), pp. 575–9. doi:

10.1128/jb.186.2.575-579.2004.

Mosquera-Rendón, J. *et al.* (2016) 'Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species', *BMC Genomics*. BioMed Central Ltd., 17(1). doi: 10.1186/s12864-016-2364-4.

Mouneimné, H. *et al.* (1999) 'Type II topoisomerase mutations in ciprofloxacin-resistant strains of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 43(1), pp. 62–6.

Mulet, X. *et al.* (2013) 'Biological markers of *Pseudomonas aeruginosa* epidemic high-risk clones', *Antimicrobial Agents and Chemotherapy*, 57(11), pp. 5527–5535. doi: 10.1128/AAC.01481-13.

Murray, J. L. *et al.* (2015) 'Intrinsic Antimicrobial Resistance Determinants in the Superbug *Pseudomonas aeruginosa.*', *mBio*. American Society for Microbiology (ASM), 6(6), pp. e01603-15. doi: 10.1128/mBio.01603-15.

Mustafa, R. A. *et al.* (2017) 'Decision making about healthcare-related tests and diagnostic test strategies. Paper 2: a review of methodological and practical challenges', *Journal of Clinical Epidemiology*, 92, pp. 18–28. doi: 10.1016/j.jclinepi.2017.09.003.

Nagaev, I. *et al.* (2001) 'Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*.', *Molecular microbiology*. England, 40(2), pp. 433–439. doi: 10.1046/j.1365-2958.2001.02389.x.

Nakano, M. *et al.* (1997) 'Mutations in the *gyr*A and *par*C genes in fluoroquinolone-resistant clinical isolates of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 41(10), pp. 2289–91.

Nebenzahl-Guimaraes, H. *et al.* (2017) 'Transmissible *Mycobacterium tuberculosis* strains share genetic markers and immune phenotypes', *American Journal of Respiratory and Critical Care Medicine*, 195(11), pp. 1519–1527. doi: 10.1164/rccm.201605-1042OC.

Nemec, A. *et al.* (2010) 'Multidrug-resistant epidemic clones among bloodstream isolates of *Pseudomonas aeruginosa* in the Czech Republic', *Research in Microbiology*, 161(3), pp. 234–242. doi: 10.1016/j.resmic.2010.02.002.

Nessar, R. *et al.* (2011) 'Genetic analysis of new 16S rRNA mutations conferring aminoglycoside resistance in *Mycobacterium abscessus*', *Journal of Antimicrobial Chemotherapy*, 66(8), pp. 1719–1724. doi: 10.1093/jac/dkr209.

Nikaido, H. and Pagès, J.-M. (2012) 'Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria', *FEMS Microbiology Reviews*, 36(2), pp. 340–363. doi: 10.1111/j.1574-6976.2011.00290.x.

Nikolayevskyy, V. *et al.* (2019) 'Role and value of whole genome sequencing in studying tuberculosis transmission', *Clinical Microbiology and Infection*, 25(11), pp. 1377–1382. doi: https://doi.org/10.1016/j.cmi.2019.03.022.

van Nimwegen, K. J. M. *et al.* (2016) 'Is the $1000 Genome as Near as We Think? A Cost Analysis of Next-Generation Sequencing', *Clinical Chemistry*, 62(11), pp. 1458–1464. doi: 10.1373/clinchem.2016.258632.

Noah, K. (2019) *Genome-wide association study of Pseudomonas aeruginosa adaptations to the cystic fibrosis lung*. Carleton University Ottawa, Ontario. doi: https://doi.org/10.22215/etd/2019-13815.

Nouri, R. *et al.* (2016) 'The role of *gyr*A and *par*C mutations in fluoroquinolones-resistant *Pseudomonas aeruginosa* isolates from Iran', *Brazilian Journal of Microbiology*. Elsevier, 47(4), pp. 925–930. doi: 10.1016/J.BJM.2016.07.016.

Nyfeler, B. *et al.* (2012) 'Identification of elongation factor g as the conserved cellular target of argyrin B', *PLoS ONE*. Edited by B. G. Vertessy, 7(9), p. e42657. doi: 10.1371/journal.pone.0042657.

O' Neil, J. (2014) 'Review on Antibiotic resisitance. Antimicrobial Resistance : Tackling a crisis for the health and wealth of nations', *Health and Wealth Nations*, (December), pp. 1–16. Available at: https://amr-review.org/sites/default/files/AMR Review Paper - Tackling a crisis for the health and wealth of nations_1.pdf.

O'Brien, S. *et al.* (2017) 'High virulence sub-populations in *Pseudomonas aeruginosa* long-term cystic fibrosis airway infections', *BMC microbiology*. BioMed Central, 17(1), p. 30. doi: 10.1186/s12866-017-0941-6.

O'Brien, T. P. *et al.* (1995) 'Efficacy of ofloxacin vs cefazolin and tobramycin in the therapy for bacterial keratitis. Report from the Bacterial Keratitis Study Research Group.', *Archives of ophthalmology (Chicago, Ill. : 1960)*. United States, 113(10), pp. 1257–1265. doi: 10.1001/archopht.1995.01100100045026.

O'Grady, J. (2019) 'A powerful, non-invasive test to rule out infection', *Nature Microbiology*. Springer US, 4(4), pp. 554–555. doi: 10.1038/s41564-019-0424-7.

O'Neill, J. (2016) *Tackling drug-resistant infections globally: final report and recommendations*. Available at: https://www.semanticscholar.org/paper/Tackling-drug-resistant-infections-globally.

Oh, H. *et al.* (2003) 'Role of Efflux Pumps and Mutations in Genes for Topoisomerases II and IV in Fluoroquinolone-Resistant *Pseudomonas aeruginosa* Strains', *Microbial Drug Resistance*, 9(4), pp. 323–328. doi: 10.1089/107662903322762743.

Oliveira, V. D. C. *et al.* (2015) 'Trends of 9,416 multidrug-resistant Gram-negative bacteria', *Revista da Associação Médica Brasileira*. Associação Médica Brasileira, 61(3), pp. 244–249. doi: 10.1590/1806-9282.61.03.244.

Oliver, A. *et al.* (2000) 'High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection', *Science*, 288(5469), pp. 1251–1253. doi: 10.1126/science.288.5469.1251.

Oliver, A. *et al.* (2015) 'The increasing threat of *Pseudomonas aeruginosa* high-risk clones', *Drug Resistance Updates*. Churchill Livingstone, pp. 41–59. doi: 10.1016/j.drup.2015.08.002.

Omae, Y. *et al.* (2017) 'Pathogen lineage-based genome-wide association study identified CD53 as susceptible locus in tuberculosis', *Journal of Human Genetics*. Nature Publishing Group, 62(12), pp. 1015–1022. doi: 10.1038/jhg.2017.82.

Oppong, Y. E. A. *et al.* (2019) 'Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance', *BMC Genomics*. BMC Genomics, 20(1), pp. 1–15. doi: 10.1186/s12864-019-5615-3.

Osawa, K. *et al.* (2019) 'Molecular characteristics of carbapenem-resistant *Pseudomonas aeruginosa* isolated from urine in Hyogo, Japan', *International Journal of Urology*, 26(1), pp. 127–133. doi: 10.1111/iju.13818.

