Inland Norway University of Applied Sciences

Faculty of Applied Ecology, Agricultural sciences and Biotechnology

Ambreen Kauser

Master's Thesis

Resistome Identification from Whole Genome Sequencing Data of Norwegian Isolates

Masters in Applied and Commercial Biotechnology

2020

Consent to lending by University College Library YES ☒ NO ☐

Consent to accessibility in digital archive Brage YES ☒ NO ☐

Acknowledgement

I would like to thank my supervisor Dr. Rafi Ahmad for providing me the opportunity to work on this project, and supporting me through out the project duration. His invaluable suggestions, expert guidance in every section of my thesis, punctuality, and regular feedback has ensured my steady progress, both in bioinformatics analysis and the thesis writing. I have learnt some of the best lessons of my life from him, and one of them is not to give up. I would also like to thank Dr. Ekaterina Avershina, my co-supervisor, for giving me valuable guidelines and insightful advices regarding my thesis, on every occasion I needed help. Her technical understanding of the project has helped me narrow down the goals of this project systematically. In addition, her kind words and constant motivation made me complete this thesis, with full conviction in putting efforts on my own, in the journey of creating something meaningful. Moreover, I would like to show my appreciation to Dr. Abdurehman Khizri for assisting me with writing with scientific precision. Here, I also extend my thanks to PhD. fellow Yusuf Khan for helping me, whenever I asked for an advice. Special thanks to Erasmas fellow; Clàudia López for her sincere contribution in the bioinformatics tasks. I am also thankful to my project fellow; Helene Bouras for her valuable work with some of wet-lab data availability and suggestions with proof-reading. Thanks to our biotechnology department coordinator Svein Birger for empathic contribution towards accomplishing my studies.

There is a special place in my heart for all those friends who have helped me come to Norway and stay here despite all the odds. I would like to show a great appreciation to my husband for his continuous love and support. Finally, I dedicate this piece of writing to my parents, who would have been the happiest now.

Hamar, 24th September, 2020

Ambreen Kauser

Abbreviations

AMR – Antimicrobial Resistance

AST – Antibiotic Susceptibility Testing

AGs – Aminoglycosides

AAC – Aminoglycoside Acetyltransferase

ART – Antimicrobial Resistance Test

ARDB – Antibiotic Resistance Genes Database

ARO – Antibiotic Resistance Ontology

ARG-ANNOT – Antibiotic Resistance Gene-ANNOTation

BLAST – Basic Local Alignment Tool

CHDLs – Carbapenem-Hydrolyzing class D β-lactamases

CBMAR – Comprehensive β-lactamase Molecular Annotation Resource

CARD – The Comprehensive Antibiotic Resistance Database

CTX-M – Cefotaxime Munich

DNA – Deoxi Ribo Nucleic Acid

DDBJ – DNA Data Bank of Japan

EAEC – Enteroaggregative Escherichia coli

EHEC – Enterohemorrhagic Escherichia coli

EIEC – Enteroinvasive Escherichia coli

EPEC – Enteropathogenic Escherichia coli

ESBL – Extended Spectrum Beta-lactamase

ETEC – Enterotoxigenic Escherichia coli

EUCAST – European Union Committee for Antimicrobial Susceptibility Testing

ExPEC – Extraintestinal Pathogenic Escherichia coli

ECOFFS – Epidemiologic Cutoff

ESBLs – Extended Spectrum Beta Lactamases

(EMBL-EBI) – European Molecular Biology Laboratory's European Bioinformatics Institute

FASTA – FAST-All

GC – Guanine Cytocine

HGT – Horizontal Gene Transfer

HMM – Hidden Markov Model

Inc.F – Incompatibility Fertility Factor

INSDC – The International Nucleotide Sequence Database Collaboration

KMA – K-mer Alignment

MS (MALDI-TOF MS) – matrix-assisted laser desorption/ionization time-of-flight

MS LC-MS – liquid chromatography

MDR – Multi Drug Resistant

MIC – Minimum Inhibitory Concentration

MS – Mass Spectroscopy

NCBI – National Center for Biotechnology Information

NDM – New Delhi Metallo-beta-lactamases

NGS – Next Generation Sequencing

NIPH – Norwegian Institute of Public Health

NORM – Norwegian Monitoring System for Antibiotic Resistance in Microbes

OXA – Oxacillinase

ONT– Oxford Nanopore technologies

pMLST– Plasmid Multilocus Sequence Typing

PLACNET– Polaris Low Acoustic Noise Exhaust Technology

PBP – Penicillin-binding Proteins

PCR – Polymerase Chain Reaction

PGM – Personal Genome Machine

qPCR– Quantitaive PCR

QUAST– Quality Assessment Tool

RGI – Resistance Gene Identifier

SMART – single molecule real time sequencing

SHV – Sulfhydryl Variable

SPADes – St. Petersburg Genome Assembler

SNP – Single Nucleotide Polymorphism

SOMs – Self- Organizing Maps

TEM – Temoniera

UTI – Urinary Tract Infection

VIM – Verona integron-encoded metallo-β-lactamase

VFDB – Virulence factor Database

WGS – Whole Genome Sequencing

WHO – World Health Organization

ZMWs – Zeromode Waveguides

Table of Contents

Abstract

Antimicrobial resistance (AMR) is considered a potential threat to global health. Norway have had a low prevalence of resistant bacteria. But in the recent years there has been an increase in resistant bacteria including, *Escherichia coli, Klebsiella pneumoniae* and *Acinetobacter baumannii*. Traditionally, clinical microbiology has used culture-based techniques to determine antimicrobial susceptibility and resistance profiles, but now whole–genome sequencing for antibiotic susceptibility (WGS-AST) has emerged as a potential alternative.

We aimed to investigate the prevalence of antimicrobial resistance genes and plasmids in WGS of 111 clinical Norwegian isolates of *E. coli, K. pneumoniae*, and *A. baumannii,* to identify correlations between phenotypic and genotypic resistance in the isolates, which are related to antibiotic resistance to β-lactam, aminoglycosides, fluoroquinolone, trimethoprim, tetracycline, and phenicol.

The most occurring drug class was β-lactam antibiotic with TEM (38%) in *E.coli,* SHV (67%) in *K. pneumoniae,* and OXA (100%) and TEM (45%) gene families in *A. baumannii. In silico* detection of plasmids with *Brooks et al* database showed plasmid p2_000837 as prominent plasmid 12% *E.coli* isolates. There were four plasmids (pIB_NDM_1, p2_W5-6, pCHL5009T-102k-mcr3, pVir_020022) in 2% *K. pneumoniae* isolates which were also shared with *E. coli*. Only one plasmid (pHZ23-1-1) was confirmed in 9% of *A. baumannii* isolates. PLSDB detected Plasmid A and plasmid 4 with the maximum percentage in *E.coli* (10%) and *K. pneumoniae* isolates (4%). In *E. coli* and *K. pneumoniae*, the presence of incompatibility groups was observed; IncFIB (64% and 27%), Col156 (74% and 27%), IncFII (43% and 15%), while IncHI-1B(pNDM-MAR) (12%) were present only in *K. pneumoniae* .

A total of 75 isolates had resistance to the tested β-lactam antibiotics, out of which 63 had the corresponding resistance genes (*ampC*, SHV, CTX-M, TEM, LEN, OXA). Only 11 *E.coli* and one *K. pneumoniae* isolates were found to have resistance genes and the plasmids on the same node to confirm plasmid mediated resistance.

This study demonstrates the utility of WGS in defining resistance elements and highlights the diversity of resistance within the selected isolates to further the diagnostics and therapeutics for the treatment of the relevant infections.

1. Introduction

1.1 Background of Research

Antimicrobial resistance has contributed immensely to the continuously growing concerns about the ineffective treatment against microbial infections (Shi et al., 2019). Overuse of antibiotics and insufficient therapy are the main causes of making AMR a global problem that leads to longer hospital stays, too costly treatments, and higher mortality rate (Elbadawi et al., 2019).

WGS is effective in tracking onward transmission of bacteria or resistance plasmid transfer between bacteria. WGS is also useful to identify trends in antibiotic resistance e.g. targeting the bacteria that are phenotypically sensitive but genotypically positive for a resistance (Köser et al., 2014). However, sensitivity of the populations and specificity of allelic variants, causing different susceptibility phenotypes, sometimes remains lower than the detection method being used, making it even more challenging (Lanza et al., 2018). This new approach requires novel microbial informatics (for development of reference databases of molecular and clinical metadata), new algorithms (for prediction of resistome and resistance phenotype from genotype), and new protocols (for global collection and sharing of high-throughput molecular epidemiology data) (McArthur and Wright, 2015).

1.2 Antimicrobial Resistance (AMR)

Bacteria are classified as antibiotic resistant when they are non-susceptible to at least one antibiotic class. It is estimated that resistant infections may kill one person every 3 seconds by the year 2050, raising the death toll worldwide to 10 million annually (Sabino et al., 2019).

In late 60s, due to presence of various antibiotics, most of the infectious bacteria remained sensitive to a great number of antibiotics being used to treat them. Since no new clinically useful structures have been discovered since 1961, the emergence of antibiotic resistance has escalated the ineffectiveness of the treatment. The reason we see the current clonal spread of resistant bacteria is because they contain the resistant gene carrying plasmids that often dump their genes into the bacterial chromosome. Species like *A.baumannii* which were never regarded as pathogens are now resistant to almost all the antibiotics. It has become the main

cause of pneumonia even in the patients who had antibiotic treatment previously (Amyes, 2000). So, antibiotic pressure increases the sequence variability in resistance genes. To measure this, metagenomics is deployed now, which allows both quantitative and qualitative analyses of resistomes.

1.3 Importance of Gram-negative bacteria in AMR Dissemination

Escherichia coli (E.coli) is a Gram-negative, rod shaped, facultative anaerobe from the family Enterobacteriaceae (Allocati et al., 2013). It resides in the large intestine of warm blooded animals including humans in the form of commensal microflora. The diseases related to *E.coli* are enteritis (caused by enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC), UTI (caused by extraintestinal *E. coli* ExPEC), septicaemia (caused by ExPEC) and neonatal meningitis (caused by *E. coli K1*) (Kaper et al., 2004).). B-lactamase (located mainly on plasmids) production in *E. coli* is the major mediator of resistance to broad spectrum of β-lactam antibiotics and multi-drug resistance (MDR) (Poirel et al., 2012).

Klebsiella pneumoniae (K. pneumoniae) are Gram-negative, encapsulated, non motile, rod shaped, anaerobic bacillus from Enterobacteriaceae, found readily in human mucosal surfaces including gastrointestinal tract and oropharynx, which can further proliferate into tissues causing serious diseases like pneumoniae, sepsis, UTI, bacteraemia, meningitis, and pyogenic liver abscesses (Bagley, 1985; Dao et al., 2014; Paczosa and Mecsas, 2016). *K. pneumonia* are among those bacteria which are responsible for the infections difficult to be treated with antimicrobial therapy (Pendleton et al., 2013) because they not only are intrinsically resistant to many antibiotics, but have also accumulated resistance to many additional drugs (de Man et al., 2018). Hundreds of AMR genes have been detected in *K. pneumoniae* (Holt et al., 2015). Two of the mechanisms for resistance used in *K. pneumoniae* are expression of ESBLs (making them resistant to cephalosporins and monobactams), and production of carbapenamases (making them resistant to all available β-lactam antibiotics including carbapenems) (Pitout et al., 2015).

Acinetobacter baumannii (A. baumannii) is a strictly aerobic, Gram-negative, non motile, nosocomial, non fermenting coccobacillus from the family Moraxellaceae (Peleg et al., 2008), which cause blood infections, pneumoniae, infections in soft tissues at surgical sites, Urinary Tract Infections (Zhao et al.), and Multi-Drug Resistance (MDR) (Harding et al., 2018; Sievert et al., 2013). Their genome is prone to mutation in stress, depicting the genetic flexibility to upregulate their natural resistance as well as acquire foreign determinants through mobile genetic elements (plasmids, integrons, and transposons) *A.baumanni* is an opportunistic pathogen and even pan-drug resistance phenotypes have been observed at unprecedented rate in recent times (Giammanco et al., 2017). Out of 33 identified genomic species of *Acinetobacter* genus (Kim et al., 2008; Nemec et al., 2009), *A. baumannii*, *Acinetobacter* genomic species 3 and 13TU have been considered as the most relevant species in clinical context (Nemec et al., 2009).

1.4 Resistance trends of gram negative bacteria in Norway

The World Health Organization (WHO) regards AMR a big threat to global health regardless of age and location. Apart from natural causes, inappropriate antibiotic prescription, and unhygienic conditions in hospitals are also important contributing factors to AMR threat (D'Costa et al., 2011). Traditionally, Scandinavia is regarded as a low incidence area for antibiotic resistance (Figure 1). In Norway, resistance to antibiotics is supervised by 3 systems; Norwegian Surveillance System for Communicable Diseases (MSIS), Norwegian Surveillance System for antimicrobial drug resistance (NORM/NORM-VET), and Norwegian Surveillance System for antimicrobial drug resistance - Veterinary Medicine (NORM/NORM-VET, 2016) to reduce antibiotic use, raise awareness about the spread of antibiotic resistance, development of new antibiotics, vaccines, and better diagnostic tools.

The percentage of *E. coli* with ESBL causing septicaemia has a ten-fold increase in the last 10 years and had an increase of 6.5 per cent of all septicaemia cases caused by *E. coli* in Norway in 2016 (NORM/NORM-VET, 2016). About 2.9% and 0.3% of healthy pregnant women were colonised by ESBL-producing or *ampC*-producing *E. coli* respectively (Rettedal et al., 2015), whereas an overall ESBL 15.8% in diarrhoea patients (273 faecal samples) with carrier rate of 10.3% in patients with no recent travel history and 56.3% in patients with a history of recent travel to Asia (Jørgensen et al., 2014a; Ulstad et al., 2016).

Since 2015, the third generation cephalosporins (ESBL) resistant *K. pneumoniae* isolates has increased from 2.9% to 5.3% in 2017, resulting in increased use of broad spectrum antibiotics. (Haug et al., 2011). National action plan on AMR in health care aims to reduce five specified groups of broad spectrum antibiotics by 30% by the end of 2020 (Ministries, 2015).

In Europe, Acinetobacter species have shown high resistance level (fluoroquinolones, aminoglycosides and carbapenems), especially in Baltic countries, Southern and South-eastern Europe (Prevention and Control, 2018).

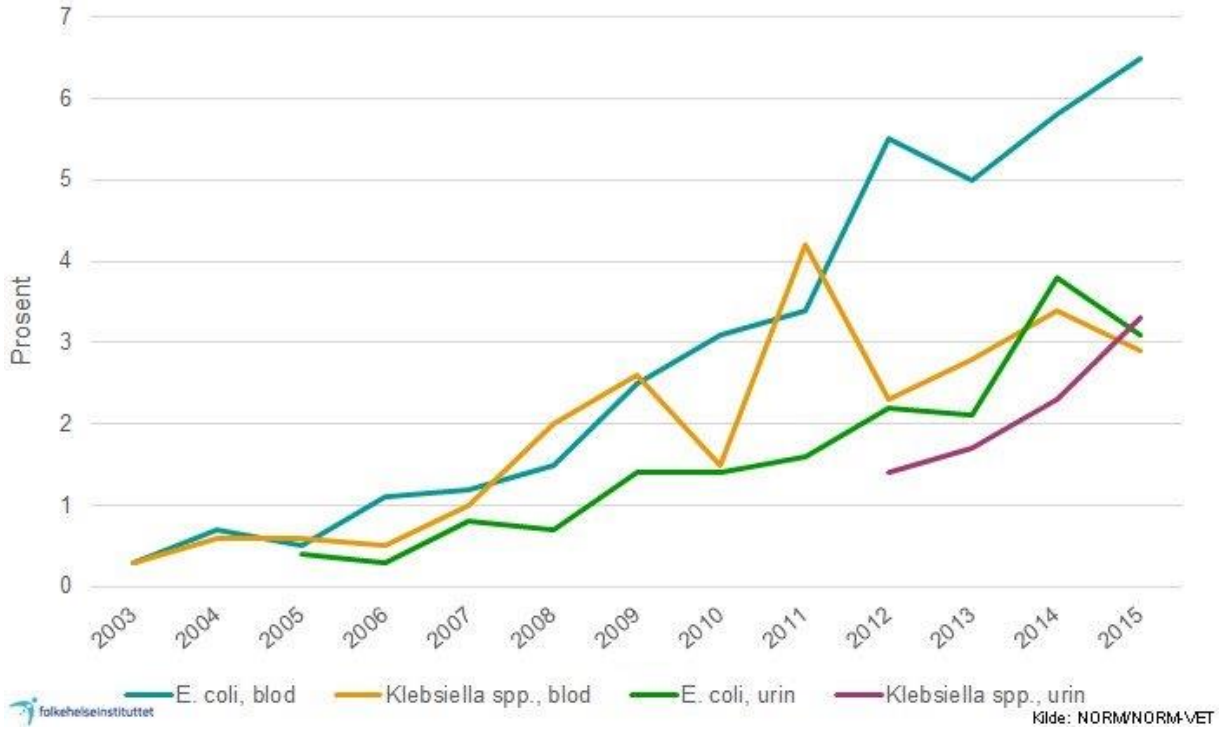


Figure 1: Proportion of ESBL-positive E. coli and K. pneumoniae in blood and urine in Norway. Retrieved from NORM/NORM-VET, 2016

1.5 Antimicrobial Agents

To treat diseases and prevent the risk of infection, the antimicrobial drugs either seize the grown of bacteria (bacteriostatic) or kill them (bactericidal) (Kohanski et al., 2010). Antibiotics like trimethoprim (disturbe the etrahydrofolate synthesis pathway), tigecycline, chloramphenicol, and tetracycline (protein synthesis inhibitors) are among bacteriostatic (*Figure 2*). Bactericidal antibacterials mainly include β-lactam antibiotics (prevent the formation of mature peptidoglycans), colistin (disrupt cell membrane), aminoglycosides (prevent protein synthesis), and quinolones (prevent

bacterial DNA replication) (Goffin and Ghuysen, 1998; Kohanski et al., 2010; Willey et al., 2011).

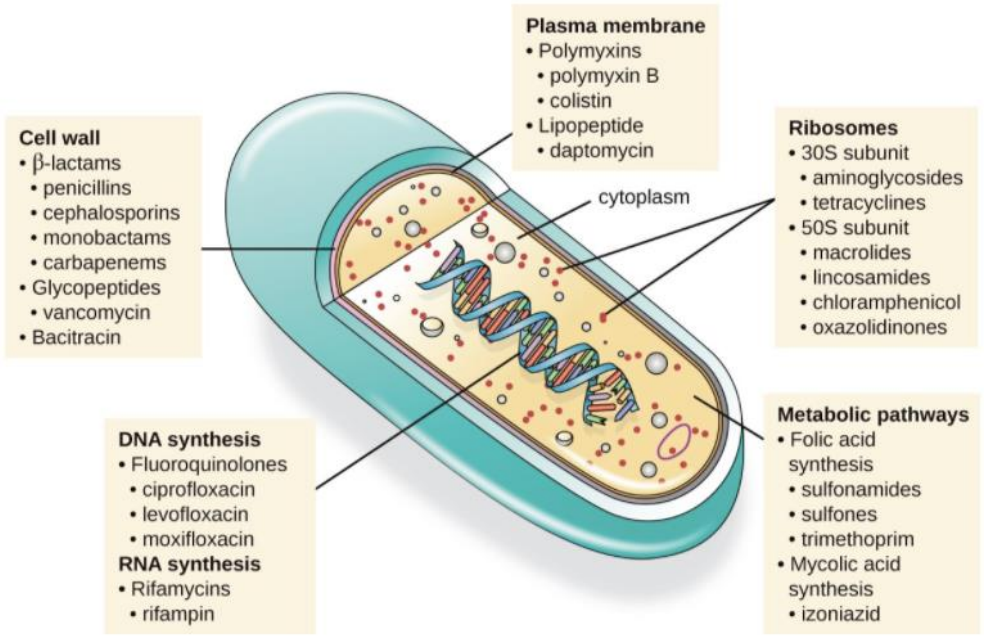


Figure 2: Target sites of Antibiotics in bacteria. Retrieved from https://courses.lumenlearning.com/microbiology/chapter/mechanisms-of-antibacterial-drugs

1.5.1 B-lactam antibiotics

β-lactam antibiotics interrupt bacterial cell-wall synthesis after they covalently bind to essential penicillin-binding proteins (PBPs), enzymes that are responsible for peptidoglycan cross-linking in both Gram-negative and Gram- positive bacteria (Bush and Bradford, 2016). The mechanism of β-lactam antibiotic action is explained by structural similarity between the β-lactam ring and the peptidoglycan building block acyl-D-alanyl-D-alanine (Tipper and Strominger, 1965). The covalent bond formed between β-lactam ring and an active site serine residue in the PBP results in the inactivation of the PBP (*Figure 3*).

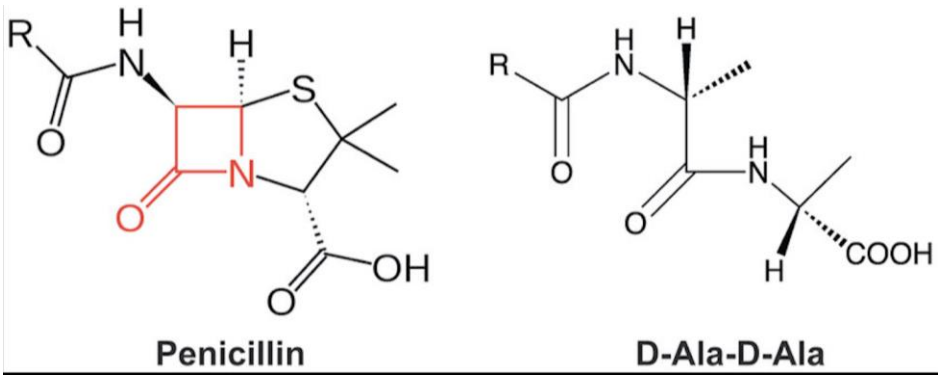


Figure 3: The four-member lactam ring in penicillin is highlighted in red. Retrieved from Zeng and Lin, 2013

The four major groups of β-lactam antibiotics are penicillins, cephalosporins, carbapenems, and monobactams (Figure 4), which are involved in affecting the bacterial cell wall synthesis. Four generations of cephalosporins have been introduced until now.

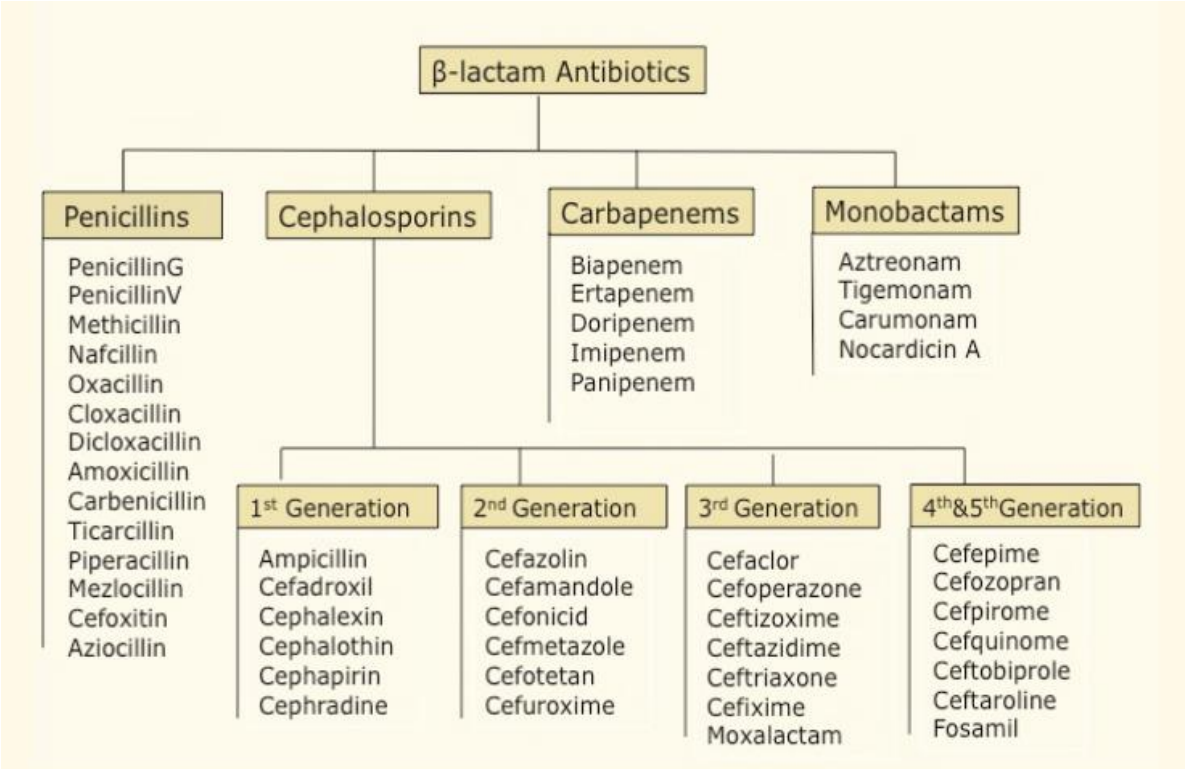


Figure 4: Classification of β-lactam abtibiotics. Retrieved from http://proteininformatics.org/mkumar/lactamasedb/lactamase.html

1.5.2 Non- B-lactam antibiotics

Aminoglycosides

Aminoglycosides are potent, broad-spectrum antibiotics that bind to A-site of 16S rRNA of 30S ribosomal subunit where codon-anticodon accuracy is analysed, resulting in disruption of protein binding in aerobic, Gram-negative bacteria (Davis et al., 2010; Hermann, 2007; Krause et al., 2016). AGs in return are inactivated by Aminoglycoside modifying enzymes, AME (Garneau-Tsodikova and Labby, 2016). For example, AAC(6') *AAC(3')-II* (N-acetyltransferases), and *aph(3')-II, aph(3')-III (*phosphotransferases) (Tolmasky, 2000; Vakulenko and Mobashery, 2003).

Tetracycline

Tetracyclines prefer to bind with 30S bacterial ribosomal subunit, and arrest the translation of highly conserved 16S ribosomal RNA (rRNA) by sterically hindering the docking of aminoacyl-transfer RNA (tRNA) to messenger RNA (mRNA)-ribosome complex during elongation (Chopra and Roberts, 2001). On the other hand, the bacteria have developed three strategies to become resistant to tetracyclines: limiting the accessibility for teracyclines to ribosomes, altering the binding site of ribosomes, production of inhibiors of tetracyclines (Speer et al., 1992). Tetracycline resistance genes could be spread by plasmids, transposons, and bacteriophages (Salyers et al., 1995). The most common tetracycline resistance mechanism in Gram-negative bacteria is by the gnes *tetA, tetB, tetC, tetD,* and *tetG*. However, *tetA* and *tetB* genes are most frequently present because they encode the most frequently used mechanism of tetracycline resistance in enterobacteriaceae; energy-dependent efflux. (Fluit et al., 2001). Mutation in *tet(A), tet(K), tet(M), and tet(X)* tetracycline resistance proteins causes tigecycline resistance. It is important to note that *tet(X3)* and *tet(X4)* inactivate all tetracyclines, including tigecycline and the newly FDA-approved eravacycline and omadacycline (He et al., 2019).

Fluoroquinolone

First generation quinolones; nalidixic acid discovered in 1962, were followed by second generation with the addition of a fluorine atom at position C-6 to the quinolone nucleus, making them fluoroquinolones (norofloxacin, ofloxacin, pefloxacin, ciprofloxacin etc). Fluoroquinolones are effective against several Gram-positive bacteria, Gram-negative bacteria, and intracellular bacteria.

