Molecular insights into the mechanisms of drug-induced resistance reversion in *Staphylococcus aureus*, *Escherichia coli* and *Acinetobacter baumannii*

By

Monique Joubert

Submitted in partial fulfilment of the requirements for the degree

MSc Bioinformatics

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

(4 February 2020)

Declaration

I, Monique Joubert declare that the dissertation, which I hereby submit for the degree MSc Bioinformatics at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: …….

DATE: …….

Table of Contents

List of Figures

Chapter 1

Chapter 2

List of Tables

Chapter 1

Chapter 2

List of Supplementary Files

Supplementary file 1

Supplementary file 2

Supplementary file 3

Supplementary file 4

Supplementary file 5

Supplementary file 6

Research Outputs

Conferences:

- The SAGS/SASBi conference (2018, Free State)
- SMRTLeiden conference (2019, Leiden, Netherland)

Publications (accepted)

Joubert, M., et al. (2019). "Assembly of Complete Genome Sequences of Negative-Control and Experimental Strain Variants of *Staphylococcus aureus* ATCC BAA-39 Selected under the Effect of the Drug FS-1, Which Induces Antibiotic Resistance Reversion." Microbiol Resour Announc **8**(30).

Korotetskiy, I. S., et al. (2019). "Complete Genome Sequence of a Multidrug-Resistant Strain, *Escherichia coli* ATCC BAA-196, as a Model for Studying Induced Antibiotic Resistance Reversion." Microbiol Resour Announc **8**(50): e01118-01119.

Korotetskiy, I. S., et al. (2020). "Complete Genome Sequence of Collection Strain *Acinetobacter baumannii* ATCC BAA-1790, Used as a Model To Study the Antibiotic Resistance Reversion Induced by Iodine-Containing Complexes." Microbiol Resour Announc **9**(3): e01467-01419.

Publications (submitted)

- "Role of chromosomal epigenetic modifications in antibiotic resistance reversion induced by the iodine-containing complex FS-1 in multidrug-resistant *Staphylococcus aureus*" in Frontiers in Microbiology
- "Epigenetic modifications in genomic DNA of the multidrug-resistant strain *Escherichia coli* ATCC BAA-196 treated with the drug FS-1 inducing antibiotic resistance reversion" in Microbiological Research

Dissertation Summary

The development of multidrug resistant bacteria is currently a great concern since the misuse of antibiotics have caused a strong selective pressure for these resistant bacteria and various treatment options are becoming ineffective. Reversion of antibiotic resistant bacteria into antibiotic sensitive phenotypes is becoming a promising approach to address this problem. This study set out to investigate the molecular mechanisms of drug induced resistance reversion in three multi-drug resistant bacteria, *S. aureus*, *E. coli*, and *A. baumannii*, after treatment with an iodine-containing drug, FS-1. Bacteria cultivated on medium containing FS-1, served as experimental bacteria (denoted as FS), and bacteria cultivated on normal medium, served as negative control cultures (denoted as NC). SMRT sequencing was used to generated long reads and high genome coverages for bacteria genomes. These sequences, together with tools for the SMRT-link software, enabled us to obtain complete bacterial genomes assemblies. The genomes were annotated using the RAST server, followed by manual corrections as needed. All genomes were submitted to the NCBI and genome announcements were published. Upon investigation of FS genomes compared to NC genomes, we found increased frameshift mutations occurring in FS-1 treated cultures. RNA sequences were also generated for these bacteria during different growth phases. This was used to investigate the effect of FS-1 on bacterial metabolism, as well as the direct effect of the drug. Bacteria were also cultivated over 10 passages to determine the effect of the drug on bacterial populations. For the immediate effect, treatment with FS-1 caused downregulation of various important pathway which consume the co-enzymes NADH and NADPH, while the metabolic processes associated with the production of the reduced species of these co-enzymes were generally up-regulated. It may be assumed that the pathways helping bacteria to withstand oxidative stress were up-regulated. Bacteria cultivated on a medium containing FS-1 regained their initial growth rate by adapting to the presence of FS-1, which required an alternative gene transcription regulation controlled

either by accumulation of specific mutations in bacterial populations, or due to epigenetic phase variations. Lastly, modified bases were detected using the generated PacBio reads, together with tools available from the SMRT-link software. In both *E. coli* and *S. aureus*, bacterial cultures treated with FS-1 had an overall increase in modified bases, while the number of methylated nucleotides remained unchanged. This was specifically observed in G and A bases. It was hypothesized that the observed sporadically modified nucleotides might be due to oxidation, especially of G bases, by the iodine contained in FS-1. Therefore, these findings conclude that treatment with FS-1 possibly leads to DNA oxidation, especially of G bases, which caused frame shift mutations, as well as alternative gene transcription regulation. This is a possible explanation of how resistant bacteria which were treated with FS-1, had increased sensitivity to antibiotics.

Acknowledgements

I would like to firstly thank my supervisor, Prof Oleg Reva, for all his guidance and support through each stage of the process. I would also like to thank the BIF team, especially Dewald, Prof Fourie Joubert, Lizette, Hannes, and Alisa for all their support, for every BIFday celebration, for every meme on the wall, and for all the laughter in the office. The morning coffee sessions and lunch breaks on the couch will be missed. I would also like to thank our systems administrator, Johann Swart, for all his technical support.

For all the lab work conducted in this study, I would like to acknowledge and thank our collaborators from the Scientific Center for Anti-Infectious Drugs (SCAID), Almaty, Kazakhstan.

I would lastly like to thank my parents and grandparents for always believing in me, and for all their unconditional love and support.

CHAPTER 1

Literature Review

Introduction

As defined by Selman A. Waksman, “An antibiotic is a chemical substance, produced by micro-organisms, which has the capacity to inhibit the growth of and even to destroy bacteria and other micro-organisms." (Waksman and Flynn, 1973). Ever since their discovery by Alexander Fleming in 1928, antibiotics have revolutionized the field of modern medicine in many respects, making them one of the most significant tools in this field (Munita and Arias, 2016, Abraham et al., 1992). Although originally developed for treating human infectious diseases, their remarkable antimicrobial properties has spread to the broad application of antibiotics in animals, plants, and also aquaculture (Barbosa and Levy, 2000). This misuse and overuse of antibiotics has led to a strong selective pressure, resulting in the survival and wide distribution of drug resistant pathogens, threatening their efficacy (Abraham et al., 1992, Barbosa and Levy, 2000).

Antibiotic discovery, their modes of action, and resistance has been an important research topic in the academic world and, until recently, in the pharmacological industry (Hughes and Karlén, 2014). As natural products, antibiotics provide challenges with respect to their chemical properties, pathways, biochemical mechanisms, and evolution (Brotz-Oesterhelt and Brunner, 2008). The latter makes the synthesis of antibiotics in the laboratory extremely challenging and costly (Nicolaou and Montagnon, 2008). Although existing compounds are continuously being modified and new antibiotic classes are being discovered, the rate of resistance continues to rise while antibiotic discovery rates have started to substantially drop (Laxminarayan, 2014). Therefore, the development of new methods, which focus on addressing the evolution of resistance mechanisms, need to be prioritized along with the discovery of new antibiotics (McClure and Day, 2014, Bush et al., 2011). It is of great importance to devise new strategies that focus on limiting, redirecting and/or reversing the development of bacterial resistance.

Antibiotics

Antibiotics allow for the body's natural defence mechanisms, such as the immune system, to actively eliminate invading micro-organisms. Antibiotics can either be cytostatic or cytotoxic, and mainly act by inhibiting the synthesis of vital structures such as bacterial cell walls, proteins, DNA, and/or RNA (Levy and Marshall, 2004). Antibiotics can be classified based on their mode of action, as described below.

Inhibition of cell wall synthesis

Bacterial cell walls serve as essential protective and shape-maintaining polysaccharide structures in nearly all bacteria, and its membrane integrity is crucially important for cell viability by protecting bacteria from osmotic shock. Gram-positive and Gram-negative bacteria both have cell walls which consist of a polymer of amino acids and sugars, known as peptidoglycan (Silhavy et al., 2010).

The biosynthesis of peptidoglycan involves the cross-linkage of glycan strands through the action of a transglycosidase (Kahne et al., 2005). These strands consist of alternating N-acetylglucosamine (Glc-NAc) and N-acetylmuramic acid (Mur-NAc), as seen in *Figure 1.1*. Bacteria encode for penicillin-binding proteins (PBPs), which are key proteins involved in the assembly of peptidoglycan (Sauvage et al., 2008).

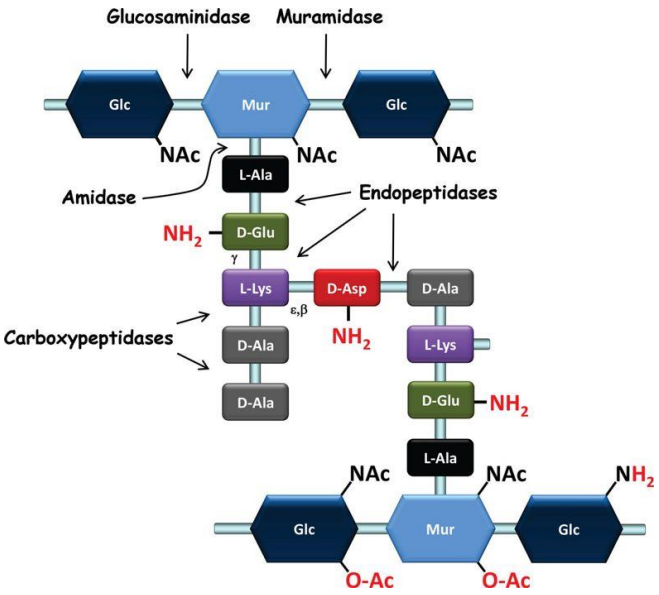


Figure 1.1: Schematic diagram of the peptidoglycan (murein) structure, retrieved from (Chapot-Chartier and Kulakauskas, 2014). This polymer consists of amino acids and sugars. β-(1,4) linked N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues make up the sugar structures.

Gram positive bacterial

In Gram-positive bacteria, such as *Staphylococcus aureus* (*S. aureus*), the cell membrane is surrounded with a thick peptidoglycan layer, which is not only essential for cell viability and growth, but is also a major antibiotic target site (*Figure 1.2-A*) (Vollmer et al., 2008). β-lactam antibiotics, such as Penicillin and Cephalosporins, are one of the most widely used antibiotic classes. Their mode of action involves impairment of bacterial cell wall synthesis by targeting PBP (Park and Strominger, 1957). β-lactams interfere with peptidoglycan synthesis due to the interaction between PBPs and the β-lactam rings, leading to a disrupted peptidoglycan layer and lysis of bacteria (Džidić et al., 2008).

Cell membrane permeability impairment

Gram negative bacterial

Gram-negative bacteria differ from Gram-positive bacteria by not only having a cytoplasmic cell membrane, but by also being surrounded by an outer membrane (OM), see *Figure 1.2-B* below. Gram-negative bacteria have a lipid species that is unique to them, called lipopolysaccharides, which serves as the major lipid component on the outer monolayer of the membrane (Beutler, 2002). The OM serves as a selective barrier with permeability properties which influences the susceptibility of antibiotics. Hydrophobic drugs can get into the interior of cells by diffusing across the lipid bilayer, while small hydrophilic drugs, such as β-lactams, make use of pore-forming β-barrel proteins, known as porins. The OM barrier plays an important role in antibiotic sensitivity, since the survival of drug-resistant strains are due to modified lipid/protein compositions in the OM (Delcour, 2009).

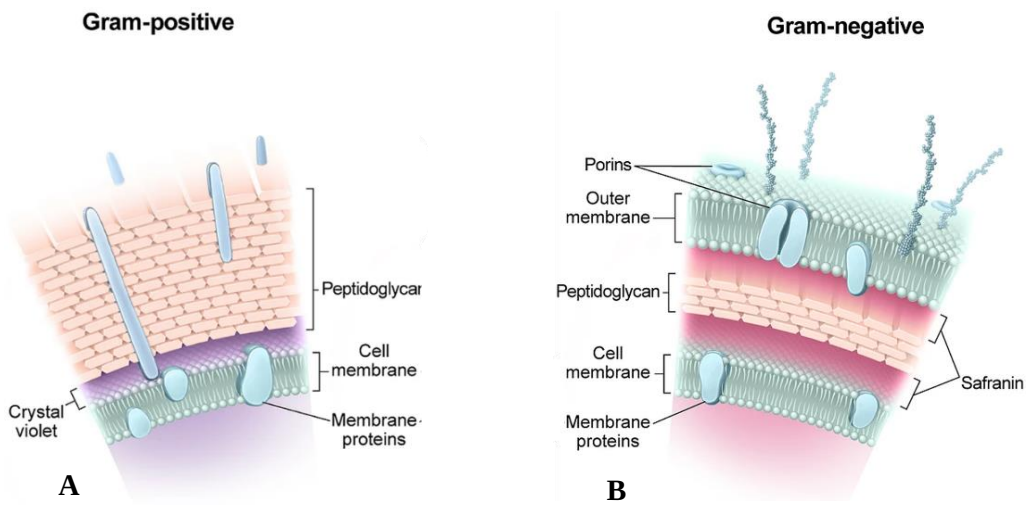


Figure 1.2: Comparative illustration of bacterial cell wall structures, adapted from (Vollmer et al., 2008). Gram-positive bacteria (A) have a thick cell wall, comprising mainly of a thick peptidoglycan layer, compared to the thin cell wall in Gram-negative bacteria (B), which have a thin peptidoglycan layer and an OM.

Inhibition of RNA/DNA synthesis

Quinolones

DNA topoisomerases are enzymes that play a central part in DNA biology and have become popular targets for antibiotics (Nitiss, 2009). DNA gyrases are part of the topoisomerase enzyme class, which is involved in maintaining DNA topology. Double stranded DNA is nicked by DNA gyrase to introduce negative supercoils. Fluoroquinolone antibiotics target bacterial DNA gyrase enzymes (Reece and Maxwell, 1991).

Inhibition of protein synthesis

The bacterial ribosome organelle, 70S, is comprised of two ribonucleoprotein subunits, 30S and 50S (Kohanski et al., 2010). Antibiotics that act to inhibit the 30S ribosomes include tetracycline and aminocyclitol. Tetracycline blocks the binding of aminoacyl-tRNAs to the ribosomes (Chopra and Roberts, 2001), whereas aminocyclitols, such as spectinomycin, bind to the 16S rRNA component of the 30S ribosome subunit to interfere with peptidyl-tRNA translocation (Kotra et al., 2000). Classes of antibiotics that act to inhibit the 50S ribosome subunits include amphenicol, lincosamide, streptogramin, and oxazolidinone.

Origin of antibiotic resistance

Antibiotic resistance occurs once a drug no longer has the ability to effectively inhibit bacterial growth. The resistant bacteria are then able to continuously multiply in the presence of the drug, while the growth of sensitive bacteria is halted. Resistant bacteria can rapidly outnumber the susceptible bacteria, causing a wide distribution of resistance throughout a population, threatening the efficacy of the drug. Higher than normal concentrations of the same drug are then required to have an effect on these resistant bacteria. A selective pressure is induced by antibiotics, and bacteria respond in association with the selective pressure (Bush and Schmidt, 2018).

Acquired Drug Resistance (ADR)

There are various factors involved in the emergence and spread of multiple resistant bacteria. These factors include mutations, which can expand the spectrum of resistance activity; the transfer of genetic material, called horizontal gene transfer (HGT); selective pressure in healthcare and community settings, which facilitates the development and distribution of multiple resistant bacteria; and, in some cases, the inability to detect emerging resistant phenotypes (Tenover and Hughes, 1996). Resistance can occur as a natural process in bacteria, either through gene level mutations, or through HGT (Lowy, 2003).

Mutations

Spontaneous mutations confer resistance in various ways, which include gene amplification (Sandegren and Andersson, 2009), modifying the target site for the antibiotic, or by up-regulation of resistance genes, such as efflux pumps or β-lactamases (Blair et al., 2015, Palmer and Kishony, 2014). Mutations in vital genes encoding for key metabolic pathways, are primary factors for development of resistance to antibiotics, such as fluoroquinolones, rifampin, and streptomycin in certain bacteria, which include *Mycobacterium tuberculosis* (Kapur et al., 1995). *Mutations in specific loci can also cause sensitive bacteria, such as Streptococcus pneumoniae, to become resistant to fluoroquinolones* (Tenover, 2001). Bacteria are known to have short generation times, enabling them to rapidly adapt to changing environments. In certain bacterial species, such as *Mycobacterium tuberculosis*, resistance to antimicrobial agents are solely mediated by mutations (Woodford and Ellington, 2007).

Mutations that occur in pre-existing resistance genes play a different role by either increasing resistance to specific antibiotics, or by expanding the spectrum of resistance. Examples include pneumococci developing increased resistance to cefotaxime and the development of extended spectrum β-lactamases (Tenover, 2001).

HGT

Resistance genes acquired through HGT can encode for specific antibiotic degradation enzymes, protection proteins, novel efflux pumps, or alternative pathways (Blair et al., 2015). Gram-positive bacteria usually exchange genetic material through transformations and transductions, whereas the most commonly recognized mode of exchange between Gram-negative bacteria is conjugation (*Figure 1.3* below) (Sadowski et al., 1979).

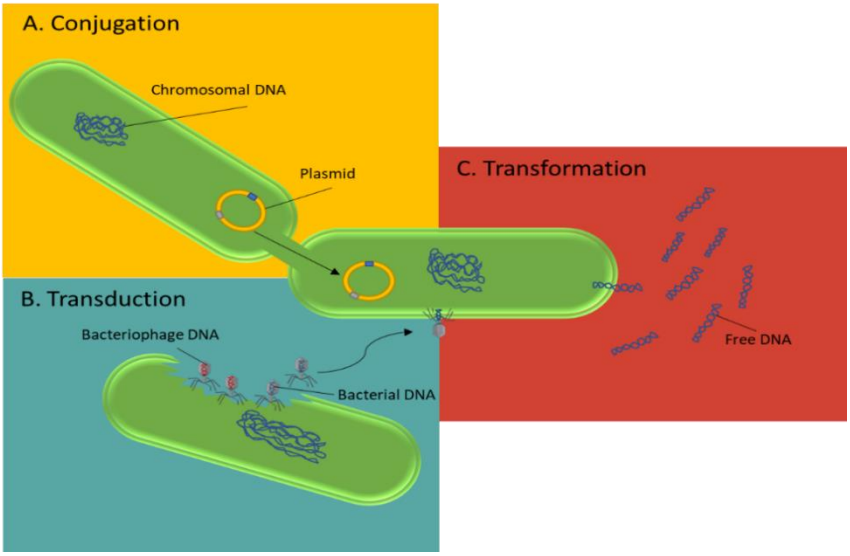


Figure 1.3: Mode of bacterial horizontal gene transfer, adapted from (von Wintersdorff et al., 2016)): A) Conjugation: Bacteria receive chromosomal DNA directly (cell to cell) from the donor bacteria (Griffiths AJF, 2000). B) Transduction: Exogenous genetic material is transferred through a bacteriophage from one bacteria to another (Britannica). C) Transformation: Foreign genetic material from the environment is taken up by bacteria (Gale, 2006).

Intrinsic Drug Resistance (IDR)

In addition to ADR, IDR is universally found within bacterial genomes. This type of resistance is not accredited to HGT, independent of antibiotic selective pressure, and may be due to the removal of drugs either by efflux pumps that are chromosomally encoded, inaccessibility into

bacterial cells, and/or presence of degrading enzymes (Cox and Wright, 2013, Kostyanev and Can, 2017).

Compared to Gram-positive bacteria, Gram-negative bacteria are intrinsically more resistant due to the presence of their OM, which functions as a permeability barrier to antibiotics, and expresses MDR efflux pumps to reduce intercellular drug concentrations (Nikaido, 1994, Zgurskaya et al., 2015). An example of IDR in Gram-negative bacteria is the MDR phenotype which causes these bacteria to be insensitive to many classes of antibiotics that are effective on Gram-positive bacteria (Nikaido, 1994). An example is the antibiotic vancomycin, which is used as treatment against Gram-positive Methicillin resistant S. aureus (MRSA), but is ineffective against Gram-negative bacteria since it is not able to penetrate the OM permeability barrier (Rice, 2012).

Selective pressure

Selective pressure is defined as a phenomenon which enhances the ability of micro-organisms to alter their behaviour and fitness within given environmental conditions. This is a driving force of evolution, as well as natural selection, whereby micro-organisms have the ability to survive environmental changes due to acquisition of new genetic material or due to spontaneous mutations (Tenover, 2001, Gale, 2006). The broad application of antimicrobial agents in and outside of hospitals, has increased the selective pressure for resistant organisms.

Resistance is generally associated with a decrease in bacterial fitness. It has been expected that the physiological cost for bacteria to maintain resistance genes in the absence of antibiotics, would be large enough to be substantially selected against. The latter has however not been generally observed for various reasons. Firstly, the fitness cost for bacteria to maintain resistance is not always large enough to select for loss of resistance alleles, therefore, even after removal of the drug, resistance can remain in the population for an extended period of time

(Sjolund et al., 2003, Andersson and Hughes, 2010). Compensatory mutations and/or genetic regulatory mechanisms can also compensate for the large fitness cost of resistance by only activating resistance in the presence of the drug (Nagaev et al., 2001). An example for the latter can be seen in *E. coli* where mutations in the *rpsL* gene, encoding for ribosomal protein S12, can confer resistance to streptomycin due to decreased rate of peptide elongation. Contrarily, mutations occurring in the *rpsD* and *rpsE* genes, which encode for ribosomal proteins S4 and S5 respectively, increases peptide elongation rate, thereby lowering the cost of resistance (Pelchovich et al., 2013, Levin et al., 2000). Lastly, resistance mutations can also provide the resistant strains with a fitness advantage by conferring increased virulence (Roux et al., 2015).

Drug induced resistance reversion

Since the occurrence of acquired drug resistance to multiple antimicrobial agents is increasing, monotherapy treatment is gradually becoming less adequate, necessitating the use of drug combination therapies. There exists a close temporal relationship between antibiotic use and the isolation of resistant bacteria. Bacteria can rapidly respond to environmental changes due to their short growth cycles, which enables them to rapidly evolve, adapt, and survive unfavourable conditions (Clatworthy et al., 2007). Currently, frequencies of outbreaks of drug resistant infections are increasing, whereas development of new antibiotic classes has dramatically slowed down over the past two decades (Spellberg et al., 2004). Drug induce reversion of antibiotic resistant pathogens into sensitive phenotypes is a prospective approach to target the mechanisms and evolution of bacterial resistance.

Strategies to combat antibiotic resistance with combination drug therapy has shown to be promising since the late 1940's, where coadministration of streptomycin and para-aminosalicylic acid showed reduced evolution of resistant *Mycobacterium tuberculosis* (Dunner et al., 1949). *Drug combinations are also currently being used in most cancer*

treatments (Lane, 2006, Bayat Mokhtari et al., 2017), *treatment of HIV infected patients* (Scourfield et al., 2011), *and the most effective treatment for malaria is artemisinin-based combinations* (Nosten and White, 2007).

Various mechanisms which reduces or inverts the selective advantage of antibiotic resistance has previously been studied. Administrating antibiotics with inhibitory molecules that inhibit specific resistance mechanisms is generally accepted as the most established approach to neutralize the evolutionary advantages of resistant strains. However, this does not necessarily create a competitive disadvantage and the relative prevalence of resistance within a bacterial population is not reduced. Negative selective pressure is required to reduce resistant strains in the population, even when the antibiotic is absent (Andersson and Hughes, 2010, Sjolund et al., 2005).

Drug interactions

Experimental and theoretical studies have recently indicated that the conflicting effects of antibiotics can be decoupled by combining specific compounds (Baym et al., 2016). There are various approaches where the evolutionary or physiological interactions between antibiotics can act to not only neutralize the selective advantage of resistance, but to pose a direct cost on it.

With notably few exceptions, bacteria that become resistant to a single drug still maintain the same sensitivity to other drugs as their drug-sensitive parents. Therefore, the shape of the growth curve for single-drug resistant bacteria will be similar in shape as for their drug-sensitive parents, except stretched towards higher drug concentrations over the axis of the drug to which the bacteria are resistant (*Figure 1.4-A*). There therefore exists a concentration regime where resistant bacteria can fully grow while the growth of sensitive bacteria is halted.

Nonetheless, there are three key approaches where drug combinations can be used to impose a direct cost on resistance, allowing selection against drug-resistant strains.

Suppressive drug interactions

Suppressive drug interactions, where drug A suppresses drug B, creates a nonmonotonic line of inhibition, and therefore a concentration regime which allows the inhibition of only resistant strains, whereas sensitive strains can continue to grow (*Figure 1.4-B*). The inhibitory line of bacteria that become resistance to drug A, is stretched further along the axis to higher concentrations of drug B, creating a concentration regime where the growth of bacteria resistant to drug A is suppressed, while the sensitive bacteria can continue to growth. Suppressive interaction can therefore, within certain concentration, invert the selective advantage of resistant bacteria (Baym et al., 2016).

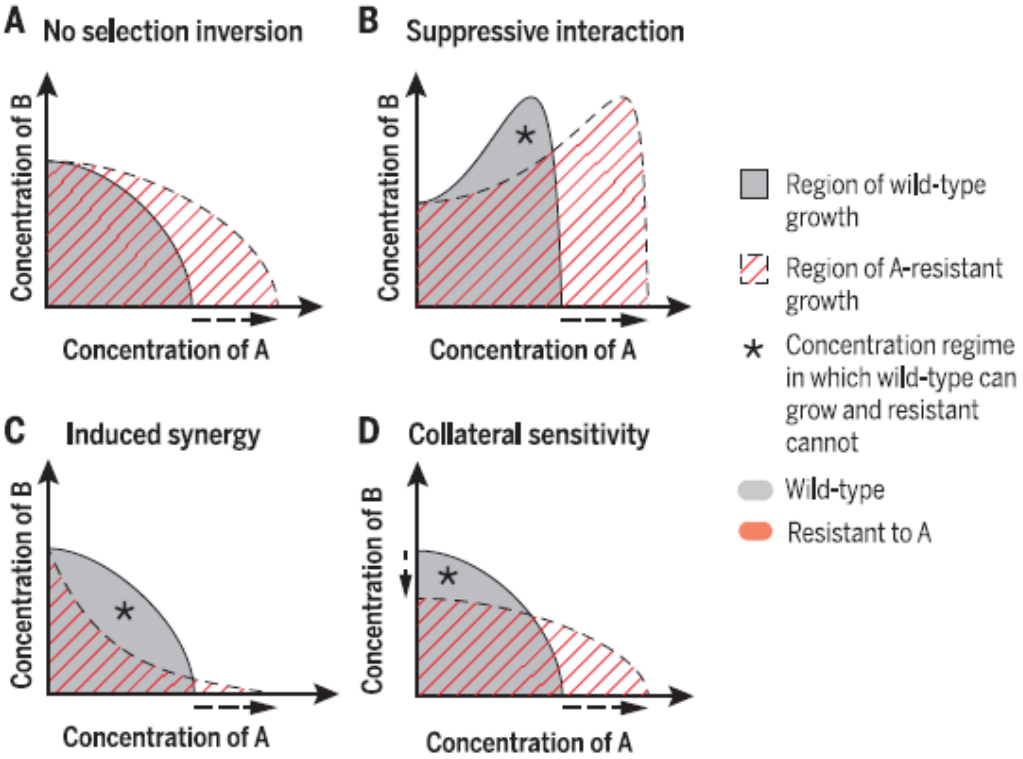


Figure 1.4: Approaches for selection inversion, adapted from (Baym et al., 2016). (A) The region of growth of drug A resistant bacteria covers the region of growth of the sensitive

bacteria completely. There no possible combination of the two drugs to create a region where sensitive bacteria can outgrow resistant bacteria. (B to D) Three primary approaches to establish a concentration regime (*) where sensitive bacteria can outgrow resistant bacteria.