Ozer, E. A., Allen, J. P. and Hauser, A. R. (2014) 'Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt', *BMC Genomics*. BioMed Central, 15(1), p. 737. doi: 10.1186/1471-2164-15-737.

Ozongwu, C. *et al.* (2017) 'The Unyvero P55 "sample-in, answer-out" pneumonia assay: A performance evaluation.', *Biomolecular detection and quantification*. Elsevier, 13, pp. 1–6. doi: 10.1016/j.bdq.2017.06.001.

Padidam, M., Sawyer, S. and Fauquet, C. M. (1999) 'Possible Emergence of New Geminiviruses by Frequent Recombination', *Virology*, 265(2), pp. 218–225. doi: 10.1006/viro.1999.0056.

Pan, Y.-P. *et al.* (2016) 'Overexpression of MexAB-OprM efflux pump in carbapenem-resistant *Pseudomonas aeruginosa*.', *Archives of microbiology*, 198(6), pp. 565–71. doi: 10.1007/s00203-016-1215-7.

Parcell, B. J. *et al.* (2018) '*Pseudomonas aeruginosa* intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing', *The Journal of hospital infection*. 2017/12/08. W.B. Saunders For The Hospital Infection Society, 98(3), pp. 282–288. doi: 10.1016/j.jhin.2017.12.005.

Patel, P. A. *et al.* (2011) 'Performance of the BD GeneOhm MRSA achromopeptidase assay for real-time PCR detection of methicillin-resistant *Staphylococcus aureus* in nasal specimens.', *Journal of clinical microbiology*. American Society for Microbiology (ASM), 49(6), pp. 2266–8. doi: 10.1128/JCM.02431-10.

Patzer, J. and Dzierzanowska, D. (1991) 'The resistance patterns and serotypes of *Pseudomonas aeruginosa* strains isolated from children.', *The Journal of antimicrobial chemotherapy*, 28(6), pp. 869–75.

Pavlova, S. I. *et al.* (2013) 'Multiple alcohol dehydrogenases but no functional acetaldehyde dehydrogenase causing excessive acetaldehyde production from ethanol by oral streptococci.', *Microbiology (Reading, England)*. Microbiology Society, 159(Pt 7), pp. 1437–46. doi: 10.1099/mic.0.066258-0.

Pehrsson, E. C. *et al.* (2013) 'Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs', *Frontiers in Microbiology*, 4, p. 145. doi:

10.3389/fmicb.2013.00145.

Pelegrin, A. C., Saharman, Y. R., Griffon, A., Palmieri, M., Mirande, C., Karuniawati, A., Sedono, R., Aditianingsih, D., Goessens, W. H. F., *et al.* (2019) 'High-risk international clones of carbapenem-nonsusceptible *Pseudomonas aeruginosa* endemic to Indonesian intensive care units: Impact of a multifaceted infection control intervention analyzed at the genomic level', *mBio*, 10(6), pp. 1–14. doi: 10.1128/mBio.02384-19.

Pena, C. *et al.* (2015) 'Influence of Virulence Genotype and Resistance Profile in the Mortality of *Pseudomonas aeruginosa* Bloodstream Infections', *Clinical Infectious Diseases*. Oxford University Press, 60(4), pp. 539–548. doi: 10.1093/cid/ciu866.

Pendleton, J. N., Gorman, S. P. and Gilmore, B. F. (2013) 'Clinical relevance of the ESKAPE pathogens', *Expert Review of Anti-infective Therapy*, 11(3), pp. 297–308. doi: 10.1586/eri.13.12.

Perdigao, J. *et al.* (2010) 'Genetic analysis of extensively drug-resistant Mycobacterium tuberculosis strains in Lisbon, Portugal', *Journal of Antimicrobial Chemotherapy*, 65(2), pp. 224–227. doi: 10.1093/jac/dkp452.

Pérez, A. *et al.* (2019) 'High incidence of MDR and XDR *Pseudomonas aeruginosa* isolates obtained from patients with ventilator-associated pneumonia in Greece, Italy and Spain as part of the MagicBullet clinical trial', *Journal of Antimicrobial Chemotherapy*, 74(5), pp. 1244–1252. doi: 10.1093/jac/dkz030.

Perry, J. A., Westman, E. L. and Wright, G. D. (2014) 'The antibiotic resistome: what's new?', *Current opinion in microbiology*, 21, pp. 45–50. doi: 10.1016/j.mib.2014.09.002.

Picard, B. *et al.* (1994) 'Genetic heterogeneity of *Pseudomonas aeruginosa* clinical isolates revealed by esterase electrophoretic polymorphism and restriction fragment length polymorphism of the ribosomal RNA gene region', *Journal of Medical Microbiology*, 40(5), pp. 313–322. doi: 10.1099/00222615-40-5-313.

Piddock, L. J. (1999) 'Mechanisms of fluoroquinolone resistance: an update 1994-1998.', *Drugs*, 58 Suppl 2, pp. 11–8.

Pirnay, J.-P. *et al.* (2002) '*Pseudomonas aeruginosa* displays an epidemic population structure', *Environmental Microbiology*, 4(12). doi: 10.1046/j.1462-2920.2002.00321.x.

Pirnay, J.-P. *et al.* (2009) '*Pseudomonas aeruginosa* population structure revisited', *PLoS ONE*, 4(11). doi: 10.1371/journal.pone.0007740.

Pitt, T. L. *et al.* (1990) 'Resistance mechanisms of multiresistant serotype 012 *Pseudomonas aeruginosa* isolated in Europe', *Journal of Antimicrobial Chemotherapy*, 26(3), pp. 319–328. doi: 10.1093/jac/26.3.319.

Polz, M. F., Alm, E. J. and Hanage, W. P. (2013) 'Horizontal gene transfer and the evolution of bacterial and archaeal population structure', *Trends in Genetics*, 29(3), pp. 170–175. doi: 10.1016/j.tig.2012.12.006.

Poole, K. (2005) 'Aminoglycoside Resistance in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*. American Society for Microbiology (ASM), 49(2), p. 479. doi: 10.1128/AAC.49.2.479-487.2005.

Poole, K. (2011) '*Pseudomonas Aeruginosa*: Resistance to the Max', *Frontiers in Microbiology*, 2, p. 65. doi: 10.3389/fmicb.2011.00065.

Poole, K. and Srikumar, R. (2001) 'Multidrug efflux in *Pseudomonas aeruginosa*: components, mechanisms and clinical significance.', *Current topics in medicinal chemistry*, 1(1), pp. 59–71.

Poonsuk, K., Tribuddharat, C. and Chuanchuen, R. (2013) 'Aminoglycoside resistance mechanisms in *Pseudomonas aeruginosa* isolates from non-cystic fibrosis patients in Thailand', *Canadian Journal of Microbiology*, 59(1), pp. 51–56. doi: 10.1139/cjm-2012-0465.

Poritz, M. A. *et al.* (2011) 'FilmArray, an Automated Nested Multiplex PCR System for Multi-Pathogen Detection: Development and Application to Respiratory Tract Infection', *PLoS ONE*. Public Library of Science, 6(10), p. e26047. doi: 10.1371/JOURNAL.PONE.0026047.

Potz, N. A. C. *et al.* (2004) 'Reliability of routine disc susceptibility testing by the British Society for Antimicrobial Chemotherapy (BSAC) method', *Journal of Antimicrobial Chemotherapy*, 53(5), pp. 729–738. doi: 10.1093/jac/dkh212.