The resistance to fluoroquinolones has emerged because of mutation in chromosomal quinolone targets (DNA gyrase and topoisomerase IV) and acquired resistance due to plasmid mediated quinolone resistance determinants (*qnr, qep, aac(6')-Ib-cr* and *oqxAB* (Veldman et al., 2011). The plasmid mediated quionolone resistance genes also have the potential to disseminate and enhance co-selection of other AMR genes (Ewers et al., 2012). First report of plasmid-mediated quinolone resistance was obtained from *K. pneumoniae* isolates in USA (Kim et al., 2009). *qnr* proteins alter quinolone target enzymes, efflux pump activation, or deficiencies in outer mebrane porins to show resistance to quinolones. *Qnr* proteins also raise the frequency level at which the quinolone resistance mutants can be selected by 100-fold (Martínez-Martínez et al., 1998). A gene variant of aminoglycoside acetyltransferase (*aac(6')-Ib-cr*) confers reduced susceptibility to ciprofloxacin and norfloxacin by N-acetylation of amino nitrogen on its piperazinyl substituent. (Robicsek et al., 2006). The *qnrA*, *qnrB* and *qnrS* genes can be found in transposons and integrons located on MDR plasmids of different incompatibility groups, which may carry multiple resistance determinants, including ESBLs and carbapenemases (Strahilevitz et al., 2009).

Phenicol

Chloramphenicol is a very specific and potent inhibitor of protein synthesis due to its affinity for peptidyltransferase of 50S ribosomal subunit of 70S ribosomes, thus preventing the peptidyl chain elongation in Gram-positive, Gram-negative, aerobic and anaerobic bacteria. Bacteria produce acetyletransferases (*catA, catB*) or phosphotransferases (*CmlA, floR*) for the enzymatic inactivation with acetylation as a mechanism of resistance to chloramphenicol. (Geisel et al., 1999; Schwarz et al., 2004). Other reasons for chloramphenicol resistance are target site mutation or modification, decreased membrane permeability, and reduction of effective intracellular drug concentration due to the presence of efflux pumps. Genes like *cmlA* and *floR* are the most commonly found genes for chloramphenicol resistance (Bissonnette et al., 1991).

1.6 Genetic mechanisms of antimicrobial resistance

1.6.1 Intrinsic resistance

Apart from environmental changes like radiation, change in light or pH, the bacteria have intrinsic resistance too (Wellington et al., 2013). Enzymes are used in intrinsic resistance to destroy or modify the drug (D'Costa et al., 2011). Bacteria can also produce inhibitors

(acetylases, phosphorylases, and adenylase) that reduce the drug's affinity for its the target sites due to steric hindrance (Munita and Arias, 2016).

1.6.2 Mutation

The binding sites of antimicrobials can be altered by one or more point mutations resulting in prevention of binding to the target by encoding abnormal target sites, which consequently increase the levels of resistance. Point mutations in β-lactamase genes have assisted in the identification of over 300 enzymes linked with a range of β-lactam antibiotic resistance phenotypes (Harbottle et al., 2006).

1.6.3 Horizontal gene transfer

Horizontal gene transfer is the ability of bacteria to exchange genes, which is responsible of spread and persistence of antibiotic resistance genes. There are three types of horizontal gene transfer; AMR gene linked with mobile genetic element, loss of gene loci in the host, and acquired AMR gene through genetic transfer (through transformation, transduction, conjugation) (Mullany et al., 2015; Pepper et al., 2018). Mobile genetic elements such as plasmids, transposons, integrons, and genomic islands harbour antibiotic resistance genes (Bennett, 2008). Many plasmids carrying resistance genes are transferred by the process of conjugation. Conjugation is a replicative process in which both donor and recipient cells have a copy of the plasmid after the process (Wilkins, 1995). Conjugative plasmids exhibit broad or narrow host range. In narrow range, the transfer is restricted generally to and between a small number of similar bacterial species. Broad range resistance plasmids are known to be associated with pathogens, for example, a resistance plasmid from *Pseudomonas aeruginosa* can be transferred to a wide variety of Gram-negative organisms. These mobile plasmids work as one of the means of acquiring resistance genes for pathogens in the environment (Bennett, 2008).

1.6.4 Production of β-lactamases

Resistance to β-lactam antibiotics is frequently mediated through the production of β-lactamase enzymes which break down β-lactam molecules. The β-lactamases bind to β-lactam antibiotics at a very fast deacylation rate resulting in the opening and thus inactivation of the β-lactam antibiotic molecule. It allows for the bacterial enzyme to return to normal functioning of forming peptidoglycan polymers (Søraas, 2014).

In the case of the Gram-positive cell, β-lactamases may either electrostatically attach with peptidoglycan layers or disseminate away into the extracellular environment (Figure 5). However, in Gram-negative bacteria, the β-lactamase is present mostly in the periplasm, but towards the permeability barrier, their ability to protect the bacteria is unpredictable (Livermore, 1995).

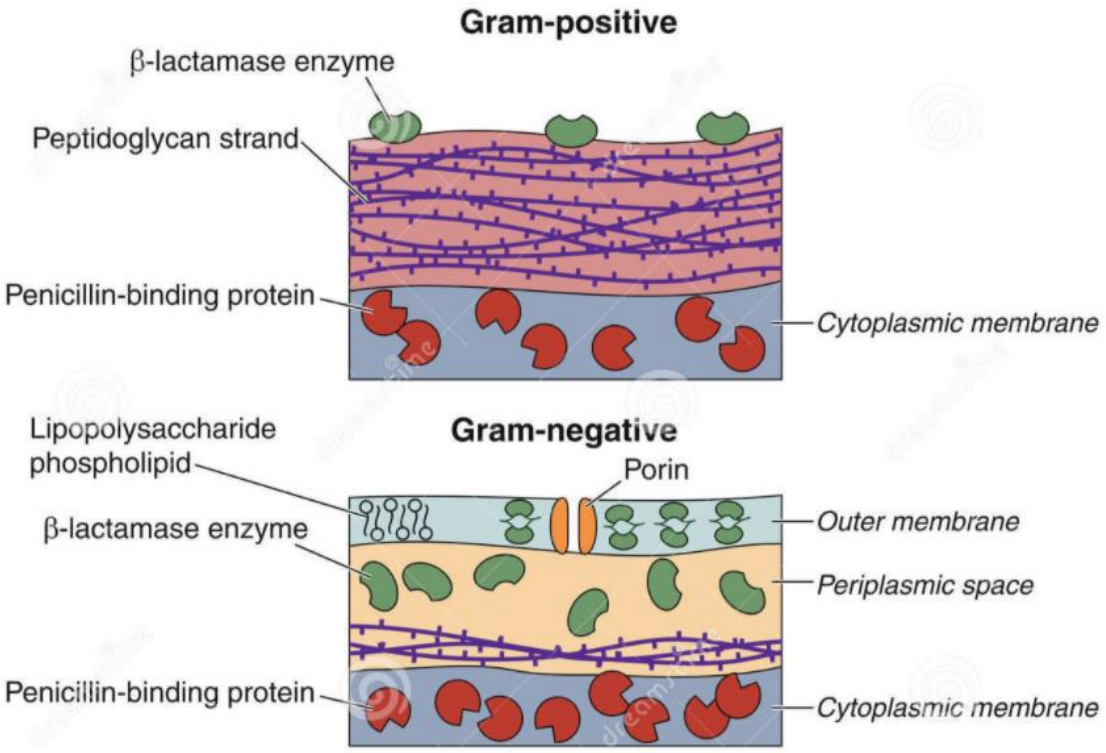


Figure 5: β-lactamases' position on both gram-negative (extracellular milieu) and gram-positive bacteria (between the outer and cytoplasmic membranes). Retrieved from https://www.dreamstime.com/stock-images-gram-positive-negative-bacteria-image13281714

1.7 Importance of ESBLs

ESBLs are plasmid-mediated β-lactamases that hydrolyse penicillins, cephalosporins (1st, 2nd, 3rd Generation), and aztreonam (Jacoby and Munoz-Price, 2005), but are susceptible to cefoxitin, carbapenems, and the β-lactam antibiotic inhibitors (clavulanic acid, tazobactams) (Bradford, 2001). The most frequently encountered ESBLs belong to the TEM, SHV (2be), and CTX-M classes (Ali et al., 2018).

When the amino acid substitutions around the active site of TEM-1/2 and SHV-1 β-lactamases started changing the configuration of the active site, the hydrolysis of oxymino-cephalosporin substrates (ceftazidime, cepodoxime, ceftriaxone, cefotaxime, monobactam, aztreonam occurred, leading to the discovery of a total of >130 TEM-type and >50 SHV-type β-lactamases (Kliebe et al., 1985).

CTX-M-type β-lactamases are capable of hydrolyzing broad-spectrum oximino- β-lactam antibiotics (cefotaxime, ceftriaxone, aztreonam), and are inhibited by clavulanate and tazobactam (Tzouvelekis et al., 2000).

Carbapenem-hydrolysing β-lactamases (carbapenemases) related to molecular class D (OXA enzymes) have appeared globally as the main mechanism causing this resistance. A phylogenetic subgroup OXA-51 has recently been found to be intrinsically present in *A. baumannii.* Since the carbapenem resistance can not be inferred from the presence of intrinsic OXA-51, alleles like OXA-23, OXA-24, and OXA-58 have been found in the *A. baumannii* isolates with acquired resistance to carbapenems (Woodford et al., 2006).

The metallo- β-lactamases like IMP and VIM have gained clinical importance, and have resistance against most β-lactamases including carbapenems (Nordmann and Poirel, 2002). In a Greek hospital, an isolate of *E.coli* with imipenem resistance was also found to have VIM β-lactamase. (Miriagou et al., 2003). The cause of resistance to cephalosporins and carbapenems in *A. baumannii* is due to the presence of metallo B-lactamases like IMP and VIM (Thomson and Bonomo, 2005).

OKP-A β-lactamases are chromosomal class A β-lactamases that confer resistance to penicillins and early cephalosporins.

1.7.1 Classification of Extended Spectrum B-lactamases

B-lactamases can be classified according to two general schemes: the Ambler molecular classification and the Bush-Jacoby-Medeiros functional classification system (Table 1) (Paterson and Bonomo, 2005). The Ambler scheme uses protein homology criterion to divide β-lactam antibiotics into four major classes: A, B, C and D. Class A, C and D utilize a transient serine acylation/deacylation at the active site. They also show structural similarities with the target of β-lactam antibiotics; the DD-peptidases, and therefore presumably come from the

same ancestral enzyme, while class B are dependent on a metal ion at the active site (Majiduddin et al., 2002).

In contrast, the Bush-Jacoby-Medeiros classification groups of β-lactamases into four main groups and multiple subgroups according to functional similarities. Both group details are mentioned in the Table 1. Most ESBLs are grouped in 2be, members of which stop the functioning of penicillins, cephalosporins, and monobactams, and are inhibited by clavulanic acid (Bush et al., 1995) .

Table 1: Main features of two general classification schemes. Retrieved from Dhillon and Clark, 2012

Bush-Jacoby-Medeiros group	Ambler molecular classification	Preferred substrate	Representative enzyme	Resistance or susceptibility to beta-lactamase inhibitor
1	C	Cephalosporins	AmpC	Resistant
2b	A	Penicillins, cephalosporins	TEM, SHV	Susceptible
2be	A	Penicillins, extended-spectrum cephalosporins, monobactams	TEM, SHV	Susceptible
2d	D	Penicillins, cloxacillin	OXA	Resistant
2e	A	Cephalosporins	Inducible cephalosporinases from *Proteus vulgaris*	Susceptible
2f	A	Penicillins, cephalosporins, carbapenems	NMC-A from *Enterobacter cloacae*	Resistant
3	B	Most beta-lactams including carbapenems	L1 from *Stenotrophomonas maltophilia*	Resistant

1.8 Plasmid Prevalence in Enterobacteriaceae and *A. baumannii*

In order to study the epidemiological relationships, classification of the plasmids needs to be understood (Datta, 1977). Resistance plasmids encode resistance to antimicrobials, for example, IncF and IncI1 plsmids are known to carry resistance genes in *E. coli, S. enterica, K.pneumoniae* and other *Enterobacteriaceae (Kaper et al., 2004)*. Moreover, the ColE plasmids encoding colicins, which have killing activity against other bacteria are also important plasmids (Hiraga et al., 1994).

Currently there are 27 Inc groups identified in Enterobacteriaceae by Plasmid Section of the National Collection Type Culture, Colindale, London (Carattoli, 2011; Couturier et al., 1988). IncFII, IncFIA, -B and -C are included in IncF group.

Interestingly, IncFIC is similar to IncFII, but still compatible. Sometimes, two plasmids in an Inc. family have distinguishable sequences, but still they appear to be incompatible (IncXI R485 and IncX2 R6K from IncX family). Within IncI-complex family of replicons, IncB/O replicons are incompatible with IncZ replicons (Jones et al., 1993), but both of them can stay together with IncI1, IncIγ and IncK replicons (which are incompatible with each other) (Praszkier and Pittard, 2005; Praszkier et al., 1991). Details of known plasmid incompatibility Inc. groups are given in Table 2.

In *A. baumannii,* the blaOXA-58 and blaOXA-23 genes encoding the OXA-58 and OXA-23 carbapenem hydrolysing oxacillinases (CHDLs) respectively, have been found in association with plasmids, gathered from various parts of world (Nordmann and Poirel, 2002). *A. baumannii* plasmids belong to a limited number of plasmid lineages and their structure is very stable, as compared to so-called mosaic plasmids. Mosaic plasmids are composed of genetic elements from distinct sources and they are highly dynamic in acquisition and loss of genes (Pesesky et al., 2019).

Table 2: List of known Incompatibility Inc. Plasmids. Retrieved from Johnson and Nolan, 2009

Inc group	Description	Example in NCBI database	Size (bp)[a]	Source or GenBank accession no.
com9	Plasmids capable of carrying transfer and MDR functions	pIP71A	85,825	Sanger Institute
IncA/C	Broad-host-range plasmids capable of carrying transfer and MDR functions	pSN254	176,473	CP000604
IncB/O	Plasmids capable of carrying transfer, MDR, and virulence functions	pTP113	96,471	Sanger Institute
IncD	Phage-associated plasmids	None		
IncFIA	Plasmids capable of carrying transfer, MDR, and virulence functions	Plasmid F	99,159	AP001918
IncFIB	Plasmids capable of carrying transfer, MDR, and virulence functions	pO157	92,077	AF074613
IncFIC	Plasmids capable of carrying transfer, MDR, and virulence functions	Plasmid F	99,159	AP001918
IncFIIA	Plasmids capable of carrying transfer, MDR, and virulence functions	pR100	94,281	AP000342
IncFIV	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU316	~77,000	P36REPA
IncFV	Plasmids capable of carrying transfer, MDR, and virulence functions	pED208	~90,000	AF411480
IncFVI	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU212	ND	X55895
IncFVII	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU221	ND	P2SINC
IncHI1	Plasmids capable of carrying transfer and MDR functions	pR27	180,641	AF250878
IncHI2	Plasmids capable of carrying transfer and MDR functions	pR478	274,762	BX664015
IncHI3	Plasmids capable of carrying transfer and MDR functions	Mip233	ND	AF192489
IncHII	Plasmids capable of carrying transfer and MDR functions	pHH1508a	ND	ECOTEHAB
IncI1	Plasmids capable of carrying transfer and MDR functions	pR64	120,826	AP005147
IncI2	Plasmids capable of carrying transfer and MDR functions	R721	75,582	AP002527
IncJ	Conjugative, self-transmitting, integrating elements	ICE R391	88,532	AY090559
IncK	Plasmids capable of carrying transfer and MDR functions	pR387	87,645	Sanger Institute
IncL/M	Plasmids capable of carrying transfer and MDR functions	pCTX-M3	89,468	AF550415
IncN	Broad-host-range plasmids capable of carrying transfer and MDR functions	R46	50,969	AY046276
IncP-α	Broad-host-range plasmids capable of carrying transfer and MDR functions	RP4	60,099	L27758
IncP-β	Broad-host-range plasmids capable of carrying transfer and MDR functions	pB4	79,370	AJ431260
IncP-γ	Broad-host-range plasmids capable of carrying transfer and MDR functions	pQKH54	69,966	AM157767
IncP-δ	Broad-host-range plasmids capable of carrying transfer and MDR functions	pEST4011	76,958	AY540995
IncP6 (IncG)	Broad-host-range plasmids capable of carrying transfer and MDR functions	Rms149	57,121	AJ877225
IncP7	Broad-host-range plasmids capable of carrying transfer and MDR functions	pCAR1	199,035	AB088420
IncP9	Broad-host-range plasmids capable of carrying transfer and MDR functions	pWWO	ND	WWODIRRPTA
IncQ1	Broad-host-range plasmids capable of carrying MDR functions	RSF1010	8,684	RSFRMRA
IncQ2	Broad-host-range plasmids capable of carrying MDR functions	pTC-F14	14,155	AF325537
IncR	Plasmids capable of carrying transfer and MDR functions	pK245	98,264	DQ449578
IncT	Large plasmids carrying transfer and DNA metabolism functions	pRts1	217,182	AP004237
IncU	Plasmids capable of carrying transfer and MDR functions	pFBAOT6	84,749	CR376602
IncV	Plasmids capable of carrying transfer and MDR functions	None		
IncW	Broad-host-range plasmids capable of carrying transfer and MDR functions	pR7K	39,792	AM901564
IncX1	Plasmids capable of carrying transfer and MDR functions	pOLA52	51,602	EU370913
IncX2	Plasmids capable of carrying transfer and MDR functions	pR6K	39,872	Sanger Institute
IncY	Phage-like plasmids	P1	94,481	AF234173

[a] ND, not determined.

1.9 Diagnostic Measures against Antimicrobial Resistance

1.9.1 Antimicrobial Susceptibility Test (AST)

Antibiotic susceptibility testing (AST) finds a dynamic antibiotic dosage and develops a form of diagnostics for protection against bacterial infections. Minimum inhibitory concentrations

of various antimicrobial susceptibility testing (AST) are classified by various international agencies. The susceptibility of microorganism towards the antibiotic is interpreted as susceptible (S), intermediate (I) and resistant (R). Most countries follow the epidemiological MIC cut-offs (ECOFFS) determined by the European Committee on Antimicrobial Susceptibility Testing (EUCAST, 2020) and/or the Clinical and Laboratory Standards Institute (CLSI, USA) (Khan et al., 2019). Presently, AST is performed using either classical manual methods or growth-dependent automated systems based on BMD testing. Other AST methods (manual and automated), commonly performed by clinical laboratories, are the conventional disk diffusion, agar dilution, antimicrobial gradient (e.g. the E-test, AB Biodisk) and automated instrumentation (Schofield, 2012).

1.9.2 PCR

Apart from culturing as the standard for diagnosing infection, sequence based approaches and quantitative PCR offer selective and sensitive way to identify a large number of Antibiotic Resistance Genes (ARG). However, qPCR requires a prior selection of targets which can overlook many important ARGs in a particular environment (Lindgreen et al., 2016; Walsh and Duffy, 2013), but is helpful to capture the non-culturable section of non clinical antibiotic resistome. However, targeting only selected genes is not enough to characterize the intrinsic resistance efflux resistance mechanisms, which are controlled by many genes (Walsh and Duffy, 2013).

1.9.3 Mass Spectrometry

Mass spectrometry (MS) has been used for microbial identification in place of conventional identification techniques (laboratory diagnostics) (Van Veen et al., 2010). Its role in AST and Antibiotic Resistance Testing (ART) has emerged recently (Hrabák et al., 2012). In contrast to conventional AST, where the response or no response of living organisms is noted upon exposure to antibiotics, ART uses the presence of biomarkers proteins, carbohydrates, lipids, and enzymatic activity to detect specific resistance mechanism. Mass spectrometry selects either resistance or susceptibility of resistance of clinical isolates. So, if we know the resistance mechanism for carbapenemase resistance (e.g. modification in drug influx or presence of carbapenemases), only the second mechanism can be selected for the inhibitors (clavulanic acid) related to it (Nordmann et al., 2012). MS techniques like matrix-assisted laser desorption/ionization time-of-flight MS (MALDI-TOF MS) and liquid chromatography-MS (LC-MS; in various forms) are currently in use (Welker and Van Belkum, 2019).

1.9.4 Whole Genome Sequencing (WGS)

WGS technology has made it possible to determine and evaluate the whole DNA sequence of a bacterium at low costs in just a few days (Punina et al., 2015). WGS not only allows *in silico* prediction of antimicrobial resistance (including resistance to compounds not routinely tested phenotypically), but also the early detection of outbreaks or their epidemiological investigation (Köser et al., 2014). Since the *in silico* prediction of resistance needs to be validated by phenotypic antimicrobial testing (Zankari et al., 2013),the combined use of phenotypic assays and techniques allowing the identification of genetic determinants of resistance can be helpful in epidemiological surveillance. Bacteria showing similar resistance patterns but different mechanisms can also be identified with WGS (Gordon et al., 2014). The unprecedented level of details of assays obtained from WGS for microbial typing and AMR surveillance can describe current trends and differentiate between emerging tendencies (Ellington et al., 2017). Moreover, Multi drug Resistance (MDR) patterns is defined with much greater precision with DNA sequence based surveillance as compared to phenotypic tests. The reason is that bioinformatics analysis goes beyond the concept of MDR as resistance to compounds from three or more drug classes, as it considers the co-carriage of particular genes behind different MDR patterns, allelic trends, their potential for horizontal transfer, and their distribution by source (Magiorakos et al., 2012).

1.10 Bioinformatic tools for WGS- based Characterization of Antimicrobial Resistance

1.10.1 Sequencing Platforms

First generation technology has remained the leading technology for decades for DNA sequencing (Sanger et al., 1977), using traditional shotgun technique that produced long low through put read sequences (500-1000 bp) at a relatively higher cost.

Second generation sequencing technology was fast and high throughput, generating short reads of 25-100 bp length (HiSeq from Illumina (https://www.illumina.com/.), 454 Life sciences from Roche (https://www.454.com/.), Solexa, and SOLiD (https://www.appliedbiosystems.). They were able to run over a few million reads in a single run with high coverage depth, cutting short the cost for DNA sequencing significantly (Butler and Grimme, 2010).

Sequencing by synthesis approach used by Illumina has made it dominate the industry in the recent years (Bentley et al., 2008), using fluorescently labeled reversible terminator nucleotides, on clonally amplified DNA templates (immobilized on acrylamide coating on the surface of glass flow-cell). In 2011, MiSeq was released which is suitable for smaller laboratories and the clinical diagnostic market (Quail et al., 2012).

One of the third generation sequencing platform; PacBio (Biosciences, 2014) has enabled single molecule real time sequencing (SMART). Here, DNA polymerase molecules, which are bound to DNA template are attached to the bottom of 50nm wide wells (zeromode waveguides (ZMWs). Second strand is synthesized by each polymerase in the presence of γ-phosphate fluorescently labeled nucleotides. When the fluorescence appears with a distinctive pulse, it means that fluorophores attached to the nucleotides are excited by the energy penetrating the waveguide at the time of addition of a new base. It produces a relatively small number of longer reads (> 10 kbp) as compared to a large number of short reads <200 bp like Illumina. However, higher cost per base, and higher sequencing error rate (15-20%) have limited their use in genome assembly (Schadt et al., 2010).

Oxford Nanopore Technologies (ONT) MinION8 uses a new technique where native DNA molecules are pulled through nanoscale pores that accept only one DNA molecule at a time. As the DNA molecule moves through the pore, followed by sensors detecting changes in the ionic current produced by each passing nucleotide. This information can be visualized in a 'squiggle plot' and provides the signal used for base calling. Resulting long read lengths significantly improve *de novo* genome assemblies and the detection of structural variations in large genomes (Deamer et al., 2016). ONT is the first technology that can deliver sequencing data from clinical samples in a timeframe that allows early de-escalation and refinement of antimicrobial treatment (Schmidt et al., 2016).

Another post 2011 NGS technique; Ion Torrent PGM (personal Genome Machine) uses Semiconductor technology, which detects the released protons as nucleotides are incorporated during synthesis. On the surface of Ion Sphere particles (3-micron diameter beads), DNA fragments with specific adaptor sequences are linked to and then emulsion PCR amplified. There are proton sensing wells fabricated on a silicon wafer for the templated beads to be loaded on, here the sequencing starts from a specific location in the adaptor sequence. The

addition of all four basis is done sequentially, base of a particular type has a particular signal after the proton gets released proportional to the number of bases incorporated (Rothberg et al., 2011).

1.10.2 AMR Detection Tools

AMR gene databases with comprehensive and accurate gene record are needed to assess AMR prevalence. Different approaches used are BLAST (Peirano et al., 2014), Hidden Markov Model (HMM) (Gibson et al., 2015), nucleotide or protein based differentiation, web interface, or operation on local servers. The researchers have to choose between the collections of resistance genes for use in HMMs (Gibson et al., 2015), or collections of nucleotides or protein sequences of individual resistance genes or resistance related mobile elements (McArthur et al., 2013; Zankari et al., 2012). Some databases focus on allelic variation of house keeping genes and their contribution to resistance, and some focus on acquired resistance mechanisms (Feldgarden et al., 2019b).

Another important factor to be considered is the bias of ARG databases towards experimentally validated genes. Thus selection of stringent cutoffs (≥ 90% per read/contig) though increases the probability of targeted functional genes, but it also omits environmentally relevant ARGs that can be more diverse. However, lowering the cutoffs to 60-80% will increase the false positives (Bengtsson-Palme et al., 2017). For the ARG characterization in metagenomic datasets, sequencing data (e.g. Illumina) can be used either without being assembled or be *de novo/reference based* assembled (Breitwieser et al., 2019; Knight et al., 2018). Although *de novo* assembly results in data loss, and needs higher genome coverage of diverse microbes with uneven taxonomic composition, it is helpful in more accurate detection of protein coding genes and exploration of upstream and downstream, unlike read-based methods (Henson et al., 2012). Moreover, with the advent of long read sequencing technologies (pacBio and Oxford Nanopore), the challenges offered by short read assemblies can be compensated by covering whole genes and even entire operons and mobile elements (Schatz et al., 2010).

Use of paired end reads in NGS technologies has made it possible to read the DNA fragment from both sides. An assembler uses both the expected distance and the orientation of the reads when reconstructing a genome. Although paired end reads are helpful for resolving repeat

regions that are longer than the length of the reads, where the one not in the repeat region helps the other to anchor correctly, but if the sequence data does not contain paired ends that span a particular repeat, then it might be impossible to assemble the data unambiguously (Treangen and Salzberg, 2011).

ResFinder is a highly cited tool among the established tools for ARG characterization in WGS data. It accepts both short reads and assembled genomes/contigs, using BLAST and/or KMA (k-mer alignment) based approaches to detect the acquired resistance, except for the resistance due to chromosomal mutations. To avoid ambiguous results, it is recommended to use 90% identity and 60% query coverage (Zankari et al., 2012).

On the other hand, Comprehensive Antibiotic Resistance database (Marini et al.) is among the tools for ARG surveillance in metagenomics sequencing data. In CARD database, molecular sequences, Resistance gene identifier (RGI), and BLAST is used for the prediction of antimicrobial resistance genes (ARGs) in metagenomics datasets, based on homology and Single nucleotide Polymorphism (SNP) models. CARD is a rigorously curated collection of characterized, peer reviewed resistance determinants, and linked antibiotics organized by the Antibiotic Resistance Ontology (ARO) and AMR gene detection models. CARD contains more than 2000 ontologically structured protein homologues, and includes intrinsic, mutation driven, and acquired resistance mechanisms. (Jia et al., 2016).
MEGARes, which is a hand curated ARG database, detects antimicrobial resistance determinants in large metagenomics datasets. Each protein and nucleotide has been validated manually with each annotation formatted in such a way that the database can be integrated into custom scripting easily. However, MEGARes focuses on previously published sequences, rather than newly discovered variants (Lakin et al., 2017).

