FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Trevor Beaudoin
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Microbiology and Immunology)
GRADE / DEGREE

Department of Microbiology and Immunology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Investigating the Role of *ndvB* in *Pseudomonas Aeruginosa* Biofilms
TITRE DE LA THÈSE / TITLE OF THESIS

Thien-Fah Mah
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Craig Lee

Sandra Ramirez

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Investigating the Role of *ndvB* in *Pseudomonas aeruginosa* Biofilms

By Trevor Wayne Beaudoin

Submitted to the Faculty of Graduate and Postdoctoral studies in partial fulfillment of the degree of Masters of Science in Microbiology and Immunology

University of Ottawa
Faculty of medicine
Department of Microbiology and Immunology

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-61220-0
Our file Notre référence
ISBN: 978-0-494-61220-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

TP

Table of Contents

List of Figures

List of Tables

Abstract

Pseudomonas aeruginosa is an opportunistic pathogen prevalent in nosocomial infections and patients with cystic fibrosis. *P. aeruginosa* shows a high degree of antibiotic tolerance which in part can be attributed to the formation of biofilms. The increased antibiotic resistance seen in biofilms can be attributed to several factors including differential gene expression within biofilms that can lead to biofilm specific mechanisims of antibiotic resistance.

The *ndvB* gene is important for biofilm specific antibiotic resistance in *P. aeruginosa*. It is important in signaling and regulation of gene expression in other pathogens. Microarray analysis comparing gene expression between wildtype and a *ndvB* deletion mutant was performed to identify genes that might be regulated by the gene product, believed to be cyclic glucans, that contribute to biofilm specific antibiotic resistance. The array analysis identified 24 genes that were differentially regulated by *ndvB*, including a response regulator, *agmR*, as well as most of the genes which it regulates. Quantitative real-time polymerase chain reactions using primers specific to *agmR* and its associated genes confirmed that they were expressed in a *ndvB* related manner. Minimal bactericidal concentration assays were performed and confirmed that these genes are important in resistance to tobramycin and ciprofloxacin.

Acknowledgements

I would like to take this opportunity to thank a number of people who helped me throughout my Masters research. Firstly, I'd like to thank Dr. Thein-Fah Mah for her guidance, encouragement and for giving me this opportunity. I would also like to thank all of the lab members with whom I have had the opportunity to work especially Li Ahng, Jenilee Gobin, Ryan Landerville and Koli Basu. I would also like to thank the mebers of my thesis advisory committee, Dr. Ken Dimock and Dr. Susan Logan for there helpful advice throughout my Masters. I would like to thank Jason Tetro and Artine Keklikian for reviewing parts of my thesis.

I would also like to thank everyone that made this such an enjoyable experience and provided support, scientific and otherwise, throughout the process. This would include Dr. Brown, Alison, Jae, Jihui, Liya, Mellisa, Nicole, Samar, Samra, Sine, Shuai and Sujeeve. You, along with many others throughout the department have made this a great life experience and have shaped my attitudes towards science, so I thank you for that.

Lastly, but just as important, I'd like to thank my family for their support throughout . Thank you everbody, so vey much!

List of Abbreviations

°	Degrees
°C	Degrees Celsius
µg	Micrograms
µL	Microliters
µM	Micromolar
23S	Part of Large Ribosomal subunit
16S	Part of Small Ribosomal Subunit
50S	Large Ribosomal Subunit (23S and 5S)
30S	Small ribosomal subunit (16S)
aRNA	Array RNA
ABC	ATP Binding Cassette
Arg	Arginine
Bp	Basepair
cDNA	Complementary DNA
Cip	Ciprofloxacin
CF	Cystic Fibrosis
CLSI	Clinical Laboratory Standards Institute
C_T	Critical Threshhold
DNA	Deoxyribonucleic acid
EPS	Exopolymeric substrate
IQR	Interquartile Region
IVT	In vitro transcribed
KO	Knockout
L	Liter
LB	Luria-Bertani
LPS	Lipopolysaccharides
M	Mismatch
M63	Minimal Media Broth
MAS 5.0	Microarray Suite 5.0
MBC	Minimal Bactericidal Concentration
mg	Milligram
mm	Millimeter

mL	Milliliter
NAD+	Oxidized Nicotinamide Adenine Dinucleotide
NUSE	Normalized Unscaled Standard Error
OD_{600}	Optical Density 600
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PM	Perfect Match
PLM	Probe level model
PQQ	Pyrolequinolonequinone
q-PCR	Quantitative Realtime PCR
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
RMA	Robust Multi-chip Average
RLE	Relative Log expression
ssDNA	Single stranded DNA
Tb	Tobramycin
Tn	Transposon insertion
Wt	Wildtype

Introduction

Introduction

Antibiotic resistance has become a major problem when dealing with various bacterial infections. The emergence of highly resistant and multidrug resistant bacteria is of great importance and concern for infected and hospitalized patients. Bacteria have several ways to become resistant to antibiotics including the expression of active transporters, usually efflux pumps, changing the target of the drug or inactivating the drug. In addition to these mechanisms, recent evidence has suggested that bacteria can grow in surface attached communities called biofilms. These diverse communities offer an additional mode of protection for the bacteria and offer intrinsic biofilm specific gene expression that can increase resistance to antibiotics.

1.0.0: Biofilms

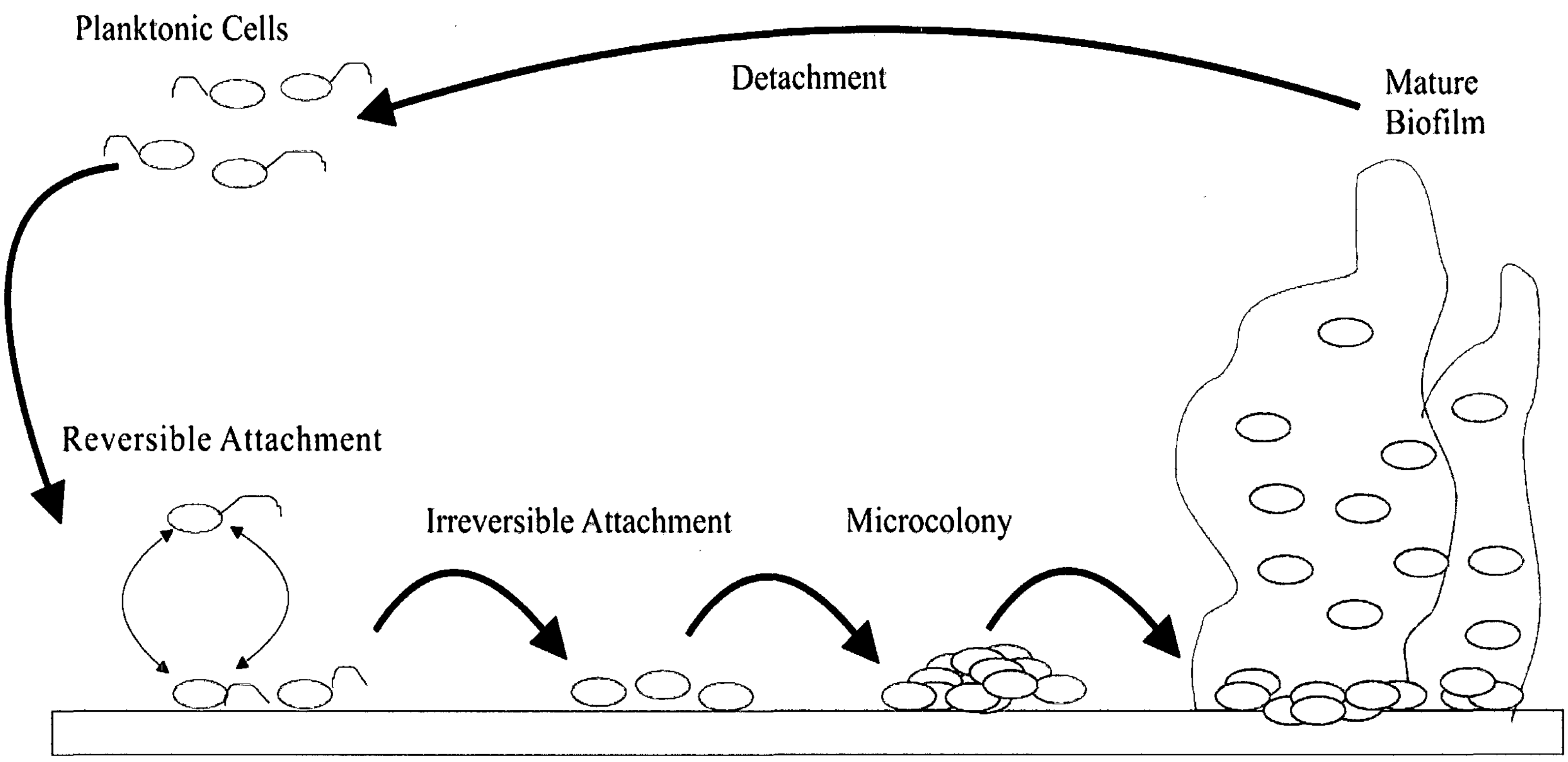
1.0.1: Introduction to Biofilms

Bacteria can grow in two general forms; either as free swimming planktonic bacteria or as surface-attached communities known as biofilms (19). Bacterial biofilms form in almost any environment containing a surface sufficient in nutrients and water. Biofilms generally consist of aggregates of bacteria that adhere to surfaces forming encased 'communities' comprised of a matrix of extracellular polymers that confers protection to the bacteria within the biofilm (5).

Several stages are involved in the production of biofilms and Figure 1 gives a brief overview of these stages. In general, a suitable surface (which can be biotic or abiotic) is preconditioned with nutrients or a carbon source that can be utilized by a bacterial species, generally by natural mechanisms such as

Figure 1: The process of formation of a biofilm community.

Biofilm formation begins with the sampling of the substrate (surface) by planktonic bacteria. The bacteria undergo reversible attachment and a change in gene expression results in irreversible attachment. The accumulation of cells at the surface results in microcolony formation and quorum sensing leads to further gene expression changes. Production of an exopolymeric substance results in the formation of mature biofilms. Detachment through dispersal or natural means occurs.


Planktonic Cells
Detachment
Mature
Biofilm
Reversible Attachment
Irreversible Attachment
Microcolony

sedimentation (5). Once the surface is conditioned, bacteria can sample the surface or be drawn to the surface nutrients via chemotaxis systems, and can interact and adhere to the surface, initially reversible through weak interactions such as van der Waals forces or hydrogen bonding (5). The flagellum of *Pseudomonas aeruginosa* plays an important role not only in bringing the bacteria to the surface, but also in forming interaction important in reversible attachment and early biofilm formation in this bacteria (8, 69)Eventually, the attachment becomes irreversible as an accumulation of weak interactions occurs and a change in gene expression begins. In the case of *P. aeruginosa* early events in biofilm formation that lead to irreversible attachment include the upregulation of type IV pili and production of glycocalyx, primarily composed of polysaccharides (34, 41). Once irreversible attachment occurs, several bacteria can aggregate at the surface. This results in the formation of microcolonies. The bacteria continue to aggregate and macrocolonies form on a lawn of bacteria at the surface.

The bacteria are firmly encased in the exopolymeric substance (EPS), also referred to as the biofilm matrix, which is composed of polysaccharides, proteins, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) (5). The polysaccharide composition of the biofilm matrix varies amongst bacteria involved and comprises the largest portion of the matrix. In *P. aeruginosa*, three types of polysaccharides have been determined: Psl, Pel and alginate. Psl (for polysaccharide synthesis locus) is deemed important for biofilm adherence to surfaces (54). The *pel* (pellicle formation) genes encode for a glucose rich polysaccharide that is important in early and late biofilm development (5, 92).

Alginate has been shown to be important in biofilm structure and function and is the best characterized of the polysaccharides in biofilms (5, 40). The composition of polysaccharides in biofilms is determined by the environment and nutrient availability (92) and is a major component that leads to the transition from aggregates of microcolonies to the development of macrocolonies and mature biofilm communities. In conjunction with this, the composition of the EPS and in particular the polysaccharides lead to many of the intrinsic properties important to biofilms. Many of the properties associated with biofilm growth are due to the EPS that acts as a barrier against the environment while providing structure and the basic architecture of the microcolonies and macrocolonies. The mature biofilm consists of clusters of these macrocolonies and microcolonies on a surface. From here, single cells or portions of biofilm can be dispersed or sloughed off respectively to colonize other areas.

The dynamics and structure of the biofilm is often determined by the physical properties surrounding the colonies such as flow rate, nutrient availability and regional oxygen or pH conditions (75). It is also important to note that these biofilms are not solid masses containing solely bacteria but rather 10-20% of the mass is bacteria while the rest is exopolymeric matrix as well as nutrients and water, which comprises the majority of the biofilm (approximately 70-80% of the mass) (5, 41). In order to provide nutrients throughout the biofilm, nutrient channels are located in between the macrocolonies in mature biofilms to allow the flow of nutrients and oxygen to different parts of the biofilm.

The composition of biofilms, whether single species or mixed species, is altogether a heterogeneous mix, both spatially and temporally. In other words, genes expressed by individual bacteria within the biofilm are different over time and depending on where they are located within the biofilm as the conditions faced by a bacterium in the center of the biofilm can be vastly different then the conditions faced by bacteria on the edge of the biofilm. The heterogeneity of biofilms leads to distinct microenvironments of growth that can have vastly different physiochemical properties, which in turn can result in distinct interactions and pockets of phenotypic and genotypic distinct populations (49). Such differences can have a major impact on the biofilm and can result in changes that are relevant to industry and medical community, which will be discussed later.

1.0.2: Experimental growth of Biofilms

Several methods for studying biofilms have been developed in order to view biofilms and elucidate the properties of these bacterial communities (5, 75). Although it is unnecessary to describe all of the methods herein, it is important to get a basic understanding of several systems and why they are used. The standard system for viewing bacterial biofilm formation is through flow cells (88). The advantage of flow cells is that they allow direct visualization through microscopy, provide a method of growing biofilms with shear force across them and provide a method that allows for biofilm growth under conditions that are easy to modify. Although this is a standard way of growing biofilms, flow cells do require a more specialized setup and do not always provide enough biofilm biomass for some downstream applications.

Biofilms can also be grown and analyzed by other methods that require less shear forces on the biofilm. One such method is the Kadouri drip-fed reactor. In this setup, fresh media is constantly pumped into the wells of a 6-well microtitre plate while waste media is removed (60). This allows for the biofilms to be grown over several days in order to accumulate a large amount of biomass. Biofilms grown in this manner can then be visualized by microscopy and provide sufficient cell mass to allow for downstream applications such as RNA extraction. Figure 2 shows a diagram of the Kadouri drip-fed reactor, which was used for most of the biofilm growth in subsequent experiments for this thesis.

Another method for biofilm growth which is rapid and provides superior images for biofilm formation is the air-liquid interface assay. In this setup, a small amount of appropriate growth media is inoculated with bacteria in a six-well plate and tilted to allow an air-liquid boundary to be formed midway up the plate. After enough time has passed for biofilm to be grown, the media can be removed and images can be taken of the biofilm (60).

Biofilm formation can also be visualized via the microtitre plate biofilm assay (60). This assay uses a standard 96-well microtitre plate that has been inoculated with a dilution of a bacteria of interest. After a suitable amount of time (as determined by experimentation) the bacteria and media is removed from the plate and replaced with an equal volume of 0.1% crystal violet solution. This solution stains bacteria purple. If the bacteria formed a biofilm on the microtitre plate, a purple ring will appear on the plate after the crystal violet solution is removed from the plate. This is a rapid method to identify if the bacteria can form biofilms under the conditions used in the experiment.

The growth of biofilms in a reproducible manner for visualization and experimental setups is not a simple task due to the inherent heterogeneity of biofilms. A number of consistent experimental procedures are required in order to compare one experiment to the next. These, along with many other techniques, have been employed by researchers to identify the unique features of biofilms.

1.0.3: Intrinsic Biofilm features and Antibiotic Resistance

The physiochemical properties of biofilms allow them to persist in a variety of different environmental conditions and survive different types of stress (17). Furthermore, many of the intrinsic properties of biofilms are a result of the physiochemical relationships provided by the densely packed bacteria as well as the phenotypic heterogeneity of genetic expression of cells within the biofilm (75). For instance, much of the persistence of biofilms on industrial and medical related devices is a result of a microbial reservoir of cells that make it hard to remove the biofilm (75). Mature biofilms cannot be washed away or easily removed from surfaces and, in addition, portions of the biofilm can detach and colonize surfaces far away from the initial attachment site (5).

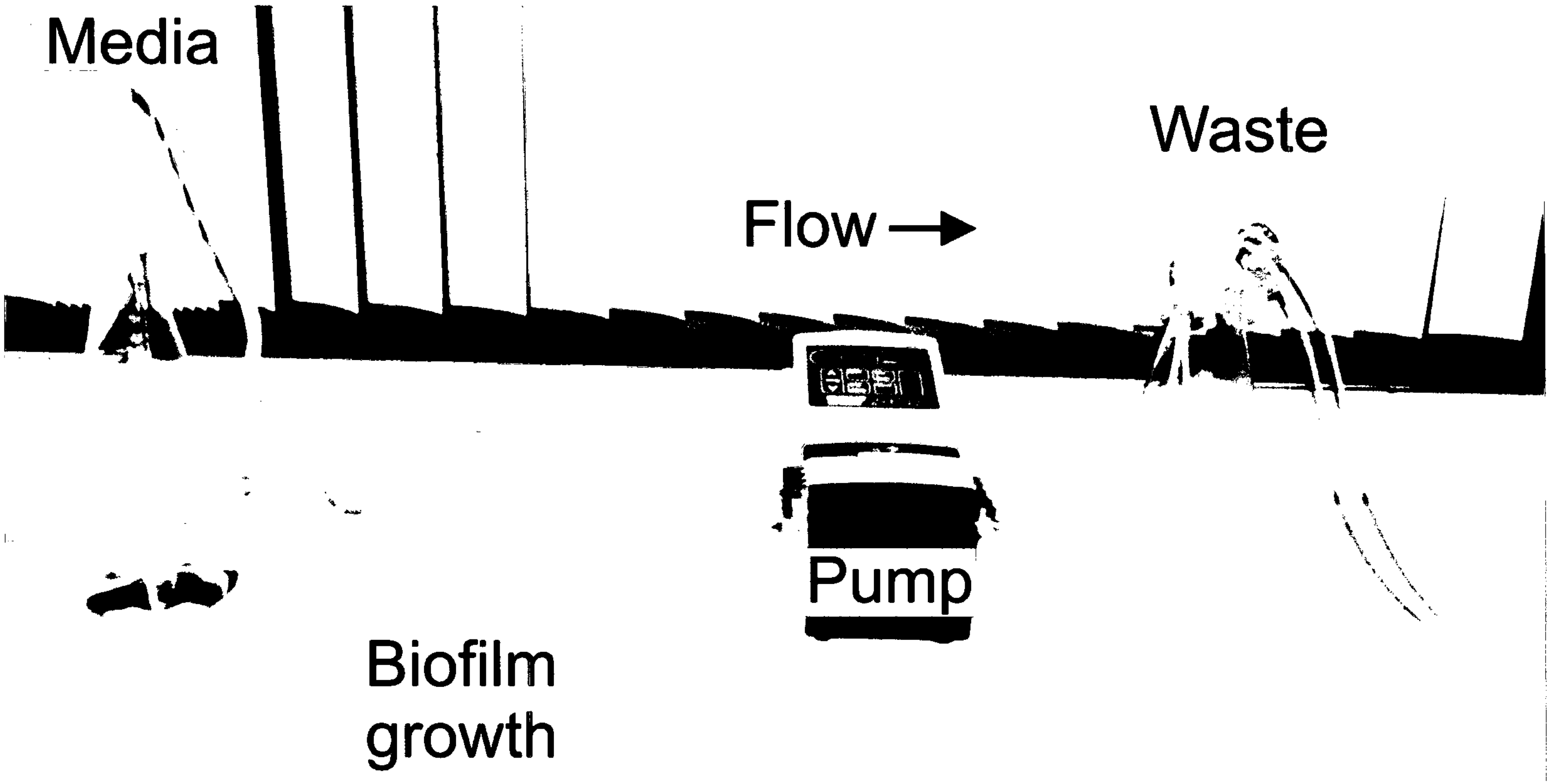
One of the hallmarks of biofilm is protection against antibiotics and immune cells of the host. Bacterial biofilms can be 10 to 1000 times more resistant to antibiotics than their planktonic counterparts (56). Antibiotic resistance is not due to mutation in bacteria grown in biofilms, since upon dispersal from biofilm, the bacteria return to planktonic levels of resistance (7, 33, 56, 70). Although there may be a role for general planktonic mechanisms, such as antibiotic inactivation, exclusion of the antibiotic or modification of targets of the antibiotic (70) in biofilm resistance, increased resistance is not be attributed to these mechanisms alone (33, 70).

The switch to a biofilm mode of growth renders many of the bacteria more resistant to antibiotics and so the formation of 'biofilm' antibiotic resistance is at least in part an intrinsic result of growing in a biofilm (6). One good example of this is the effect that the biofilm matrix has on increasing resistance to antibiotics. Although the matrix of the biofilm is mostly water and does not provide a barrier to diffusion, molecules within the matrix can prevent antibiotics from reaching their site of action (23). Many of the polysaccharides and proteins have negative charges which can interact with positively charged antibiotics, such as aminoglycosides, thus making penetration through the biofilm matrix slower (64).

The expression of genes in biofilms, but not in planktonic bacteria, can lead to increased resistance in biofilm-associated bacteria. One such example is the production of cyclic glucans in *P. aeruginosa* via *ndvB* that act as antibiotic traps in the periplasm of biofilm grown cells but are not present in planktonic counterparts (57). The expression of efflux systems preferentially expressed in biofilms has also recently been observed (102). The expression of these type of genes produce a biofilm specific phenotype of antibiotic resistance within the

Figure 2: Experimental biofilm growth

The Kadouri drip-fed method of growing biofilm. The flask on the left is the fresh 1XM63 media reservoir. *P. aeruginosa* biofilms grow in the 6-well plate next to the media reservoir. The pump removes media from the 6-well plate where biofilms grow, and replenishes it with fresh media. The removed media is discarded in the flask on the right which is the waste container.


Media
Flow →
Waste
Pump
Biofilm
growth

bacterial population that is not induced by mutation but rather through differential gene expression within the biofilm. Thus, the expression of biofilm specific gene that can lead to an increased resistance is an exciting prospect of study.

In addition to biofilm specific resistance due to altered gene expression in biofilms, resistance to antibiotics is also greatly affected by the physiological environment encountered by the bacteria in a biofilm. Reduced oxygen rate or nutritional availability results in slow growth of bacteria in some parts of the biofilm. In general, bacteria in biofilms grow at slower rates than bacteria growing in planktonic conditions (29). Antibiotics are typically taken up by and affect fast-growing bacteria, while slow-growing bacteria are protected (15, 29).

The persistence of cells in a bacterial population after a stress is introduced has obvious implications in antibiotic resistance. A role for the general stress response has also been implicated in the formation of biofilms; the *P. aeruginosa* mutant *rpoS* (a regulator of stress response) produces biofilms that have less of a thickness compared to the wildtype form (32). RpoS is a vital control for bacteria against different stress conditions, and the fact that it is important in biofilm formation warrants further study.

Bacterial cells growing in biofilms exhibit differential gene expression that leads to the formation of persister cells within the population, leading to multi-drug tolerance in the biofilm population (49). How tolerance is achieved through persister populations in a biofilm is not fully understood, but it has been hypothesized that a population of cells in the biofilm show reduced gene expression such that the targets of many antibiotics are not expressed and thus protection is granted to this subset population (49). Additionally, a subset of cells within the biofilm undergo 'no growth', and thus the major mode of action of antibiotics does not affect these cells: the bacteria gain tolerance at the expense of cellular proliferation (49)

These mechanisms, as well as several others not discussed here, lead to the observed increase in antibiotic resistance seen in bacterial biofilms. The culmination of intrinsic properties of biofilms makes them difficult to eradicate in any setting, both industrial and medical. The following section deals with the prevalence of biofilms in disease.

1.1.0: *Pseudomonas aeruginosa*

P. aeruginosa is a gram negative, motile unipolar flagellated bacterium that is a member of the *Pseudomonadaceae* family (21, 42). *P. aeruginosa* is ubiquitous in nature, commonly found in soil and water systems and can live in a diverse range of environments. It has the ability to infect a wide range of species including plants, insects and vertebrates (74, 87). *P. aeruginosa* can grow on a variety of carbon sources, utilizing different carbon metabolism pathways depending on the level of oxygen in the environment; although *Pseudomonads* are typically thought to be aerobic bacteria, *P. aeruginosa* can grow in low oxygen conditions or under complete oxygen depletion by utilizing nitrate as a terminal oxygen acceptor (31, 44).

The genetic versatility of these bacteria makes it an important bacterium in industry as well as in medical infections (3, 5, 59, 75, 85, 101). The versatility exhibited by *P. aeruginosa* is arises from the large number of genes that can be expressed by the bacteria in different conditions. The complete genome sequence was released in August of 2000 where it was described as 6.3 million basepair circular genome that houses approximately 5,570 open reading frames making it significantly more complex than other bacterial genomes sequenced to date (90). Having the complete genome available makes *P. aeruginosa* an ideal system to study. It has also been well studied for its ability to form biofilms and produce complex bacterial communities.

P. aeruginosa is of particular interest in medical settings because of its ability to act as an opportunistic pathogen in humans causing urinary tract infections, ear infections, ventilator associate pneumonia and infections in patients with cystic fibrosis (CF) (22, 63, 94).

1.1.1: PA14

P. aeruginosa PA14, a clinical variant of the common PA01 strain, is also well characterized and used for studying important virulence factors in *P. aeruginosa*. PA14 was isolated from a burn patient and has since been used to study *P. aeruginosa* that resembles more closely a pathogenic strain (18). PA14 has approximately 300 more genes than PA01 with most contributing to increased virulence, and is well studied for its ability to infect a wide range of hosts (18). In order to understand the ability for this bacteria to act as a human pathogen and cause persistent infection, it is important to understand the basic modes of growth for *P. aeruginosa.*

1.1.2: Clinical relevance of *Pseudomonas aeruginosa* in disease:

P. aeruginosa has been deemed an important pathogen in nosocomial infections including burn wounds, surgical implant devices, ventilator associated pneumonias and urinary tract infections via catheter contamination. Conservative estimates link *P. aeruginosa* to 10% of all hospital acquired infections and even a greater problem in intensive care units (2, 30). The presence of *P. aeruginosa* infection can led to extended hospitalization periods, increase associated cost and mortality amongst patients (30). In addition to these infections, *P. aeruginosa* is a major problem for people suffering form the genetic disease cystic fibrosis.

1.1.3: Cystic fibrosis and *Pseudomonas aeruginosa*

Cystic fibrosis (CF) is the most common genetic disorder among young Canadians and primarily affects the lungs and breathing of patients (1). CF results from a mutation in the CF transmembrane conductance regulator (CFTR), causing defective functioning and consequently reduced mucociliary clearance (MCC) in epithelial tissues (38, 91). The CFTR protein creates a chloride ion channel in the outer membrane of epithelial cells. A defective CFTR allows for the accumulation of chloride ions in lung epithelial cells. This reduced solute transport leads to a buildup of abnormal airway surface liquid that tends to be high in salt and low in water/oxygen content resulting in a thick mucosal layer in the lungs (38, 91). Since MCC is one of the most important innate mechanisms for clearing bacteria from the lung and preventing infection, its reduced function in CF patients makes them susceptible to various opportunistic infections (46). *P.*

aeruginosa infection in CF patients leads to further reduction of lung function and results in increased mortality in these patients (47).

The conditions present in the CF lung, mainly a thick mucus high in salt content but low in water content resulting in anaerobic like conditions, are suitable for *P. aeruginosa* growth but preclude many other pathogens (68). In addition, chronic infections often occur increasing the presence of inflammatory cells in the lung which further exacerbates the situation for patients (47). The presence of chronic *P. aeruginosa* and induced inflammation leads to the decreased pulmonary function in CF patients.

Current treatment consisting of aminoglycosides antibiotics, such as tobramycin, and fluoroquinolones, such as ciprofloxacin, serves to combat the infection. These drugs are also used as a prophylaxis in an attempt to prevent initial infection (25). However, even with current antibiotic regimens, *P. aeruginosa* infections are notoriously difficult to treat once established in these patients. Several studies indicate that *P. aeruginosa* actually forms biofilm in lung epithelial resulting in the reduced ability to clear or treat infections (12, 68, 86, 95).

1.1.4: Mode of Antibiotic Action

The current method of treatment for patients with CF is the administration of a single, or combination of, bactericidal drugs (25). Common sites of action for bactericidal drugs are the inhibition of DNA replication/repair, inhibition of protein synthesis, and inhibiting cell wall turnover (96). Tobramycin, an aminoglycoside, belongs to the aminocyclitol group of 30S ribosome inhibitors which induce protein mistranslation (20). Gentamicin is another aminoglycoside that generally shows a reduced ability to kill *P. aeruginosa* biofilms when studied (57, 67). Ciprofloxacin is a fluoroquinolone which binds to DNA gyrase and inhibits DNA replication and repair (27).

1.2.0: Biofilms in Disease

Biofilms offer a mode of growth for bacteria that allow them to survive and indeed thrive in a host (24, 39). Recent figures released from the Center for Disease Control (CDC) in the United States attribute biofilms to 65% of all bacterial infections (39, 89). Because of the properties of biofilms in general, and due to the genetic diversity of cells within a biofilm, treatment of biofilm infections has become a major obstacle in hospitals. In general, biofilms have been associated with a slew of human disease including otitis media (78), periodontal disease (58) and as previously described, in patients with CF. In addition to this, biofilms can be found attached to medical devices such as urinary catheters, contact lens cases and prosthetic heart valves (23).

Because of the difficulty in treating bacterial biofilms, quite often these infections become chronic and recurring, or in the case of medical devices, require the removal of devices leading to further stress on the patients (23, 56). The properties of biofilms provide several advantages that lead to increased infection and difficulty in treatment. The ability to detach and colonize a larger area increases the amount of area that can be infected by these biofilms (23). The biofilm also provides protection against host immune cells and their associated reactive oxygen species while providing a niche for bacteria in the host (23).

Furthermore, the increased resistance to antibiotics and ability to have subset populations that can persist and survive treatment make these biofilms difficult to treat and clear (6). Therefore it is very important to study mechanisms that allow biofilms to resist the action of antibiotics.

1.2.1: Search for biofilm specific mechanisms of antibiotic resistance

In order to identify genes that could be important to biofilm cells in resisting antibiotics, Mah et al. screened a library of approximately 10,000 mutants in *P. aeruginosa* to identify genes that were important to biofilm-specific antibiotic resistance. In this screen, biofilms formed by the mutants were exposed to tobramycin to see which mutants would have an increased sensitivity to the drug (57). Mutants of interest would then have an increased sensitivity to antibiotics when grown in biofilms compared to the parental strain PA14, but show no difference when grown planktonically. Additionally, these mutants showed no hindrance in biofilm growth or architecture (57).

This study identified seven genes that were found to fit within these parameters. One of the genes that was identified was *ndvB*. Wildtype bacterial cells showed a 50X increase in tobramycin resistance between planktonic and biofilm states of growth, while the *ndvB* mutant only showed a 6.25X fold increase in resistance (57). It was clear that this mutant was more susceptible to antibiotics, but this increased susceptibility was not due to any structural defects with in the biofilm. Thus, this study presented a biofilm-specific phenotype due to a specific gene that was preferentially expressed in biofilms, leading to a biofilm-specific mechanism of antibiotic resistance.

1.3.0: NdvB and cyclic glucans

Plant pathogens of the family *Rhizobiaceae* have been shown to produce a variety of cyclic glucans that can be found in the periplasmic space, extracellularly, as well as attached to the surface of the cell (14). While the majority of studies focus on *Rhizobium meliloti* and *Agrobacterium tumefaciens* production of β-(1→2)-D-glucans, *Bradyrhizobium japonicum* produces β-(1→3),(1→6)-glucans (10). *P. aeruginosa* has the gene PA1163 which has been identified as 58% similar in sequence homology to the *ndvB* gene in *B. japonicum* (57). The *ndvB* gene encodes a glucosyl transferase that is responsible for the cyclization of glucose into branched cyclic glucans typically found in the periplasm and extracellular space of *B. japonicum* (14).