Poulsen, B. E. *et al.* (2019) 'Defining the core essential genome of *Pseudomonas aeruginosa*', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 116(20), pp. 10072–10080. doi: 10.1073/pnas.1900570116.

Power, R. A., Parkhill, J. and de Oliveira, T. (2017) 'Microbial genome-wide association studies: lessons from human GWAS', *Nature Reviews Genetics*, 18(1), pp. 41–50. doi: 10.1038/nrg.2016.132.

Prakash, O. *et al.* (2007) 'Polyphasic approach of bacterial classification - An overview of recent advances.', *Indian journal of microbiology*. Springer, 47(2), pp. 98–108. doi: 10.1007/s12088-007-0022-x.

Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) 'FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments', *PLoS ONE*. Edited by A. F. Y. Poon. Public Library of Science, 5(3), p. e9490. doi: 10.1371/journal.pone.0009490.

Purssell, A. and Poole, K. (2013) 'Functional characterization of the NfxB repressor of the *mex*CD-oprJ multidrug efflux operon of *Pseudomonas aeruginosa*.', *Microbiology (Reading, England)*, 159(Pt 10), pp. 2058–73. doi: 10.1099/mic.0.069286-0.

Quainoo, S. *et al.* (2017) 'Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis', *Clinical Microbiology Reviews*, 30(4), p. 1015 LP-1063. doi: 10.1128/CMR.00016-17.

Quick, J. *et al.* (2016) 'Real-time, portable genome sequencing for Ebola surveillance', *Nature*, 530(7589), pp. 228–232. doi: 10.1038/nature16996.

Quick, J. *et al.* (2017) 'Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples', *Nature Protocols*, 12(6), pp. 1261–1276. doi: 10.1038/nprot.2017.066.

Quick, J., Quinlan, A. R. and Loman, N. J. (2014) 'A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer', *GigaScience*, 3(1), p. 22. doi: 10.1186/2047-217X-3-22.

Radlinski, L. C. *et al.* (2019) 'Chemical Induction of Aminoglycoside Uptake Overcomes Antibiotic Tolerance and Resistance in *Staphylococcus aureus*', *Cell Chemical Biology*. Elsevier Ltd., 26(10), p. 1355–1364.e4. doi: 10.1016/j.chembiol.2019.07.009.

Raja, N. S. and Singh, N. N. (2007) 'Antimicrobial susceptibility pattern of clinical isolates of *Pseudomonas aeruginosa* in a tertiary care hospital.', *Journal of microbiology, immunology, and*

infection, 40(1), pp. 45–9.

Ramirez, M. S. and Tolmasky, M. E. (2017) 'Amikacin: Uses, Resistance, and Prospects for Inhibition', *Molecules (Basel, Switzerland)*. MDPI, 22(12), p. 2267. doi: 10.3390/molecules22122267.

Ramond, E. *et al.* (2014) 'Glutamate Utilization Couples Oxidative Stress Defense and the Tricarboxylic Acid Cycle in Francisella Phagosomal Escape', *PLoS Pathogens*. Edited by D. Weiss, 10(1), p. e1003893. doi: 10.1371/journal.ppat.1003893.

Ranjbar, R. *et al.* (2014) 'Typing methods used in the molecular epidemiology of microbial pathogens: a how-to guide.', *The new microbiologica*, 37(1), pp. 1–15. Available at: http://www.newmicrobiologica.org/PUB/allegati_pdf/2014/1/1.pdf.

Read, T. D. and Massey, R. C. (2014) 'Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: A new direction for bacteriology', *Genome Medicine*, 6(11), pp. 1–11. doi: 10.1186/s13073-014-0109-z.

Recio, R. *et al.* (2018) 'Bacteraemia due to extensively drug-resistant *Pseudomonas aeruginosa* sequence type 235 high-risk clone: Facing the perfect storm', *International Journal of Antimicrobial Agents*, 52(2), pp. 172–179. doi: 10.1016/j.ijantimicag.2018.03.018.

Recio, R. *et al.* (2019) 'Pathogenic characteristics of *Pseudomonas aeruginosa* bacteraemia isolates in a high-endemicity setting for ST175 and ST235 high-risk clones', *European Journal of Clinical Microbiology and Infectious Diseases*. European Journal of Clinical Microbiology & Infectious Diseases. doi: 10.1007/s10096-019-03780-z.

Redfern, J. *et al.* (2019) 'Biofilm Associated Genotypes of Multidrug-Resistant *Pseudomonas aeruginosa*, *bioRxiv*, p. 713453. doi: 10.1101/713453.

Redgrave, L. S. *et al.* (2014) 'Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success', *Trends in Microbiology*, 22(8), pp. 438–445. doi: 10.1016/j.tim.2014.04.007.

Rehman, A., Patrick, W. M. and Lamont, I. L. (2019) 'Mechanisms of ciprofloxacin resistance in *Pseudomonas aeruginosa*: New approaches to an old problem', *Journal of Medical Microbiology*, 68(1), pp. 1–10. doi: 10.1099/jmm.0.000873.

Reuter, S. *et al.* (2013) 'Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology', *JAMA Internal Medicine*. doi: 10.1001/jamainternmed.2013.7734.

Reynolds, R., Hope, R. and Williams, L. (2008) 'Survey, laboratory and statistical methods for the BSAC Resistance Surveillance Programmes', *Journal of Antimicrobial Chemotherapy*. Oxford University Press, 62 (Supplement 2), pp. ii15-ii28. doi: 10.1093/jac/dkn349.

Rice, K. C. *et al.* (2003) 'The *Staphylococcus aureus* cidAB operon: Evaluation of its role in regulation of murein hydrolase activity and penicillin tolerance', *Journal of Bacteriology*, 185(8), pp. 2635–2643. doi: 10.1128/JB.185.8.2635-2643.2003.

Richardot, C. *et al.* (2016) 'Amino Acid Substitutions Account for Most MexS Alterations in Clinical nfxC Mutants of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 60(4), pp. 2302–10. doi: 10.1128/AAC.02622-15.

Ritchie, N. D. and Evans, T. J. (2019) ' Dual RNA-seq in *Streptococcus pneumoniae* Infection Reveals Compartmentalized Neutrophil Responses in Lung and Pleural Space ', *mSystems*, 4(4), pp. 1–17. doi: 10.1128/msystems.00216-19.

Rommens, J. M. *et al.* (1989) 'Identification of the cystic fibrosis gene: chromosome walking and jumping', *Science*, 245(4922), p. 1059 LP-1065. doi: 10.1126/science.2772657.

Rossen, J. W. A. *et al.* (2018) 'Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology', *Clinical Microbiology and Infection*, 24(4), pp. 355–360. doi: 10.1016/j.cmi.2017.11.001.

Roy Chowdhury, P., Scott, M. J. and Djordjevic, S. P. (2017) 'Genomic islands 1 and 2 carry multiple antibiotic resistance genes in *Pseudomonas aeruginosa* ST235, ST253, ST111 and ST175 and are globally dispersed', *Journal of Antimicrobial Chemotherapy*. Narnia, 72(2), pp. 620–622. doi: 10.1093/jac/dkw471.