The Bacterial Antimicrobial Resistance Reference Gene Database (AMRFinder) is derived from β-lactamase alleles, quinolone resistance protein alleles, ResFinder, and CARD. Since the AMR gene nomenclature is defined by protein identity and similarity, this phenomenon is used as a base of AMRFinder database. Within this framework using protein based HMMs, can be helpful to discover potentially novel AMR genes. This database contains over 560 AMR HMMs, and over 4579 curated AMR protein sequences to identify AMR genes from sequence data. Both AMR HMMs and AMR protein sequences are put together in a

hierarchical framework of gene families, symbols, and names in collaboration with groups like CARD (Feldgarden et al., 2019b).

With the exception of CARD and ResFinder , most of the ARG databases lack effective and sustainable curation strategies making them outdated (Lal Gupta et al., 2020).

Since the best hit approach of Next Generation Sequencing produces a high rate of false negatives, the Machine Learning (ML) approach considers the similarity distribution of sequences in the ARG database, instead of only the best hit. Due to disregard of cutoffs in ML gene prediction, there is a great reduction in false negatives, as well as maintaining high positive rate associated with traditional best hit approach by expanding the available ARGs individually available in the databases like CARD, ARDB, UNIPROT etc (Arango-Argoty et al., 2018).

1.10.3 *In Silico* Plasmid Detection Tools

Plasmids primarily contain the genes related with environmental fitness of the host, catabolism, and resistance (Carattoli, 2013; Zhang et al., 2011), leading them to contribute to horizontal gene transfer between different species (Thomas and Nielsen, 2005). However, assemblies generated using Illumina sequencing do not produce complete genomes, which affects the efforts to characterize the plasmid content of samples.
This happens because the plasmids tend to contain repeat sequences with sizes greater than sequences generated by Illumina technology (Arredondo-Alonso et al., 2017).

The need for *in silico* plasmid detection also emerged from the difficulty of plasmid DNA purification if they are longer than 50kbp (Smalla et al., 2015). Moreover, since the metagenomes usually are biased towards chromosomal content as compared to plasmids, many plasmid sequences remain unidentified in sequenced metagenomes, making it a complex process (Dib et al., 2015).

Most of the *in silico* plasmid detection methods are aimed at recovering circular contigs from de Bruijn assembly graphs (Jørgensen et al., 2014b; Rozov et al., 2017). However, even if plasmids are assembled directly from WGS by short read sequencing platform, still they have repeat region sequences that prohibit complete assembly of the plasmids, and they rely on

laborious and computationally intensive methods (De Toro et al., 2015; Kristiansson et al., 2011).

De Bruijn graph based plasmid prediction is done by Recycler (Rozov et al., 2017) and PlasmidSPAdes (Antipov et al., 2016). PlasmidSPAdes first calculates the median coverage from the SPAdes assembly graph to estimate a chromosome coverage, then it builds a second assembly graph which considers only those contigs which have a read contig coverage differing from chromosome coverage (Antipov et al., 2016; Bankevich et al., 2012). These second assemblies are regarded as putative plasmids after repeat resolution by ExSPAnder (Prjibelski et al., 2014). However, the read contig coverage dependency of PlasmidSPADes makes large and low copy plasmids nearly indistinguishable from the chromosome. This dependency is not applied by the databases like PlasmidFinder, cBar, and MOB-suite for resistance analysis (Page et al., 2018b).

PlasFlow is a neural network model, that is trained to separate chromosomal and plasmid sequences (short-length) (Vollmers et al., 2017) from different phyla by finding hidden structures in highly complicated biological data (Angermueller et al., 2016). A total of 9565 FASTA sequences were used to compile it, including 1961 chromosomes and plasmids 7604 of organisms from the kingdom Bacteria (Krawczyk et al., 2018). Unlike PlasmidSpades and Recycler, which output full length plasmid sequence predictions, based on their circularity or differential sequencing coverage, PlasFlow can predict the plasmid origin of the contigs even if it does not cover the whole plasmid sequence. That clarifies PlasFlow usage in the type of analysis that does not require full plasmid sequences with precise taxonomic information (Arredondo-Alonso et al., 2017; Krawczyk et al., 2018).

The plasmid detection programs that try to determine the plasmid origin of contigs include PlasmidFinder and cBar. cBar predicts plasmid-derived sequences (using self organizing maps: SOMs), on the basis of genomic signatures (k-mer composition) in full length sequences (Zhou and Xu, 2010), while PlasmidFinder tool detects the plasmid replicons and assigns the query plasmids to the respective Inc. group in Enteobacteriaceae (Orlek et al., 2017). Since two plasmids sharing the same replication mechanism can not co-exist within the same cell, the plasmids are put into different incompatibility groups (Carattoli et al., 2014). However, the size of PlasmidFinder database and its limitation only to Enterobacteriaceae replicons limits its usage for metagenomics studies.

In Plasmid Constellation Network (PLACNET), BLAST is used to compare sequences against reference databases to reconstruct plasmids through network analysis. Plasmid prediction by PLACNET depends on the expertise of the researchers because it needs scaffold linking and coverage information, replication initiator proteins (Rip) and relaxase proteins (Rel), and similarity of the sequences with non redundant plasmid sequences from NCBI. In addition, it relies on manual curation of obtained sequence clusters, which prevents its use in automatic annotation pipeline (Lanza et al., 2014).

Another plasmid database; PLSDB has an extensive set of complete bacterial plasmids from the NCBI database covering records from RefSeq and INSDC (DDBJ, EMBL-EBI, and Gen-Bank). All the plasmids present in the database are annotated using ARG- ANNOT (Gupta et al., 2014), CARD (Jia et al., 2016), ResFinder (Zankari et al., 2012) and VFDB (Chen et al., 2005), and characterized by PlasmidFinder and pMLST (Carattoli et al., 2014).

A comprehensive plasmid database; *Brooks et al* database contains 10,892 complete plasmid sequences and related metadata from NCBI and all available annotated bacterial genomes. (Brooks et al., 2019).

1.11 Aim of the study

The aim of this study was to perform *in silico* detection of AMR genes and plasmids in the selected WGS of Norwegian isolates from *E. coli*, *K. pneumoniae* and *A. baumanii*

The goal was accomplished by achieving the following secondary objectives:

- Annotation of *de novo* assembled WGS (for resistance genes and incompatibility groups) as well as plasmid only contigs (for plasmid detection)
- Assessing the prevalence of most abundant ESBLs in the isolates to evaluate their relevance to Norwegian background of the isolates
- Correlating genotypic antibiotic resistance with phenotypic expression for concordance purposes
- Narrowing down to the same contig number for both resistance gene and plasmids in order to predict the presence of plasmid mediated resistance amongst the isolates

2. Study Design

Schematic workflow of the study is represented in *Figure 6*.

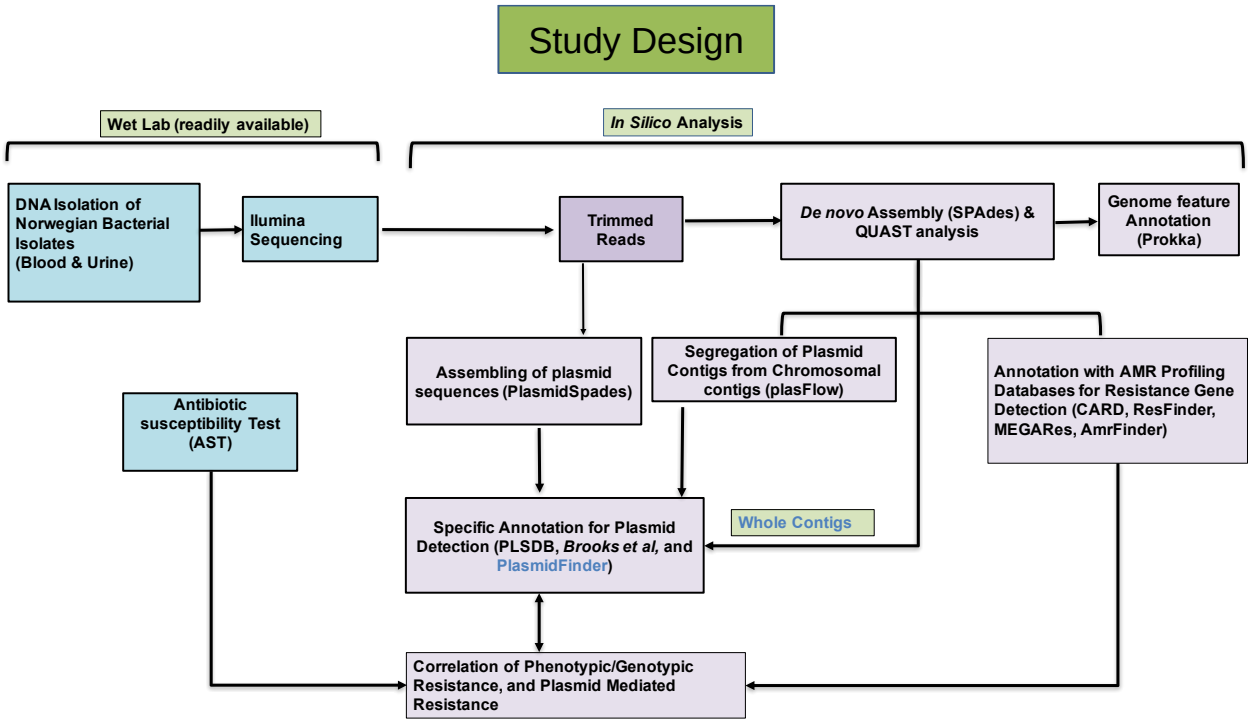


Figure 6:: Workflow of the study

3. Materials and Method

This work is a part of the ongoing bilateral Norway-India project AMR-Diag: A Novel Diagnostic Tool for Sequence Based Prediction of Antimicrobial Resistance funded by the Research Council of Norway.

3.1 Clinical isolates

The sample collection comprised of the details mentioned in Table 3. *E. coli* and *K. pneumoniae* samples belonged to Norwegian patients. *A. baumanii* samples collection was based on carbapenem resistance and selected according to the guidelines of the Reference Centre of Antimicrobial Resistance.

Table 3: Details of clinical Isolates collected for the study

Microorganism	Source		
	Oslo University Hospital		National competence service for the detection of antibiotic resistance (K-Res)
	Blood	Urine	blood, pus, respiratory secretions, abdominal cavity fluid and spinal fluid
E. coli (n=58)	53 (sample numbers100-152)	5 (152-157)	
K. pneumoniae (n=41)	38 (sample numbers 200-236, 240)	4 (241-244)	
A. baumannii (n=11)			11 (sample numbers 301-311)

3.2 Antimicrobial susceptibility testing (AST)

Phenotypic antibiotic resistance profiles for the isolates were received from the fellow master student (Helene Bouras) working on the AMR Diag project. Briefly, antibiotic resistance was assessed using the Sensititre system (ThermoFischer) in the laboratories of NIPH-FHI (Norwegian institute for public health/Folkehelseinstitutet). The results from quality strains were accepted only if they were within EUCAST range of acceptance. Each isolate was classified as sensitive-intermediate or resistant (Alcock et al., 2020) against given antibiotics

according EUCAST (European Committee for Antimicrobial Susceptibility Testing of the European Society of Clinical) guidelines (v 10.0, January 2020).

3.3 Whole Genome Sequencing

All library preparations and Illumina sequencing was performed at Oslo University Hospital Ullevål using the MiSeq platform. The generated output was fastq files with pair-end reads of 300 bp length. These reads were subsequently used for *de novo* genome assembly and annotation. Three isolates with <5X coverage were re-sequenced and included in the final dataset.

3.4 Genome Assembly

3.4.1 Quality Control of Illumina Outputs

Raw sequencing paired-end reads were quality controlled using FastQC v0.11.8 (Andrews, 2010). Using Trimmomatic (Bolger, Lohse, & Usadel, 2014), adaptors and low-quality (with <15 per base quality) sequences were removed. Average quality score threshold of 25 within sliding window of 4 bases was set (if the average quality score over any consecutive four bases drops below 25, the tool will cut the leftmost position in the window and remove the rest of the read). The trimmed reads were subsequently reassessed by FastQC before further analysis. Both FastQC and Trimmomatic were used as part of the Omics box tool (https://www.biobam.com/omicsbox, March 3, 2019).

3.4.2 *De novo* Assembly

Genomes were *de novo* assembled using SPAdes v3.13.1 (Bankevich et al., 2012) using default settings. Contigs < 500 bp were discarded. The command used was:

'~spades.py –k 21, 33, 55, 77 -1 [path to forward reads] -2 [path to reverse reads] –o [path to output file]

Assemblies were assessed by QUAST v 4.6.0 (Gurevich et al.). All statistics were based on contigs of size ≥ 500 bp unless otherwise noted. The command used was:

python quast.py [options] <contig_file(s)>--o <output_dir>

De novo plasmid assembly from WGS in a few isolates (104, 125, 142, 211, 225, 240) was performed using PlasmidSpades v 3.9 with default settings (Antipov et al., 2016). The

resulting plasmid assemblies were further BLAST searched for plasmids. The command for SPADes was used with the addition 'plasmid' flag:

'~spades.py --plasmid –k 21, 33, 55, 77 -1 [path to forward reads] -2 [path to reverse reads] –o [path to output file]

3.5 Genome Annotation

3.5.1 Genome Features Annotation

Annotation of features with the *de novo* assembled genome was performed with Prokka (version1.12) (Seemann, 2014) using default parameters. Counts of features (Genes, CDS, tmRNA, tRNAs, Bases, and repeat regions) were identified along with products of the genes. The command used was:

prokka<input_file.fasta> --outdir <output_directory_name>

3.5.2 *In Silico* Plasmid Identification

Presence of plasmids in *Enterobacteriaceae* genomes was assessed using PlasmidFinder (Carattoli et al., 2014) and mlPlasmids (Arredondo-Alonso et al., 2018). In the first step of the analysis, PlasmidFinder database was used to identify plasmid replicons after 95 % identity as a cutoff.

Next, to improve the plasmid replicon detection and to identify all contigs representing a plasmid, the mlPlasmids (web interface) tool was used. The best-matching hits in each genome for each replicon sequence were given as output, using 0.5 as posterior probability of belonging to the plasmid or chromosomal class and 1000 bp being the minimum sequence length.

Since plasmidFinder and mlPlasmids databases do not include *A.baumanii* genome, the comparison was restricted to Enterobacteriaceae. In this regard, the *de novo* assemblies were separated into plasmid and chromosomal contigs using a neural network model; PlasFlow (Krawczyk et al., 2018) based on the genome signatures of chromosomes and plasmids sequences. The resulting plasmid only sequences of all three bacteria were BLAST searched for the most similar/reference plasmids in *Brooks et al* and PLSDB databases with 95% identity as a cut-off. The output files were filtered with the selection of only those contigs

which showed over 80% contig coverage and a length between 500-100000 bp length, as the plasmids rarely exceed 100 kb size (Smillie et al., 2010).

3.5.3 Resistance Gene Identification

Since the transferrable ARGs are typically of greater concern, ResFinder (version 2.1) (Zankari et al., 2012), which focuses on acquired ARGs, was used for the *in silico* prediction of acquired antibiotic resistance genes in the current study. In addition, Comprehensive Antibiotic Resistance Database – CARD (Alcock et al., 2020) was also employed to search for AMR genes. Two other resistance gene databases; MEGARes and AmrFinder (done by Erasmas fellow: Clàudia López) along with genome annotation with Prokka (without identity and matching length details) were also used to extract resistance genes. Minimum 60% of gene length coverage and a sequence identity of 95% was used as criteria to select the genes from AmrFinder and MEGARes.

All the BLASTN commands for both ARG databases and the plasmid databases were as follows:

blastn –<query> -db <database> -outfmt '6 qseqid sseqid salltitles length qstart qend sstart send' –perc_identity 95 –word_size 28

3.6 Correlation Analysis

3.6.1 Identifying Plasmids Hosting the Resistance Genes

The prediction of location of antibiotic resistance genes and plasmids on the same contig was performed to confirm plasmid mediated resistance. It was accomplished by combining the results of resistance prediction and plasmid detection outputs. Contigs identified by CARD/ResFinder were tallied with those which were carrying plasmid replicons detected by PlasmidFinder, Brooks et al, and PLSDB.

3.6.2 Contig Source Comparison for Plasmid Detection

It was important to see which contig source (plasmid contigs or WGS contigs) was a better option to be used as an input for plasmid detection tools. For this purpose, in a few isolates (104, 125, 142, 211, 225, 240), WGS were assembled with PlasmidSPADes. Furthermore assembled plasmids for the same isolates using plasmid exclusive contigs (SPADes, PlasFlow)

were generated. At the end, results from both assembly techniques were used as input for plasmid detection.

4. Results

4.1 Phenotypic Antimicrobial Resistance

Forty-one percent of *E. coli* isolates were resistant to ampicillin. Resistance to ceftazidime (13%), trimethoprim (22%), ciprofloxacin (14%), and gentamicin (7%) was also found, as shown in Figure 7. Only 2 % of *E. coli* isolates were resistant to chloramphenicol. No isolates were able to grow in the presence of colistin, meropenem, and tigecycline. In *K.pneumoniae*, resistance to ampicillin (98%), trimethoprim (16%). Moreover, resistance to ceftazidime (13%), cefotaxime and tigecycline (11% each), ciprofloxacin and gentamicin (9% each) was also present. Resistance to colistin was not shown by any of the isolates. In *A.baumannii*, all 11 isolates (100%) were resistant to ciprofloxacin, and meropenem, while ten of them (91%) were found to be reistant to gentamicin.

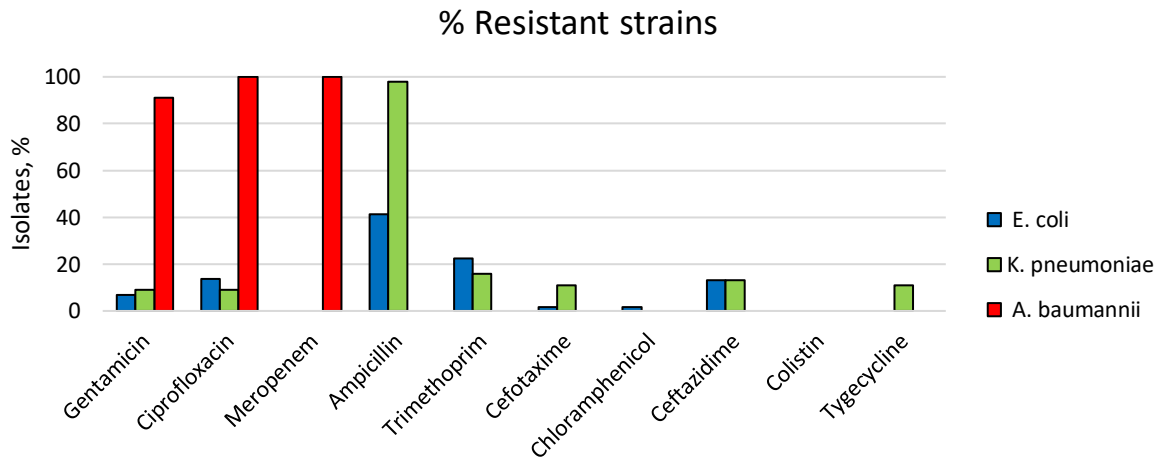


Figure 7: Prevalence (%) of resistant Ecoli (n=58), K.pneumoniae (n=42), and A.baumannii (n=11) strains (%)

Regarding resistance patterns, 28% of *E.coli* isolates were not resistant to any of the antibiotics tested (Figure 8). Resistance to at-least one antibiotic (15%) (either of ampicillin, ciprofloxacin, trimethoprim, and chloramphenicol), resistance to at-least two antibiotics (10%) (ampicillin combined with either of gentamicin, ciprofloxacin, and trimethoprim), and resistance to at-least three antibiotics (4%) (ampicillin combined with either of gentamicin and trimethoprim, ciprofloxacin and trimethoprim, gentamicin and trimethoprim) was prevalent in *E.coli* isolates. Only 1% *E.coli* isolates were resistant to at-least four antibiotics (ampicillin, cefotaxime, ciprofloxacin, and trimethoprim) (Figure 8).

Resistance to one antibiotic (ampicillin) was the most common in *K. pneumoniae* isolates (26%). Resistance to at-least two antibiotics (10%) (ampicillin with either of trimethoprim, tigecycline, and chloramphenicol), resistance to at-least three antibiotics (4%) (ampicillin with either tigecyline and trimethoprim, or cefotaxime and ceftazidime), resistance to at-least five antibiotics (2%) (ampicillin, cefotaxime, ceftazidime, ciprofloxacin, and gentamicin), resistance to at-least six antibiotics (1%) (ampicillin, cefotaxime, ceftazidime, chloramphenicol, ciprofloxacin, and gentamicin), and resistance to at-least seven antibiotics (1%) (ampicillin, ceftazidime, chloramphenicol, ciprofloxacin, gentamicin, tigecycline, and trimethoprim) was also found.

In *A.baumanii*, 91% of the isolates were resistant to at-least 3 antibiotics (ciprofloxacin, gentamicin, meropenem) , and the rest (9%) were resistant to at-least two antibiotics (ciprofloxacin, meropenem) (Figure 8).

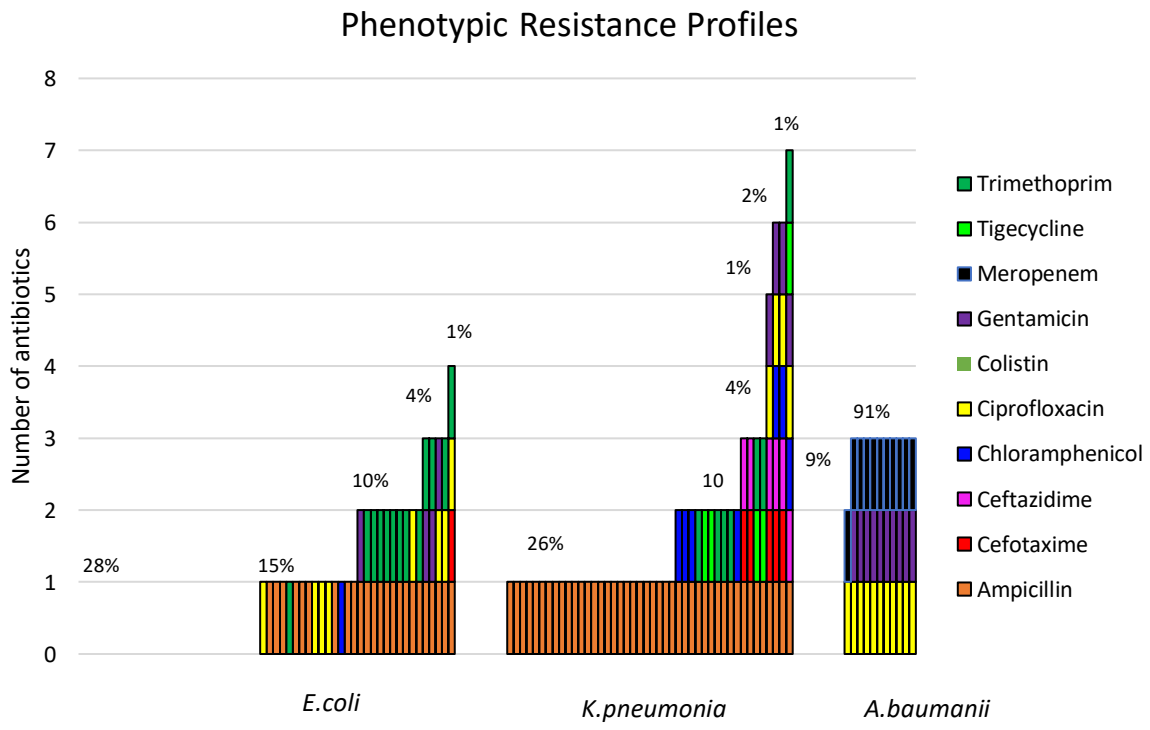
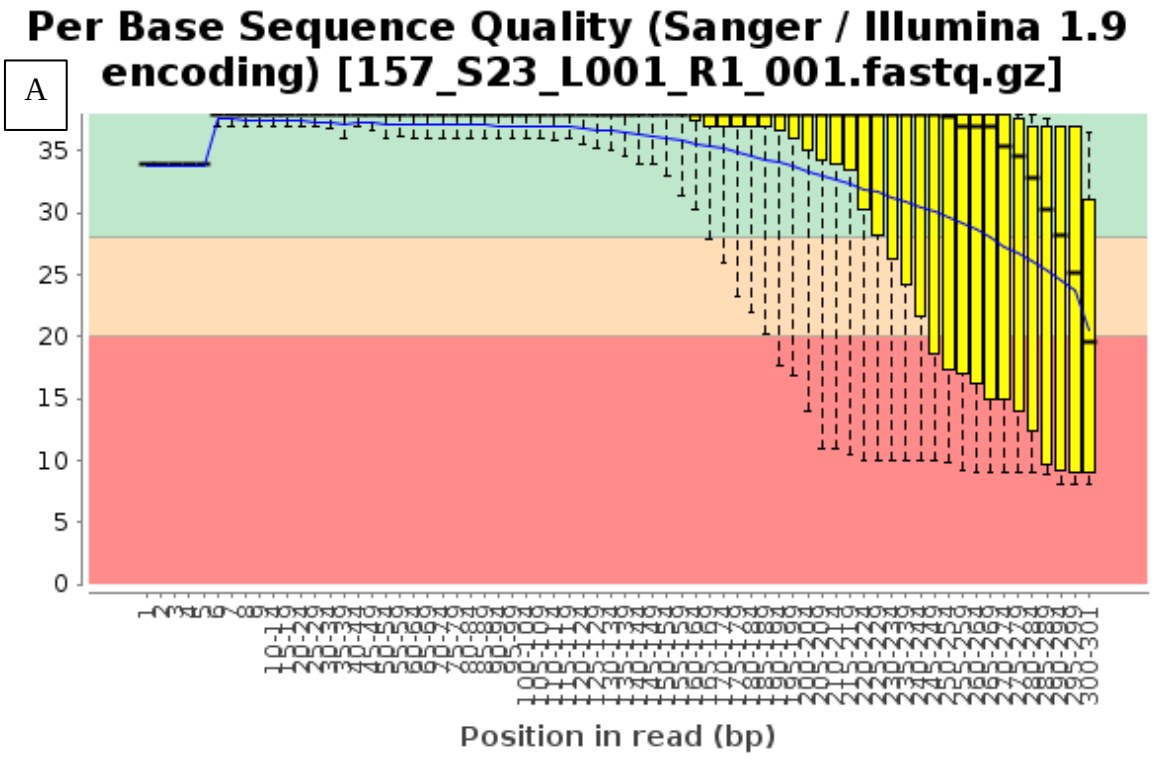


Figure 8: Resistance profiles of all the isolates (n=111) showing resistance to antibiotics ranging from 1 to 6. The colored bars show the type of resistance.

4.2 Trimming of Illumina Sequence Reads

The pre-processing step illustrates that bases at the end of reads tend to have lower quality. The quality trimming step improves the read quality leading to a higher average quality (*Figure 9* A & B). The quality score for each base ranges from -5 to 40, and in our study, the reads for each sample were of highest quality (>30) between 75-225 bp and the quality dropped at 5'. Despite that, the overall quality of the reads remained towards high. The processing of reads prior to assembly removed overrepresented sequences (only 8 samples failed in overrepresented sequences and the rest were either passed or with warning), thus decreased the duplication. The sequencing resulted in sequence data comprising average 704571 ±570319 reads per file (supplementary file 10). With trimming, the number of both surviving reads were reduced to average 660694± 548544, as can be seen in *Figure 10*.