Synergistic drug interactions

Drug interactions can also be synergistic, where the overall therapeutic effect of the combined drugs are larger compared to the sum of the effects on the individual components (Sucher, 2014). Synergistic interactions can avoid toxicity and slow down the development of drug resistance, since a lower drug dosage is required while the drug efficacy is increased/maintained (Chou, 2006). These drug interactions can cause a concentration regime where sensitive bacteria can outcompete resistant bacteria (*Figure 1.4-C*). However, this approach is challenging to apply to antimicrobial therapy due to the need for fine tuning of the drug concentrations and in currently only hypothetical (Baym et al., 2016).

Collateral sensitivity

Collateral sensitivity is a phenomenon in which bacteria within a two-drug concentration, that select for resistance to one drug, become more sensitive to the other (*Figure 1.4-D*) (Pal et al., 2015). This can either happen directly through mutations in drug target sites, or indirectly through mutations in important cellular mechanisms. Collateral sensitivity has been observed in DNA gyrase inhibiting drug combinations, where the drugs target the same proteins (Lázár et al., 2014). The use of this approach to select against mutations that lead to resistance has recently become a focus (Imamovic and Sommer, 2013, Oz et al., 2014, Baym et al., 2016).

Iodine-containing nanomolecular complex FS-1 as a potential inducer of antibiotic resistance reversion

It is widely recognized that acquired drug resistance is a major dilemma, which is only further complicated by the lack of new antibiotic development (Tenover, 2001). Tuberculosis (TB) still remains a major threat to public health, especially due to the wide distribution of multi-drug resistant bacteria. First line treatment antibiotics for TB are nearly fifty years old and new drugs has only recently been proposed, which include Bedaquiline and FS-1 (Mahajan, 2013). Resistance to Bedaquiline has been already been reported (Hoffmann et al., 2016).

Iodine-containing drugs are a viable option to overcome the given problem of antibiotic resistance, since they have strong antimicrobial properties and no acquired resistance to iodine has been reported (Murdoch and Lagan, 2013). Halogenation plays an essential part in organic reactions, where halogens can act as the electrophiles (Lewis acids) and organic compounds, which are electron-rich, as the nucleophiles (Lewis bases) (Moulay, 2013). Molecular iodine is an effectively mild and nontoxic Lewis acid catalyst (Dandia et al., 2014) that can easily penetrate the cells of micro-organisms and pass through the bi-lipid layers of cell membranes. Since iodine can easily penetrate cell membranes, its application is more efficient for infections caused by intracellular bacteria, including *Salmonella*, *Mycobacterium*, *Nocardia*, and *Listeria*. Iodine is known to have a high bioactivity and thus far no resistance development to iodine has been seen in bacteria and viruses (Ilin et al., 2017a).

FS-1 is a new antibacterial agent developed by Ilin et al. (2017b) et al., that induces the reversion of drug resistant pathogens, making them more sensitive to antibiotics. The development of FS-1 was intended to supplement treatment of infectious bacterial diseases, including nosocomial infections and multidrug resistant tuberculosis (MDR-TB) (Kalykova et al., 2016). FS-1 is a polymer complex, in the form of an aqueous solution, comprising coordinated molecular iodine atoms and metal ions, incorporated into a n alpha-dextrin and polypeptide moiety (Ilin et al., 2017a, Kalykova et al., 2016). FS-1 has passed both preclinical as well as clinical trials and was approved as a new anti-multidrug resistant (MDR)/ extensively

drug resistant (XDR) drug in Kazakhstan in 2015 (Ilin, 2014, Ilin et al., 2017b). However, the therapeutic mechanisms of FS-1 is still unclear.

In a study done by Ilin et al. (2017b) et al., they aimed at identifying genetic changes induced by FS-1 in bacterial populations. This study involved a laboratory experiment where animals, 105 guinea pigs, were infected with an XDR-TB strain. The animals were divided into five groups; 1. Stock culture; 2. positive control; 3. complex of anti-tuberculosis antibiotics (CAA); 4. CAA with low concentration FS-1; 5. CAA with high concentration FS-1, see *Table 1.1*. After 45 days of treatment, isolates were collected from each of the five groups. Increased susceptibility to antibiotics was observed in isolates collected from animals that received treatment of CAA supplemented with FS-1 when compared to isolates from the positive control group, and even more so when compared to those only treated with CAA (*Table 1.1*). After a period of 60 days, untreated animals died while animals subjected to the combinatorial treatment recovered. Animals that received only antibiotics, showed a transient effect and during the 30 days of recovery, the disease reoccurred.

Table 1.1 adapted from (Ilin et al., 2017b): List of antibiotics and antibiotic resistance (in percentages).

Antibiotics	Stock culture (%)	Group 2, positive control (%)	Group 3, CAA (%)	Group 4, CAA+FS-1 (2.5 mg/kg) (%)	Group 5, CAA+FS-1 (4.0 mg/kg) (%)
Isoniazid	100	100	100	75.0–83.3	75.0–80.0
Rifampicin	100	100	100	83.3–87.5	*50.0–60.0
Streptomycin	100	100	100	100	80.0–75.0
Pyrazinamide	80.0–92.0	83.3	100	*33.3–37.5	*20.0–25.0
Ethambutol	88.0–100	83.3	100	66.7–75.0	75.0–80.0
Amikacin	0.0–4.0	0.0	*16.7–20.0	0.0	0.0
Kanamycin	60.0–80.0	66.7–83.3	83.3–90.0	50.0	*20.0–25.0
Capreomycin	48.0–60.0	33.3–66.7	80.0–83.3	62.5–66.7	50.0–60.0
Ofloxacin	56.0–60.0	66.7	33.3	0.0–12.5	50.0–60.0
Cycloserine	16.0–20.0	33.3	66.7–70.0	0.0–25.0	0.0
Ethionamide	60.0–72.0	50.0	83.3–90.0	0.0–12.5	*0.0

In this study they hypothesized that drug resistance reversion was induced due to the therapeutic effect of the FS-1 drug. Their study demonstrated drug resistance reversion without any direct interaction between antibiotics and FS-1.

Model Bacteria

Staphylococcus aureus

Staphylococcus aureus, a cocci-shaped Gram-positive bacterial human pathogen, is responsible for a wide variety of clinical manifestations, including both community-acquired and nosocomial infections. Although these bacteria can colonize people asymptomatically (30% of people are carriers), these carriers are a source of spread and therefore still pose a major health threat to other individuals (Kluytmans et al., 1997). *Staphylococcus aureus* can be transmitted through direct contact with either an infected person, a carrier, or even a contaminated object (McConnell et al., 2013)(McConnell et al., 2013)(McConnell et al., 2013).

Although *S. aureus* can acquire resistance through HGT, other mechanisms, such as selection and chromosomal mutations, also have an important role (Stapleton and Taylor, 2002). Before the 1950s, *S. aureus* was treated by a β-lactam antibiotic, Benzylpenicillin. However, by the late 1950s resistance to this antibiotic already started to emerge. These bacteria become resistant by producing β-lactamases, which can inactivate the β-lactam ring in antibiotics. Later on, penicillin derivatives were synthesised to combat the resistance to β-lactams, leading to the synthesis of methicillin (Chambers, 1997). Unfortunately, not soon after the clinical use of methicillin, methicillin-resistant *S. aureus* (MRSA) strains developed, due to an alternative PBP2a, acquired from a different species, being expressed.

Over the years, various types of antibiotics were used to treat *S. aureus* infections, ultimately causing selective pressure, leading to multi-resistant MRSA strains, due to mutations in target protein coding genes (Livermore, 2000). A great challenge that we are currently facing, is that resistant strains are no longer only contained within hospitals, but modify and adjust to become

fit and virulent pathogens, which are able to arise within communities (Chambers and Deleo, 2009).

Escherichia coli

Escherichia coli is a Gram-negative bacterium, which colonizes the intestinal tract of both humans and animals. Many *E. coli* strains are not harmful and are essential for the synthesis of vitamin K, an important clotting factor. The rate antibiotic resistance in these bacteria have raised rapidly over the years, and surprisingly are acquired from within the community rather than hospitals (Laupland et al., 2008).

In *E. coli* the main mechanism of resistance to β-lactams is mediated by degrading enzymes, known as β-lactamases. Resistance to quinolones mainly involves altered target sites associated with mutations occurring in the resistance-determining regions of genes, such as *gyrA* and *parC*. Various gene have also been reported to be involved in tetracycline resistance, including *tetA* – *tetE*, and *tetI* among others. Resistance to aminoglycoside is due to aminoglycoside-modifying enzymes.

In *E. coli*, most of the adenines are methylated at the N6 position in the sequence *5'-GATC-3* by DNA adenine methyltransferase (dam). These sequences are mainly found in regions of the genome that are translated, with the exception of some tRNA genes and rRNA clusters. There are eleven *GATC* sites in the origin of replication (oriC) in *E. coli* (Bae et al., 2003). In a previous study were the importance of the epigenome of bacteria under antibiotic stress was explored, they observed that the *GATC* methylome provided bacteria with structural support to survive antibiotic stress. They found that *GATC* sites without adenine methylation servery compromised the survival of *E. coli* under antibiotic stress (Cohen et al., 2016).

The development of antibiotic resistance of *E. coli* in the community has rapidly increased. It is important to better understand the development of these resistance mechanisms, to successfully combat multidrug resistance infections.

Acinetobacter baumannii

Acinetobacter baumannii (*A. baumannii*) are Gram-negative coccobacilli, responsible for various infections, which include skin infections, urinary tract infections, as well as soft skin infections. According to the world health organization (WHO), *A. baumannii* is considered to be "one of the most serious ESKAPE organisms (*Enterococcus faecium*, *Staphylococcus aureus, Klebsiella pneumoniae*, *A. baumannii*, *Pseudomonas aeruginosa*, and Enterobacter species) that effectively escape the effects of antibacterial drugs" (Boucher et al., 2009).

Rapid development of antibiotic resistance has been reported for *A. baumannii*, and multidrug resistant variants have also been isolated (McConnell et al., 2013). Some of the known resistance mechanisms in *A. baumannii* include drug degrading enzymes, target alterations, efflux pumps, as well as permeability defects (Gordon and Wareham, 2010). *Acinetobacter baumannii* have outer membrane proteins, called porins, which assist in modulating cell permeability. One of the most abundant porins is a β-barrel porin, OmpA, that is a well-characterized virulence factor with biological properties of interest.

From the above discussions, it is clear that treatment options for multidrug resistant bacteria are becoming limited. Development of novel and rational strategies to combat the development of resistance is of great importance and are urgently needed.

This study set out to investigate the molecular mechanisms of drug induced resistance reversion of three different bacteria when cultivated on a medium containing a sub-lethal dose of the FS-1 drug. With the use of single molecule real-time (SMRT) PacBio sequencing and RNA

sequencing, epigenetic changes as well as regulation of gene expression could be studied. This allowed us to better understand the induced drug resistance reversion observed in the bacteria.

Objectives

- Assemble complete genome sequences for each of the strains in the negative control (NC) and experimental (FS) variants using long PacBio sequences.
- Annotate the genomes using the automated RAST server and then manually correct the annotations.
- Identify differentially expressed gene between NC and FS cultures after 5 min exposure of bacteria to FS-1.
- Identify differentially expressed genes between FS and NC after cultivation of bacteria over 10 passages with and without the drug, respectively.
- Determine metabolic pathways affected by differentially expressed genes in the experimental and negative control conditions.
- Investigate mechanisms of adaptation of bacteria to the presence of FS-1 through genetic and epigenetic variations in bacterial populations.
- Use Python scripts to create pipelines, which include various tools from the SMRT-link software, to work with the obtained SMRT sequencing data.

References

ABRAHAM, E. P., CHAIN, E., FLETCHER, C. M., FLOREY, H. W., GARDNER, A. D., HEATLEY, N. G. & JENNINGS, M. A. 1992. Further observations on penicillin. 1941. *Eur J Clin Pharmacol,* 42**,** 3-9.

ANDERSSON, D. I. & HUGHES, D. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol,* 8**,** 260-71.

BAE, S. H., CHEONG, H. K., CHEONG, C., KANG, S., HWANG, D. S. & CHOI, B. S. 2003. Structure and dynamics of hemimethylated GATC sites: implications for DNA-SeqA recognition. *J Biol Chem,* 278**,** 45987-93.

BARBOSA, T. M. & LEVY, S. B. 2000. The impact of antibiotic use on resistance development and persistence. *Drug Resist Updat,* 3**,** 303-311.

BAYAT MOKHTARI, R., HOMAYOUNI, T. S., BALUCH, N., MORGATSKAYA, E., KUMAR, S., DAS, B. & YEGER, H. 2017. Combination therapy in combating cancer. *Oncotarget,* 8**,** 38022-38043.

BAYM, M., STONE, L. K. & KISHONY, R. 2016. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science,* 351**,** aad3292.

BEUTLER, B. 2002. Review paper: LPS in microbial pathogenesis: promise and fulfilment. *Journal of Endotoxin Research,* 8**,** 329-335.

BLAIR, J. M., WEBBER, M. A., BAYLAY, A. J., OGBOLU, D. O. & PIDDOCK, L. J. 2015. Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol,* 13**,** 42-51.

BOUCHER, H. W., TALBOT, G. H., BRADLEY, J. S., EDWARDS, J. E., GILBERT, D., RICE, L. B., SCHELD, M., SPELLBERG, B. & BARTLETT, J. 2009. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis,* 48**,** 1-12.

BRITANNICA, T. E. O. E. *Transduction* [Online]. MICROBIOLOGY. Available: https://www.britannica.com/science/transduction-microbiology [Accessed 04/08/2019].

BROTZ-OESTERHELT, H. & BRUNNER, N. A. 2008. How many modes of action should an antibiotic have? *Curr Opin Pharmacol,* 8**,** 564-73.

BUSH, K., COURVALIN, P., DANTAS, G., DAVIES, J., EISENSTEIN, B., HUOVINEN, P., JACOBY, G. A., KISHONY, R., KREISWIRTH, B. N., KUTTER, E., LERNER, S. A., LEVY, S., LEWIS, K., LOMOVSKAYA, O., MILLER, J. H., MOBASHERY, S., PIDDOCK, L. J., PROJAN, S., THOMAS, C. M., TOMASZ, A., TULKENS, P. M., WALSH, T. R., WATSON, J. D., WITKOWSKI, J., WITTE, W., WRIGHT, G., YEH, P. & ZGURSKAYA, H. I. 2011. Tackling antibiotic resistance. *Nat Rev Microbiol,* 9**,** 894-6.

BUSH, L. M. & SCHMIDT, C. E. 2018. *Overview of Bacteria* [Online]. Merck Manual. [Accessed 15/08/2018].

CHAMBERS, H. F. 1997. Methicillin resistance in staphylococci: molecular and biochemical basis and clinical implications. *Clin Microbiol Rev,* 10**,** 781-91.

CHAMBERS, H. F. & DELEO, F. R. 2009. Waves of resistance: Staphylococcus aureus in the antibiotic era. *Nature reviews. Microbiology,* 7**,** 629-641.

CHAPOT-CHARTIER, M.-P. & KULAKAUSKAS, S. 2014. Cell wall structure and function in lactic acid bacteria. *Microbial Cell Factories,* 13**,** S9.

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

CHOU, T.-C. 2006. Theoretical Basis, Experimental Design, and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies. *Pharmacological Reviews,* 58**,** 621-681.

CLATWORTHY, A. E., PIERSON, E. & HUNG, D. T. 2007. Targeting virulence: a new paradigm for antimicrobial therapy. *Nature Chemical Biology,* 3**,** 541.

COHEN, N. R., ROSS, C. A., JAIN, S., SHAPIRO, R. S., GUTIERREZ, A., BELENKY, P., LI, H. & COLLINS, J. J. 2016. A role for the bacterial GATC methylome in antibiotic stress survival. *Nature genetics,* 48**,** 581-586.

COX, G. & WRIGHT, G. D. 2013. Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology,* 303**,** 287-292.

DANDIA, A., GUPTA, S. L. & MAHESHWARI, S. 2014. Molecular Iodine: Mild, Green, and Nontoxic Lewis Acid Catalyst for the Synthesis of Heterocyclic Compounds. *In:* AMETA, K. L. & DANDIA, A. (eds.) *Green Chemistry: Synthesis of Bioactive Heterocycles.* New Delhi: Springer India.

DELCOUR, A. H. 2009. Outer Membrane Permeability and Antibiotic Resistance. *Biochimica et biophysica acta,* 1794**,** 808-816.

DUNNER, E., BROWN, W. B. & WALLACE, J. 1949. The effect of streptomycin with para-amino salicylic acid on the emergence of resistant strains of tubercle bacilli. *Dis Chest,* 16**,** 661-6.

DŽIDIĆ, S., ŠUŠKOVIĆ, J. & KOS, B. 2008. *Antibiotic resistance mechanisms in bacteria: Biochemical and genetic aspects*.

GALE, T. 2006. *Selective Pressure* [Online]. Bookrags. Available: http://www.bookrags.com/research/selective-pressure-wob/#gsc.tab=0 [Accessed 2018/09/03].

GORDON, N. C. & WAREHAM, D. W. 2010. Multidrug-resistant Acinetobacter baumannii: mechanisms of virulence and resistance. *Int J Antimicrob Agents,* 35**,** 219-26.

GRIFFITHS AJF, M. J., SUZUKI DT, ET AL. 2000. Bacterial transformation. *Introduction to Genetic Analysis.* 7th ed.

HOFFMANN, H., KOHL, T. A., HOFMANN-THIEL, S., MERKER, M., BECKERT, P., JATON, K., NEDIALKOVA, L., SAHALCHYK, E., ROTHE, T., KELLER, P. M. & NIEMANN, S. 2016. Delamanid and Bedaquiline Resistance in Mycobacterium tuberculosis Ancestral Beijing Genotype Causing Extensively Drug-Resistant Tuberculosis in a Tibetan Refugee. *Am J Respir Crit Care Med,* 193**,** 337-40.

HUGHES, D. & KARLÉN, A. 2014. Discovery and preclinical development of new antibiotics. *Upsala journal of medical sciences,* 119**,** 162-169.

ILIN, A., BAHKYTZHAN, K. & G, Y. 2017a. *Action Mechanism of Molecular Iodine Complex with Bioorganic Ligands, Magnesium and Lithium Halogenides on Human Tuberculosis Strain With Multiple Drug Resistance*.

ILIN, A. I., AND KULMANOV, M. E. 2014. *Antibacterial Agent for Treating Infectious Diseases of Bacterial Origin*.

ILIN, A. I., KULMANOV, M. E., KOROTETSKIY, I. S., ISLAMOV, R. A., AKHMETOVA, G. K., LANKINA, M. V. & REVA, O. N. 2017b. Genomic Insight into Mechanisms of Reversion of Antibiotic Resistance in Multidrug Resistant Mycobacterium tuberculosis Induced by a Nanomolecular Iodine-Containing Complex FS-1. *Front Cell Infect Microbiol,* 7**,** 151.

IMAMOVIC, L. & SOMMER, M. O. A. 2013. Use of Collateral Sensitivity Networks to Design Drug Cycling Protocols That Avoid Resistance Development. *Science Translational Medicine,* 5**,** 204ra132-204ra132.

KAHNE, D., LEIMKUHLER, C., LU, W. & WALSH, C. 2005. Glycopeptide and lipoglycopeptide antibiotics. *Chem Rev,* 105**,** 425-48.

KALYKOVA, A., KUSTOVA, T., SAKIPOVA, Z., IBRAGIMOVA, N., ISLAMOV, R., VETCHÝ, D. & ILIN, A. 2016. Acute and subchronic toxicity studies of the original drug FS-1. *Acta Veterinaria Brno,* 85**,** 9-16.

KAPUR, V., LI, L. L., HAMRICK, M. R., PLIKAYTIS, B. B., SHINNICK, T. M., TELENTI, A., JACOBS, W. R., JR., BANERJEE, A., COLE, S., YUEN, K. Y. & ET AL. 1995. Rapid Mycobacterium species assignment and unambiguous identification of mutations associated with antimicrobial resistance in Mycobacterium tuberculosis by automated DNA sequencing. *Arch Pathol Lab Med,* 119**,** 131-8.

KLUYTMANS, J., VAN BELKUM, A. & VERBRUGH, H. 1997. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev,* 10**,** 505-20.

KOHANSKI, M. A., DWYER, D. J. & COLLINS, J. J. 2010. How antibiotics kill bacteria: from targets to networks. *Nature reviews. Microbiology,* 8**,** 423-435.

KOSTYANEV, T. & CAN, F. 2017. Chapter 1 - The Global Crisis of Antimicrobial Resistance. *In:* PULCINI, C., ERGÖNÜL, Ö., CAN, F. & BEOVIĆ, B. (eds.) *Antimicrobial Stewardship.* Academic Press.

KOTRA, L. P., HADDAD, J. & MOBASHERY, S. 2000. Aminoglycosides: Perspectives on Mechanisms of Action and Resistance and Strategies to Counter Resistance. *Antimicrobial Agents and Chemotherapy,* 44**,** 3249-3256.

LANE, D. 2006. Designer combination therapy for cancer. *Nat Biotechnol,* 24**,** 163-4.

LAUPLAND, K. B., CHURCH, D. L., VIDAKOVICH, J., MUCENSKI, M. & PITOUT, J. D. D. 2008. Community-onset extended-spectrum β-lactamase (ESBL) producing Escherichia coli: Importance of international travel. *Journal of Infection,* 57**,** 441-448.

LAXMINARAYAN, R. 2014. Antibiotic effectiveness: balancing conservation against innovation. *Science,* 345**,** 1299-301.

LÁZÁR, V., NAGY, I., SPOHN, R., CSÖRGŐ, B., GYÖRKEI, Á., NYERGES, Á., HORVÁTH, B., VÖRÖS, A., BUSA-FEKETE, R., HRTYAN, M., BOGOS, B., MÉHI, O., FEKETE, G., SZAPPANOS, B., KÉGL, B., PAPP, B. & PÁL, C. 2014. Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nature Communications,* 5**,** 4352.

LEVIN, B. R., PERROT, V. & WALKER, N. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics,* 154**,** 985-997.

LEVY, S. B. & MARSHALL, B. 2004. Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine,* 10**,** S122.

LIVERMORE, D. M. 2000. Antibiotic resistance in staphylococci. *Int J Antimicrob Agents,* 16 Suppl 1**,** S3-10.

LOWY, F. D. 2003. Antimicrobial resistance: the example of Staphylococcus aureus. *Journal of Clinical Investigation,* 111**,** 1265-1273.

MAHAJAN, R. 2013. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. *International Journal of Applied and Basic Medical Research,* 3**,** 1-2.

MCCLURE, N. S. & DAY, T. 2014. A theoretical examination of the relative importance of evolution management and drug development for managing resistance. *Proc Biol Sci,* 281.

MCCONNELL, M. J., ACTIS, L. & PACHON, J. 2013. Acinetobacter baumannii: human infections, factors contributing to pathogenesis and animal models. *FEMS Microbiol Rev,* 37**,** 130-55.

MOULAY, S. 2013. Molecular iodine/polymer complexes. *Journal of Polymer Engineering.*

MUNITA, J. M. & ARIAS, C. A. 2016. Mechanisms of Antibiotic Resistance. *Microbiology spectrum,* 4**,** 10.1128/microbiolspec.VMBF-0016-2015.

MURDOCH, R. & LAGAN, K. M. 2013. The role of povidone and cadexomer iodine in the management of acute and chronic wounds. *Physical Therapy Reviews,* 18**,** 207-216.

MUTO, C. A., JERNIGAN, J. A., OSTROWSKY, B. E., RICHET, H. M., JARVIS, W. R., BOYCE, J. M. & FARR, B. M. 2003. SHEA guideline for preventing nosocomial transmission of multidrug-resistant strains of Staphylococcus aureus and enterococcus. *Infect Control Hosp Epidemiol,* 24**,** 362-86.

NAGAEV, I., BJÖRKMAN, J., ANDERSSON, D. I. & HUGHES, D. 2001. Biological cost and compensatory evolution in fusidic acid-resistant Staphylococcus aureus. *Molecular Microbiology,* 40**,** 433-439.

NICOLAOU, K. C. & MONTAGNON, T. 2008. Molecules That Changed the World.

NIKAIDO, H. 1994. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science,* 264**,** 382-8.

NITISS, J. L. 2009. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer,* 9**,** 338-50.

NOSTEN, F. & WHITE, N. J. 2007. Artemisinin-based combination treatment of falciparum malaria. *Am J Trop Med Hyg,* 77**,** 181-92.

OZ, T., GUVENEK, A., YILDIZ, S., KARABOGA, E., TAMER, Y. T., MUMCUYAN, N., OZAN, V. B., SENTURK, G. H., COKOL, M., YEH, P. & TOPRAK, E. 2014. Strength of Selection Pressure Is an Important Parameter Contributing to the Complexity of Antibiotic Resistance Evolution. *Molecular Biology and Evolution,* 31**,** 2387-2401.

PAL, C., PAPP, B. & LAZAR, V. 2015. Collateral sensitivity of antibiotic-resistant microbes. *Trends Microbiol,* 23**,** 401-7.

PALMER, A. C. & KISHONY, R. 2014. Opposing effects of target overexpression reveal drug mechanisms. *Nat Commun,* 5**,** 4296.

PARK, J. T. & STROMINGER, J. L. 1957. Mode of Action of Penicillin. *Biochemical Basis for the Mechanism of Action of Penicillin and for Its Selective Toxicity,* 125**,** 99-101.

PELCHOVICH, G., SCHREIBER, R., ZHURAVLEV, A. & GOPHNA, U. 2013. The contribution of common rpsL mutations in Escherichia coli to sensitivity to ribosome targeting antibiotics. *Int J Med Microbiol,* 303**,** 558-62.

REECE, R. J. & MAXWELL, A. 1991. DNA Gyrase: Structure and Function. *Critical Reviews in Biochemistry and Molecular Biology,* 26**,** 335-375.

RICE, L. B. 2012. Mechanisms of resistance and clinical relevance of resistance to beta-lactams, glycopeptides, and fluoroquinolones. *Mayo Clin Proc,* 87**,** 198-208.

ROUX, D., DANILCHANKA, O., GUILLARD, T., CATTOIR, V., ASCHARD, H., FU, Y., ANGOULVANT, F., MESSIKA, J., RICARD, J. D., MEKALANOS, J. J., LORY, S., PIER, G. B. & SKURNIK, D. 2015. Fitness cost of antibiotic susceptibility during bacterial infection. *Sci Transl Med,* 7**,** 297ra114.

SADOWSKI, P. L., PETERSON, B. C., GERDING, D. N. & CLEARY, P. P. 1979. Physical Characterization of Ten R Plasmids Obtained from an Outbreak of Nosocomial Klebsiella pneumoniae Infections. *Antimicrobial Agents and Chemotherapy,* 15**,** 616-624.

SANDEGREN, L. & ANDERSSON, D. I. 2009. Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat Rev Microbiol,* 7**,** 578-88.

SAUVAGE, E., KERFF, F., TERRAK, M., AYALA, J. A. & CHARLIER, P. 2008. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS microbiology reviews,* 32**,** 234-258.

SCOURFIELD, A., WATERS, L. & NELSON, M. 2011. Drug combinations for HIV: what's new? *Expert Rev Anti Infect Ther,* 9**,** 1001-11.

SILHAVY, T. J., KAHNE, D. & WALKER, S. 2010. The bacterial cell envelope. *Cold Spring Harbor perspectives in biology,* 2**,** a000414-a000414.

SJOLUND, M., TANO, E., BLASER, M. J., ANDERSSON, D. I. & ENGSTRAND, L. 2005. Persistence of resistant Staphylococcus epidermidis after single course of clarithromycin. *Emerg Infect Dis,* 11**,** 1389-93.

SJOLUND, M., WREIBER, K., ANDERSSON, D. I., BLASER, M. J. & ENGSTRAND, L. 2003. Long-term persistence of resistant Enterococcus species after antibiotics to eradicate Helicobacter pylori. *Ann Intern Med,* 139**,** 483-7.