Cyclic glucans play a role in hyperosmotic regulation and adaptation in these bacteria. In *P. aeruginosa,* a model of antibiotic trapping was proposed (57), where cyclic glucans isolated from the periplasm were able to sequester tobramycin away from its site of action in the cytoplasm thus increasing resistance to the antibiotic (57). It was also shown that *ndvB* is preferentially expressed in biofilms, which makes for an exciting prospect such that a specific gene expressed in a biofilm can be responsible for the increased antibiotic resistance phenotype (57). The increased resistance present in biofilms was not completely abolished with mutation in this gene however, and so a multifactorial mechanism might be at play in the overall resistance to antibiotics.

Several studies have implicated the cyclic glucans of *B. japonicum* as being important signaling molecules that can affect the gene expression of its soybean host (10, 11, 61). In order to identify genes that might be under the control of cyclic glucans in *P. aeruginosa* I carried out microarrays to compare wildtype PA14 with a *ndvB* knock-out strain of PA14 to identify genes differentially expressed under these conditions. These experiments, and the results obtained from theme, will be discussed in chapters 2 and 3.

1.4.0: Searching for targets using Affymetrix Genechip microarray

Genes differentially expressed from one condition to another can be identified using a genome-wide approach such as microarray analysis comparing the different conditions. This method uses mRNA expression of genes under different conditions to produce a signal which can be scanned and analyzed for a relative fold expression. Affymetrix has a Genechip available for *P. aeruginosa* PA01 that is suitable for the present study. The PA01 Genechip includes 5,549 protein coding sequences as well as 179 protein coding regions prevalent in *P. aeruginosa* strains other then PA01(65). The Genechip is a 3'-*in vitro* transcribed (IVT) chip that relies on first strand cDNA synthesis from the RNA samples, creating cDNA with a T7 promoter. This is followed by second strand cDNA synthesis which is then labelled through IVT with biotin tagged uridines to produce arrayRNA (aRNA). This aRNA is then hybridized to the Genechip, which then undergoes scanning.

The Affymetrix Genechip technology uses a perfect match (PM)/mismatch (MM) signal score to generate a list of genes present, absent or marginal in each of the conditions. Each coding sequences in *P. aeruginosa* is broken down into 13 probe sets of 25 basepair oligomers, representing the PM sequences for the chip. Each MM sequence is identical to the PM sequence with a single nucleotide change in the middle of the probe set sequence. This accounts for non-specific hybridizations to the sequence. A gene is deemed present in a given condition if the signal for the PM minus the signal from the MM is greater then a threshold signal determined from internal controls on the chip. After the chips are scanned the data undergo statistical analysis to determine a list of differentially expressed genes (26).

1.4.1 Statistical analysis

The statistical analysis procedures of any microarray experiment are of great importance when identifying genes that are differentially regulated. The scanning of the microarray produces an unprocessed chip, and ensuring that the data from these unprocessed chips is of good quality is important before proceeding with analysis. Microarray analysis of variance (MAANOVA) of the various biological replicates/arrays is one way to tell if the data are of good quality (13, 37, 73). Good quality data will show little variance between like conditions. In addition to the statistical analysis of the preprocessed data, visualizing the chip for anomalies of plots of the intensities can be done to ensure that the chips are of good quality (37).

Once the preprocessed data are checked for quality, one must choose the proper method for normalizing the data in order to generate a list of significantly differentially expressed genes. Many different methods are available to do this including Microarray suite 5.0 (MAS 5.0) algorithms or Robust Multi-chip Averaging (RMA,) the drawbacks, limitations and advantages of which are discussed in many publications (13, 43, 53) and quiet beyond the scope of this introduction. Most methods rely on quantile normalization, making all of the probe level data have a distribution of one. From here genes are either upregulated or downregulated. Fitting several methods to your dataset and taking the overlap of the identified genes provides a greater confidence that the genes identified from the procedure are the actual genes that are differentially regulated between the conditions. Any method used for normalization requires that your chip data be of comparable quality in order to ensure high degrees of confidence.

Once the data have been normalized, statistical test can be applied to the data. Typical analysis, such as t-test or ANOVA, must be adjusted to ensure that genes defined as statistically significant actually are. A moderated t-test is suitable for establishing a p-value of statistically significant genes at a cut off of greater than 2 fold change within the array. The t-test can be paired or unpaired with equal or unequal variance. Although most arrays have equal variance and normal distributions after normalization, intensities at low values can have high variance and spots with high expression can be overlooked. The p-value must be modified however in order to generate a true list of statistically significant genes, because a large population of genes is being studied under a small sample size (generally 3 replicates for each condition). Holms step-wise correction is often a good choice, allowing for a p-value that is provides a high interval of confidence for the statistical analysis used in microarray studies, as well as Bayesian statistics (13, 26, 53).

1.5.0: Rationale for current study

The nature of this study revolves around the fact that certain genes that are expressed preferentially by bacteria in biofilms can lead to the increased resistance that is seen in bacterial biofilm communities. This novel idea was first demonstrated by Mah et al. when they showed the *ndvB* is more highly expressed in biofilms and a deletion of this gene reduces the biofilm specific resistance (57). The current study was undertaken in order to understand how cyclic glucans could affect the gene expression in *P. aeruginosa* biofilms and to determine if the genes regulated by these glucans could play a role in the multifactorial mechanism of biofilm resistance to antibiotics.

I have proposed that the expression of *ndvB* in biofilms and the presence of cyclic glucans produced by the NdvB protein lead to differential gene expression within the biofilms that can in turn lead to increased resistance to antibiotics. Since it has been observed that the cyclic glucans are important in signaling in other organisms and since *ndvB* has been shown to be crucial in the production of these molecules, I carried out an affymetrix PA01 microarray analysis of a PA14 *ndvB* knockout strain and compared the expression to the PA14 wildtype. This design allowed for 5500 *P. aeruginosa* genes present on the affymetrix array to be scanned for differential expression between the different conditions. PA14 was chosen as an experimental model because the knockout strain was available (57) and the previous studies done in the lab revolved around this more clinically relevant strain. Although this would result in a less complete list of genes due to the fact that some genes present in the PA14 genome would not be on the array and some genes that are actually present might be deemed marginal do to minor sequence variation between PA14 and PA01, I deemed this a necessary and ultimately minor limitation. I expected to generate a list of genes that would be differentially expressed between the two conditions and confirm that they were differentially expressed through the use of quantitative real time polymerase chain reaction (q-PCR) and determine the phenotype of antibiotic sensitivity.

Hypothesis and Objectives

Hypothesis

The production of *ndvB* derived cyclic glucans will alter the gene expression in *Pseudomonas aeruginosa* biofilms and the genes that are regulated will have an impact on biofilm specific antibiotic resistance. NdvB is an important protein in the production of cyclic glucans and abolishing it through a gene deletion has been shown to reduce the amount of glucans in *P. aeruginosa* (57). Using a *ndvB* gene deletion strain of *P. aeruginosa* and comparing it to gene expression in a wildtype strain of *P. aeruginosa* one can identify if any genes are differentially expressed between the two conditions using microarrays. By studying mutants of the genes identified in the microarray analysis and using antibiotic sensitivity assays, it can be shown that cyclic glucans play an important role in expression of genes important for biofilm specific antibiotic resistance.

Objectives

1) Confirm *ndvB* gene deletion via polymerase chain reactions (PCR), antibiotic sensitivity of mutant via minimal bactericidal concentration (MBC) assay and preferential gene expression of *ndvB* in biofilms by quantitative real time polymerase chain reactions (q-PCR).
2) Create a protocol for growing biofilms, harvesting and purifying RNA for microarray analysis.
3) Perform microarray of wildtype and *ndvB* knockout biofilms. Analyze data to generate a list of differentially expressed genes
4) Perform MBCs on mutants of genes from the above list to determine the antibiotic phenotype, perform q-PCR to confirm gene expression in biofilms vs. planktonic state.

Materials and Methods

Materials and Methods

2.1.0: Primers:

Primers were required for a number of applications including polymerase chain reactions (PCR) and quantitative real-time polymerase chain reactions (q-PCR). Table 1 has a complete list of primers. All primers were ordered from Invitrogen (Carlsbad, CA.) and were 50 mM desalted primers. The construction of the primers was done using Primer 3, a freely available program accessible on the internet (76). PCR primers bind to sequences found in the 500 basepairs upstream (L) and 500 basepairs downstream (R) of the gene of interest in the PA14 genome and were chosen based on similar melting temperatures of the pair, GC content and minimal interactions. q-PCR primers were chosen within the gene of interest and selected using the same criteria as well as creating a product of between 100-150 basepairs, which is optimal for q-PCR.

2.2.0: Bacterial Strains:

The *P.aeruginosa* strain used during this study was PA14 (74), herein referred to as the wildtype (Wt) strain. The PA14 *ndvB* in-frame deletion was constructed by T.F. Mah (57), herein referred to as the *ndvB* knockout (KO) strain. Both of these strains were used for Affymetrix microarray, quantitative polymerase chain reaction (q-PCR) experiments and minimum bactericidal concentration (MBC) assays.

2.3.0: Confirming *ndvB* Knockout strain

2.3.1: Chromosomal DNA extraction

Chromosomal DNA was extracted using phenol-chloroform extraction. Bacteria were grown up overnight in 5 mL of rich media [Luria Bertani (LB) broth] from Difco (Sparks, MD.) at 37 °C. One mL of suspended bacteria was collected in a 1.5 mL centrifuge tube (Diamed, Mississauga, ON.) and spun down at 12,000 x g for 1 minute at room temperature. The supernatant was removed and the pellet collected and resuspended in a 0.5 ml solution of 150 mM NaCl, 50mM Tris-HCl, 10mM EDTA . Cells were lysed with 1% weight/volume SDS and contaminated RNA was degraded with the addition of 50 µg/ml RNAseA (Invitrogen). The sample was incubated for 10 minutes at 65°C. 550 µL of phenol (Sigma, St. Louis MO.) was added and the sample was centrifuged at 12000xg for 10 minutes at room temperature. The aqueous phase was collected in a fresh microcentrufuge tube. Phenol was added to half the volume of the collected fraction and the same volume of chloroform (Fisher Scientific, Pittsburgh, PA.) was added and the mixture was spun down at 12000xg for 10 minutes at room temperature. The aqueous phase was collected in a fresh centrifuge tube and an equal volume of chloroform was added to the tube. The sample was spun down at 12000xg for 10 minutes at room temperature and the aqueous phase was collected in a fresh centrifuge tube. Twice the volume of cold 96% volume/volume ethanol (EtOH) was added and the sample was spun down at 1000xg for 5 minutes. The supernatant was removed and the pellet was washed with 250 µl of 70% EtOH and allowed to dry for 10 minutes. The pellet was then resuspended in 60 µl of water. Five ul of sample was run on 1% weight/volume agarose gel (0.4 grams of agarose in 40µl of TE buffer) to check for contamination and the concentration of the DNA was determined via spectrophotometry (Bio-Rad, Hercules, CA.).

2.3.2: Polymerase chain reaction (PCR) of *ndvB* Knockout

To confirm the KO of *ndvB,* the chromosomal DNA extracted from the wildtype and *ndvB* knockout strains was then used in a PCR. The PCR solution contained 35 ul of water, 10 uL of 10X PCR buffer (Invitrogen), 5 µL of $MgCl_2$ (Invitrogen), 2 µL of dNTP mix (Invitrogen), 2 µL of DNA, 2 µL of primer *ndvB*-L at 1mg/mL and 2µL of primer *ndvB*-R at 1mg/mL (Table 1), 1.5 µL of Taq DNA Polymerase (Invitogen) and 5.5 µL of DMSO for a final volume of 65 µL. The PCR reaction cycles were as follows: one cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 150 seconds followed by one cycle of 72 °C for 10 minutes. The PCR products were then run on a 1% weight/volume agarose gel to visualize the bands produced.

2.3.3: Confirming expression of ndvB in biofilms

The expression of *ndvB* in biofilm versus planktonic *P.aeruginosa* cells was confirmed via q-PCR analysis of wildtype and *ndvB* knockouct cDNA.

2.4.0: Quantitative real time polymerase chain reactions (q-PCR)

2.4.1: Total RNA extraction

2.4.1.1: Planktonic Culture

Stationary phase cultures of strains of interest were prepared by overnight growth in LB media at 37°C. The PureLink Micro-to-Midi total RNA purification system (Invitrogen) was used according to the manufacturer's instructions. Cells were lysed with a 10ml syringe and 20 gauge needle. On column DNase digest was performed (with one unit of DNAase from Sigma). RNA was resuspended in 50 µL of RNAse free water (Invitrogen). RNA purity and quality were checked by running 5 µl of product on a 1% weight/volume agarose gel and the concentration was determined by a spectrophotometer (Bio-Rad) reading. If deemed necessary a second in tube DNAse step was performed to remove any residual DNA contamination. Extracted RNA was stored at -80°C. Total RNA from planktonic culture was used for q-PCR.

2.4.1.2: Biofilm Culture

Biofilm cells were prepared in a 6-well plates (Falcon, Franklin Lakes, NJ.) using the Kadouri drip-fed system as described below (Section 2.5.1). The spent media were removed from the wells and 1ml of fresh 1XM63 was added. The bottom of the well was vigorously scraped with a plastic pipette tip to detach the cells from each well. The cells were transferred into a 1.5 ml microfuge tubes and spun down at 12,000xg. The supernatant was removed from centrifuge tube and the well was scraped again with 1 mL of M63 and added to the same tube. Cells were spun down at 12,000xg and collected. The total RNA was extracted using the same method that was used for the planktonic cell total RNA. Biofilm total RNA was used for q-PCR and Affymetrix microarray.

2.4.2: q-PCR

Reverse transcription and first-strand cDNA synthesis were performed on 2 μg of total RNA using the iScript reverse transcriptase first-strand cDNA synthesis system (Bio-Rad). RNA was primed with random primers to produce full-length cDNA transcripts. Reactions were incubated at 42°C for 45 minutes followed by a termination step at 72°C for 10 minutes. Quantification by q-PCR was carried out using SYBR-green detection of PCR products in real time using the MyiQ single-color detection system (Bio-Rad). In each experiment the reference standard, *rpoD,* was amplified using the primers *rpoD*-Q-L and *rpoD*-Q-R. Each q-PCR reaction (20 μL) contained 2 μL of cDNA, 10 μL of 2′ SYBR-green Master Mix (Applied Biosystems, Foster City, CA), including Amplitaq polymerase (Perkin-Elmer, Waltham, MA), and primers (Invitrogen) at a final concentration of 20 pM. All reactions were performed using the following thermal cycler conditions: 95°C for 5 minutes followed by 45 cycles of a three-step reaction consisting of denaturation at 95°C for 30 seconds, annealing at 56°C for 30 seconds, and extension and data collection at 72°C for 30 seconds. Q-PCR primers for the genes of interest can be found in table 1. Statistical analysis was performed with Bio-Rad gene expression analysis software via the Livak method (51). All experiments were run in triplicate with two biological samples.

2.5.0: Biofilm Growth

2.5.1: Kadouri Drip-fed reactor

Biofilms were grown in 6 well plates (Falcon) using the Kadouri Drip-fed reactor method (60). Figure 3 shows the set up of this system. One well of the plate acts as a chamber in which biofilms can be grown on the bottom and sides of the chamber. One 20 gauge needle was inserted into a polypropylene fitting made in the top of the 6 well plate for in flow of media into the chamber. Media were removed via a different 20 gauge needle and pumped into a waste container. The media were pumped into 4 wells via 4 channels on an 8 channel Watson Marlow pump (Wilmington, MA.). Four separate wells were used for each experiment. Between each experiment, the system was sterilized with 600 mL of 0.05% weight/volume sodium hypochlorite solution followed by 1L of sterile water to remove any excess sodium hypochlorite.The sterilization was carried out at an approximate rate of 150 mL/hour. New sterile 6 well plates and needles were used for each experiment and the tops with polypropylene fittings were sterilized with ethanol each time.

2.5.2: Inoculating the plates and running the reactor

Three mL of minimal media (1xM63) containing 3 g of KH_2PO_4, 7 g KH_2HPO_4 and 2 g $(NH_4)_2SO_4$ in 1L water supplemented with arginine (0.4% weight/volume) and $MgSO_4$ (1mM) was added to each well and 50 ul of overnight culture of bacteria was diluted into the media. Bacteria were allowed to attach to the surface of the wells for 4 hours at 37 °C. After this, the media were removed and 1 mL of fresh 1xM63 was added to each well. The plates were then attached to the reactor and sterilized 1xM63 was pumped in and out of the chambers at a constant rate of 35 mL/hour. Biofilms were grown for 48 hours, after which plates were removed from the reactor, washed 2x with 1mL of 1x M63 and then visualized under the microscope to check for confluence of biofilms on the plate. After this biofilms could be used for desired downstream applications.

2.5.3: Inserting green fluorescence protein (GFP) plasmid for visualization

Biofilms grown under the Kadouri drip-fed reactor were visualized under the Leica 6000 microscope after 0 (wash only) 8, 16,24 and 48 hours in order to determine the length of time needed to grow biofilms to produce sufficient biomass for downstream applications. The plasmid pSMC21 which contains the GFP gene was transformed into PA14 via electroporation. PA14 wildtype cells were grown to early log phase (A_{540}= 0.3 to 0.5) in 5 mL of LB. 1.5 mL of these cells were spun down in 1.5 mL centrifuge tube at 12,000xg for 3 minutes at 4°C and resuspended in 1.5 mL of 300 mM sucrose. These cells were then centrifuged as before and resuspended 750 µL of 300 mM sucrose. This step was repeated once more and cells were resuspended in a volume of 15 µL and chilled on ice for 30 minutes. 2 µL of plasmid DNA was then added to the cells and the total preparation was electroporated with a Bio-Rad Micropulser using preprogrammed *E. coli* setting 1. Cells were then added to 500 µL of LB and incubated for 1 hour at 37 °C. This solution was then plated on LB agar plates with kanamycin at 20 ug/mL. Individual colonies were selected and grown overnight in LB at 37 °C, checked for GFP insert using Leica 6000 microscope fluorescence and made into glycerol stocks as previously described.

2.5.4: Air-liquid interface assays (ALI):

Air-liquid interface (ALI) images were captured to observe any differences in the architecture of the various biofilms. An overnight culture of bacteria was grown in 5 mL rich medium (LB) overnight at 37°C and diluted 1/100 in 3 mL of 1xM63 media. A 6 well microtitre plate (Falcon) was placed in the 37°C incubator on an angle of approximately 30° and the well was filled with 1.5 mL of the diluted culture. This forms a line of liquid about half-way up the plate. After 24 hours of growth, the media was removed and the well was washed three times with water. One mL of water was left in the well to prevent drying and the biofilm air-liquid interface was visualized under the microscope (Leica CMI 6000B) and images were taken with the microscope and related software (Leica advanced fluorescence)

2.5.5: Crystal violet assays:

Crystal violet assays were conducted to determine if biofilm formation of the strains of interest occurred after a 24 hour incubation at 37°C. 1 in 100 dilutions of stationary phase cultures were prepared for the strains of interest in 1xM63. Wells of a 96-well microtiter plate (Falcon) were inoculated with 100 ul of the diluted cultures. The plate was incubated at 37°C for 24 hours followed by the removal of media and rinsing of the wells with water. 132 μl of 0.1% weight by volume crystal violet dye was then added to each inoculated well, and the plate was incubated for 20 minutes at room temperature. The dye was rinsed out twice with water and the plate was allowed to air dry. The presence or absence of a purple ring was observed in the wells. If a ring was present, then biofilms had formed in the initial 24 hour incubation. Six wells were inoculated with WT PA14, known to form biofilms in this assay, and six wells were inoculated with fresh M63 media without bacteria and viewd as controls for this experiment.

2.6.0: Microarray analysis

2.6.1: Preparing the RNA by concentration, precipitation and washing

RNA extracted from biofilms and used for microarray analysis was concentrated using ethanol precipitation. 150 μL of 96% volume/volume ethanol was added to 50 μL of RNA solution and 27.3 μl of 1 M sodium acetate was added to a final concentration of 0.12 M. The solution was mixed by inversion and stored at -20°C overnight. The solution was then centrifuged at 12,000xg for 15 minutes at 4°C and supernatant was removed. One hundred μL of 70% volume/volume ethanol was added to the pellet and left at room temperature for 5 minutes. The ethanol was removed and the pellet was left to air dry in centrifuge tube for 10 minutes. 25 µl of RNAse free water (Invitrogen) was used to dissolve the pellet, 5 µl was used to determine purity on 1% weight/volume agarose gel and 1 µl was used for concentration determination on a Bio-Rad spectrophotometer.

DNA, RNA or cDNA (ssDNA) used for down stream applications had its concentration determined via Bio-Rad Smartec Plus spectrophometer. One µl of sample was added to 99 µl of distilled water. The spectrophotometer was blanked with water and samples were read for concentration and A260/A280 ratio. Pre-programmed settings were used for determining concentration For dsDNA, an absorbance of 1 at 260 nm was equivalent to 50 µg/mL, for RNA an absorbance of 1 at 260 nm was equivalent to 40 µg/mL, and for ssDNA an absorbance of 1 at 260 nm was equivalent to 37 µg/mL.

2.6.2: Bio-Rad gel electrophoresis for RNA purity

RNA for microarray analysis was analyzed for concentration and purity by Bio-Rad gel electrophoresis system at the StemCore laboratories located at the OHRI at the Ottawa General Hospital Campus. RNA was extracted from 48 hour biofilms grown via Kadouri drip-fed method as previously described, concentrated by ethanol precipitation and resuspended in RNAse free water to a concentration of 2 µg/ul, as determined by spectrophometer. 1.2 µl of this solution was added to 28.8 µL of RNAse free water to create a solution of 50 ng/µL of RNA. Ten µL of this solution was sent in a 1.5 mL centrifuge tube for analysis via Bio-Rad gel electrophoresis system.

2.6.3: Affymetrix microarray

The Affymetrix (Santa Clara, CA.) microarray was run by StemCore laboratories at the OHRI Ottawa General Hospital Campus. Ten µl of 2µg/µl total RNA was used for each sample. The conditions studied were wildtype biofilms vs. *ndvB* knockout biofilms. Biofilms were grown for 48 hours via Kadouri drip-fed reactors in parallel. Two wells from 2 separate plates were used to grow 2 biofilms of wildtype and 2 biofilms of *ndvB* knockout each time. After 48 hours, total RNA was extracted from biofilm cultures as previously described and the two wildtype wells were combined to form a single wildtype sample and the two biofilms from the *ndvB* knockout were combined to make a single sample. After the RNA was collected, it was concentrated by ethanol precipitation and stored at -80°C until it was ready to be sent to StemCore labratories for processing. In total 6 wildtype biofilms were harvested and combined into 3 biological replicates for microarray analysis (named WT1, WT2 and WT3) and 6 *ndvB* knockout biofilms were collected into 3 biological replicate samples (named KO1, KO2 and KO3 for 'knockout sample') and sent for analysis. StemCore ran the samples on standard *P. aeruginosa* PA01 Affymetrix Genechips purchased from Affymetrix, using a standard scale protocol. A detailed protocol for the hybridization, washing and staining of the array can be found online (65)

2.6.4: Analysis of microarray data

Analysis of microarray data was performed using R statistical language and environment, V 2.6.2 (9) and statistical packages from Bioconductor, including 'simplyaffy' and 'affy packages'. The R code was adapted from previous published work (37). "Code chunks" used for the analysis can be found in the appendix of this thesis. The approach for this analysis was to normalize the data using robust multichip average (RMA) and Microarray Suite 5.0 (MAS 5.0) and to generate a list of significantly expressed genes from these two methods using standard t tests with a cutoff value of $p \leq 0.000001$ for a fold change of greater than 2. Genes that met both these criteria were included in the final list of genes to study. In addition, genes belonging to operons of the genes that were found on these lists were also added for study. RMA analysis coupled with Bayesian distribution was used to create a volcano plot and confirm the genes selected from the above method. A heatmap were created from the list of genes through R code.

2.7.0: Confirming the microarray data

2.7.1: Transposon insertion library

The roles of the genes of interest identified in the microarray experiments in relation to antibiotic resistance was studied using minimal bactericidal concentration (MBC). Assays were run using transposon insertion mutants in 7 genes identified from the microarray. Theses genes were *pqqB, pqqC, pqqE, pqqF, exaA, exaC* and *exaE.* These genes were chosen because they were identified in the array and we had transposon mutants available in the lab. The transposon insertion library was given to us by George O'Toole from the Dartmouth Medical School via the Ausubel Lab at Harvard University and is described in Liberati et al. (50). In brief, the library consists of non-redundant Tn7 mariner (MAR2xT7) transposon insertions into genes of the PA14 genome. Each gene that was successfully disrupted with this transposon was clonally isolated, grown up in culture and sequenced. Because of this, the location of each transposon within the gene is known and included in the set. The mariner transposon carries a gentamicin resistant cassette for colony selection and two T7 promoters, having a total length of 994 basepairs. The strains obtained were stored in 96-well microtitre plates at -80°C in 25% volume/volume glycerol. Strains of interest were picked from the plates, grown overnight in 5 mL of LB at 37 °C and then diluted 1/1000 and streaked out on LB agar plates. These plates were incubated overnight at 37 °C to allow colonies to grow. Single colonies were picked, and subsequently grown up overnight in 5 ml of LB. Seven hundred µl of the overnight culture was added to a 2 mL cryogenic vials (Sarstedt, Numbrecht, Germany) tube and an equal volume of 50% volume/volume glycerol was added to make a stock of the strain of interest for use at a later time.

2.7.2: Confirming the transposon insertion by PCR

A pair of primers (Invitrogen) specific for each of the genes studied was used to amplify the DNA of transposon insertion strains (prepared as described previously in section 2.3.1). Table 1 shows a list of primer pairs used for the PCR of the mutant genes. For example *exaE-F-L* (*exaE*-Full length-Left) and *exaE-F-R* (*exaE*-Full length-Right) were used to PCR the *exaE* gene. The full length primers were chosen from the 500 base pair upstream and down stream region so as to ensure that the whole of the gene was amplified. The primers were also used to amplify the DNA in the *P. aeruginosa* wildtype strain. Amplification was performed by PCR in a volume of 60 µl with Taq polymerase (Invitrogen) and the following incubations: 4 minutes at 95°C, followed by 30 cycles (with 1 cycle consisting of 30 seconds at 95°C, 2 minutes and 30 seconds at 53.5°C, and 1 minutes and 30 seconds at 72°C), and a final step of 7 minutes at 72°C. The PCR products were analyzed by 1% weight/volume agarose gel electrophoresis to determine the length of the gene in the wildtype strain versus the transposon insertion mutant strains. Table 2 shows PCR running conditions for all experiments done for this thesis.

2.7.3: Growth curves

To ensure that the rate of growth of each of the mutant strains used in the study was similar and that the mutants were able to grow as well as wildtype in planktonic cultures, growth curves were conducted. Overnight strains of bacteria were diluted into 60 mL of fresh LB broth and the OD600 was taken at t=0 and recorded. Bacteria were grown at 37° C with shaking and 3 ml aliquots were removed every 30 minutes, the OD was recorded and the culture was returned to the flask. The OD600 was measured until an OD was reached that was consistent for at least three time points (indicating stationary phase). Plots of OD vs time were generated. This experiment was repeated three times and the average OD at each time point was plotted,

2.7.4: Minimum Bactericidal Concentration Planktonic/Biofilm (MBC P/B)

To determine if a gene is important in biofilm specific antibiotic resistance, Minimal Bactericidal Concentration (MBC) assays of cells grown in planktonic (P) and biofilm (B) conditions were performed as previously described (57). These assays allow for a direct comparison between planktonic and biofilm antibiotic resistance in bacterial cells. For the MBC-B protocol, bacteria were diluted 1/50 in 1X M63, 100 μL was inoculated into the wells of a 96 well microtitre plate and allowed to form a biofilm for 24 hours at 37°C. Spent 1X M63 media was replaced with 90 μL of fresh media and serial dilutions of 10 μL of the antibiotic of interest were added. Bacteria were incubated for 24 hours at 37°C. Following this growth period, media containing the antibiotic was replaced with 100 μL of fresh media and the bacteria that survived antibiotic treatment were incubated for 24 hours at 37°C. Live bacteria were transferred onto an LB agar plate via a multi-prong device (Dan-Kar Corp, Woburn, MA) and grown overnight at 37°C in order to determine the concentration of drug that had killed the biofilm grown cells completely. For the MBC-P assay, bacteria were diluted 1/50 in 1X M63, 90 uL was inoculated into the wells of a 96 well microtitre plate and serial dilutions of 10 uL of the antibiotic of interest were added to the wells. The bacteria were incubated for 24 hours in the presence of antibiotic and then transferred to an LB agar plate to determine the concentration of antibiotic that killed the planktonic bacteria.

2.7.5: Confirming expression q-PCR

Microarray gene expression data was confirmed by q-PCR for the genes of interest on the list. Q-PCR was done as described previously and the list of primers can be found in Table 1. A graph of the relative fold change via q-PCR vs Microarray gene fold change was plotted for the 7 genes.

Table 1: List of primers used for this study.

List of all the primers used in this thesis for PCR and q-PCR reactions as well as transposon insertion confirmation. All primers were ordered from Invitrogen and were 50 mM desalted.

Primer Name	DNA Sequence	Application
ndvB-L	5’-CAACCACCTCTGAAACCGAATCGCGGATTG-3’	PCR
ndvB-R	5’-GACCAGCTTCCAGGTTTCCAACCTGAAC-3’	PCR
pqqB-F-L	5’-CTACCAAATCAGGGGGAAAA CT -3’	PCR
pqqB-F-R	5’-TCCTCGAGGATACGCTGGAT-3’	PCR
exaA-F-L	5’-AG GTTGGCTTTTTCCAGGTT-3’	PCR
exaA-F-R	5’-GGCCTGAACATCTTCTTCACC-3’	PCR
pqqB-Q-L	5’-GACGAGATGCT GGTCTGC-3’	q-PCR
pqqB-Q-R	5’-ATTGGTGTTGTTGATGTGGA-3’	q-PCR
exaA-Q-L	5’-CAAGGCGTTGTTCT TGTTCA-3’	q-PCR
exaA-Q-R	5’-AGGTTGGCTTTTTCCAGGTT-3’	q-PCR
ndvB-Q-R	5’-GATCTTGCCGACCTTGA AGAC-3’	q-PCR
ndvB-Q-L	5’-GGCCTGAACATCTTCTTCACC-3’	q-PCR
rpoD-Q-L	5’-CATCCGCATGATC AACGACA-3’	q-PCR
rpoD-Q-R	5’- GATCGATATAGCCGCTGCTGAGG-3’	q-PCR
agmR-Q-L	5’-AGGTAGACGTTGCCGTTGAG-3’	q-PCR
agmR-Q-R	5’-CATTCCCGTGGTTATCGTCT-3’	q-PCR
pqqC-Q-R	5’-CGATAGTGCTCCAGGGTGAT-3’	q-PCR
pqqC-Q-L	5’-GTCGACGCCTACGTCAACTT-3’	q-PCR
pqqD-Q-L	5’-ACTGCAGCTCGATCCAGAAT-3’	q-PCR
pqqD-Q-R	5’-GACTGCCATCTGCTGCTCTA-3’	q-PCR
pqqE-Q-L	5’-GGCACAACATCGACAACATC-3’	q-PCR

pqqE-Q-R	5’-GTAGTAGTCGGGGGTGCACGA-3’	q-PCR
pqqF-Q-L	5’-GCTGCACCAACTGTCTCCATA-3’	q-PCR
pqqF-Q-R	5’-GGCGTTGCTGTTGTTCTGT-3’	q-PCR
exaC-Q-L	5’-ACGACCACACAGGAGAGCTT-3’	q-PCR
exaC-Q-R	5’- TTGGTAGGGACTGCCTGTTC-3’	q-PCR
exaE-Q-L	5’-GGTCAAGGGCCAGTACTTCA-3’	q-PCR
exaE-Q-R	5’-GCGATCTTCAGCAGGATGTT-3’	q-PCR
pqqC-F-L	5’- CGCAAGCTCAAGGTACACAA-3’	PCR
pqqC-F-R	5’-CGACAGGTCGACAACTACCA-3’	PCR
pqqD-F-L	5’-ATCACCCTGGAGCACTATCG-3’	PCR
pqqD-F-R	5’- GGACAGCGATAGGTCAGCTC-3’	PCR
pqqE-F-L	5’- ATGAGCATGGCCTACGAACT-3’	PCR
pqqE-F-R	5’- AGTCGCCGTAGTCGACCATT-3’	PCR
pqqF-F-L	5’- CCAGGACAACCTGTTCCACT-3’	PCR
pqqF-F-R	5’- TCGTCCTTGTCGTACAGCAG-3’	PCR
exaE-F-L	5’-ATCTCTACCGCCTGTTGCAG-3’	PCR
exaE-F-R	5’-CCAACCTGCAGGAAGTCAAG-3’	PCR
exaC-F-L	5’-AAATTCCTTCGCTGGTCATC-3’	PCR
exaC-F-R	5’-AAGAACAGCACACGCAACTG-3’	PCR

Table 2: PCR running conditions.