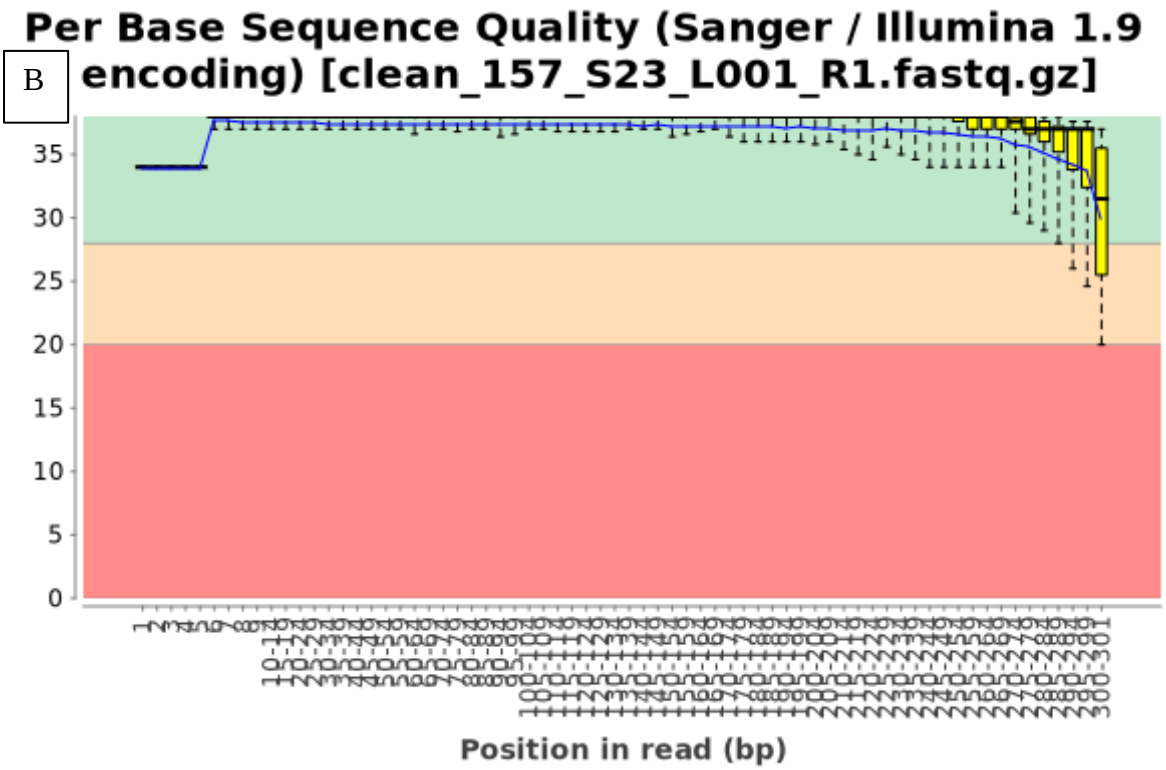


Figure 9: FastQC visualization of per base sequence quality of E. coli isolate 157 before (A) and after (B) trimming of adapters and low quality reads.

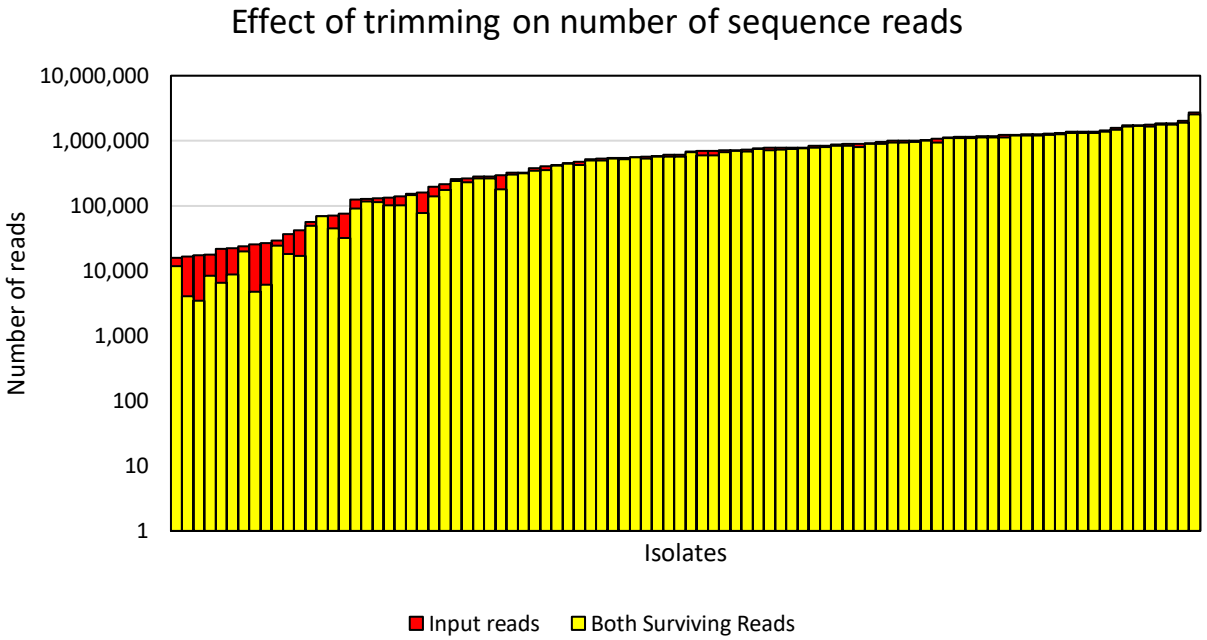


Figure 10: Influence of quality-based trimming on sequencing reads. Yellow bars indicate reads after trimming. Red bars indicate the number of reads before trimming.

4.3 *De novo* Assembly

Trimmed sequences were *de novo* assembled with SPAdes (Bankevich et al., 2012). The contigs were further put to quality check using Quast (Gurevich et al., 2013). The results from the final quality assessment are shown in Table 4.

Table 4: Mean ± Standard deviation of SPAdes' de novo assembly of 58 E. coli, 41 K.pneumoneae, and 11 Acinetobacter isolates, visualized by QUAST. All statistics were based on contigs length ≥500 bp, minimum alignment length of 65, and ambiguity one.

Statistics	***E. coli (n=58)***	***K. pneumoniae (n=42)***	***A. baumannii (n=11)***
Number of contigs	471±622	840±961	347±293
Largest contig (bp)	373475±229473	332403±377228	219493±150134
Total assembly length (bp)	5058946±1000581	4156185±2314802	4060152±92430
GC %	50±1	56±2	39±0
N50	146114±106360	128977±148029	82805±78310

Both *E. coli* and *K. pneumoniae* displayed average GC% ≥50, while A. baumannii had 39%. Nearly half of *E. coli* isolates (48 %) had largest contig length in the range of 500 Kbp followed by 31% isolates in 800 Kbp range. In *K. pneumoniae*, most of the isolates (26%) had contig length in the range of 10 Kbp followed by 17% isolates having contig length in the range of 500 Kbp and 800 Kbp each. *K. pneumoniae* isolates (10% and 14%) also had contig length in the range of 20 Kbp and 1000 Kbp each and 50 Kbp respectively *A. baumanii* displayed 73 % of *A. baumannii* isolates displayed length within the range of 500 Kbp followed by 18% and 9% isolates in the range of 100 Kbp and 50 Kbp respectively. Regarding contiguity, 12, 18, and 3 of *E.coli, K. pneumoniae*, and *A. baumanii* isolates had number of contigs above 500 as shown in Figure 19.

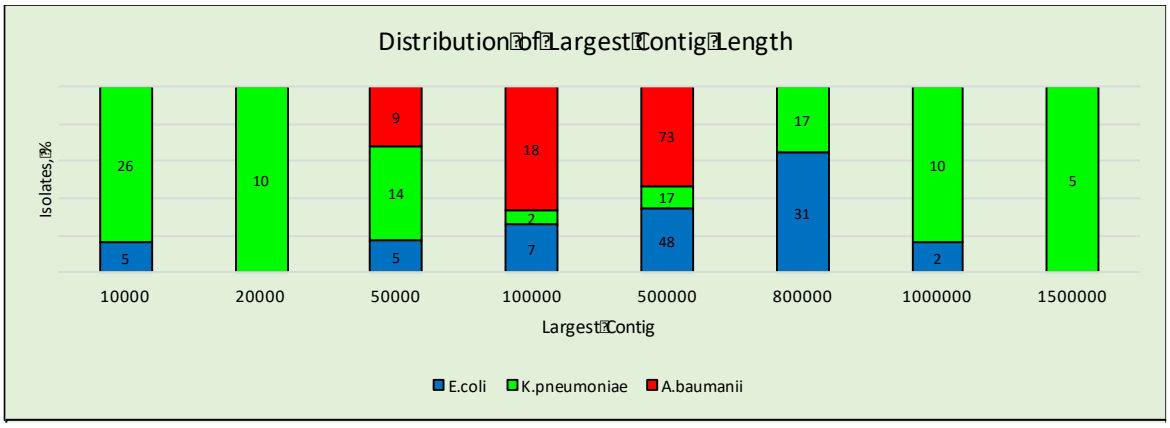


Figure 11: Graphical representation of distribution of largest contig length

4.4 Genome Features Annotation

Total genome size in most of the isolates ranged between 4-5 Mbp for *E.coli* and *K.pneumoniae* while a few isolates fell into either below 2 Mbp or above 6 Mbp (supplementary file 5 and *Figure 18*). Number of CDS in *E.coli* and *K.pneumoniae* ranged between 4500-5500 in most of the cases while a few isolates had below 2000 or above 6000 genes. tmRNA were found to be either one or two in most of *E.coli* and *K.pneumoniae* isolates while a few had no tRNAs. Total genome size in the form of base pairs was between 4 - 4 Mbp for 7 of *A.baumanii* isolates, while 4 isolates were found to be in the genome size of above 5 Mbp. Number of CDS in *A.baumanii* ranged between 3500-4000 in 7 cases, while 4 isolates had above 5000 genes. tRna in 7 *A.baumanii* isolates ranged between 60-70 with 4 isolates having tRNA above 80. tmRNA were one in number in 10 *A.baumanii* isolates with only one isolate having two tmRNA. There were 21, 9, and 2 isolates of *E.coli, K.pneumoniae, A.baumanii* where the repeat regions were present while 37, 33, and 9 isolates of all three bacteria had no repeat regions.

4.5 Prevalence of Antimicrobial Resistance Genes

WGS sequences were screened for AMR genes in different resistance gene databases as shown in (*Figure 12*). According to the results obtained from CARD, the most occurring drug class in which AMR genes were detected was β-lactam.

In *E.coli*, the most prevalent gene was TEM (38%) with other β-lactamases in small percentages (SHV (5%), OXA (7%), VIM (2%), and CTX-M (3%). *ampC* being inherent to

E.coli was found in 93% of the isolates. Two β-lactam representative genes LEN and OKP-A were not present at all. The most abundant non- β-lactamase gene found was in 28% of *E. coli* isolates i.e. *aac/aph* representing gentamicin followed by 16% isolates with *dfr* gene for trimethoprim resistance and 14% isolates with *tet* gene (tetracycline, tigecycline). Fluoroquinolone (*qnr*) and phenicol genes *(cat/Cml/floR)* were found in only 3% of isolates.

In *K.pneumoniae*, the β-lactam antibiotic resistance gene with highest percentage of isolates was SHV (67%) while other β-lactamase resistance genes were TEM (19%), OXA (7%), LEN (5%), OKP-A (2%), CTX-M (7%) and *ampC* (5%). Non-β-lactam antibiotic resistance occurred mainly with the genes (*aac/aph*) for aminoglycoside in 26% of *K.pneumoniae* isolates, while resistance to other non-β-lactam was in 10%, 7%, 7%, and 5% isolates for chloramphenicol, trimethoprim, fluoroquinolone, and tetracycline/tigecycline respectively.

Regarding β-lactam resistance gene prevalence in *A.baumanii*, the most abundant gene was OXA (100% isolates) followed by *ampC* (91% isolates)), and TEM (45% isolates)). Nine percent of isolates had both SHV and CTX-M. The most abundant non β-lactam gene found in *A.baumanii* was *aac/aph* followed by *tet* (55% isolates)) and chloramphenicol genes (*cat/Cml/floR*) (36% isolates).

Similar gene pattern was detected by other databases (ResFinder, MEGARes, AmrFinder, and Prokkka) as well, but with different prevalence percantage. For details of all the databases see supplementary file 1.

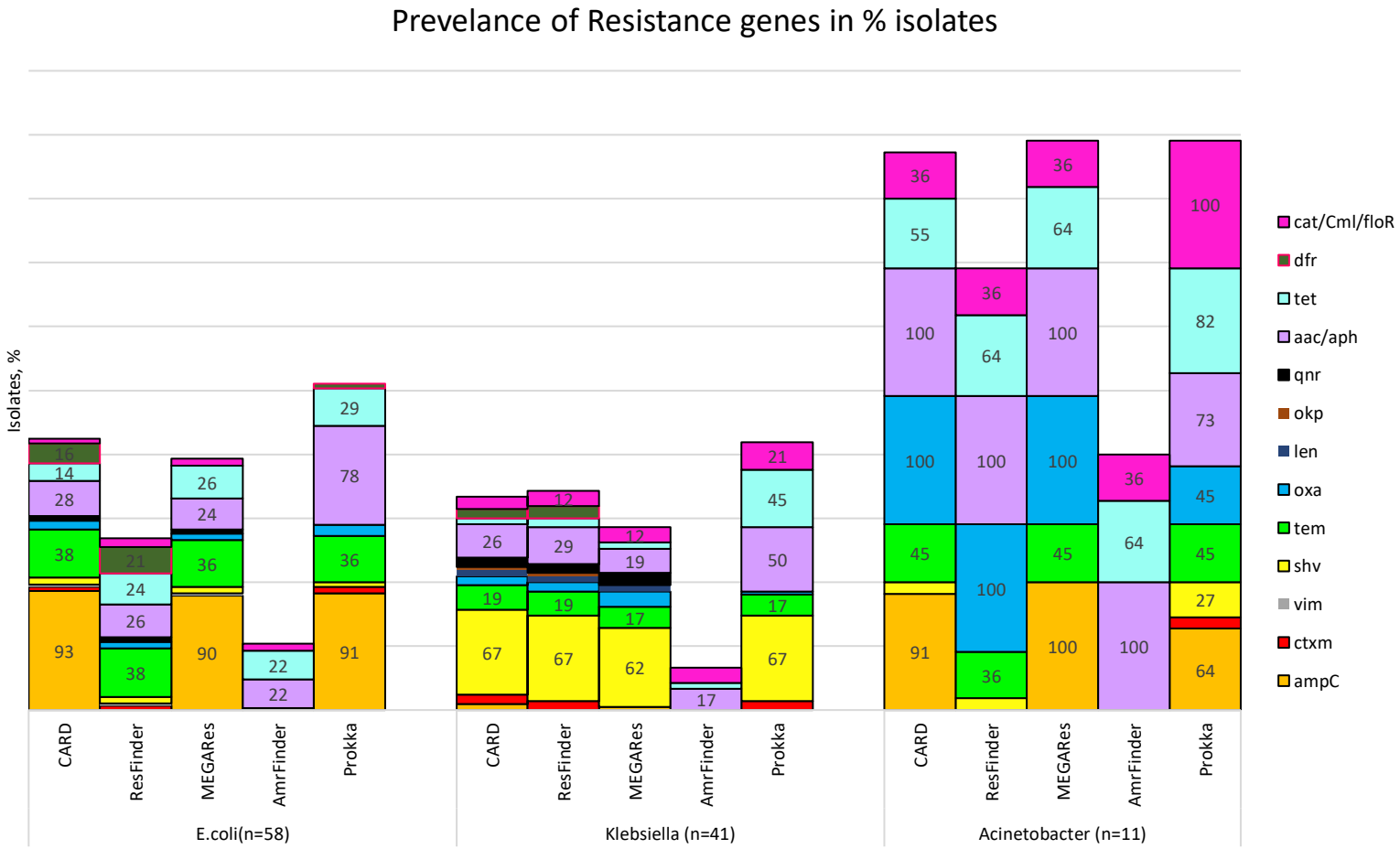


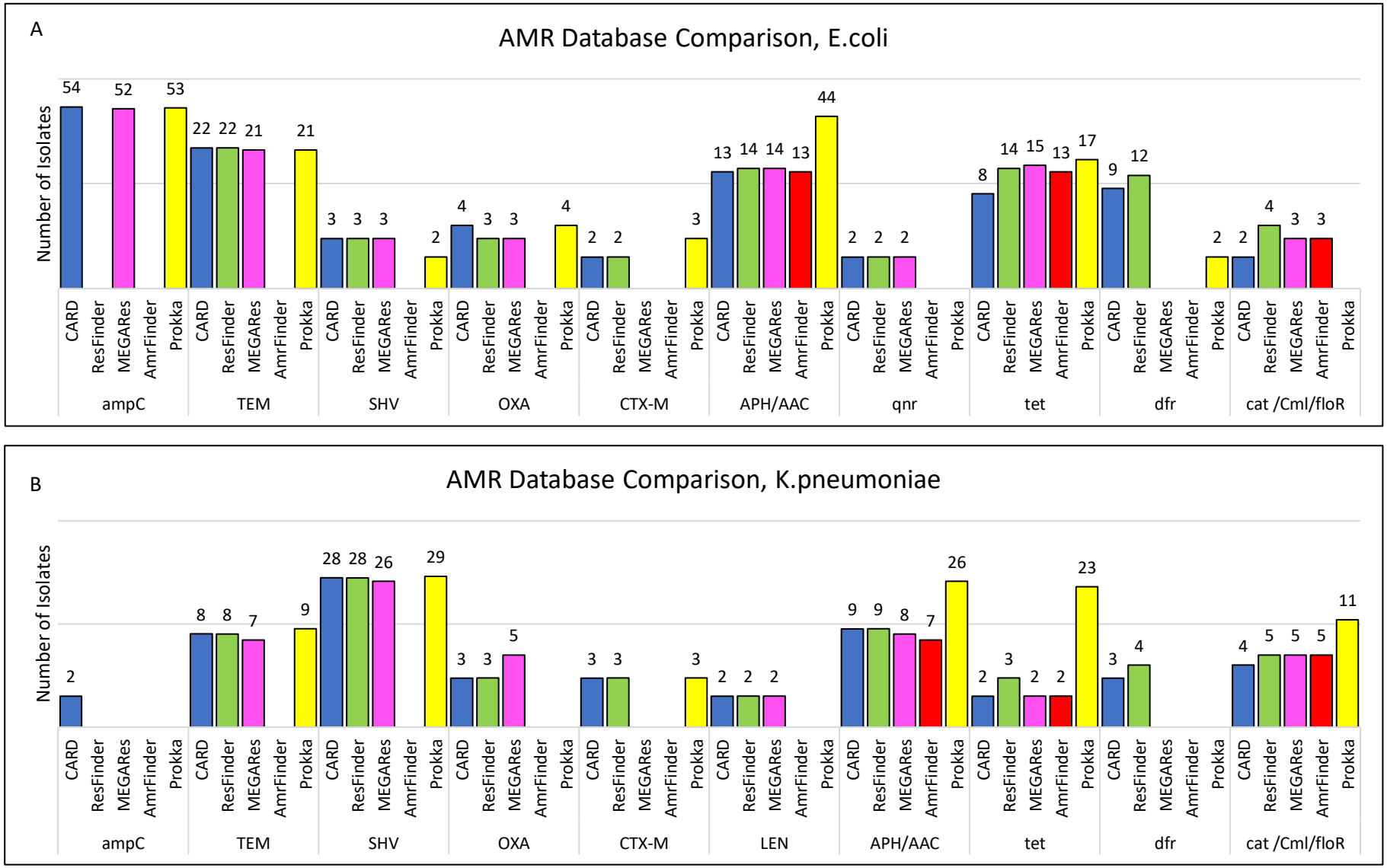
Figure 12: Relative abundance of antimicrobial resistance genes in E. coli, K. pneumoniae and A. baumannii detected using different databases

4.6 AMR Database Performance Evaluation

The comparison of AMR databases was performed, based on the values detected in at-least two isolates in at-least one database.

CARD database detected five β-lactamases (*ampC*, SHV, TEM, OXA, CTX-M) along with genes of all other classes in *E. coli*. Six β-lactamases (*ampC*, SHV, TEM, OXA, CTX-M, LEN) along with resistance genes of aminoglycoside, tetracycline, chloramphenicol, trimethoprim, fluoroquinolone in *K. pneumoniae*, while three β-lactamases (*ampC*, TEM, OXA) along with aminoglycoside, tetracycline and phenicol genes were detected in *A. baumannii* as shown in *Figure 13* and in supplementary file 9.

ResFinder detected all the genes, as did CARD. However, ResFinder did not detect *ampC*, as it was intrinsically present. In agreement with CARD and ResFinder, MEGARes detected all the genes except dfr and CTX-M in both *E. coli* and *K. pneumoniae*. However, it was in agreement with CARD in the detection of *ampC* in *E. coli* and *A. baumannii*. AmrFinder detected three non-βlactamases resistance genes (*aac/aph tet*, *cat/Cml/floR*) in all three bacteria. Prokka was in agreement with CARD, ResFinder, and MEGARes in the detection of TEM, *aac/aph, tet,* and CTX-M in all three bacteria. Regarding *ampC*, it was in agreement with CARD and MEGARes in *E. coli* and *A. baumannii, while* the gene SHV was only detected by Prokka in *A. baumannii.*


A
AMR Database Comparison, E.coli
Number of Isolates
54 52 53 22 22 21 21 3 3 3 2 4 3 3 4 2 2 3 13 14 14 13 44 2 2 2 8 14 15 13 17 9 12 2 2 4 3 3
CARD ResFinder MEGARes AmrFinder Prokka
ampC TEM SHV OXA CTX-M APH/AAC qnr tet dfr cat /Cml/floR
B
AMR Database Comparison, K.pneumoniae
Number of Isolates
2 8 8 7 9 28 28 26 29 3 3 5 3 3 3 2 2 2 9 9 8 7 26 2 3 2 2 23 3 4 4 5 5 5 11
CARD ResFinder MEGARes AmrFinder Prokka
ampC TEM SHV OXA CTX-M LEN APH/AAC tet dfr cat /Cml/floR

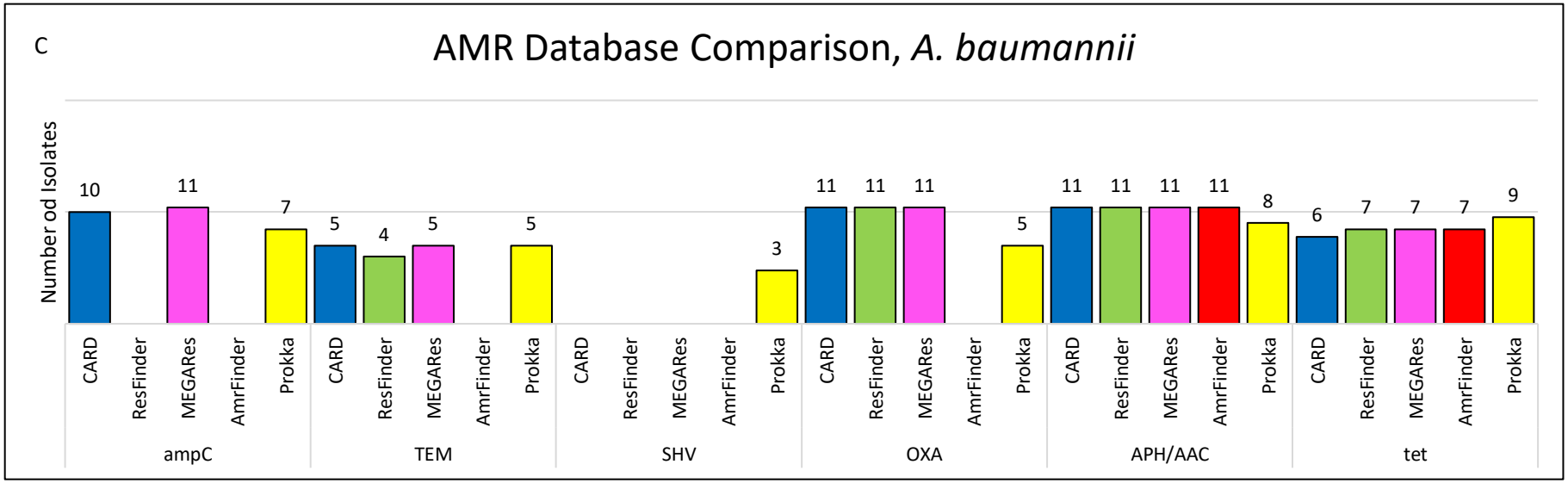


Figure 13:. AMR database comparison on the basis of presence of resistance genes in number of isolates by CARD, ResFinder, MEGARes, AmrFinder, and Prokka. Reported genes were detected in at-least two isolates in at-least one databases

4.7 *In Silico* Dectection of Plasmids

PlasmidFinder was used to detect plasmid replicons, while Brooks et al and PLSDB was also used to search plasmids as shown in *Figure 14*. Regarding plasmids detected by Brooks et al database, 12% of *E. coli* strains had plasmid p2_000837, while the rest of 11 identified plasmids were found in below 10% of isolates. There were four plasmids (pIB_NDM_1, p2_W5-6, pCHL5009T-102k-mcr3, pVir_020022) in 2% *K. pneumoniae* isolates which were also shared with *E. coli*. Only one plasmid (pHZ23-1-1) was present in 9% of *A. baumannii* isolates. Unknown plasmids in all three bacteria can be viewed in supplementary file 2.

PLSDB detected many plasmids in *E. coli* isolates including Plasmid A with the maximum percentage (10%) while the rest of the 20 plasmids were present in under 4% isolates. In *K. pneumoniae*, only 9 plasmids were identified with plasmid 4 in as many as 4% isolates while the rest of the plasmids were in less than 3 % isolates. Only plasmid A and plasmid B were present in both *E. coli* and *K. pneumoniae*. No identified plasmid was present in *A. baumannii* although unknown plasmids are shown in supplementary file 2.

PlasmidFinder database detected 31 previously known plasmids in *E. coli* isolates. IncFIB (64%), Col156 (47%), IncFII (21%), IncFII(29) (22%), Col8282 (10%), Col(BS512) (17%) were prominent plasmids while the rest of the plasmids were found in below 10% of the isolates. In *K. pneumoniae*, IncFIB(K) plasmids were present in the highest percentage (27%) while 15% of the isolates had both IncFII(K) and Col(MG828). Moreover, Col(8282) and IncHI-1B(pNDM-MAR) both were present in 12% of the isolates. The rest of the plasmids were represented by below 10% *K. pneumoniae* isolates. Contigs of plasmid origin, that could not be assigned to one particular plasmid/plasmid class with high degree of certainty (because of a lot of hits to different plasmids of different classes) can be viewed in supplementary file 2.