Royer, G. *et al.* (2020) 'Local outbreak of extended-spectrum b -lactamase SHV2a-producing *Pseudomonas aeruginosa* reveals the emergence of a new specific sub-lineage of the international ST235 high-risk clone', *Journal of Hospital Infection*. Elsevier Ltd, 104(1), pp. 33–39. doi: 10.1016/j.jhin.2019.07.014.

Ruesen, C. *et al.* (2018) 'Large-scale genomic analysis shows association between homoplastic

genetic variation in *Mycobacterium tuberculosis* genes and meningeal or pulmonary tuberculosis', *BMC Genomics*. BMC Genomics, 19(1), pp. 1–11. doi: 10.1186/s12864-018-4498-z.

Ruimy, R. *et al.* (2001) 'Genetic diversity of *Pseudomonas aeruginosa* strains isolated from ventilated patients with nosocomial pneumonia, cancer patients with bacteremia, and environmental water.', *Infection and immunity*. American Society for Microbiology, 69(1), pp. 584–8. doi: 10.1128/IAI.69.1.584-588.2001.

Rules, E. (2018) 'The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone Version 9.0, 2019.', pp. 0–99. doi: 10.1111/j.0021-8782.2005.00369.x.

Ryan, A. *et al.* (2010) 'Reaction mechanism of azoreductases suggests convergent evolution with quinone oxidoreductases.', *Protein & cell*, 1(8), pp. 780–90. doi: 10.1007/s13238-010-0090-2.

Ryan, A. *et al.* (2011) 'Activation of nitrofurazone by azoreductases: multiple activities in one enzyme.', *Scientific reports*, 1(1), p. 63. doi: 10.1038/srep00063.

Ryan, A. *et al.* (2014) 'Identification of NAD(P)H Quinone Oxidoreductase Activity in Azoreductases from *P. aeruginosa*: Azoreductases and NAD(P)H Quinone Oxidoreductases Belong to the Same FMN-Dependent Superfamily of Enzymes', *PLoS ONE*. Edited by L. Es. Netto, 9(6), p. e98551. doi: 10.1371/journal.pone.0098551.

Rybenkov, V. V (2014) 'Maintenance of chromosome structure in *Pseudomonas aeruginosa*', *FEMS microbiology letters*. 2014/06/12, 356(2), pp. 154–165. doi: 10.1111/1574-6968.12478.

Sabat, A. J. *et al.* (2017) 'Targeted next-generation sequencing of the 16S-23S rRNA region for culture-independent bacterial identification - increased discrimination of closely related species', *Scientific Reports*. Nature Publishing Group, 7(1), p. 3434. doi: 10.1038/s41598-017-03458-6.

SALASPURO, M. (1997) 'INVITED REVIEW Microbial metabolism of ethanol and acetaldehyde and clinical consequences', *Addiction Biology*, 2(1), pp. 35–46. doi: 10.1080/13556219772840.

Salipante, S. J. *et al.* (2015) 'Large-scale genomic sequencing of extraintestinal pathogenic *Escherichia coli* strains', *Genome Research*, 25(1), pp. 119–128. doi: 10.1101/gr.180190.114.

Salma, R. *et al.* (2013) 'gyrA and parC mutations in quinolone-resistant clinical isolates of *Pseudomonas aeruginosa* from Nini Hospital in north Lebanon', *Journal of Infection and Chemotherapy*. Springer Japan, 19(1), pp. 77–81. doi: 10.1007/s10156-012-0455-y.

Samuelsen, O. *et al.* (2010) 'Molecular Epidemiology of Metallo- -Lactamase-Producing *Pseudomonas aeruginosa* Isolates from Norway and Sweden Shows Import of International Clones and Local Clonal Expansion', *Antimicrobial Agents and Chemotherapy*, 54(1), pp. 346–352. doi: 10.1128/AAC.00824-09.

Sanz-García, F., Hernando-Amado, S. and Martínez, J. L. (2018) 'Mutational Evolution of *Pseudomonas aeruginosa* Resistance to Ribosome-Targeting Antibiotics', *Frontiers in Genetics*. Frontiers, 9, p. 451. doi: 10.3389/fgene.2018.00451.

Schadt, E. E., Turner, S. and Kasarskis, A. (2010) 'A window into third-generation sequencing', *Human Molecular Genetics*, 19(R2), pp. R227–R240. doi: 10.1093/hmg/ddq416.

Schatz, A., Bugle, E. and Waksman, S. A. (1944) 'Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.', *Experimental Biology and Medicine*. SAGE PublicationsSage UK: London, England, 55(1), pp. 66–69. doi: 10.3181/00379727-55-14461.

Schaumburg, F. *et al.* (2017) 'Susceptibility of MDR *Pseudomonas aeruginosa* to ceftolozane/tazobactam and comparison of different susceptibility testing methods', *Journal of Antimicrobial Chemotherapy*, 72(11), pp. 3079–3084. doi: 10.1093/jac/dkx253.

Schito, A. M. and Schito, G. C. (2004) 'Levofloxacin, a broad spectrum anti-infective: from *Streptococcus pneumoniae* to *Pseudomonas aeruginosa*.', *Journal of chemotherapy (Florence, Italy)*. England, 16 Suppl 2, pp. 3–7. doi: 10.1080/1120009x.2004.11782365.

Schmidt, K. *et al.* (2017) 'Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing', *Journal of Antimicrobial Chemotherapy*. Narnia, 72(1), pp. 104–114. doi: 10.1093/jac/dkw397.

Schmieder, R. and Edwards, R. (2012) 'Insights into antibiotic resistance through metagenomic approaches', *Future Microbiology*, 7(1), pp. 73–89. doi: 10.2217/fmb.11.135.

Schmitz, J. E. and Tang, Y.-W. (2018) 'The GenMark ePlex ® : another weapon in the syndromic

arsenal for infection diagnosis', *Future Microbiology*, 13(16), pp. 1697–1708. doi: 10.2217/fmb-2018-0258.

Schniederjans, M., Koska, M. and Häussler, S. (2017) 'Transcriptional and Mutational Profiling of an Aminoglycoside-Resistant *Pseudomonas aeruginosa* Small-Colony Variant', *Antimicrobial Agents and Chemotherapy*, 61(11). doi: 10.1128/AAC.01178-17.

Schweizer, H. P. (2003) 'Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions.', *Genetics and molecular research : GMR*, 2(1), pp. 48–62.

Segatore, B. *et al.* (1999) 'Italian survey on comparative levofloxacin susceptibility in 334 clinical isolates of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*, 43(2), pp. 428–431.

Segatore, B. *et al.* (2000) 'Bactericidal activity of levofloxacin and ciprofloxacin on clinical isolates of different phenotypes of *Pseudomonas aeruginosa*', *International Journal of Antimicrobial Agents*, 13(3), pp. 223–226. doi: https://doi.org/10.1016/S0924-8579(99)00119-3.

Seth-Smith, H. M. B. *et al.* (2013) 'Whole-genome sequences of *Chlamydia trachomatis* directly from clinical samples without culture.', *Genome research*. Cold Spring Harbor Laboratory Press, 23(5), pp. 855–66. doi: 10.1101/gr.150037.112.

Severiano, A. *et al.* (2011) 'Adjusted Wallace coefficient as a measure of congruence between typing methods.', *Journal of clinical microbiology*. American Society for Microbiology Journals, 49(11), pp. 3997–4000. doi: 10.1128/JCM.00624-11.