SPELLBERG, B., POWERS, J. H., BRASS, E. P., MILLER, L. G. & EDWARDS, J. J. E. 2004. Trends in Antimicrobial Drug Development: Implications for the Future. *Clinical Infectious Diseases,* 38**,** 1279-1286.

STAPLETON, P. D. & TAYLOR, P. W. 2002. Methicillin resistance in Staphylococcus aureus: mechanisms and modulation. *Science progress,* 85**,** 57-72.

SUCHER, N. J. 2014. Searching for synergy in silico, in vitro and in vivo. *Synergy,* 1**,** 30-43.

TENOVER, F. C. 2001. Development and Spread of Bacterial Resistance to Antimicrobial Agents: An Overview. *Clinical Infectious Diseases,* 33**,** S108-S115.

TENOVER, F. C. & HUGHES, J. M. 1996. The challenges of emerging infectious diseases: Development and spread of multiply-resistant bacterial pathogens. *JAMA,* 275**,** 300-304.

VOLLMER, W., BLANOT, D. & DE PEDRO, M. A. 2008. Peptidoglycan structure and architecture. *FEMS Microbiol Rev,* 32**,** 149-67.

VON WINTERSDORFF, C., PENDERS, J., M. VAN NIEKERK, J., DOMINIC MILLS, N., MAJUMDER, S., VAN ALPHEN, L., SAVELKOUL, P. & WOLFFS, P. 2016. *Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer*.

WAKSMAN, S. A. & FLYNN, J. E. 1973. History of the word 'antibiotic'. *J Hist Med Allied Sci,* 28**,** 284-6.

WOODFORD, N. & ELLINGTON, M. J. 2007. The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection,* 13**,** 5-18.

ZGURSKAYA, H. I., LOPEZ, C. A. & GNANAKARAN, S. 2015. Permeability Barrier of Gram-Negative Cell Envelopes and Approaches To Bypass It. *ACS Infect Dis,* 1**,** 512-522.

CHAPTER 2

Complete genome sequencing and annotation of bacterial genomes

Abstract

Over the years, NGS technologies have become an attractive choice to generate vast amounts of genomic data, making whole-genome sequencing projects routine procedures. However, due to their short-read length, NGS still have some limitations. This resulted in many unfinished daft genomes. With the use of long reads generated by the newest sequencing technologies, SMRT and nanopore sequencing, these limitations can be overcome, and genomes can more easily be assembled into complete reference genomes. For this study, multidrug resistant reference cultures *S. aureus*, *E. coli*, and *A. baumannii* obtained from the ATCC, were cultivated over 10 passages on a medium containing the FS-1 drug (FS culture) and on normal medium (NC culture). Thereafter, DNA was extracted and sequenced using PacBio SMRT sequencing. Long reads were generated with coverages ranging between 600x and 1000x of the expected genome lengths. Tools available in the SMRT-link software were used to do *de novo* assemblies for bacterial genomes. The genomes were annotated using the RAST server, followed by manual editing. Complete bacterial genomes were obtained for all strains and submitted to the NCBI. Genome announcements were published for all three bacteria.

Introduction

With recent advances in next-generation sequencing (NGS) technologies, generating reads that are highly accurate can be done in a short time and at a low-cost. This led to numerous genomes being sequenced, generating vast amounts of genomic data, and changed whole-genome sequencing projects into routine procedures. Although significant improvements in high-throughput NGS technologies, such as Illumina and Roche 454, has led to a more cost effective and comprehensive way of studying many new genomes, there are still some limitations, especially with their short read lengths, that make particular biological problems, such as genome assemblies challenging (Schadt et al., 2010).

With the short-read lengths generated by NGS technologies, microbial genomes can usually be resolved into high-quality draft assemblies. Repetitive regions are abundant in some microbial genomes and pose the greatest challenge for assemblers, especially when the read lengths are shorter than the repetitive regions (Treangen and Salzberg, 2011, Brown et al., 2014). Accurate *de novo* assemblies are challenging with short read lengths, resulting in fragmented genomes with ambiguous regions (Schatz et al., 2010, Alkan et al., 2011). During short-read library preparation, DNA amplification can also introduce chimeric reads (Guan and Sung, 2016).

Genome assembly tools do not produce perfectly aligned genomes, they are designed to rather generate a manageable number of contigs from the sequencing reads, ready for finishing. Finishing a genome is a process that involves assembling a set of contigs or scaffolds into complete sequences, e.g. closing all the gaps; doing error correction; and verifying all low coverage regions (Phillippy et al., 2008). Since this can take up to several months to complete and due to the required effort, complete genome finishing is rarely achieved. This has resulted in the submission of many unfinished draft assemblies, which are fragmented and may contain

regions that are misassembled, have incorrect gene calls, or other artefacts (Phillippy et al., 2008, Treangen and Salzberg, 2011).

With third generation sequencing (TGS) technologies, where single molecules can be sequenced in real time without having to pause between read steps, many of the discussed limitations can be overcome (Rhoads and Au, 2015). Although long reads are more error prone compared to short reads, various error correction algorithms can be implemented to reduce the error rates associated with long read lengths.

Pacific Biosciences developed single molecule real time (SMRT) sequencing, which generates longer reads and consist of two main technological components, a zero-mode waveguide (ZMW), and fluorescently labelled phospho-linked nucleotides. The ZMW is a nanostructure with holes (~ 100 nm in diameter), that are fabricated in a metal cladding, such as aluminium or gold, and deposited on a transparent floor (Foquet et al., 2008). ZMWs serve as nanophotonic chambers to visualize single polymerization reactions and provide the smallest available volume for light detection (Eid et al., 2009). Sequencing-by-synthesis sequencing approaches mostly utilize base-labelled nucleotides, in which the fluorophores are directly attached to the bases. This causes fluorophores to permanently become part of the DNA strand as each nucleotide is incorporated. This labelling approach cannot be used for SMRT sequencing, since the build-up of multiple dye molecules in the growing DNA chain can lead to steric hindrance, which limits enzymatic activity, as well as increase background noise. Alternatively, SMRT sequencing uses phospho-linked nucleotides, since during the nucleotide incorporation, the phosphate chains are cleaved .

SMRT sequencing generates single-pass reads which are more error-prone, having a median error rate of about 11%, and are mainly either deletions or insertions. Pacific Biosciences developed a mapping tool, BLASR (2018a), which takes into account all of these characteristics of SMRT sequencing and has been optimized to allow confident mapping of

reads generated with SMRT technology. A consensus sequence is generated through 'vertically' averaging sequencing information for each reference position. They also developed a consensus tool, called Quiver, that builds high-quality consensus sequences (2018b).

SMRT sequencing achieves higher consensus accuracy than other sequencing methods due to the random nature of the errors, as shown in *Figure 2.1*. In various studies SMRT sequencing has been used to validated SNPs that have been previously discovered with other platforms (Carneiro et al., 2012, Pugh et al., 2012). This highlights the importance of consensus accuracy as opposed to single-pass error rates.

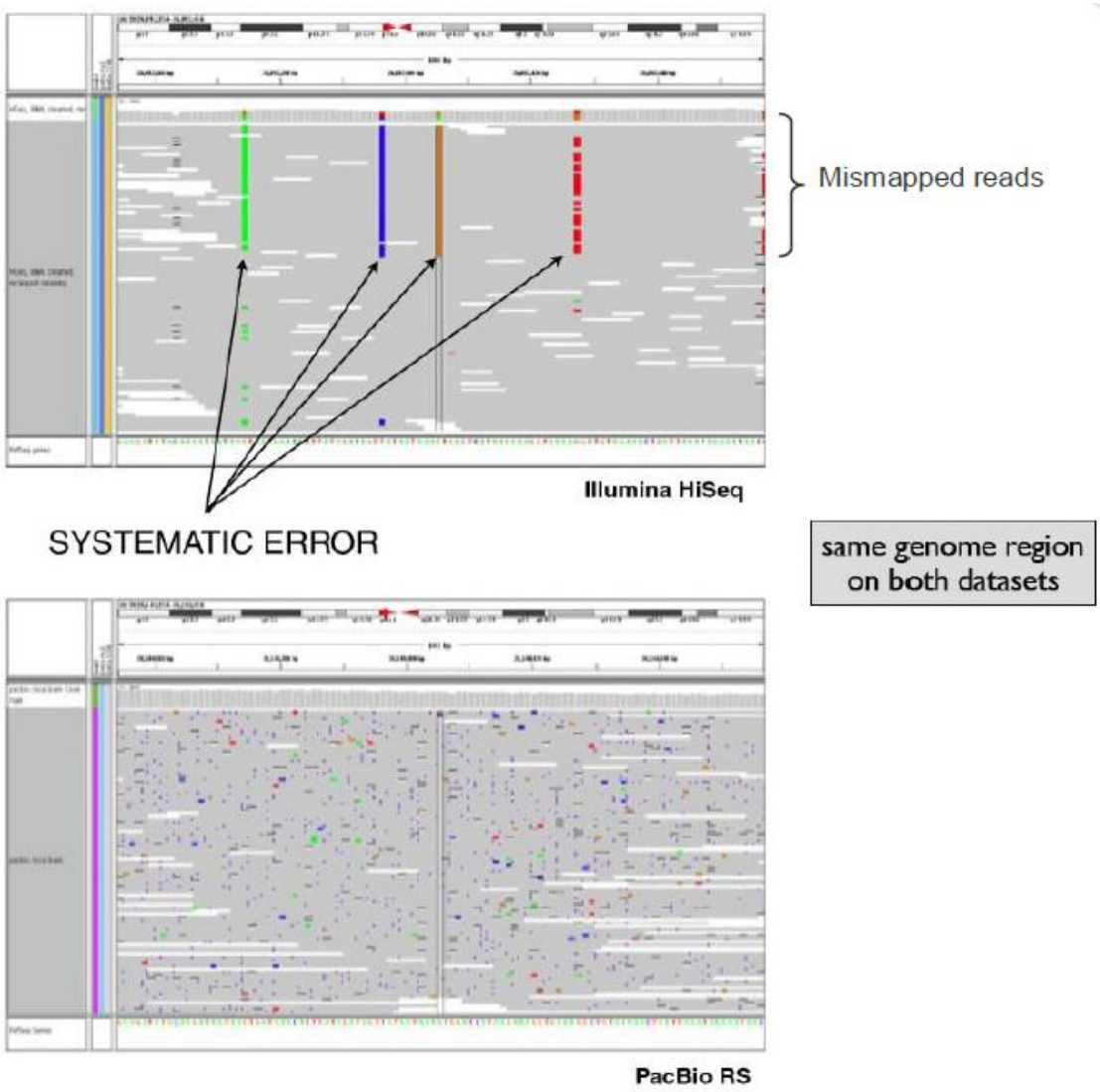


Figure 2.1, adapted from (Carneiro et al., 2012): Representation of systematic errors in Illumina HiSeq reads vs the random nature of errors in PacBio RS reads.

The length of sequence reads is important, since it can directly affect the accuracy of read assembly. Sequence reads can still be uninformative and ambiguous if they cannot be correctly mapped to a reference genome, even if they are 100% accurate. When a read is too short to span repetitive regions, its origin cannot be determined unambiguously, meaning that any variation observed, will be ambiguous with respect to where it occurred in the genome.

This chapter aimed to use long reads generated with PacBio sequencing, together with tools available in the SMRT-link software, to assemble and annotate complete bacterial genomes for the multidrug resistant reference bacteria *Staphylococcus aureus*, *Escherichia coli*, and *Acinetobacter baumannii* obtained from ATCC.

Materials and methods

It should be noted that all experimental procedures with bacterial cultures were performed by our collaborators in this project from the Scientific Center for Anti-Infectious Drugs (SCAID), Almaty, Kazakhstan. These methods are mentioned here to better represent the overall experimental design of the project. All bioinformatics procedures aimed at complete genome sequence assembly, annotation and analysis were performed in the Centre for Bioinformatics and Computational Biology (CBCB) at the University of Pretoria.

Sample collection and bacterial cultivation

Multidrug resistant strains *Staphylococcus aureus* ATCC BAA-39, *Escherichia coli* ATCC BAA-196, and *Acinetobacter baumannii* ATCC BAA-1790 were obtained from the American Type Culture Collection (ATCC) and used as model organisms for this study. All bacteria were inoculated into test-tubes that contained 10 ml of the Mueller-Hinton Broth medium. For experimental conditions (denoted as FS), the medium was supplemented with the FS-1 drug (1/2 of the minimal bactericidal concentration [MBC] experimentally determined for each bacterial culture), and for control conditions (denoted as NC), the same medium was used without the supplementation of the drug. The test-tubes were incubated at 37°C for 24 h, thereafter 0.1 ml aliquots of the cultures were transferred to fresh tubes with the corresponding media. After 10 passages, experimental and control samples were collected for *E. coli* and *A. baumanni*, whereas for *S. aureus*, the experimental and control bacteria were cross-inoculated into tubes with drug-containing and drug-free media for further overnight incubation, in 3 repeats (*Figure 2.2*).

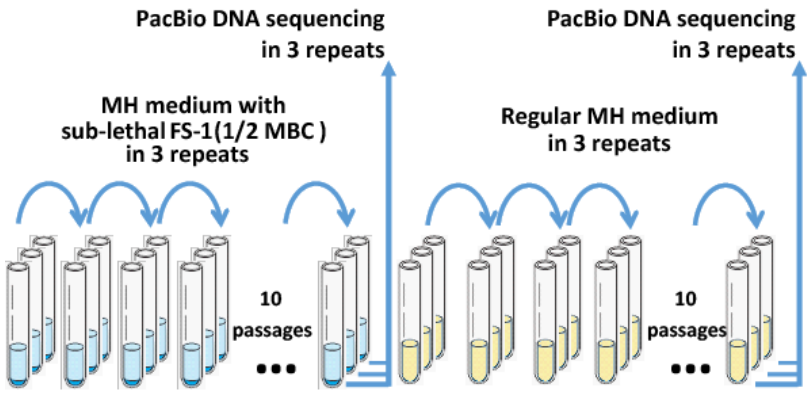


Figure 2.2: Representation of the experimental scheme for the cultivation of the reference bacterial strains in a medium containing the FS-1 drug (experimental samples) and in regular medium without the drug (control samples). After 10 passages, *E. coli* and *A. baumannii* cultures were sequenced, whereas for *S. aureus*, experimental and control bacteria were further cultivated in medium containing FS-1 and normal medium.

DNA extraction and PacBio sequencing

DNA extraction was performed with the PureLink Genomic DNA Kits (Publication Number: MAN0000601, Revision 2.0). SMRTbell templates of 10kb were prepared from the samples and sequenced on the PacBio RS II sequencer.

Genome assembly and annotation

The long PacBio reads generated for the two genome variants, NC and FS, were used to do a *de novo* assembly using the SMRT-link pipelines. The SMRT-link software makes use of the hierarchical genome assembly process (HGAP) for genome assemblies. HGAP proceeds in 2 main steps; Seed read construction and pre-assembly (Greg Concepcion, 2016) (see *Figure 2.3*).

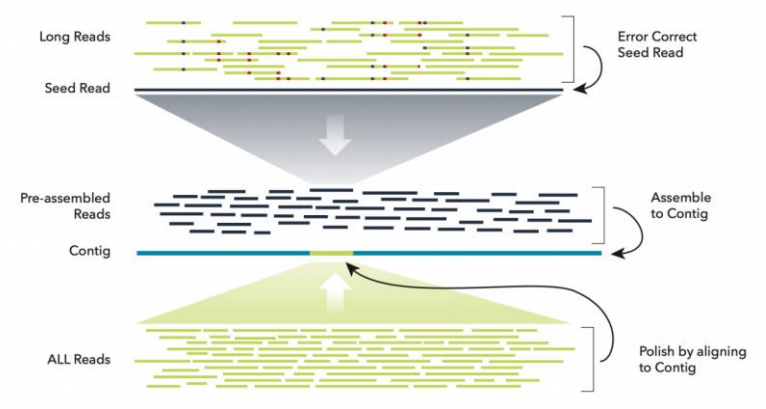


Figure 2.3: Illustration of steps involved in the HGAP assembly used in the SMRT-link software.

Firstly, accurate consensus sequences are generated by selecting the longest reads in the dataset, called seed reads, and then aligning shorter reads against them. Thereafter, the pre-assembled reads are assembled into contigs.

The raw PacBio sequences for this study were received in *BAX.H5* format and stored on our local server at the University of Pretoria. The SMRT-link software v5.0.1 (https://www.pacb.com/support/software-downloads/) was installed on our local server and various tools available in the software were used. All tools and pipelines were used with standard (default) parameters, since on the PacBio website they state that "for most microbes, closed genomes with accessory plasmids can be assembled with 50-fold coverage of SMRT Sequencing data using the default settings of our assembly pipeline" . Firstly, "HdfSubreadSets" were created using the "dataset create --type HdfSubreadSet" tool and command. The raw PacBio reads were used as input, and this generated *XML* files. These *XML* files were then converted to "subreadsets", which could easily be manipulated in further downstream analyses, using the "pbsmrtpipe pipeline-id pbsmrtpipe.pipelines.sa3_hdfsubread_to_subread" pipeline. These "subreadsets" were then used as input for the de novo assembly pipeline. The pipeline used was the "polished_falcon_fat" pbsmrtpipe pipeline. The *de novo* assembly pipeline makes use of

HGAP v4, with a default minimum seed length of 6 kb. The full commands used for the assemblies are available in *Supplementary File 1*.

To evaluate the completeness of the bacterial genomes, the Benchmarking Universal Single-Copy Orthologous (BUSCO) software (Simao et al., 2015) was used. The assembled reference genomes were used as input. The genomes were then automatically annotated using the Rapid Annotations using Subsystems Technology (RAST) Server (Aziz et al., 2008), and thereafter manually corrected.

The program SeqWord Genomic Island Sniffer (Bezuidt et al., 2009) was used to determine locations of horizontally transferred genomic islands, as well as to identify the replication origin and terminus on the bacterial chromosomes, with the use of GC-skews between leading and lagging strands.

Variant calling

PacBio read alignments against genome assemblies stored in *SAM* files were used for identification of genetic polymorphisms using the program bcftools. The command below was used to do variant calling:

$ bcftools mpileup -f reference.fa alignments.bam | bcftools call -mv -Ob -o calls.bcf

The first "mpileup" part generates genotype likelihoods at each genomic position with coverage. The second part makes the actual calls. The "-m" option tells the program to use the default calling method, the "-v" option outputs only variant sites, finally the "-O" option selects the output format. In this example we choose binary compressed BCF files, which is the optimal starting format for further processing, such as filtering.

Results

The results discussed below for *S. aureus*, *E. coli*, and *A. baumannii* have been published in the *Microbiology Resource Announcements* journal. All the genomes have been submitted to the NCBI; accession numbers are listed in *Table 2.2* for each of the genomes.

Complete genome assembly and annotation

The bacteria used in this study were cultivated in test-tubes containing a medium supplemented with a sub-lethal concentration of the FS-1 drug. In parallel, bacterial cultures were also cultivated on the same medium without FS-1, to serve as a negative control (NC). After the bacteria were cultivated, genomic DNA samples were collected in three repeats and sequenced using the PacBio RS II sequencer. The long PacBio reads with coverages ranging between 600x and 1000x were generated for this study and used to do *de novo* genome assemblies of the NC and FS variants of the bacteria. Since we had such high coverage for all genomes, complete genomes were obtained with default parameters used for all SMRT-link pipelines. Genomes were then automatically annotated using RAST (Aziz et al., 2008) followed by manual corrections and verification by aligning Ion Torrent RNA reads to the genomes. The completeness of the final assemblies was evaluated using the BUSCO software. The obtained result for each of the bacteria are discussed below.

General statistics of PacBio reads, quality control and filtering

Below some general statistics are described and illustrated for the PacBio reads used for the assembly of *S. aureus* genomes, NC and FS. These statistics were generated for all bacterial genomes, however, they will not be discussed in detail below to avoid repetitiveness. Statistics for *E. coli* and *A. baumannii* can be seen is Supplementary file 2.

In *Table 2.1* below, some general statistics about PacBio reads generated for *S. aureus* NC and FS are shown. For the NC genome, a total of 2 922 373 632 bases were generated, 122 578

reads, with an average read length of 12 900 bp. For FS, 3 152 009 426 bases were generated, 323 376 reads, and a read length of 14 315. Although the FS genome had more bases, reads, and a larger read length compared to NC, the NC genome had a higher average polymerase read quality of 0.852 compared to FS, which had a polymerase read quality of 0.850. The polymerase read quality is defined as "a trained prediction of a read's mapped accuracy based on its pulse and base file characteristics" .

Table 2.1: General statistics of PacBio reads generated for NC and FS S. aureus genomes.

Cell Name	Polymerase Read Bases	Polymerase Reads	Polymerase Read N50	Polymerase Read Length	Polymerase Read Quality
NC1	629,785,144	45,729	22,607	13,771	0.855
NC2	1,529,903,991	102,380	23,962	14,942	0.846
NC3	762,684,497	76,367	17,630	9,986	0.856
FS1	779,242,292	55,003	22,652	14,167	0.852
FS2	1,360,199,488	99,414	23,462	13,681	0.845
FS3	1,012,567,646	67,061	26,878	15,098	0.854
TOTAL	6,074,383,058	445,954	22,865	13,608	0.851

Figure 2.4 shows the distribution of the mapped concordance vs read lengths. As can be seen, there are higher concordance in shorter reads, whereas with longer reads the concordances decreases.

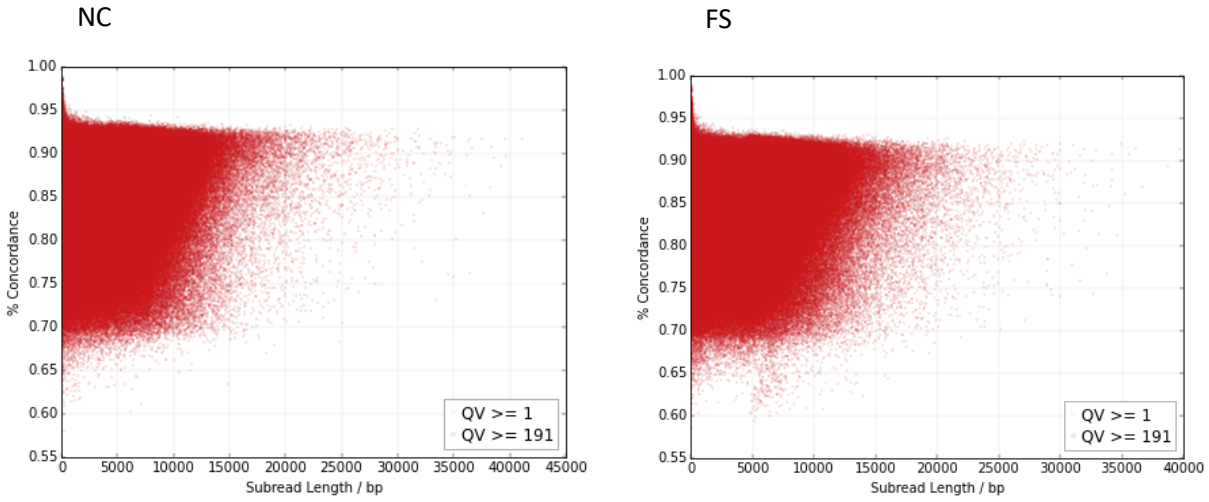


Figure 2.4: Mapped concordance vs read length of PacBio reads generated for *S. aureus* NC and FS genomes.

In *Figure 2.5*, the read lengths vary from exceptionally long reads, over 40 000 bp, to shorter reads of only a few hundred base pairs. As seen in these figures, there were significantly fewer extremely long reads generated, compared to the shorter reads.

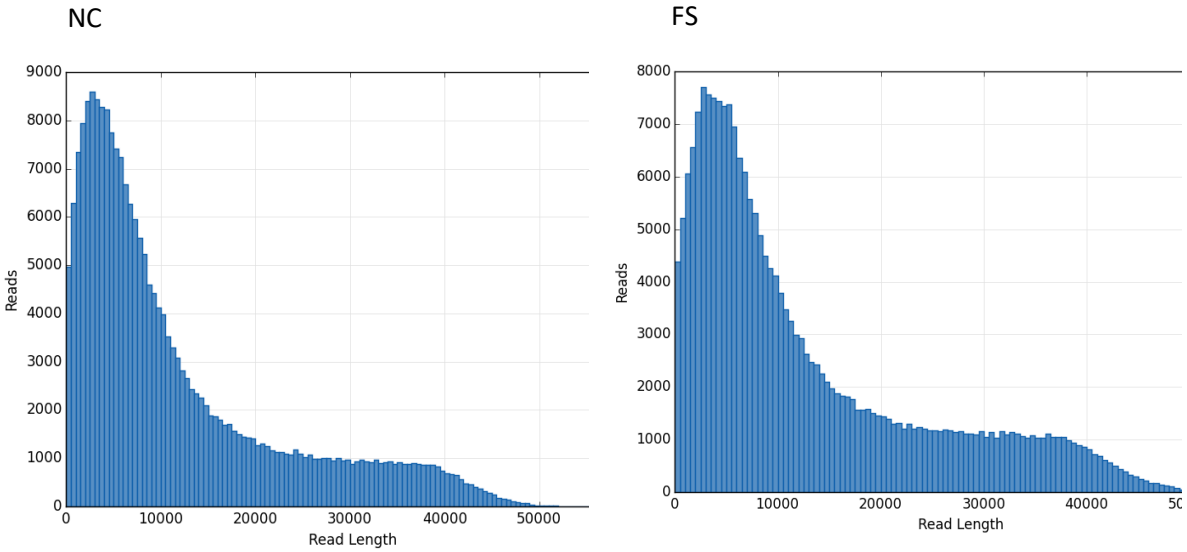


Figure 2.5: Distribution of read lengths of PacBio reads generated for *S. aureus* NC and FS genomes.

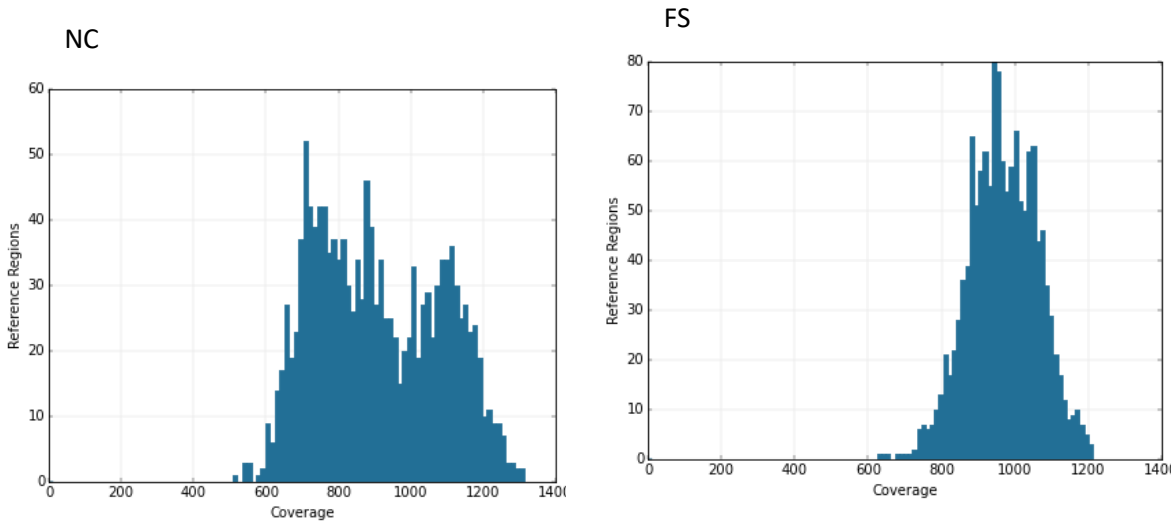


Figure 2.6: Histogram representing coverage of reference regions in *S. aureus* NC and FS genomes.

As illustrated in *Figure 2.6* above, both the NC and FS genomes of *S. aureus* had significantly high coverages of between 600x to 1200x. For NC, over 50% of the genome had a coverage of approximately 700x. For FS, over 80% of the genome had a coverage of more than 900x.