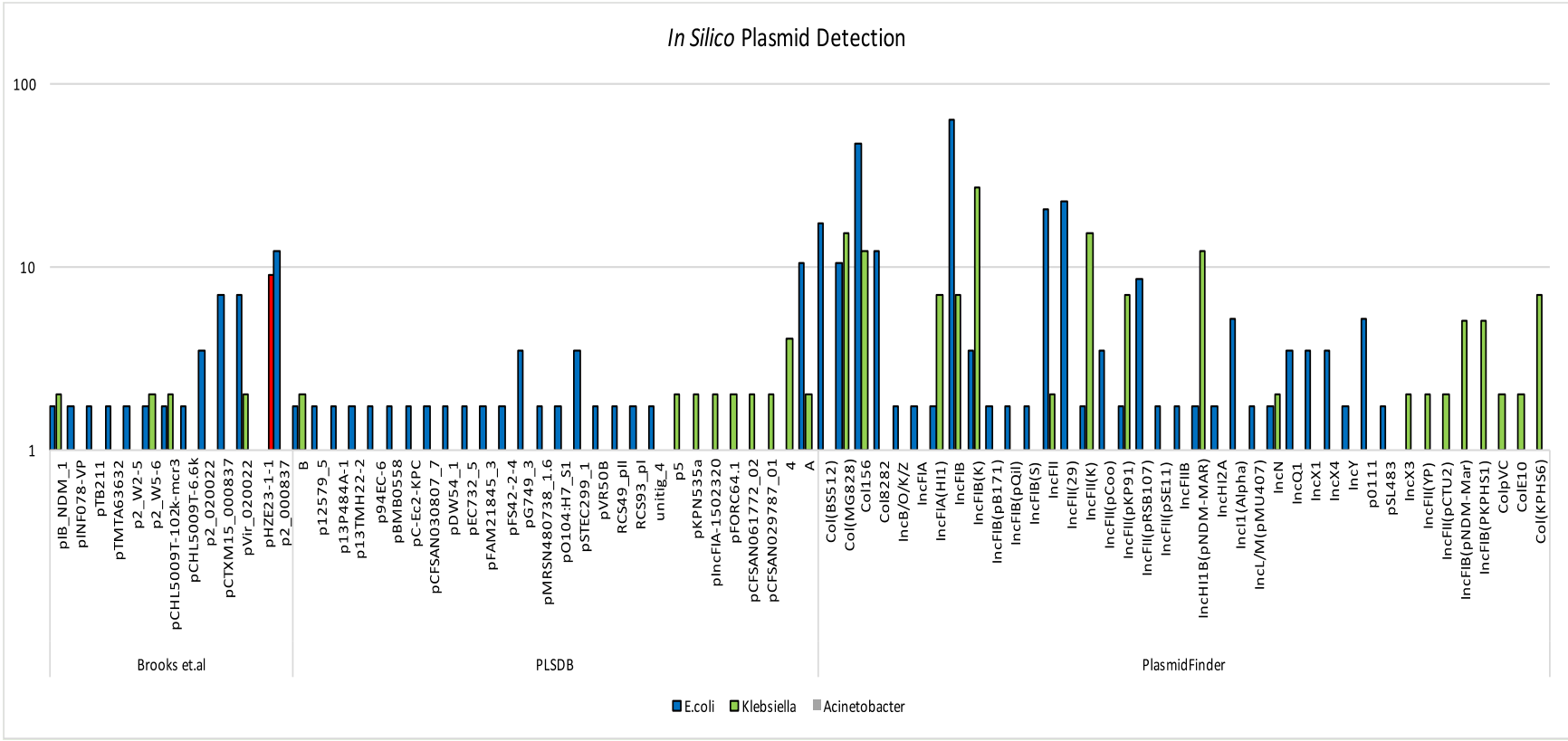


Figure 14: Plasmid Detection by multiple databases; Brooks et al, PLSDB, and PlasmidFinder

4.8 Correlation between detected resistance genes, Phenotype resistance profile, and plasmids

ResFinder detects only acquired genes and ignores chromosomal mutations (Xavier et al., 2016) wherease CARD includes chromosomal genes and mutations. Therefore resistance genes from CARD database were correlated with both the phenotypic resistance and the plasmid found in all the isolates.

Table 5 shows a total of 75 isolates that had resistance to β-lactam antibiotics (ampicillin, cefotaxime, ceftazidime, meropenem), out of which 63 had the corresponding resistance genes either alone or in combination (*ampC*, SHV, CTX-M, TEM, LEN, OXA). *ampC* and the phenotypic resistance to β lactam antibiotics matched 34 times, and 33 times the gene *ampC* was present without the isolate displaying and phenotypic resistance. TEM was found in 36 isolates which was complimented by 31 isolates having phenotypic resistance to β lactam antibiotics. SHV gene was present in 32 isolates which was again reciprocated by 31 isolates having phenotypic resistance to β lactam antibiotics. A total number of three isolates had CTX-M gene which displayed resistance to β lactam antibiotics in all three isolates. LEN gene was present in two isolates, and out of these two only one showed resistance to β lactam. 18 isolates had OXA genes with 16 isolates matching the expected phenotypic resistance to β lactam antibiotics. A total number of 33, 30, and 39 isolates had the presence of both the resistance to β lactam antibiotics and different types of plasmid; Col, IncFII, and IncFIB respectively.

The antibiotic resistance against aminoglycosides (gentamicin) was found in 18 isolates, and either of *aac/aph* genes was found in 17 of the respective isolates. On the other hand, 15 isolates were found to have the genes *aac/aph* but no resistance phenotypically was detected. There were 5, 4, 6 times when the isolates had both the resistance to gentamicin and the plasmid Col (different types), IncFII (different types), and InFIB (different types) respectively.

There were 12 dfr genes that matched with the phenotypic resistance against trimethoprim in 20 isolates. On 14, 9, 15 occasions, the isolates had both the resistance to trimethoprim and the plasmid Col (different types), IncFII and InFIB (different types) respectively.

A sum of 23 isolates were found with fluoroquinolone resistance (ciprofloxacin), out of which 3 times there was a corresponding *qnr* gene, while 2 isolates had *qnr* gene without any phenotypic resistance to ciprofloxacin. Col (different types), IncFII (different types), and IncFIB (different types) plasmids in combination with ciprofloxacin resistance were found in 6, 5, and 9 isolates.

Eight isolates found with chloramphenicol resistance were being reciprocated with either of *cat/Cml/floR* genes in only two isolates, while 7 isolates had these genes but with no expressed resistance to chloramphenicol. There were 5, 3, and 3 isolates had Col (different types), IncFII (different types), and IncFIB (different types) along with chloramphenicol resistance.

Five isolates had tigecycline resistance but no corresponding *tet* gene was present although *tet* gene was present in 19 isolates without showing resistance to tetracycline or tigecycline antibiotic. It occurred 4, 2, and 1 time that the isolates had Col (different types), IncFII (different types), and IncFIB (different types) along with tigecycline resistance.

Table 5: Concordance between phenotype and genotype for predictions made using a database of resistance determinants and the plasmids. Red color coded genes and antibiotics have concordance. Detailed table can be viewed in appendix Table 8. .

Isolates	Antibiotic Resistance Profile	In Silico Resistance Genes (CARD)	In Silico detected Plasmids
100	Ciprofloxacin	ampC	IncFII, Col(BS512), IncFIA
102	Ampicillin, Gentamicin, Trimethoprim	ampC, TEM, aac, *dfr*	IncFIB, Col156
103	Ampicillin, Gentamicin	ampC, TEM, *aac*	IncFII(29), IncFIB, Col156
106	Ampicillin, Trimethoprim	ampC , TEM	IncFII(29), IncFIB, IncI1(Alpha)
107	Ampicillin, Trimethoprim	ampC , TEM, dfr, *tet, qnr*	IncFII(29), IncFIB, Col156
108	Ampicillin	ampC, SHV	IncFIB, Col156
110	Ampicillin, Gentamicin, Trimethoprim	AmpC, SHV	IncFIB, Col156
112	Ampicillin	ampC, TEM	IncFII, IncFIB, Col156 , B, pO104:H7_S1
117	Ampicillin, Ciprofloxacin, Gentamicin	ampC, OXA, *aac*	IncFII(pCoo), IncFIB
120	Ampicillin	ampC,TEM, *aph*	IncFII, IncFIB, Col156 , p94EC-6, pCFSAN030807_7, pEC732_5,
121	Ampicillin, Trimethoprim	ampC, TEM, *dfr, aph*	IncFII, IncFIB, Col(MG828), IncB/O/K/Z, IncQ1,
123	Ampicillin, Trimethoprim	ampC, TEM, *dfr , aph*	IncFII(29), IncFIB, Col156.
127	Ampicillin, Trimethoprim	ampC, TEM, *dfr*	IncFII(29), IncFIB, Col(BS512), Col8282, Col(MG828), Col156, p2_020022, p2_000837, pFAM21845_3, p12579_5,
128	Ampicillin	ampC, , TEM, *aph, tet*	IncFII(pRSB107), IncFII(pSE11), IncFIB(pB171), Col156, Col(BS512), pVir_020022, pCTXM15_000837, pINF078-VP, pCHL5009T-102k-mcr3, RCS93_pI, pSTEC299_1, p13P484A-1, p13TMH22-2, pG749_3, A
129	Ampicillin	ampC, TEM, *aph*	IncFII(29), IncFIB, Col156, IncQ1,
130	Ampicilline, Ciprofloxacin, Trimethoprim	ampC, TEM, *dfr, tet*	IncFII(pRSB107), IncFIB, Col156, Col(MG828), pIB_NDM_1, pCTXM15_000837, p2_000837, pTMTA63632
132	Ciprofloxacin	ampC	IncFIB, Col8282, p2_W5-6, IncY,
134	Ampicillin, Trimethoprim	ampC, TEM, *dfr*	IncFII(29), IncFIB, Col156, pCHL5009T-6.6k
136	Ampicillin, Trimethoprim	ampC, *aph,* TEM	IncFII(29), IncFIB, Col156, Col156
142	Ciprofloxacin	ampC	IncFII, B, IncI1(Alpha), pTB211, pVir_020022, pC-Ec2-KPC,
143	Ampicillin, Ciprofloxacin	ampC, TEM	IncFIB, IncFIB(K), p2_000837,
144	Ciprofloxacin	ampC, OXA, TEM	IncFII(K), IncFIB, Col156, Col(BS512), IncFIB(pQil), IncHI1B(pNDM-MAR), pFS42-2-4,
146	Ampicilline, Cefotaxime, Ciprofloxacin, Trimethoprim	ampC, CTX-M, *dfr , aph*	IncFII(pRSB107), IncFIB, pVir_020022, p2_000837
147	Ampicillin	ampC, TEM	IncFII, Col(MG828), Col156
155	Trimethoprim	TEM, *dfr, tet*	IncFII, IncFII(pKP91), IncFIB(K), IncFIA(HI1)
156	Ampicillin	ampC, TEM, *tet*	IncFII, IncFII(29), IncFIB, p2_W2-5, pCTXM15_000837, unitig_4,
157	Ampicillin	TEM, OXA, CTX-M, *aph, aac, tet, cat*	IncFII(pCoo)
200	Ampicillin	SHV	Col(MG828)
201	Ampicillin	SHV	IncHI1B(pNDM-MAR)
202	Ampicillin	SHV	IncFIB(K), Col(MG828)
203	Ampicillin	SHV	IncFII(pKP91), IncFIB(K), Col(MG828)
204	Ampicillin	SHV	ColE10
205	Ampicillin	SHV	Unknown
206	Ampicillin	SHV	IncFII(pKP91), IncFIB(K), Col(MG828)
207	Ampicillin	SHV	Unknown
208	Ampicillin	LEN, SHV	IncFII(K), IncFIB(K)
209	Ampicillin	SHV,	IncFII(K), IncFIB(K), IncFIA(HI1)
210	Ampicillin	SHV	IncFIB(K)
211	Ampicilline, Cefotaxime, Ceftazidime, Chloramphenicol, Ciprofloxacin, Gentamicin	TEM, SHV, *aac, qnr, floR*	Col(MG828), 4, pIncFIA-1502320, pCFSAN061772_02, IncX3, IncFIA(HI1)
212	Ampicilline, Cefotaxime, Ceftazidime, Chloramphenicol, Ciprofloxacin, Gentamicin	TEM, SHV, *aac,, floR*	IncFII(K), IncFIB(K)
213	Ampicilline, Cefotaxime, Ceftazidime, Ciprofloxacin, Gentamicin, Trimethoprim	SHV, CTX-M, TEM, OXA, *aac, aph, qnr, dfr*	p2_W5-6
216	Ampicillin, Cefotaxime, Ceftazidime, Tigecycline, Trimethoprim	SHV, CTX-M, TEM, *dfr, qnr, aph*	IncFII(K), IncFIB(K), Col(KPHS6)
217	Ampicillin, Chloramphenicol, Tigecycline	SHV	IncFIB, Col(KPHS6)
218	Ampicillin	SHV	Col(KPHS6)
220	Ampicilline, Cefotaxime, Ceftazidime, Trimethoprim	SHV, CTX-M, TEM, *aph, dfr*	IncFIB(K)
221	Ampicillin	SHV, TEM	IncFIB(pNDM-Mar), IncN, IncHI1B(pNDM-MAR)
224	Ampicillin	OKP-A	IncFIB
225	Ampicillin, Chloramphenicol, Tigecycline	SHV	IncFII(pKP91), Col156
226	Ampicillin, Chloramphenicol, Tigecycline, Trimethoprim	SHV, TEM	p5, Col156
227	Ampicillin	SHV	IncFII(YP), Col156, pKPN535a
228	Ampicillin	TEM, SHV	IncFII(K), IncFIB(K)
232	Ampicillin	*aac*	IncFII(pCTU2)
240	Ampicillin, Ceftazidime, Chloramphenicol, Ciprofloxacin, Gentamicin, Trimethoprim	TEM, SHV, *aac*	IncFII, IncFIB(pKPHS1), IncFIB(pNDM-Mar), IncHI1B(pNDM-MAR), ColpVC
241	Ampicillin	ampC, SHV, OXA, *aac, aph tet, cat*	IncHI1B(pNDM-MAR)
243	Ampicillin	ampC, SHV, OXA, *aph, aac, tet, cat*	IncFII(K), IncFIB(K), Col156
244	Ampicillin	SHV	IncHI1B(pNDM-MAR)
301	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aac, aph*	Unknown
302	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, TEM, *aac, aph, tet*	Unknown
303	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, TEM, aac, *cat*	Unknown
304	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aac, aph*	Unknown
305	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aph, aac, tet*	Unknown
306	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aph*	Unknown
307	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aph*	Unknown
308	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, TEM, *aph, tet*	Unknown
309	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, *aph, aac, tet, cat*	Unknown
310	Ciprofloxacin, Meropenem	ampC, OXA, TEM, *aph, tet*	Unknown
311	Ciprofloxacin, Gentamicin, Meropenem	ampC, OXA, TEM, *aph, tet, Cml*	pHZE23-1-1

In order to confirm the plasmid mediated resistance, contigs of plasmid origin on which AMR gene was also detected are shown in Table 6. IncFII (different types) plasmids were present in 6 isolates with the resistance gene pattern; TEM, TEM/*aac/aph*, and TEM/*aac/aph*/dfr. These 6 isolates had resistance to either ampicillin or trimethoprim or both. IncI1(Alpha) was found in the isolate 106 sharing the same node with TEM with phenotypic resistance to ampicillin and trimethoprim. In two isolates, resistance to ampicillin and trimethoprim antibiotics, and antibiotic resistance genes *aac/aph* were located on contigs assigned to IncQ1 plasmid. Isolate 130 exhibiting resistance towards trimethoprim, had two plasmids (pTMTA63632, pIB_NDM_1) sharing contigs with *tet* and *dfr* resistance genes respectively. The isolate 211 also had 2 plasmids (IncX3, pCFSAN061772_02) sharing contigs with *qnr*S1, TEM, *aac/aph* resistance genes, and showing phenotypic resistance to ampicillin, cefotaxime, ceftazidime, chloramphenicol, Ciprofloxacin, and gentamicin antibiotics.

Table 6*: Common nodes between resistance genes and plasmid*

Plasmids and Resistance Genes Sharing the same Nodes													
			Corresponding ResFinder Genes					Corresponding CARD Genes						
Plasmid Database	Isolate	Identified Plasmids(n=15)	TEM	aph/aac	dfr	tet	qnrS1	TEM	aph/aac	dfr	tet	qnrS1	Phenotype Resistance	Common Nodes
	130	pTMTA63632											Ampicillin, Ciprofloxacin, Trimethoprim	NODE_98_length_1862_cov_81.187029
Brooks et al	130	pIB_NDM_1											Ampicillin, Ciprofloxacin, Trimethoprim	NODE_75_length_5841_cov_112.714561
	106	IncI1(Alpha)											Ampicillin, Trimethoprim	NODE_60_length_28072_cov_8.245303
	112	IncFII											Ampicillin	NODE_35_length_48973_cov_11.701797
	120	IncFII											Ampicillin	NODE_57_length_26638_cov_15.176455
	121	IncQ1											Ampicillin, Trimethoprim	NODE_49_length_4640_cov_18.198759
	123	IncFII(29)											Ampicillin, Trimethoprim	NODE_51_length_28633_cov_16.121132
	129	IncQ1											Ampicillin	NODE_342_length_4639_cov_14.851950
	134	IncFII(29)											Ampicillin, Trimethoprim	NODE_26_length_65751_cov_24.106379
	136	IncFII(29)											Ampicillin, Trimethoprim	NODE_77_length_11338_cov_21.658282
	147	IncFII											Ampicillin	NODE_5_length_57294_cov_9.913761
	157	IncFII(pCoo)											Ampicillin	NODE_28_length_63384_cov_68.357004
PlasmidFinder	211	IncX3											Ampicillin, Cefotaxime, Ceftazadime, Chloramphenicol, Ciprofloxacin, Gentamicin	NODE_27_length_40222_cov_56.692580
PLSDB	211	pCFSAN061772_02											Ampicillin, Cefotaxime, Ceftazadime, Chloramphenicol, Ciprofloxacin, Gentamicin	NODE_47_length_3694_cov_94.104850

The graphical representation of the contigs harbouring both the resistance genes and the plasmids is shown in *Figure 15*.

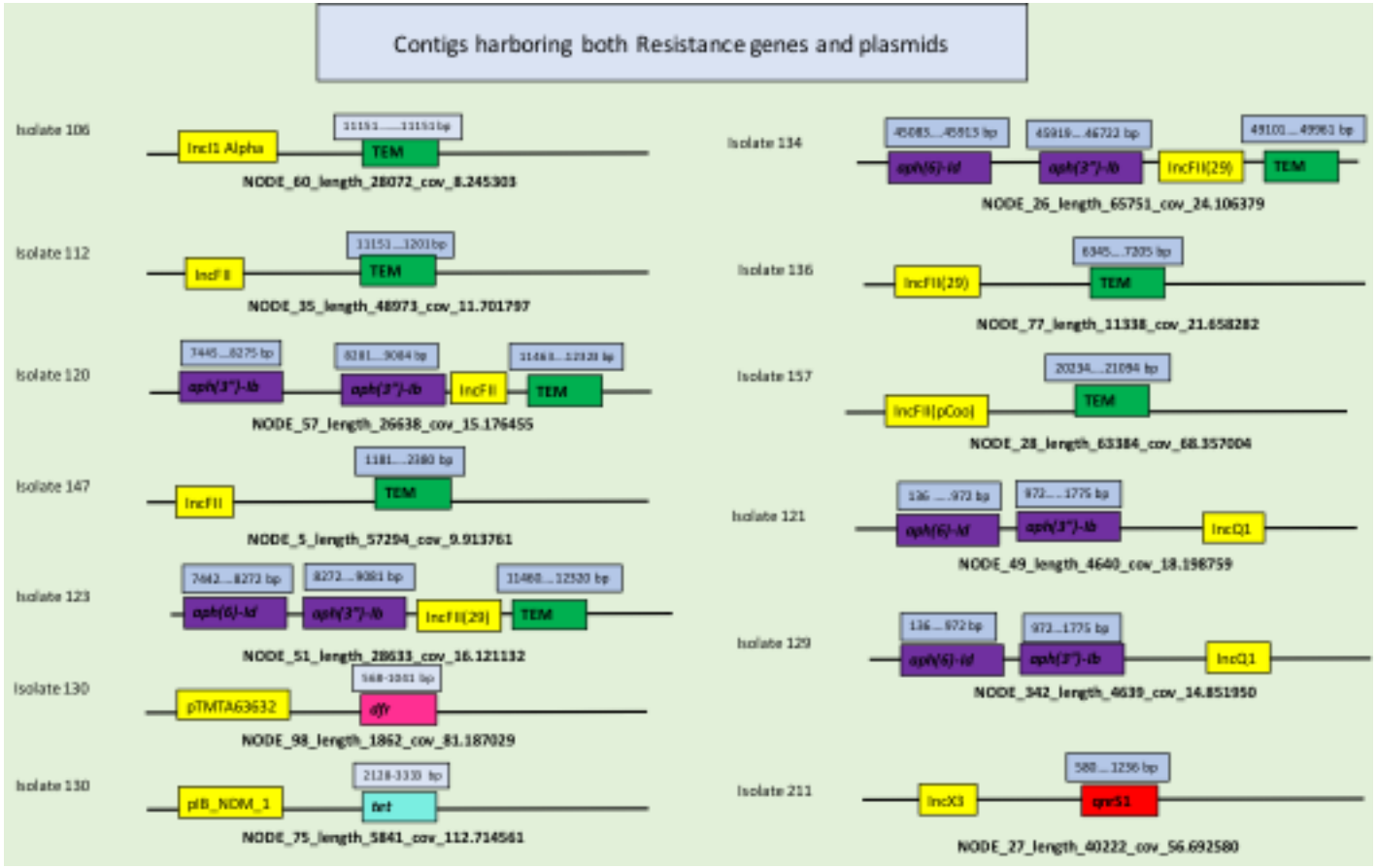


Figure 15: Graphical representation of the contigs being shared by the resistance gene and the plasmids along with their position on the contig

When PlasmidSPAdes assembled contigs were searched for plasmids, and was compared with the SPAdes assemblies (PLasFlow segregated), there were only PlasmidFinder detected plasmids (IncFII(29), IncX3, ColpVC, IncFIB(pkPHS1), IncFII) in 104, 125, 211, and 240 that were common between them (Table 7). Regarding database detection, isolate 104 had only PlasmidFinder detected plasmid, while isolate 125 and 142 had plasmids detected by all three databases. In isolate 211, there were plasmids detected by PlasmidFinder and PLSDB, and no *Brooks et al* plasmids were there. Contigs of plasmid origin, that could not be assigned to one particular plasmid/plasmid class with high degree of certainty (because of a lot of hits to different plasmids of different classes) can be viewed in the supplementary file 3.

Table 7: Plasmid detection in PlasmidSPAdes assembled WGS contigs, and its comparison with plasmids detected in SPAdes assembled and PlasFlow segregated plasmid contigs.

	PlasmidSPAdes Assemblies		SPADes Assemblies	
Isolate	**Database**	**Plasmid**	**Database**	**Plasmid**
104	PlasmidFinder	IncFII(29)	PlasmidFinder	IncFII(29)
104			PlasmidFinder	Col156
104			PlasmidFinder	IncFIB
125	PlasmidFinder	IncFII(29)	PlasmidFinder	IncFII(29)
125	*brooks et. al*	pEC732_5	PlasmidFinder	IncFIB
125	PLSDB	pEC732_5	PlasmidFinder	Col156
125			PlasmidFinder	Col8282
125			PlasmidFinder	Col(BS512)
125			PlasmidFinder	IncX1
125			PlasmidFinder	Col(MG828)
142	*brooks et. al*	NZ_CP012736.1 (unnamed)	*brooks et. al*	pTB211
142	PlasmidFinder	IncFIA	*brooks et. al*	pVir_020022
142	PLSDB	pAR-0428-2'	PlasmidFinder	IncFIIB
142	PLSDB	p009_C	PlasmidFinder	IncI1(Alpha)
142	PLSDB	pL73-3	PLSDB	pC-Ec2-KPC
211	PlasmidFinder	IncX3	PlasmidFinder	IncX3
211	PLSDB	pKOR-e3cb	PlasmidFinder	IncFIA(HI1)
211	PLSDB	pC51_001	PlasmidFinder	Col(MG828)
211			PLSDB	4
211			PLSDB	pIncFIA-1502320
211			PLSDB	pCFSAN061772_02
211			PlasmidFinder	Col156
211			PlasmidFinder	IncFII(pKP91)
240	PlasmidFinder	ColpVC	PlasmidFinder	ColpVC
240	PlasmidFinder	IncFIB(pKPHS1)	PlasmidFinder	IncFIB(pKPHS1)
240	PlasmidFinder	IncFII	PlasmidFinder	IncFII
240			PlasmidFinder	IncHI1B(pNDM-MAR)
240			PlasmidFinder	IncFIB(pNDM-Mar)

5. Discussion

5.1 Quality assessment of sequencing data

As a rule of thumb, 5-10X coverage is suggested in order to support sequence assembly and genome- reconstruction (Kunin et al., 2008). In this regard, only 27 out of 111 isolates had sequencing depth below 5X in our study. This along with range of sequencing depth can be viewed in Figure 16 and Supplementary file 6. One explanation of low sequencing depth can be the fact that sequencing depth is influenced by errors at many stages during DNA processing and library preparation e.g. amplification error, DNA quality, and target region complexity (Jennings et al., 2017; Ma et al., 2019; Quail and Smith), and with peaks in sequencing error shows marked drops in coverage (Ekblom et al., 2014). Another reason for low sequencing depth could be that Illumina sequencing platform favours GC-balanced regions that have fewer reads in GC poor regions, which usually results in uneven sequencing depth across genome (Sims et al., 2014). However, average GC% in *E. coli* (50%), and *K. pneumoniae* (56%), *and A. baumannii* (39%) in our study does not agree with the statement above. Abrupt ARG spread across different contigs can result in low coverage in some isolates which is responsible for discordance between ARGs detection and phenotype (Clausen et al., 2016). In agreement to this finding, nine *E. coli* and *K. pneumoniae* isolates with low coverage (1X) displayed phenotypic resistance without corresponding resistance genes in our study (Figure 17).

5.2 *De novo* Assembly

Regarding contiguity, 12, 18, and 3 isolates of *E. coli, K. pneumoniae,* and *A. baumannii* had more than 500 contigs (Figure 19 and see supplementary file 4). De Bruijn graphing techniques like SPAdes specifically look for the exact features that repetitive elements create within a graph such as convergent, divergent or cyclic paths (Ricker et al., 2012), and therefore terminate at these repetitive elements to avoid them to be overly compressed in the final assembly. Since repeat regions were detected in a few isolates as shown by the high number of contigs (see supplementary file 5), this resulted in a more fragmented assembly for these isolates. However, resistance genes were found despite the fragmented assemblies in our study.

5.3 AMR Database Comparison

CARD, ResFinder, and MEGARes have been in agreement with each other in most of the antibiotic resistance gene detection in our study (*Figure 13*). On all occasions, where CARD database detected a resistance gene, ResFinder also did (although only HGT resistance genes, thus fewer variants). However, CARD remained on top for the most number of predictions (both acquired and mutation- based resistance genes with multiple variants). A similar study, related with consolidating and exploring antibiotic resistance genes data resources, has also proved that using whole-genome sequences and metagenomic sequencing data, CARD not only performed better than the rest of the databases used (ResFinder, Antibiotic Resistance Genes Database; ARDB, and Comprehensive β-lactamase Molecular Annotation Resource; CBMAR), but it also reported the most number of correct predictions (Xavier et al., 2016). On the other hand, cyclical annotation graphs like the ARO (such as used by CARD) can result in falsely inflated counts for the conflation of assignments in sequence classification (Lakin et al., 2017)

Apart from MEGARes, no single resource currently enables structured, comprehensive and statistically appropriate analysis of metagenomics data for all types of antimicrobial compounds, including biocides and metals (McArthur and Tsang, 2017). However, the main focus of MEGARes is not to be an alternative choice for CARD and ResFinder users, but be available as a foundation for the development of resistome-centered analytical methods, such as sequence classifiers and hierarchical statistical models. However, MEGARes focuses on previously published sequences, rather than newly discovered variants (Lakin et al., 2017). In our study, MEGARes was in agreement with CARD in the detection of all the genes with the exception of CTX-M and *dfr*. Another study compared ARG-miner and MEGARes with CARD, and these databases didn't accurately detect all mutants that were detected by CARD, suggesting CARD is better suited for detecting chromosomal mutations compared to other available databases. Moreover, due to different nomenclature strategies, some discrepancies have been noted in ARG annotation whith MEGARes as compared to CARD, where just the name of the gene appeared and not the variant number (Lal Gupta et al., 2020).

