The role of bacteriophages in mixed microbial communities and populations of *Pseudomonas aeruginosa*

Muhammad Hafiz Ismail

2019

Muhammad Hafiz Ismail. (2019). The role of bacteriophages in mixed microbial communities and populations of *Pseudomonas aeruginosa*. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/83268>

<https://doi.org/10.32657/10220/48129>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**THE ROLE OF BACTERIOPHAGES IN
MIXED MICROBIAL COMMUNITIES AND
POPULATIONS OF *PSEUDOMONAS
AERUGINOSA***

MUHAMMAD HAFIZ BIN ISMAIL

School of Biological Sciences

Singapore Centre for Environmental Life Sciences Engineering

2019

**THE ROLE OF BACTERIOPHAGES IN MIXED MICROBIAL
COMMUNITIES AND POPULATIONS OF *PSEUDOMONAS*
*AERUGINOSA***

MUHAMMAD HAFIZ BIN ISMAIL

B.Sc. Biological Sciences (Hons)

Nanyang Technological University

School of Biological Sciences

Singapore Centre for Environmental Life Sciences Engineering

A thesis submitted to the Nanyang Technological University in partial
fulfilment of the requirement for the degree of

Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

25th February 2019

.....
Date

.....
Muhammad Hafiz Bin Ismail

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

25th February 2019

.....

Date

.....

Scott A. Rice

Authorship Attribution Statement

Thesis title:

The role of bacteriophages in mixed microbial communities and populations of *Pseudomonas aeruginosa*

I have excluded the Authorship Attribution Statement declaration because the thesis does not contain any published material.

.....
Muhammad Hafiz Bin Ismail

Date : 25th February 2019

.....
Scott A. Rice

Date: 25th February 2019

Acknowledgements

I would like to express my utmost gratitude to my supervisor, Assoc. Prof Scott A. Rice, for giving me the opportunity to work on this project. I thank him for the continuous support from when I was an undergraduate till now as I submit my PhD thesis. This project has been challenging and I appreciate that Scott has been encouraging and have provided guidance throughout my candidature. It has been a pleasant and fruitful experience and I hope that I have achieved the expectations you have of me.

I also thank Assoc. Prof Diane McDougald for having been my mentor and the contributions she had to my projects. The feedback and comments of my writing has helped me in the completion of this thesis and have made it much better.

I am extremely grateful to Dr Martin Tay, for having taught and mentoring me from when I was an undergraduate. Your guidance and insights when I just started has been really helpful. I thank you for your advice as I went along this journey and also the friendship that we have. Thank you for also introducing me to Bioinformatics and initiating me on this skill. Gosh, it was really difficult to learn. Clearly this project would not have been possible without it.

To Dr Grant Tan, I am extremely grateful for your patience and willingness to guide me throughout my project. I will not forget the time and effort you have taken in mentoring me, as well as the friendly banters we have in both the lab and outside. Without those, and also without your help in the operation of the bioreactors, this project would have been much more difficult.

I am also very grateful to Dr Chan Siew Heng Stan for his efforts during the operation of the bioreactors and for your help in teaching me the basics of the bioprocesses involved. Stan has been a great team player and working on the bioreactors with you has made it much more enjoyable. I also appreciate the friendship that we have.

I would also like to specially thank Dr Kelvin Lee for your insights in RNA work and transcriptomics analysis.

I would like to thank the School of Biological Sciences, NTU and Singapore Centre for Environmental Life Sciences Engineering for having given me the opportunity and resources to do my PhD in such a conducive environment.

I thank the PUB Environment and Water Industry Programme office for having awarded me my scholarship and this chance to pursue my PhD.

To my office table neighbours Kelvin Chong and Adeline Yong, thank you for your friendship and support during our candidature. This is it. We're submitting and our PhD journey is coming to a close. Congratulations to the two of you too! Not forgetting Teh Wooi Keong, Renee Nicole Ng, Casandra Tan and Dr Ding Yichen. You all have been awesome neighbours. To the glass panel beside mine and Kelvin's office table, thank you for being a "whiteboard" for all discussions. A good amount of ideas came out of the discussions and scribbles we have on there.