Shaffer, L. (2019) 'Inner Workings: Portable DNA sequencer helps farmers stymie devastating viruses.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 116(9), pp. 3351–3353. doi: 10.1073/pnas.1901806116.

Shahab Qureshi, MD, FACP; Chief Editor: Michael Stuart Bronze, M. (2020) *Pseudomonas aeruginosa Infections Treatment & Management, emedicine, Medscape*. Available at: https://emedicine.medscape.com/article/226748-treatment.

Sharma, C. M. and Vogel, J. (2014) 'Differential RNA-seq: The approach behind and the biological insight gained', *Current Opinion in Microbiology*, 19(1), pp. 97–105. doi:

10.1016/j.mib.2014.06.010.

Shaughnessy, A. F. (2007) ‘Clinical Epidemiology: A Basic Science for Clinical Medicine’, *BMJ*, 335(7623), p. 777 LP-777. doi: 10.1136/bmj.39349.461678.43.

Shearer, B. G. and Legakis, N. J. (1985) ‘*Pseudomonas aeruginosa*: Evidence for the Involvement of Lipopolysaccharide in Determining Outer Membrane Permeability to Carbenicillin and Gentamicin’, *Journal of Infectious Diseases*, 152(2), pp. 351–355. doi: 10.1093/infdis/152.2.351.

Sheehan, D. *et al.* (2001) ‘Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily.’, *The Biochemical journal*. Portland Press Ltd, 360(Pt 1), pp. 1–16.

Sheppard, S. K. *et al.* (2013) ‘Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter’, *Proceedings of the National Academy of Sciences of the United States of America*, 110(29), pp. 11923–11927. doi: 10.1073/pnas.1305559110.

Shibata, N. *et al.* (2010) ‘Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B12 analogs and substrates.’, *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 285(34), pp. 26484–93. doi: 10.1074/jbc.M110.125112.

Silby, M. W. *et al.* (2011) ‘*Pseudomonas* genomes: diverse and adaptable’, *FEMS Microbiology Reviews*. John Wiley & Sons, Ltd (10.1111), 35(4), pp. 652–680. doi: 10.1111/j.1574-6976.2011.00269.x.

Singh, P. B., Saini, H. S. and Kahlon, R. S. (2016) ‘*Pseudomonas*: The Versatile and Adaptive Metabolic Network’, in *Pseudomonas: Molecular and Applied Biology*. Cham: Springer International Publishing, pp. 81–126. doi: 10.1007/978-3-319-31198-2_3.

Sirgel, F. A. *et al.* (2012) ‘Mutations in the *rrs* A1401G Gene and Phenotypic Resistance to Amikacin and Capreomycin in *Mycobacterium tuberculosis*’, *Microbial Drug Resistance*, 18(2), pp. 193–197. doi: 10.1089/mdr.2011.0063.

Smith, C. A. and Baker, E. N. (2002) ‘Aminoglycoside antibiotic resistance by enzymatic deactivation.’, *Current drug targets. Infectious disorders*, 2(2), pp. 143–60.

Smith, J. (1992) ‘Analyzing the mosaic structure of genes’, *Journal of Molecular Evolution*. Springer-Verlag, 34(2), pp. 126–129. doi: 10.1007/BF00182389.

Smith, J. M. *et al.* (1993) ‘How clonal are bacteria?’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 90(10), pp. 4384–8. doi: 10.1073/pnas.90.10.4384.

Snyder, L. A. *et al.* (2013) ‘Epidemiological investigation of *Pseudomonas aeruginosa* isolates from a six-year-long hospital outbreak using high-throughput whole genome sequencing.’, *Euro surveillance : bulletin Europeen sur les maladies transmissibles, European communicable disease bulletin*, 18(42).

Sobel, M. L. *et al.* (2005) ‘Mutations in PA3574 (*nal*D) lead to increased MexAB-OprM expression and multidrug resistance in laboratory and clinical isolates of *Pseudomonas aeruginosa.*’, *Antimicrobial agents and chemotherapy*, 49(5), pp. 1782–6. doi: 10.1128/AAC.49.5.1782-1786.2005.

Sobel, M. L., Neshat, S. and Poole, K. (2005) ‘Mutations in PA2491 (*mex*S) promote MexT-dependent *mex*EF-oprN expression and multidrug resistance in a clinical strain of *Pseudomonas aeruginosa.*’, *Journal of bacteriology*, 187(4), pp. 1246–53. doi: 10.1128/JB.187.4.1246-1253.2005.

Sobral, D. *et al.* (2012) ‘A new highly discriminatory multiplex capillary-based MLVA assay as a tool for the epidemiological survey of *Pseudomonas aeruginosa* in cystic fibrosis patients’, *European Journal of Clinical Microbiology & Infectious Diseases*. Springer-Verlag, 31(9), pp. 2247–2256. doi: 10.1007/s10096-012-1562-5.

Sommer, M. O. A., Dantas, G. and Church, G. M. (2009) ‘Functional Characterization of the Antibiotic Resistance Reservoir in the Human Microflora’, *Science*, 325(5944), pp. 1128–1131. doi: 10.1126/science.1176950.

Spencer, M. D. *et al.* (2019) ‘Whole Genome Sequencing detects Inter-Facility Transmission of Carbapenem-resistant *Klebsiella pneumoniae*’, *Journal of Infection*. W.B. Saunders, 78(3), pp. 187–199. doi: 10.1016/J.JINF.2018.11.003.

Springer, B. *et al.* (2001) ‘Mechanisms of Streptomycin Resistance: Selection of Mutations in

the 16S rRNA Gene Conferring Resistance', *Antimicrobial Agents and Chemotherapy*, 45(10), pp. 2877–2884. doi: 10.1128/AAC.45.10.2877-2884.2001.

SPSS.V21, I. C. (no date) 'IPM SPSS'. IBM Corp. Released 2013. IBM SPSS Statistics for W; Version 21.0. Armonk; NY: IBM Corp.

Srikumar, R., Li, X. Z. and Poole, K. (1997) 'Inner membrane efflux components are responsible for beta-lactam specificity of multidrug efflux pumps in *Pseudomonas aeruginosa*.', *Journal of bacteriology*, 179(24), pp. 7875–81.

Srikumar, R., Paul, C. J. and Poole, K. (2000) 'Influence of mutations in the *mex*R repressor gene on expression of the MexA-MexB-oprM multidrug efflux system of *Pseudomonas aeruginosa*.', *Journal of bacteriology*, 182(5), pp. 1410–4.

Stanton, J.-A. L. *et al.* (2019) 'Rapid extraction of DNA suitable for NGS workflows from bacterial cultures using the PDQeX', *BioTechniques*. Future Science Ltd London, UK , 66(5), pp. 208–213. doi: 10.2144/btn-2019-0006.

Steiner, P. A. *et al.* (2019) 'Highly variable mRNA half-life time within marine bacterial taxa and functional genes', *Environmental Microbiology*, 21(10), pp. 3873–3884. doi: 10.1111/1462-2920.14737.

Stoesser, N. *et al.* (2013) 'Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data.', *The Journal of antimicrobial chemotherapy*. Oxford University Press, 68(10), pp. 2234–44. doi: 10.1093/jac/dkt180.