Assembly and annotation of complete genomes of *S. aureus* ATCC BAA-39

The *MRSA* ATCC BAA-39 strain, isolated from a nasal clinical sample, has previously been published in the NCBI (2010) and comprised of 83 contigs (BioProject accession number PRJNA50533). Long PacBio sequences were generated for this study and had high genome coverages of 966x and 906x, for the FS and NC variants respectively.

The assembled genome variants were obtained as single-contig sequences with no gaps or ambiguities. The length of the NC and FS variants were 2 791 218 bp and 2 792 888 bp (*Figure 2.7*), respectively, with an average GC content of 32.9%. The obtained genome lengths were shorter compared to the previous assembly. It may be explained by a possible redundancy of genomic sequences in the multiple contigs of the previous assembly based on relatively short reads generated by Roche 454. The two genome variants were published at the NCBI under the BioProject accession number PRJNA480363. No plasmids were identified for either the NC or FS variants.

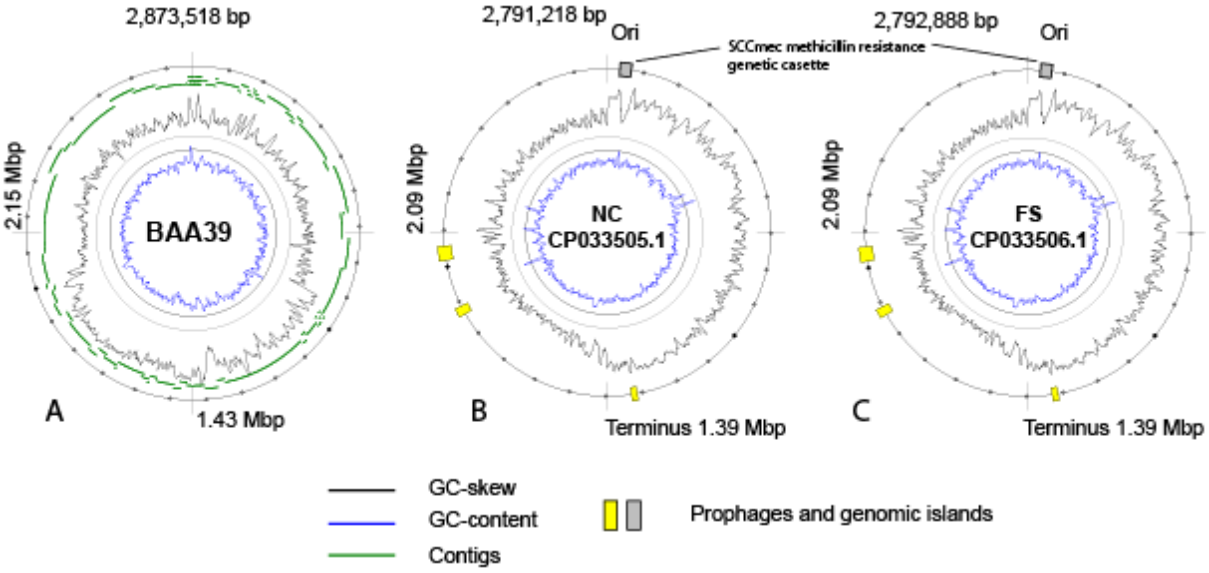


Figure 2.7: Circular genomic maps of *S. aureus* genomes. (A) Previous assembly which had 83 contigs and a "messy" GC-skew. (B) Newly assembled NC genome obtained as a single

contig with an improved GC-skew. (C) Newly assembled FS genome also obtained as a single contig and an improved GC-skew.

Figure 2.8 represents a phylogenetic tree of *S. aureus* sequences published at the NCBI. Our strains, ATCC BAA-39, are highlighted in red. The closest related strains to our strains are strains from the National Collection of Type Cultures (NCTC), NCTC4137 and NCTC4163. These stains were both published in the NCBI in 2018, under BioProject number PRJEB6403, and were obtained as single contigs with genome sizes of 2 750 707 bp and 2 759 837 bp, for NCTC4137 and NCTC4163 respectively.

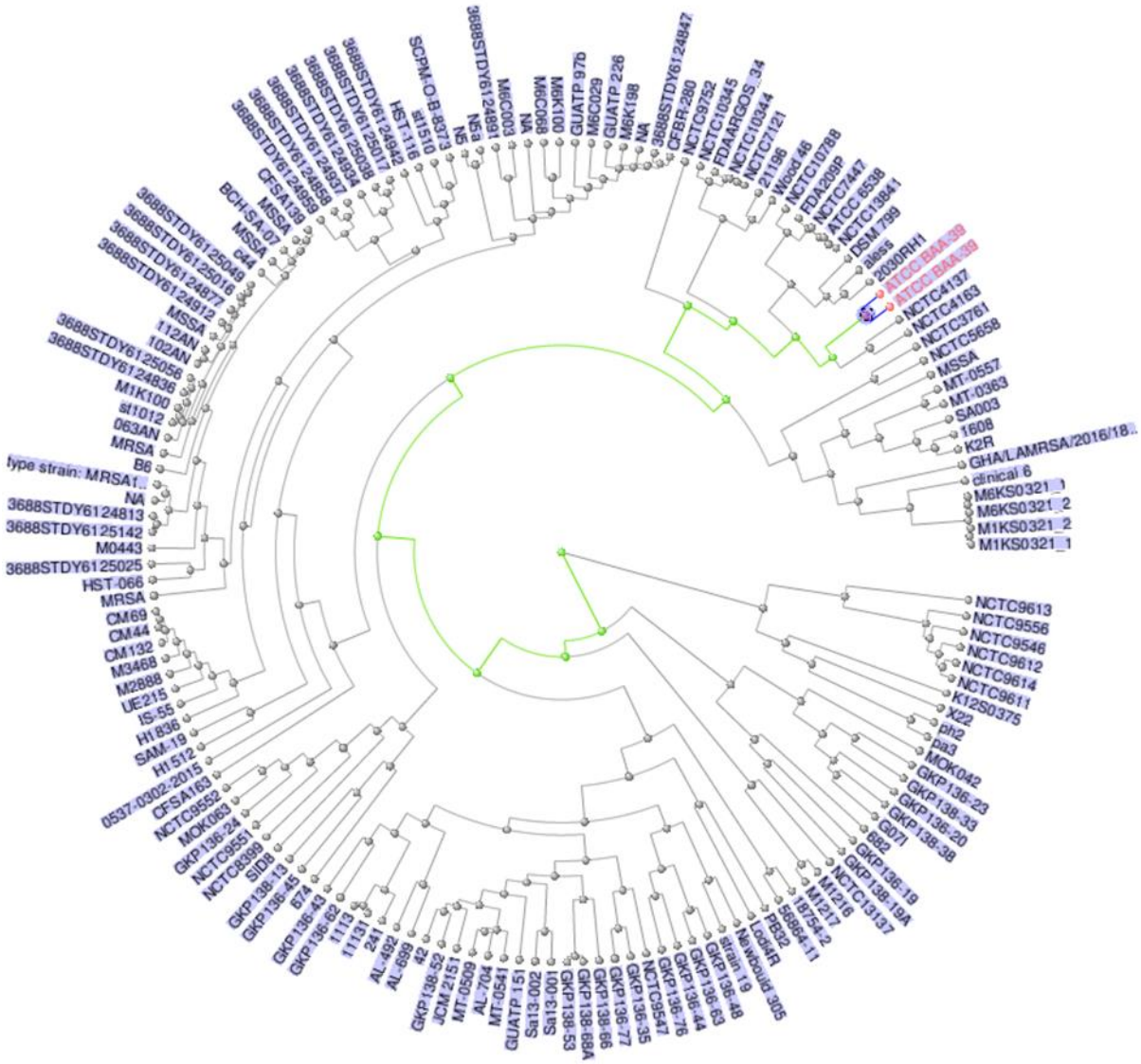


Figure 2.8: Phylogenetic tree of *S. aureus* sequences available on the NCBI. Sequences obtained and published during this study are highlighted in red.

When analysing the annotated genomes, we observed frameshift mutations, where neighbouring genes were annotated as parts of the same longer protein. The use of variant calling allowed us to identify polymorphic sites in the genomes. In *Figure 2.9*, the distribution of the polymorphic sites, as well as frameshift mutations are illustrated. More frameshift mutations were observed in the FS-1 treated cultures as well as different distribution of polymorphic sites.

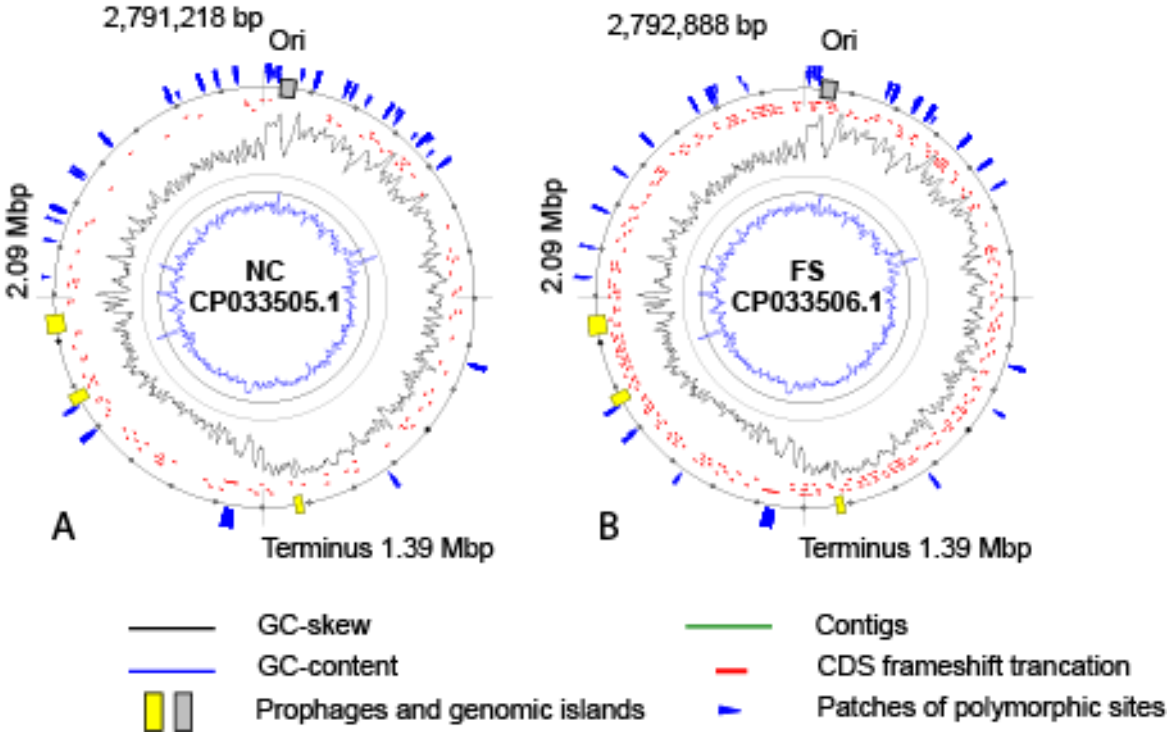


Figure 2.9: Distribution of polymorphic sites and CDS frameshift truncations in the NC (A) and FS (B) genomes of *S. aureus*. Blue triangles represent patches of polymorphic sites and red lines represent CDS frameshift truncations.

Assembly and annotation of complete genomes of *E.coli ATCC BAA-196*

The strain *E. coli* ATCC BAA-196 was selected in 1988 in Boston from a number of clinical isolates showing an extended spectrum of resistance to β-lactam antibiotics and identified as *Klebsiella pneumonia* (Rice et al., 1993). This strain was not previously assembled and published. The long PacBio reads generated for our research had genome coverages of 691x and 701x for the FS and NC variants, respectively. The genome variants were assembled into singe contig sequences with no gaps or ambiguities. The length of the two obtained variants were 4 682 561 bp for NC and 4 682 572 bp for FS (*Table* 2.2). One plasmid was found in the NC variant, with a length of 266 396 bp, and three plasmids were found in the FS variant, with lengths of 279 992 bp, 44 240 bp, and, 11 153 bp. Horizontally transferred genomic islands (*Figure 2.10*) were identified by SeqWord genomic Island Sniffer software (Bezuidt et al., 2009).

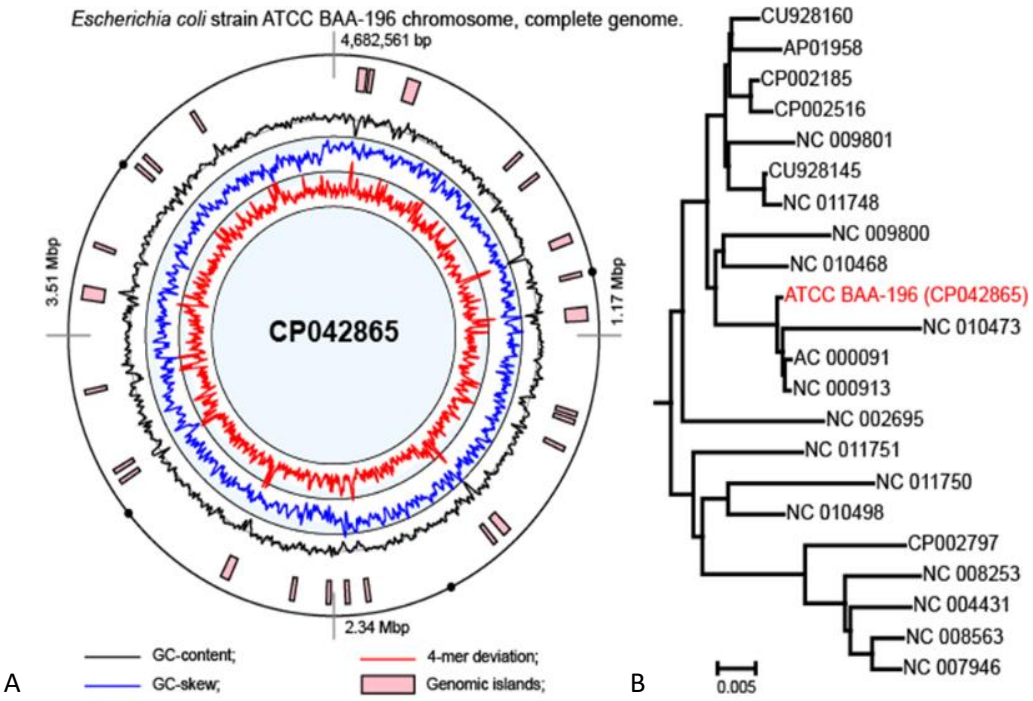


Figure 2.10: Circular map of the chromosome for the NC variant. (A) Predicted genomic islands and histograms of the GC-content, GC-skew and tetranucleotide pattern deviations

calculated in an 8 kbp sliding window stepping 2 kbp. (B) A Neighbour-Joining phylogenetic tree of the sequenced BAA-196 strain and other reference *E. coli* genomes.

Assembly and annotation of complete genomes of *A. baumannii* ATCC BAA-1790

The multidrug resistant *A. baumannii* ATCC BAA-1790 strain was isolated from a sputum sample in 2008. For *A. baumannii*, the FS genome could not be assembled due to sample contamination. The long reads generated for the NC variant, had a genome coverage of 647x. The NC genome obtained had a length of 3 963 235 bp (*Table 2.2*) and one plasmid with a length of 67 023bp. Genomic islands were identified by SeqWord genomic Island Sniffer software (F*igure 2.11*). The identified genomic islands contained genetic determinants associated with antibiotic resistance.

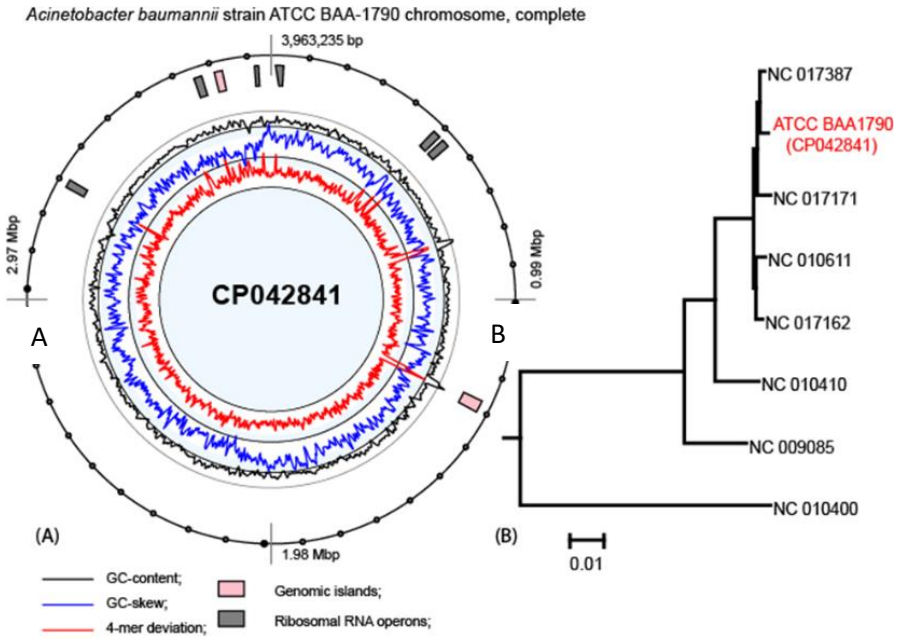


Figure 2.11: (A) Circular map representing the obtained chromosome for *A. baumannii* with predicted genomic islands, rRNA clusters and histograms of the GC-content, GC-skew and tetranucleotide pattern deviations calculated in an 8 kbp sliding window stepping 2 kbp. (B)

Neighbour-Joining phylogenetic tree shows relations of the sequenced strain BAA-1790 with reference genomes of *A. baumannii*.

BUSCO

The Benchmarking Universal Single-Copy Orthologs (BUSCO v3.0.2) (Simao et al., 2015) software was used with default parameters to assess the completeness of the final assemblies. *Table 2.2* shows the BUSCO scores obtained for each of the assembled genomes. Genomes were then annotated using the Rapid Annotations using Subsystems Technology (RAST) server (Aziz et al., 2008), followed by manual editing of the annotations.

Table 2.2.: Table summary of bacterial genome assemblies.

	S. aureus	*E. coli*	*A. baumannii*
Length NC	2,791,218	4,682,561	3,963,235
Length FS	2,792,888	4,682,572	-
BUSCO score NC	89.9	85,5	79.0
BUSCO score FS	86.5	86,7	-
GC content	32.9%.	50.5%	40%
NCBI Accession NC	CP033505	CP042865	CP042841
NCBI Accession FS	CP033506	CP042867	-
BioProject	PRJNA493710	PRJNA557356	PRJNA557366
BioSample NC	SAMN09635552	SAMN12395313	SAMN12399660
BioSample FS	SAMN09635551	SAMN12395320	-

Discussion

To better understand the molecular mechanisms of microorganisms, including the development of drug resistance, it is important to accurately determine their genomic sequences. Although recent advances in sequencing technologies have greatly assisted to make obtaining sequence information rapid and cost effective, accurate reference strains are still needed to determine significant differences between isolates. With the use of long read sequencing technologies complex bacterial genomes can now be resolved more easily.

PacBio sequencing generates reads which do not have a set read length, but rather generates a distribution of read lengths which are dependent on how long each individual polymerase is active. *Table 2.1* and *Figure 2.5* showed that with PacBio sequencing, a large number of reads, with varying read lengths could be obtained for NC and FS genomes for *S. aureus*. Since the bacterial genomes used in our study had relatively small genome sizes, ranging between 2,7 – 4.6 Mbp, the polymerases used during sequencing were able to sequences the genomes various times, allowing many reads to be generated, and ultimately high coverage of the genomes. This was especially required for the *de novo* assemblies, since we used a non-hybrid approach.

S. aureus ATCC BAA-39 genome assembly analysis

In 2010 the MRSA strain, *S. aureus* ATCC BAA-39, was isolated from a nasal clinical sample (PRJNA50533), sequenced with Roche 454 technology and assembled into 83 contigs. The assembly was published in the NCBI under accession number GCA_000146385.1. In our study, two genome variants of the *S. aureus* ATCC BAA-39 strain, NC and FS, were obtained as single contig sequences without gaps and ambiguities. The BUSCO output for the genomes revealed that 90% of the expected conserved bacterial genes were identified as complete without fragmentation. The completeness of our assemblies is an improvement compared to the previous assembly, GCA_000146385.1. Lengths of the NC and FS genomes were 2 798

153 bp and 2 798 095 bp, respectively. The obtained genomes were shorter compared to the previous assembly, which had 2 865 318 bp excluding gaps. The difference in lengths could have resulted from either a redundancy of the previous assembly or due to selection of bacterial variants with faster growth rates and shorter genomes during the daily transference of the cultures to fresh medium under experimental conditions. The staphylococcal methicillin-resistance chromosome cassette (SCCmec), containing the *mecA* gene, which encodes for the PBP2a-protein [3, 5], was found in both genomes, 50 000 bp downstream from the replication origin. The genomes contained three other large prophages, of which most genes were transcriptionally silent. No plasmids were identified in the *S. aureus* genomes.

Investigation of phylogenetic trees generated from *S. aureus* sequences available from the NCBI reveal our strain to be in the same clade as two other strains, NCTC4163 and NCTC4137. These strains are also methicillin resistant strains.

Analyses of the genome annotations revealed frameshift mutations, where neighbouring genes were annotated as part of the same longer protein. Variant calling allowed us to identify polymorphic sites in the genomes. More frameshift mutations were observed in the FS-1 treated cultures as well as different distribution of polymorphic sites compared to the NC cultures. These frameshift mutations, which caused truncated and possibly less functional proteins, could be responsible for the observed increased sensitivity to antibiotics when bacteria are treated with FS-1.

Assembly and annotation of complete genomes of *E.coli* ATCC BAA-196

The complete genome sequences of both the experimental and negative control strains were assembled and annotated for the multidrug resistant strain, *E. coli* ATCC BAA-196. The complete genome assembly of this *E. coli* strain was not previously published. For this study, the size of the obtained chromosomes of the NC and FS genomes were 4 682 572 bp and 4 682

561 bp, respectively. Large plasmids of 266 396 bp (NC) and 279 992 bp (FS) were discovered and showed sequence similarity of between 90 and 99% with a plasmid from the bacteria *Klebsiella pneumonia*, pKP64477b. The FS variant had a larger plasmid due to an insertion of a prophage flanked with two copies of *insH* transposases, which inserts the sequence element *IS5*, responsible for enhancing gene transcription when placed on either side of the promoter for a target gene . Furthermore, the FS variant also contained two smaller plasmids of 44 240 bp and 11 153 bp which comprised transposable elements and genes mobilized from the large plasmid implying a plasmid destabilizing effect of the treatment with FS-1. The sequences of small plasmids were aligned against the large plasmid, as seen in *Figure 2.12* below. Small plasmids contain *tra* operons ensuring their mobility by conjugation.

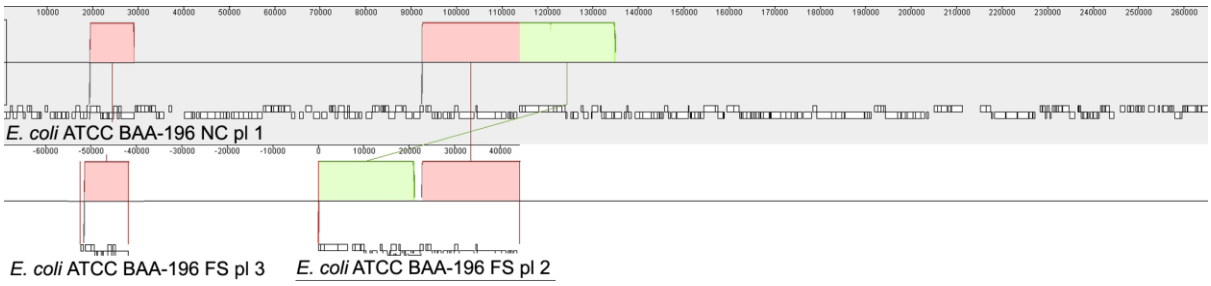


Figure 2.12: Representation of alignment of smaller plasmids to larger plasmid.

The identified genomic islands and plasmids comprised genetic determinants associated with antibiotic resistance, including β-lactamases of classes A and D, which utilizes serine for the hydrolysis β-lactams (Richmond and Sykes, 1973), tellurium resistance operon *terABCDW*, arsenic resistance gene *arsR*, which encodes a regulatory protein to control the expression of the *ars* operon (Branco et al., 2008), chloramphenicol and aminoglycoside acetyltransferases, drug resistance regulators, and efflux proteins of MdtEF-TolC, EmrAB-OMF, EmrE-QacE and EmrKY-TolC families. The mobile genetic elements, including fimbrial adhesins *ecpD* and *fimHBGFE* operon, provided the strain with multiple adhesion virulent factors (Low et al., 2006).

Assembly and annotation of complete genome of *A. baumannii* ATCC BAA-1790

The genome of *A. baumannii* ATCC BAA-1790, a multidrug resistant strain isolated in 2008 from sputum, has not been previously published. For this study, only the negative control (NC) sample was used, since the experiment sample (FS) was contaminated. For *A. baumannii*, 326 117 long PacBio reads were generated, with an average length of 10 000 bp. An ungapped chromosome of 3 963 235 bp and a contig N50 of 3 963 222 were obtained, as well as a 67 023 bp plasmid which had an average GC-content of 39.06%. The identified genomic islands contained the carbapenem-hydrolyzing class D β-lactamase OXA-23, responsible for hydrolyzing and conferring resistance to oxacillin and penicillin (Richmond and Sykes, 1973). Sul1 sulfonamide-resistance protein was also identified, which contributed to the antibiotic resistance of the strain. An integrase insertion, followed by aadA1 aminoglycoside 3'-adenyltransferase and aadC1 aminoglycoside N(3')-acetyltransferase I, was also observer. This renders the strain with resistance to aminoglycosides (Karah et al., 2016). Multiple genes for drug efflux pumps and β-lactamases of OXA-23, OXA-82, and ADC-25 families rendering the resistance to cephalosporins and penams (subclass of the broader β-lactam family of antibiotics) were identified by the Resistance Gene Identifier (RGI) server (Jia et al., 2017). Examination of 2 588 orthologous proteins revealed that the *BAA-1790* strain is phylogenetically closest to the *A. baumannii TCDC-AB0715* strain (*Figure 4 B*) (Chen et al., 2011). The identified plasmid is common in *A. baumannii* isolates and shows more than 99% DNA sequence identity with the plasmids pMAL-2 (KX230794.1) and FDAARGOS_493 (CP033857.1).

General assessment of results of SMRT-link assembly of PacBio reads generated from bacterial genomic DNA

Non-hybrid *de novo* assembly approaches, such as HGAP (used in this study) and the PacBio Corrected Reads (PBcR) pipeline , are “self-correction” approaches and have successfully been

used to complete complex microbial genomes (Koren et al., 2013, Chin et al., 2013). These hybrid approaches do not require short highly accurate reads to perform error corrections, however, they do require sequences coverage of 80x-100x (Utturkar et al., 2014). The key part of HGAP is to exploit the advantages of PacBio reads by pre-assembling accurate long overlapping sequences to create a consensus and then correcting errors on the longest reads by using reads from the same library that are shorter (Chin et al., 2013).

Systematic errors affect the accurateness of consensus sequences. Bases that are systematically read incorrectly, will incorrectly be called in the consensus sequence and cannot be corrected for by adding more coverage. SMRT sequencing achieves a >99,999% consensus accuracy due to the single-pass errors that are randomly distributed and can rapidly be 'washed' out when building a consensus. Multiple publications have verified this theoretically and experimentally (Carneiro et al., 2012, Koren et al., 2012).