List of the PCR runs for each gene and the conditions used for the runs.

PCR reaction	Cycle Information
ndvB Wildtype and *ndvB* knockout	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 150 seconds followed by one cycle of 72 °C for 10 minutes.
pqqB wildtype and *pqqB*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 130 seconds followed by one cycle of 72 °C for 10 minutes.
pqqC wildtype and *pqqC*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 54 °C for 50 seconds, 72 °C for 180 seconds followed by one cycle of 72 °C for 10 minutes.
pqqE wildtype and *pqqE*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 130 seconds followed by one cycle of 72 °C for 10 minutes.
pqqF wildtype and *pqqF*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 54 °C for 50 seconds, 72 °C for 200 seconds followed by one cycle of 72 °C for 10 minutes.
exaA wildtype and *exaA*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 54 °C for 50 seconds, 72 °C for 210 seconds followed by one cycle of 72 °C for 10 minutes.
exaC wildtype and *exaC*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 210 seconds followed by one cycle of 72 °C for 10 minutes.
exaE wildtype and *exaE*tn	One cycle 95°C for 5 minutes followed by 35 cycles of 95 °C for 30 seconds, 56 °C for 50 seconds, 72 °C for 210 seconds followed by one cycle of 72 °C for 10 minutes.

Results

Results

In order to fulfill the objectives of this thesis, experiments were performed to confirm the expression of *ndvB* in a biofilm specific manner as well as confirm the *ndvB* gene deletion so that this strain could be used in further studies. Once this was accomplished, microarray analysis was performed using RNA from 3 wildtype biological samples and 3 *ndvB* knockout samples at StemCore laboratories located at the Ottawa Health Research Institute (OHRI) at the Ottawa General Hospital. A list of differentially expressed genes between the two conditions was generated and 8 of the genes from the microarray were chosen for further study. Validation of the array data using quantitative real-time polymerase chain reaction (q-PCR) and antibiotic resistance profile assays (Minimum bactericidal concentration (MBC) planktonic (P) and biofilm (B) assays) was done.

3.1.0: Preparing for the microarray

3.1.1: Confirming the knockout *ndvB* strain

Prior to conducting RNA extraction for microarray analysis it was important to confirm the knockout of the *ndvB* that was previously created (57). The knockout was prepared by three step ligation and approximately 1000 basepairs of the *ndvB* gene was removed which eliminated the production of the *ndvB* gene product. Because there was a 1000 basepair deletion, this knockout could be confirmed using polymerase chain reaction of the knockout *ndvB* (KO) strain and comparing it to the wildtype (Wt) strain PCR product on a 1% weight/volume agarose gel (Fig. 3A). The size of the bands was determined by generating an equation for the line of best fit for a standard log molecular weight vs. migration distance. A standard 1kb+ ladder was used and the known basepairs were multiplied by 660 to obtain the molecular weight in Daltons. The log of this number was plotted against the migration distance of the band from an arbitrary line as shown in Figure 3B. Once a line was generated and an equation determined, the migration distance of the bands of interested was inserted into the equation and the number of base pairs was determined. A sample calculation can be found in appendix A. The expected size of the Wt *ndvB* was 3689 basepairs (bp) and the band was around 3560 bp, while the KO was supposed to have a product of around 2679 bp and the band was visualized at about 2430 bp.

3.1.2: Confirming gene expression of *ndvB*

ndvB was previously described as being expressed preferentially in biofilm grown cells (57). In order to confirm this, and determine the gene expression level in the KO strain, I ran a quantitative real time polymerase chain reaction (q-PCR) of the Wt and KO strains grown in biofilms and planktonic culture. RNA was extracted from biofilms grown for 48 hours in a Kadouri drip-fed reactor. RNA was extracted from 3 mL planktonic culture grown overnight in 5 mL of LB. Both the Wt and *ndvB* KO strains were used for analysis. Figure 4 shows the results of this assay and it is clear that there is a significant increase in expression of *ndvB* between biofilm and planktonic conditions. This was confirmed statistically by a two-tailed student T-test. In addition to this, the KO strain showed no expression of the *ndvB* in either the planktonic or biofilm states.

These data confirm that the *ndvB* knockout strain does not express *ndvB*. This confirms data already previously published (57).

3.1.3: Growing the biofilms and extraction of RNA for microarray analysis

In order to produce sufficient amounts of RNA to be used for microarray analysis I had to choose a method for growing and extracting RNA that would give a high yield of pure RNA. I chose to grow biofilms in the Kadouri drip-fed reactor as described because it is a small scale reactor for biofilm growth and allows extended periods for growth due to the addition of fresh media and removal of waste products. Figure 5 depicts the growth of biofilms over time in this reactor. Forty-eight hour biofilms were chosen because they gave the most amount of biofilm biomass over a reasonable amount of time, and allowed for efficient extraction of RNA at high yield and purity.

Efficient extraction of RNA from biofilms was important in order to proceed with downstream microarray analysis. Table 3 shows the concentration and purity of the RNA extracted from biofilms and used for microarray analysis. A higher concentration of RNA was needed than originally obtained from the extraction, although the amount of RNA was sufficient. For this reason, I chose to keep the conditions of biofilm growth and extraction the same, and concentrate the RNA via ethanol precipitation. This allowed me to achieve the concentration required for microarray analysis as defined by the protocol used by StemCore technicians. Table 3 shows the final concentration of RNA used for the microarray analysis.

The purity of the RNA samples was also checked via Bio-Rad experion electrophoresis system at StemCore laboratories. Figure 6 shows the 1% weight/volume agarose gel of the samples used and Figure 7 shows the results of the Bio-Rad experion experiment. The 23S (larger subunit, around 3900 bp) and the 16S (smaller around 1500 bp) ribosomal subunits can be seen on the gel in Figure 6 and are represented as the major peaks in the graphs of Figure 7. The expected ratio is about 2 to 1 for 23S to 16S subunits and all of the RNA samples had a ratio between 1.8 and 2.0 as determined by the Bio-Rad experion system. The concentration was also accurately determined by the Bio-Rad experion system and purity and concentration of the samples was deemed to be sufficient for microarray analysis and therefore these samples were used to conduct the experiments.

3.2.0: Microarray analysis of samples

The majority of this project focused on the differential expression of genes between wildtype (Wt) and knockout *ndvB* (KO) conditions. Because *ndvB* is preferentially expressed in biofilms and is important in biofilm specific mechanisms, I believe it is responsibe for regulating genes that are important for biofilm specific resistance to antibiotics. In order to test this hypothesis, microarray analysis of the two conditions, deemed Wt and KO conditions, was performed.

3.2.1: Confirming the array quality

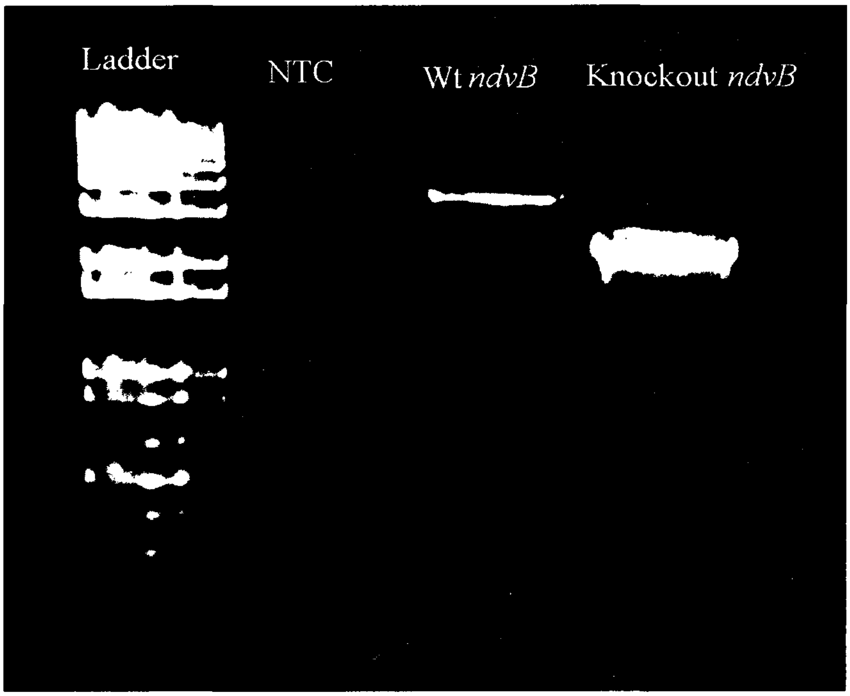
All of the array analysis was done using R statistical coding environment and Bioconductor packages freely available, using code modified from published projects (37). The code for the operations performed can be found in appendix B. In order to ensure that the data were of sufficient quality, different quality control steps were used to check the probe level data and the chip quality. Table 5 shows the general parameters of the array chips as given by StemCore laboratories. The number of genes present was comparable to other studies that showed a range of 89% to 70% genes present depending on conditions used (52, 97). Affymetrix suggests that chips are comparable if the noise of each chip is within 2 standard deviations of one another and the scale factors are within 4 fold of each other (13, 37, 73). The data in Table 4 from StemCore suggests that these chips were comparable to each other. As a result, it is reasonable to proceed with a more comprehensive analysis of the chip data.

The raw chip data can be visualized as probe points to identify chips that might give problematic results for analysis, if the chip looks radically more light or dark when compared to ships of the same sample (i.e. WT1 vs. WT2) then the data would not be suitable for comparison. Figure 8 gives a representation of the perfect match (PM) and mismatch (MM) intensities of the preprocessed (before normalization or standardization) Affymetrix Genechip from the experiment. Figure 8A shows a representative array of the Wt condition and 8B shows a representative array of the KO condition. The other chips can be seen in appendix C. Along with the raw probe data of the Genechips Figure 8 shows histograms and density plots of the array. Chips of similar conditions (i.e. Wt chips compared to KO chips) should have a similar distribution of intensities if they are of similar quality and are suitable for comparison.

Figure 3: Confirming the *ndvB* knockout with PCR.

A) The *ndvB* gene deletion was confirmed using full length PCR of the wildtype and comparing it to the full length PCR product from the mutant using *ndvB* specific primers. The expected size of the ndvB wildtype PCR product was 3689 and the expected PCR product for the knockout *ndvB* is 2279. The first lane held the 1kb+ ladder, the second lane held the no template control (NTC), the third lane held the wildtype (Wt) *ndvB* PCR product and the fourth lane held the knockout *ndvB* product.
B) Determination of the PCR product size of *ndvB* wildtype and *ndvB* knockout via migration distance. The size of the PCR product was determined by generating a line of best fit for the log of the molecular weight of the standard 1kb+ plus vs. migration distance of the bands and using the equation of the line. The band size of the wildtype was approximately 3560 and the knockout product was approximately 2439 basepairs

A


Ladder
NTC
Wt ndvB
Knockout ndvB

B

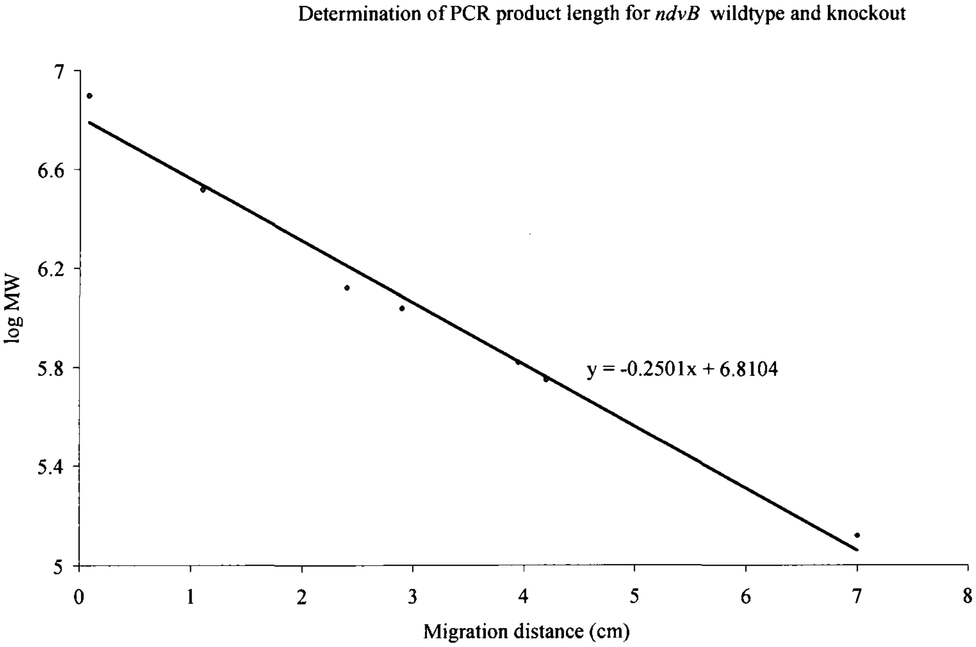

Determination of PCR product length for ndvB wildtype and knockout
log MW
7
6.6
6.2
5.8
5.4
5
y = -0.2501x + 6.8104
0
1
2
3
4
5
6
7
8
Migration distance (cm)

Figure 4: q-PCR of *ndvB* gene expression
The gene expression of *ndvB* in biofilm and planktonic states was confirmed using q-PCR. RNA was extracted from planktonic cultures in stationary phase or 48 hour biofilms. cDNA was synthesized using random primers from Bio-Rad script cDNA synthisaze kit and q-PCR was performed with Bio-Rad SYBR-green supermix. RNA was extracted from wildtype and knockout *ndvB* strains to confirm elimination of *ndvB* RNA expression in the knockout. Statistical significance between means was confirmed using the student t-test.

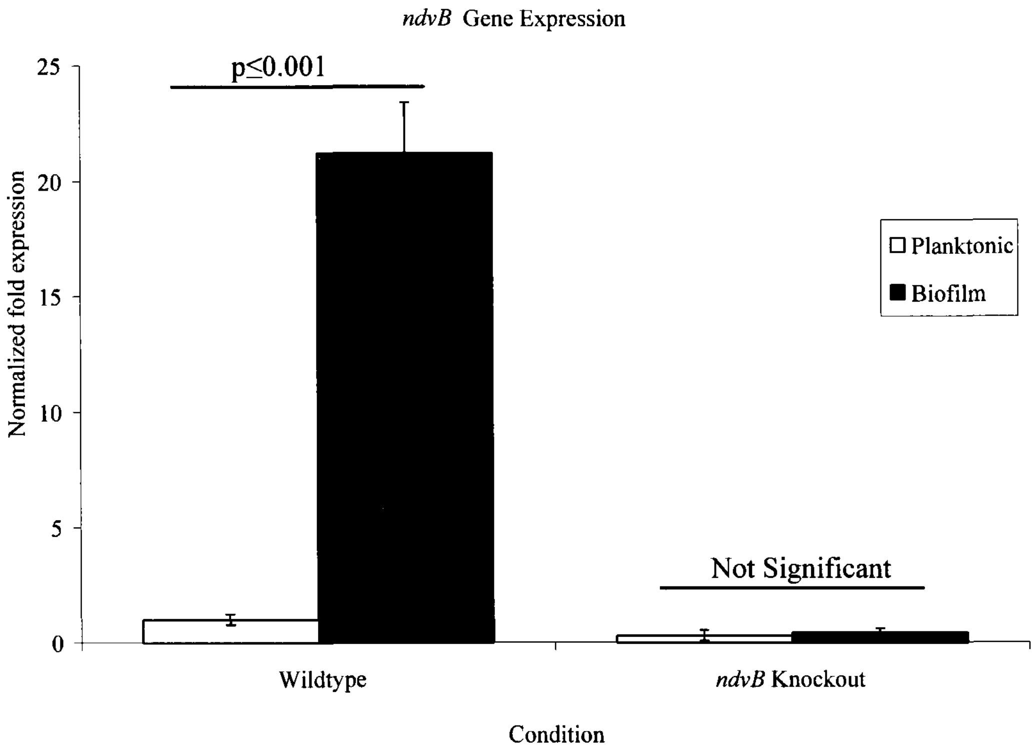
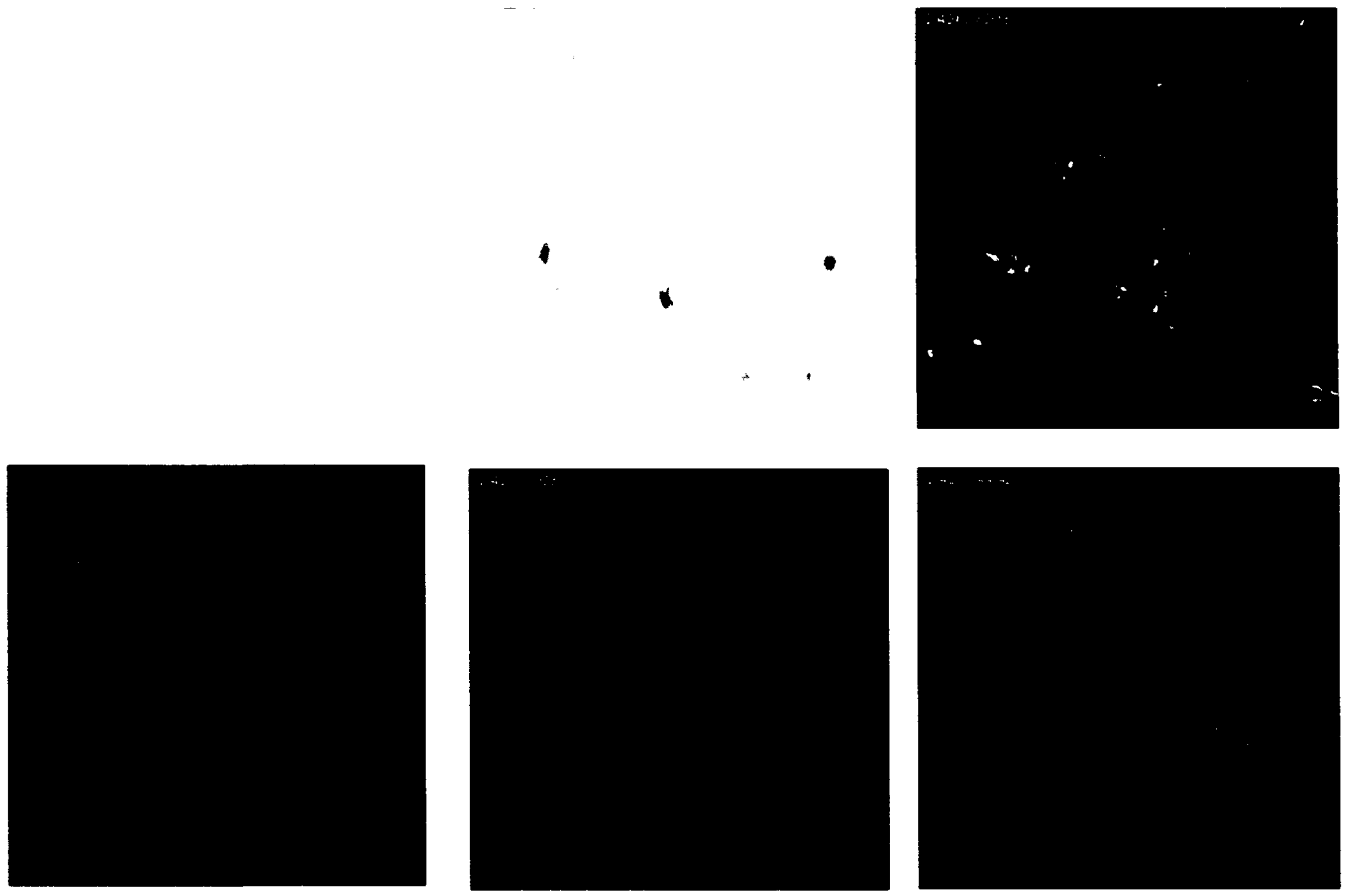

ndvB Gene Expression
Normalized fold expression
0
5
10
15
20
25
p≤0.001
Not Significant
☐ Planktonic
■ Biofilm
Wildtype
ndvB Knockout
Condition

Figure 5: Biofilm growth in Kadouri drip-fed reactor.

Biofilm growth via the Kadouri drip-fed reactor was monitored visually by the Leica 6000B microscope. The reactor was inoculated with a GFP strain of PA14 *Pseudomonas aeruginosa* and visualized at the time points (0, 4, 16,24 and 48 hours) under phase and fluorescence settings under the microscope at 200X magnification.



O Hours

24 Hours

48 Hours

Figure 6: Agarose gel of RNA for microarray analysis.

RNA of wildtype and knockout *ndvB* strains of *P. aeruginosa* was extracted from 48 hour biofilms via the Invitrogen Micro to Midi RNA extraction kit and run on a 1% agrose gel. The RNA was designated WT1, WT2 and WT3 for wildtype grown RNA biological replicate 1, 2 or 3 and KO1, KO2 and KO3 for knockout *ndvB* biofilm RNA biological replicate 1, 2 or 3.


Ladder
WT1
WT2
WT3
KO1
KO2
KO3

Table 3: Concentration, purity of RNA.

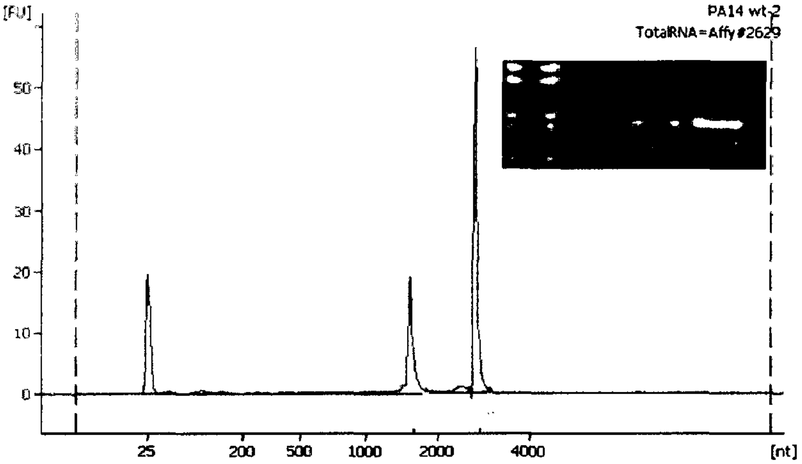
The concentration of each of the RNA samples used for microarray analysis and purity based on A260/A280 ratio. The concentration and purity of each sample was taken before and after ethanol precipitation

	Before Concentration			After Concentration		
	Concentration (μg/mL)	A_{260}/A_{280}	μg RNA	Concentration (μg/mL)	A_{260}/A_{280}	μg RNA
WT1	177.0	1.56	17.8	988	1.94	15.4
WT2	180.2	1.73	18.2	907	2.01	13.6
WT3	191.0	1.48	19.6	918	2.21	14.2
KO1	161.0	1.56	16.4	807	1.98	12.1
KO2	188.0	1.65	19.0	986	2.13	15.3
KO3	178.6	1.87	18.0	933	2.11	14.0

Figure 7: Bio-Rad gel electrophoresis for RNA purity and concentration.

The purity of the RNA was confirmed using Bio-Rad gel electrophoresis system. The system uses a small amount of RNA and guides it through a series of channels in a microfluidic chip by applying various voltages. From this an expression pattern is generated with pure RNA generating a 23S (at 3000 basepairs and 16S (at 1500 basepairs) ribosomal RNA peak in a 2 to 1 ratio.
A) Representative gel of Wildtype RNA samples. The inset in the first diagram of A shows 1% agarose gel of WT RNA and KO1 RNA . Lane 1 shows the ladder, Lane 2 shows empty, Lane 3 is WT1 RNA, Lane 4 is KO1 RNA
B) *ndvB* knockout representative gel.

A


PA14 wt-2
TotalRNA=Affy#2629
[FU]
50
40
30
20
10
0
25
200
500
1000
2000
4000
[nt]

B

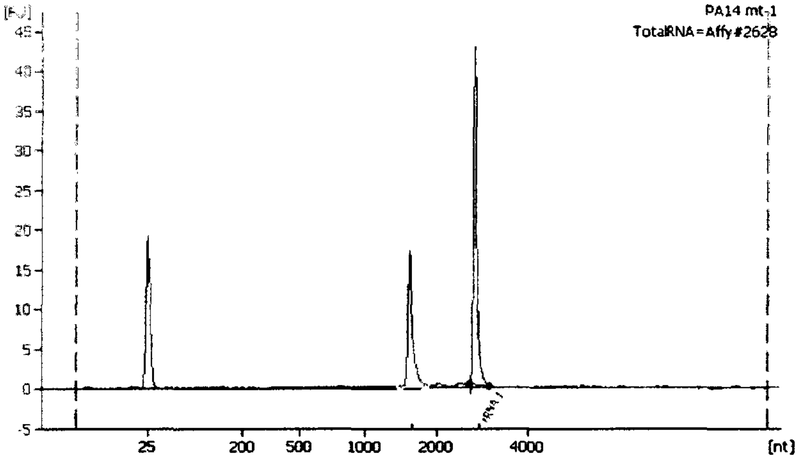

PA14 mt-1
TotalRNA=Affy#2628
[FU]
45
40
35
30
25
20
15
10
5
0
-5
25
200
500
1000
2000
4000
[nt]

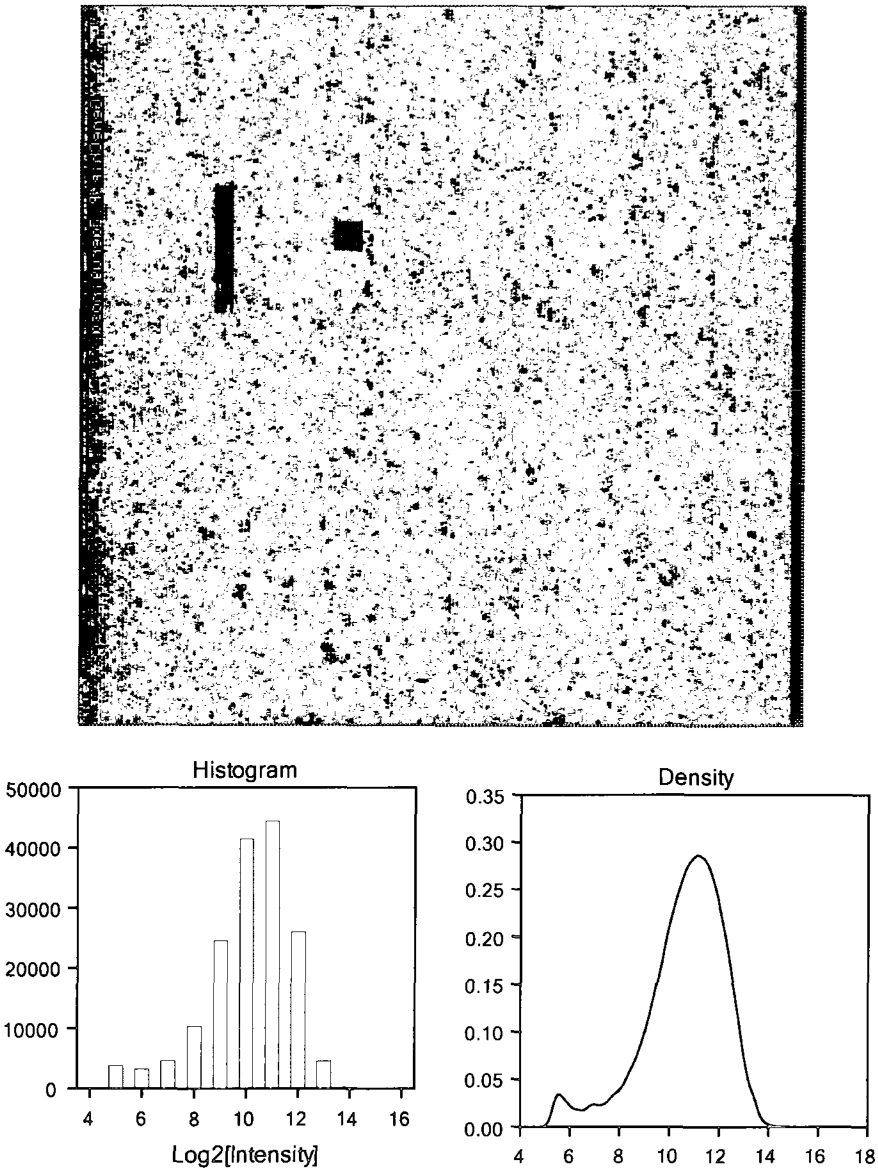
Figure 8: Microarray slide images and histograms.

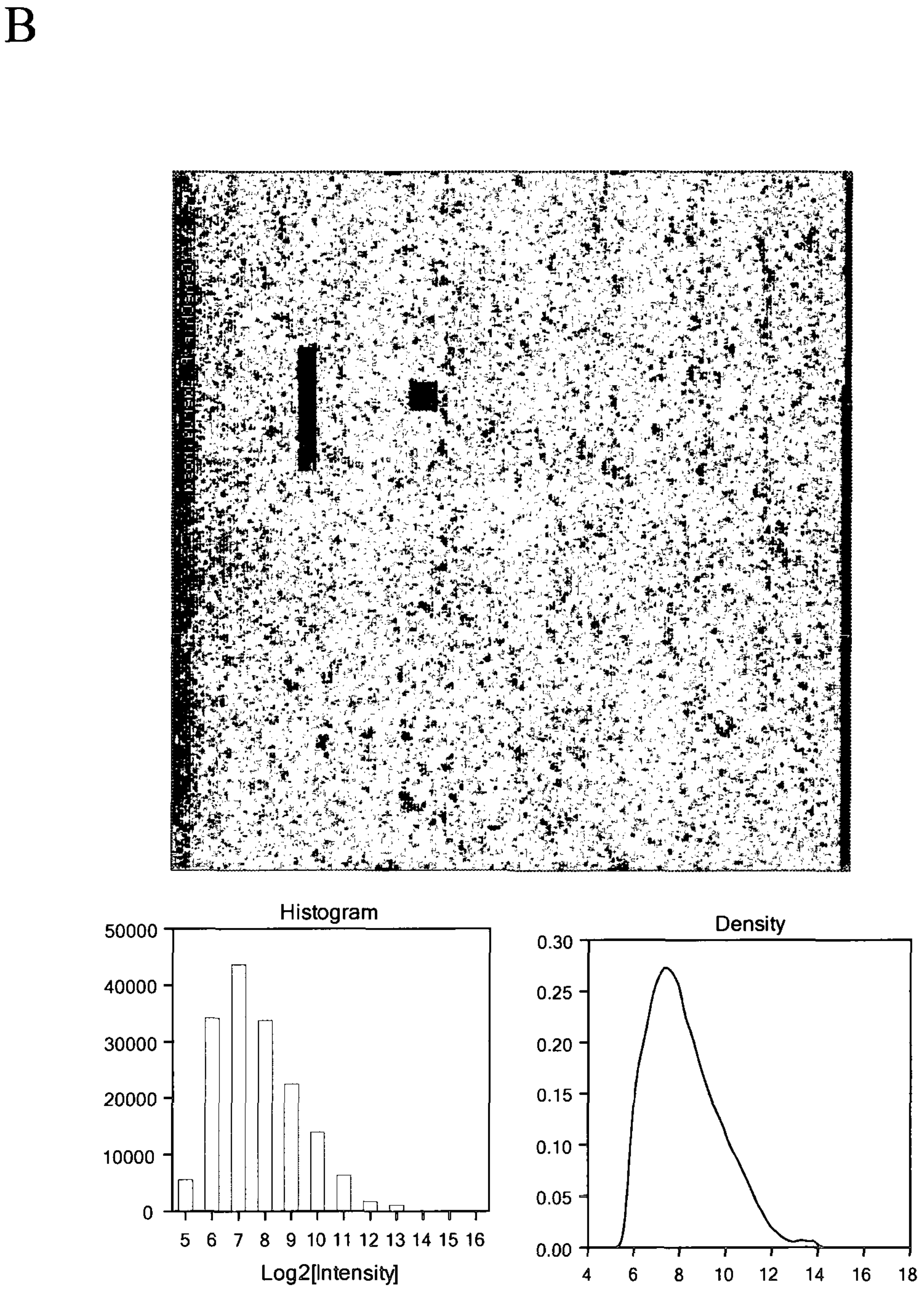
Image plots of perfect match (PM) and mismatch (MM) probe intensities (log base 2) for two arrays along with density plots of probe-level data. chip data for WT2,WT3, KO2 and KO3 chips can be found in Appendix C. Good quality chips show similar intensity patterns among the conditions studied. A darker or lighter chip may indicate problems. The histogram shows the number of probes at a given $\log_2$ intensity value. The density plot shows the percentage of probes that are at a given $\log_2$ intensity for the chips.