Stokes, H. W. *et al.* (2012) 'Class 1 integron-associated spread of resistance regions in *Pseudomonas aeruginosa*: Plasmid or chromosomal platforms?', *Journal of Antimicrobial Chemotherapy*, pp. 1799–1800. doi: 10.1093/jac/dks116.

Subedi, D. *et al.* (2018) 'Comparative genomics of clinical strains of *Pseudomonas aeruginosa* strains isolated from different geographic sites', *Scientific Reports*. Nature Publishing Group, 8(1). doi: 10.1038/s41598-018-34020-7.

Suresh, M. *et al.* (2018) 'Mutational analyses of regulatory genes, *mex*R, *nal*C, *nal*D and *mex*Z of *mex*AB-oprM and *mex*XY operons, in efflux pump hyperexpressing multidrug-resistant

clinical isolates of *Pseudomonas aeruginosa*', *World Journal of Microbiology and Biotechnology*, 34(6), p. 83. doi: 10.1007/s11274-018-2465-0.

Sutherland, C. A., Verastegui, J. E. and Nicolau, D. P. (2016) 'In vitro potency of amikacin and comparators against *E. coli*, *K. pneumoniae* and *P. aeruginosa* respiratory and blood isolates', *Annals of Clinical Microbiology and Antimicrobials*, 15(1), p. 39. doi: 10.1186/s12941-016-0155-z.

Sutton, D. *et al.* (2019) 'Genome-Wide Identification of Myxobacterial Predation Genes and Demonstration of Formaldehyde Secretion as a Potentially Predation-Resistant Trait of *Pseudomonas aeruginosa*', *Frontiers in Microbiology*, 10 (November), pp. 1–9. doi: 10.3389/fmicb.2019.02650.

Taber, H. W. *et al.* (1987) 'Bacterial uptake of aminoglycoside antibiotics.', *Microbiological reviews*, 51(4), pp. 439–57.

Tankovic, J. *et al.* (1996) 'Contribution of mutations in *gyr*A and *par*C genes to fluoroquinolone resistance of mutants of *Streptococcus pneumoniae* obtained in vivo and in vitro.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 40(11), pp. 2505–10.

Teixeira Rodrigues, A. *et al.* (2013) 'Understanding physician antibiotic prescribing behaviour: a systematic review of qualitative studies', *International Journal of Antimicrobial Agents*. Elsevier, 41(3), pp. 203–212. doi: 10.1016/J.IJANTIMICAG.2012.09.003.

Tettelin, H. *et al.* (2008) 'Comparative genomics: the bacterial pan-genome', *Current Opinion in Microbiology*, 11(5), pp. 472–477. doi: 10.1016/j.mib.2008.09.006.

Thrane, S. W., Taylor, V. L., *et al.* (2016) 'Application of Whole-Genome Sequencing Data for O-Specific Antigen Analysis and In Silico Serotyping of *Pseudomonas aeruginosa* Isolates.', *Journal of clinical microbiology*. American Society for Microbiology Journals, 54(7), pp. 1782–1788. doi: 10.1128/JCM.00349-16.

Thrane, S. W., Taylor, V. L., *et al.* (2016) 'Application of whole-genome sequencing data for o-specific antigen analysis and in silico serotyping of *Pseudomonas aeruginosa* isolates', *Journal of Clinical Microbiology*, 54(7), pp. 1782–1788. doi: 10.1128/JCM.00349-16.

Torres, M. P. *et al.* (2016) 'A multiplex real-time PCR assay for identification of multi-drug

resistant Gram-negative pathogens', *BioTechniques*, 61(4). doi: 10.2144/000114464.

Towner, K. J. and Cockayne, A. (1993) 'An introduction to microbial identification and typing methods', in *Molecular Methods for Microbial Identification and Typing*. First edit. Chapman and Hall.

Tramper-Stranders, G. A. *et al.* (2012) 'Initial *Pseudomonas aeruginosa* infection in patients with cystic fibrosis: characteristics of eradicated and persistent isolates', *Clinical Microbiology and Infection*, 18(6), pp. 567–574. doi: https://doi.org/10.1111/j.1469-0691.2011.03627.x.

Treepong, P. *et al.* (2017) 'Global emergence of the widespread *Pseudomonas aeruginosa* ST235 clone', *Clinical Microbiology and Infection*. doi: 10.1016/j.cmi.2017.06.018.

Tsoy, O., Ravcheev, D. and Mushegian, A. (2009) 'Comparative genomics of ethanolamine utilization.', *Journal of bacteriology*. American Society for Microbiology Journals, 191(23), pp. 7157–64. doi: 10.1128/JB.00838-09.

Tsuji, A. *et al.* (2005) 'An epidemiological study of the susceptibility and frequency of multiple-drug-resistant strains of *Pseudomonas aeruginosa* isolated at medical institutes nationwide in Japan', *J Infect Chemother*, 11, pp. 64–70. doi: 10.1007/s10156-005-0377-z.

Tuite, N. *et al.* (2014) 'Rapid nucleic acid diagnostics for the detection of antimicrobial resistance in Gram-negative bacteria: is it time for a paradigm shift?', *Journal of Antimicrobial Chemotherapy*, 69(7), pp. 1729–1733. doi: 10.1093/jac/dku083.

Turton, J. F. *et al.* (2010) 'Evaluation of a nine-locus variable-number tandem-repeat scheme for typing of *Pseudomonas aeruginosa*', *Clinical Microbiology and Infection*. Blackwell Publishing Ltd, 16(8), pp. 1111–1116. doi: 10.1111/j.1469-0691.2009.03049.x.

US CDC (2019) 'Antibiotic resistance threats in the United States', *Centers for Disease Control and Prevention*, pp. 1–150.

Vaez, H. *et al.* (2014) 'Efflux pump regulatory genes mutations in multidrug resistance *Pseudomonas aeruginosa* isolated from wound infections in Isfahan hospitals.', *Advanced biomedical research*, 3(1), p. 117. doi: 10.4103/2277-9175.133183.

Valot, B. *et al.* (2015) 'What It Takes to Be a *Pseudomonas aeruginosa*? The Core Genome of

the Opportunistic Pathogen Updated', *PLOS ONE*. Edited by G. F. Kaufmann. Public Library of Science, 10(5), p. e0126468. doi: 10.1371/journal.pone.0126468.

Varin, A. *et al.* (2017) 'High prevalence and moderate diversity of *Pseudomonas aeruginosa* in the U-bends of high-risk units in hospital', *International Journal of Hygiene and Environmental Health*, 220(5), pp. 880–885. doi: 10.1016/j.ijheh.2017.04.003.

Vatansever, C. *et al.* (2020) 'Co-existence of OXA-48 and NDM-1 in colistin resistant *Pseudomonas aeruginosa* ST235', *Emerging microbes & infections*, 9(1), pp. 152–154. doi: 10.1080/22221751.2020.1713025.

Vaziri, F. *et al.* (2011) 'The prevalence of aminoglycoside-modifying enzyme genes (aac (6')-I, aac (6')-II, ant (2")-I, aph (3')-VI) in *Pseudomonas aeruginosa* ', *Clinics* . scielo , pp. 1519–1522.

Ventola, C. L. (2015) 'The antibiotic resistance crisis: part 1: causes and threats.', *P & T : a peer-reviewed journal for formulary management*. MediMedia, USA, 40(4), pp. 277–83.