Schmid et al. (Schmid et al., 2018) recently studied long repeat regions in prokaryotic genomes and illustrated how the use of very long reads together with assembly algorithms, are able to resolve long, nearly identical repeats. In this study we demonstrated de novo assemblies of bacterial genomes with the use of PacBio reads which had sufficient coverages. All three bacteria had genome coverages of over 600x for NC and FS genomes. With these high coverages, our PacBio reads could be "self-corrected" and short reads were not required. All the assembled genomes were complete (no gaps) and can be used to study important molecular mechanisms in these bacteria to improve our understanding, especially concerning the development of resistance. There are various other projects that have recently revealed how complete circular bacterial genomes can easily be resolved with the use of long read sequencing technologies (Liao et al., 2019, Wick et al., 2017).

De novo genome assemblies enables a more comprehensive view of not only prokaryotic, but also eukaryotic genomes, and provides important genetic information which cannot be easily

revealed with other approaches, such as resequencing. With the fast advances in sequencing technologies, generating reference grade genomes can soon become routine procedures.

References

Complete microbial genomes with ease and confidence [Online]. PacBio. Available: https://www.pacb.com/applications/whole-genome-sequencing/microbial/ [Accessed].

EcoCyc [Online]. Available: https://biocyc.org/gene?orgid=ECOLI&id=G6134 [Accessed 2August2019].

Pacific Biosciences Develops Transformative DNA Sequencing Technology *PacBio*.

PBcR Assembler [Online]. Available: http://wgs-assembler.sourceforge.net/wiki/index.php?title=PBcR [Accessed].

Post Run QC Analysis [Online]. PacBio. Available: https://www.pacb.com/training/PostRunQCAnalysis/story_content/external_files/Post%20Run%20QC%20Analysis.pdf [Accessed].

2018a. *BLASR: The PacBio long read aligner* [Online]. GitHub: Pasific Biosciences. Available: https://github.com/PacificBiosciences/blasr [Accessed 24 Seeptember 2018].

2018b. *PacBio variant and consensus caller* [Online]. GitHub: Pacific Biosciences. Available: https://github.com/PacificBiosciences/GenomicConsensus [Accessed 24 Spetember 2018].

ALKAN, C., SAJJADIAN, S. & EICHLER, E. E. 2011. Limitations of next-generation genome sequence assembly. *Nat Methods,* 8**,** 61-5.

AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics,* 9**,** 75.

BEZUIDT, O., LIMA-MENDEZ, G. & REVA, O. 2009. SeqWord Gene Island Sniffer: A program to study the lateral genetic exchange among bacteria. *World Academy of Science, Engineering and Technology,* 58**,** 1169-1174.

BRANCO, R., CHUNG, A.-P. & MORAIS, P. V. 2008. Sequencing and expression of two arsenic resistance operons with different functions in the highly arsenic-resistant strain Ochrobactrum tritici SCII24T. *BMC Microbiology,* 8**,** 95.

BROWN, S. D., NAGARAJU, S., UTTURKAR, S., DE TISSERA, S., SEGOVIA, S., MITCHELL, W., LAND, M. L., DASSANAYAKE, A. & KOPKE, M. 2014. Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of Clostridium autoethanogenum and analysis of CRISPR systems in industrial relevant Clostridia. *Biotechnol Biofuels,* 7**,** 40.

CARNEIRO, M. O., RUSS, C., ROSS, M. G., GABRIEL, S. B., NUSBAUM, C. & DEPRISTO, M. A. 2012. Pacific biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics,* 13**,** 375-375.

CHEN, C. C., LIN, Y. C., SHENG, W. H., CHEN, Y. C., CHANG, S. C., HSIA, K. C., LIAO, M. H. & LI, S. Y. 2011. Genome sequence of a dominant, multidrug-resistant Acinetobacter baumannii strain, TCDC-AB0715. *J Bacteriol,* 193**,** 2361-2.

CHIN, C. S., ALEXANDER, D. H., MARKS, P., KLAMMER, A. A., DRAKE, J., HEINER, C., CLUM, A., COPELAND, A., HUDDLESTON, J., EICHLER, E. E., TURNER, S. W. & KORLACH, J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods,* 10**,** 563-9.

EID, J., FEHR, A., GRAY, J., LUONG, K., LYLE, J., OTTO, G., PELUSO, P., RANK, D., BAYBAYAN, P., BETTMAN, B., BIBILLO, A., BJORNSON, K., CHAUDHURI, B., CHRISTIANS, F., CICERO, R., CLARK, S., DALAL, R., DEWINTER, A., DIXON, J., FOQUET, M., GAERTNER, A., HARDENBOL, P., HEINER, C., HESTER, K., HOLDEN, D., KEARNS, G., KONG, X., KUSE, R., LACROIX, Y., LIN, S., LUNDQUIST, P., MA, C., MARKS, P., MAXHAM, M., MURPHY, D., PARK, I., PHAM, T., PHILLIPS, M., ROY, J., SEBRA, R., SHEN, G., SORENSON, J., TOMANEY, A., TRAVERS, K., TRULSON, M., VIECELI, J., WEGENER, J., WU, D., YANG, A., ZACCARIN, D., ZHAO, P., ZHONG, F., KORLACH, J.

& TURNER, S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science,* 323**,** 133-8.

FOQUET, M., SAMIEE, K. T., KONG, X., CHAUDURI, B. P., LUNDQUIST, P. M., TURNER, S. W., FREUDENTHAL, J. & ROITMAN, D. B. 2008. Improved fabrication of zero-mode waveguides for single-molecule detection. *Journal of Applied Physics,* 103**,** 034301.

GREG CONCEPCION, S. K., CHRIS DUNN, JASON CHIN. 2016. *FALCON* [Online]. Available: https://pb-falcon.readthedocs.io/en/latest/about.html [Accessed 19 September 2019].

GUAN, P. & SUNG, W.-K. 2016. Structural variation detection using next-generation sequencing data: A comparative technical review. *Methods,* 102**,** 36-49.

JIA, B., RAPHENYA, A. R., ALCOCK, B., WAGLECHNER, N., GUO, P., TSANG, K. K., LAGO, B. A., DAVE, B. M., PEREIRA, S., SHARMA, A. N., DOSHI, S., COURTOT, M., LO, R., WILLIAMS, L. E., FRYE, J. G., ELSAYEGH, T., SARDAR, D., WESTMAN, E. L., PAWLOWSKI, A. C., JOHNSON, T. A., BRINKMAN, F. S., WRIGHT, G. D. & MCARTHUR, A. G. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res,* 45**,** D566-d573.

KARAH, N., DWIBEDI, C. K., SJOSTROM, K., EDQUIST, P., JOHANSSON, A., WAI, S. N. & UHLIN, B. E. 2016. Novel Aminoglycoside Resistance Transposons and Transposon-Derived Circular Forms Detected in Carbapenem-Resistant Acinetobacter baumannii Clinical Isolates. *Antimicrob Agents Chemother,* 60**,** 1801-18.

KOREN, S., HARHAY, G. P., SMITH, T. P., BONO, J. L., HARHAY, D. M., MCVEY, S. D., RADUNE, D., BERGMAN, N. H. & PHILLIPPY, A. M. 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol,* 14**,** R101.

KOREN, S., SCHATZ, M. C., WALENZ, B. P., MARTIN, J., HOWARD, J., GANAPATHY, G., WANG, Z., RASKO, D. A., MCCOMBIE, W. R., JARVIS, E. D. & PHILLIPPY, A. M. 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nature biotechnology,* 30**,** 693-700.

LIAO, Y.-C., CHENG, H.-W., WU, H.-C., KUO, S.-C., LAUDERDALE, T.-L. Y. & CHEN, F.-J. 2019. Completing Circular Bacterial Genomes With Assembly Complexity by Using a Sampling Strategy From a Single MinION Run With Barcoding. *Frontiers in Microbiology,* 10.

LOW, A. S., DZIVA, F., TORRES, A. G., MARTINEZ, J. L., ROSSER, T., NAYLOR, S., SPEARS, K., HOLDEN, N., MAHAJAN, A., FINDLAY, J., SALES, J., SMITH, D. G. E., LOW, J. C., STEVENS, M. P. & GALLY, D. L. 2006. Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic Escherichia coli O157:H7. *Infection and immunity,* 74**,** 2233-2244.

PHILLIPPY, A. M., SCHATZ, M. C. & POP, M. 2008. Genome assembly forensics: finding the elusive mis-assembly. *Genome Biology,* 9**,** R55.

PUGH, T. J., WEERARATNE, S. D., ARCHER, T. C., POMERANZ KRUMMEL, D. A., AUCLAIR, D., BOCHICCHIO, J., CARNEIRO, M. O., CARTER, S. L., CIBULSKIS, K., ERLICH, R. L., GREULICH, H., LAWRENCE, M. S., LENNON, N. J., MCKENNA, A., MELDRIM, J., RAMOS, A. H., ROSS, M. G., RUSS, C., SHEFLER, E., SIVACHENKO, A., SOGOLOFF, B., STOJANOV, P., TAMAYO, P., MESIROV, J. P., AMANI, V., TEIDER, N., SENGUPTA, S., FRANCOIS, J. P., NORTHCOTT, P. A., TAYLOR, M. D., YU, F., CRABTREE, G. R., KAUTZMAN, A. G., GABRIEL, S. B., GETZ, G., JÄGER, N., JONES, D. T. W., LICHTER, P., PFISTER, S. M., ROBERTS, T. M., MEYERSON, M., POMEROY, S. L. & CHO, Y.-J. 2012. MEDULLOBLASTOMA EXOME SEQUENCING UNCOVERS SUBTYPE-SPECIFIC SOMATIC MUTATIONS. *Nature,* 488**,** 106-110.

RHOADS, A. & AU, K. F. 2015. PacBio Sequencing and Its Applications. *Genomics, Proteomics & Bioinformatics,* 13**,** 278-289.

RICE, L. B., MARSHALL, S. H., CARIAS, L. L., SUTTON, L. & JACOBY, G. A. 1993. Sequences of MGH-1, YOU-1, and YOU-2 extended-spectrum beta-lactamase genes. *Antimicrobial Agents and Chemotherapy,* 37**,** 2760-2761.

RICHMOND, M. H. & SYKES, R. B. 1973. The beta-lactamases of gram-negative bacteria and their possible physiological role. *Adv Microb Physiol,* 9**,** 31-88.

SCHADT, E. E., TURNER, S. & KASARSKIS, A. 2010. A window into third-generation sequencing. *Human Molecular Genetics,* 19**,** R227-R240.

SCHATZ, M. C., DELCHER, A. L. & SALZBERG, S. L. 2010. Assembly of large genomes using second-generation sequencing. *Genome Research,* 20**,** 1165-1173.

SCHMID, M., FREI, D., PATRIGNANI, A., SCHLAPBACH, R., FREY, J. E., REMUS-EMSERMANN, M. N. P. & AHRENS, C. H. 2018. Pushing the limits of de novo genome assembly for complex prokaryotic genomes harboring very long, near identical repeats. *Nucleic Acids Res,* 46**,** 8953-8965.

SIMAO, F. A., WATERHOUSE, R. M., IOANNIDIS, P., KRIVENTSEVA, E. V. & ZDOBNOV, E. M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics,* 31**,** 3210-2.

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

UTTURKAR, S. M., KLINGEMAN, D. M., LAND, M. L., SCHADT, C. W., DOKTYCZ, M. J., PELLETIER, D. A. & BROWN, S. D. 2014. Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences. *Bioinformatics,* 30**,** 2709-16.

WICK, R. R., JUDD, L. M., GORRIE, C. L. & HOLT, K. E. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology,* 13**,** e1005595.

CHAPTER 3

Investigating the effect of FS-1 on gene regulation patters to investigate the reversion of antibiotic resistance

Abstract

Currently, the development and spread of multidrug resistance is becoming a major concern to the public health. Reversion of antibiotic resistant bacteria into antibiotic sensitive phenotypes is becoming a promising approach to address this problem. It is therefore important to better understand how bacteria regulate genes and pathways to develop resistance. This study set out to investigate gene regulation patterns in the multiple drug resistant strains *Staphylococcus aureus*, *Escherichia coli*, and *Acinetobacter baumannii*, cultivated in a sub-lethal dose of FS-1. It was hypothesized that gene regulation could be affected by both the specific activity of the iodine-containing drug, and/or by the selection of genetic variants better adapted to the presence of the drug. Bacteria were cultivated in drug-free and drug-containing medium to study the effect of FS-1 on bacterial metabolism, as well as the direct effect of the drug. Bacteria were also cultivated over 10 passages to determine the effect of the drug on bacterial populations. For the immediate effect, treatment with FS-1 caused down-regulation of various important pathway which consume the co-enzymes NADH and NADPH, while the metabolic processes associated with the production of the reduced species of these co-enzymes generally were up-regulated. It may be assumed that the pathways helping bacteria to withstand oxidative stress were up-regulated. After cultivating bacteria over 10 passages in the medium with FS-1, they regained the initial growth rate by adapting to the presence of FS-1, which required an alternative gene transcription regulation controlled either by accumulation of specific mutations in bacterial populations, or due to epigenetic phase variations.

Introduction

The central dogma of molecular biology is an important concept to better understand how the flow of genetic information, stored within genes, are transcribed into RNA, and ultimately translated into proteins (Crick, 1970). An organism's phenotype is characterized by the expression of this genetic information, since the transcription of genes into RNA molecules determine the cell's biological activities and regulation and it can be influenced by environmental factors. It is therefore important to better understand an organism's transcriptome to help us to interpret and understand how the functional elements within the genome are involved with disease development and treatment (Kukurba and Montgomery, 2015).

The development of multidrug resistant bacteria is currently a great concern (Levy and Marshall, 2004), since the misuse of antibiotics have caused a strong selective pressure for these resistant bacteria and various treatment options are becoming ineffective (Wang and Lipsitch, 2006). There are various factors involved in the acquisition of multidrug resistance in bacteria. It is generally recognized that bacteria may gain resistance to antibiotics either due to mutations, which modify the target proteins and expand the spectrum of resistance activity, or by acquisition of antibiotic resistance genes (Tenover and Hughes, 1996).

The short generation times of bacteria enable them to rapidly adapt to changing environments, as is currently being observed by the development of antibiotic resistance (Woodford and Ellington, 2007). Although bacteria can regulate gene expression levels to survive antibiotics and other stresses, regulation is not always perfect, and the regulation of certain important genes may not be influenced by a given stress, ultimately disrupting proper growth and survival of the bacteria under these stresses (Gottesman, 2017).

Common mutations have been found in clinical isolates and in evolutionary experiments, which confer resistance to antibiotics through altered gene expression of resistance genes (Bergstrom and Normark, 1979, Koutsolioutsou et al., 2005). Resistance genes contribute to specific antibiotic resistance mechanisms, which include drug degrading enzymes, such as β-lactamases, and transcriptional regulators of stress response. Other genes can also contribute to the intrinsic function of a drug's mechanism of action, such as drug targets themselves, which include *folA* found in *Escherichia coli*, enzymes which can activate a prodrug, and porins that mediate cellular entry of certain antibiotic molecules (Tamae et al., 2008, Girgis et al., 2009). Although the development of resistance due to mutations that change gene expression levels have previously been studied, the extent of this phenomenon has not been systematically characterized. Understanding how bacteria regulate gene expression during treatment, could open new insights into the development of resistance and how to reverse it.

The first aim of the research covered in this chapter was to investigate the immediate effect of FS-1 on the gene regulation patterns in the multidrug resistant strains *Staphylococcus aureus* (*S. aureu)*, *Escherichia coli* (*E. coli*), and *Acinetobacter baumannii* (*A. baumannii*) cultivated in a sub-lethal dose of FS-1 (FS variant) in comparison to the negative control strain (NC) cultivated on Mueller-Hinton (MH) liquid medium without the drug. Thereafter, the metabolic effect of the drug on the bacteria was studied. Secondly, the effect of the drug was investigated on *S. aureus* and *E. coli* cultures after cultivation over 10 passages, to determine the effect on bacterial populations. These findings will improve our understanding of the observed reversion of resistant bacteria to drug sensitive phenotype due to alternative gene regulation patterns.

Materials and methods

It should be noted that all experimental procedures with bacterial cultures were performed by our collaborators in this project from the Scientific Center for Anti-Infectious Drugs (SCAID), Almaty, Kazakhstan. These methods are mentioned here to better represent the overall experimental design of the project. All bioinformatics procedures were performed in the Centre for Bioinformatics and Computational Biology (CBCB) at the University of Pretoria.

Multidrug resistant strains of *Staphylococcus aureus* (S. aureus) ATCC BAA-39, *Escherichia coli* (E. coli) ATCC BAA-196, and *Acinetobacter baumannii (A. baumannii)* ATCC BAA-1790 were obtained from the American Type Culture Collection (ATCC) and used as model organisms for this study.

Immediate Effect

Culture Cultivation With The FS-1 Drug

To determine the immediate effect of the drug, bacterial inoculants were incubated for 2.5 hours (end of lag phase) and 9 hours (mid of the exponential growth phase) at 37°C, see *Figure 3.1*. Thereafter, the experimental cultures were suplemented with FS-1 for 5 min cultivation, whereas for the negative control samples, the same volume of physiological saline was applied to the bacterial cultures. All experiments were performed in three replicates.

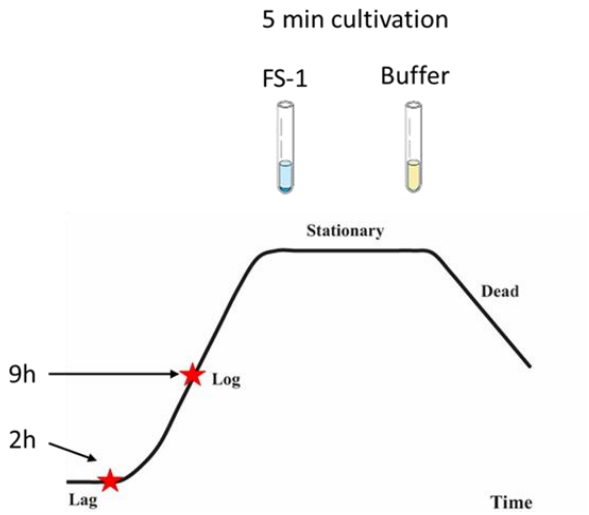


Figure 3.1: Bacterial growth curve indicating the lag phase and log phase.

Long Cultivation

Sample Collection and Bacterial Cultivation

All bacteria were inoculated into test-tubes that contained 10 ml of the Mueller-Hinton Broth medium, for experimental conditions (denoted as FS) the medium was supplemented with the FS-1 drug, and for control conditions (denoted as NC) the same medium was used without the drug. The test-tubes were incubated at 37°C for 24 h, thereafter 0.1 ml aliquots of the cultures were transferred to fresh tubes with the corresponding media. After 10 passages, experimental and control samples were collected for *E. coli*, whereas, for *S. aureus* the experimental and control bacteria were cross-inoculated into tubes with drug-containing and drug-free media for further overnight incubation in three repeats. Thereafter, DNA extraction followed, see experimental scheme in *Figure 3.2*.

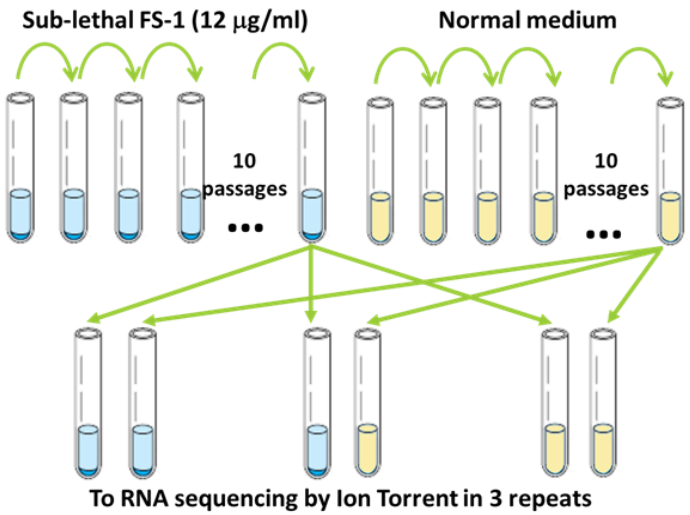


Figure 3.2: Experimental scheme for cultivation of bacteria in the medium containing FS-1 (experimental samples) and in regular medium without the drug (control samples).

RNA Library Preparation and Sequencing

Total RNA was isolated from the cultures with the use of the RiboPure Bacteria Kit (Ambion, Lithuania) as instructed in the developer's guidelines. Afterwards, the quantity and quality of the isolated RNA was verified with use of the NanoDrop 2000c spectrophotometer (Thermo Scientific, USA) at optical wavelengths of 260 and 280 nm. The MICROBExpress Bacterial mRNA Purification Kit (Ambion, Lithuania) was used according to the developer's guidelines to do purification of total RNA, thereafter the effectiveness of the purification was determined on the Bioanalyzer 2100 (Agilent, Germany) with the RNA 6000 Nano LabChip Kit (Agilent Technologies, Lithuania).

The library preparation for the extracted RNA samples included enzymatic fragmentation with the use of the Ion Total RNA Seq kit V2. Thereafter, the Ion Xpress RNA-Seq Barcode 01-16 Kit was used to do barcoding. RNA sequencing was then done using the Ion 318 Chip Kit V2 on the Ion Torrent PGM sequencer (Life Technologies, USA).

Differential Expression Analyses

The differential expression was done using the R-3.4.4 software. Firstly, a reference index was built for each reference genome using the "buildindex" function available in the *Rsubreads* package (Bioconducter). For each bacterium, the obtained RNA fragments were aligned to the relevant reference genome with the use of the "*align*" function. The aligned *BAM* files and relevant *GFF* annotation files were then used as input for the *featureCounts* function to obtain gene counts. The R packages *DESeq2* (Bioconducter) and *GenomicFeatures* was then used in R studio for the differential expression analyses. The full commands used are available in *Supplementary File 3* and *Supplementary File 4*.

Thereafter, an inhouse Python script was used to generate expression plots. This script required as input a *GFF* file and the output generated by *DESeq2* which contains the gene counts.

Metabolic Pathway Analyses

With the use of the significantly identified differentially expressed genes, pathways and reactions influenced by those gene can be investigated. To evaluate the overall metabolic pathways affected by FS-1, the Pathway Tools software (Karp et al., 2002) was used. Firstly, the "*pathologic*" tool was used to generate a new Pathway/Genome Database (PGDB) which contains the predicted metabolic pathways and operons of the bacteria, given a Genbank or *GFF* file as input. Thereafter, a "smart table" was created containing all genes. The "smart tables" were then transformed by "reactions of genes". The "smart tables" were then exported for more in depth analysis.

Results

Immediate effect

To investigate the immediate effect of FS-1 on bacterial metabolism, bacteria were exposed to the FS-1 drug for 5 minutes. After 5 min incubation of the bacterial strains with FS-1 during different growth phases, a killing buffer was used to stop all metabolic processes in bacterial cells, followed by total RNA extraction and purification.

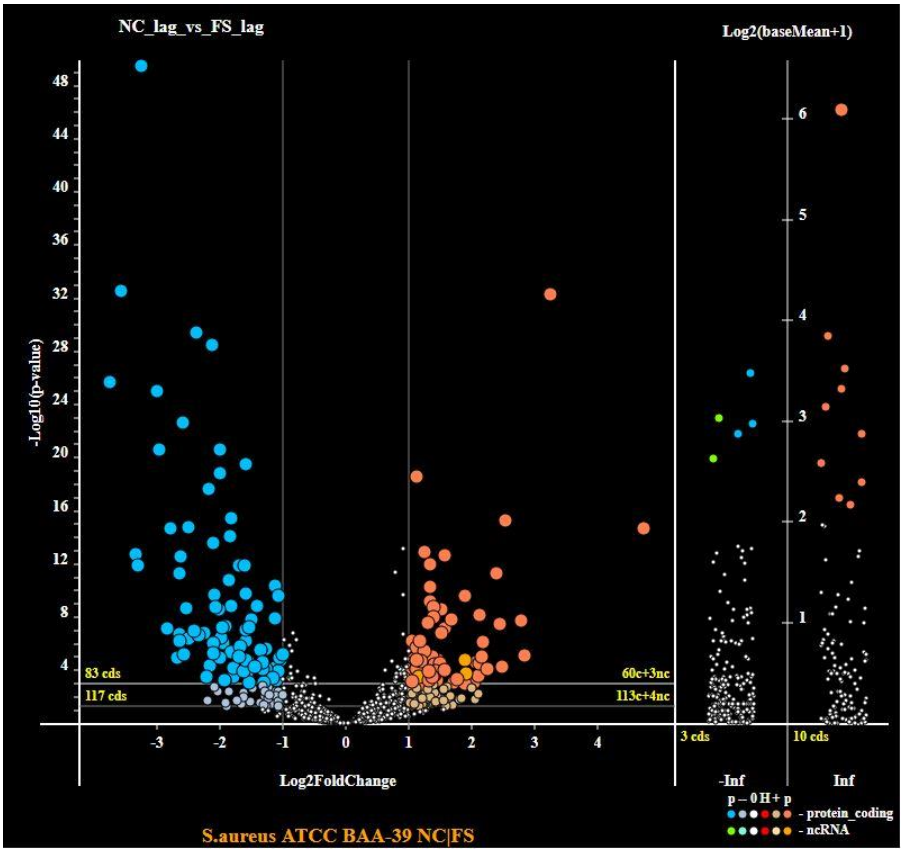
Investigating gene regulation of bacteria exposed to FS-1 at different growth phases

Regulation of gene expression due to treatment with FS-1 was studied at the end of the lag phase and in the mid of the exponential growth (log) phase. During the log phase, the regulation of gene expression was less affected compared to during the lag phase. This could be due to low level of gene expression at the beginning of the culture growth, and the 5 min exposure of the culture to FS-1 was able to impact the gene expression stronger than during the log phase where many genes were over-expressed. However, the overall trends of the gene expression regulation were similar between the log and lag phase. Only results obtained during the lag growth phase will be discussed below.

The effect of treatment with FS-1 on gene expression regulation was studied and is shown in *Figure 3.3*. A common trend can be seen in all three bacteria (*Figure 3.3A-C*), which is that FS-1 caused suppression of gene expression in bacteria treated with the drug compared to the control cultures. The *-Inf* and *Inf* plots, next to the log fold change in *Figure 3.3*, represents gene only expressed in the NC culture (*-Inf*) and gene only expressed in the experimental cultures (*Inf*).

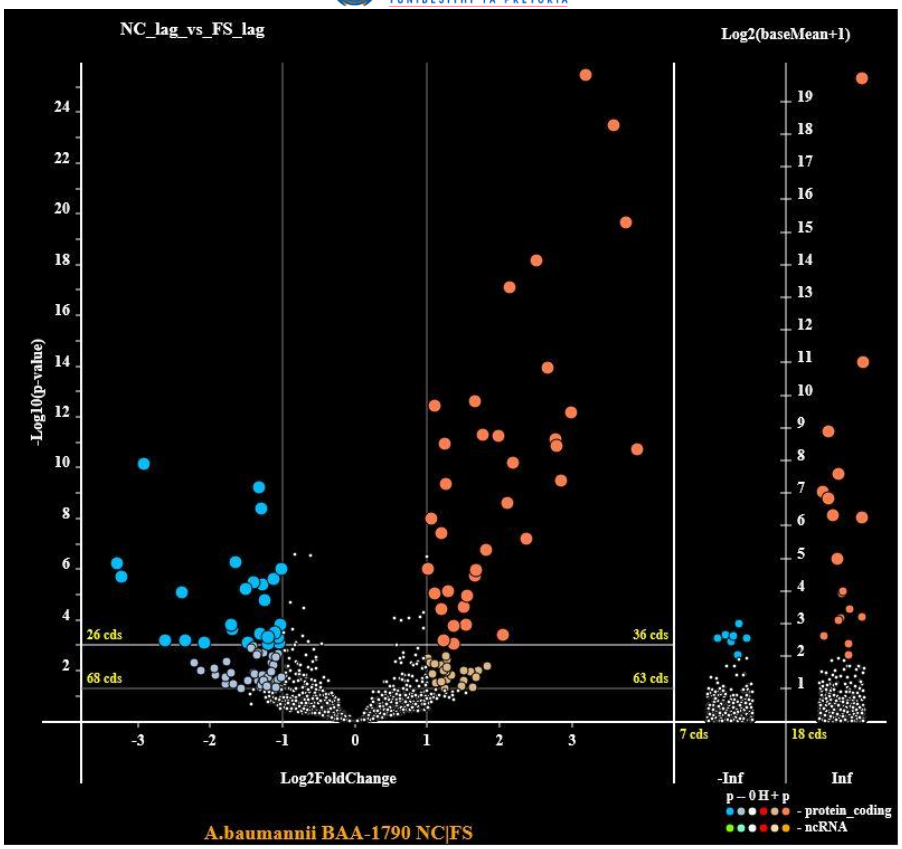
As seen in *Figure 3.3-A*, 117 protein coding genes were suppressed whereas 113 protein coding and 4 non-coding genes were induced in *S. aureus*. The figure also shows that 3 protein coding and 2 non-coding genes were expressed only in the control culture, compared to 10 protein

coding genes only expressed in the FS culture. In *Figure 3.3-B*, 68 protein coding genes were suppressed, and 63 protein coding genes were induced in *A. baumannii*, and 7 protein coding genes were expressed only in the control culture, compared to 18 protein coding genes only expressed in the FS culture. In *Figure 3.3-C*, 35 protein coding genes were suppressed, and 31 protein coding genes were induced in *E. coli*. The figure also shows that 35 protein coding genes were expressed only in the control culture, compared to 2 protein coding genes only expressed in the FS culture.