Although AMRFinder did not detect any β-lactamase resistance gene, and detected only three non-β-lactamase antibiotic resistance genes (aminoglycoside, tetracycline, chloramphenicol), it was in agreement with CARD, ResFinder, and MEGARes. In this regard, it is important to

note that AMRFinder does not attempt to assert the effects of detected proteins found to have a clinical resistance phenotype too, as the factors responsible for the expression of those proteins are outside the current coverage of AMRFinder (Feldgarden et al., 2019b). Moreover, gene symbol output disagreement (8.8%) was also noted between 2017 version of ResFinder and AMRFinder. Since HMM and BLAST-based approaches are used by AMRFinder and ResFinder, both approaches need to be synchronized to minimize inconsistent outputs due to algorithmic differences (Feldgarden et al., 2019a). However, HMM approaches may have poor specificity, producing high number of false positive predictions and sometimes may not be able to distinguish between ARGs with closely related functions (Lal Gupta et al., 2020).

5.4 Phenotypic-Genotypic Relationship

In our study, *E. coli* isolates showed 56% cumulative resistance to three β-lactam antibiotics (ampicillin and cefotaxime, ceftazidime), followed by trimethoprim (22%), and fluoroquinolone (14%), while resistance to aminoglycosides, and chloramphenicol was low (7%, 2% respectively) (
Figure 7). This β-lactam antibiotic resistance phenotype is complemented by isolates having β-lactamase genes; *ampC* (93%), TEM (38%), SHV (5%), OXA (7%),VIM (2%), and CTX-M (3%) in *E. coli* (Table 5 and *Figure 12*). Low percentage (9%) of *E. coli* isolates being resistant to gentamicin in our study is in contrast to a study where 69% (total 44 isolates) of ESBL producing *E. coli* bacteria were resistant to gentamicin (Ojdana et al., 2018). However, it complies with the finding in Norway where the gentamicin non- susceptibility among the *E. coli* (109 isolates) was 4% in the isolates collected in 2009 (Lindemann et al., 2012). One explanation of low resistance to gentamicin in our study can be the clinical use of aminoglycosides below 10% of sales (total 5,450 kg) of antibiotics since 2016 in Norway (NORM/NORM-VET, 2016). Trimethoprim resistance rate in *E. coli* isolates (22%) in our study was comparable to 14.1% in ECO.SENS study (Kahlmeter, 2003), and 18-26% *E. coli* isolates from human clinical samples in Lithuania (Šeputienė et al., 2010).

In *K. pneumoniae*, β-lactam antibiotic resistance was mainly towards ampicillin (98%), ceftazidime (13%), and cefotaxime (11%), as compared to non β-lactam antibiotics; trimethoprim (16%), tigecycline (11%), ciprofloxacin and gentamicin (9% each). *K. pneumoniae* isolates being resistant to gentamicin (9%) in this study are similar to the 5% (11 isolates) in the west Norwegian *K. pneumoniae* isolates (Ambaye et al., 1997).

High percentage of phenotypic ampicillin resistance was shown in our study (*E. coli* 41% and *K. pneumoniae* 91%). Similar results were found in a study, where *E. coli* isolated from outpatient population (urine samples) in Bosnia and Herzegovina showed the highest antimicrobial resistance to ampicillin (82.79%) (Vranic and Uzunovic, 2016). Moreover, 100% ampicillin resistant isolates of both *E. coli* and *K. pneumoniae* (urine) were obtained in another study from India, where *E. coli* and *K. pneumoniae* comprised 60% and 15% of total 20 identified microorganisms (Agarwal et al., 2015). One reason of such high ampicillin resistance could be the high rate of penicillin (both β-lactamase sensitive and extended-spectrum) prescription as human medicine in Norway (NORM/NORM-VET, 2016).

Multi Drug Resistance (MDR) means the ability of the microorganism to resist at least one drug from three different antimicrobial classes (Magiorakos et al., 2012). In our study, *E. coli* was found to have MDR in 5 isolates (9%). The resistance patterns was towards three (isolate 105, 124, 138, 112) and four (isolate 121) antibiotics. This is in contrast to many other studies, for example a higher percent (33.2%) of *E. coli* isolates were reported to be MDR in another study in North-western Libya (Abujnah et al., 2015).

In ten isolates (24%) of *K. pneumoniae*, MDR phenotype was expressed with the resistance to three (215, 217, 225), four (220, 226), five (216), and six (211, 212, 213, 240) antibiotics respectively. In a similar study, higher percentage of MDR resistant *K. pneumoniae* (46% of 116 isolates) was observed (Moini et al., 2015). Moreover, presence of complimentary β-lactamase resistance genes (*ampC*, SHV, TEM, OXA, LEN, OKP, CTXM) supports the high β-lactam antibiotic resistance found in *K. pneumoniae*, where SHV and TEM together were represented by 86% of the isolates, making them the dominant β-lactamases antibiotic resistance genes in *K. pneumoniae* (Table 5 and*Figure 12*).

Like *E. coli* and *K. pneumoniae*, *A. baumannii* also displayed MDR; ciprofloxacin, gentamicin, and meropenem in 10 isolates (91%). Similar finding was observed in a study where 78 (80%) out of 97 *A. baumannii* clinical isolates were resistant to three or more classes of antimicrobial compounds, and thus considered MDR (Taitt et al., 2014). The most common mechanism responsible for carbapenem resistance in *A. baumannii* is mediated by the acquired oxacillinases OXA-23-like, OXA-24-like, OXA-58-like, OXA-143-like and OXA-235. Metallo-βlactamases, such as VIM have only rarely been found in *A. baumannii* (Krizova et al., 2012). In this regard, we found OXA-23 in nine *A. baumanii* isolates (302, 303, 304, 306,

307, 308, 309, 310, 311), OXA-24 in one (305), OXA-58 in one (301), while OXA-235 along with VIM enzyme was not found in any of the isolates.

In our study, isolates with resistance phenotypes with no AMR genes were identified, as were the isolates with susceptible phenotypes that carried resistance genes (Table 5). The resistance phenotypes without the corresponding genes included ciprofloxacin (*E. coli:* 8, *K. pneumoniae:* 2), trimethoprim (*E. coli*: 4, *K. pneumoniae:* 4), chloramphenicol (*E. coli*: 2, *K. pneumoniae: 5*), gentamicin (*E. coli*: 2), and ciprofloxacin in all eleven *A. baumannii* isolates. So was the case with another study, where every antimicrobial outcome had some isolates with a resistant phenotype, but no genetic explanation (Rosengren et al., 2009). On the other hand, the susceptible phenotypes with AMR genes in *E. coli* included β-lactamase (*ampC* in 32 isolates, TEM in 4 isolates, and OXA in 1 isolate), gentamicin (*aac/aph* in10 isolates), tetracycline (*tet* in 9 isolates), fluoroquinolone (*qnr* in 2 isolates), and phenicol (*cat* in 2 isolates).

In *K. pneumoniae,* the susceptible phenotypes with AMR genes were aminoglycosides (*aac/aph* in 5 isolates), fluoroquinolone and βlactam (*qnr* and LEN in one isolate each), and tetracycline (Nordmann et al.) and phenicol (*cat/Cml/floR*) in 2 isolates each.

In case of *A. baumannii,* tetracycline, aminoglycosides and phenicol related susceptibilities were found with AMR genes (*tet*, *aph*, and *cat/Cml/floR* genes in six, one and three isolates).

Another study, where *E. coli* genotypic resistance was compared with phenotype (Do Nascimento et al., 2017), reported discrepancies mainly referring to phenotypically-susceptible isolates harbouring a resistance gene. This evident contradiction of susceptible isolates carrying resistance genes can be because of unexpressed resistance genes, if they are far from or associated with a weak promoter in an integron. Similarly, the free gene cassettes, which are not a part of an integron are silent because the integron's promoter is required for expression (Carattoli, 2001). Alternatively, isolates could be wrongly represented as susceptible, if the MIC breakpoint is higher than the resistance communicated by the gene (Boerlin et al., 2005).

Three different enzymes, CTX-M-15, -16, -19 and, recently, CTX-M-27 have been reported to be linked with ceftazidime hydrolysis (Bonnet et al., 2001). This is in agreement with our

finding where three *K. pneumoniae* isolates (213, 216, 220) had CTX-M gene as well as phenotypic resistance to ceftazidime.

ESBL harbouring *K. pneumoniae* isolates have been found to be resistant to other antibiotics, especially, fluroquinolones (Lautenbach et al., 2001) In a study done by Tumbarello et al., in Italy, 32% of ESBL producing isolates of *K. pneumoniae* were resistant to ciprofloxacin (Tumbarello et al., 2006). Considering our bacterial samples as pathogenic, at all the occasions where ciprofloxacin resistance was present (Table 5), it co-existed with phenotypic β-lactam antibiotics resistance (*E. coli*: 4, *K. pneumoniae*: 4, *A. baumannii*:11 isolates).

Another aspect to consider is the co-existence of carbapenem and aminoglycoside resistance phenotype in 91% of *A. baumannii* isolates in our study (
Figure 7), which is confirmed by another study done on multidrug-resistant clinical isolates of *A. baumannii* from Krakow, Poland, where genes conferring resistance to carbapenems and aminoglycosides coexisted in 44.3% (61 isolates) clinical strains of *A. baumannii* (Nowak et al., 2014).

Seventeen SHV variants are exclusively found in clinical *K. pneumoniae*: blaSHV-6, blaSHV-13, blaSHV-16, blaSHV-18, blaSHV-23, blaSHV-45, blaSHV-64, blaSHV-66, blaSHV-86, blaSHV-90, blaSHV-91, blaSHV-98, blaSHV-99, blaSHV-100, blaSHV−104, blaSHV-105, and blaSHV-134 (Liakopoulos et al., 2016). These variants are mostly associated with plasmids. All these variants were found in our data also (supplementary file 1). A variant blaSHV-27 has been detected on different plasmids in *E. coli* and *K. pneumoniae (Corkill et al., 2001)*. In our study, four *E. coli* isolates, and 26 *K. pneumoniae* isolates had SHV-27 along with plasmids except for two isolates (212, 219). As SHV-27 confers resistance to cefotaxime, ceftazidime and aztreonam (Corkill et al., 2001), six *K. pneumoniae* isolates (211, 212, 213, 216, 220, 240) from our study had SHV-27 and exhibited resistance to either cefotaxime or ceftazidime, or both simultaneously (Table 5). SHV-12 has been reported as the most prevalent enzyme within SHV family all over the world in *K. pneumoniae* and in *E. coli* from community patients (Valverde et al., 2004). In agreement to the statement above, we found SHV-12 in isolates of of *E. coli* (2) and *K. pneumoniae* (29) respectively (Supplementary file 1).

TEM-1 is the most commonly encountered β-lactamase in gram-negative bacteria. Up to 90% of ampicillin resistance in *E. coli* is due to the production of TEM-1 (Livermore, 1995). Regarding TEM-1 in our study, there isolates of *E. coli* (21), *K. pneumoniae* (9), and *A. baumannii* (https://www.454.com/.) with the variant (supplementary file 1). TEM-3 and TEM-4 also seem to be widespread, and to be associated with different clones of *K. pneumoniae* in ICUs (Asensio et al., 2000). TEM-52 is also widespread in Europe, and is associated with *E. coli* from urinary tract infections (Caccamo et al., 2006). It was also seen in our study that TEM-3 and TEM-52 was represented by isolates of *E. coli* (20), *K. pneumoniae* (8), and *A. baumannii* (https://www.454.com/.) respectively.

CTX-M-15 prevalence has been increasing all over Europe (Livermore et al., 2007). Moreover, we found CTX-M-15 only in four *K. pneumoniae* isolates (215, 216, 220, 213) (Table 5). International spread of blaCTX-M-15 seems to be linked with IncFII plasmids (Lavollay et al., 2006), and we observed the presence of both of them together on two occasions in our study (213, 216) although not on common nodes. Similarly, blaCTX-M-32, which has association with IncN plasmids (Cottell et al., 2013), was present in three isolates (213, 216, 220), but its presence with IncN plasmids was not seen. Another variant blaCTX-M-9, which is associated with IncHI2 (Novais et al., 2006), was present in two isolates (146, 147) along IncHI2 plasmid. blaCTX-M-1 is the most often identified gene on IncI plasmid (Rozwandowicz et al., 2018), but it was not found in any of the isolates in our study.

5.5 Plasmid mediated Resistance

Overall, the presence of the different types of plasmid Col and IncFIB in the isolates which expressed phenotypic resistance to β-lactam antibiotics (33 and 39 isolates) (ampicillin, cefotaxime, ceftazidime, meropenem) and trimethoprim (14 and 15 isolates) was observed (Table 5). Moreover, IncFII (different types) were also found in the isolates that expressed phenotypic resistance to β- lactam antibiotics (30 isolates).

The most frequently described resistance genes on IncF plasmids are related to carbapenemases, aminoglycoside and plasmid-mediated quinolone resistance (PMQR) genes (Rozwandowicz et al., 2018). In our study, in all the isolates, where *aac/aph* and *qnr* resistance genes were present in *E.coli* and *K. pneumoniae*, IncF (different types) plasmids were also

present, while regarding carbapenemases, it was not observed in *A. baumannii* isolates, which could be due to the low instances of carbapenamases in Norway.

In silico detection using PlasmidFinder and ResFinder on WGS data explores the opportunity to associate replicons with antimicrobial resistance genes on the same DNA fragment because the exact position of genes and the plasmids is available in these tools. However, uncertainty prevails in deciding whether genetic elements, which have been identified on different contigs, are located on the same plasmid too (Carattoli et al., 2014). Our plasmid detection results had only a few instances (*E.coli:* 11, *K. pneumoniae*: 1), where the resistance genes and the plasmids actually shared the same contig (Table 6 and Figure 15). Moreover, the plasmids detected by PLSDB and *Brooks et.al* were numerous in number and types (80-90 thousand hits per isolate before filtering). However, regardless of employed algorithm for plasmid detection and identification, there were multiple plasmids found on the same contig that made the selection of one confirmed plasmid difficult. This problem can be explained with the fact that if the plasmids are sequenced along the rest of the genome, they can rarely be completely assembled from Illumina reads, making it difficult to separate the contigs of the plasmids from the rest of the genome (Page et al., 2018a). A helpful thing in this regard was to use mlPlasmids to confirm the plasmid containing contigs with plasmid databases used (supplementary file 7).

Usually, if the same contig or set of contigs match several plasmids, we can select the plasmid that matches over the greatest length of the plasmid with the highest sequence identity (Hall, 2018). In our case however, the several plasmids with the same assigned contig had identical matched plasmid length and identity, which lowered the certainty in identifying plasmids. Otherwise, the resistance genes can actually be present on the chromosome and not on the plasmid. These two factors contributed to low number of isolates with plasmid mediated resistance.

WGS based *in silico* analysis of resistance genes and their plasmid context is performed in a unified way on a large number of isolates (Carrër et al., 2010). However, Plasmid detection from WGS can be challenging to understand considering the presence of multiple plasmids or a single plasmid containing multiple replicons (Johnson et al., 2007).

6. Conclusion

In conclusion, detected ESBLs and their resistance specific variants confirm the importance of selected pathogens in the spread of antimicrobial resistance in Norway. TEM, SHV, and OXA remain the most dominant ESBLs in our study. A total of 63 isolates (57%) had concordance between antibiotic resistance phenotype and corresponding resistance genes (to keep the analysis brief, only CARD detected genes were taken), which explains how well curated databases ensure a high concordance between phenotype and genotype resistance. CARD, ResFinder, and MEGARes performance in resistance gene detection was in agreement with each other, and thus reliable. There is however, a significant need for standardization of pipelines and databases as well as phenotypic predictions based on the genomic data.

IncFIB, IncFII, and Col remain the dominant types of incompatibility groups of plasmids in *Enterobacteriaceae* isolates. PlasmidFinder was not only more accurate in plasmid detection than PLSDB and *Brooks et al* database, but it also predicted most number of plasmids that were hosting antibiotic resistance genes. This happened due to the noise created by huge number of hits and the presence of multiple plasmids on the same contig in PLSDB and *Brooks et al* database. Since such contigs were excluded from the plasmid detection analysis, it impacted their plasmid mediated resistance analysis.

Important conclusion is that Norway has low level of resistance based on the AMR genes and AST data, which is good for the health care in Norway. This could also be attributed to the regulations and guidelines for antibiotic use in Norway.

The wide adoption of WGS has proved to be useful in describing AMR genes and plasmids in priority pathogens; *E.coli, K. pneumoniae*, and *A. baumannii* in Noway.

7. Future Prospects

In future studies, choice of hybrid assemblies with the inclusion of non-Norwegian isolates can be explored to detect the plasmids with high, comparable certainty in less fragmented assemblies. The predicted resistance determinants and the related risks for human health using WGS technology should play an important part in future risk assessment policies to combat AMR spread in Norway.

This study can be extended to investigate the resistance mechanisms used by antimicrobial determinants to identify the correct antibiotic treatment. Moreover, molecular extraction of plasmids followed by sequencing can also be tried in order to see if the *in silico* plasmid prediction improves

8. References

Abujnah, A.A., A. Zorgani, M.A. Sabri, H. El-Mohammady, R.A. Khalek, and K.S. Ghenghesh. 2015. Multidrug resistance and extended-spectrum β-lactamases genes among0 Escherichia coli from patients with urinary tract infections in Northwestern Libya. *Libyan Journal of Medicine*. 10.

Agarwal, M., L. Gupta, and K. Bala. 2015. Identification and antibiotic susceptibility testing of urinary isolates in a tertiary care hospital. *Der Pharmacia Lettre*. 7:270-279.

Alcock, B.P., A.R. Raphenya, T.T.Y. Lau, K.K. Tsang, M. Bouchard, A. Edalatmand, W. Huynh, A.V. Nguyen, A.A. Cheng, S. Liu, S.Y. Min, A. Miroshnichenko, H.K. Tran, R.E. Werfalli, J.A. Nasir, M. Oloni, D.J. Speicher, A. Florescu, B. Singh, M. Faltyn, A. Hernandez-Koutoucheva, A.N. Sharma, E. Bordeleau, A.C. Pawlowski, H.L. Zubyk, D. Dooley, E. Griffiths, F. Maguire, G.L. Winsor, R.G. Beiko, F.S.L. Brinkman, W.W.L. Hsiao, G.V. Domselaar, and A.G. McArthur. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 48:D517-d525.

Ali, T., I. Ali, N.A. Khan, B. Han, and J. Gao. 2018. The growing genetic and functional diversity of extended spectrum beta-lactamases. *BioMed research international*. 2018.

Allocati, N., M. Masulli, M.F. Alexeyev, and C. Di Ilio. 2013. Escherichia coli in Europe: an overview. *International journal of environmental research and public health*. 10:6235-6254.

Ambaye, A., P.C. Kohner, P.C. Wollan, K.L. Roberts, G.D. Roberts, and F. Cockerill. 1997. Comparison of agar dilution, broth microdilution, disk diffusion, E-test, and BACTEC radiometric methods for antimicrobial susceptibility testing of clinical isolates of the Nocardia asteroides complex. *Journal of clinical microbiology*. 35:847-852.

Amyes, S.G. 2000. The rise in bacterial resistance is partly because there have been no new classes of antibiotics since the 1960s. *BMJ*. 320:199-200.

Andrews, S. 2010. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. *URL:* https://www/. *bioinformatics. babraham. ac. uk/projects/fastqc/(accessed 06.12. 2018)*.

Angermueller, C., T. Pärnamaa, L. Parts, and O. Stegle. 2016. Deep learning for computational biology. *Molecular systems biology*. 12:878.

Antipov, D., N. Hartwick, M. Shen, M. Raiko, A. Lapidus, and P.A. Pevzner. 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *bioRxiv*:048942.

Arango-Argoty, G., E. Garner, A. Pruden, L.S. Heath, P. Vikesland, and L. Zhang. 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*. 6:1-15.

Arredondo-Alonso, S., M.R.C. Rogers, J.C. Braat, T.D. Verschuuren, J. Top, J. Corander, R.J.L. Willems, and A.C. Schurch. 2018. mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. *Microb Genom*. 4.

Arredondo-Alonso, S., R.J. Willems, W. van Schaik, and A.C. Schürch. 2017. On the (im) possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microbial genomics*. 3.

Asensio, A., A. Oliver, P. González-Diego, F. Baquero, J.C. Perez-Diaz, P. Ros, J. Cobo, M. Palacios, D. Lasheras, and R. Cantón. 2000. Outbreak of a multiresistant Klebsiella

pneumoniae strain in an intensive care unit: antibiotic use as risk factor for colonization and infection. *Clinical Infectious Diseases*. 30:55-60.

Bagley, S.T. 1985. Habitat association of Klebsiella species. *Infection Control & Hospital Epidemiology*. 6:52-58.

Bankevich, A., S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, and A.D. Prjibelski. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*. 19:455-477.

Bengtsson-Palme, J., D.J. Larsson, and E. Kristiansson. 2017. Using metagenomics to investigate human and environmental resistomes. *Journal of Antimicrobial Chemotherapy*. 72:2690-2703.

Bentley, D.R., S. Balasubramanian, H.P. Swerdlow, G.P. Smith, J. Milton, C.G. Brown, K.P. Hall, D.J. Evers, C.L. Barnes, and H.R. Bignell. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *nature*. 456:53-59.

Biosciences, P. 2014. Data Release: 54x Long-Read Coverage for PacBio-only De Novo Human Genome Assembly.

Bissonnette, L., S. Champetier, J.P. Buisson, and P.H. Roy. 1991. Characterization of the nonenzymatic chloramphenicol resistance (cmlA) gene of the In4 integron of Tn1696: similarity of the product to transmembrane transport proteins. *Journal of Bacteriology*. 173:4493-4502.

Boerlin, P., R. Travis, C.L. Gyles, R. Reid-Smith, N.J.H. Lim, V. Nicholson, S.A. McEwen, R. Friendship, and M. Archambault. 2005. Antimicrobial resistance and virulence genes of Escherichia coli isolates from swine in Ontario. *Applied and environmental microbiology*. 71:6753-6761.

Bonnet, R., C. Dutour, J. Sampaio, C. Chanal, D. Sirot, R. Labia, C. De Champs, and J. Sirot. 2001. Novel cefotaximase (CTX-M-16) with increased catalytic efficiency due to substitution Asp-240→ Gly. *Antimicrobial agents and chemotherapy*. 45:2269-2275.

Bradford, P.A. 2001. Extended-spectrum β-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clinical microbiology reviews*. 14:933-951.

Breitwieser, F.P., J. Lu, and S.L. Salzberg. 2019. A review of methods and databases for metagenomic classification and assembly. *Briefings in bioinformatics*. 20:1125-1136.

Brooks, L., M. Kaze, and M. Sistrom. 2019. A Curated, Comprehensive Database of Plasmid Sequences. *Microbiol Resour Announc*. 8:e01325-01318.

Bush, K., and P.A. Bradford. 2016. β-Lactams and β-lactamase inhibitors: an overview. *Cold Spring Harbor perspectives in medicine*. 6:a025247.

Bush, K., G.A. Jacoby, and A.A. Medeiros. 1995. A functional classification scheme for beta-lactamases and its correlation with molecular structure. *Antimicrobial agents and chemotherapy*. 39:1211.

Butler, D., and S. Grimme. 2010. Human genome at ten: science after the sequence. *Nature*. 465:1000-1001.

Caccamo, M., M. Perilli, G. Celenza, G. Bonfiglio, G. Tempera, and G. Amicosante. 2006. Occurrence of extended spectrum β-lactamases among isolates of Enterobacteriaceae from urinary tract infections in southern Italy. *Microbial Drug Resistance*. 12:257-264.

Carattoli, A. 2001. Importance of integrons in the diffusion of resistance. *Veterinary research*. 32:243-259.

Carattoli, A. 2011. Plasmids in Gram negatives: molecular typing of resistance plasmids. *International Journal of Medical Microbiology*. 301:654-658.

Carattoli, A. 2013. Plasmids and the spread of resistance. *International Journal of Medical Microbiology*. 303:298-304.

Carattoli, A., E. Zankari, A. Garcia-Fernandez, M. Voldby Larsen, O. Lund, L. Villa, F. Moller Aarestrup, and H. Hasman. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother*. 58:3895-3903.

Carrër, A., L. Poirel, M. Yilmaz, Ö.A. Akan, C. Feriha, G. Cuzon, G. Matar, P. Honderlick, and P. Nordmann. 2010. Spread of OXA-48-encoding plasmid in Turkey and beyond. *Antimicrobial agents and chemotherapy*. 54:1369-1373.

Chen, L., J. Yang, J. Yu, Z. Yao, L. Sun, Y. Shen, and Q. Jin. 2005. VFDB: a reference database for bacterial virulence factors. *Nucleic acids research*. 33:D325-D328.

Chopra, I., and M. Roberts. 2001. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and molecular biology reviews*. 65:232-260.

Clausen, P.T., E. Zankari, F.M. Aarestrup, and O. Lund. 2016. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *Journal of Antimicrobial Chemotherapy*. 71:2484-2488.

Corkill, J.E., L.E. Cuevas, R.Q. Gurgel, J. Greensill, and C.A. Hart. 2001. SHV-27, a novel cefotaxime-hydrolysing β-lactamase, identified in Klebsiella pneumoniae isolates from a Brazilian hospital. *Journal of Antimicrobial Chemotherapy*. 47:463-465.

Cottell, J.L., N. Kanwar, L. Castillo-Courtade, G. Chalmers, H.M. Scott, B. Norby, G.H. Loneragan, and P. Boerlin. 2013. bla_{CTX-M-32} on an IncN Plasmid in Escherichia coli from Beef Cattle in the United States. *Antimicrobial Agents and Chemotherapy*. 57:1096.

Couturier, M., F. Bex, P.L. Bergquist, and W.K. Maas. 1988. Identification and classification of bacterial plasmids. *Microbiological reviews*. 52:375.

D'Costa, V.M., C.E. King, L. Kalan, M. Morar, W.W. Sung, C. Schwarz, D. Froese, G. Zazula, F. Calmels, and R. Debruyne. 2011. Antibiotic resistance is ancient. *Nature*. 477:457-461.

Dao, T.T., D. Liebenthal, T.K. Tran, B.N.T. Vu, D.N.T. Nguyen, H.K.T. Tran, C.K.T. Nguyen, H.L.T. Vu, A. Fox, and P. Horby. 2014. Klebsiella pneumoniae oropharyngeal carriage in rural and urban Vietnam and the effect of alcohol consumption. *PLoS One*. 9:e91999.

Datta, N. 1977. Classification of plasmids as an aid to understanding their epidemiology and evolution. *Journal of Antimicrobial Chemotherapy*. 3:19-23.

Davis, M.A., K.N. Baker, L.H. Orfe, D.H. Shah, T.E. Besser, and D.R. Call. 2010. Discovery of a gene conferring multiple-aminoglycoside resistance in Escherichia coli. *Antimicrobial agents and chemotherapy*. 54:2666-2669.

de Man, T.J., J.D. Lutgring, D.R. Lonsway, K.F. Anderson, J.A. Kiehlbauch, L. Chen, M.S. Walters, M. Sjölund-Karlsson, J.K. Rasheed, and A. Kallen. 2018. Genomic analysis of a pan-resistant isolate of Klebsiella pneumoniae, United States 2016. *MBio*. 9.

De Toro, M., M. Pilar Garcillán-Barcia, and F. De La Cruz. 2015. Plasmid diversity and adaptation analyzed by massive sequencing of Escherichia coli plasmids. *Plasmids: Biology and Impact in Biotechnology and Discovery*:219-235.

Deamer, D., M. Akeson, and D. Branton. 2016. Three decades of nanopore sequencing. *Nature biotechnology*. 34:518-524.

Dib, J.R., M. Wagenknecht, M.E. Farías, and F. Meinhardt. 2015. Strategies and approaches in plasmidome studies—uncovering plasmid diversity disregarding of linear elements? *Frontiers in Microbiology*. 6:463.

Do Nascimento, V., M.R. Day, M. Doumith, K.L. Hopkins, N. Woodford, G. Godbole, and C. Jenkins. 2017. Comparison of phenotypic and WGS-derived antimicrobial resistance profiles of enteroaggregative Escherichia coli isolated from cases of diarrhoeal disease in England, 2015–16. *Journal of Antimicrobial Chemotherapy*. 72:3288-3297.