Verder, E. and Evans, J. (1961) 'A Proposed Antigenic Schema for the Identification of Strains of *Pseudomonas aeruginosa*', *The Journal of Infectious Diseases*. Oxford University Press, pp. 183–193. doi: 10.2307/30099339.

Verma, J. S. *et al.* (2014) 'Evaluation of gidB alterations responsible for streptomycin resistance in *Mycobacterium tuberculosis*', *Journal of Antimicrobial Chemotherapy*, 69(11), pp. 2935–2941. doi: 10.1093/jac/dku273.

Viedma, E. *et al.* (2009) 'Nosocomial spread of colistin-only-sensitive sequence type 235 *Pseudomonas aeruginosa* isolates producing the extended-spectrum β-lactamases GES-1 and GES-5 in Spain', *Antimicrobial Agents and Chemotherapy*, 53(11), pp. 4930–4933. doi: 10.1128/AAC.00900-09.

Vogne, C. *et al.* (2004) 'Role of the Multidrug Efflux System MexXY in the Emergence of Moderate Resistance to Aminoglycosides among *Pseudomonas aeruginosa* Isolates from Patients with Cystic Fibrosis', *Antimicrobial Agents and Chemotherapy*, 48(5), pp. 1676–1680. doi: 10.1128/AAC.48.5.1676-1680.2004.

Vu-Thien, H. *et al.* (2007) 'Multiple-locus variable-number tandem-repeat analysis for

longitudinal survey of sources of *Pseudomonas aeruginosa* infection in cystic fibrosis patients.', *Journal of clinical microbiology*. American Society for Microbiology, 45(10), pp. 3175–83. doi: 10.1128/JCM.00702-07.

Walker, G. T. *et al.* (2016) 'Analytical Performance of Multiplexed Screening Test for 10 Antibiotic Resistance Genes from Perianal Swab Samples.', *Clinical chemistry*. Clinical Chemistry, 62(2), pp. 353–9. doi: 10.1373/clinchem.2015.246371.

Walker, G. T. *et al.* (2019) 'Predicting Antibiotic Resistance in Gram-Negative Bacilli from Resistance Genes.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology Journals, 63(4), pp. e02462-18. doi: 10.1128/AAC.02462-18.

Walker, J. and Moore, G. (2015) '*Pseudomonas aeruginosa* in hospital water systems: Biofilms, guidelines, and practicalities', *Journal of Hospital Infection*. Elsevier Ltd, 89(4), pp. 324–327. doi: 10.1016/j.jhin.2014.11.019.

Wang, Y.-T., Lee, M.-F. and Peng, C.-F. (2014) 'Mutations in the quinolone resistance-determining regions associated with ciprofloxacin resistance in *Pseudomonas aeruginosa* isolates from Southern Taiwan', *Biomarkers and Genomic Medicine*. Elsevier, 6(2), pp. 79–83. doi: 10.1016/J.BGM.2014.03.003.

Ward, C. *et al.* (2015) 'Performance evaluation of the Verigene® (Nanosphere) and FilmArray® (BioFire®) molecular assays for identification of causative organisms in bacterial bloodstream infections.', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*, 34(3), pp. 487–96. doi: 10.1007/s10096-014-2252-2.

Wardell, S. J. T. *et al.* (2019) 'A Large-Scale Whole-Genome Comparison Shows that Experimental Evolution in Response to Antibiotics Predicts Changes in Naturally Evolved Clinical *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 63(12), pp. e01619-19. doi: 10.1128/AAC.01619-19.

Wattam, A. R. *et al.* (2017) 'Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center', *Nucleic Acids Research*, 45(D1), pp. D535–D542. doi: 10.1093/nar/gkw1017.

Watts, G. S. *et al.* (2017) '16S rRNA gene sequencing on a benchtop sequencer: accuracy for identification of clinically important bacteria', *Journal of applied microbiology*. 2017/11/07. John Wiley and Sons Inc., 123(6), pp. 1584–1596. doi: 10.1111/jam.13590.

Webber, M. A. and Piddock, L. J. V. (2003) 'The importance of efflux pumps in bacterial antibiotic resistance', *Journal of Antimicrobial Chemotherapy*. Oxford University Press, 51(1), pp. 9–11. doi: 10.1093/jac/dkg050.

Wei, Q. *et al.* (2011) 'Phenotypic and genome-wide analysis of an antibiotic-resistant small colony variant (SCV) of *Pseudomonas aeruginosa*', *PLoS ONE*, 6(12). doi: 10.1371/journal.pone.0029276.

Westfall, L. W. *et al.* (2006) 'mvaT mutation modifies the expression of the *Pseudomonas aeruginosa* multidrug efflux operon *mex*EF-oprN.', *FEMS microbiology letters*, 255(2), pp. 247–54. doi: 10.1111/j.1574-6968.2005.00075.x.

WHO (2016) 'WHO | Antimicrobial resistance: global report on surveillance 2014', *Antimicrobial resistance: global report on surveillance 2014*. World Health Organization.

Wiegand, I., Hilpert, K. and Hancock, R. E. W. (2008) 'Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances', *Nature Protocols*, 3(2), pp. 163–175. doi: 10.1038/nprot.2007.521.

Wiehlmann, L. *et al.* (2007) 'Population structure of *Pseudomonas aeruginosa*', *Proceedings of the National Academy of Sciences of the United States of America*, 104(19). doi: 10.1073/pnas.0609213104.

Wilhelm, S. *et al.* (2007) 'The autotransporter esterase EstA of *Pseudomonas aeruginosa* is required for rhamnolipid production, cell motility, and biofilm formation', *Journal of Bacteriology*, 189(18), pp. 6695–6703. doi: 10.1128/JB.00023-07.

Wilhelm, S., Tommassen, J. and Jaeger, K. E. (1999) 'A novel lipolytic enzyme located in the outer membrane of *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 181(22), pp. 6977–6986. doi: 10.1128/jb.181.22.6977-6986.1999.

Winand, R. *et al.* (2020) 'Targeting the 16s rRNA gene for bacterial identification in complex mixed samples: Comparative evaluation of second (illumina) and third (oxford nanopore

technologies) generation sequencing technologies', *International Journal of Molecular Sciences*, 21(1), pp. 1–22. doi: 10.3390/ijms21010298.

Winsor, G. L. *et al.* (2009) '*Pseudomonas* Genome Database: facilitating user-friendly, comprehensive comparisons of microbial genomes', *Nucleic Acids Research*. Narnia, 37(Database), pp. D483–D488. doi: 10.1093/nar/gkn861.

Winsor, G. L. *et al.* (2016) 'Enhanced annotations and features for comparing thousands of *Pseudomonas*genomes in the *Pseudomonas* genome database', *Nucleic Acids Research*, 44(D1), pp. D646–D653. doi: 10.1093/nar/gkv1227.

Wishart, D. S. *et al.* (2006) 'DrugBank: a comprehensive resource for in silico drug discovery and exploration', *Nucleic Acids Research*. Oxford University Press, 34(90001), pp. D668–D672. doi: 10.1093/nar/gkj067.

Wong, A., Rodrigue, N. and Kassen, R. (2012) 'Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*.', *PLoS genetics*. Public Library of Science, 8(9), p. e1002928. doi: 10.1371/journal.pgen.1002928.