A) Example of wildtype (WT1)

B) Example of *ndvB* knockout (KO1).

A


Histogram
50000
40000
30000
20000
10000
0
4
6
8
10
12
14
16
Log2[Intensity]
Density
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4
6
8
10
12
14
16
18


B
Histogram
50000
40000
30000
20000
10000
0
5 6 7 8 9 10 11 12 13 14 15 16
Log2[Intensity]
Density
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4 6 8 10 12 14 16 18

The preprocessed chip data were shown to be suitable for comparison and of good quality. The MA plots in Figure 9 show a plot of the difference in log expression (M) against the average log intensities (A) of the condition chip (eg. WT1) vs a median synthetic chip created by pairwise comparisons of all data points before processing for each of the chips. Poor quality chips will result in widely different patterns and vastly different Lowess lines. Because this was unprocessed data, the spread of the Lowess line will be slightly more dramatic then with normalized data, and the cluster cloud in the graph is less centered on zero. Both Figure 8, in conjunction with the chip data found in the appendix, and Figure 9 indicate that the preprocessed chip data are comparable and allowed for confident generation of genes differentially expressed between the two conditions studied.

It is important to confirm the accuracy of the normalized data. Figure 10, showing the relative log expression (RLE) boxplots and Figure 11, showing the normalized unscaled standard error (NUSE) boxplots are indicators of the quality of the normalization procedure. These figures were generated from the Robust Multichip Average (RMA) normalization procedure, which uses quantile normalization of the probe level data, which makes the distribution of all probe intensities amongst the arrays equal. A high quality normalization procedure will produce RLE plots that are centered around 0 and show minimal spread of the box plots amongst the arrays (13, 37). Figure 10 shows that the RMA procedure does produce a good RLE for the arrays. Similarly, NUSE box plots should be centered around 1 with minimal spread amongst the box plots (13, 37). Figure 11 shows a NUSE for the RMA normalization of the arrays and there does not seem to be any vastly different boxplots amongst the arrays, suggesting that the normalization procedure produces accurate correction for the arrays. This provides reassurance that the genes identified to be differentially expressed under these conditions are actually differentially regulated. Taken together, the arrays run using the WT and KO samples I prepared were used to generate a list of genes that were used for further study.

3.2.3: Microarray Results

The microarray analysis revealed 24 genes that had a fold change over 2 and a p-value of less then 0.000001. A list of these genes can be found in Table 5. *ndvB* itself had a p-value slightly higher than this, but had a fold change of 6.46 greater expression in the wildtype (Wt) condition. Twenty of the genes were upregulated in the Wt condition and 4 genes were down regulated. Several hypothetical genes were identified; some small molecule transporters, transcriptional regulators and genes involved in glucose metabolism were identified.

Interestingly, 6 genes involved in ethanol oxidation, as well as the regulator of this network were identified in the microarray. The metabolism genes are *pqqB/C/D/E* , *exaA* and *exaE* and the regulator of these genes is *agmR*. All were shown to be upregulated in the Wt condition through the microarray. The regulatory network has been identified and is represented in Figure 21. The location of these genes in PA14 genome is shown in Figure 22. Because these genes are all related and identified as significant in the microarray, along with genes associated with them, were chosen for further study.

Figure 12 shows a Heatmap representing of the different genes expressed in the microarray. Various shades of green represent increasing fold expression within a condition, black depicts no change in expression. Bright green represents the highest fold expression, being about 6x more abundant. Bright red represents the most downregulated gene in a condition. The other genes identified in the microarray belong to a variety of different classes, including transcriptional regulators, small molecule transport and outer membrane proteins. Due to time constraints, it was not possible to confirm the regulation and antibiotic profiles of all of these genes, but this will be further studied.

3.3.0: Confirming microarray expression with q-PCR

In order to confirm that the genes were differentially regulated, I performed quantitative real time polymerase chain reactions (q-PCR) on a subset of the genes from the microarray. Because *ndvB* is preferentially expressed in biofilms, and because I believe that *ndvB* is important in regulating these genes, I performed q-PCR on Wt RNA and KO *ndvB* RNA extracted from planktonic bacteria and biofilm bacteria. All runs were normalized to *rpoD,* a housekeeping gene in *P. aeruginosa* that has been deemed suitable for normalization in q-PCR experiments (79). Each gene was run with two biological replicates of planktonic and biofilm RNA, repeated in triplicate within each experiment. The mean of the

Table 4: Gene chip data.

The table shows the standard quality parameters given for each Affymetrix Gene chip run by StemCore. Noise and scale factors of each chip should be within 3 fold of each other to have chips that are comparable.

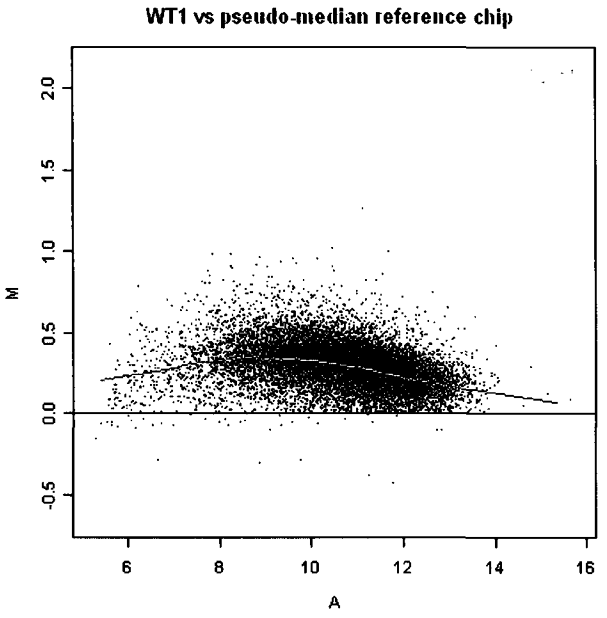
Chip	Noise	Scale Factor	% genes present
WT 1	4.2	0.841	93.5
WT2	4.33	1.043	93.9
WT 3	3.75	2.25	92.5
KO 1	3.02	3.00	76.6
KO 2	3.02	2.23	72.2
KO 3	2.30	2.48	60.8

Figure 9: Median vs Average plots of experimental chips and a pseudo median chip.

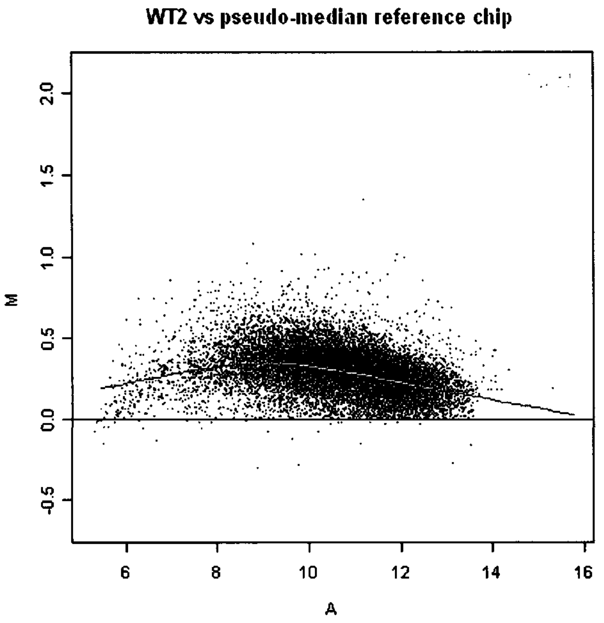
Quality assessment of chip data by a plot of the difference of their log intensities for each probe on each gene (M) plotted against their average intensities (A). This is compared to a synthetic array generated by taking the pair wise comparison of all probe level data. Each plot has a lowess regression line added to the plot. IQR= Interquartile region

A) WT vs Mock median chip

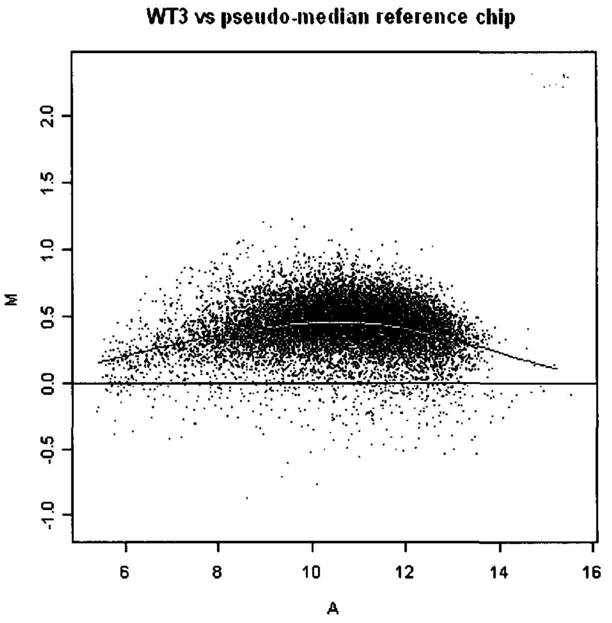
B) KO MA plots vs Median mock chip.



Median: 0.394
IQR: 0.161

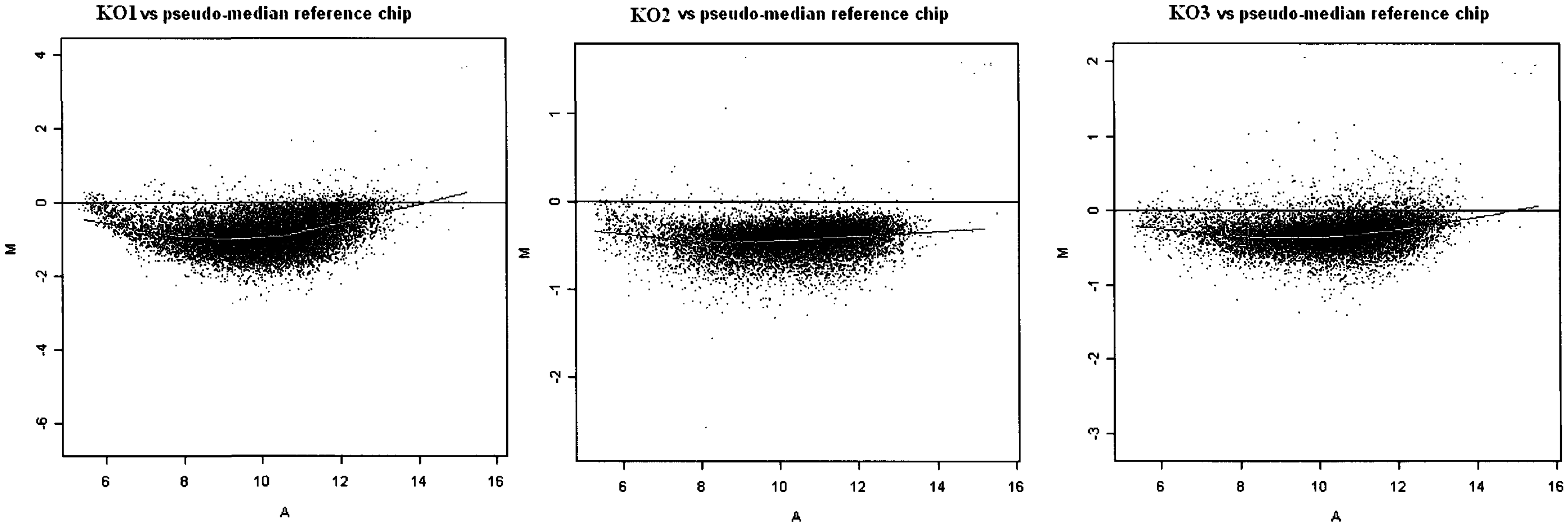


Median: 0.494
IQR: 0.176



Median:0.429
IQR:0.224

B



Median: -0.694
IQR: 0.191

Median: -0.411
IQR: 0.183

Median:-0.215
IQR:0.200

Figure 10: Relative Log Expression (RLE) box plots.

Relative log expression plots of the 6 arrays for quality control of probe level models (PLM). This model estimates for probe sets and chips (arrays) on a. gene by gene basis. RMA was used for background correction and quantile normalization was used for to generate this data. Appropriate normalization results in boxplots centered around 0 with minimal spread and error associated from plot to plot.

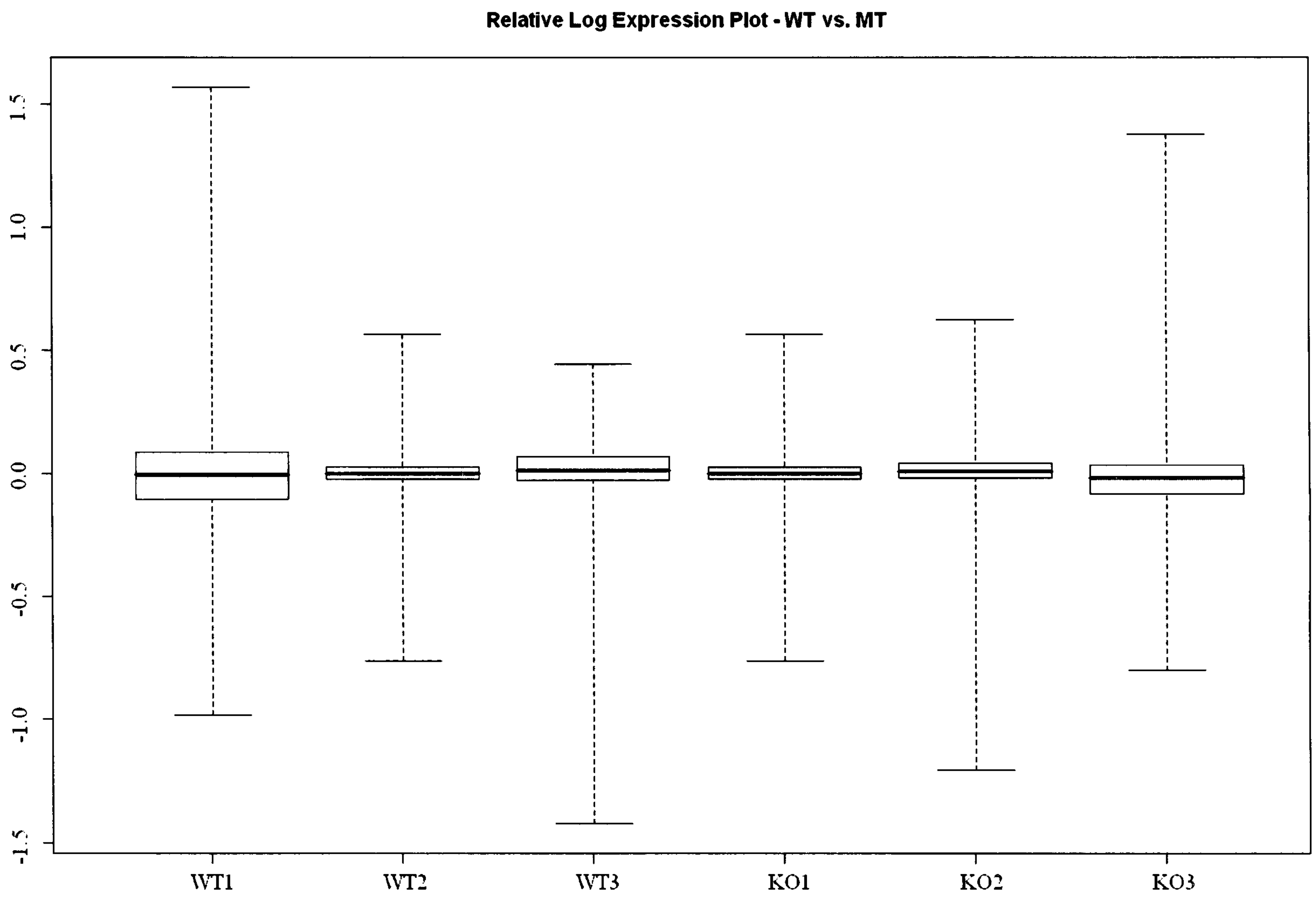

Relative Log Expression Plot - WT vs. MT
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
WT1
WT2
WT3
KO1
KO2
KO3

Figure 11: Normalized Unscaled Standard Error (NUSE) plots

Plots of standardized standard error estimates such that the median standard error across arrays is 1 for each gene. Appropriate normalization results in boxplots centered on 1 with minimal spread and error associated from plot to plot.

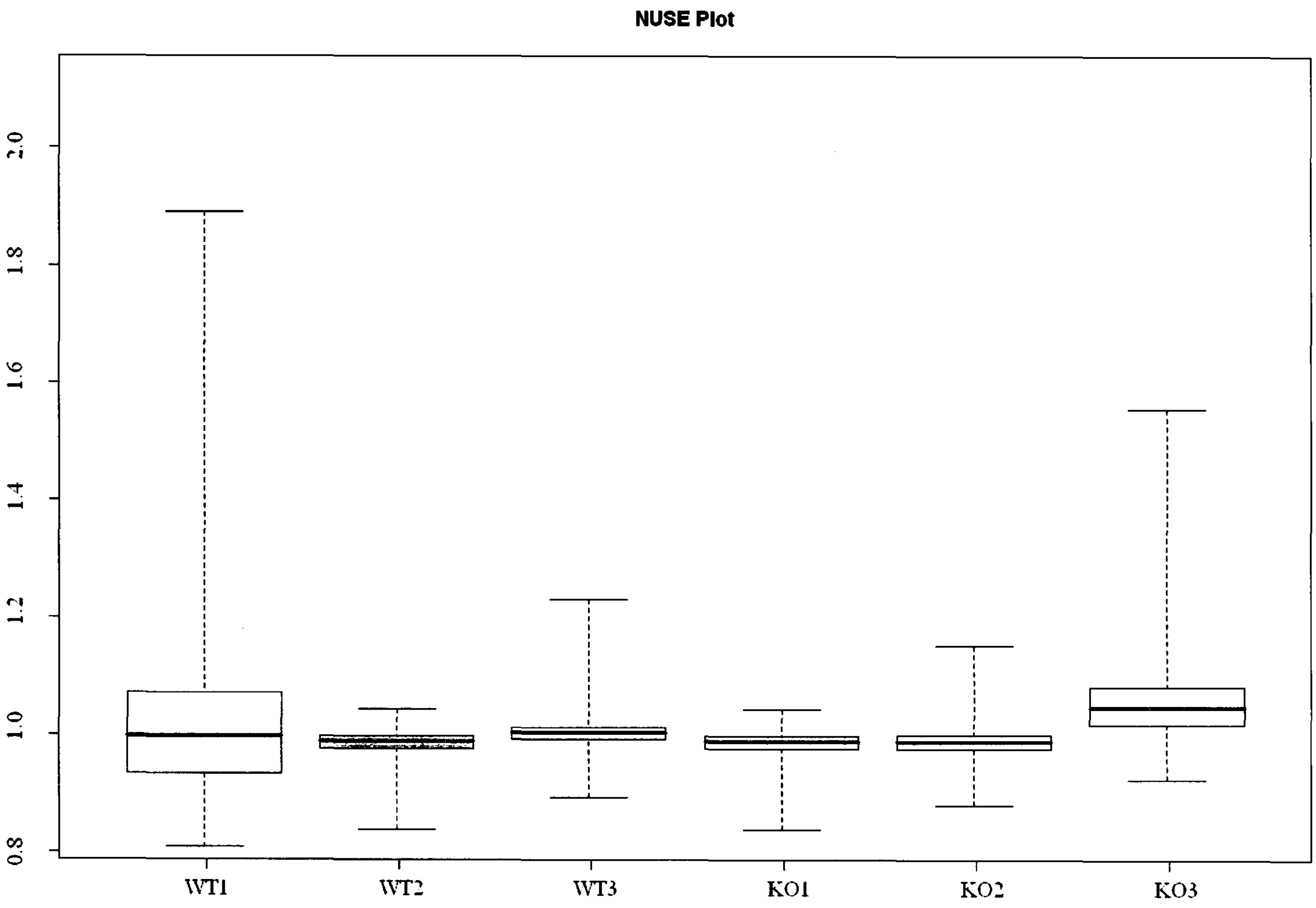

NUSE Plot
2.0
1.8
1.6
1.4
1.2
1.0
0.8
WT1
WT2
WT3
KO1
KO2
KO3

Figure 12: Microarray heatmap of differentially expressed genes.

Heatmap of differentially expressed genes in the three knockout conditions, denoted KO1, KO2 and KO3. Upregulated genes in a given condition are coloured green, while downregulated genes are coloured red. The relative fold expression between the conditions is visualized by the shade of green or red used, the brighter the shade, the larger the relative fold expression of the gene. A histogram and scale for the colour shading is provided in the upper left hand corner. Genes in blue rectangular boxes are related to an ethanol oxidation pathway regulated by *agmR*.

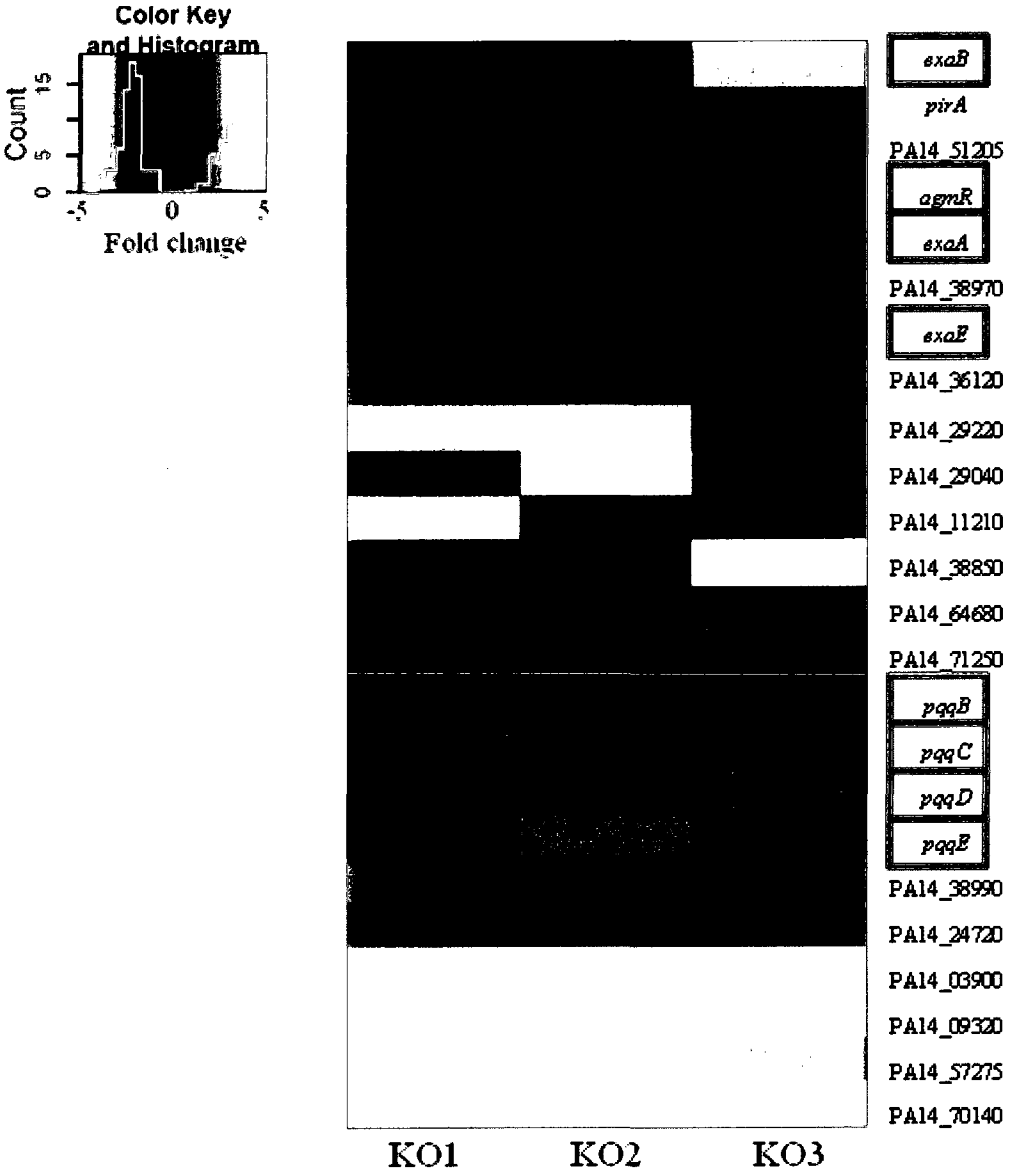

Color Key
and Histogram
Count
0
5
15
-5
0
5
Fold change
exaB
pirA
PA14_51205
agmR
exaA
PA14_38970
exaE
PA14_36120
PA14_29220
PA14_29040
PA14_11210
PA14_38850
PA14_64680
PA14_71250
pqqB
pqqC
pqqD
pqqE
PA14_38990
PA14_24720
PA14_03900
PA14_09320
PA14_57275
PA14_70140
KO1
KO2
KO3

Table 5: List of genes differentially regulated in WT and KO microarrays.

A list of genes differentially regulated between the wildtype and knockout *ndvB* conditions as generated by Bioconductor. Significance was set at $p \leq 0.000001$ to establish 'significantly different" along with a fold change of at least 2.

Gene	Fold Change	P-Value ($X10^{-5}$)	Function
PA01_ 1163	6.23	12.01	*ndvB*
PA14_03900	-2.36	4.48	Biotin biosynthetic process
PA14_38850	2.95	2.51	*exaB*
PA14_52230	2.13	2.98	*pirA*- Siderophore-iron transmembrane transporter activity
PA14_51205	2.10	6.51	Probable transcriptional regulator
PA14_38930	2.89	1.89	*agmR*
PA14_38860	3.75	5.67	*exaA*
PA14_38990	3.01	2.03	Hypothetical
PA14_38970	3.89	2.89	Two-component sensor
PA14_38900	3.36	3.37	Two-component response Reg.(*exaE*)
PA14_38820	3.75	3.86	*pqqB*
PA14_36120	2.99	1.33	Probable MFS transporter
PA14_29220	2.43	1.41	Porin *OpdB*-Small molecule transport
PA14_29040	2.10	1.04	Probable ferredoxin
PA14_24720	2.76	1.07	Probable two-component sensor
PA14_38780	2.82	1.11	*pqqE*
PA14_38800	2.11	1.09	*pqqC*
PA14_11210	2.19	7.48	Amino acid transporter
PA14_09320	-2.76	8.57	ATP-binding component of ABC transporter
PA14_57275	-2.30	5.14	Cell division protein *FtsZ*
PA14_38850	2.53	6.15	Hypothetical
PA14_64240	5.10	4.73	*pqqD*
PA14_64680	2.74	7.54	Hypothetical

six expressions was taken and statistical analysis of the means was done using a two tailed student t-test in Excel. Figure 13 shows the q-PCR reaction of *agmR* , the response regulator of the system. There is a clear increase in *agmR* expression in the biofilm Wt vs KO expression, that is statistically significant, and the *ndvB* KO condition returns to levels similar to planktonic expression. This confirms the pattern that is suggested by the microarray data, suggesting that *ndvB* plays a role in the regulation of *agmR* gene expression.

The genes *exaAB/B/C/D/E* were also checked for correlation with the array data via q-PCR through analysis of the representative genes *exaA/C/E.* Figure 14A-C shows the q-PCR results of *exaA/C/E* gene expression. *exaA/C/E* were chosen to represent *exaA/B/C/D/E* for several reasons. First, all of the genes studied have transposon insertion mutants available, while *exaB* and *exaD* do not, which is important for further experiments. In addition to this, *exaB* and *exaC* are co-transcribed while *exaD* and *exaE* are co-transcribed (36).

The q-PCR data for *exaA/C/E* show similar results to *agmR.* All of the genes show a statistical increase in biofilm gene expression when compared to planktonic gene expression. This increase in expression is abolished when *ndvB* is not present in the bacterial strains, indicating that the expression is biofilm specific and dependant on *ndvB.* Whether this is directly due to the fact that the glucans control the regulator of the system (*agmR*) or they directly affect the expression of these operons was not determined in this thesis.

Finally, genes belonging to the pyrolequinolonequinone (PQQ) operon were selected for confirmation of the array data. Figure 15A-E shows the results from the q-PCR of the pyrolequinolonequinone (PQQ) *operon. Figure 15A-D show the results of the PQQ genes in the operon regulated by agmR* and they are similar to both the *agmR* and *exaA/C/E* results. All of the genes in the operon share a promoter that is located in front of *pqqA* and is regulated by *agmR* (35, 36). All of the genes in the operon show significant increases in their biofilm expression in Wt conditions with return to planktonic levels in biofilms grown without the *ndvB* gene. *pqqF* is not in the same operon as the rest of the genes and has not been implicated as being under control of *agmR,* so it was of interest to see if this gene would follow the same pattern as the rest. Figure 15 E shows that the expression of *pqqF* does have an increase in expression in biofilms when compared to planktonic state, but this increase seems to be independent of *ndvB* as it occurs in both the Wt and KO *ndvB* conditions.

The correlation of the array data and the q-PCR data was graphed and an R^2 correlation coefficient was generated (Fig. 16). The microarray fold expression was plotted against the q-PCR fold expression for the eight genes studied and an R^2 correlation test was done on the data. A perfect positive correlation would be 1 and the R^2 test gave a correlation of 0.952, showing that the q-PCR data and the microarray data have a strong correlation.

3.4.0: Confirming antibiotic sensitivity of mutant genes

3.4.1 Confirming the presence of the transposon in the transposon insertion mutants.

The role of *exaA/C/E* and *pqqB/C/E/F* genes in antibiotic resistance was studied using transposon insertion (tn) mutants of these genes. The library of mutants have a single copy of the mariner Tn7 transposon (994 basepairs (bp) inserted in each gene resulting in a non-functional gene in most genes present in the *P. aeruginosa* PA14 genome (50)(as described in section 2.7.1). Since it is a large library, covering more then 5000 genes it could be prone to error or contamination. As such the transposon insertion in each gene was confirmed as well as some of the basic properties, such as biofilm formation and growth rates, of the tn mutants, were determined in order to perform antibiotic sensitivity assays on the mutant. The tn insertions were confirmed via polymerase chain reactions (PCR) of the Wt chromosomal DNA and chromosomal DNA from the tn mutant of interest. Figure 17 shows an example of this PCR run on a 1% weight/volume agarose gel. The image was produced by Koli Basu for her honours thesis. The first lane contains a 1 kb+ ladder, lanes 2-4 represent the non template control, *pqqB* wt and *pqqBtn* respectively. The inserted 994 bp transposon is clearly discernible in the PCR product of *pqqB*tn making this an adequate method for confirming the transposon insertion. Similar PCR reactions were performed using primers specific for each gene and Table 6 shows the expected length of the wildtype and tn mutant PCR products and the actual length observed on the gel. The increased size of the tn mutant PCR products confirms the presence of the insertion. In all cases the transposon insertion was found in the gene.

In addition to confirming the transposon insertion in the mutants, it was important to confirm that the mutants grew well and formed biofilms. Thus, I grew biofilms in six well plates using an air-liquid interface (ALI) assay and in 96 well plates to perform crystal violet assays (CV). The results of these assays can be seen in Figure 18. These images show that all of the mutants are able to form biofilms under standard assay conditions. In addition to the biofilm assays, growth curves of the mutants were done to establish if they grew the same as wildtype strain. Figure 19 shows these growth curves. The bacteria were grown in LB for 24 hours, the absorbance was taken at 600 nm ever half an hour for the first 8 hours and then again at 12 and 24 hour time points. There was no difference in the growth of any of the strains.

3.5.0: Antibiotic resistance of transposon mutants

In order to determine the sensitivity of the *exaA/C/E* and *pqqB/C/E/F* transposon mutants to antibiotics, I performed minimal bactericidal concentration (MBC) assays using antibiotics that are of clinical importance in *P. aeruginosa* infections. MBC assays allow for the direct comparison of the concentration of drug that it takes to kill bacteria in planktonic (P) and biofilm (B) states.