A

B



C

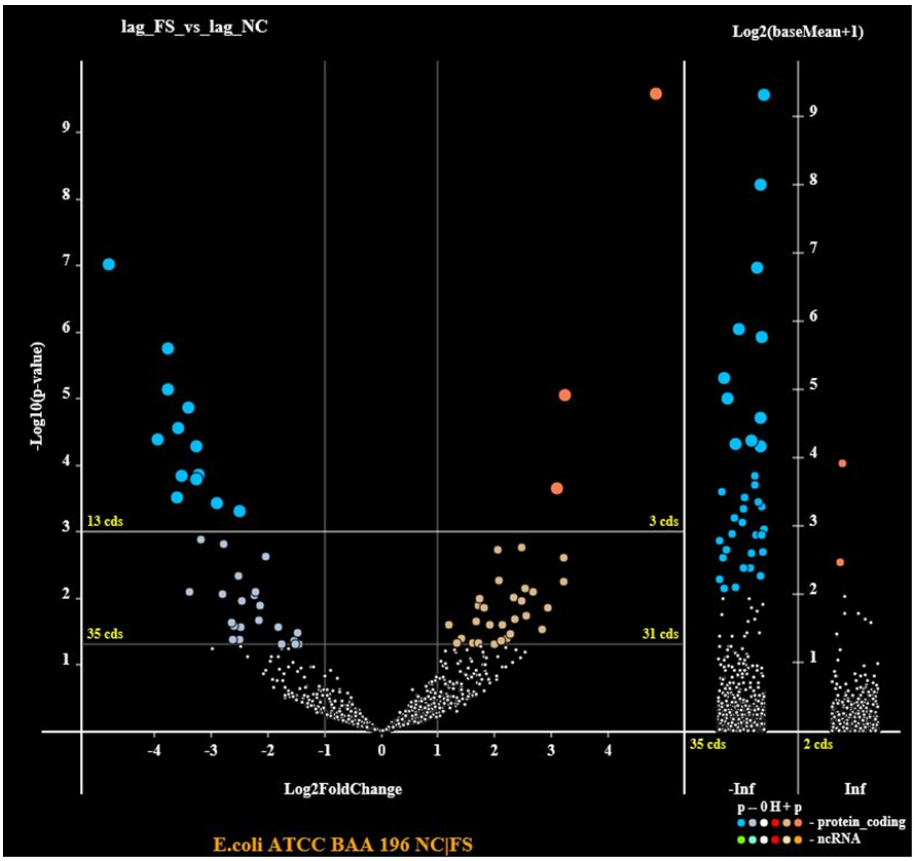


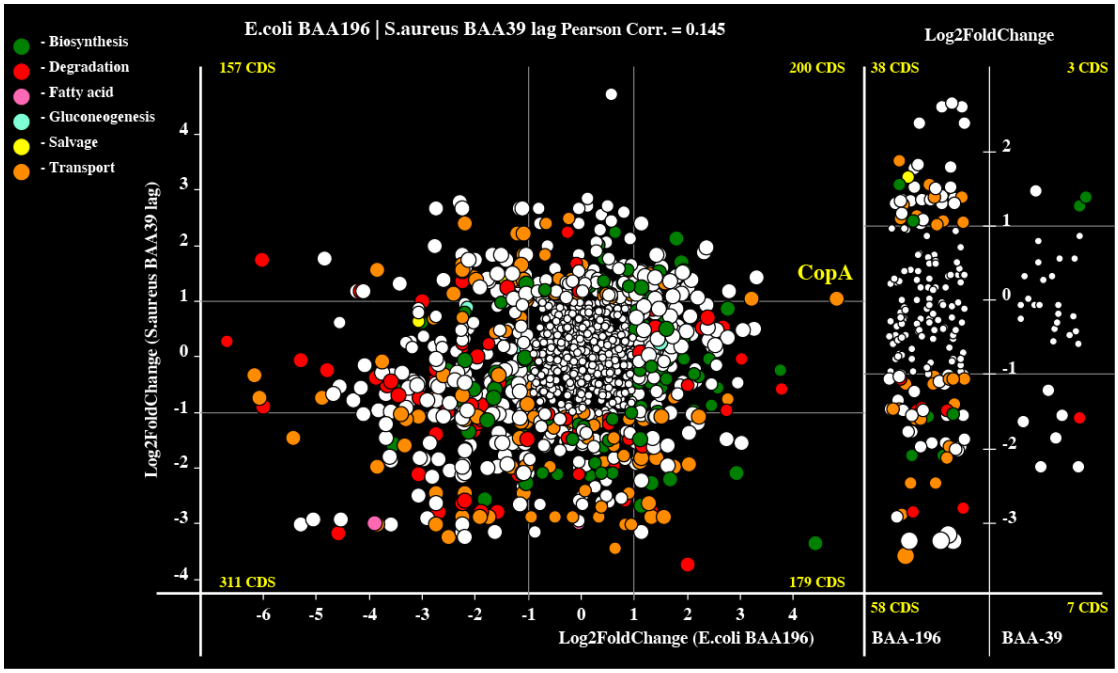
Figure 3.3: Volcano plots generated for the three bacteria during the lag growth phase for NC and FS bacterial cultures. The horizontal lines on the plot represent p-values of 0.05 and 0.01.

A*) S. aureus* cultivated on normal medium vs medium with FS-1; B) *A. baumannii* cultivated on normal medium vs medium with FS-1; C) *E. coli* cultivated on normal medium vs medium with FS-1.

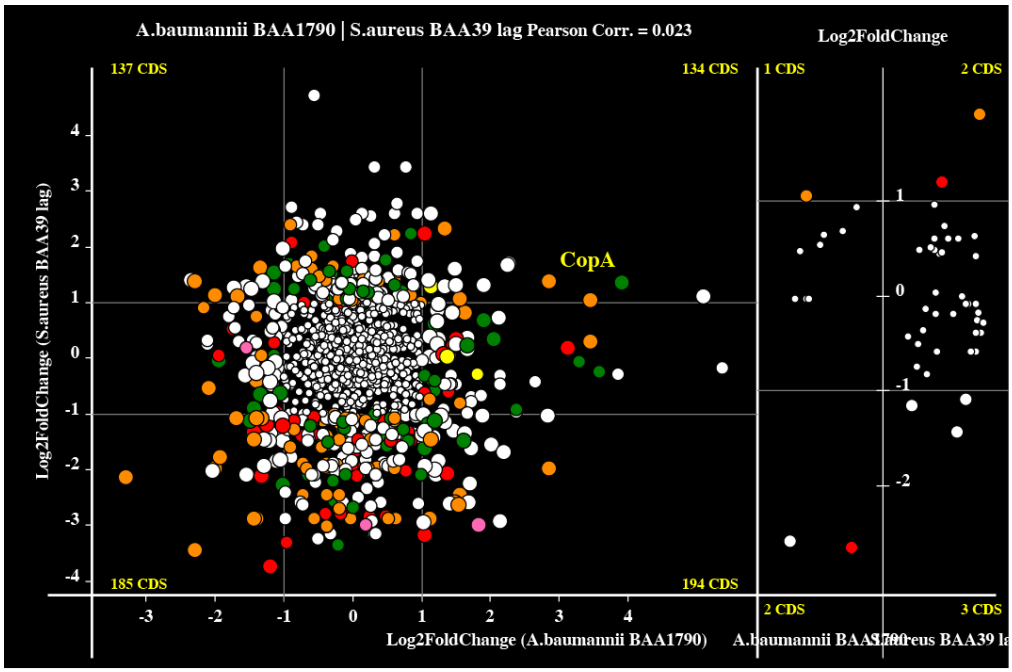
Does addition of FS-1 to the medium cause similar gene regulation response in all bacteria?

To address this research question, only pathways that were affected in at least two of the three bacteria were investigated. With the use of an inhouse Python script, plots in *Figure 3.4* were generated. These plots represent genes involved in similar pathways regulated in the bacteria. Each plot represents gene regulation between two bacteria represented by log2 (fold_change) expression values in experimental conditions (treatment with FS-1) compared to the negative control condition. Different plots represent different combinations of the bacteria. As shown in the plots, the Pearson correlation is low since the bacteria are not closely related. Comparing the regulation of expression in all three bacteria, only one gene was up-regulated by FS-1 in all bacteria, *copA*, which encodes for heavy metal ion efflux protein.

A



B



C

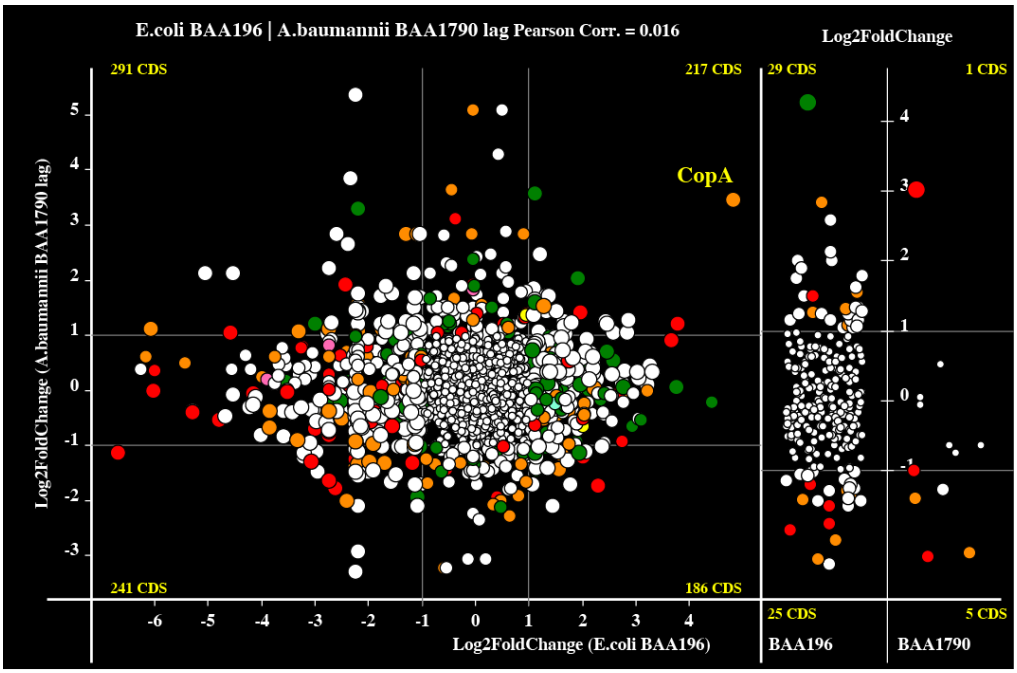


Figure 3.4: Plots representing genes regulated by FS-1 treatment in different model bacteria. Each dot represents a gene, which is coloured based on the functional category it represents: biosynthesis (green); degradation (red); salvage of important molecules (yellow); biosynthesis of cell wall and membrane compounds (pink); gluconeogenesis (blue); transmembrane transportation (brown); and unidentified or other categories (white). A) Regulation of genes in

S. aureus and *E. coli*; B) Regulation of genes in *A. baumannii* and *S. aureus;* C) Regulation of genes in *A. baumannii* and *S. aureus*. The plot on the right represent genes only up or down-regulated in individual bacteria.

A general observation between the three different bacteria was that various important genes were found to be upregulated in the FS-1 treated cultures compared to the NC cultures during the lag phase. See *Supplementary file 5* and *6* for full tables of genes and pathways. The main pathways affected include pathways involved in amino acid biosynthesis and degradation, which produces NADH and NADPH as predicted by the Pathway Tools modelling (*Figure 3.5*). These are vital co-enzymes that are involved in the generation of reactive oxygen species (ROS), anti-oxidative defence mechanisms, and the biosynthesis of crucial cellular components, which include DNA and lipids (Wos and Pollard, 2009, Begley et al., 2001).

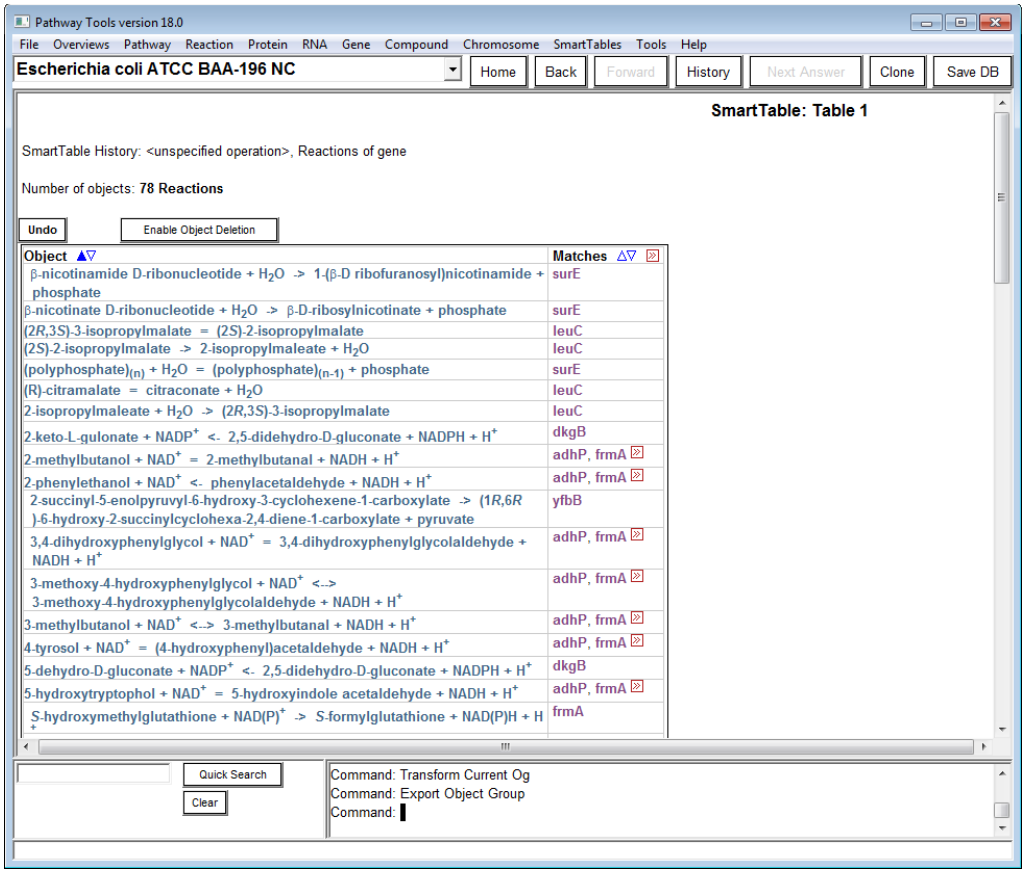


Figure 3.5: Metabolic reactions catalysed by genes, which were up-regulated in model bacteria after treatment with FS-1.

Some of the major pathways affected by suppression of gene expression include the tricarboxylic acid (TCA) cycle (*Figure 3.7*), fatty-acid β-oxidation (*Figure 3.6*), biotin biosynthesis, and reactive oxygen species degradation, and superoxide radical degradation pathways. These pathways play important roles in energy production and oxidative stress in microbial cells. In prokaryotes, fatty acids are broken down in the cytosol by the catabolic process, beta-oxidation. This process generates acetyl coenzyme A (acetyl-CoA), which enters the TCA cycle (as seen in *Figure 3.7*), as well as NADH and FADH2, which are important the electron transport chain (Janßen and Steinbüchel, 2014, Alteri et al., 2012).

The TCA cycle is a crucial metabolic pathway, which is involved in the biosynthesis of important cellular intermediates for anabolic reactions. It is also an important and sensitive pathway with respect to ROS (Baldwin and Krebs, 1981). A previous study has shown that the application of FS-1 resulted in oxidative stress in bacteria (Korotetskiy et al., 2017).

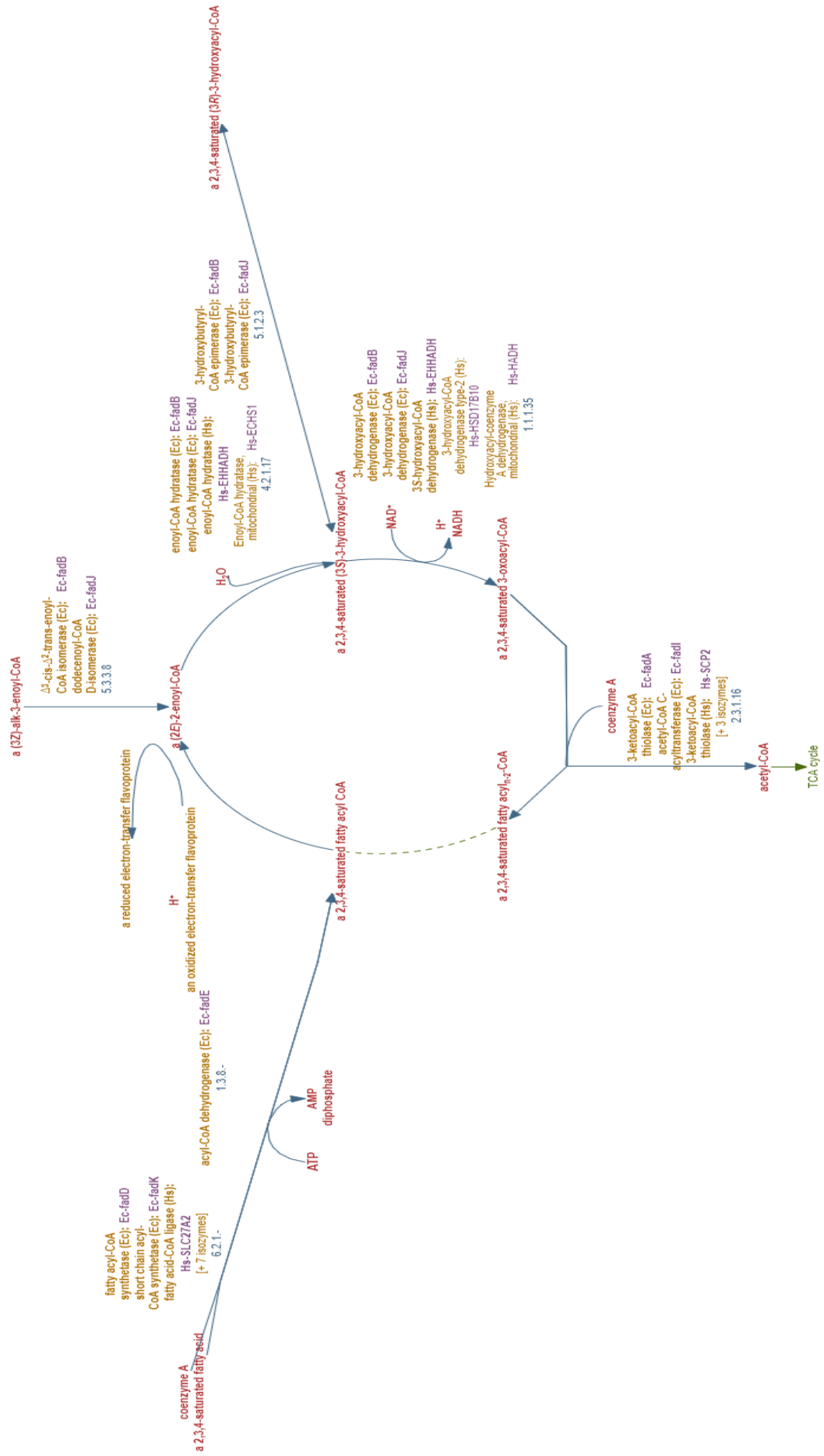


Figure 3.6: adapted from MetaCyc . Pathway for fatty acid β-oxidation in *E. coli*. Important to note in this figure is the production of acetyl-coA, which enters the TCA cycle.

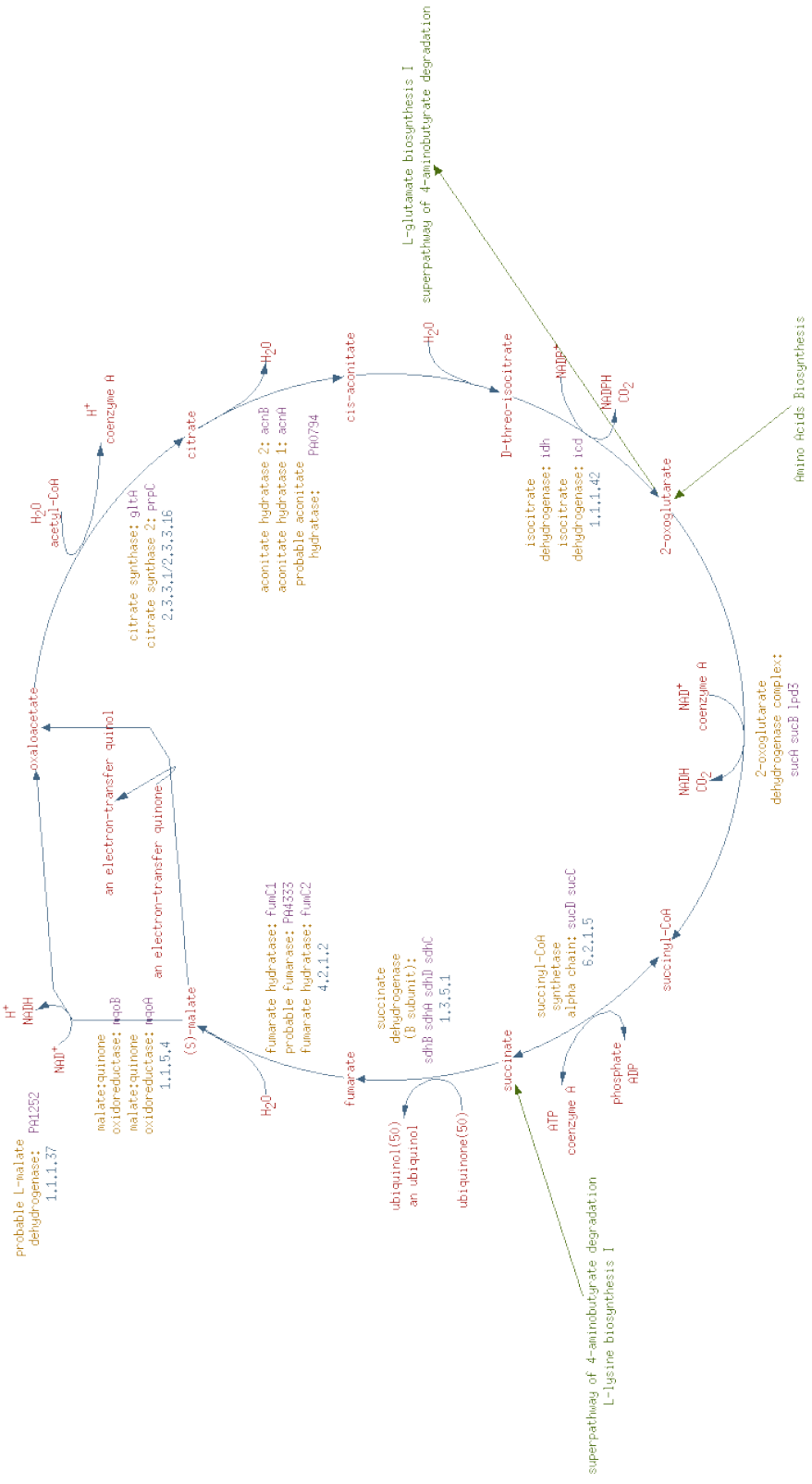


Figure 3.7, adapted from MetaCyc . TCA cycle showing the production of NADH and NADPH.

Another pathway affected by down-regulation of genes is the Entner-Doudoroff pathway. This pathway is one of the major sources of NADH and NADPH, as well as precursors of many metabolic pathways including nucleotide biosynthesis through phosphoribosyl pyrophosphate (PRPP), see *Figure 3.8*.

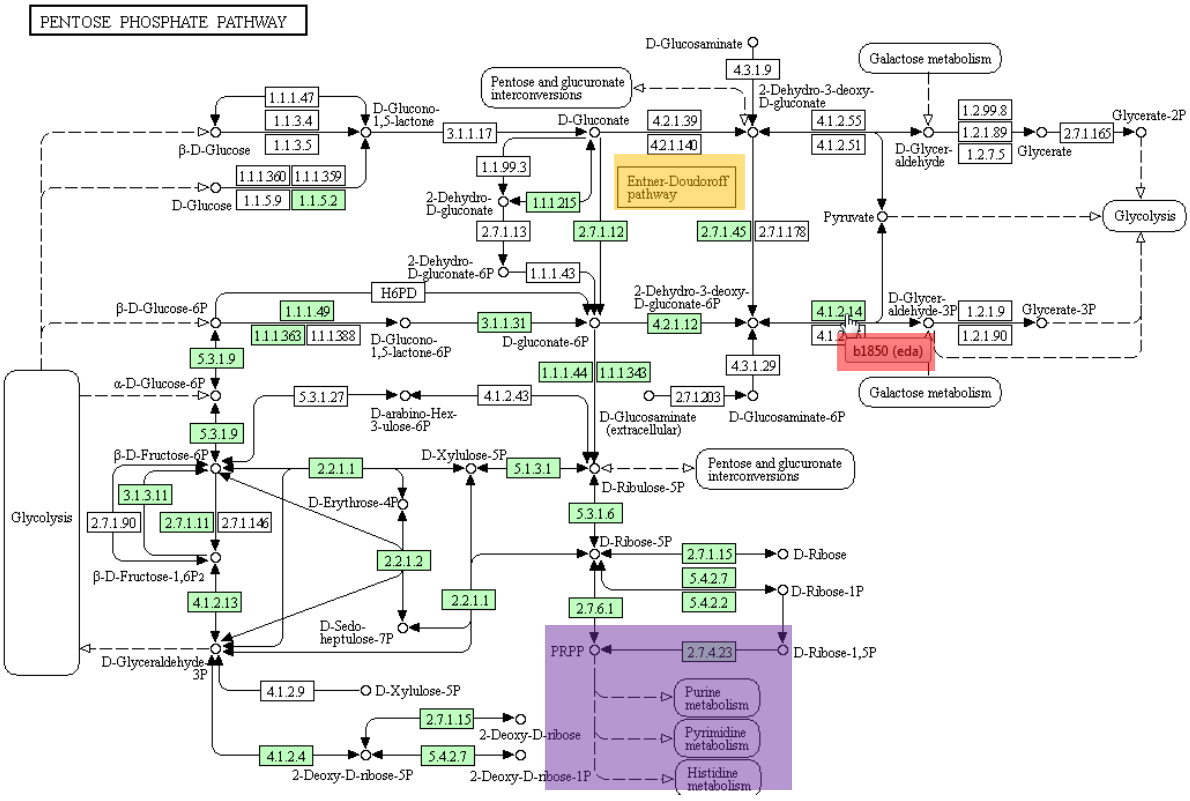


Figure 3.8: Entner-Doudoroff pathway shown in the yellow block. The *eda* gene, shown in red, is down-regulated in all three bacteria. This blocks the path to glycolysis, and drives it towards nucleotide biosynthesis, shown in purple.