To my other friends in SCELSE, Tay Wei Hong (who is also submitting, congratulations!), Dr Tan Shi Ming, Choo Pei Yi, Dr Poh Wee Han, Goh Yu Fen, Brenda Tien, Wong Jun Jie, Zeus Jaren Nair, all my students and so many others. Thank you for making SCELSE a great place to be in. I learned a lot.

I would like to also thank my other friends, Adrian Thian, Chen Junzheng, Jolene Ong, Wahyudin Hassan, Leonard Lim, Lee Wen Qin, Sarah Tan and others for your moral support and encouragements.

Lastly, I would like to thank my family without whom I wouldn't even have been able to start this journey. To my father, Ismail Padri, my mother, Salinah Ngaspan, my brothers Ridhwan and Hazwan, my sister, Nurfatin Ayuni, my grandmother, Nor, and all my other family. Thank you for your unwavering support and love through this. The completion of this PhD is also for all of you.

Table of Contents

Acknowledgements	xi
Table of Contents	xiii
List of Tables	xvii
List of Figures	xix
List of Abbreviations	xxii
Abstract	xxiv
Chapter 1 – Literature Review	1
1.1. Introduction to bacterial biofilms	1
1.1.1. Ubiquity of biofilms	2
1.2. The use of activated sludge in water reclamation	5
1.2.1. Enhanced biological phosphorus removal	6
1.2.2. Simultaneous nitrification and denitrification	8
1.2.3. Simultaneous nitrification-denitrification and phosphorus removal ..	9
1.2.4. Aerobic granulation in biological wastewater reclamation	10
1.2.4.1. Non-biological and biological factors impacting the formation of aerobic granules	12
1.3. Predators of bacteria in the environment and in water reclamation systems	16
1.3.1. Protozoa	16
1.3.2. Predatory bacteria	18
1.3.3. Bacteriophages	19
1.3.3.1. Bacteriophages and their potential roles in water reclamation systems	21
1.4. <i>Pseudomonas aeruginosa</i> biofilms	23
1.4.1. Role of bacteriophage in biofilm development, dispersal, and variant formation	24
1.4.1.1. The bacteriophage lytic and lysogenic switch	26
1.5. Thesis aims	28

Chapter 2 – The role of bacteriophage in the formation of aerobic granules

.....	31
2.1. Introduction	31
2.2. Materials and methods.....	33
2.2.1. Sequencing batch reactor operation	33
2.2.2. Sampling and performance studies of the sequencing batch reactors	35
2.2.3. Concentration of the viral fraction	36
2.2.4. Bacteriophage nucleic acid extraction and multiple displacement amplification	38
2.2.5. Total genomic DNA extraction from aerobic granular sludge.....	39
2.2.6. Total metagenome analysis of the aerobic granular sludge.....	39
2.2.7. Viral fraction analysis	40
2.3. Results.....	41
2.3.1. Development of aerobic granular sludge from activated floccular sludge.....	41
2.3.2. Microbial community composition of floccular and granular sludge	44
2.3.3. Virome dynamics of a granulating sludge	50
2.4. Discussion.....	53
2.4.1. The aerobic granulation process	53
2.4.2. Changes in the sludge microbial community.....	54
2.4.3. Population dynamics of the bacteriophage community during granulation	57

Chapter 3 – The role of Pf filamentous phage in *Pseudomonas aeruginosa* PAO1 biofilm development.....