The MBC is defined as the concentration of antibiotic that kills bacteria and is represented by the concentration where no growth of bacteria occurs on the LB agar plate. As an example, Figures 20 A and B show the MBC-P to tobramycin for Wt, *ndvB* KO and *pqqC*tn, *exaA*tn, respectively and Figures 20 C and D show the MBC-B plates of Wt, *ndvB* KO and *pqqC*tn, *exaA*tn. Each experiment was done at least in triplicate.

Figure 13: q-PCR gene expression analysis of *agmR*.

The gene expression of *agmR* in biofilm and planktonic states was confirmed using q-PCR. RNA was extracted from planktonic cultures in stationary phase or 48 hour biofilms produced by Wt and *ndvB* KO strains. cDNA was synthesized using random primers from Bio-Rad script cDNA synthisase kit and q-PCR was performed with Bio-Rad SYBR-green mastermix. Statistical significance of the means was determined using the student t-test.

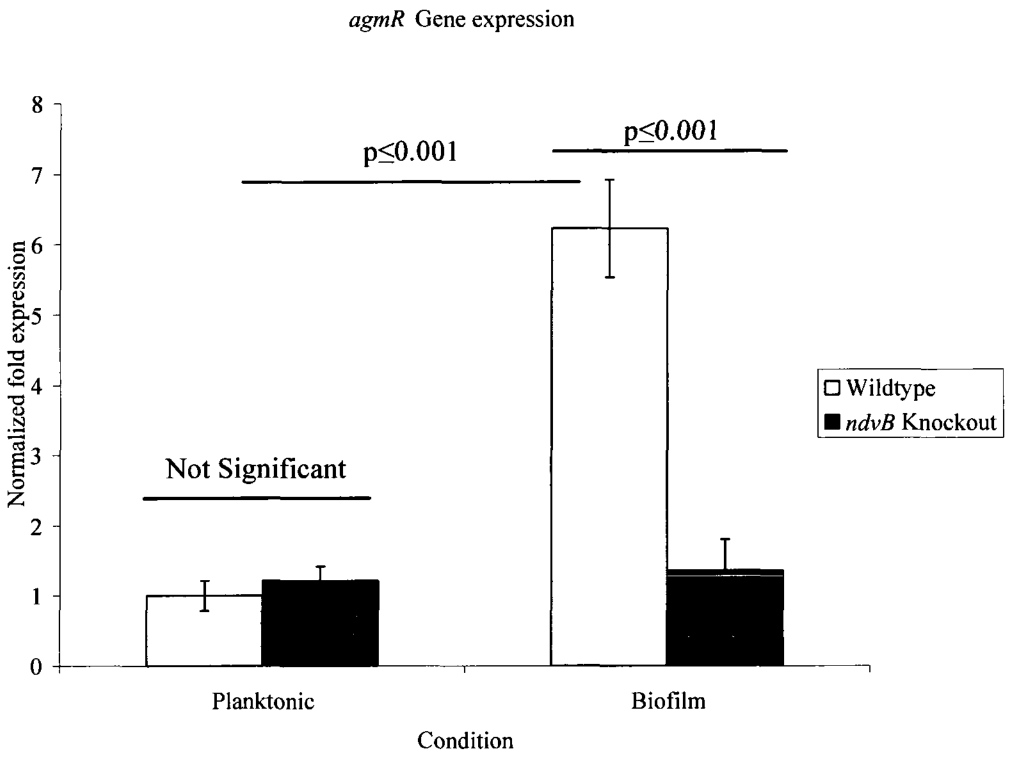
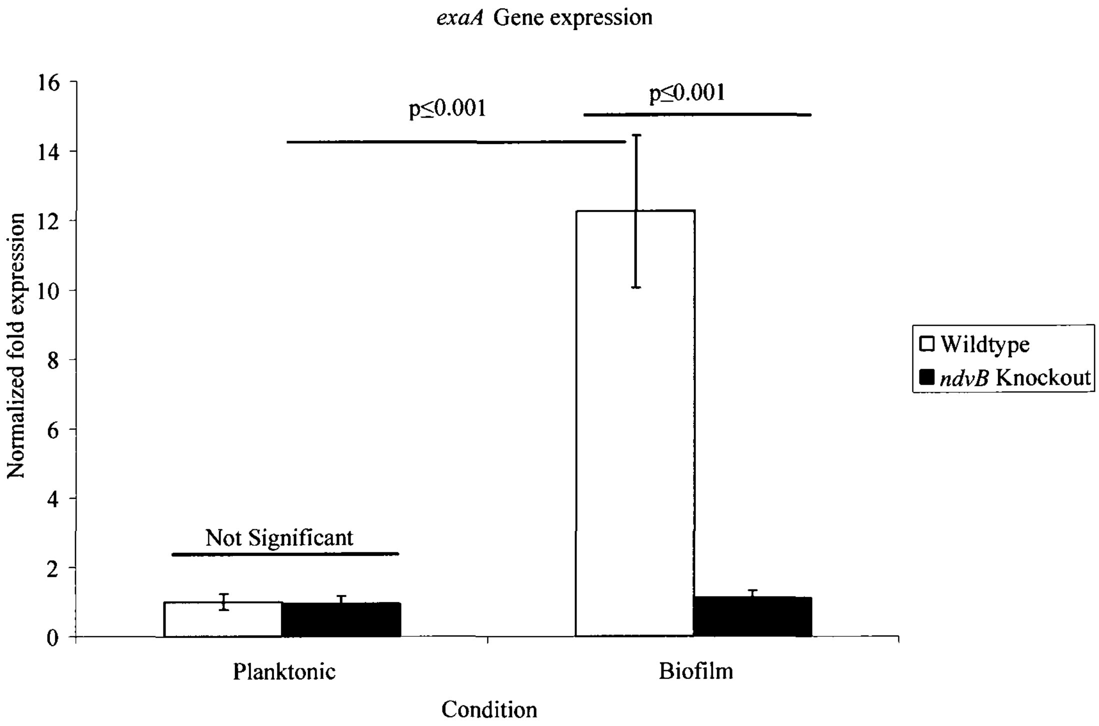

agmR Gene expression
8
7
6
5
4
3
2
1
0
Normalized fold expression
p≤0.001
p≤0.001
Not Significant
☐ Wildtype
■ ndvB Knockout
Planktonic
Biofilm
Condition

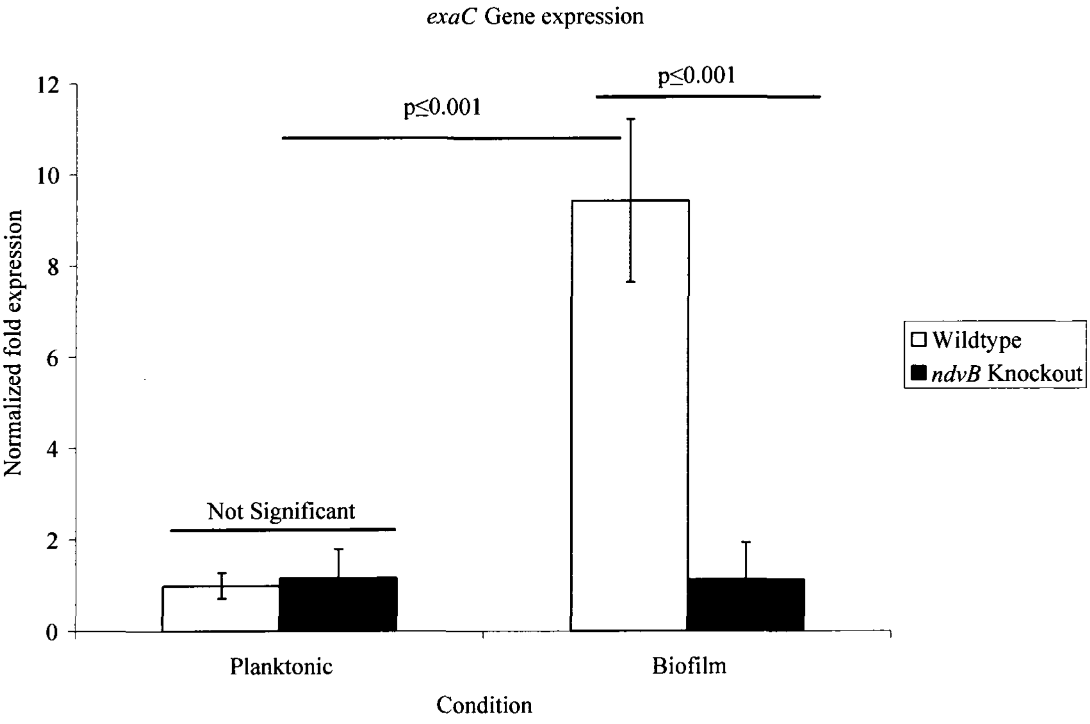
Figure 14: q-PCR of A) *exaA*, B) *exaC* and C) *exaE*.

The gene expression of *exaA, exaC* and *exaE* in biofilm and planktonic states was confirmed using q-PCR. RNA was extracted from planktonic cultures in stationary phase or 48 hour biofilms produced by Wt and *ndvB* KO strains. cDNA was synthesized using random primers from Bio-Rad script cDNA synthisase kit and q-PCR was performed with Bio-Rad SYBR-green mastermix. Statistical significance of the means was determined using the student t-test.

A


exaA Gene expression
p≤0.001
p≤0.001
Not Significant
Normalized fold expression
0
2
4
6
8
10
12
14
16
Planktonic
Biofilm
Condition
☐ Wildtype
■ ndvB Knockout

B


exaC Gene expression
p≤0.001
p≤0.001
Not Significant
Normalized fold expression
0
2
4
6
8
10
12
Planktonic
Biofilm
Condition
☐ Wildtype
■ ndvB Knockout

C

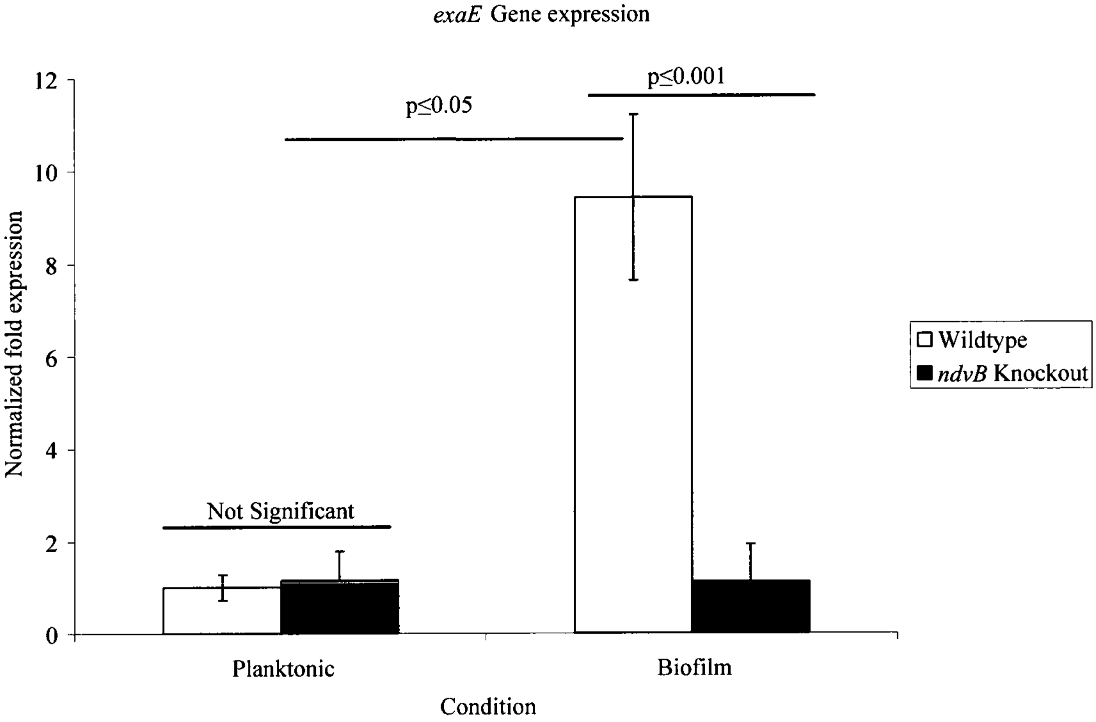
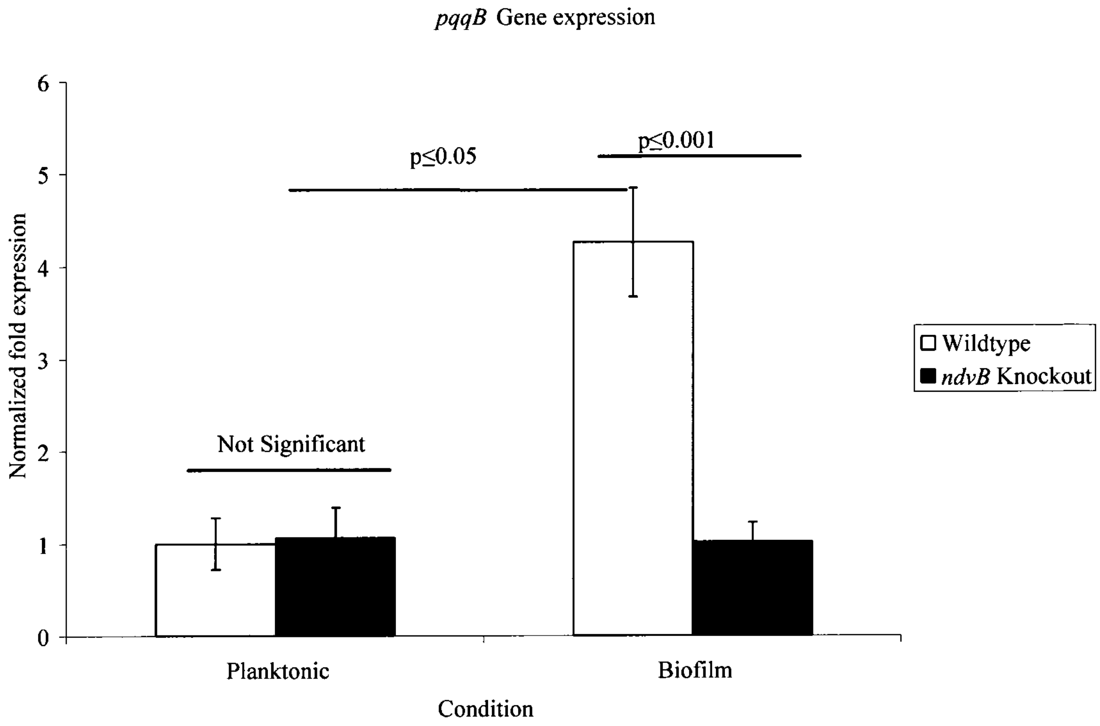

exaE Gene expression
Normalized fold expression
12
10
8
6
4
2
0
p≤0.05
p≤0.001
Not Significant
Planktonic
Biofilm
Condition
□ Wildtype
■ ndvB Knockout

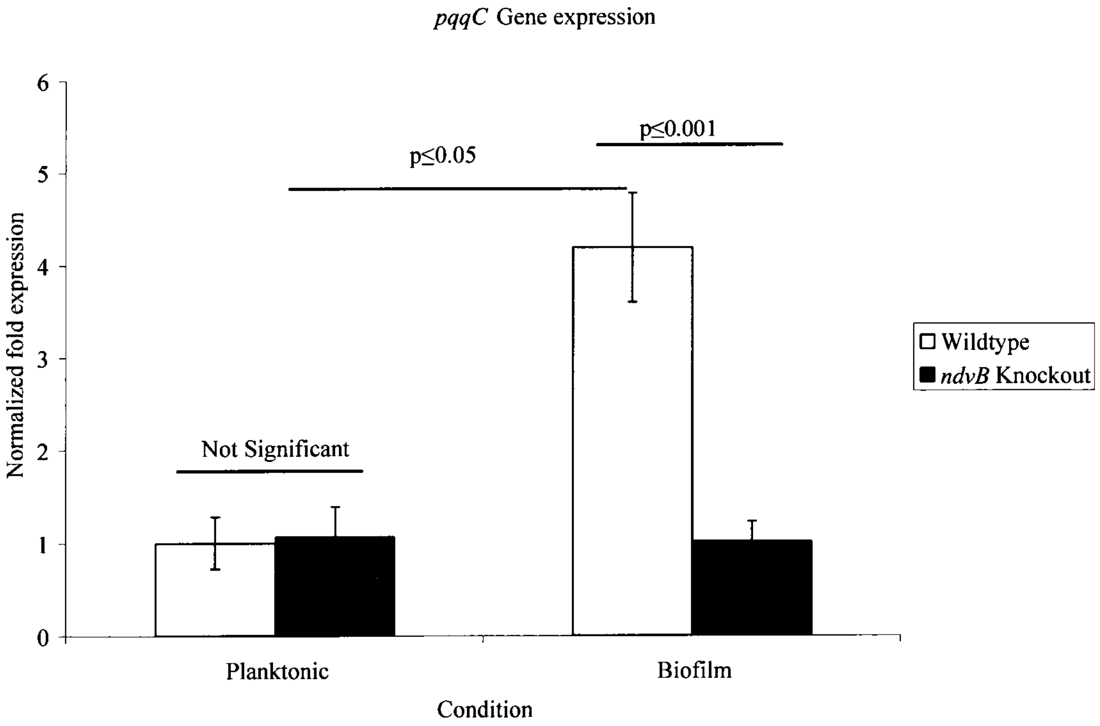
Figure 15: q-PCR of A) *pqqB* B) *pqqC* C) *pqqD* D) *pqqE* and E) *pqqF*.

The gene expression of *pqqB, pqqC, pqqD, pqqE* and *pqqF* in biofilm and planktonic states was confirmed using q-PCR. RNA was extracted from planktonic cultures in stationary phase or 48 hour biofilms produced by Wt and *ndvB* KO strains.. cDNA was synthesized using random primers from Bio-Rad script cDNA synthisase kit and q-PCR was performed with Bio-Rad SYBR-green mastermix. Statistical significance of the means was determined using the student t-test.

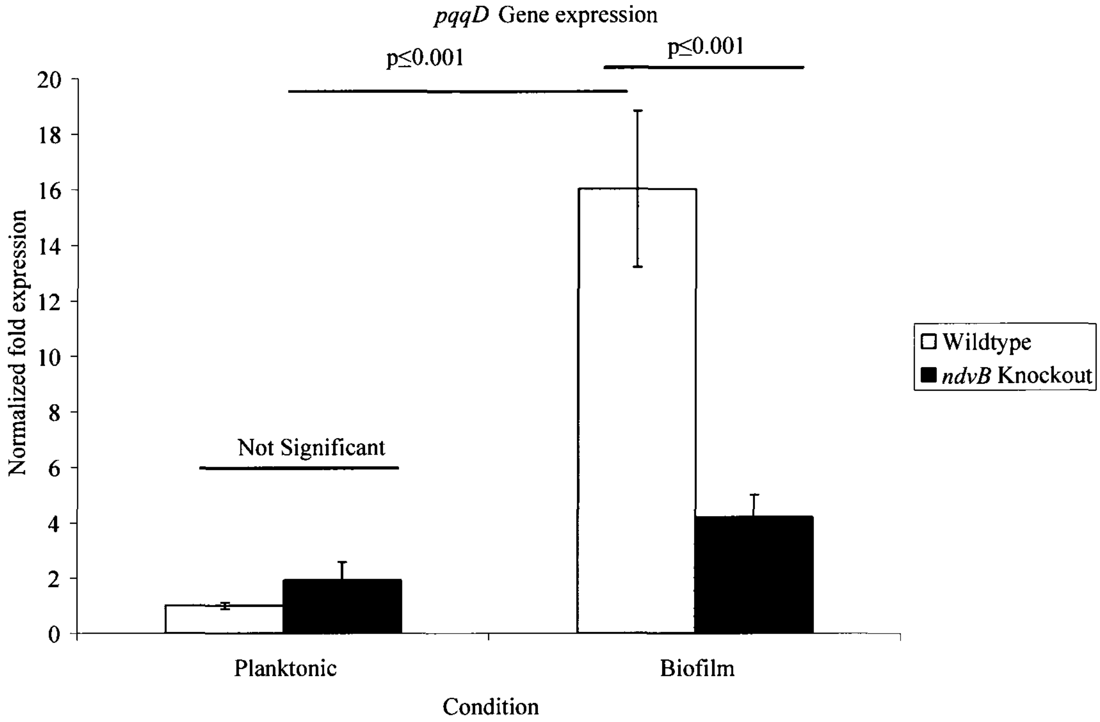
A


pqqB Gene expression
Normalized fold expression
0
1
2
3
4
5
6
p≤0.05
p≤0.001
Not Significant
Planktonic
Biofilm
Condition
□ Wildtype
■ ndvB Knockout

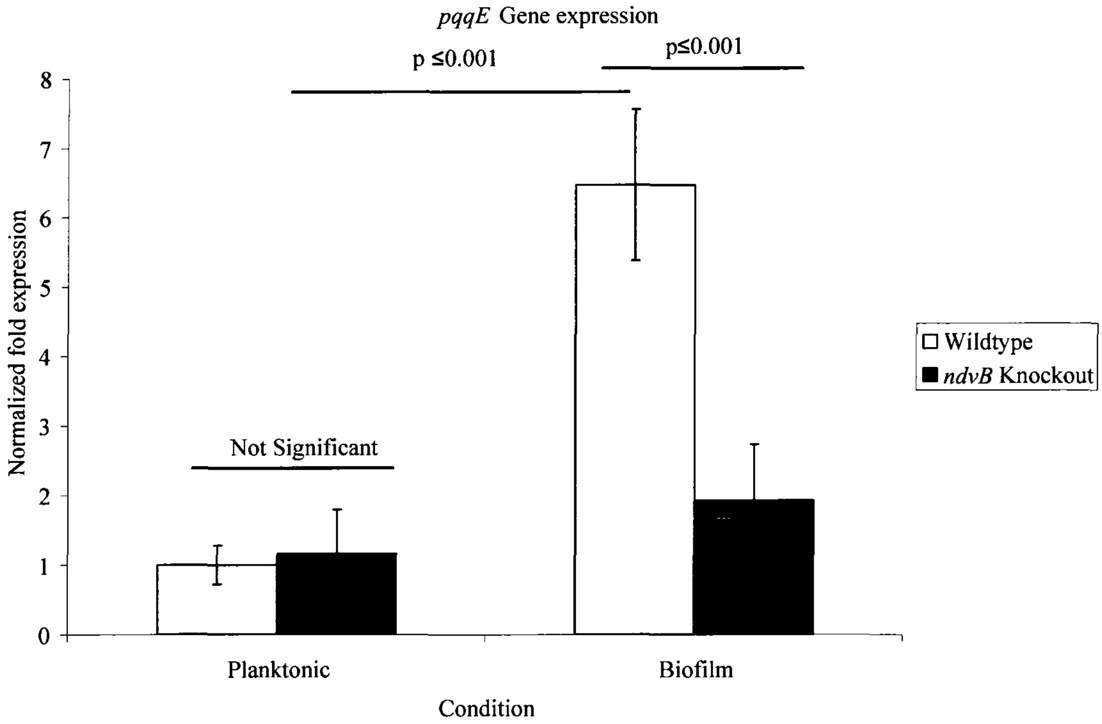
B


pqqC Gene expression
Normalized fold expression
0
1
2
3
4
5
6
p≤0.05
p≤0.001
Not Significant
Planktonic
Biofilm
Condition
□ Wildtype
■ ndvB Knockout

C


pqqD Gene expression
p≤0.001
p≤0.001
Not Significant
Normalized fold expression
0
2
4
6
8
10
12
14
16
18
20
Planktonic
Biofilm
Condition
☐ Wildtype
■ *ndvB* Knockout

D


pqqE Gene expression
p ≤0.001
p≤0.001
Not Significant
Normalized fold expression
0
1
2
3
4
5
6
7
8
Planktonic
Biofilm
Condition
☐ Wildtype
■ *ndvB* Knockout

E

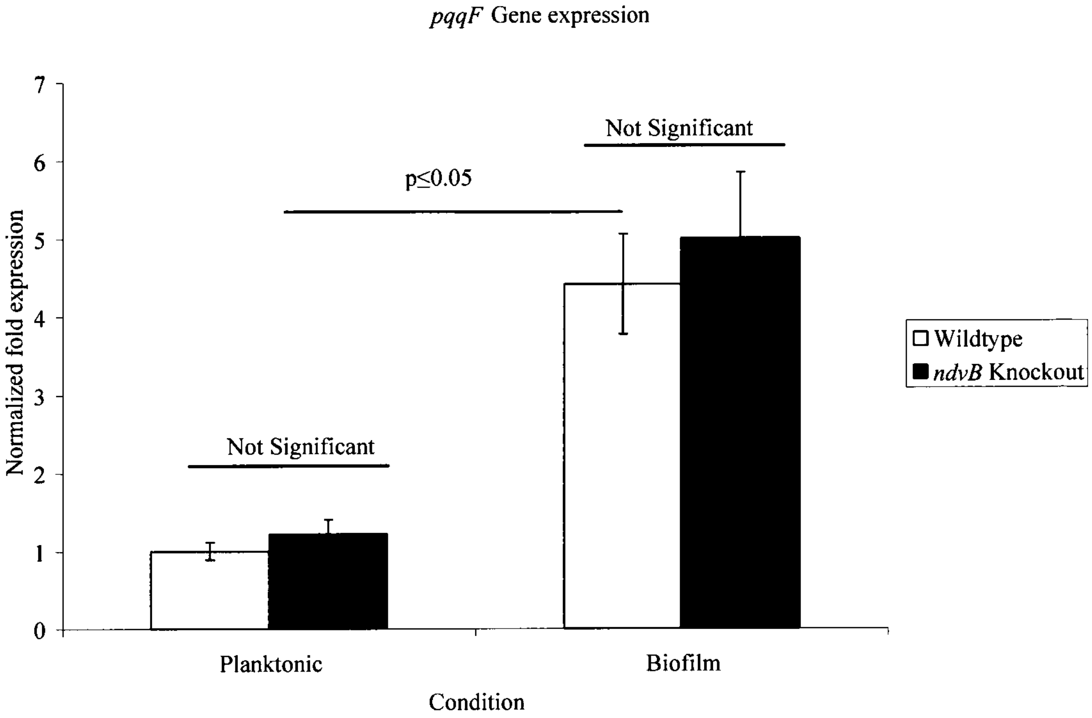

pqqF Gene expression
Normalized fold expression
0
1
2
3
4
5
6
7
Not Significant
p≤0.05
Not Significant
Planktonic
Biofilm
Condition
□ Wildtype
■ *ndvB* Knockout

Figure 16: Correlation between microarray data and q-PCR data.

Microarray gene expression and q-PCR gene expression of 8 genes from *Pseudomonas aeruginosa* and identified as significantly different through microarray analysis were plotted against each other. A R-squared test was run to determine if a correlation existed. A positive correlation of 0.952 was observed

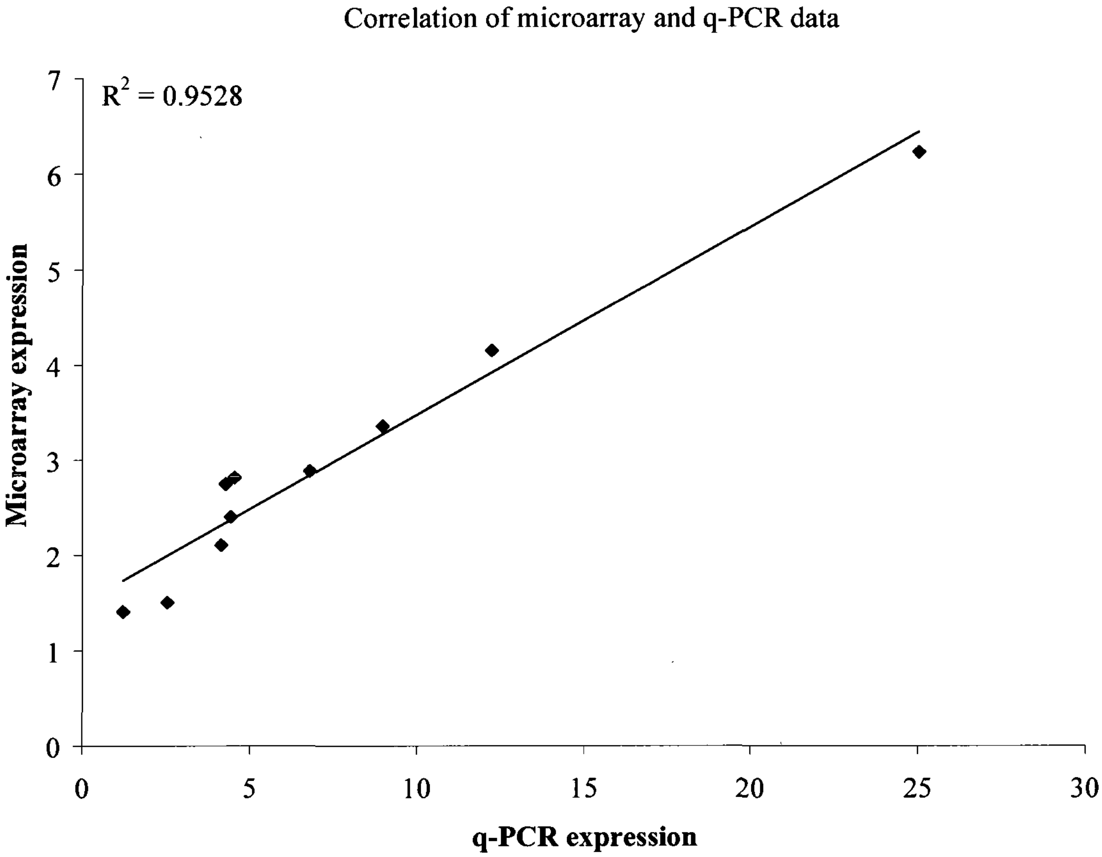

Correlation of microarray and q-PCR data

Table 7 shows the results of the MBC assays with tobramycin, an aminoglycoside, and ciprofloxacin, a fluoroquinolone, *exaA* and *pqqC* show the largest increase in susceptibility to tobramycin in the MBC-B, matching the level of the *ndvB* knockout strain. The other transposon mutants show increased sensitivity to tobramycin compared to the wildtype in the MBC-B assay, though not as much as *exaA* and *pqqC*. None of the mutants are as susceptible to ciprofloxacin as the *ndvB* knockout, and *pqqF* and *exaC* show no difference to the wildtype in the MBC-B assay. The other mutants showed a slight increase in sensitivity when compared to wildtype in the MBC-B assay. All of the MBC-P assay results were comparable for the wildtype and tn mutants for both ciprofloxacin and tobramycin.

Figure 17: Confirming transposon insertion via PCR

Transposon insertion mutants were confirmed via PCR of wildtype chromosomal DNA and chromosomal DNA extracted from the mutant strains. A 1% agarose gel electrophoresis of the full gene PCR products of the wildtype (WT) and *pqqBtn* strains of *pqqB* and wildtype and *exaAtn* strains of *exaA* is shown as a representative.. No template controls (NTC) for both sets of primers are also shown. The sizes of select components of the 1kb$^+$ ladder used as a marker are indicated on the right.