The *eda* gene is a gene involved in the degradation of glucose via the Entner-Doudoroff pathway, which helps to source further anabolic processes of biosynthesis of important classes of metabolites, such as nucleic acids, histidine and several other bioactive cyclic compounds. Down-regulation of this gene blocks the shunt to glycolysis, bypassing the Entner-Doudoroff pathway, and driving it towards nucleotide biosynthesis which may be damaged by FS-1.

Long cultivation experiment

In this experiment, bacteria were cultivated on either normal drug free MH medium (negative control), or on the medium containing FS-1. The scheme of the experiment is shown in *Figure 3.2*. This allows the investigation of the effect of FS-1 of the bacterial population.

Adaption of bacteria to the presence of FS-1

To investigate the adaptation of bacteria to the presence of FS-1, bacteria were cultivated on media with FS-1, as well as on normal media without the drug for 10 passages. The general level of gene expression for the bacteria are the same when comparing NC and FS bacterial cultures, however, they differed significantly by the patterns of highly expressed and suppressed genes, see *Figure 3.9*.

A B

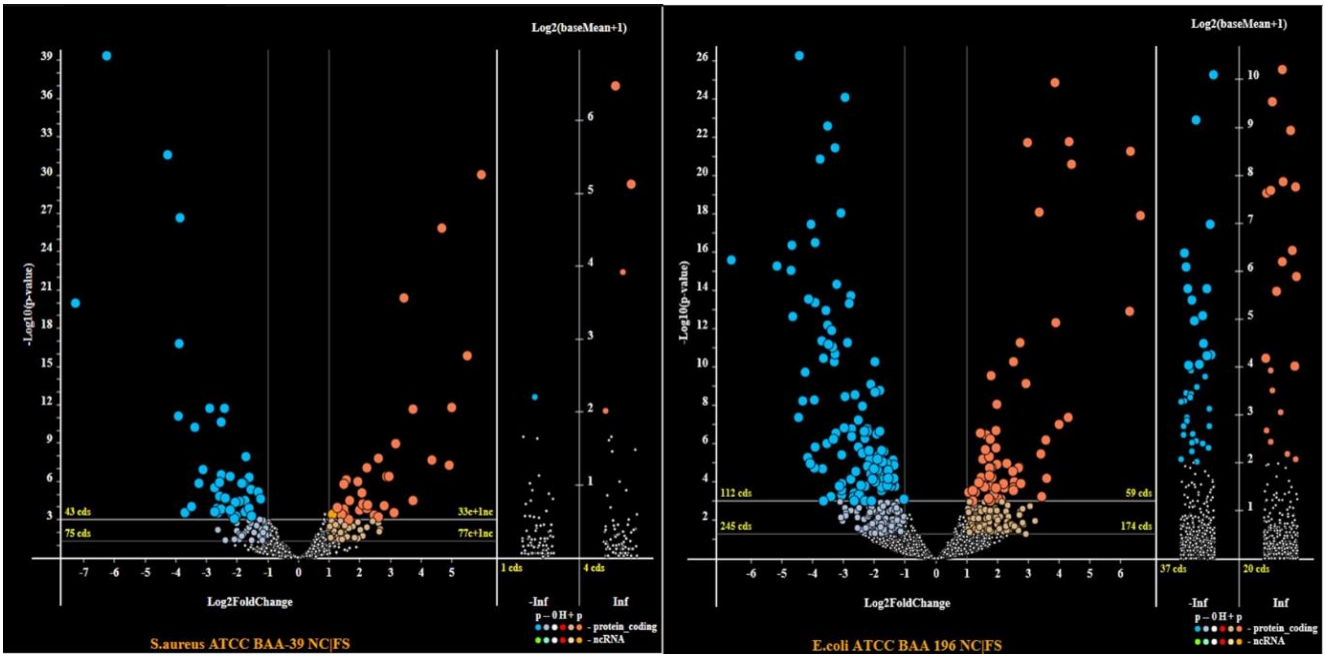
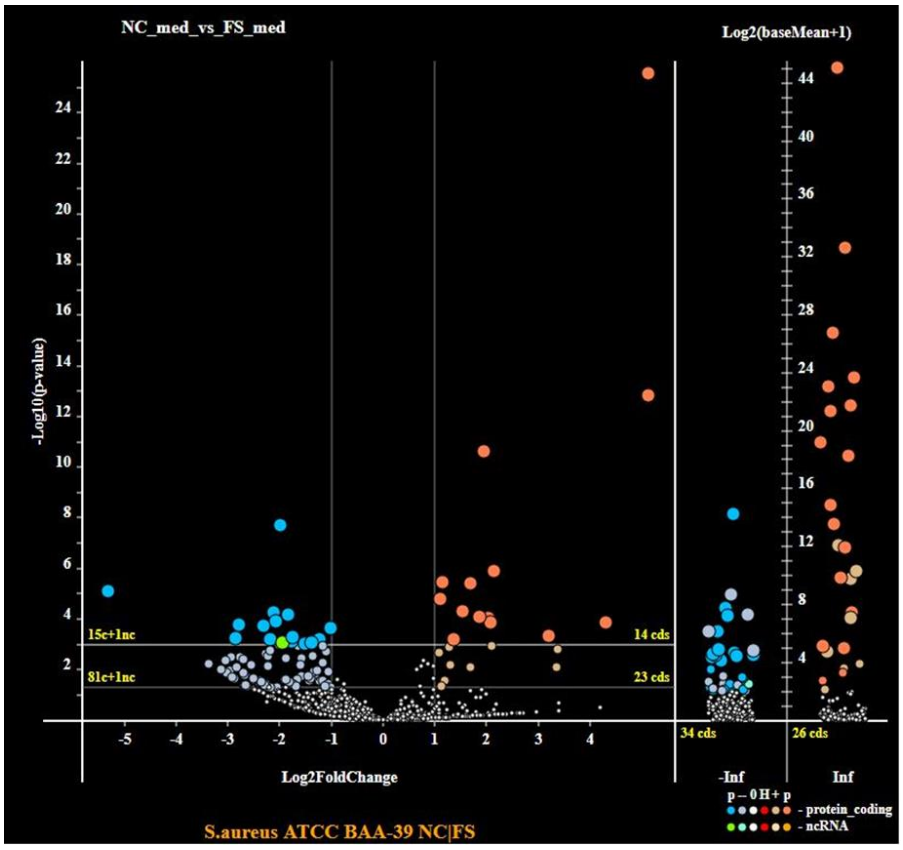


Figure 3.9: Volcano plots representing differential gene expression of A) *S. aureus* and B) *E. coli* cultivated on normal drug free medium and medium containing FS-1.

Effect on bacteria when the cultivation medium is changed

For *S. aureus*, after 10 passages, the experimental and control bacteria were cross-inoculated into tubes with drug-containing and drug-free media for further overnight incubation. *Figure 3.10* shows the volcano plots when changing growth media conditions.



A

B

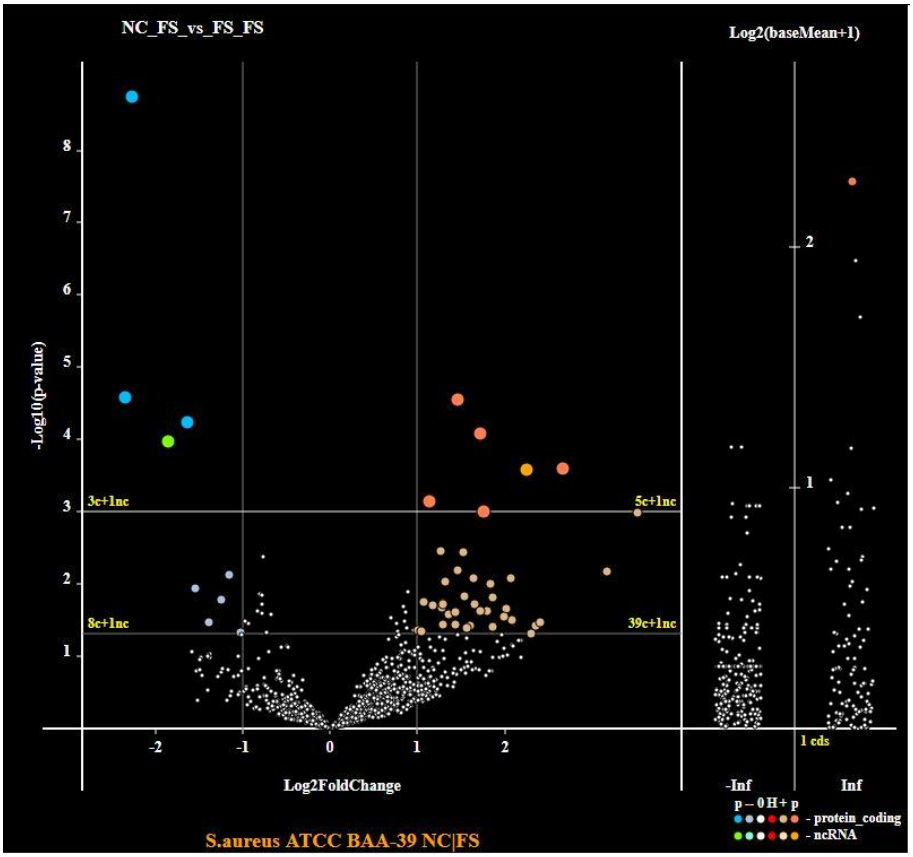


Figure 3.10: Volcano plots representing differential gene expression after A) Substituting the medium of bacteria cultivated on normal medium, to medium with FS-1. B) Substituting the medium of bacteria cultivated on FS-1, to normal drug free medium.

After transferring the FS culture from the drug containing medium to the medium without the drug, strong stress and growth suppression was observed in this culture. Many protein coding genes, which include the methicillin resistance gene *mecR* and *mecA* encoded on SCCmec, were down-regulated. When the NC culture was cultivated on the medium with FS-1, gene expression was also supressed, including *mecR* and *mecA*. Comparing the gene expression patterns of NC and FS cultures grown either on the regular medium or on the medium with FS-1 confirmed that the stress associated with the change of medium was the most important factor shaping the gene expression at these conditions.

Discussion

The observed antibiotic resistance reversion in bacteria cultivated with FS-1 was hypothesized to be induced either due to a direct activity of the iodine-containing drug, or due to bacteria selecting for better adapted genetic variants. In a previous study done by Ilin, Kulmanov [14] et al., the therapeutic effect of FS-1 has been demonstrated. However, the molecular mechanisms of this drug still remain unclear.

Immediate effect experiment

Analysis of changes in gene expression profiles under the influence of FS-1 at different growth phases

A general observation found during the lag phase when the bacterial cultures were exposed to the FS-1 drug for 5 min cultivation, was suppression of gene expression in all three bacteria. Another observation was up-regulation of various important genes. These genes affected pathways involved in amino acid biosynthesis and degradation, which resulted in increased production of important co-enzyme, NADH and NADPH. These co-enzymes are involved in redox reactions, and in conditions were oxidative stress is continuous, important macromolecules can be damaged and cause serious consequences for bacterial physiology (Chiang and Schellhorn, 2012). Oxidation of DNA bases, especially guanine, can lead to mutagenic modifications, DNA strand breaks, as well as cell death (Simandan et al., 1998, Kino et al., 2017).

On the other hand, major pathways that were affected by suppression of gene expression included the TCA cycle, fatty acid β-oxidation, biotin biosynthesis, and reactive oxygen species degradation and superoxide radical degradation pathways. These pathways are important when it comes to energy production and oxidative stress in bacterial cells. Fatty acid β-oxidation generates acetyl-CoA as well as NADH and FADH2. Acetyl-CoA is a crucial

molecule needed in the TCA cycle. The TCA cycle plays an important role in the catabolism of organic molecules to generate energy and reducing power. Furthermore, it is also the first step in generating precursors for biosynthesis . The pathway also produces NADH and NADPH which will transfer their electrons into the electron transport chain to generate ATP through oxidative phosphorylation.

Another important finding was down-regulation of *eda* gene that prevents shifting metabolites from the anabolic Entner-Doudoroff pathway to metabolic glycolysis. The Entner-Doudoroff pathway is another major source of NADH and NADPH, as well as precursors of many metabolic pathways including nucleotide biosynthesis through PRPP.

Redox homeostasis is important and required to harness reducing power, which is generated by catabolic processes, and used in the anabolism of molecules such as DNA, proteins, and lipids. Other studies reported that overflow metabolism and ROS formation are an inherent cellular responses to antibiotic lethality in *E. coli* (Dwyer et al., 2014).

Ultimately, FS-1 caused down-regulation of various important pathways which consume the co-enzymes NADH and NADPH in the bacteria. This could explain the response of bacteria to up-regulate genes involved in amino acid biosynthesis and degradation, to ultimately produce these important co-enzymes, probably in order to withstand an oxidative stress. A previous study also investigated the effect of FS-1 on *S. aureus*, and found that treatment with this drug caused oxidative stress in bacteria (Korotetskiy et al., 2017). Investigating functions of the regulated genes in this study also suggested that bacteria responded primarily to an increased oxidative stress.

Does addition of FS-1 to the medium cause similar gene regulation response in all bacteria?

As shown in *Figure 3.4*, one gene, *copA*, was positively regulated in all three bacteria. This gene encodes for heavy metal ion efflux proteins. This gene was possibly up-regulated in

bacteria to contribute to the removal of heavy anions, since the FS-1 drug contains iodine. These proteins could possibly play a role in halogen resistance. These proteins can also be considered as a target for supplementary drugs to increase the effect of FS-1. Other commonly up- and down-regulated genes were involved in various pathways, particularly in the Entner-Doudoroff pathway. Regulation of this pathway was most likely associated with a redirection of the bacterial metabolism towards producing excess reduced species of co-enzymes NADH and NADPH to cope with the oxidative stress, and towards synthesis of cell membrane proteins and nucleotides, which could have been damaged by direct halogenation with iodine.

Long cultivation experiment

Adaption of bacteria to the presence of FS-1

When comparing the NC and FS cultures cultivated over 10 passages, the gene expression levels were similar in both strains, however, the patterns of highly expressed and suppressed genes differed significantly. This difference demonstrates the adaptation of *S. aureus* and *E. coli* to sub-lethal doses of FS-1. This adaptation required an alternative gene transcription regulation which could either be due to accumulation of specific mutations, or due to epigenetic modifications.

Effect on bacteria when the cultivation medium is changed

After 10 passages, *S. aureus* cultures were cross-inoculated into tubes with drug-containing and drug-free medium for further overnight incubation. In *Figure 3.*10, the FS culture was transferred from the drug containing medium to the medium without the drug for overnight cultivation. Removal of the drug surprisingly caused strong stress and growth suppression of the culture, as 81 protein coding, which includes the methicillin resistance genes *mecR* and *mecA*, and 1 non-coding genes were downregulated in the FS culture. Substituting the medium of NC cultures, to medium containing FS-1, also caused gene suppression in the culture.

Treatment with FS-1 ultimately leads to suppression of gene, including important resistance gene, such as *mecR* and *mecA*. Down-regulation of these genes can allow antibiotics to effectively kill both susceptible and resistant bacteria.

References

MetaCyc [Online]. Available: https://biocyc.org/META/new-image?object=FAO-PWY [Accessed 8 June 2019].

ALTERI, C. J., HIMPSL, S. D., ENGSTROM, M. D. & MOBLEY, H. L. T. 2012. Anaerobic Respiration Using a Complete Oxidative TCA Cycle Drives Multicellular Swarming in Proteus mirabilis. *mBio,* 3**,** e00365-12.

BALDWIN, J. E. & KREBS, H. 1981. The evolution of metabolic cycles. *Nature,* 291**,** 381-2.

BEGLEY, T. P., KINSLAND, C., MEHL, R. A., OSTERMAN, A. & DORRESTEIN, P. 2001. The biosynthesis of nicotinamide adenine dinucleotides in bacteria. *Vitamins & Hormones.* Academic Press.

BERGSTROM, S. & NORMARK, S. 1979. Beta-lactam resistance in clinical isolates of Escherichia coli caused by elevated production of the ampC-mediated chromosomal beta-lactamase. *Antimicrob Agents Chemother,* 16**,** 427-33.

BIOCONDUCTER. *DESeq2* [Online]. Available: https://bioconductor.org/packages/release/bioc/html/DESeq2.html [Accessed 22 March 2019].

BIOCONDUCTER. *Rsubread* [Online]. Available: https://bioconductor.org/packages/release/bioc/html/Rsubread.html [Accessed 10 March 2019].

CHIANG, S. M. & SCHELLHORN, H. E. 2012. Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Archives of Biochemistry and Biophysics,* 525**,** 161-169.

CRICK, F. 1970. Central dogma of molecular biology. *Nature,* 227**,** 561-3.

DWYER, D. J., BELENKY, P. A., YANG, J. H., MACDONALD, I. C., MARTELL, J. D., TAKAHASHI, N., CHAN, C. T., LOBRITZ, M. A., BRAFF, D., SCHWARZ, E. G., YE, J. D., PATI, M., VERCRUYSSE, M., RALIFO, P. S., ALLISON, K. R., KHALIL, A. S., TING, A. Y., WALKER, G. C. & COLLINS, J. J. 2014. Antibiotics induce redox-related physiological alterations as part of their lethality. *Proc Natl Acad Sci U S A,* 111**,** E2100-9.

GIRGIS, H. S., HOTTES, A. K. & TAVAZOIE, S. 2009. Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One,* 4**,** e5629.

GOTTESMAN, S. 2017. Post-transcriptional Regulation and the Bacterial Response to Stress. *The FASEB Journal,* 31**,** 22.1-22.1.

JANßEN, H. J. & STEINBÜCHEL, A. 2014. Fatty acid synthesis in Escherichia coli and its applications towards the production of fatty acid based biofuels. *Biotechnology for Biofuels,* 7**,** 7.

KARP, P. D., PALEY, S. & ROMERO, P. 2002. The Pathway Tools software. *Bioinformatics,* 18 Suppl 1**,** S225-32.

KINO, K., HIRAO-SUZUKI, M., MORIKAWA, M., SAKAGA, A. & MIYAZAWA, H. 2017. Generation, repair and replication of guanine oxidation products. *Genes and environment : the official journal of the Japanese Environmental Mutagen Society,* 39**,** 21-21.

KOROTETSKIY, I., SHILOV, S., SHVIDKO, S., JUMAGAZIYEVA, A., SULDINA, N. A., KOROTETSKAYA, N. V., ILIN, A. & REVA, O. 2017. *Transcriptional response of the multidrug resistant Staphylococcus aureus following FS-1 exposure*.

KOUTSOLIOUTSOU, A., PENA-LLOPIS, S. & DEMPLE, B. 2005. Constitutive soxR mutations contribute to multiple-antibiotic resistance in clinical Escherichia coli isolates. *Antimicrob Agents Chemother,* 49**,** 2746-52.

KUKURBA, K. R. & MONTGOMERY, S. B. 2015. RNA Sequencing and Analysis. *Cold Spring Harbor protocols,* 2015**,** 951-969.

LEVY, S. B. & MARSHALL, B. 2004. Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine,* 10**,** S122.

SIMANDAN, T., SUN, J. & DIX, T. A. 1998. Oxidation of DNA bases, deoxyribonucleosides and homopolymers by peroxyl radicals. *Biochem J,* 335 (Pt 2)**,** 233-40.

TAMAE, C., LIU, A., KIM, K., SITZ, D., HONG, J., BECKET, E., BUI, A., SOLAIMANI, P., TRAN, K. P., YANG, H. & MILLER, J. H. 2008. Determination of antibiotic hypersensitivity among 4,000 single-gene-knockout mutants of Escherichia coli. *J Bacteriol,* 190**,** 5981-8.

TENOVER, F. C. & HUGHES, J. M. 1996. The challenges of emerging infectious diseases: Development and spread of multiply-resistant bacterial pathogens. *JAMA,* 275**,** 300-304.

WANG, Y. C. & LIPSITCH, M. 2006. Upgrading antibiotic use within a class: tradeoff between resistance and treatment success. *Proc Natl Acad Sci U S A,* 103**,** 9655-60.

WOODFORD, N. & ELLINGTON, M. J. 2007. The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection,* 13**,** 5-18.

WOS, M. L. & POLLARD, P. C. 2009. Cellular nicotinamide adenine dinucleotide (NADH) as an indicator of bacterial metabolic activity dynamics in activated sludge. *Water Sci Technol,* 60**,** 783-91.

Detection of epigenetic modifications in the multidrug resistant strains *S. aureus* and *E. coli* treated with the drug, FS-1

Abstract

This misuse of antibiotics as well as inappropriate prescribing and overuse of antibiotics, has led to a strong selective pressure, resulting in the survival and wide distribution of multidrug resistant pathogens. Drug induce reversion of antibiotic resistant bacteria into sensitive phenotypes is a prospective approach to target the mechanisms and evolution of bacterial resistance. The model multidrug resistant bacteria, *Escherichia coli* ATCC BAA-196 and *Staphylococcus aureus* ATCC BAA-39, was used in this study. Bacteria were cultivated on medium with and without the FS-1 drug over 10 passages. Afterwards DNA was extracted and sequenced using the PacBio RS II sequencer. Modified bases were then detected using tools available from the SMRT-link software. The program can detect any difference between methylated nucleotides and other modifications of unknown nature. In both *E. coli* and *S. aureus*, bacterial cultures treated with FS-1 had an overall increase in modified bases, while the number of methylated nucleotides remained unchanged. This was specifically observed in G and A bases. It was hypothesized that the observed sporadically modified nucleotides might be due to oxidation, especially of G bases, by the iodine contained in FS-1.

Introduction

Over the years, the misuse and overuse of antibiotics has created a strong selective pressure, resulting in the survival and wide distribution of drug resistant microorganisms, threatening the effectiveness of antibiotics, which is becoming a major concern to the public health (Abraham et al., 1992, Barbosa and Levy, 2000). There are various factors involved in the emergence and spread of multiple resistant bacteria. These factors include mutations, horizontal gene transfer (HGT), selective pressure in healthcare and community settings, and, in some cases, the inability to detect emerging resistance phenotypes (Tenover and Hughes, 1996).

Bacteria are constantly faced with the challenge to maintain fitness in unfavourable conditions, which include oxidative stress, DNA damage, and heat. In response, they alter their phenotypes by modulating gene expression levels. An important factor involved in the generation of genetic variation, is mutation of an organism's DNA. Phase variation is a heritable, usually reversible, process that can help bacteria to rapidly adapt to changing environments. It is a form of gene regulation that involves altering between low and high levels of gene expression and can generate diverse bacterial populations in a relatively short period of time (Willems et al., 1990, Seifert, 1996). Bacteria have phase variable genes, referred to as contingency genes, that are highly mutable, compared to housekeeping genes, that have lower mutation rates (Moxon et al., 1994). Various molecular mechanisms influence phase variation, either genetically, involving a change in the DNA sequence, or epigenetically, involving methylation of DNA at specific loci (van der Woude and Bäumler, 2004).

Slipped-strand mispairing (SSM) is a mechanism that can promote phase variation. SSM results in mispairing of short repeat sequences between mother and daughter strands during DNA replication (Henderson et al., 1999, Levinson and Gutman, 1987, Kim et al., 2000). Alterations in these short sequence repeat (SSR) regions that are located either within a gene (influencing

translational reading frames) or upstream of a gene (altering transcription), can alter gene expression (Chandler and Fayet, 1993, van Belkum et al., 1998). SSM is not the only mechanism of phase variation. There are other less common mechanisms, such as homologous recombination, insertion/excision of transposons, site specific recombination, and differential DNA methylation.

DNA methylation enables bacteria to epigenetically control the reversible ON/OFF switching of important genes. An example of this includes the pyelonephritis associated pili (*pap*) operon, controlled by DNA adenine methylase (Dam). Switching between the ON/OFF state of the *papBA* genes determines the binding of two proteins at two *GATC* sites, before and after the promotor. If these proteins bind to a site, methylation at that site is blocked. The operon is turned to the ON state when methylation occurs proximal to the promotor, and vice versa (van der Woude and Bäumler, 2004, Hernday et al., 2002). Another example includes the *flu* gene, encoding for the outer membrane protein Ag43 in *E. coli*. This gene is regulated at three *GATC* sites and its expression is repressed by the oxidative stress response protein, OxyR. Binding of OxyR to a *GATC* site, masks this site, therefore blocking methylation by Dam, and turning expression to the OFF state (Haagmans and van der Woude, 2000).

Studies involving a new iodine-containing drug, FS-1, have reported the reversion of antibiotic resistance in extensively drug-resistant strains of *Mycobacterium tuberculosis* after treatment with this drug (Ilin et al., 2017). This drug has been introduced into clinical practice as a supplement for antibiotic treatment. Previous studies have recently indicated that the conflicting effects of antibiotics can be decoupled by combining specific compounds (Baym et al., 2016). With an increasing occurrence in multiple drug resistant bacteria, monotherapy treatment is gradually becoming less effective, necessitating the use of drug combination therapies. Strategies to combat antibiotic resistance with combination drug therapy has shown to be promising since the late 1940's, where coadministration of streptomycin and para-

aminosalicylic acid showed reduced evolution of resistant *Mycobacterium tuberculosis* (Dunner et al., 1949). Drug combinations are also currently being used in most cancer treatments (Lane, 2006, Bayat Mokhtari et al., 2017), treatment of HIV infected patients (Scourfield et al., 2011), and generally the most effective treatment for malaria is artemisinin-based combinations (Nosten and White, 2007).

Therefore, it is of great importance to devise strategies that focus on the application of supplementary drugs to increase susceptibility to regular antibiotics by inhibiting bacterial growth while reversing the selection for resistance.

Although FS-1 was primarily developed to supplement treatment against multidrug resistant *Mycobacterium tuberculosis*, further laboratory experiments demonstrated the antibiotic resistance reversion on several other multidrug-resistant bacteria, which include *S. aureus* (Joubert et al., 2019) and *E. coli* (Korotetskiy et al., 2017). *S. aureus* and *E. coli* multidrug-resistant isolates were used as model organisms in this study, since they are more convenient laboratory models to study antibiotic resistance reversion rather than the extensively drug-resistant *M. tuberculosis* clinical isolates.

The aim of this chapter was to use the DNA reads generated by PacBio, together with tools available in the SMRT-link software, to detect epigenetically modified bases between bacteria cultivated on medium with and without the FS-1 drug.

Materials and Methods

Bacterial cultures

The model multidrug resistant bacteria *Escherichia coli* ATCC BAA-196 and *Staphylococcus aureus* ATCC BAA-39 was obtained from the ATCC collection and kept in a freezer at −80°C. Bacteria were cultivated on Mueller-Hinton (MH) liquid medium (Himedia, India) supplemented with ceftazidime 10 µg/ml as recommended in the ATCC product sheet.

Cultivation with FS-1

For experimental cultures, bacteria were inoculated into test-tubes with 10 ml of liquid MH medium supplemented with the FS-1 drug. As a negative control, the cultures were cultivated in the same medium without the drug. Test-tubes were incubated at 37°C for 24 h and then 0.1 ml aliquots of the cultures were transferred to fresh tubes with the corresponding media. After 10 passages, the experimental and control bacteria were cross-inoculated in three repeats into tubes with drug-containing and drug-free media for further overnight incubation, followed by DNA extraction.

DNA extraction and PacBio sequencing

DNA samples were extracted from bacterial cells using PureLink Genomic DNA Kits (Publication Number: MAN0000601, Revision 2.0) following the manufacturer's recommendations. Samples were prepared according to a guide for preparation of SMRTbell templates for sequencing on the PacBio RS II System. The samples were sequenced at Macrogene (South Korea) on SMRT Cell 8Pac V3 cells using the DNA Polymerase Binding Kit P6 and DNA Sequencing Reagent 4.0 v2 following the SMRTbell 20-kb library preparation protocol.