Ekblom, R., L. Smeds, and H. Ellegren. 2014. Patterns of sequencing coverage bias revealed by ultra-deep sequencing of vertebrate mitochondria. *BMC genomics*. 15:467.

Elbadawi, H.S., K.M. Elhag, E. Mahgoub, H.N. Altayb, and M.M.A. Hamid. 2019. Antimicrobial resistance surveillance among gram-negative bacterial isolates from patients in hospitals in Khartoum State, Sudan. *F1000Research*. 8:156.

Ellington, M.J., O. Ekelund, F.M. Aarestrup, R. Canton, M. Doumith, C. Giske, H. Grundman, H. Hasman, M.T.G. Holden, K.L. Hopkins, J. Iredell, G. Kahlmeter, C.U. Köser, A. MacGowan, D. Mevius, M. Mulvey, T. Naas, T. Peto, J.M. Rolain, Ø. Samuelsen, and N. Woodford. 2017. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical Microbiology and Infection*. 23:2-22.

EUCAST, T.E.C.o.A.S.T.-. 2020.

Ewers, C., A. Bethe, T. Semmler, S. Guenther, and L. Wieler. 2012. Extended-spectrum β-lactamase-producing and AmpC-producing Escherichia coli from livestock and companion animals, and their putative impact on public health: a global perspective. *Clinical Microbiology and Infection*. 18:646-655.

Feldgarden, M., V. Brover, D.H. Haft, A.B. Prasad, D.J. Slotta, I. Tolstoy, G.H. Tyson, S. Zhao, C.-H. Hsu, and P.F. McDermott. 2019a. Using the NCBI AMRFinder tool to determine antimicrobial resistance genotype-phenotype correlations within a collection of NARMS isolates. *BioRxiv*:550707.

Feldgarden, M., V. Brover, D.H. Haft, A.B. Prasad, D.J. Slotta, I. Tolstoy, G.H. Tyson, S. Zhao, C.-H. Hsu, and P.F. McDermott. 2019b. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrobial agents and chemotherapy*. 63:e00483-00419.

Fluit, A.C., M.R. Visser, and F.-J. Schmitz. 2001. Molecular detection of antimicrobial resistance. *Clinical microbiology reviews*. 14:836-871.

Garneau-Tsodikova, S., and K.J. Labby. 2016. Mechanisms of resistance to aminoglycoside antibiotics: overview and perspectives. *Medchemcomm*. 7:11-27.

Geisel, R., F.J. Schmitz, L. Thomas, G. Berns, O. Zetsche, B. Ulrich, A.C. Fluit, H. Labischinsky, and W. Witte. 1999. Emergence of heterogeneous intermediate vancomycin resistance in Staphylococcus aureus isolates in the Düsseldorf area. *Journal of Antimicrobial Chemotherapy*. 43:846-848.

Giammanco, A., C. Calà, T. Fasciana, and M.J. Dowzicky. 2017. Global assessment of the activity of tigecycline against multidrug-resistant Gram-negative pathogens

between 2004 and 2014 as part of the Tigecycline Evaluation and Surveillance Trial. *Msphere*. 2.

Gibson, M.K., K.J. Forsberg, and G. Dantas. 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *The ISME journal*. 9:207-216.

Goffin, C., and J.-M. Ghuysen. 1998. Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. *Microbiology and molecular biology reviews*. 62:1079-1093.

Gordon, N., J. Price, K. Cole, R. Everitt, M. Morgan, J. Finney, A. Kearns, B. Pichon, B. Young, and D. Wilson. 2014. Prediction of Staphylococcus aureus antimicrobial resistance by whole-genome sequencing. *Journal of clinical microbiology*. 52:1182-1191.

Gupta, S.K., B.R. Padmanabhan, S.M. Diene, R. Lopez-Rojas, M. Kempf, L. Landraud, and J.-M. Rolain. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial agents and chemotherapy*. 58:212-220.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 29:1072-1075.

Hall, B.G. 2018. Identifying Plasmids in Bacterial Genome Assemblies. *bioRxiv*:332049.

Harbottle, H., S. Thakur, S. Zhao, and D. White. 2006. Genetics of antimicrobial resistance. *Animal biotechnology*. 17:111-124.

Harding, C.M., S.W. Hennon, and M.F. Feldman. 2018. Uncovering the mechanisms of Acinetobacter baumannii virulence. *Nature Reviews Microbiology*. 16:91.

Haug, J.B., D. Berild, M. Walberg, and Å. Reikvam. 2011. Increased antibiotic use in Norwegian hospitals despite a low antibiotic resistance rate. *Journal of antimicrobial chemotherapy*. 66:2643-2646.

He, T., R. Wang, D. Liu, T.R. Walsh, R. Zhang, Y. Lv, Y. Ke, Q. Ji, R. Wei, Z. Liu, Y. Shen, G. Wang, L. Sun, L. Lei, Z. Lv, Y. Li, M. Pang, L. Wang, Q. Sun, Y. Fu, H. Song, Y. Hao, Z. Shen, S. Wang, G. Chen, C. Wu, J. Shen, and Y. Wang. 2019. Emergence of plasmid-mediated high-level tigecycline resistance genes in animals and humans. *Nature Microbiology*. 4:1450-1456.

Henson, J., G. Tischler, and Z. Ning. 2012. Next-generation sequencing and large genome assemblies. *Pharmacogenomics*. 13:901-915.

Hermann, T. 2007. Aminoglycoside antibiotics: old drugs and new therapeutic approaches. *Cellular and molecular life sciences*. 64:1841-1852.

Hiraga, S.-i., T. Sugiyama, and T. Itoh. 1994. Comparative analysis of the replicon regions of eleven ColE2-related plasmids. *Journal of bacteriology*. 176:7233-7243.

Holt, K.E., H. Wertheim, R.N. Zadoks, S. Baker, C.A. Whitehouse, D. Dance, A. Jenney, T.R. Connor, L.Y. Hsu, J. Severin, S. Brisse, H. Cao, J. Wilksch, C. Gorrie, M.B. Schultz, D.J. Edwards, K.V. Nguyen, T.V. Nguyen, T.T. Dao, M. Mensink, V.L. Minh, N.T.K. Nhu, C. Schultsz, K. Kuntaman, P.N. Newton, C.E. Moore, R.A. Strugnell, and N.R. Thomson. 2015. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. *Proceedings of the National Academy of Sciences*. 112:E3574-E3581.

Hrabák, J., V. Študentová, R. Walková, H. Žemličková, V. Jakubů, E. Chudáčková, M. Gniadkowski, Y. Pfeifer, J.D. Perry, and K. Wilkinson. 2012. Detection of NDM-1, VIM-1, KPC, OXA-48, and OXA-162 carbapenemases by matrix-assisted laser

desorption ionization–time of flight mass spectrometry. *Journal of clinical microbiology*. 50:2441-2443.

https://www.454.com/. 454 Life Sciences.

https://www.appliedbiosystems/. SOLiD System from Applied Biosystems.

https://www.biobam.com/omicsbox. March 3, 2019. OmicsBox – Bioinformatics Made Easy, BioBam Bioinformatics.

https://www.illumina.com/. Illumina Sequencing Technology.

Jacoby, G.A., and L.S. Munoz-Price. 2005. The new β-lactamases. *New England Journal of Medicine*. 352:380-391.

Jennings, L.J., M.E. Arcila, C. Corless, S. Kamel-Reid, I.M. Lubin, J. Pfeifer, R.L. Temple-Smolkin, K.V. Voelkerding, and M.N. Nikiforova. 2017. Guidelines for validation of next-generation sequencing–based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. *The Journal of molecular diagnostics*. 19:341-365.

Jia, B., A.R. Raphenya, B. Alcock, N. Waglechner, P. Guo, K.K. Tsang, B.A. Lago, B.M. Dave, S. Pereira, and A.N. Sharma. 2016. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic acids research*:gkw1004.

Johnson, T.J., Y.M. Wannemuehler, S.J. Johnson, C.M. Logue, D.G. White, C. Doetkott, and L.K. Nolan. 2007. Plasmid replicon typing of commensal and pathogenic Escherichia coli isolates. *Applied and environmental microbiology*. 73:1976-1983.

Jørgensen, S.B., Ø. Samuelsen, A. Sundsfjord, S.A. Bhatti, I. Jørgensen, T. Sivapathasundaram, and T.M. Leegaard. 2014a. High prevalence of faecal carriage of ESBL-producing Enterobacteriaceae in Norwegian patients with gastroenteritis. *Scandinavian journal of infectious diseases*. 46:462-465.

Jørgensen, T.S., Z. Xu, M.A. Hansen, S.J. Sørensen, and L.H. Hansen. 2014b. Hundreds of circular novel plasmids and DNA elements identified in a rat cecum metamobilome. *PloS one*. 9:e87924.

Kahlmeter, G. 2003. An international survey of the antimicrobial susceptibility of pathogens from uncomplicated urinary tract infections: the ECO·SENS Project. *Journal of Antimicrobial Chemotherapy*. 51:69-76.

Kaper, J.B., J.P. Nataro, and H.L. Mobley. 2004. Pathogenic escherichia coli. *Nature reviews microbiology*. 2:123-140.

Khan, Z.A., M.F. Siddiqui, and S. Park. 2019. Current and Emerging Methods of Antibiotic Susceptibility Testing. *Diagnostics (Basel)*. 9:49.

Kim, D., K.S. Baik, M.S. Kim, S.C. Park, S.S. Kim, M.S. Rhee, Y.S. Kwak, and C.N. Seong. 2008. Acinetobacter soli sp. nov., isolated from forest soil. *The Journal of Microbiology*. 46:396-401.

Kim, E.S., J.-Y. Jeong, J.-B. Jun, S.-H. Choi, S.-O. Lee, M.-N. Kim, J.H. Woo, and Y.S. Kim. 2009. Prevalence of aac(6')-Ib-cr Encoding a Ciprofloxacin-Modifying Enzyme among Enterobacteriaceae Blood Isolates in Korea. *Antimicrobial Agents and Chemotherapy*. 53:2643-2645.

Kliebe, C., B. Nies, J. Meyer, R. Tolxdorff-Neutzling, and B. Wiedemann. 1985. Evolution of plasmid-coded resistance to broad-spectrum cephalosporins. *Antimicrobial agents and chemotherapy*. 28:302-307.

Knight, R., A. Vrbanac, B.C. Taylor, A. Aksenov, C. Callewaert, J. Debelius, A. Gonzalez, T. Kosciolek, L.-I. McCall, and D. McDonald. 2018. Best practices for analysing microbiomes. *Nature Reviews Microbiology*. 16:410-422.

Kohanski, M.A., D.J. Dwyer, and J.J. Collins. 2010. How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology*. 8:423-435.

Köser, C.U., M.J. Ellington, and S.J. Peacock. 2014. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet*. 30:401-407.

Krause, K.M., A.W. Serio, T.R. Kane, and L.E. Connolly. 2016. Aminoglycosides: an overview. *Cold Spring Harbor perspectives in medicine*. 6:a027029.

Krawczyk, P.S., L. Lipinski, and A. Dziembowski. 2018. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic acids research*. 46:e35-e35.

Kristiansson, E., J. Fick, A. Janzon, R. Grabic, C. Rutgersson, B. Weijdegård, H. Söderström, and D.J. Larsson. 2011. Pyrosequencing of antibiotic-contaminated river sediments reveals high levels of resistance and gene transfer elements. *PloS one*. 6:e17038.

Krizova, L., R.A. Bonnin, P. Nordmann, A. Nemec, and L. Poirel. 2012. Characterization of a multidrug-resistant Acinetobacter baumannii strain carrying the blaNDM-1 and blaOXA-23 carbapenemase genes from the Czech Republic. *J Antimicrob Chemother*. 67:1550-1552.

Kunin, V., A. Copeland, A. Lapidus, K. Mavromatis, and P. Hugenholtz. 2008. A Bioinformatician's Guide to Metagenomics. *Microbiology and Molecular Biology Reviews*. 72:557-578.

Lakin, S.M., C. Dean, N.R. Noyes, A. Dettenwanger, A.S. Ross, E. Doster, P. Rovira, Z. Abdo, K.L. Jones, and J. Ruiz. 2017. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic acids research*. 45:D574-D580.

Lal Gupta, C., R. Kumar Tiwari, and E. Cytryn. 2020. Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes. *Environment International*. 138:105667.

Lanza, V.F., F. Baquero, J.L. Martínez, R. Ramos-Ruíz, B. González-Zorn, A. Andremont, A. Sánchez-Valenzuela, S.D. Ehrlich, S. Kennedy, E. Ruppé, W. van Schaik, R.J. Willems, F. de la Cruz, and T.M. Coque. 2018. In-depth resistome analysis by targeted metagenomics. *Microbiome*. 6:11.

Lanza, V.F., M. de Toro, M.P. Garcillán-Barcia, A. Mora, J. Blanco, T.M. Coque, and F. de la Cruz. 2014. Plasmid flux in Escherichia coli ST131 sublineages, analyzed by plasmid constellation network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. *PLoS Genet*. 10:e1004766.

Lautenbach, E., B.L. Strom, W.B. Bilker, J.B. Patel, P.H. Edelstein, and N.O. Fishman. 2001. Epidemiological investigation of fluoroquinolone resistance in infections due to extended-spectrum β-lactamase—producing Escherichia coli and Klebsiella pneumoniae. *Clinical Infectious Diseases*. 33:1288-1294.

Lavollay, M., K. Mamlouk, T. Frank, A. Akpabie, B. Burghoffer, S.B. Redjeb, R. Bercion, V. Gautier, and G. Arlet. 2006. Clonal dissemination of a CTX-M-15 β-lactamase-producing Escherichia coli strain in the Paris area, Tunis, and Bangui. *Antimicrobial agents and chemotherapy*. 50:2433-2438.

Liakopoulos, A., D. Mevius, and D. Ceccarelli. 2016. A review of SHV extended-spectrum β-lactamases: neglected yet ubiquitous. *Frontiers in microbiology*. 7:1374.

Lindemann, P.C., K. Risberg, H.G. Wiker, and H. Mylvaganam. 2012. Aminoglycoside resistance in clinical Escherichia coli and Klebsiella pneumoniae isolates from Western Norway. *Apmis*. 120:495-502.

Lindgreen, S., K.L. Adair, and P.P. Gardner. 2016. An evaluation of the accuracy and speed of metagenome analysis tools. *Scientific reports*. 6:19233.

Livermore, D.M. 1995. beta-Lactamases in laboratory and clinical resistance. *Clinical microbiology reviews*. 8:557-584.

Livermore, D.M. 1997. β-Lactamases: quantity and resistance. *Clinical Microbiology and Infection*. 3:4S10-14S19.

Livermore, D.M., R. Canton, M. Gniadkowski, P. Nordmann, G.M. Rossolini, G. Arlet, J. Ayala, T.M. Coque, I. Kern-Zdanowicz, and F. Luzzaro. 2007. CTX-M: changing the face of ESBLs in Europe. *Journal of Antimicrobial Chemotherapy*. 59:165-174.

Ma, X., Y. Shao, L. Tian, D.A. Flasch, H.L. Mulder, M.N. Edmonson, Y. Liu, X. Chen, S. Newman, and J. Nakitandwe. 2019. Analysis of error profiles in deep next-generation sequencing data. *Genome biology*. 20:1-15.

Magiorakos, A.-P., A. Srinivasan, R. Carey, Y. Carmeli, M. Falagas, C. Giske, S. Harbarth, J. Hindler, G. Kahlmeter, and B. Olsson-Liljequist. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clinical microbiology and infection*. 18:268-281.

Majiduddin, F.K., I.C. Materon, and T.G. Palzkill. 2002. Molecular analysis of beta-lactamase structure and function. *International Journal of Medical Microbiology*. 292:127-137.

Marini, G., M. Dalcin Luchese, A. Paula Correa Argondizzo, A.C. Góes, R. Galler, T. Alves, M. Medeiros, and A. Leites Larentis. 2014. Experimental design approach in recombinant protein expression: Determining medium composition and induction conditions for expression of pneumolysin from Streptococcus pneumoniae in Escherichia coli and preliminary purification process. 1 pp.

Martínez-Martínez, L., A. Pascual, and G.A. Jacoby. 1998. Quinolone resistance from a transferable plasmid. *The Lancet*. 351:797-799.

McArthur, A.G., and K.K. Tsang. 2017. Antimicrobial resistance surveillance in the genomic age. *Annals of the New York Academy of Sciences*. 1388:78-91.

McArthur, A.G., N. Waglechner, F. Nizam, A. Yan, M.A. Azad, A.J. Baylay, K. Bhullar, M.J. Canova, G. De Pascale, and L. Ejim. 2013. The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy*. 57:3348-3357.

McArthur, A.G., and G.D. Wright. 2015. Bioinformatics of antimicrobial resistance in the age of molecular epidemiology. *Current Opinion in Microbiology*. 27:45-50.

Ministries, N. 2015. National Strategy against Antibiotic Resistance 2015–2020. *Norwegian Ministry of Health and Care Services*:1-36.

Miriagou, V., E. Tzelepi, D. Gianneli, and L.S. Tzouvelekis. 2003. Escherichia coli with a self-transferable, multiresistant plasmid coding for metallo-β-lactamase VIM-1. *Antimicrobial agents and chemotherapy*. 47:395-397.

Moini, A.S., B. Soltani, A.T. Ardakani, A. Moravveji, M. Erami, M.H. Rezaei, and M. Namazi. 2015. Multidrug-resistant Escherichia coli and Klebsiella pneumoniae isolated from patients in Kashan, Iran. *Jundishapur journal of microbiology*. 8.

Mullany, P., E. Allan, and A.P. Roberts. 2015. Mobile genetic elements in Clostridium difficile and their role in genome function. *Research in microbiology*. 166:361-367.

Nemec, A., M. Musilek, M. Maixnerova, T. De Baere, T. Van der Reijden, M. Vaneechoutte, and L. Dijkshoorn. 2009. Acinetobacter beijerinckii sp nov and Acinetobacter gyllenbergii sp nov., haemolytic organisms isolated from humans. *International journal of systematic and evolutionary microbiology*. 59:118-124.

Nordmann, P., L. Dortet, and L. Poirel. 2012. Carbapenem resistance in Enterobacteriaceae: here is the storm! *Trends in molecular medicine*. 18:263-272.

Nordmann, P., and L. Poirel. 2002. Emerging carbapenemases in Gram-negative aerobes. *Clinical Microbiology and Infection*. 8:321-331.

NORM/NORM-VET. 2016. Usage of Antimicrobial Agents and Occurrence of Antimicrobial Resistance in Norway. Tromsø/Oslo 2017.

Novais, Â., R. Cantón, A. Valverde, E. Machado, J.-C. Galán, L. Peixe, A. Carattoli, F. Baquero, and T.M. Coque. 2006. Dissemination and Persistence of bla_{CTX-M-9} Are Linked to Class 1 Integrons Containing CR1 Associated with Defective Transposon Derivatives from Tn402 Located in Early Antibiotic Resistance Plasmids of IncHI2, IncP1-α, and IncFI Groups. *Antimicrobial Agents and Chemotherapy*. 50:2741.

Nowak, P., P. Paluchowska, and A. Budak. 2014. Co-occurrence of carbapenem and aminoglycoside resistance genes among multidrug-resistant clinical isolates of Acinetobacter baumannii from Cracow, Poland. *Medical science monitor basic research*. 20:9-14.

Ojdana, D., A. Sieńko, P. Sacha, P. Majewski, P. Wieczorek, A. Wieczorek, and E. Tryniszewska. 2018. Genetic basis of enzymatic resistance of E. coli to aminoglycosides. *Advances in medical sciences*. 63:9-13.

Orlek, A., N. Stoesser, M.F. Anjum, M. Doumith, M.J. Ellington, T. Peto, D. Crook, N. Woodford, A.S. Walker, and H. Phan. 2017. Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Frontiers in microbiology*. 8:182.

Paczosa, M.K., and J. Mecsas. 2016. Klebsiella pneumoniae: going on the offense with a strong defense. *Microbiology and Molecular Biology Reviews*. 80:629-661.

Page, A.J., A. Wailan, Y. Shao, K. Judge, G. Dougan, E.J. Klemm, N.R. Thomson, and J.A. Keane. 2018a. PlasmidTron: assembling the cause of phenotypes and genotypes from NGS data. *Microb Genom*. 4:e000164.

Page, A.J., A. Wailan, Y. Shao, K. Judge, G. Dougan, E.J. Klemm, N.R. Thomson, and J.A. Keane. 2018b. PlasmidTron: assembling the cause of phenotypes and genotypes from NGS data. *Microbial genomics*. 4.

Paterson, D.L., and R.A. Bonomo. 2005. Extended-spectrum β-lactamases: a clinical update. *Clinical microbiology reviews*. 18:657-686.

Peirano, G., A.K. van der Bij, J.L. Freeman, L. Poirel, P. Nordmann, M. Costello, V.L. Tchesnokova, and J.D. Pitout. 2014. Characteristics of Escherichia coli sequence type 131 isolates that produce extended-spectrum β-lactamases: global distribution of the H30-Rx sublineage. *Antimicrobial agents and chemotherapy*. 58:3762-3767.

Peleg, A.Y., H. Seifert, and D.L. Paterson. 2008. Acinetobacter baumannii: emergence of a successful pathogen. *Clinical microbiology reviews*. 21:538-582.

Pendleton, J.N., S.P. Gorman, and B.F. Gilmore. 2013. Clinical relevance of the ESKAPE pathogens. *Expert review of anti-infective therapy*. 11:297-308.

Pepper, I.L., J.P. Brooks, and C.P. Gerba. 2018. Antibiotic resistant bacteria in municipal wastes: is there reason for concern? *Environmental science & technology*. 52:3949-3959.

Pesesky, M.W., R. Tilley, and D.A. Beck. 2019. Mosaic plasmids are abundant and unevenly distributed across prokaryotic taxa. *Plasmid*. 102:10-18.

Pitout, J.D., P. Nordmann, and L. Poirel. 2015. Carbapenemase-producing Klebsiella pneumoniae, a key pathogen set for global nosocomial dominance. *Antimicrobial agents and chemotherapy*. 59:5873-5884.

Poirel, L., R.A. Bonnin, and P. Nordmann. 2012. Genetic support and diversity of acquired extended-spectrum β-lactamases in Gram-negative rods. *Infection, Genetics and Evolution*. 12:883-893.

Prevention, E.C.f.D., and Control. 2018. Surveillance of antimicrobial resistance in Europe: annual report of the European Antimicrobial Resistance Surveillance Network (EARS-Net) 2017. *S. ECDC; Editor*.

Prjibelski, A.D., I. Vasilinetc, A. Bankevich, A. Gurevich, T. Krivosheeva, S. Nurk, S. Pham, A. Korobeynikov, A. Lapidus, and P.A. Pevzner. 2014. ExSPAnder: a universal repeat resolver for DNA fragment assembly. *Bioinformatics*. 30:i293-i301.

Punina, N., N. Makridakis, M. Remnev, and A. Topunov. 2015. Whole-genome sequencing targets drug-resistant bacterial infections. *Human genomics*. 9:19.

Quail, M., and M. Smith. Coupland Petal (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics*. 13:341.

Quail, M.A., M. Smith, P. Coupland, T.D. Otto, S.R. Harris, T.R. Connor, A. Bertoni, H.P. Swerdlow, and Y. Gu. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 13:341.

Rettedal, S., I. Löhr, E. Bernhoff, O. Natås, A. Sundsfjord, and K. Øymar. 2015. Extended-spectrum β-lactamase-producing Enterobacteriaceae among pregnant women in Norway: prevalence and maternal–neonatal transmission. *Journal of Perinatology*. 35:907-912.

Ricker, N., H. Qian, and R.R. Fulthorpe. 2012. The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. *Genomics*. 100:167-175.

Robicsek, A., J. Strahilevitz, G.A. Jacoby, M. Macielag, D. Abbanat, C.H. Park, K. Bush, and D.C. Hooper. 2006. Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase. *Nature medicine*. 12:83-88.

Rosengren, L.B., C.L. Waldner, and R.J. Reid-Smith. 2009. Associations between antimicrobial resistance phenotypes, antimicrobial resistance genes, and virulence genes of fecal Escherichia coli isolates from healthy grow-finish pigs. *Applied and environmental microbiology*. 75:1373-1380.

Rothberg, J.M., W. Hinz, T.M. Rearick, J. Schultz, W. Mileski, M. Davey, J.H. Leamon, K. Johnson, M.J. Milgrew, and M. Edwards. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 475:348-352.

Rozov, R., A. Brown Kav, D. Bogumil, N. Shterzer, E. Halperin, I. Mizrahi, and R. Shamir. 2017. Recycler: an algorithm for detecting plasmids from de novo assembly graphs. *Bioinformatics*. 33:475-482.

Rozwandowicz, M., M. Brouwer, J. Fischer, J. Wagenaar, B. Gonzalez-Zorn, B. Guerra, D. Mevius, and J. Hordijk. 2018. Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *Journal of Antimicrobial Chemotherapy*. 73:1121-1137.

Sabino, Y.N.V., M.F. Santana, L.B. Oyama, F.G. Santos, A.J.S. Moreira, S.A. Huws, and H.C. Mantovani. 2019. Characterization of antibiotic resistance genes in the species of the rumen microbiota. *Nature communications*. 10:1-11.

Salyers, A.A., N.B. Shoemaker, and L.-Y. Li. 1995. In the driver's seat: the Bacteroides conjugative transposons and the elements they mobilize. *Journal of bacteriology*. 177:5727.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences*. 74:5463-5467.

Schadt, E.E., S. Turner, and A. Kasarskis. 2010. A window into third-generation sequencing. *Human molecular genetics*. 19:R227-R240.

Schatz, M.C., A.L. Delcher, and S.L. Salzberg. 2010. Assembly of large genomes using second-generation sequencing. *Genome research*. 20:1165-1173.

Schmidt, K., S. Mwaigwisya, L. Crossman, M. Doumith, D. Munroe, C. Pires, A. Khan, N. Woodford, N. Saunders, and J. Wain. 2016. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy*. 72:104-114.

Schofield, C.B. 2012. Updating antimicrobial susceptibility testing methods. *Clinical Laboratory Science*. 25:233.

Schwarz, S., C. Kehrenberg, B. Doublet, and A. Cloeckaert. 2004. Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS microbiology reviews*. 28:519-542.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 30:2068-2069.

Šeputienė, V., J. Povilonis, M. Ružauskas, A. Pavilonis, and E. Sužiedėlienė. 2010. Prevalence of trimethoprim resistance genes in Escherichia coli isolates of human and animal origin in Lithuania. *Journal of medical microbiology*. 59:315-322.

Shi, J., Y. Yan, M.G. Links, L. Li, J.-A.R. Dillon, M. Horsch, and A. Kusalik. 2019. Antimicrobial resistance genetic factor identification from whole-genome sequence data using deep feature selection. *BMC Bioinformatics*. 20:535.

Sievert, D.M., P. Ricks, J.R. Edwards, A. Schneider, J. Patel, A. Srinivasan, A. Kallen, B. Limbago, and S. Fridkin. 2013. Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009–2010. *Infection control and hospital epidemiology*. 34:1-14.

Sims, D., I. Sudbery, N.E. Ilott, A. Heger, and C.P. Ponting. 2014. Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*. 15:121-132.

Smalla, K., S. Jechalke, and E.M. Top. 2015. Plasmid detection, characterization, and ecology. *Plasmids: Biology and Impact in Biotechnology and Discovery*:445-458.

Smillie, C., M.P. Garcillán-Barcia, M.V. Francia, E.P. Rocha, and F. de la Cruz. 2010. Mobility of plasmids. *Microbiology and Molecular Biology Reviews*. 74:434-452.

Søraas, A.V.L. 2014. Extended spectrum beta-lactamase producing Enterobacteriaceae. Aspects on carriage, infection and treatment.