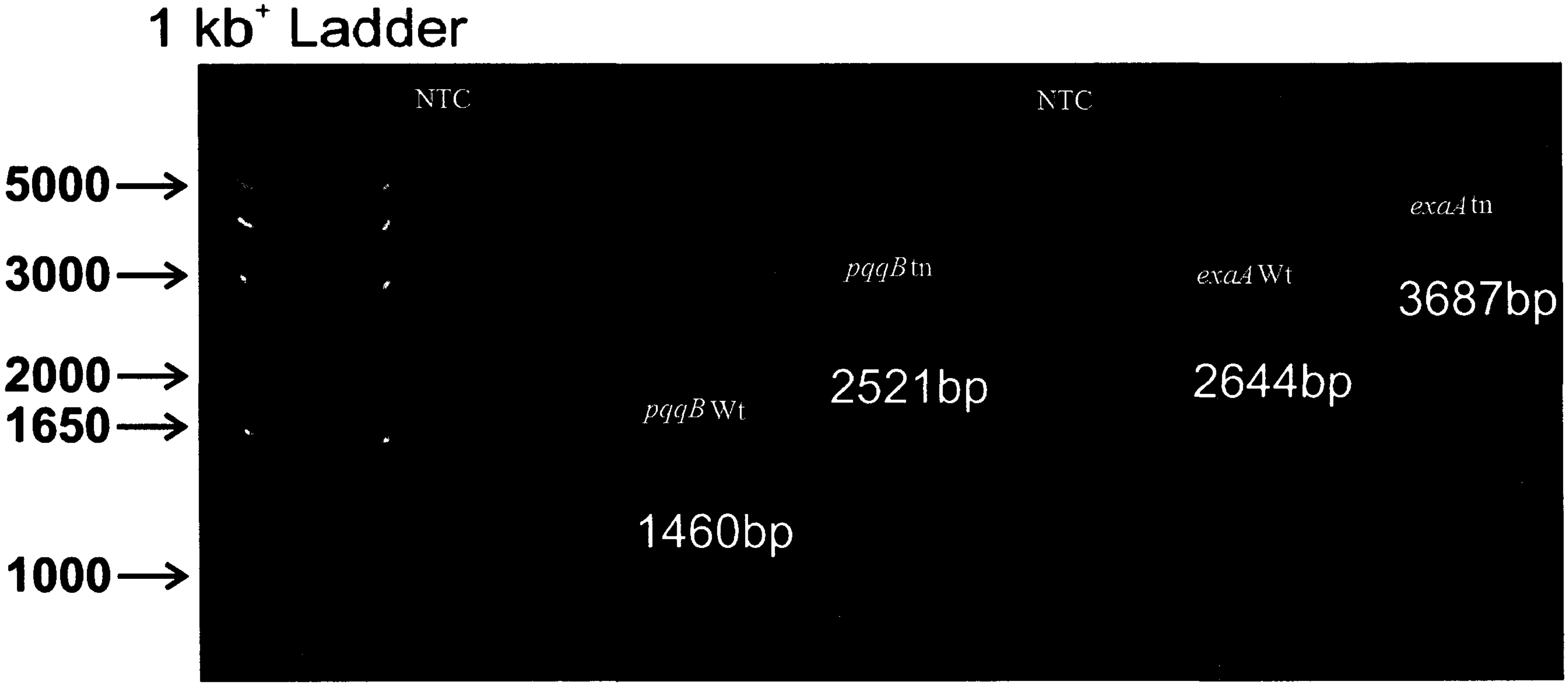

1 kb+ Ladder
NTC
NTC
5000
3000
2000
1650
1000
pqqB Wt
1460bp
pqqB tn
2521bp
exaA Wt
2644bp
exaA tn
3687bp

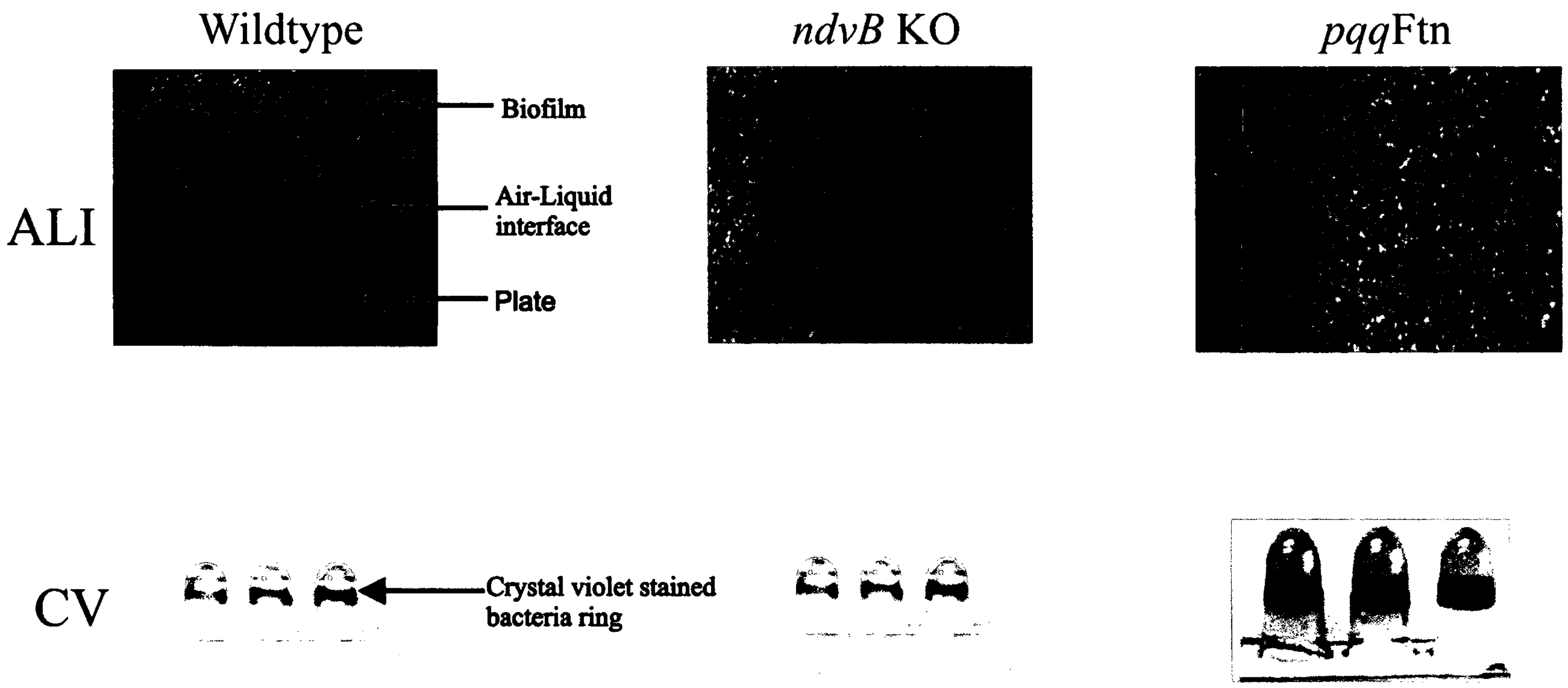
Table 6:Expected and actual band length of PCR products.

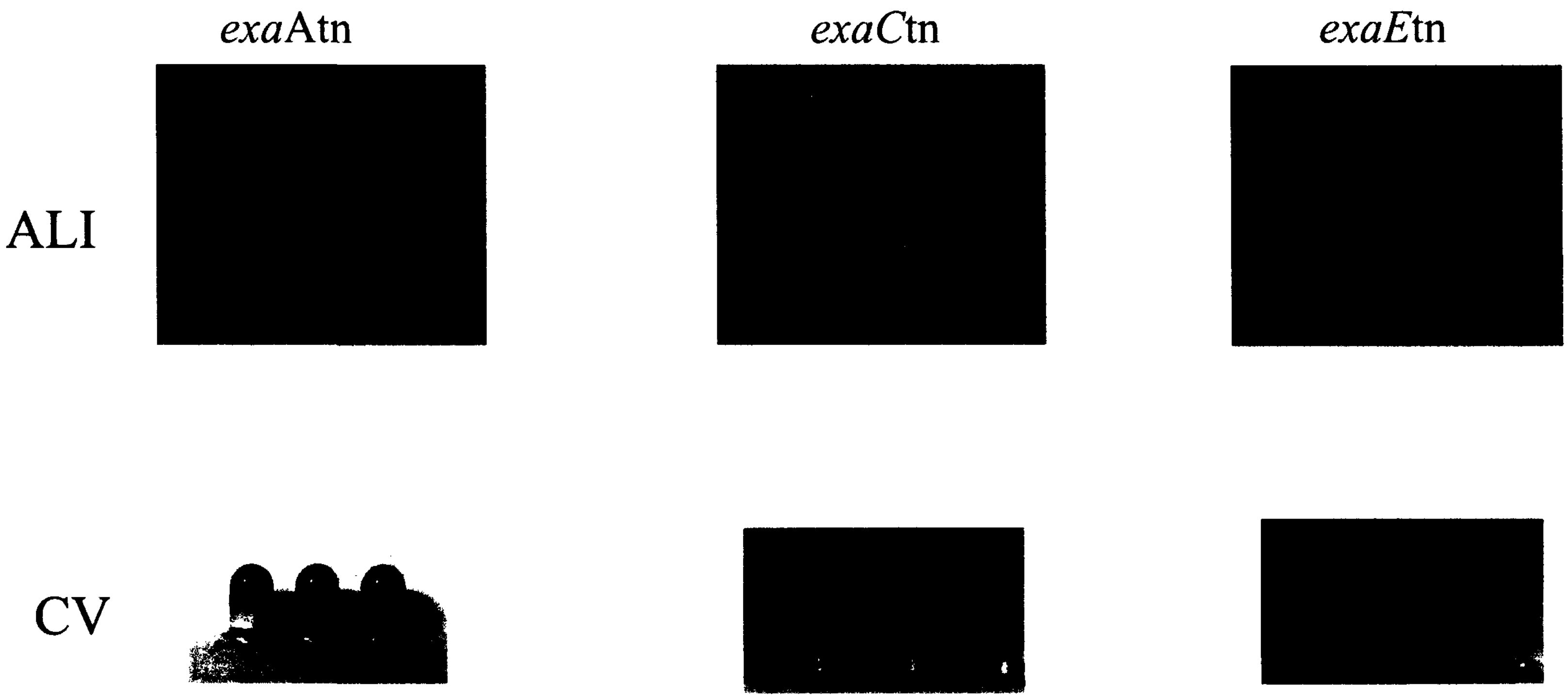
The expected length of the wildtype PCR product of the wildtype and transposon insertion mutant and the actual length of the PCR products that was visualized on a 1% agarose gel.

Gene	Wildtype		Transposon Insertion Mutant	
	Expected	Actual	Expected	Actual
pqqB	1390	1460	2386	2551
pqqC	1121	1050	2115	2000
pqqE	1340	1200	2336	2100
pqqF	2620	2300	2714	2600
exaA	2350	2521	3346	3687
exaC	1721	1800	2715	2400
exaE	1000	1100	1994	2000

Figure 18: Biofilm growth of mutants.

Transposon insertion mutants were studied for there ability to grow biofilms under standard laboratory conditions. Each mutant was grown on an angled 6 well microtitre plate for 24 hours and observed under the microscope in an air-liquid interface assay (ALI) experiment. Biofilm formation was also confirmed with a crystal violet (CV) assay. Wildtype and knockout *ndvB strains were used for controls. pqqB* transposon insertion (tn) (*pqqB*tn),*pqqC*tn, *pqqE*tn, *pqqF*tn, *exaA*tn, *exaC*tn and *exaE*tn were observed for their ability to form biofilms.


Wildtype
ndvB KO
pqqFtn
ALI
Biofilm
Air-Liquid interface
Plate
CV
Crystal violet stained bacteria ring


*exa*Atn
*exaC*tn
*exaE*tn
ALI
CV

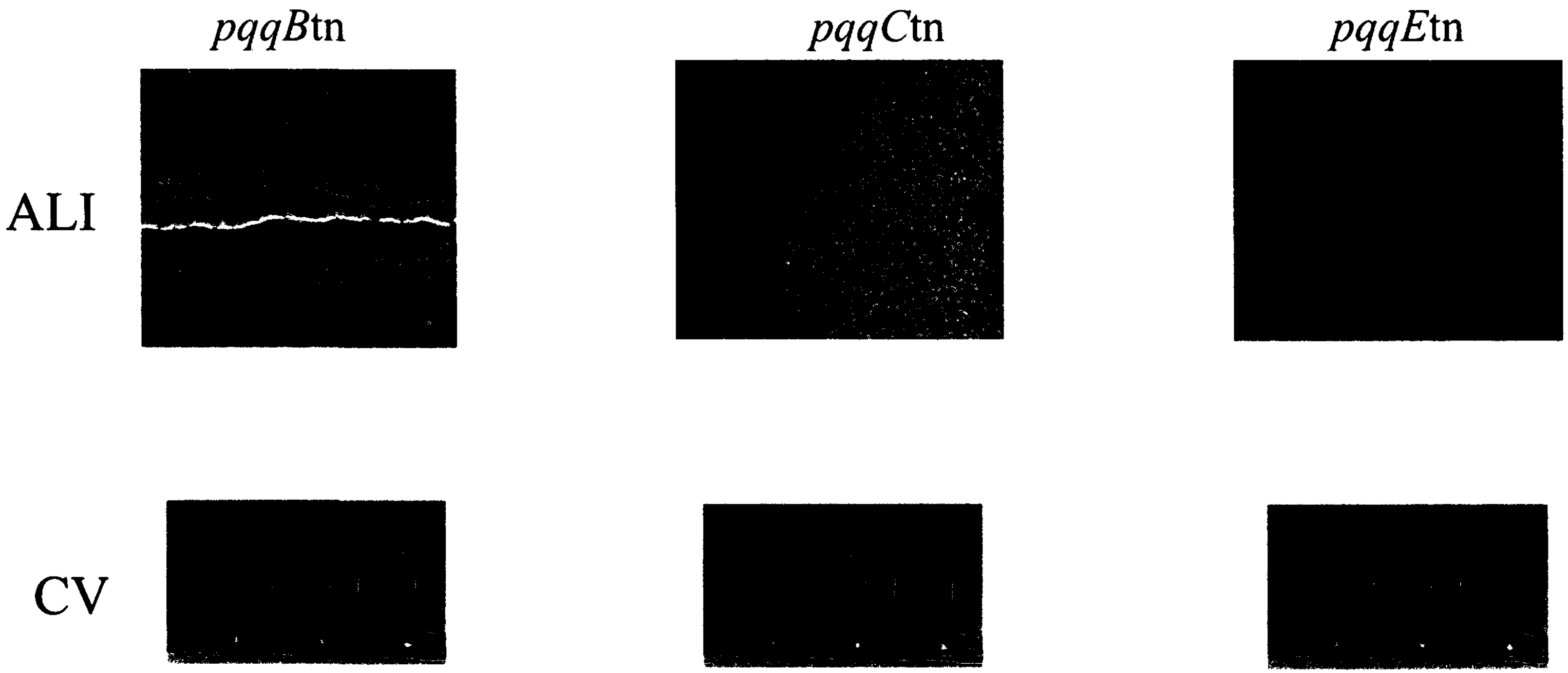

*pqqB*tn
*pqqC*tn
*pqqE*tn
ALI
CV

Figure 19: Growth curves of transposon mutants.

Growth curves of the transposon insertion mutants were performed to confirm viability of the mutants under standard conditions.
An overnight bacterial culture was diluted 1/1 000 000 in LB and allowed to grow for eight hours, with an absorbance being taken every half hour. A time point was also taken at 12 and 24 hours.

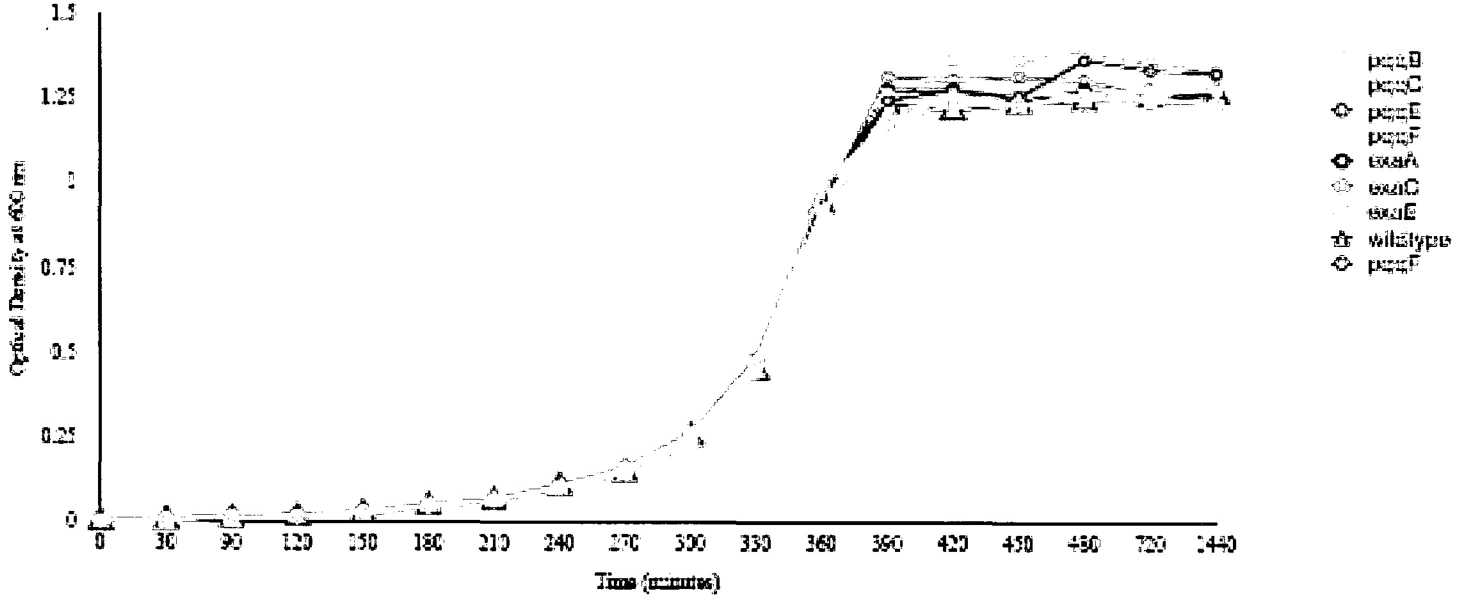

Optical Density at 600 nm
1.5
1.25
1
0.75
0.5
0.25
0
0
30
90
120
150
180
210
240
270
300
330
360
390
420
450
480
720
1440
Time (minutes)
pqqB
pqqC
pqqE
pqqF
exaA
exaC
exaE
wildtype
pqqF

Figure 20: MBC P and MBC B plates.

Minimal bactericidal concentration (MBC) of planktonic (P) and biofilm (B) to tobramycin. The plates shown are examples of plates from the MBCP/B assays performed for tobramycin. Bacterial growth occurs at the drug concentration insufficient to kill the bacteria.

A) MBC-P Wildtype and knockout *ndvB*
B) MBC-P *pqqC*tn and *exaA*tn
C) MBC-B Wildtype and knockout *ndvB*
D)MBC-B *pqqC*tn and *exaA*tn

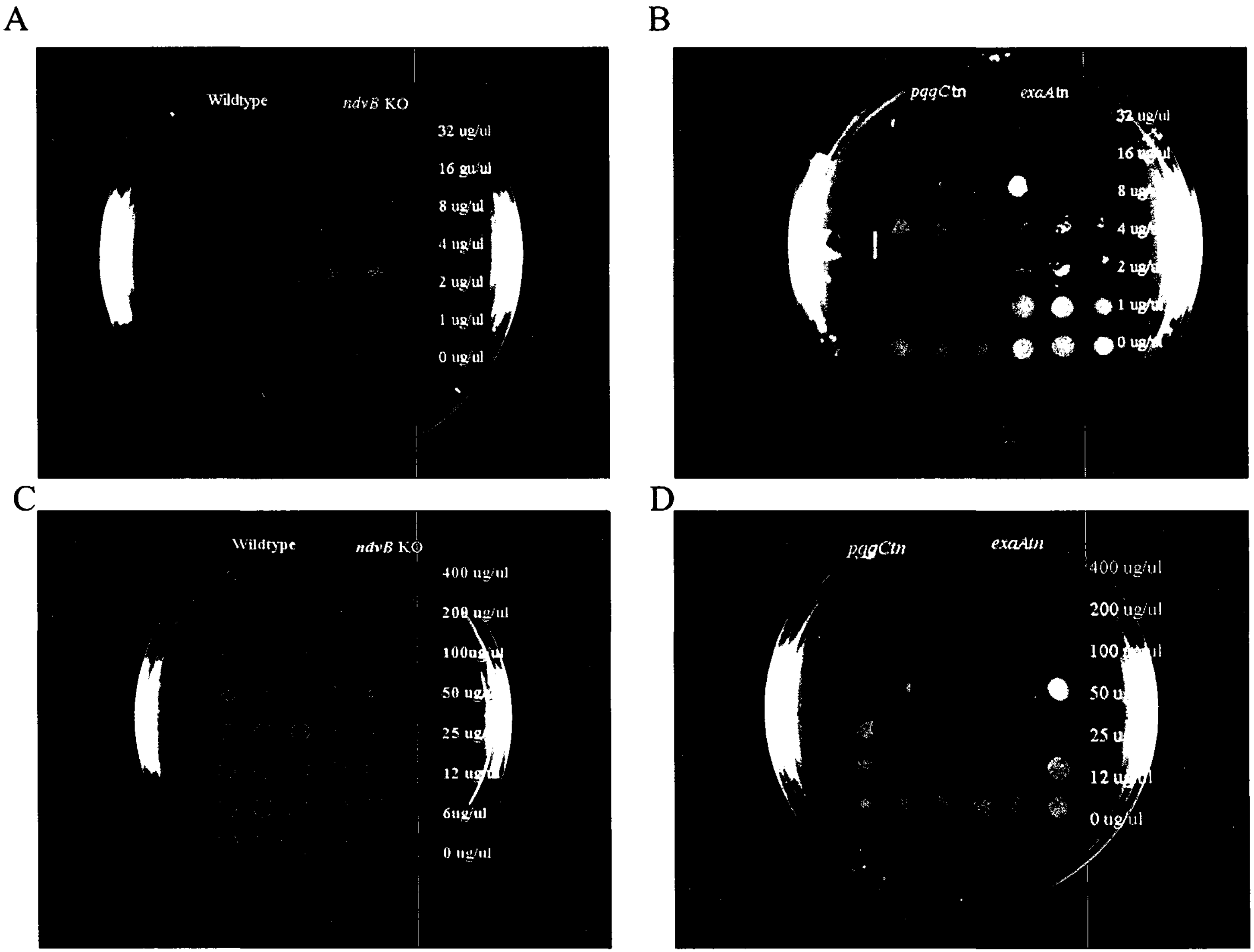

A
Wildtype
ndvB KO
32 ug/ul
16 gu/ul
8 ug/ul
4 ug/ul
2 ug/ul
1 ug/ul
0 ug/ul
B
pqqCtn
exaAtn
32 ug/ul
16 ug/ul
8 ug/ul
4 ug/ul
2 ug/ul
1 ug/ul
0 ug/ul
C
Wildtype
ndvB KO
400 ug/ul
200 ug/ul
100ug/ul
50 ug/ul
25 ug/ul
12 ug/ul
6ug/ul
0 ug/ul
D
pqqCtn
exaAtn
400 ug/ul
200 ug/ul
100 ug/ul
50 ug/ul
25 ug/ul
12 ug/ul
0 ug/ul

Table 7: MBC-P and B results.

Minimal bactericidal concentration (MBC) of planktonic (P) and biofilm (B) *P. aeruginosa* cells for various antibiotics. Isolates were grown overnight in LB media at 37°, diluted 1/50 into 1XM63 media supplemented with 1mM MgSO4 and 0.4 % arginine and transferred to 96 well microtitre plates. Tb=Tobramycin, Cip=Ciprofloxacin.

Strain	Tobramycin (ug/mL)		Ciprofloxacin (ug/mL)	
	Planktonic	Biofilm	Planktonic	Biofilm
Wildtype	16	400	2	25
ndvB KO	16	50	2	4
*pqqB*tn	16	200	2	12.5
*pqqC*tn	16	50	2	12.5
*pqqE*tn	16	200	2	12.5
*pqqF*tn	16	200	2	25
*exaA*tn	16	50	2	12.5
*exaC*tn	16	200	2	25
*exaE*tn	16	100	2	12.5

Discussion

Discussion

One of the hallmarks of biofilm growth is the dramatic increase in antibiotic resistance shown by bacteria within the biofilm. Many mechanisms have been proposed to account for this increased resistance and one of the most exciting is the expression of genes preferentially in biofilms that can lead to increased biofilm specific antibiotic resistance. For quite some time it has been known that biofilms express different genes than their planktonic counterparts (97) and it has been documented that some of these genes can led to antibiotic resistance (57). The regulation and control of these genes is not well known, but elucidating how these genes are expressed in a biofilm specific manner and what regulates them could led to better treatment of disease of biofilm based infections.

4.1.0: Microarray analysis

The primary experiments of this thesis are based upon the findings of differential gene expression between two *P. aeruginosa* PA14 strains, wildtype (Wt) and knockout (KO) *ndvB* PA14, grown as biofilms. This approach allowed me to look at the global regulation of genes in an *ndvB* dependant manner in order to determine if the cyclic glucans produced by NdvB can effect gene expression.

The results of the microarray analysis were dependent upon good experimental design, consistent experimental conditions and using appropriate measures for analysis and interpretation of results. Having quality RNA for microarray analysis was very important and the extraction of this RNA from biofilms can be an added difficulty. The results in figures 3,5,6 and 7 showed that this goal was achieved.

The analysis of Affymetrix Genechips preprocessed chips (before normalization and standardization) allows for identification of poor quality or problem chips. Figure 8 and 9 confirmed that the data, from the chip point of view, were acceptable. Having good quality array data increases the confidence that the genes identified from the array are truly differentially expressed.

The final step in the microarray analysis was just as crucial as the first step. Choosing methods to normalize and standardize the data on the chips, compare the means and generate the list of differentially expressed genes can make all the difference in generating a true list that one can be confident reflects the true nature of things. The advantage of using Affymetrix Genechips is that there is a wide array of tools available for the analysis of data. I used R statistical packages developed specifically for Affymetrix microarrays (37). This allowed for accurate normalization and standardization of data. Using two standard background normalization algorithms and taking the overlap of genes expressed (RMA and MAS 5 (13, 43, 53) as well as moderated t- tests [Bayesian, Holms Stepwise correction, (26)] and a stringent p-value ($p \leq 0.000001$) allowed me to be confident that these genes are actually differentially expressed in these conditions. Of course using these methods, it could be possible that some genes which actually are differentially expressed between these conditions have been left out, but because an adequate list of genes (Table 5, Figure 12) I felt these measures were reasonable.

4.2.0: Transcriptional regulator and its targets identified

The hypothesis of this study was that *ndvB* produces cyclic glucans that may play a role in signaling in *P. aeruginosa*. Although the exact mechanisms of this signal have yet to be elucidated, I have identified 24 genes that were differentially regulated between the wildtype and knockout *ndvB* biofilm conditions. A transcriptional regulator, known as *agmR*, as well as 6 genes that it regulates were all found to be upregulated in the wildtype PA14 when compared to the knockout condition, confirming the array experiments.

4.2.1 AgmR

AgmR is a response regulator important in the control of the quinoprotein ethanol dehydrogenase system in *P. aeruginosa* (35). AgmR (PA1978) that was initally identified as a glycerol metabolism activator which acts as an environmentally responsive gene directly involved in glycerol metabolism (84). Recent studies have identified *agmR* as the response regulator controlling genes involved in ethanol oxidation pathway in *P. aeruginosa* (99). Together with the genes that AgmR regulates, AgmR forms a cluster of proteins important for ethanol metabolism in *P. aeruginosa* through a series of oxidation reductions that occur in the periplasm. AgmR regulates the expression of three operons that are essential for *P. aeruginosa* growth on ethanol: the *exaBC* operon, the *exaDE* operon and the *pqqABCDE* operon (82, 84). Figure 20 represents a schematic breakdown of AgmR control, and Figure 21 depicts the location of these genes on the *P. aeruginosa* genome.

The identification of *agmR* as being responsive to presence of *ndvB*, and the fact that its regulated genes were also identified in the microarray made it an ideal candidate for further study. The analysis of *agmR* gene expression, as seen in Figure 13, showed that it was upregulated in biofilms in an *ndvB* dependant fashion. Since it is a regulator of the other genes, I would expect them to show a similar profile to this (See below). Unfortunately, no mutant is yet available for PA14 *agmR* and due to time constraints I was unable to produce a knockout of *agmR* in PA14, but this would allow for further phenotypic study of the antibiotic resistance provided by this gene.

4.2.2: *exaABCDE*

exaB (PA1983) encodes a soluble cytochrome c550 found in the periplasmic space of *P. aeruginosa* and is co-transcribed with *exaC* (PA1984) which encodes an NAD+-dependent acetaldehyde dehydrogenase. *exaBC* is situated in an operon that is controlled by the promoter for *exaB* (82). *exaB* codes for a 15 kDa soluble c_{550} cytochrome that is responsible for the transfer of electrons between QEDH and ferrocyinide during the reduction of ethanol while *exaC* codes for NAD+ acetaldehyde dehydrogenase, which reduces the alcohol into an aldehyde during metabolism (83). The gene expression results for *exaC* from Figure 14B show a similar trend as *agmR*, confirming the array data and indicating that *ndvB* plays a role in the biofilm specific expression of these two genes.

Furthermore, *exaC*tn mutant showed increased sensitivity to tobramycin and ciprofloxacin (Table 7). Both of these drugs are important for the treatment of *P. aeruginosa* infection in cystic fibrosis patients (25). Tobramycin is an aminoglycoside and ciprofloxacin is a fluoroquinolone, a completely different class of antibiotic, suggesting that the method of protection conferred by these genes, or the system, provides multidrug resistance (resistance to two or more classes of drugs).

exaA (PA1982) codes for a quinone-dependent ethanol dehydrogenase (QEDH) that is under the control of a two-component regulatory system, encoded by *exaDE*, (PA1979 and PA1980 respectively) (83)

QEDH, along with a pyrolequinolonequinone (PQQ), are responsible for the metabolism of glycerol in the periplasam of *P. aeruginos*a. *exaD* encodes a protein of 272 amino acids, which bears similarities to the C-terminal domains of sensor kinases of the histidine type, while *exaE* encodes a protein of 225 amino acids containing a histidine receiving module motif (77). This system, along with PQQ, allows for *P. aeruginosa* to grow on ethanol as it sole carbon source, expanding the type of environments in which it can survive.

The q-PCR results (Figure 14 A and C) for *exaA* and *exaE* were predictably similar to that of *agmR,* and as expected the q-PCR confirmed that both of these genes are expressed preferentially in biofilms that are able to produce cyclic glucans *(ndvB+).* Whether this q-PCR profile is a result of *agmR* being downregulated in the absence of cyclic glucans or the fact that cyclic glucans directly affect the expression of this system has not yet been established. Further experimentation on how *ndvB* is controlling the gene expression is required.

Figure 21: Regulation scheme of AgmR.

The regulation of genes under the control of AgmR in *P. aeruginosa.* (35) AgmR has been identified as a regulator of genes important for glycerol metabolisim.