Profiling of epigenetic modifications

A pipeline, which makes use of various tools available in the SMRT-link 6.0.0 software (Vorobiof et al., 2001), was created and used with an in-house Python script to do base-call kinetic analysis on the PacBio reads generated from chromosomal DNA. This pipeline consists of the following steps; The complete genome consensus sequence in *FASTA* format, obtained by assembling PacBio reads, was indexed by the program *samtools*, and used as the reference sequence for PacBio read alignments; The *bax2bam* tool was used to convert the original PacBio reads, obtained in *BAX.H5* format, to *BAM* format. Reads stored in *BAM* files were aligned against the indexed reference sequence by the tool *blasr*. The aligned reads in *BAM* format were sorted by locations and indexed by *samtools sort* and *index* functions. For the following step, the sorted and indexed *BAM* files were analyzed by the tool *ipdSummary* to estimate base call kinetics for every nucleotide in the reference genome (the output file **_kinetics.csv*). Also, for every nucleotide position with a significant base call delay in multiple overlapped reads, the tool calculated several statistical parameters such as IPD ratio of the average base call time to the expectation; and the quality values (QV) score. QV score 14 corresponds to p-value 0.05, and QV score 21 to a p-value 0.01 of statistical reliability of base modification identification. The program stores all the estimated parameters together with context sequences into an output file **_basemods.gff*. Contextual motifs of base modifications were searched by the tool *motifMaker*. Thereafter, the epigenetic profiles of the studied genomes were visualized using an in-house Python script which uses the **_kinetics.csv* and **_basemods.gff* output files.

Results

With the use of PacBio sequencing, epigenetically modified bases could be detected by using the base call kinetics analysis pipeline. This is a useful technique to investigate base modification profiling of bacterial genomes. The software calculates the likelihood of an epigenetic modification at a given site as Phred scores, where values above 21 corresponds to the confidence of a p-value equal to 0.01 or smaller. Long-read PacBio sequencing was performed in three repeats for *S. aureus* and *E. coli*, and average scores were calculated for every nucleotide site in the genome. The SMRT-link software classifies modified bases into 6 categories: methylated sites m6A and m4C, and modified bases A, T, G and C of unknown nature. Diagrams were generated for QV scores and coverages for each bacterium. Only bases which had QV scores of 20 or more, and that were present in all three repeats were taken into consideration.

Detection of epigenetic modifications in *S. aureus*

The QV modification score versus coverage scatterplots calculated for *S. aureus* NC and FS genomes are shown in *Figure 4.1-A* and 4.1-*B*, respectively. Base modifications were predicted independently in the three repeated samples generated from the NC and FS genomes. In *Figure 4.1*, modified bases were plotted as dots according to their average values of QV scores and coverage values in the three repeats. Modifications of different types were denoted by dots of different styles.

As seen in *Figure 4.1* below, treatment with FS-1 increased the overall number of modified bases. The total number of modified bases in NC was 18 114, which increased to 22 133 in the FS cultures. However, the NC modified sites generally had higher QV scores compared to FS. Cultures treated with FS-1 also had the highest increase in adenosine (A) and guanosine (G) modified bases compared to the NC cultures.

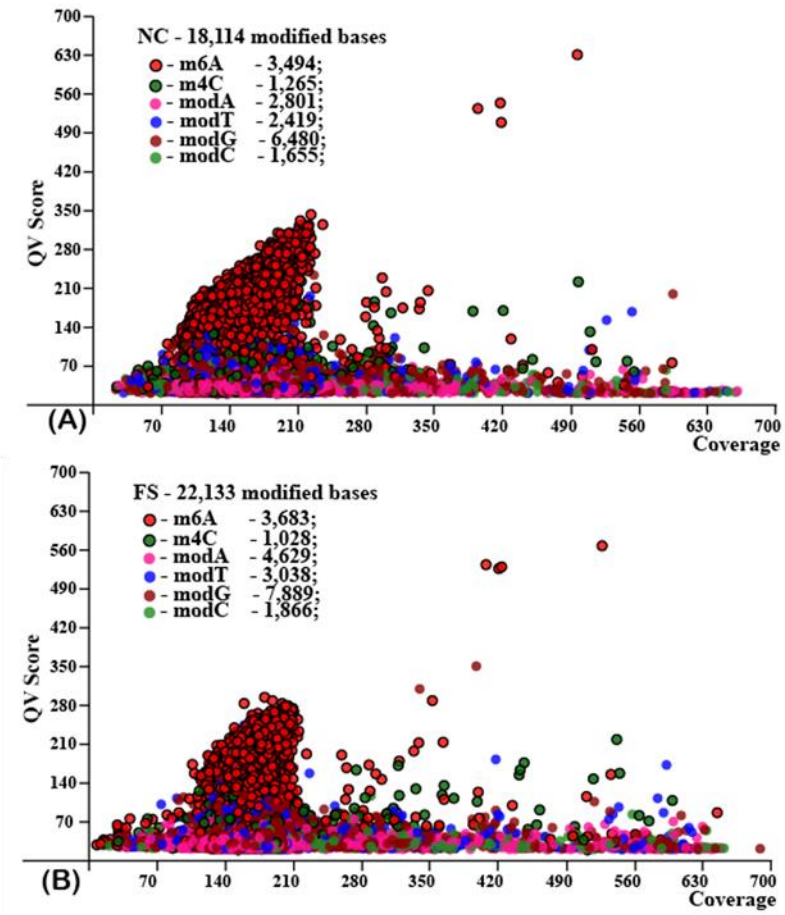


Figure 4.1: Scatterplots representing QV modification scores versus coverage for *S. aureus* A) NC and B) FS genomes. Each dot represents a base modification and is plotted by the average value of the QV modification scores and coverage values in three repeats. Nucleotide modifications of different types are denoted by dots of different styles as shown in the legend together with the frequencies of each modification type.

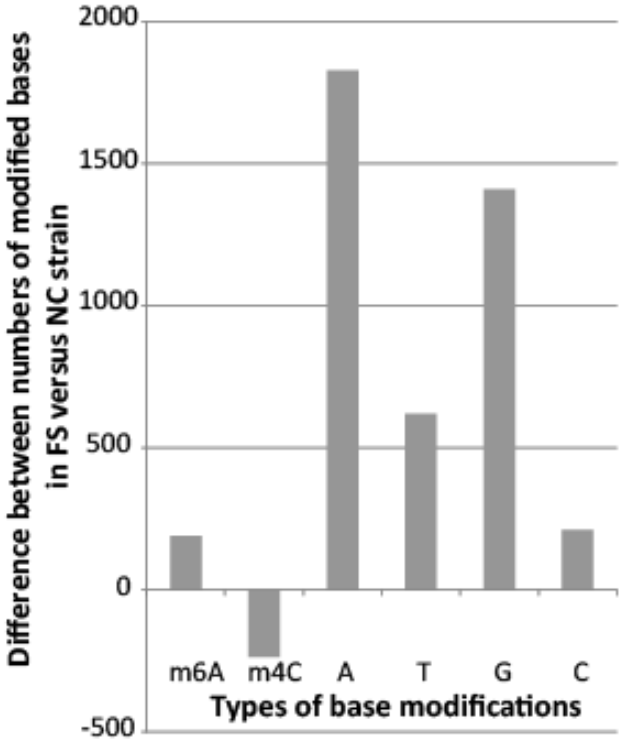


Figure 4.2: The differences in absolute numbers of modified bases of different types in FS versus NC.

In *Figure 4.1* and *Figure 4.2* it can be seen that the level of DNA methylation between the NC and FS cultures were similar in both genomes. The number of m6A residues has increased insignificantly in FS and the level of m4C methylation has decreased. However, the level of base modifications of unknown nature has increased significantly under the effect of cultivation with FS-1.

Detection of epigenetic modifications in *E. coli*

Rank diagrams of the distribution of modification scores in NC and FS *E. coli* strains, are shown in *Figure 4.3* below. In the NC genome (*Figure 4.3-A*), there is a progressive decrease in the number of higher scored G residues, while the A residue line shows an increase of nucleotides with modification scores in the range 110 – 260. *Figure 4.3-B*, representing the FS culture, illustrates how treatment with FS-1 influenced the line of guanine scores by a significant increase in the number of modified residues.

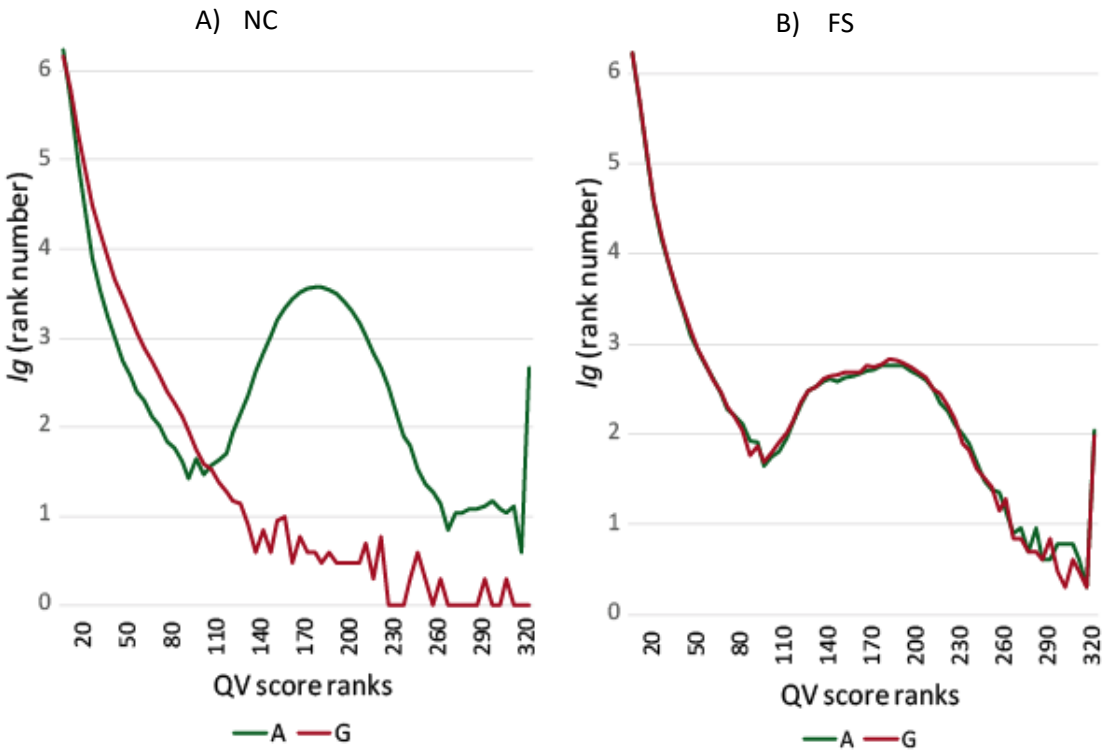


Figure 4.3: Distribution of modified nucleotides in *E. coli* after treatment with FS-1 (FS) and in the negative control variant (NC). A) and B) represent QV modification score distribution diagrams in the NC and FS genomes, respectively. Axis X depicts score ranks stepping 5 score units. Axis Y indicates decimal logarithms of rank numbers.

In both *S. aureus* and *E. coli*, an increase in modified A and G bases were observed when cultures were treated with FS-1. These modifications could be caused by oxidation (*Figure 4.4*) or iodination of bases, especially G bases.

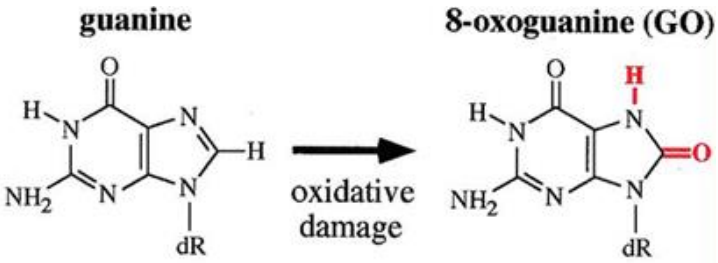


Figure 4.4, adapted from (Lanier and Williams, 2017): Reaction illustration the oxidation of guanine, forming 8-oxoguanine.

Discussion

Experimental and theoretical studies have recently indicated that the conflicting effects of antibiotics can be decoupled by combining specific compounds (Baym et al., 2016). With an increasing occurrence in multiple drug resistant bacteria, monotherapy treatment is gradually becoming less adequate, necessitating the use of drug combination therapies. Therefore, it is of great importance to devise strategies that focus on the application of supplementary drugs to increase susceptibility to regular antibiotics by inhibiting bacterial growth while reversing the selection for resistance. Drug induce reversion of antibiotic resistant pathogens into sensitive phenotypes is a prospective approach to target the mechanisms and evolution of bacterial resistance (Baym et al., 2016).

Cultures treated with FS-1 showed significant epigenetic changes. The detected modified bases are most likely integral parts of the induction of antibiotic resistance reversion. The first observation was that the overall total number of modified nucleotides has increased in cultures treated with FS-1 compared to the negative control. Another finding was an increased in the number of A and G modifications after treatment with the drug observed in both *S. aureus* and *E. coli* FS cultures.

When comparing the level of DNA methylation between the NC and FS cultures, only minor differences were observed. The FS-1 treated cultures had 189 more modified m6A residues, whereas NC cultures had 237 more modified m4C residues. However, the level of base modifications of unknown nature has increased significantly under the effect of cultivation with FS-1. The total number of modified bases increased in FS, but mostly owing to the sites with low QV scores.

The adaptation for bacteria to the presence of FS-1 possibly required an alternative gene regulation acquired either from accumulation of specific mutations, or due to epigenetic

modifications. It was hypothesized that the nature of these modifications could be iodination and/or oxidation of G and A residues to 8-oxoguanosine and 8-oxoadenosine. The latter modifications have been shown to delay the SMRT basecalling (Rehm, 2013).

Previous studies shown that the application of FS-1 caused an oxidative stress in bacteria (Korotetskiy et al., 2017). Changing of the redox potential in proximity with the chromosomal DNA may promote the oxidation of nucleotides. An alternative hypothesis may be that modG modifications correspond to O-6-guanine methylation which resulted from an abnormal activity of DNA methylases. O-6-methylguanosine is highly mutagenic, provoking single nucleotide mismatches during replication (Gu et al., 2017). However, delayed SMRT basecalling by O-6-MeG has not been previously reported.

The resulted damage of nucleotides in genomic DNA, and possibly in mRNA molecules, may reduce general fitness of bacteria and thus make them more sensitive to antibiotics. Furthermore, epigenetic modifications of nucleotides that are irregular in bacterial genomes may lead to interference of gene regulation by transcriptional factors and non-coding regulatory RNA. It may be concluded that the treatment with FS-1 significantly increases the frequency of sporadically distributed modified nucleotides while the general pattern of the global genome methylation remained stable.

References

ABRAHAM, E. P., CHAIN, E., FLETCHER, C. M., FLOREY, H. W., GARDNER, A. D., HEATLEY, N. G. & JENNINGS, M. A. 1992. Further observations on penicillin. 1941. *Eur J Clin Pharmacol,* 42**,** 3-9.

BARBOSA, T. M. & LEVY, S. B. 2000. The impact of antibiotic use on resistance development and persistence. *Drug Resist Updat,* 3**,** 303-311.

BAYAT MOKHTARI, R., HOMAYOUNI, T. S., BALUCH, N., MORGATSKAYA, E., KUMAR, S., DAS, B. & YEGER, H. 2017. Combination therapy in combating cancer. *Oncotarget,* 8**,** 38022-38043.

BAYM, M., STONE, L. K. & KISHONY, R. 2016. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science,* 351**,** aad3292.

CHANDLER, M. & FAYET, O. 1993. Translational frameshifting in the control of transposition in bacteria. *Mol Microbiol,* 7**,** 497-503.

DUNNER, E., BROWN, W. B. & WALLACE, J. 1949. The effect of streptomycin with para-amino salicylic acid on the emergence of resistant strains of tubercle bacilli. *Dis Chest,* 16**,** 661-6.

GU, S., XIONG, J., SHI, Y., YOU, J., ZOU, Z., LIU, X. & ZHANG, H. 2017. Error-prone bypass of O(6)-methylguanine by DNA polymerase of Pseudomonas aeruginosa phage PaP1. *DNA Repair (Amst),* 57**,** 35-44.

HAAGMANS, W. & VAN DER WOUDE, M. 2000. Phase variation of Ag43 in Escherichia coli: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol Microbiol,* 35**,** 877-87.

HENDERSON, I. R., OWEN, P. & NATARO, J. P. 1999. Molecular switches--the ON and OFF of bacterial phase variation. *Mol Microbiol,* 33**,** 919-32.

HERNDAY, A., KRABBE, M., BRAATEN, B. & LOW, D. 2002. Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci U S A,* 99 Suppl 4**,** 16470-6.

ILIN, A. I., KULMANOV, M. E., KOROTETSKIY, I. S., ISLAMOV, R. A., AKHMETOVA, G. K., LANKINA, M. V. & REVA, O. N. 2017. Genomic Insight into Mechanisms of Reversion of Antibiotic Resistance in Multidrug Resistant Mycobacterium tuberculosis Induced by a Nanomolecular Iodine-Containing Complex FS-1. *Front Cell Infect Microbiol,* 7**,** 151.

JOUBERT, M., REVA, O. N., KOROTETSKIY, I. S., SHVIDKO, S. V., SHILOV, S. V., JUMAGAZIYEVA, A. B., KENESHEVA, S. T., SULDINA, N. A. & ILIN, A. I. 2019. Assembly of Complete Genome Sequences of Negative-Control and Experimental Strain Variants of Staphylococcus aureus ATCC BAA-39 Selected under the Effect of the Drug FS-1, Which Induces Antibiotic Resistance Reversion. *Microbiology Resource Announcements,* 8**,** e00579-19.

KIM, D. W., GAZOURIAN, L., A., Q. S., ROMIEU-MOUREZ, R., SHERR, D. H. & SONENSHEIN, G. E. 2000. The RelA NF-kappaB subunit and the aryl hydrocarbon receptor (AhR) cooperate to transactivate the c-myc promoter in mammary cells. *Oncogene,* 19**,** 5498-506.

KOROTETSKIY, I., SHILOV, S., SHVIDKO, S., JUMAGAZIYEVA, A., SULDINA, N. A., KOROTETSKAYA, N. V., ILIN, A. & REVA, O. 2017. *Transcriptional response of the multidrug resistant Staphylococcus aureus following FS-1 exposure*.

LANE, D. 2006. Designer combination therapy for cancer. *Nat Biotechnol,* 24**,** 163-4.

LANIER, K. & WILLIAMS, L. 2017. The Origin of Life: Models and Data. *Journal of Molecular Evolution,* 84.

LEVINSON, G. & GUTMAN, G. A. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol,* 4**,** 203-21.

MOXON, E. R., RAINEY, P. B., NOWAK, M. A. & LENSKI, R. E. 1994. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr Biol,* 4**,** 24-33.

NOSTEN, F. & WHITE, N. J. 2007. Artemisinin-based combination treatment of falciparum malaria. *Am J Trop Med Hyg,* 77**,** 181-92.

REHM, H. L. 2013. Disease-targeted sequencing: a cornerstone in the clinic. *Nature reviews. Genetics,* 14**,** 295-300.

SCOURFIELD, A., WATERS, L. & NELSON, M. 2011. Drug combinations for HIV: what's new? *Expert Rev Anti Infect Ther,* 9**,** 1001-11.

SEIFERT, H. S. 1996. Questions about gonococcal pilus phase- and antigenic variation. *Mol Microbiol,* 21**,** 433-40.

TENOVER, F. C. & HUGHES, J. M. 1996. The challenges of emerging infectious diseases: Development and spread of multiply-resistant bacterial pathogens. *JAMA,* 275**,** 300-304.

VAN BELKUM, A., SCHERER, S., VAN ALPHEN, L. & VERBRUGH, H. 1998. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev,* 62**,** 275-93.

VAN DER WOUDE, M. W. & BÄUMLER, A. J. 2004. Phase and Antigenic Variation in Bacteria. *Clinical Microbiology Reviews,* 17**,** 581-611.

VOROBIOF, D. A., SITAS, F. & VOROBIOF, G. 2001. Breast cancer incidence in South Africa. *Journal of clinical oncology,* 19**,** 125s-127s.

WILLEMS, R., PAUL, A., VAN DER HEIDE, H. G., TER AVEST, A. R. & MOOI, F. R. 1990. Fimbrial phase variation in Bordetella pertussis: a novel mechanism for transcriptional regulation. *Embo j,* 9**,** 2803-9.

Chapter 5

Concluding remarks

Continued misuse and overuse of antibiotics over the years has led to the development of a strong selective pressure, which resulted in the survival and wide distribution of drug resistant microorganisms, threatening the effectiveness of antibiotics. This is currently becoming a major concern to the public health (Abraham et al., 1992, Barbosa and Levy, 2000). Although existing compounds are continuously being modified and new antibiotic classes are being discovered, the rate of resistance continues to rise while antibiotic discovery rates have started to substantially drop (Laxminarayan, 2014). Therefore, it is of great importance to devise new strategies that focus on limiting, redirecting, and/or reversing the development of bacterial resistance. This study set out to investigate the molecular mechanisms of drug induced resistance reversion, which was observed in three different bacteria after treatment with a new iodine-containing drug, FS-1.

With recent advances in next-generation sequencing (NGS) technologies, reads that are highly accurate can now be generated in a short period of time and at a much lower cost. Although vast amounts of data can be generated with NGS, due to their short-read lengths, there still are some limitations and challenges. With the use of third-generation sequencing (TGS) technologies these limitations can more easily be overcome. TGS has been widely used in genomic studies since the successful launch of the PacBio RS II sequencer. With continued improvements and upgrading, PacBio recently launched the Sequel sequencing system. Some advantages of the sequel system over the RS II system, is that it allows higher-throughput sequencing within a shorter time at a lower cost (Abrams, 2001).

With SMRT sequencing most bacterial genomes are able to be completely resolved, since reads of sufficient length are generated and are able to span complex repeat regions. This allows complete assemblies to be obtained without the requirement of manual finishing (Koren et al., 2013).

For this study, long PacBio reads were generated for the bacterial genomes of multidrug resistant reference cultures *S. aureus*, *E. coli*, and *A. baumannii*. Coverages ranging between 600x and 1000x of the expected genome lengths were obtained and used for de novo assemblies. Complete genomes were obtained for all bacteria and annotated using RAST. Thereafter, genome annotations were manually corrected as needed. All genomes were submitted to the NCBI and have published genome announcements.

This study then set out to investigate gene regulation patterns of the bacteria after cultivation in a sub-lethal dose of FS-1 during different growth phases in comparison to negative control strains. For the immediate effect, treatment with FS-1 caused down-regulation of various important pathways which consume the co-enzymes NADH and NADPH, while the metabolic processes associated with the production of the reduced species of these co-enzymes generally were up-regulated. It may be assumed that the pathways helping bacteria to withstand oxidative stress were upregulated. After cultivation of bacteria in the medium with FS-1 over 10 passages, they regained the initial growth rate by adapting to the presence of FS-1, which required an alternative gene transcription regulation controlled either by accumulation of specific mutations in bacterial populations, or due to epigenetic phase variations.

Lastly, the project focused on epigenetic modifications in bacterial genomes. Various Python scripts were created and used in this study to create pipelines, which included various tools from the SMRT-link software, to work with the obtained SMRT sequencing data. This was not only useful for the current project but can also be useful for future projects which are based on PacBio sequencing. Previously, epigenetic modifications were not well studied, especially in bacteria, due to the lack of efficient methods to detect epigenetic modifications in DNA. One of the remarkable feature of SMRT sequencing is that modifications on every sequenced nucleotide can be detected (Clark et al., 2012). The DNA reads generated by PacBio, together with tools available in the SMRT-link software, were used to detect epigenetically modified

bases between bacteria cultivated on medium with and without the FS-1 drug. Epigenetic modifications in *A. baumannii* were not investigated, since the FS reads provided by our collaborators appeared to be contaminated and were not suitable for further use. In both *E. coli* and *S. aureus*, bacterial cultures treated with FS-1 had an overall increase in modified bases, while the number of methylated nucleotides remained unchanged. This was specifically observed in G and A bases. It was hypothesized that the observed sporadically modified nucleotides might be due to oxidation, especially of G bases, by the iodine contained in FS-1. This can be dangerous for bacteria as the oxidation product, *7,8-dihydro-8-oxoguanine*, is strongly mutagenic due to its complementation with both cytosine and adenine (Neeley and Essigmann, 2006).

Oxidized guanosine residues on the chromosome can lead to double-strand DNA breaks and cell death. Lethal double-strand DNA breaks and RNA mistranslation can be caused by closely spaced 8-oxo-deoxyguanosine lesions, ultimately leading to the cytotoxicity of many broadly used antibiotics, which include β-lactams, quinolones, and aminoglycosides (Foti et al., 2012). a Significant increase of frameshift mutations was observed while investigating the assembled genomes of the FS-1 treated bacterial cultures. Furthermore, we observed the adaptation of bacteria to the presence of FS-1, which required an alternative gene transcription regulation, either due to accumulation of specific mutations, or due to epigenetic modifications. Therefore, these findings conclude that treatment with FS-1 possibly leads to DNA oxidation, especially of G bases, which caused frame shift mutations, as well as alternative gene transcription regulation. This is a possible explanation of how resistant bacteria which were treated with FS-1, had increased sensitivity to antibiotics. FS-1 is inherited by bacterial daughter cells in several generations, and increased the sensitivity to antibiotics are remained, even when FS-1 is removed from the medium.

These are however only theoretical findings, for future studies these findings need to be tested in the laboratory. It will be important that future research focus on further investigation of the observed epigenetic modifications. Since DNA modification detection of *A. baumannii* could not be done in this study due to contamination, for future studies it would be interesting to also detect modified bases in *A. baumannnii* after treatment with FS-1, and compare it to our results obtained for *S. aureus* and *E. coli*.

Treating resistant bacteria in response to their resistance phenotype at a given time is no longer a sufficient or sustainable option. Understanding the evolutionary potential of bacterial infections will help the development of strategies which not only treat resistance, but also anticipate and prevent further development of multidrug resistance. Bacteria keep adapting to new treatments, we must therefore also evolve our strategies to combat resistance.

References

ABRAHAM, E. P., CHAIN, E., FLETCHER, C. M., FLOREY, H. W., GARDNER, A. D., HEATLEY, N. G. & JENNINGS, M. A. 1992. Further observations on penicillin. 1941. *Eur J Clin Pharmacol,* 42**,** 3-9.

ABRAMS, J. S. 2001. Adjuvant therapy for breast cancer — results from the USA consensus conference. *Breast Cancer,* 8**,** 298-304.

BARBOSA, T. M. & LEVY, S. B. 2000. The impact of antibiotic use on resistance development and persistence. *Drug Resist Updat,* 3**,** 303-311.

CLARK, T. A., MURRAY, I. A., MORGAN, R. D., KISLYUK, A. O., SPITTLE, K. E., BOITANO, M., FOMENKOV, A., ROBERTS, R. J. & KORLACH, J. 2012. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. *Nucleic Acids Res,* 40**,** e29.

FOTI, J. J., DEVADOSS, B., WINKLER, J. A., COLLINS, J. J. & WALKER, G. C. 2012. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science,* 336**,** 315-9.

KOREN, S., HARHAY, G. P., SMITH, T. P., BONO, J. L., HARHAY, D. M., MCVEY, S. D., RADUNE, D., BERGMAN, N. H. & PHILLIPPY, A. M. 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol,* 14**,** R101.

LAXMINARAYAN, R. 2014. Antibiotic effectiveness: balancing conservation against innovation. *Science,* 345**,** 1299-301.

NEELEY, W. L. & ESSIGMANN, J. M. 2006. Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem Res Toxicol,* 19**,** 491-505.