Speer, B.S., N.B. Shoemaker, and A.A. Salyers. 1992. Bacterial resistance to tetracycline: mechanisms, transfer, and clinical significance. *Clinical microbiology reviews*. 5:387-399.

Strahilevitz, J., G.A. Jacoby, D.C. Hooper, and A. Robicsek. 2009. Plasmid-mediated quinolone resistance: a multifaceted threat. *Clinical microbiology reviews*. 22:664-689.

Taitt, C.R., T.A. Leski, M.G. Stockelman, D.W. Craft, D.V. Zurawski, B.C. Kirkup, and G.J. Vora. 2014. Antimicrobial resistance determinants in Acinetobacter baumannii isolates taken from military treatment facilities. *Antimicrobial agents and chemotherapy*. 58:767-781.

Thomas, C.M., and K.M. Nielsen. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature reviews microbiology*. 3:711-721.

Thomson, J.M., and R.A. Bonomo. 2005. The threat of antibiotic resistance in Gram-negative pathogenic bacteria: β-lactams in peril! *Current opinion in microbiology*. 8:518-524.

Tipper, D.J., and J.L. Strominger. 1965. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proceedings of the National Academy of Sciences of the United States of America*. 54:1133.

Tolmasky, M.E. 2000. Bacterial resistance to aminoglycosides and beta-lactams: the Tn1331 transposon paradigm. *RNA*. 5:11.

Treangen, T.J., and S.L. Salzberg. 2011. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet*. 13:36-46.

Tumbarello, M., T. Spanu, M. Sanguinetti, R. Citton, E. Montuori, F. Leone, G. Fadda, and R. Cauda. 2006. Bloodstream infections caused by extended-spectrum-β-lactamase-producing Klebsiella pneumoniae: risk factors, molecular epidemiology, and clinical outcome. *Antimicrobial agents and chemotherapy*. 50:498-504.

Tzouvelekis, L.S., E. Tzelepi, P.T. Tassios, and N.J. Legakis. 2000. CTX-M-type β-lactamases: an emerging group of extended-spectrum enzymes. *International Journal of Antimicrobial Agents*. 14:137-142.

Ulstad, C.R., M. Solheim, S. Berg, M. Lindbæk, U.R. Dahle, and A.L. Wester. 2016. Carriage of ESBL/AmpC-producing or ciprofloxacin non-susceptible Escherichia coli and Klebsiella spp. in healthy people in Norway. *Antimicrobial Resistance & Infection Control*. 5:57.

Vakulenko, S.B., and S. Mobashery. 2003. Versatility of aminoglycosides and prospects for their future. *Clinical microbiology reviews*. 16:430-450.

Valverde, A., T.M. Coque, M.P. Sánchez-Moreno, A. Rollán, F. Baquero, and R. Cantón. 2004. Dramatic increase in prevalence of fecal carriage of extended-spectrum β-lactamase-producing Enterobacteriaceae during nonoutbreak situations in Spain. *Journal of clinical microbiology*. 42:4769-4775.

Van Veen, S., E. Claas, and E.J. Kuijper. 2010. High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *Journal of clinical microbiology*. 48:900-907.

Veldman, K., L.M. Cavaco, D. Mevius, A. Battisti, A. Franco, N. Botteldoorn, M. Bruneau, A. Perrin-Guyomard, T. Cerny, and C. De Frutos Escobar. 2011. International collaborative study on the occurrence of plasmid-mediated quinolone resistance in Salmonella enterica and Escherichia coli isolated from animals, humans, food and

the environment in 13 European countries. *Journal of Antimicrobial Chemotherapy*. 66:1278-1286.

Vollmers, J., S. Wiegand, and A.-K. Kaster. 2017. Comparing and evaluating metagenome assembly tools from a microbiologist's perspective-not only size matters! *PloS one*. 12:e0169662.

Vranic, S.M., and A. Uzunovic. 2016. Antimicrobial resistance of escherichia coli strains isolated from urine at outpatient population: a single laboratory experience. *Materia socio-medica*. 28:121.

Walsh, F., and B. Duffy. 2013. The culturable soil antibiotic resistome: a community of multi-drug resistant bacteria. *PloS one*. 8:e65567.

Welker, M., and A. Van Belkum. 2019. One system for all: is mass spectrometry a future alternative for conventional antibiotic susceptibility testing? *Frontiers in Microbiology*. 10:2711.

Wellington, E.M., A.B. Boxall, P. Cross, E.J. Feil, W.H. Gaze, P.M. Hawkey, A.S. Johnson-Rollings, D.L. Jones, N.M. Lee, and W. Otten. 2013. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. *The Lancet infectious diseases*. 13:155-165.

Willey, J.M., L. Sherwood, and C.J. Woolverton. 2011. Prescott's microbiology. McGraw-Hill New York.

Woodford, N., M.J. Ellington, J.M. Coelho, J.F. Turton, M.E. Ward, S. Brown, S.G.B. Amyes, and D.M. Livermore. 2006. Multiplex PCR for genes encoding prevalent OXA carbapenemases in Acinetobacter spp. *International Journal of Antimicrobial Agents*. 27:351-353.

Xavier, B., A. Das, G. Cochrane, S. Ganck, S. Kumar-Singh, F. Aarestrup, H. Goossens, and S. malhotra-kumar. 2016. Consolidating and Exploring Antibiotic Resistance Gene Data Resources. *Journal of Clinical Microbiology*. 54.

Zankari, E., H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, F.M. Aarestrup, and M.V. Larsen. 2012. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*. 67:2640-2644.

Zankari, E., H. Hasman, R.S. Kaas, A.M. Seyfarth, Y. Agersø, O. Lund, M.V. Larsen, and F.M. Aarestrup. 2013. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *Journal of Antimicrobial Chemotherapy*. 68:771-777.

Zhang, T., X.-X. Zhang, and L. Ye. 2011. Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. *PloS one*. 6:e26041.

Zhao, S., Z. Liang, V. Demko, R. Wilson, W. Johansen, O.-A. Olsen, and K.J.B.e.b. Shalchian-Tabrizi. 2012. Massive expansion of the calpain gene family in unicellular eukaryotes. 12:193.

Zhou, F., and Y. Xu. 2010. cBar: a computer program to distinguish plasmid-derived from chromosome-derived sequence fragments in metagenomics data. *Bioinformatics*. 26:2051-2052.

9. Appendix

9.1 Appendix A

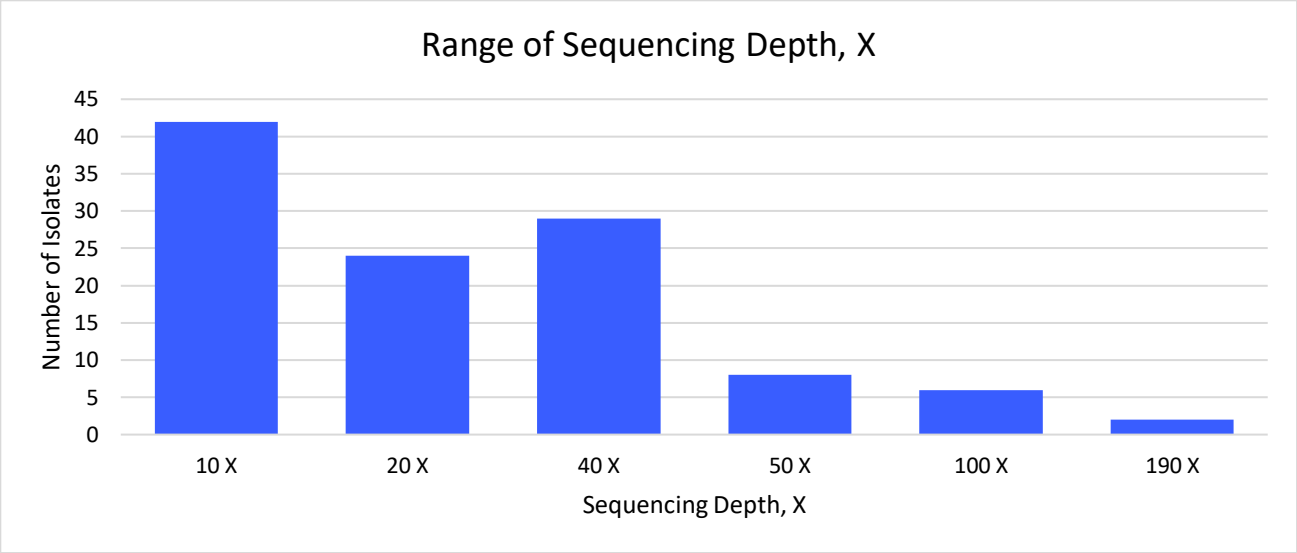


Figure 16: Range of sequencing depth distribution among the isolates (n=111)

9.2 Appendix B

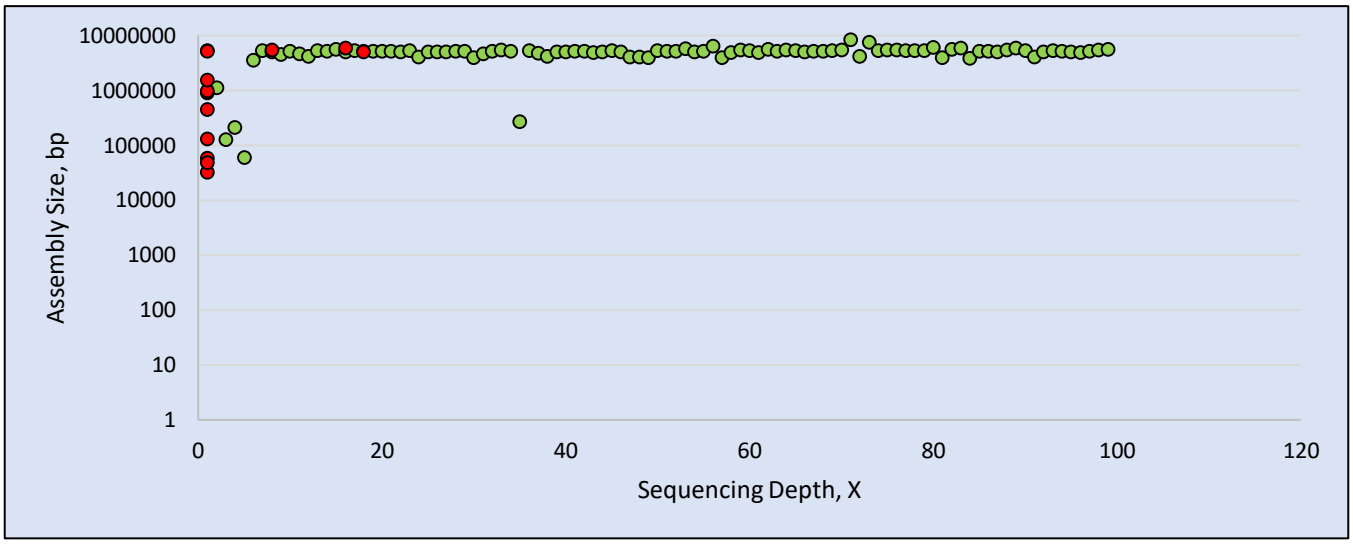


Figure 17 : Graphical Representation of the effect of sequencing depth on the genotype resistance detection of the isolates. The red colored dots represent the isolates with phenotypic resistance without any resistance genes. Two E. coli and ten K. pneumoniae isolates displayed phenotypic resistance without corresponding resistance genes. All these isolates had sequencing depth between 1X- 16X.

9.3 Appendix C

(Note for supplementary file 5)

Genome annotation with Prokka Genome size of isolates (142, 233, 112, 119 , 131, 150, 105, 140, 234, 222, 236, 230, 231, 235, , 200, 215, 223, 232) fell into either below 2 Mbp or above 6 Mbp (156, 157, 141, 136, 220, 225, 241, 243, 307, 302, 304). A few isolates had genes below 2000 (142, 112, 119, 131, 105, 150, 233, 142, 234, 236, 222, 230, 231, 235, 200, 215, 223) or above 6000 (141, 136, 157, 204, 225, 220, 243, 304).

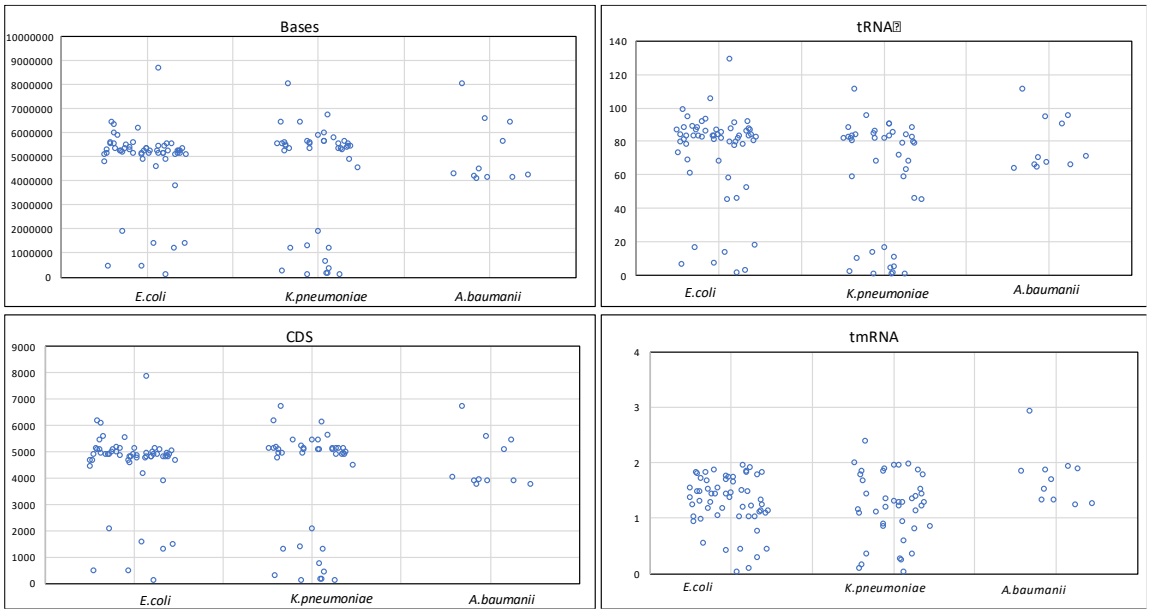


Figure 18: Prokka statistics for general features of E. coli, K. pneumoniae, and Acinetobacter genome.

9.4 Appendix D

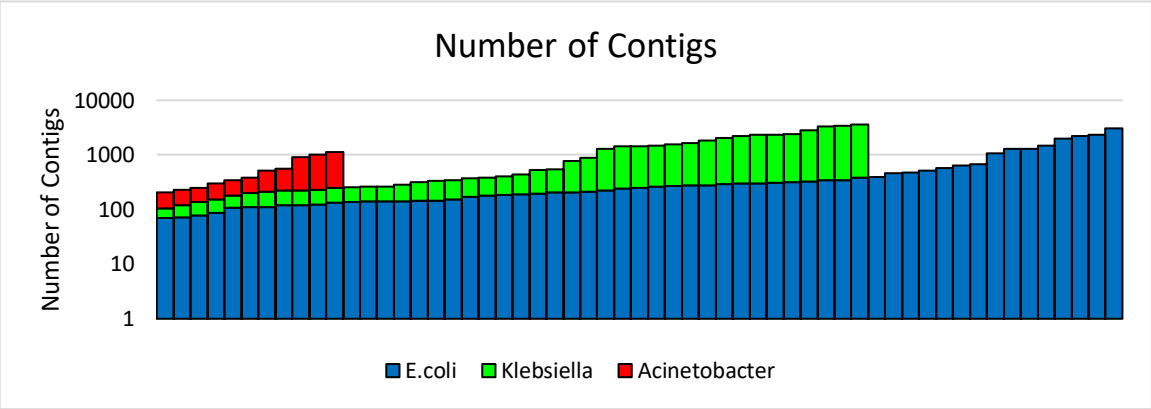


Figure 19: Number of contigs from SPAdes assembly assessment by Quast

9.5 Appendix E

Table 8: Concordance between phenotype and genotype for predictions made using a database of resistance determinants and the plasmids. Red color coded genes and antibiotics have concordance.

Isolates	Antibiotic Resistance Profile	In Silico Resistance Genes (CARD)	In Silico detected Plasmids
100	Ciprofloxacin	ampC	IncFII , Col(BS512) , IncFIA
101		ampC	IncFII, IncFIB
102	Ampicillin, Gentamicin, Trimethoprim	ampC, TEM, aac, *dfr*	IncFIB, Col156
103	Ampicillin, Gentamicin	ampC, TEM, *aac*	IncFII(29) , IncFIB, Col156
104		ampC	IncFII(29) , IncFIB, Col156,
105		ampC	IncFII , IncFIB, Col156, Col(BS512)
106	Ampicillin, Trimethoprim	ampC , TEM	IncFII(29) , IncFIB, IncI1(Alpha)
107	Ampicillin, Trimethoprim	ampC , TEM, dfr,*tet, qnr*	IncFII(29) , IncFIB, Col156
108	Ampicillin	ampC, SHV	IncFIB, Col156
109		ampC	IncFII , I ncFIB
110	Gentamicin, Trimethoprim		IncFIB, Col156
111		ampC	
112	Ampicillin	ampC, TEM	IncFII, IncFIB , Col156 , B, pO104:H7_S1
113		ampC	IncFIB, Col156
114		ampC	IncX4, IncI1(Alpha)
115			IncFIB, pMRSN480738_1.6,
116		ampC, *aph, tet*	IncFIB , pCTXM15_000837
117	Ampicillin, Ciprofloxacin, Gentamicin	ampC, OXA, *aac*	IncFII(pCoo), IncFIB
118	Ampicillin	ampC	
119	Trimethoprim		
120	Ampicillin	ampC,TEM, *aph*	IncFII, IncFIB , Col156 , p94EC-6, pCFSAN030807_7, pEC732_5,
121	Ampicillin, Trimethoprim	ampC , TEM, *dfr, aph*	IncQ1,
122		ampC	IncFII, IncFIB, RCS49_pII, p0111,
123	Ampicillin, Trimethoprim	ampC, TEM, *dfr , aph*	IncFII(29) , IncFIB, Col156 .
124		ampC, *cat*	IncFII(pRSB107) , IncFIB, Col(BS512), Col156, p0111
125		ampC, TEM	IncFII(29), IncFIB, Col156, Col8282, Col(BS512), Col(MG828) , IncX1
126		ampC, TEM	IncFII(29), Col156, Col8282, Col(BS512), IncX4, IncX1
127	Ampicillin, Trimethoprim	ampC, TEM, *dfr*	IncFII(29), IncFIB, Col(BS512), Col8282, Col(MG828), Col156, p2_020022, p2_000837, pFAM21845_3, p12579_5,
128	Ampicillin	ampC, , TEM, *aph, tet*	IncFII(pRSB107) , IncFII(pSE11), IncFIB(pB171), Col156, Col(BS512), pVir_020022, pCTXM15_000837, pINF078-VP, pCHL5009T-102k-mcr3, RCS93_pI, pSTEC299_1, p13P484A-1, p13TMH22-2, pG749_3, A
129	Ampicillin	ampC , TEM, *aph*	IncFII(29) , IncFIB, Col156, IncQ1,
130	Ampicilline, Ciprofloxacin, Trimethoprim	ampC, TEM, *dfr, tet*	IncFII(pRSB107) , IncFIB , Col156, Col(MG828), pIB_NDM_1, pCTXM15_000837, p2_000837, pTMTA63632
131		ampC	Col8282, Col(BS512) , IncL/M(pMU407),
132	Ciprofloxacin	ampC	IncFIB, Col8282, p2_W5-6, IncY,
133		ampC	
134	Ampicillin, Trimethoprim	ampC, TEM, *dfr*	6.6k
135		ampC	IncFIB , Col8282
136	Ampicillin, Trimethoprim	ampC, *aph*, TEM	IncFII(29), IncFIB , Col156, Col156
137		ampC	Col(MG828), p2_000837
138		ampC	Unknown
139		ampC	IncFII, IncFIB, pG749_3, pBMB0558
140		ampC, OXA, *aac*	IncFIB, IncFIB(S) , Col156, IncHI2A,
141		ampC, VIM, SHV, *tet, qnr*	IncFIB , IncN, pVir_020022, A,
142	Ciprofloxacin	ampC	IncFII , B, IncI1(Alpha), pTB211, pVir_020022, pC-Ec2-KPC,
143	Ampicillin, Ciprofloxacin	ampC, TEM	IncFIB, IncFIB(K), p2_000837,
144	Ciprofloxacin	ampC, OXA, TEM	IncFII(K), IncFIB, Col156, Col(BS512), IncFIB(pQil),
145		ampC	IncFII(pRSB107), IncFIB
146	Ampicilline, Cefotaxime , Ciprofloxacin, Trimethoprim	ampC , CTX-M, *dfr , aph*	IncFII(pRSB107) , IncFIB, pVir_020022, p2_000837
147	Ampicillin	ampC, TEM	IncFII , Col(MG828), Col156
148		ampC	IncFIB, Col8282, p2_020022
149		ampC	Col156, pSL483
150		ampC	IncFII(29), Col156, Col(BS512)
151	Chloramphenicol	ampc	
152		ampc	Unknown
153		ampC, *tet*	p0111
154		ampC	p2_000837, pVR50B, A
155	Trimethoprim	TEM, *dfr, tet*	IncFIA(HI1)
156	Ampicillin	ampC, TEM, *tet*	IncFII, IncFII(29), IncFIB, p2_W2-5, pCTXM15_000837, unitig_4,
157	Ampicillin	TEM, OXA, CTX-M, *aph, aac, tet, cat*	IncFII(pCoo)

Isolates	Antibiotic Resistance Profile	In Silico Resistance Genes (CARD)	In Silico detected Plasmids
200	Ampicillin	SHV	Col(MG828)
201	Ampicillin	SHV	IncHI1B(pNDM-MAR)
202	Ampicillin	SHV	IncFIB(K), Col(MG828)
203	Ampicillin	SHV	IncFII(pKP91), IncFIB(K), Col(MG828)
204	Ampicillin	SHV	ColE10
205	Ampicillin	SHV	Unknown
206	Ampicillin	SHV	IncFII(pKP91) , IncFIB(K), Col(MG828)
207	Ampicillin	SHV	Unknown
208	Ampicillin	LEN, SHV	IncFII(K), IncFIB(K)
209	Ampicillin	SHV,	IncFII(K), IncFIB(K), IncFIA(HI1)
210	Ampicillin	SHV	IncFIB(K)
211	Ampicilline, Cefotaxime, Ceftazidime, Chloramphenicol, Ciprofloxacin,	TEM, SHV, *aac, qnr, floR*	Col(MG828), 4, pIncFIA-1502320, pCFSAN061772_02, IncX3, IncFIA(HI1)
212	Ampicilline, Cefotaxime, Ceftazidime, Chloramphenicol, Ciprofloxacin,	TEM, SHV , *aac,, floR*	IncFII(K) , IncFIB(K)
213	Ciprofloxacin, Gentamicin, Trimethoprim	SHV, CTX-M, TEM, OXA, *aac, aph, qnr, dfr*	p2_W5-6
214	Ampicillin		IncFIB(PKPHS1), pIB_NDM_1, pCHL5009T-102k-mcr3, pVir_020022, pFORC64.1, pSTEC299_1, pCFSAN029787_01, A, B
215	Ampicillin, Tigecycline, Trimethoprim		
216	Ampicillin, Cefotaxime, Ceftazidime, Tigecycline, Trimethoprim	SHV, CTX-M, TEM, *dfr, qnr, aph*	IncFII(K), IncFIB(K) , Col(KPHS6)
217	Ampicillin, Chloramphenicol,	SHV	IncFIB, Col(KPHS6)
218	Ampicillin	SHV	Col(KPHS6)
219	Ampicillin	SHV	
220	Ampicilline, Cefotaxime, Ceftazidime, Trimethoprim	SHV, CTX-M, TEM, *aph, dfr*	IncFIB(K)
221	Ampicillin	SHV, TEM	MAR)
222	Ampicillin		Col(MG828)
223	Ampicillin		
224	Ampicillin	OKP-A	IncFIB
225	Ampicillin, Chloramphenicol,	SHV	IncFII(pKP91), Col156
226	Ampicillin , Chloramphenicol, Tigecycline, Trimethoprim	SHV, TEM	p5, Col156
227	Ampicillin	SHV	IncFII(YP) , Col156, pKPN535a
228	Ampicillin	TEM, SHV	IncFII(K), IncFIB(K)
229		LEN	IncFIA(HI1)
230	Ampicillin		IncFIB
231	Ampicillin		
232	Ampicillin	*aac*	IncFII(pCTU2)
233	Ampicillin		
234	Ampicillin		
235	Ampicillin		
236	Ampicillin, Trimethoprim		Col156
240	Ampicillin, Ceftazidime, Chloramphenicol,	TEM, SHV, *aac*	IncFII, IncFIB(pKPHS1), IncFIB(pNDM-Mar), IncHI1B(pNDM-MAR), ColpVC
241	Ampicillin	ampC, SHV, OXA, *aac, aph tet, cat*	IncHI1B(pNDM-MAR)
242	Ampicillin , Chloramphenicol	SHV	
243	Ampicillin	ampC, SHV, OXA, a*ph, aac, tet, cat*	IncFII(K) , IncFIB(K), Col156
244	Ampicillin	SHV	IncHI1B(pNDM-MAR)
Isolates	**Antibiotic Resistance Profile**	**In Silico Resistance Genes (CARD)**	**In Silico detected Plasmids**
301	Ciprofloxacin, Gentamicin,	ampC, OXA, *aac, aph*	Unknown
302	Ciprofloxacin, Gentamicin,	ampC, OXA, TEM, *aac, aph, tet*	Unknown
303	Ciprofloxacin, Gentamicin,	ampC, OXA, TEM, aac, *cat*	Unknown
304	Ciprofloxacin, Gentamicin,	ampC, OXA, *aac, aph*	Unknown
305	Ciprofloxacin, Gentamicin,	ampC, OXA, *aph, aac, tet*	Unknown
306	Ciprofloxacin, Gentamicin,	ampC, OXA, *aph*	Unknown
307	Ciprofloxacin, Gentamicin,	ampC, OXA, *aph*	Unknown
308	Ciprofloxacin, Gentamicin,	ampC, OXA, TEM, *aph, tet*	Unknown
309	Ciprofloxacin, Gentamicin,	ampC, OXA, *aph, aac, tet, cat*	Unknown
310	Ciprofloxacin, Meropenem	ampC, OXA, TEM, *aph, tet*	Unknown
311	Meropenem	ampC, OXA, TEM, *aph, tet, Cml*	pHZE23-1-1

9.6 Supplementary Files

1. **S1- Gene variants and nodes CARD/ResFinder**
   Description: Gene variants along with nodes, gene length and identity from CARD and ResFinder. Details of MEGARes AMRFinder and Prokka are also included.

2. **S2- Plasmid detection *in silico***
   Description: Includes WGS for PlasmidFinder detection and PlasFlow segregated WGS probed by Brooks et al and PLSDB for plasmids

3. **S3- Plasmidspades assemblies for plasmid detection**
   Description: WGS assembled with PlasmidSPAdes and probed through PLSDB, Brooks et al and PlasmidFInder

4. **S4- Quast statistics for SPAdes assemblies**
   Description: SPAdes assemblies evaluated with QUAST for assembly quality

5. **S5- Prokka stats supplementary file**
   Description: Genomic feature annotation with Prokka

6. **S6- Sequencing coverage**
   Description: Sequencing coverage for all the isolates

7. **S7- mlplasmid predictions**
   Description: Plasmid contig prediction through probability value
8. **S8-AMR database 18 Sep 2020 comparison**
   Description: Describes AMR database performance evaluation for AMR deection
9. **S9- AMR database comparison supplementary file**
   Description: CARD, ResFinder, MEGARes, AMRFinder, Prokka comparison
10. **S10- Trimming data**
    Description: Input reads and surviving reads comparison
11. **S11- After trimming sequence quality**
    Description: FASTQC