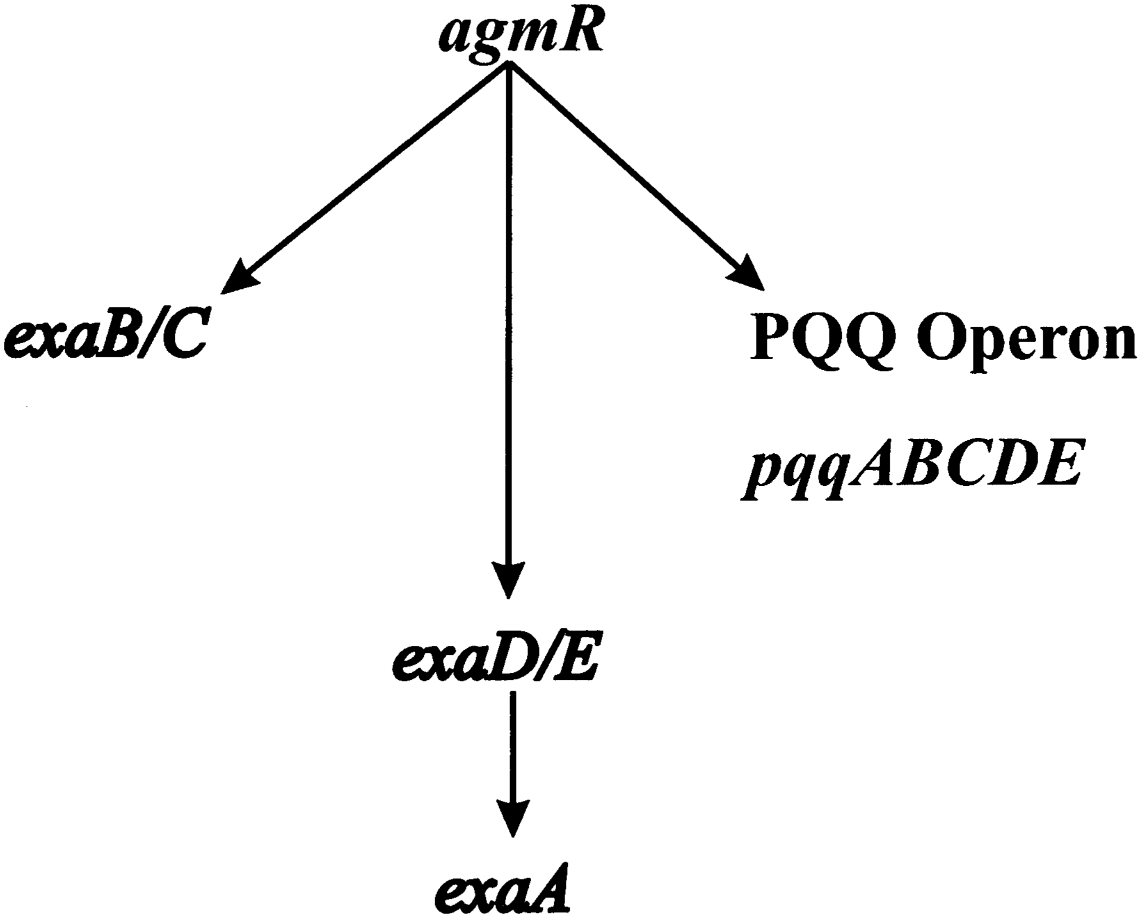
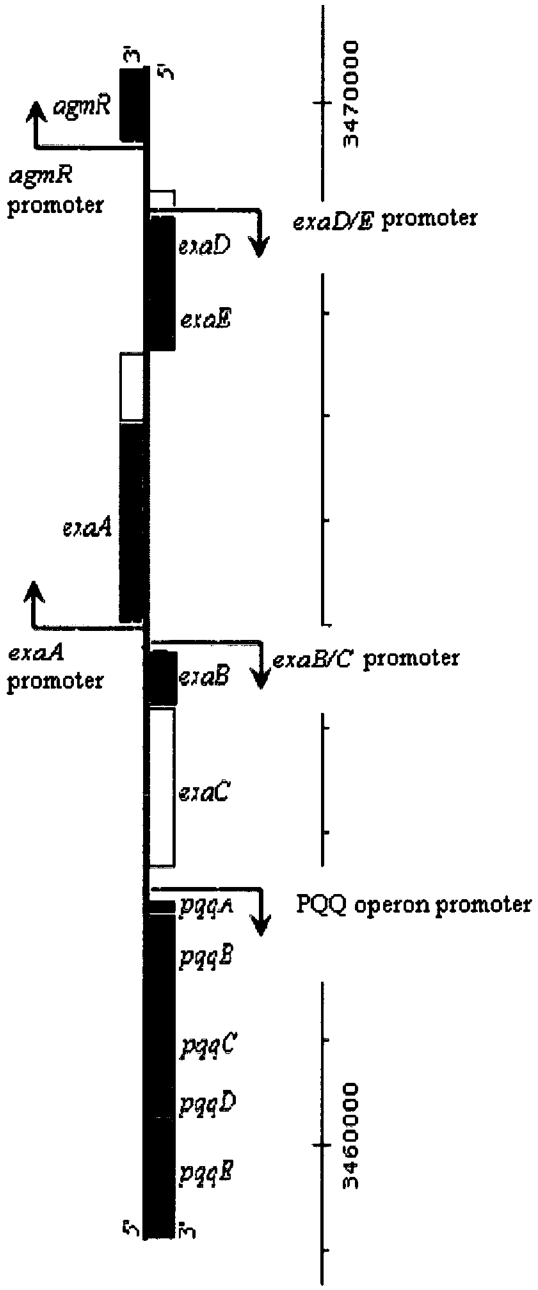

agmR
exaB/C
PQQ Operon
pqqABCDE
exaD/E
exaA

Figure 22: The location of agmR and the genes it regulates on the PA14 genome.

The location of *agmR* along with *pqqABCDE* and *exaABCDE* in the PA14 *Pseudomonas aeruginosa* genome. The scale and location of the genes in the PA14 genome is given along the bottom. The 5' end of the genes is on the left hand side of the page. Genes above the line are oriented in a 5'-3' direction and genes below the line are oriented in a 3' to 5'. Genes in blue were found to be significantly up-regulated in the microarray and the black genes are the other members of the ethanol oxidation pathway.


3'
5'
agmR
agmR promoter
exaD/E promoter
exaD
exaE
exaA
exaA promoter
exaB/C promoter
exaB
exaC
PQQ operon promoter
pqqA
pqqB
pqqC
pqqD
pqqE
5'
3'
3470000
3460000

MBC-B assays showed that *exaA* has an increased sensitivity towards tobramycin when grown in the biofilm state (Table 7), as well as ciprofloxacin, at levels that also kill the *ndvB* knockout mutant (57). Although most of the other genes involved in this system showed some decrease in resistance, the fact that *exaA* showed the same profile as the *ndvB* knockout is interesting. Because *exaA* encodes for the electron transporter (QEDH) of the system, not having it present could leave *P. aeruginosa* more susceptible to reactive oxygen species. The mechanisms for how it is protecting the bacteria from these drugs, while not trivial, would be an interesting project to explore. *exaC* also showed reduced resistance to the antibiotics when grown in a biofilm state suggesting it might be part of a multifactorial mode of biofilm specific antibiotic resistance.

4.2.3: *pqqABCDEF* and PQQ

As stated in the previous section, pyrroloquinoline quinone (PQQ) is the active part of the quinoprotein dehydrogenase that allows transfer of electrons between the cytochrome and the substrate (alcohol) during its metabolism (72). PQQ is present in the periplasm of a number of Gram-negative bacteria including *Klebsiella pneumoniae*, *Gluconobacter* sp. and *P. aeruginosa* (72). In addition to its role in ethanol and metabolism of alcohols, PQQ is important in the reactions of free radicals found in bacteria (62, 84).

The synthesis of PQQ in *P. aeruginosa* involves 6 genes, 5 of which are in the operon *pqqABCDE,* and *pqqF,* located 14 kilobases upstream of the this operon. The *pqqABCDE* operon has been found to have a single promoter associated with *pqqA* (82). The protein PqqA is a relatively small peptide of about 20-40 peptides in length that has a conserved Glu-X-X-X-Tyr motif that acts as a precursor for PQQ (36) .

PqqB is a 300 amino acid peptide that is not essential for the production of PQQ but rather is required for its movement from the cytosol to the periplasm and may be required for the release of PQQ from PqqC in the final step of its formation (55). In addition to these features, PqqB has a similar structure to metallo-β-lactamases which can inactivate antibiotics in the β-lactam superfamily II (99).

PqqC is a 250 amino acid peptide that is important in the final step of forming PQQ, being an oxidase that does not require a cofactor (4). Because of this, PqqC can readily form oxygen radicals, which would be damaging to the host cell, and so it requires a transporter to remain inactive (thought to be PqqB) (4).

PqqD is a small peptide with an unknown function. PqqE is a 300 amino acid protein important in C-C bond formation during the formation of PQQ via linkage of glutamate and tyrosine. PqqF is the largest protein of the group at 760 residues and is the only one found outside of the operon, not under the control of *agmR*. It is a metalloendopeptidase that is involved in the processing of glutamate and tyrosine of PqqA (72).

Together, *pqqABCDE* and *pqqF* allow for the production of PQQ from glutamate and tyrosine precursors on PqqA and facilitate its transport to the periplasm where it is free to interact with QEDH and metabolize ethanol (72). Figure 15A-E shows the q-PCR of *pqqBCDE* and *pqqF*. The gene expression of all of the genes within the operon and under control of *agmR* showed that they are preferentially expressed in biofilms producing cyclic glucans via *ndvB*. This result confirmed the microarray data and would seem to fit the profile that these genes are regulated by *agmR* in a manner dependent on *ndvB* gene expression.

pqqF was not found to be differentially expressed in the microarray and is not under the control of *agmR* but I was interested to see if it would be regulated in a similar manner. The q-PCR (Figure 15 E) results for *pqqF* showed that there was a significant increase in the gene expression between the planktonic and biofilm states, but this does not seem to be linked to *ndvB* as no real difference was found between the Wt and KO *ndvB*. This leads one to wonder what is controlling the expression of this gene in a biofilm specific manner. This question, while interesting, was not the focus of the current work.

4.2.4: PQQ and antibiotic resistance.

Although the primary role of PQQ is in ethanol metabolism, it has also been shown to be a strong free radical scavenger. PQQ has been found in bacteria lacking the dehydrogenases that require PQQ as a cofactor and so the role of PQQ in these organisms was found to be protective against free radicals and reactive oxygen species (62). Furthermore, expression of PQQ in *Escherichia coli* which typically lacks PQQ has led to the observation that *E. coli* expressing this factor has increased survival when faced with oxidative stress through scavenging of free radicals (62). In conjunction with this, it has been proposed that bactericidal drugs induce cellular death through the production of free hydroxyl radicals. This study's findings (48) indicate that, independent of the drug-target interaction of particular antibiotics, bactericidal drugs end up killing the bacteria through the production of, and accumulation of hydroxyl radicals. Although no studies have linked the production of hydroxyl radical during antibiotic treatment with the protection that might be provided by PQQ against such effects, it is an intriguing aspect that could be investigated in the future.

The transposon insertion mutants of the PQQ genes were more sensitive to antibiotic treatment (Table 7), particularly with tobramycin where the mutants showed a 4 to 8-fold increase in sensitivity. The fact that the absence of any one of these genes leads to a reduction of the antibiotic resistance seen in biofilms seems to indicate that the final PQQ product maybe important for conferring protection, however at this point I cannot rule out the fact that individual components, as well as the final PQQ product maybe important for conferring resistance to antibiotics. By producing double and triple knockouts of genes within the operon, a clearer picture will be obtained in the future.

Although no such studies have implicated a role for QEDH (*exaA*) or any of the genes in that system with increased sensitivity to antibiotics or free radical scavenging, the fact that QEDH is an electron acceptor in a similar manner as PQQ, suggests that the resistance mechanism might be similar to the one provided by PQQ.

4.3.0: Implications

The experiments I performed for this thesis strongly suggest that cyclic glucans, produced via *ndvB* play a role in biofilm specific gene expression. Furthermore, I have demonstrated that a subset of the genes which are regulated by these cyclic glucans have a role in biofilm specific resistance. Antibiotic resistance is of major concern in hospital environments and to patients infected with these bacteria, minor infections suddenly can become serious problems. This is compounded in patients that are immune compromised, surgical patients, or patients susceptible to infection, such as patients with cystic fibrosis. While the acquisition of mutation or planktonic mechanisms have been a major are of study, recent evidence suggest that biofilms confer antibiotic resistance and contribute to a significant number of infections, up to 65% of all hospital infections in fact (39). While the physical properties of biofilms cannot be ignored nor the population of persister cells, it is clear that genes intrinsically expressed in a biofilm specific manner play a role to the antibiotic resistance profile of biofilms (57).

Many studies have shown that bacteria have a differential expression of genes when they are growing in biofilms as compared to planktonic states (66, 97) so its not hard to imagine that genes expressed by biofilms can contribute to their antibiotic resistance profiles. Therefore, by directly comparing the resistance of bacteria in planktonic and biofilm conditions the importance of certain genes in biofilm specific resistance can be monitored. By identifying genes that are differentially expressed in biofilms, it might be possible to target these genes in order to eliminate the biofilm infection and improve the condition of patients. Understanding the type of genes regulated is also of importance.

The role of many of the genes identified in the microarray include transcriptional regulators, genes involved in metabolism of carbohydrates or genes of unknown function (Table 5). Even though the primary role of the genes may not involve a direct link to antibiotic resistance, they may play a role in increasing resistance and they could be good targets for antibiotics if they are specific to bacterial systems.

The increase in antibiotic resistant bacteria has become a major problem for healthcare providers and patients. The old strategies for bacterial clearance with antibiotics is becoming obsolete. Novel approaches and finding new targets for drugs is of great importance in the arms race against bacteria. Targeting virulence, biofilm formation or antibiotic resistance mechanisms could prove to be vital in the treatment of bacterial infections. Targeting regulators that have an effect on antibiotic resistance in biofilms has great potential for future treatment of bacterial infections (71, 93, 100).

4.4: Future Work

There are several interesting avenues with which to continue this work. Primarily it would be important to produce an *agmR* knockout in PA14 in order to complete antibiotic phenotype assays (MBCP/B) on the mutant. It would also be beneficial to check the q-PCR expression patterns of the genes regulated by *agmR* with this mutant. The creation of a knockout and subsequent experiments are currently under way.

It would also be advantageous to purify glucans from *P. aeruginosa* and grow *ndvB* knockout biofilms supplemented with these glucans. If the presence of glucans in *ndvB* knockout biofilms can reestablish the expression of *agmR* and the related genes, then a link between gene regulation and cyclic glucans would be established. In addition to this, understanding the mechanisms of signalling, export and transport of cyclic glucans in the exopolymeric matrix is important. Finding a receptor for the cyclic glucans and potential binding partners, if they exist, would also be of interest.

Determine the antibiotic resistance profiles of transposon mutants should be done for several more drugs and it would be interesting to see the effect of bacterostatic drugs on the PQQ mutants. Along these same lines, it would be interesting to look at oxidative stress and free radical accumulation/hydroxyl radical effects on the PQQ mutants to see if they actually are scavenging free radicals produced during the use of bactericidal drugs.

Finally, the current work only focused on 8 genes from a list of 20 genes identified to be differentially regulated between Wt and *ndvB* KO biofilm conditions. Confirming the rest of the array data with q-PCR analysis of these genes and checking for the antibiotic resistance profiles with MBC-B/P test of mutants or knockouts of these genes would also be important for completeness of this study.

Conclusions

Conclusions

The gene expression of *ndvB* in *P. aeruginosa* is preferential for biofilm grown bacteria and the knockout shows reduced biofilm but not planktonic antibiotic resistance when compared to a wildtype strain of PA14. Both the knockout and wildtype strains produced viable biofilms when grown in plastic with minimal essential media and produce biofilms of sufficient biomass for adequate RNA extraction after 48 hours. This RNA was of sufficient quality for Affymetrix Genechip array analysis.

ndvB, which is responsible for the production of cyclic glucans, was implicated in the expression of 20 genes. Many of these genes are related to transcriptional regulation or metabolism, most notably the *agmR* response regulator and 6 of the genes it regulates were identified as being significantly more expressed in wildtype compared to knockout *ndvB* biofilms. I confirmed this observation with q-PCR analysis of the genes and this showed good correlation with the microarray data. These genes were also shown to be important in antibiotic resistance as MBC-B assays showed that the wildtype biofilm was more resistant to tobramycin and ciprofloxacin than the mutants of these genes. The mechanisms of this increased resistance due to expression of these genes were not studied.

List of References

List of References

1. **2008 Canadian Cystic Fibrosis Foundation.** 2008. January 3, 2008. Canadian Cystic Fibrosis Statistics. [Online.] www.cysticfibrosis.ca.

2. **Aeschlimann, J. R.** 2003. The role of multidrug efflux pumps in the antibiotic resistance of *Pseudomonas aeruginosa* and other gram-negative bacteria. Insights from the Society of Infectious Diseases Pharmacists. Pharmacotherapy **23:**916-924.

3. **Agodi, A., M. Barchitta, R. Cipresso, L. Giaquinta, M. A. Romeo, and C. Denaro.** 2007. *Pseudomonas aeruginosa* carriage, colonization, and infection in ICU patients. Intensive Care Med. **33:**1155-1161.

4. **Aizenman, E., K. A. Hartnett, C. Zhong, P. M. Gallop, and P. A. Rosenberg.** 1992. Interaction of the putative essential nutrient pyrroloquinoline quinone with the N-methyl-D-aspartate receptor redox modulatory site. J. Neurosci. **12:**2362-2369.

5. **Allison DG, Gilbert, P, Lappin-Scott HM and Wilson M. (ed.).** 2000. Commuinity Structure and co-operation in biofilms. Cambridge University Press, Cambridge, Great Britain.

6. **Anderson, G.G and George O'Toole.** 2008. Innate and Induced Resistance Mechanisims of Bacterial Biofilms, p. 85-105. *In* T. Romeo (ed.), Bacterial Biofilms. Springer Berlin Heidleberg, Berlin.

7. **Anwar, H., T. van Biesen, M. Dasgupta, K. Lam, and J. W. Costerton.** 1989. Interaction of biofilm bacteria with antibiotics in a novel in vitro chemostat system. Antimicrob. Agents Chemother. **33:**1824-1826.

8. **Barken, K. B., S. J. Pamp, L. Yang, M. Gjermansen, J. J. Bertrand, M. Klausen, M. Givskov, C. B. Whitchurch, J. N. Engel, and T. Tolker-Nielsen.** 2008. Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in *Pseudomonas aeruginosa* biofilms. Environ. Microbiol. **10:**2331-2343.

9. **Bengtsson, H.** 2003. The R.oo package - Object-Oriented Programming with References Using Standard R Code *In* K. Hornik, F. Leisch and A. Zeileis (ed.), Proceedings of the 3rd International Workshop on Distributed Statistical Computing.

10. **Bhagwat, A. A., K. C. Gross, R. E. Tully, and D. L. Keister.** 1996. Beta-glucan synthesis in *Bradyrhizobium japonicum*: characterization of a new locus (ndvC) influencing beta-(1-->6) linkages. J. Bacteriol. **178:**4635-4642.

11. **Bhagwat, A. A., A. Mithofer, P. E. Pfeffer, C. Kraus, N. Spickers, A. Hotchkiss, J. Ebel, and D. L. Keister.** 1999. Further studies of the role of cyclic beta-glucans in symbiosis. An NdvC mutant of *Bradyrhizobium japonicum* synthesizes cyclodecakis-(1-->3)-beta-glucosyl. Plant Physiol. **119:**1057-1064.

12. **Bjarnsholt, T., P. O. Jensen, M. J. Fiandaca, J. Pedersen, C. R. Hansen, C. B. Andersen, T. Pressler, M. Givskov, and N. Hoiby.** 2009. *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. Pediatr. Pulmonol. .

13. **Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed.** 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19:**185-193.

14. **Breedveld, M. W., and K. J. Miller.** 1994. Cyclic beta-glucans of members of the family *Rhizobiaceae*. Microbiol. Rev. **58:**145-161.

15. **Brown, M. R., D. G. Allison, and P. Gilbert.** 1988. Resistance of bacterial biofilms to antibiotics: a growth-rate related effect? J. Antimicrob. Chemother. **22:**777-780.

16. **Caiazza, N. C., and G. A. O'Toole.** 2004. SadB is required for the transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas aeruginosa* PA14. J. Bacteriol. **186:**4476-4485.

17. **Chang, W. S., M. van de Mortel, L. Nielsen, G. Nino de Guzman, X. Li, and L. J. Halverson.** 2007. Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. J. Bacteriol. **189:**8290-8299.

18. **Choi, J. Y., C. D. Sifri, B. C. Goumnerov, L. G. Rahme, F. M. Ausubel, and S. B. Calderwood.** 2002. Identification of virulence genes in a pathogenic strain of *Pseudomonas aeruginosa* by representational difference analysis. J. Bacteriol. **184:**952-961.

19. **Davey, M. E., and G. A. O'toole.** 2000. Microbial biofilms: from ecology to molecular genetics. Microbiol. Mol. Biol. Rev. **64:**847-867.

20. **Davis, B. D.** 1987. Mechanism of bactericidal action of aminoglycosides. Microbiol. Rev. **51:**341-350.

21. **Deley, J.** 1964. *Pseudomonas* and Related Genera. Annu. Rev. Microbiol. **18:**17-46.

22. **Doggett, R. G., G. M. Harrison, and E. S. Wallis.** 1964. Comparison of some Properties of *Pseudomonas Aeruginosa* Isolated from Infections in Persons with and without Cystic Fibrosis. J. Bacteriol. **87:**427-431.

23. **Donlan, R. M.** 2002. Biofilms: microbial life on surfaces. Emerg. Infect. Dis. **8:**881-890.

24. **Donlan, R. M., and J. W. Costerton.** 2002. Biofilms: survival mechanisms of clinically relevant microorganisms. Clin. Microbiol. Rev. **15:**167-193.

25. **Doring, G., S. P. Conway, H. G. Heijerman, M. E. Hodson, N. Hoiby, A. Smyth, and D. J. Touw.** 2000. Antibiotic therapy against *Pseudomonas aeruginosa* in cystic fibrosis: a European consensus. Eur. Respir. J. **16:**749-767.

26. **Draghici, S.** 2003. Data Analysis Tools for DNA micro Arrays. Chapman & Hall, New York.

27. **Drlica, K., and X. Zhao.** 1997. DNA gyrase, topoisomerase IV, and the 4-quinolones. Microbiol. Mol. Biol. Rev. **61:**377-392.

28. **Duine, J. A.** 1999. The PQQ story. J. Biosci. Bioeng. **88:**231-236.

29. **Dunne, W. M.,Jr.** 2002. Bacterial adhesion: seen any good biofilms lately? Clin. Microbiol. Rev. **15:**155-166.

30. **Eagye, K. J., D. P. Nicolau, and J. L. Kuti.** 2009. Impact of superinfection on hospital length of stay and costs in patients with ventilator-associated pneumonia. Semin. Respir. Crit. Care. Med. **30:**116-123.

31. **Favero, M. S., L. A. Carson, W. W. Bond, and N. J. Petersen.** 1971. *Pseudomonas aeruginosa*: growth in distilled water from hospitals. Science **173:**836-838.

32. **Foley, I., P. Marsh, E. M. Wellington, A. W. Smith, and M. R. Brown.** 1999. General stress response master regulator rpoS is expressed in human infection: a possible role in chronicity. J. Antimicrob. Chemother. **43:**164-165.

33. **Fux, C. A., J. W. Costerton, P. S. Stewart, and P. Stoodley.** 2005. Survival strategies of infectious biofilms. Trends Microbiol. **13:**34-40.

34. **Giltner, C. L., E. J. van Schaik, G. F. Audette, D. Kao, R. S. Hodges, D. J. Hassett, and R. T. Irvin.** 2006. The *Pseudomonas aeruginosa* type IV pilin receptor binding domain functions as an adhesin for both biotic and abiotic surfaces. Mol. Microbiol. **59:**1083-1096.

35. **Gliese, N., V. Khodaverdi, M. Schobert, and H. Gorisch.** 2004. AgmR controls transcription of a regulon with several operons essential for ethanol oxidation in *Pseudomonas aeruginosa* ATCC 17933. Microbiology **150:**1851-1857.

36. **Gorisch, H.** 2003. The ethanol oxidation system and its regulation in *Pseudomonas aeruginosa.* Biochim. Biophys. Acta **1647:**98-102.

37. **Gregory Alvord, W., J. A. Roayaei, O. A. Quinones, and K. T. Schneider.** 2007. A microarray analysis for differential gene expression in the soybean genome using Bioconductor and R. Brief Bioinform **8:**415-431.

38. **Guggino, W. B.** 2001. Cystic fibrosis salt/fluid controversy: in the thick of it. Nat. Med. **7:**888-889.

39. **Hall-Stoodley, L., J. W. Costerton, and P. Stoodley.** 2004. Bacterial biofilms: from the natural environment to infectious diseases. Nat. Rev. Microbiol. **2:**95-108.

40. **Hentzer, M., G. M. Teitzel, G. J. Balzer, A. Heydorn, S. Molin, M. Givskov, and M. R. Parsek.** 2001. Alginate overproduction affects *Pseudomonas aeruginosa* biofilm structure and function. J. Bacteriol. **183:**5395-5401.

41. **Hoiby, N., H. Krogh Johansen, C. Moser, Z. Song, O. Ciofu, and A. Kharazmi.** 2001. *Pseudomonas aeruginosa* and the in vitro and in vivo biofilm mode of growth. Microbes Infect. **3:**23-35.

42. **Holloway, B. W.** 1969. Genetics of Pseudomonas. Bacteriol. Rev. **33:**419-443.

43. **Hubbell, E., W. M. Liu, and R. Mei.** 2002. Robust estimators for expression analysis. Bioinformatics **18:**1585-1592.

44. **Hunt, J. C., and P. V. Phibbs Jr.** 1983. Regulation of alternate peripheral pathways of glucose catabolism during aerobic and anaerobic growth of *Pseudomonas aeruginosa*. J. Bacteriol. **154:**793-802.

45. **Kanner, D., N. N. Gerber, and R. Bartha.** 1978. Pattern of phenazine pigment production by a strain of *Pseudomonas aeruginosa*. J. Bacteriol. **134:**690-692.

46. **Knowles, M. R., and R. C. Boucher.** 2002. Mucus clearance as a primary innate defense mechanism for mammalian airways. J. Clin. Invest. **109:**571-577.

47. **Knudsen, P. K., H. V. Olesen, N. Hoiby, M. Johannesson, F. Karpati, B. N. Laerum, P. Meyer, T. Pressler, A. Lindblad, and Scandinavian CF Study**

Consortium (SCFSC). 2009. Differences in prevalence and treatment of *Pseudomonas aeruginosa* in cystic fibrosis centres in Denmark, Norway and Sweden. J. Cyst Fibros **8:**135-142.

48. **Kohanski, M. A., D. J. Dwyer, B. Hayete, C. A. Lawrence, and J. J. Collins.** 2007. A common mechanism of cellular death induced by bactericidal antibiotics. Cell **130:**797-810.

49. **Lewis, K.** 2005. Persister cells and the riddle of biofilm survival. Biochemistry (Mosc) **70:**267-274.

50. **Liberati, N. T., J. M. Urbach, S. Miyata, D. G. Lee, E. Drenkard, G. Wu, J. Villanueva, T. Wei, and F. M. Ausubel.** 2006. An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. Proc. Natl. Acad. Sci. U. S. A. **103:**2833-2838.

51. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T) Method. Methods **25:**402-408.

52. **Lyczak, J. B., C. L. Cannon, and G. B. Pier.** 2000. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. Microbes Infect. **2:**1051-1060.

53. **Ma, J., and Z. S. Qin.** 2007. Different normalization strategies for microarray gene expression traits affect the heritability estimation. BMC Proc. **1 Suppl 1:**S154.

54. **Ma, L., M. Conover, H. Lu, M. R. Parsek, K. Bayles, and D. J. Wozniak.** 2009. Assembly and development of the *Pseudomonas aeruginosa* biofilm matrix. PLoS Pathog. **5:**e1000354.

55. **Magnusson, O. T., H. Toyama, M. Saeki, A. Rojas, J. C. Reed, R. C. Liddington, J. P. Klinman, and R. Schwarzenbacher.** 2004. Quinone biogenesis: Structure and mechanism of PqqC, the final catalyst in the production of pyrroloquinoline quinone. Proc. Natl. Acad. Sci. U. S. A. **101:**7913-7918.

56. **Mah, T. F., and G. A. O'Toole.** 2001. Mechanisms of biofilm resistance to antimicrobial agents. Trends Microbiol. **9:**34-39.

57. **Mah, T. F., B. Pitts, B. Pellock, G. C. Walker, P. S. Stewart, and G. A. O'Toole.** 2003. A genetic basis for *Pseudomonas aeruginosa* biofilm antibiotic resistance. Nature **426:**306-310.

58. **Marsh, P. D.** 2005. Dental plaque: biological significance of a biofilm and community life-style. J. Clin. Periodontol. **32 Suppl 6:**7-15.

59. **Mena, K. D., and C. P. Gerba.** 2009. Risk assessment of *Pseudomonas aeruginosa* in water. Rev. Environ. Contam. Toxicol. **201:**71-115.

60. **Merritt, J. H., D. E. Kadouri, and G. A. O'Toole.** 2005. Growing and analyzing static biofilms. Curr. Protoc. Microbiol. **Chapter 1:**Unit 1B.1.

61. **Miller, K. J., J. A. Hadley, and D. L. Gustine.** 1994. Cyclic [beta]-1,6-1,3-Glucans of Bradyrhizobium japonicum USDA 110 Elicit Isoflavonoid Production in the Soybean (Glycine max) Host. Plant Physiol. **104:**917-923.

62. **Misra, H. S., N. P. Khairnar, A. Barik, K. Indira Priyadarsini, H. Mohan, and S. K. Apte.** 2004. Pyrroloquinoline-quinone: a reactive oxygen species scavenger in bacteria. FEBS Lett. **578:**26-30.

63. **Moore, B.** 1966. An outbreak of urinary *Pseudomonas aeruginosa* infection acquired during urological operations. Lancet **2:**929-931.

64. **Nichols, W. W., S. M. Dorrington, M. P. Slack, and H. L. Walmsley.** 1988. Inhibition of tobramycin diffusion by binding to alginate. Antimicrob. Agents Chemother. **32:**518-523.

65. **Not Available.** 2006. Affymetrix GeneChip$^{\circledR}$ Expression Analysis Technical Manual. Affymetrix Inc., Santa Carla, California.

66. **Ochsner, U. A., P. J. Wilderman, A. I. Vasil, and M. L. Vasil.** 2002. GeneChip expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. Mol. Microbiol. **45:**1277-1287.

67. **O'May, C. Y., D. W. Reid, and S. M. Kirov.** 2006. Anaerobic culture conditions favor biofilm-like phenotypes in *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. FEMS Immunol. Med. Microbiol. **48:**373-380.

68. **O'May, C. Y., D. W. Reid, and S. M. Kirov.** 2006. Anaerobic culture conditions favor biofilm-like phenotypes in *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. FEMS Immunol. Med. Microbiol. **48:**373-380.

69. **O'Toole, G. A., and R. Kolter.** 1998. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. Mol. Microbiol. **30:**295-304.

70. **Patel, R.** 2005. Biofilms and antimicrobial resistance. Clin. Orthop. Relat. Res. **(437):**41-47.

71. **Pierce, C. G., D. P. Thomas, and J. L. Lopez-Ribot.** 2009. Effect of tunicamycin on *Candida albicans* biofilm formation and maintenance. J. Antimicrob. Chemother. **63:**473-479.

72. **Puehringer, S., M. Metlitzky, and R. Schwarzenbacher.** 2008. The pyrroloquinoline quinone biosynthesis pathway revisited: a structural approach. BMC Biochem. **9:**8.

73. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat. Genet. **32 Suppl:**496-501.

74. **Rahme, L. G., E. J. Stevens, S. F. Wolfort, J. Shao, R. G. Tompkins, and F. M. Ausubel.** 1995. Common virulence factors for bacterial pathogenicity in plants and animals. Science **268:**1899-1902.

75. **Romeo, T. (ed.).** 2008. Bacterial Biofilms. Springer Berlin Heidelberg, Berlin.

76. **Rozen, S., and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers, p. 365-386. *In* S. Krawetz and S. Misener (ed.), *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Human Press, Totowa, NJ.

77. **Rupp, M., and H. Gorisch.** 1988. Purification, crystallisation and characterization of quinoprotein ethanol dehydrogenase from *Pseudomonas aeruginosa*. Biol. Chem. Hoppe Seyler **369:**431-439.

78. **Saidi, I. S., J. F. Biedlingmaier, and P. Whelan.** 1999. In vivo resistance to bacterial biofilm formation on tympanostomy tubes as a function of tube material. Otolaryngol. Head Neck Surg. **120:**621-627.

79. **Savli, H., A. Karadenizli, F. Kolayli, S. Gundes, U. Ozbek, and H. Vahaboglu.** 2003. Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of *Pseudomonas aeruginosa* by real-time quantitative RT-PCR. J. Med. Microbiol. **52:**403-408.

80. **Schleheck, D., N. Barraud, J. Klebensberger, J. S. Webb, D. McDougald, S. A. Rice, and S. Kjelleberg.** 2009. *Pseudomonas aeruginosa* PAO1 preferentially grows as aggregates in liquid batch cultures and disperses upon starvation. PLoS ONE **4:**e5513.

81. **Schnider, U., C. Keel, C. Voisard, G. Defago, and D. Haas.** 1995. Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHA0: mutational inactivation of the genes results in overproduction of the antibiotic pyoluteorin. Appl. Environ. Microbiol. **61:**3856-3864.

82. **Schobert, M., and H. Gorisch.** 1999. Cytochrome c550 is an essential component of the quinoprotein ethanol oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of the genes encoding cytochrome c550 and an adjacent acetaldehyde dehydrogenase. Microbiology **145 (Pt 2):**471-481.

83. **Schobert, M., and H. Gorisch.** 2001. A soluble two-component regulatory system controls expression of quinoprotein ethanol dehydrogenase (QEDH) but not expression of cytochrome c(550) of the ethanol-oxidation system in *Pseudomonas aeruginosa*. Microbiology **147:**363-372.