Woo, P.C.; Lau, S.K.; Teng, J.L.; Tse, H.; Yuen, K. Y. (2008) 'Then and now: Use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Disease*, 14, pp. 908–934.

Woodford, N., Turton, J. F. and Livermore, D. M. (2011) 'Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance', *FEMS Microbiology Reviews*. Blackwell Publishing Ltd, 35(5), pp. 736–755. doi: 10.1111/j.1574-6976.2011.00268.x.

Wright, G. D. (1999) 'Aminoglycoside-modifying enzymes.', *Current opinion in microbiology*, 2(5), pp. 499–503.

Wright, G. D. and Poinar, H. (2012) 'Antibiotic resistance is ancient: implications for drug discovery', *Trends in Microbiology*, 20(4), pp. 157–159. doi: 10.1016/j.tim.2012.01.002.

Wright, L. L. *et al.* (2015) 'Genetic environment of metallo-β-lactamase genes in *Pseudomonas aeruginosa* isolates from the UK', *The Journal of antimicrobial chemotherapy*, 70(12), pp.

3250–3258. doi: 10.1093/jac/dkv263.

Wright, M. S. *et al.* (2015) 'SISPA-Seq for rapid whole genome surveys of bacterial isolates.', *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. NIH Public Access, 32, pp. 191–8. doi: 10.1016/j.meegid.2015.03.018.

Yamane, K. *et al.* (2007) 'New Plasmid-Mediated Fluoroquinolone Efflux Pump, QepA, Found in an Escherichia coli Clinical Isolate', *Antimicrobial Agents and Chemotherapy*, 51(9), pp. 3354–3360. doi: 10.1128/AAC.00339-07.

Yan, B. *et al.* (2018) 'SMRT-Cappable-seq reveals complex operon variants in bacteria', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-05997-6.

Yang, L. *et al.* (2019) 'Metagenomic identification of severe pneumonia pathogens in mechanically-ventilated patients: A feasibility and clinical validity study', *Respiratory Research*. Respiratory Research, 20(1), pp. 1–12. doi: 10.1186/s12931-019-1218-4.

Yang, P. *et al.* (2020) 'Association between the rate of fluoroquinolones-resistant gram-negative bacteria and antibiotic consumption from China based on 145 tertiary hospitals data in 2014.', *BMC infectious diseases*, 20(1), p. 269. doi: 10.1186/s12879-020-04981-0.

Yokoyama, K. *et al.* (2003) 'Acquisition of 16S rRNA methylase gene in *Pseudomonas aeruginosa*', *The Lancet*, 362(9399), pp. 1888–1893. doi: 10.1016/S0140-6736(03)14959-8.

Yoneyama, H., Sato, K. and Nakae, T. (1991) 'Aminoglycoside Resistance in *Pseudomonas aeruginosa* due to Outer Membrane Stabilization', *Chemotherapy*, 37(4), pp. 239–245. doi: 10.1159/000238861.

Yonezawa, M. *et al.* (1995) 'DNA gyrase gyrA mutations in quinolone-resistant clinical isolates of *Pseudomonas aeruginosa*.', *Antimicrobial agents and chemotherapy*. American Society for Microbiology (ASM), 39(9), pp. 1970–2. doi: 10.1128/AAC.39.9.1970.

Yordanov, D. and Strateva, T. (2009) '*Pseudomonas aeruginosa* – a phenomenon of bacterial resistance', *Journal of Medical Microbiology*, 58(9), pp. 1133–1148. doi: 10.1099/jmm.0.009142-0.

Yoshida, H. *et al.* (1990) 'Quinolone resistance-determining region in the DNA gyrase gyrA

gene of Escherichia coli.', *Antimicrobial agents and chemotherapy*, 34(6), pp. 1271–2.

Yuan, J., & Zhao, Z. (1963) *Study on the grouping of Pseudomonas aeruginosa by cellular antigens, People's Mil. Med, 2,*.

Zakham, F. *et al.* (2019) 'Whole-genome sequencing for rapid, reliable and routine investigation of *Mycobacterium tuberculosis* transmission in local communities', *New Microbes and New Infections*. Elsevier, 31, p. 100582. doi: 10.1016/J.NMNI.2019.100582.

Zamudio, R. *et al.* (2019) 'Phylogenetic analysis of resistance to ceftazidime/avibactam, ceftolozane/tazobactam and carbapenems in piperacillin/tazobactam-resistant *Pseudomonas aeruginosa* from cystic fibrosis patients', *International Journal of Antimicrobial Agents*. Elsevier B.V., 53(6), pp. 774–780. doi: 10.1016/j.ijantimicag.2019.02.022.

Zankari, E. *et al.* (2012) 'Identification of acquired antimicrobial resistance genes.', *The Journal of antimicrobial chemotherapy*. Oxford University Press, 67(11), pp. 2640–4. doi: 10.1093/jac/dks261.

Zhang, R., Wu, R., Collart, F., Joachimiak, A. (2004) 'RCSB PDB - 1XDZ: Crystal Structure of Gram_Positive Bacillus subtilis Glucose inhibited Division protein B (gidB), Structural genomics, MCSG'. doi: 10.2210/pdb1XDZ/pdb.

Zhang, J.-J. *et al.* (2009) 'Identification and characterization of catabolic para-nitrophenol 4-monooxygenase and para-benzoquinone reductase from *Pseudomonas sp*. strain WBC-3.', *Journal of bacteriology*. American Society for Microbiology Journals, 191(8), pp. 2703–10. doi: 10.1128/JB.01566-08.

Zhang, Y. *et al.* (2016) 'Validation of a High-Throughput Multiplex Genetic Detection System for Helicobacter pylori Identification, Quantification, Virulence, and Resistance Analysis.', *Frontiers in microbiology*. Frontiers Media SA, 7, p. 1401. doi: 10.3389/fmicb.2016.01401.

Zhang, Y. *et al.* (2020) 'Bronchoscope-related *Pseudomonas aeruginosa* pseudo-outbreak attributed to contaminated rinse water', *American Journal of Infection Control*. Elsevier Inc., 48(1), pp. 26–32. doi: 10.1016/j.ajic.2019.06.013.

Zhao, G. *et al.* (2013) 'Structure and function of *Escherichia coli* RimK, an ATP-grasp fold, l-glutamyl ligase enzyme', *Proteins: Structure, Function and Bioinformatics*, 81(10), pp. 1847–

1854. doi: 10.1002/prot.24311.

Zhao, Q. *et al.* (1998) ‘Contribution of outer membrane efflux protein OprM to antibiotic resistance in *Pseudomonas aeruginosa* independent of MexAB.’, *Antimicrobial agents and chemotherapy*, 42(7), pp. 1682–8.

Zheng, R. *et al.* (2018) ‘Genome-wide association study identifies two risk loci for tuberculosis in Han Chinese’, *Nature Communications*. Springer US, 9(1). doi: 10.1038/s41467-018-06539-w.

Ziha-Zarifi, I. *et al.* (1999) ‘In vivo emergence of multidrug-resistant mutants of *Pseudomonas aeruginosa* overexpressing the active efflux system MexA-MexB-OprM.’, *Antimicrobial agents and chemotherapy*, 43(2), pp. 287–91.