84. **Schweizer, H. P.** 1991. The *agmR* gene, an environmentally responsive gene, complements defective *glpR*, which encodes the putative activator for glycerol metabolism in *Pseudomonas aeruginosa*. J. Bacteriol. **173:**6798-6806.

85. **September, S. M., F. A. Els, S. N. Venter, and V. S. Brozel.** 2007. Prevalence of bacterial pathogens in biofilms of drinking water distribution systems. J. Water. Health. **5:**219-227.

86. **Singh, P. K., A. L. Schaefer, M. R. Parsek, T. O. Moninger, M. J. Welsh, and E. P. Greenberg.** 2000. Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. Nature **407:**762-764.

87. **Starkey, M., and L. G. Rahme.** 2009. Modeling *Pseudomonas aeruginosa* pathogenesis in plant hosts. Nat. Protoc. **4:**117-124.

88. **Sternberg, C., and T. Tolker-Nielsen.** 2006. Growing and analyzing biofilms in flow cells. Curr. Protoc. Microbiol. **Chapter 1:**Unit 1B.2.

89. **Stoodley, P., K. Sauer, D. G. Davies, and J. W. Costerton.** 2002. Biofilms as complex differentiated communities. Annu. Rev. Microbiol. **56:**187-209.

90. **Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature **406:**959-964.

91. **Thelin, W. R., and R. C. Boucher.** 2007. The epithelium as a target for therapy in cystic fibrosis. Curr. Opin. Pharmacol. **7:**290-295.

92. **Vasseur, P., I. Vallet-Gely, C. Soscia, S. Genin, and A. Filloux.** 2005. The *pel* genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. Microbiology **151:**985-997.

93. **Veesenmeyer, J. L., A. R. Hauser, T. Lisboa, and J. Rello.** 2009. *Pseudomonas aeruginosa* virulence and therapy: evolving translational strategies. Crit. Care Med. **37:**1777-1786.

94. **Victorin, L.** 1967. An epidemic of otitis in newborns due to infection with *Pseudomonas aeruginosa*. Acta Paediatr. Scand. **56:**344-348.

95. **Wagner, V. E., and B. H. Iglewski.** 2008. *P. aeruginosa* Biofilms in CF Infection. Clin. Rev. Allergy Immunol. **35:**124-134.

96. **Walsh, C.** 2003. Where will new antibiotics come from? Nat. Rev. Microbiol. **1:**65-70.

97. **Whiteley, M., M. G. Bangera, R. E. Bumgarner, M. R. Parsek, G. M. Teitzel, S. Lory, and E. P. Greenberg.** 2001. Gene expression in *Pseudomonas aeruginosa* biofilms. Nature **413:**860-864.

98. **Williams, I., W. A. Venables, D. Lloyd, F. Paul, and I. Critchley.** 1997. The effects of adherence to silicone surfaces on antibiotic susceptibility in *Staphylococcus aureus.* Microbiology **143 (Pt 7):**2407-2413.

99. **Winsor, G. L., T. Van Rossum, R. Lo, B. Khaira, M. D. Whiteside, R. E. Hancock, and F. S. Brinkman.** 2009. *Pseudomonas* Genome Database:

facilitating user-friendly, comprehensive comparisons of microbial genomes. Nucleic Acids Res. **37:**D483-8.

100. **Yacoby, I., and I. Benhar.** 2007. Targeted anti bacterial therapy. Infect. Disord. Drug Targets **7:**221-229.

101. **Yang, S. S., J. Y. Lin, and Y. T. Lin.** 1998. Microbiologically induced corrosion of aluminum alloys in fuel-oil/aqueous system. J. Microbiol. Immunol. Infect. **31:**151-164.

102. **Zhang, L., and T. F. Mah.** 2008. Involvement of a novel efflux system in biofilm-specific resistance to antibiotics. J. Bacteriol. **190:**4447-4452.

Appendix

Appendix A-Sample Calculations

Log molecular weight:
1 basepair (bp) is approximately 660 Daltons (Da)
12,000 bp (660 Da/bp)= 7 920 000 da
Log (7 920 000) = 6.89

Migration distance calculation
$y=m(x)+b$, equation of a line, where y is the log molecular weight and x is the migration distance of the band in centimeters
$y= -0.25x +6.81$
$y= -0.25 (1.2) + 6.82$
$y= 6.51$
Band size
Molecular weight= 10^{y}
$= 10^{6.51}$
= 3 235 936 Da
Basepairs= Molecular weight/ 660 Da/bp
= 3 235 936 Da /660Da/bp
= 4902

Interquartile region (IQR):
IQR=Q3-Q1

Livak method:
Fold expression= $2^{-\Delta\Delta C_T}$
Where
$\Delta\Delta C_T=(C_{TTarget} - C_{TrpoD})Time_{(x)} - (C_{TTarget} - C_{TrpoD})Time_{(0)}$.

Appendix B: R Source Code Chunks

```
# Load ‘affy’ from Bioconductor
library('affy')


# Read in 6 .CEL files into an object called pae.ab
pae.ab <- ReadAffy('<pathname>/A3.CEL’,
                   '<pathname>/B3.CEL',
                   '<pathname>/C3-3.CEL',
                   '<pathname>/A4.CEL',
                   '<pathname>/B4-3.CEL',
                   '<pathname>/C4-3.CEL')

# Check contents of pae.ab object
# It is an AffyBatch object with 61170 affyids at this point
pae.ab

# Check sample names of pae.ab object
sampleNames(pae.ab)

# Rename the sampleNames: 'hr.a3.12' refers to the original
# A3.CEL file, which is Hawaii/Resistant;
# 'ts.a4.12' refers to the original A4.CEL file, which is
# Taiwan/Susceptible.
new.sampleNames <- c('WT1','WT2',’WT3’,'KO1','KO2',’KO3’)
sampleNames(pae.ab) <- new.sampleNames

# Check the sample names of the pae.ab object to confirm changes
sampleNames(pae.ab)


# Start preparation for phenoData slot in AffyBatch object
pd <- data.frame(population = c(1,1,1,2,2,2), replicate =
c(1,2,3,1,2,3)

# Display contents of pd
pd

# Assign the sampleNames(pae.ab) to the rownames of pd
rownames(pd) <- sampleNames(pae.ab)

# Display contents of pd again, notice change in rownames
pd

# Continue preparation for phenoData slot
metaData <- data.frame(labelDescription =
c(‘condition’,’replicate’)

# Establish new phenoData slot
phenoData(pae.ab) <- new(‘AnnotatedDataFrame’, data = pd,
varMetadata = metaData)

# Display pData(pae.ab)
pData(pae.ab)
```

```
# Display phenoData(pae.ab)
phenoData(pae.ab)

# Construct image plots for quality control assessment
# Note: that these plots appear dark; they are not log
transformed
par(oma = c(1,1,3,1)
par(mfrow = c(2,3)
palette.gray <- c(rep(gray(0:10/10), times = seq(1,41, by = 4))
image(pae.ab[,1], transfo = function(x) x, col = palette.gray)
image(pae.ab[,2], transfo = function(x) x, col = palette.gray)
image(pae.ab[,3], transfo = function(x) x, col = palette.gray)
image(pae.ab[,4], transfo = function(x) x, col = palette.gray)
image(pae.ab[,5], transfo = function(x) x, col = palette.gray)
image(pae.ab[,6], transfo = function(x) x, col = palette.gray)
mtext (“> image(pae.ab[,i=1:6], transfo = function(x) x, col =
palette.gray)”, side = 3, outer = T, cex = .8)

# Construct individual image plots using log intensities
# Note: that these plots appear lighter; they are log transformed
par(mfrow = c(1,1)
palette.gray <- c(rep(gray(0:10/10), times = seq(1,41, by = 4))
image(pae.ab[,1],  col = palette.gray)
title(sub = "> image(pae.ab[,1],  col = palette.gray)")

# Repeat above commands with .. pae.ab[,i], i = 2-6

# Display all six log-transformed image plots in one plot
par(oma = c(3,1,3,1)
par(mfrow = c(2,3)
image(pae.ab[,1],  col = palette.gray)
image(pae.ab[,2],  col = palette.gray)
image(pae.ab[,3],  col = palette.gray)
image(pae.ab[,4],  col = palette.gray)
image(pae.ab[,5],  col = palette.gray)
image(pae.ab[,6],  col = palette.gray)
mtext(‘Image Plots -WT vs MT’, side = 3, outer = T)
mtext(‘Top Row:WT - Bottom Row: MT’, side = 1, outer = T)

# Reset to single screen.
par(mfrow = c(1,1)

# Construct color boxplots
library('RColorBrewer')
brewer.cols <- brewer.pal(6, 'Set1')
boxplot(pae.ab, col = brewer.cols, ylab = 'Unprocessed log (base
2) scale Probe Intensities', xlab = 'Array Names')
title('Box Plots of PA Probe Level Data')

# Construct density plots
hist(pae.ab, col = brewer.cols, lty = 1, xlab = 'Log (base 2)
Intensities', lwd = 3)
samp.leg.names <- new.sampleNames
legend(5, .6, legend = samp.leg.names, lty = 1, col =
brewer.cols, lwd = 1)
```

```
title('Density plots - probe level data for four PA arrays', cex
= 1)

111# Construct MAplots
par(oma = c(1,1,3,1)
par(mfrow = c(2,3)
MAplot(pae.ab, cex = .5)
mtext('M', side = 2, outer = TRUE)
mtext('A', side = 1, outer = TRUE)
mtext('Six MA-Plots vs. the Median Mock Array', side = 3, outer =
TRUE)
par(mfrow = c(1,1)
# Fit a probe-level model to the pae.ab probe-level data
library(‘affyPLM’)
Pset1 <- fitPLM(pae.ab)

# Display probe-level quality diagnostics for array # 1
par(mfrow = c(2,2)
par(oma = c(3,1,3,1)
image(pae.ab[,1], col = palette.gray)
image(Pset1, type = ‘weights’, which = 1)
image(Pset1, type = ‘resids’, which = 1)
image(Pset1, type = ‘sign.resids’, which = 1)
mtext(‘Probe Level Models - QC Checks’, side = 3, outer = T)
par(mfrow = c(1,1)
# Repeat above commands with .. pae.ab[,i], i = 2-6; which = i


# Construct Relative Log Expression (RLE) Plot
library(affyPLM)
Mbox(Pset1, col = brewer.cols, names = NULL, main = “Relative Log
Expression Plot - WT vs. MT”)

# Add ylim = c(-.5, .5) after observing previous plot
Mbox(Pset1, col = brewer.cols, ylim = c(-.5,.5),  main =
'Relative Log Expression Plot - H/R vs. T/S')


# Construct Normalized Unscaled Standard Error (NUSE) Plot
boxplot(Pset1, col = brewer.cols, main = “NUSE Plot”, ylab =
“NUSE - Normalized Unscaled Standard Error”)

# Add ylim = c(.9, 1.1) after observing previous plot
boxplot(Pset1, ylim = c(.9, 1.1), col = brewer.cols, main = 'NUSE
Plot', ylab = 'NUSE - Normalized Unscaled Standard Error', las =
2)

# Execute RMA procedure on pae.ab object; assign to object eset
eset <- rma(pae.ab)

# Display eset
eset

# Create an object that contains exprs(eset), named exprs.eset
exprs.eset <- exprs(eset)
```

```
# Create index values - Index1 for WT,
# Index2 for KO
Index1 <- 1:3
Index2 <- 4:6
# Compute Difference vector for rowMeans between WT and KO
Difference <- rowMeans(exprs.eset[,Index1]) -
rowMeans(exprs.eset[,Index2])

# Also compute the Average for each row
Average <- rowMeans(exprs.eset)

# Create data frame for matrix exprs.eset
exprs.eset.df <- data.frame(exprs.eset)

# Construct expression set boxplots following RMA
par(oma = c(1,1,3,1)
boxplot(exprs.eset.df, col = brewer.colors)
mtext('RMA expression Data', side = 3, outer = T)

# Construct MA-Plot, Difference vs. Average
plot(Average, Difference)
lines(lowess(Average, Difference), col = 'red', lwd = 4)
abline( h = -2)
abline( h = 2)
title(sub = "> lines(lowess(Average, Difference), col = 'red',
lwd = 4)")
mtext('MA-Plot, Difference vs. Average, PA (WT & MT)', outer = T,
side = 3)


# Compute ordinary t statistics
library('genefilter')
tt <- rowttests(exprs.eset, factor(eset$population)
names(tt)
# Check that length of tt$statistic is 5600
length(tt$statistic)

# Start the construction of the Volcano Plot
# Construct the lod scores...
lod <- -log10(tt$p.value)
o1 <- order(abs(Difference), decreasing = TRUE)[1:50]
o2 <- order(abs(tt$statistic), decreasing = TRUE)[1:50]

# Construct union and intersection of sets
o <- union(o1, o2)
i <- intersect(o1, o2)
# Construct simple Volcano plot to get an idea of the range
# of the Difference values and lod values
plot(Difference, lod)

# Construct more sophisticated set of commands for a nicer
Volcano plot
par(mfrow = c(1,1)
plot(Difference[-o], lod[-o], cex = .25, xlim = c(-5,5), ylim =
range(lod)
points(Difference[o1], lod[o1], pch = 18, col = 'blue', cex =
1.5)
```

```
points(Difference[o2], lod[o2], pch = 1, col = 'red', cex = 2,
lwd = 2)
abline(h = 3)
text(-2, 3.8, “Red circles: 70 genes”)
text(-2, 3.6, “with lowest p-values”)
text(-2, 3.4, “=====>>>”)
text(-2, 3.15, “p < 0.00001”)
title(“Volcano Plot: lowest p values and largest fold
differences”)


# Prepare to calculate attenuated t statistics using limma
package
# Load limma package
library(limma)

# Construct population.groups for model design
population.groups <- factor(c(rep(‘MT’,2), rep('WT',2))
design <- model.matrix( ~ population.groups )

# Display design
design

# Fit linear model
fit <- lmFit(eset, design)
# Prepare to calculate attenuated t statistics using limma
package
# Load limma package
library(limma)

# Construct population.groups for model design
population.groups <- factor(c(rep(‘MT’,2), rep('WT',2))
design <- model.matrix( ~ population.groups )

# Display design
design

# Fit linear model
fit <- lmFit(eset, design)

# Check the dimension of fit$coeff; should be 5600x 2
dim(fit$coeff)

# Execute eBayes to obtain attenuated t statistics
fit.eBayes <- eBayes(fit)

# Check dimension of fit.eBayes$t; should be 5600 x 2
dim(fit.eBayes$t)

# Compute statistics necessary for Volcano Plot
# using attenuated t statistics
lodd <- -log10(fit.eBayes$p.value[,2])
oo2 <- order(abs(fit.eBayes$t[,2]), decreasing = TRUE)[1:50]
oo <- union(o1, oo2)
ii <- intersect(o1, oo2)

# Display ii
```

```
ii

# Construct Volcano plot using attenuated t statistics
plot(Difference[-oo], lodd[-oo], cex = .25, xlim = c(-3,3), ylim
= range(lodd), xlab = ‘Average (log) Fold-change’, ylab = ‘LOD
score - Negative log10 of P-value’)
points(Difference[o1], lodd[o1], pch = 18, col = 'blue', cex =
1.5)
points(Difference[oo2], lodd[oo2], pch = 1, col ='red',cex =
2,lwd = 2)
abline(h = 3)
title('Volcano Plot with moderated t statistics')
text(-2, 3.2, 'p < 0.001')
text(1, 4, 'Nine intersects')

# Construct the appropriate matrices/data frames for
# construction of the heat map.
ii.mat <- exprs.eset[ii,]
ii.df <- data.frame(ii.mat)

# Dislplay ii.df
ii.df

# Construct heatmap using reds and blues
library('RColorBrewer')
library('genefilter')
hmcol <- colorRampPalette(brewer.pal(10, 'RdBu')(256)
tv.MT <- dimnames(ii.mat)[[2]][1:3]
spcol <- ifelse(dimnames(ii.mat)[[2]] == tv.MT, 'skyblue’,
‘goldenrod’)
heatmap(ii.mat, col = hmcol, ColSideColors = spcol, margins =
c(10,15)

# Construct heatmap using a grey (gray) scale
library('RColorBrewer')
library('genefilter')
hmcol <- colorRampPalette(brewer.pal(9, 'Greys')(256)
tv.Hawaii.Resistant <- dimnames(ii.mat)[[2]][1:3]
spcol <- ifelse(dimnames(ii.mat)[[2]] == tv.Hawaii.Resistant,
'grey10’, ‘grey80’)
heatmap(ii.mat, col = hmcol, ColSideColors = spcol, margins =
c(10,15)
```

More code:

```
library(simpleaffy)
raw.data <- read.affy("covdesc.txt")
x.rma <- call.exprs(raw.data, "rma")
write.exprs(x.rma, file="rma_affy_all2.xls")
results <- pairwise.comparison(x.rma, "Condition", c("a", "b"),
raw.data)
write.table(data.frame(means(results), fc(results), tt(results),
calls(results), file="my_comp.xls", sep="\t", quote=F, col.names
= NA)
significant <- pairwise.filter(results, fc=log2(4),
min.present.no=1, tt= 0.00001, present.by.group=FALSE)
sig.rma <- data.frame(means(significant), fc(significant),
tt(significant), calls(significant)
```

```
write.table(sig.rma, file="sig.rma2.xls", sep="\t", quote=F,
col.names = NA)
```

```
library(simpleaffy)
raw.data <- read.affy("covdesc.txt")
x.mas <- call.exprs(raw.data, "mas5")
write.exprs(x.mas, file="mas_affy_all.xls")
results <- pairwise.comparison(x.mas, "Condition", c("a", "b"),
raw.data)
write.table(data.frame(means(results), fc(results), tt(results),
calls(results), file="my_comp2.xls", sep="\t", quote=F, col.names
= NA)
significant <- pairwise.filter(results, fc=log2(4),
min.present.no=1, tt= 0.00001, present.by.group=FALSE)
sig.mas5 <- data.frame(means(significant), fc(significant),
tt(significant), calls(significant)
write.table(sig.mas5, file="sig.mas52.xls", sep="\t",quote=F,
col.names = NA)
pdf(file="my_comp_mas.pdf"); plot(significant,type="scatter");
dev.off()
```

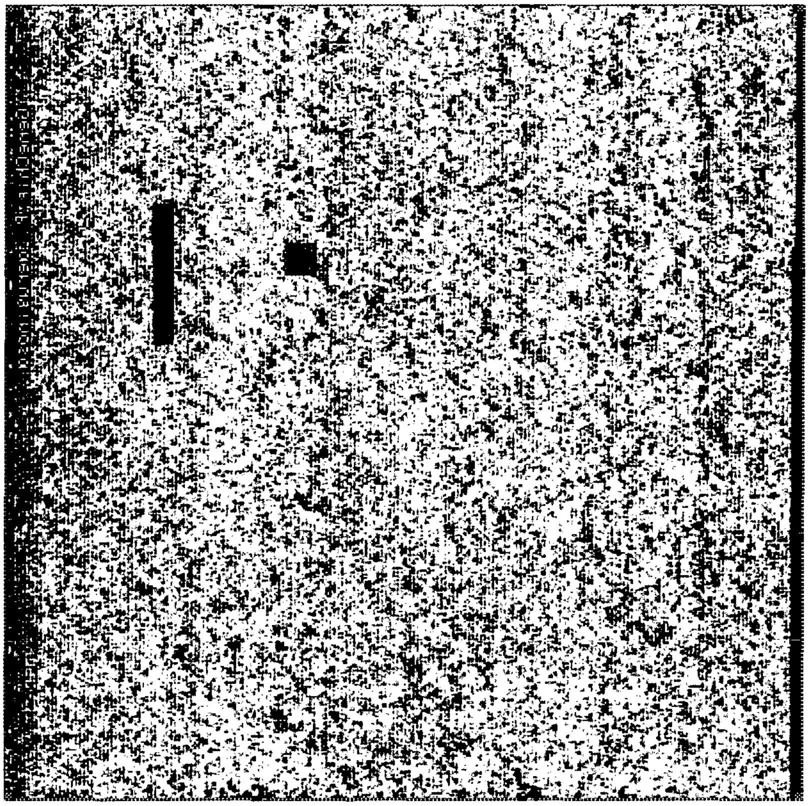
```
overlap <-(merge(sig.rma, sig.mas5, by.x = "row.names", by.y =
"row.names", all = FALSE)
write.table(overlap, file="overlap.xls", sep="\t",quote=F,
col.names = NA)
```

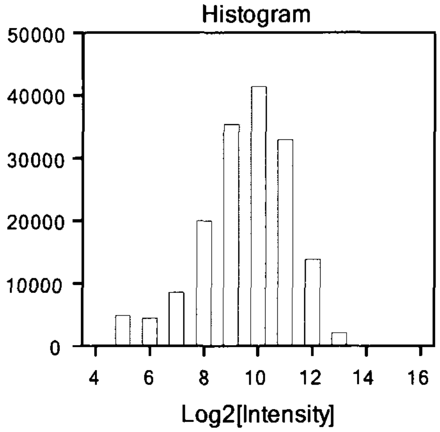
Appendix C: Preprocessed Chip data

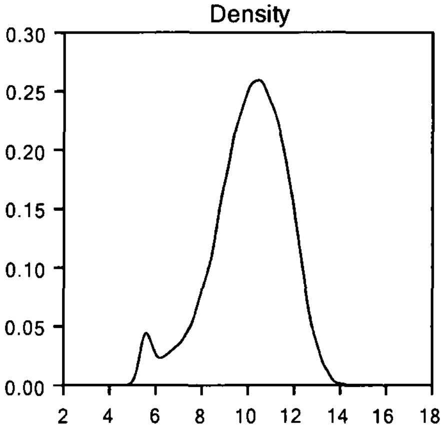
Figure 23: Preprocessed chip data of microarrays.

Preprocessed chip data of WT2,WT3, KO2 and KO3 microarray slides along with density plots and histograms.

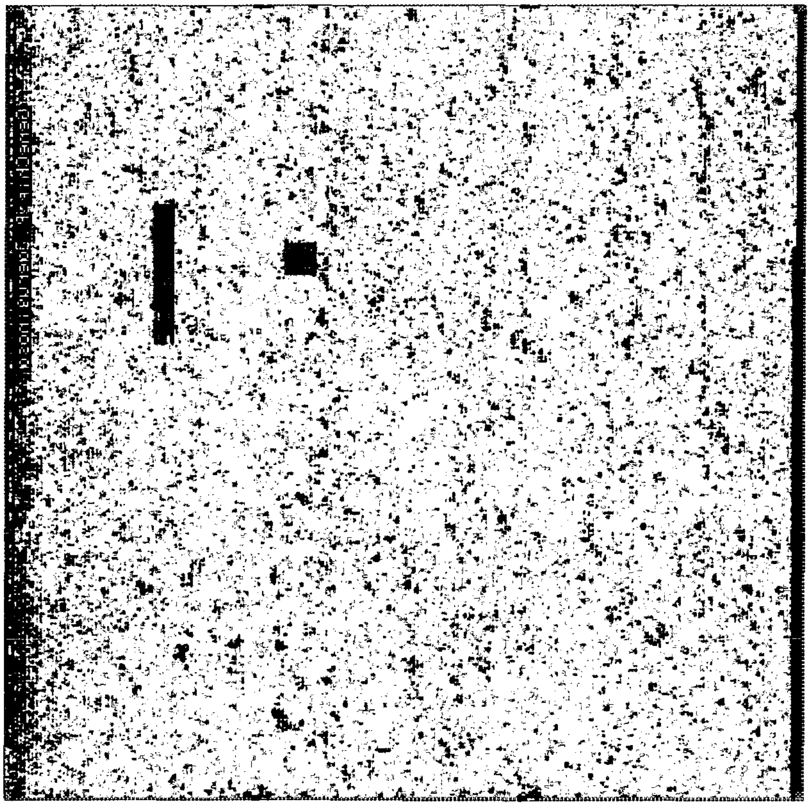
WT2

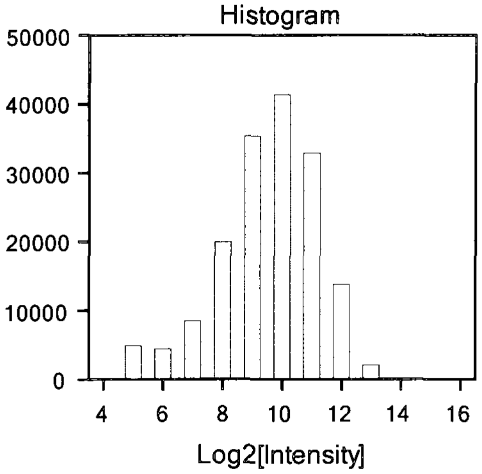


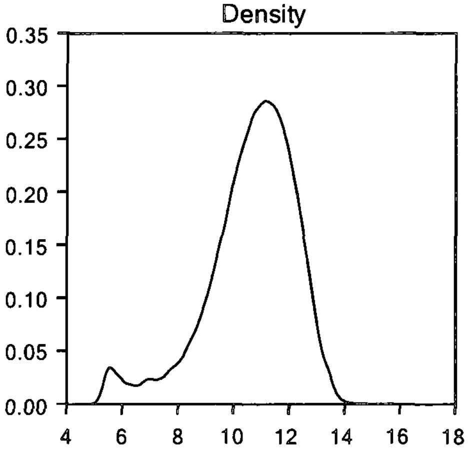

Histogram
50000
40000
30000
20000
10000
0
4
6
8
10
12
14
16
Log2[Intensity]


Density
0.30
0.25
0.20
0.15
0.10
0.05
0.00
2
4
6
8
10
12
14
16
18

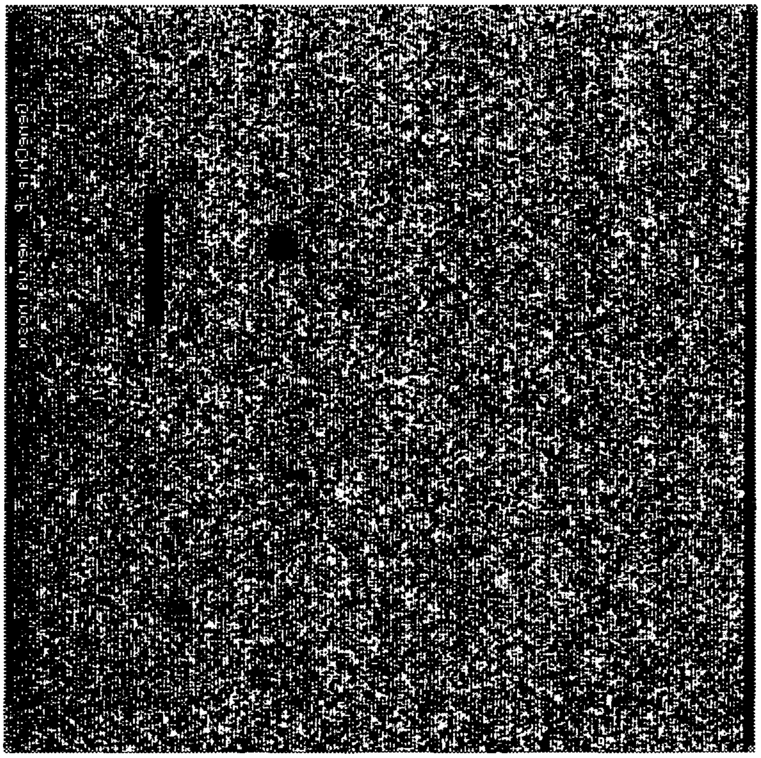
WT3

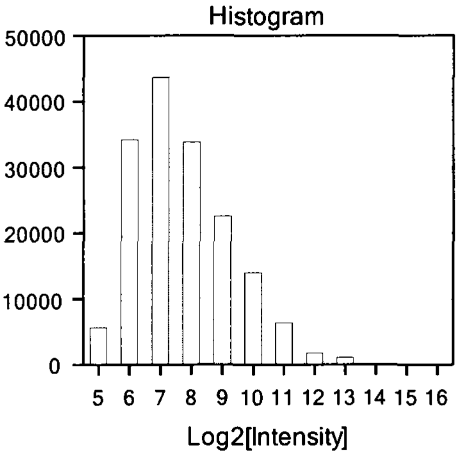


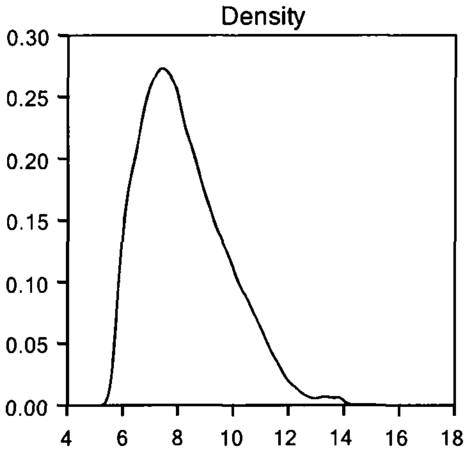

Histogram
50000
40000
30000
20000
10000
0
4
6
8
10
12
14
16
Log2[Intensity]


Density
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4
6
8
10
12
14
16
18

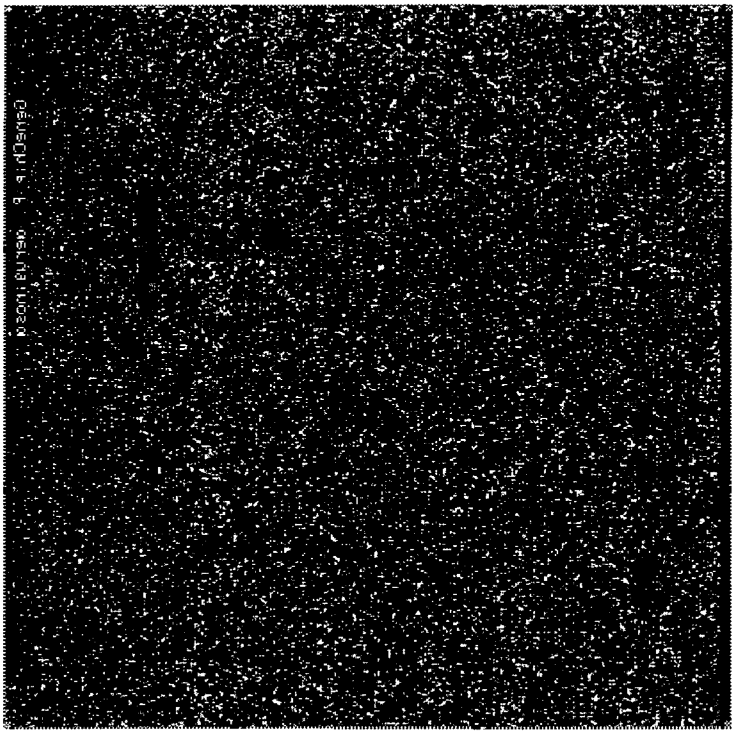
KO2

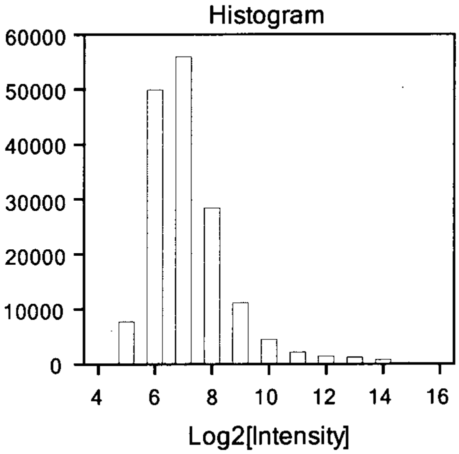


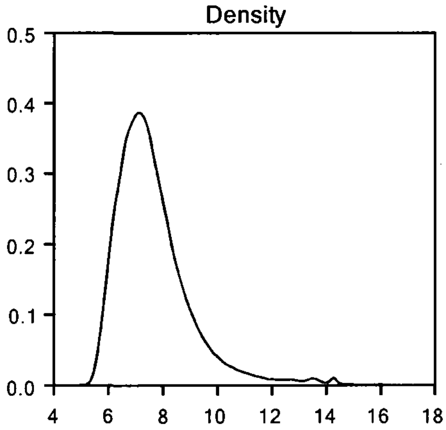

Histogram
50000
40000
30000
20000
10000
0
5
6
7
8
9
10
11
12
13
14
15
16
Log2[Intensity]


Density
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4
6
8
10
12
14
16
18

KO3




Histogram
60000
50000
40000
30000
20000
10000
0
4
6
8
10
12
14
16
Log2[Intensity]


Density
0.5
0.4
0.3
0.2
0.1
0.0
4
6
8
10
12
14
16
18