3.1. Introduction	62
3.2. Materials and methods.....	64
3.2.1. Bacterial strains and culture conditions.....	64
3.2.2. Gel electrophoresis	65
3.2.3. Genomic DNA extraction.....	65
3.2.4. Polymerase chain reaction	65
3.2.5. Generation of Pf6 mutant via Lambda RED mediated recombination	67

3.2.6. Screening for Pf6 knockout mutants	71
3.2.7. Construction of quantitative PCR quantification standards	72
3.2.8. Tubing biofilm system.....	73
3.2.9. Quantification of phage titre in biofilm effluent	74
3.2.10. Extraction of DNA from phage in biofilm effluent.....	75
3.2.11. Determination of phage population dynamics via qPCR.....	75
3.2.12. Transcriptomic analysis of PAO1 biofilms	76
3.3. Results	78
3.3.1. Generation and screening of Pf6 knockout mutants	78
3.3.2. Phage population dynamics	81
3.3.3. Transcriptomic comparison of <i>P. aeruginosa</i> PAO1 and PAO1ΔPf4 biofilms	83
3.4. Discussion.....	102
3.4.1. The Pf6 prophage may be essential.....	102
3.4.2. Pf phage population dynamics in PAO1 and PAO1ΔPf4 biofilms	104
3.4.3. Transcriptome of the wild type <i>P. aeruginosa</i> PAO1 and the Pf4 knockout biofilms.....	106
Chapter 4 – The role of the repressor C of the <i>Pseudomonas aeruginosa</i> PAO1 filamentous phages in global gene regulation	110
4.1. Introduction	110
4.2. Materials and methods.....	113
4.2.1. Bacterial strains and culture conditions.....	113
4.2.2. Genomic DNA and plasmid DNA extraction.....	116
4.2.3. Polymerase chain reaction	116
4.2.4. Gel electrophoresis	118
4.2.5. Site-directed mutagenesis (SDM) of repressor <i>c</i>	119
4.2.6. Electrophoretic mobility shift assay (EMSA) of the RepC	120
4.2.7. Construction and expression of C-terminal hexahistidine-tagged RepC	122
4.2.8. <i>In vitro</i> immunity assay of repressor C protein.....	123
4.2.9. Quantification of <i>repC</i> expression by qRT-PCR.....	124
4.2.10. Chromatin immunoprecipitation sequencing (ChIPseq) sample preparation	125

4.2.11. ChIPseq analysis	127
4.2.12. Tubing biofilm system	128
4.2.13. Transcriptomic analysis of <i>repC</i> complemented biofilms	128
4.3. Results	129
4.3.1. Plasmid construction	129
4.3.2. Electrophoretic mobility shift assay of the repressor C proteins ..	131
4.3.3. The repressor <i>c</i> confers immunity to superinfection	132
4.3.4. ChIPseq analysis of the Pf4 and Pf6 RepC	134
4.3.4.1. The Pf4 RepC binding targets	137
4.3.4.2. The Pf6 RepC binding targets	141
4.3.5. Differential gene expression by complementation of Pf4 RepC and sdmRepC	144
4.4. Discussion.....	157
4.4.1. The role of RepC in immunity and phage conversion to a superinfective state	157
4.4.2. The potential regulon of Pf4 RepC.....	161
4.4.3. The potential regulon of Pf6 RepC.....	162
4.4.4. Conclusions.....	163
Chapter 5 – General discussion and conclusions	164
5.1. General discussion	164
5.1.1. The role of bacteriophages in granulation.....	165
5.1.2. The role of the Pf4 filamentous phage and its repressor C in biofilm development.....	168
5.1.2.1. Proposed model	170
5.2. Limitations of this study.....	173
5.3. Conclusions	175
Contributions	176
Permissions	178
References	179
Appendix Figures and Tables	209

List of Tables

Table 2.1. Pearson's correlation between virus family and particle size.	52
Table 3.1. Bacterial strains and plasmids.....	66
Table 3.2. Primer sequences.....	70
Table 3.3. Test for the effect of sample conditions on gene expression.	86
Table 3.4. Permutation test for homogeneity of multivariate dispersion.....	87
Table 3.5. Top 20 genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 2.	90
Table 3.6. Virulence associated genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 on day 2.....	92
Table 3.7. Top 20 genes upregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 2.	94
Table 3.8. Top 20 genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 7.	96
Table 3.9. Top 20 genes upregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 7.	98
Table 3.10. Virulence associated genes upregulated in PAO1ΔPf4 biofilms relative to wild type biofilms on day 7.....	100
Table 4.1. Bacterial strains and plasmids used.	114
Table 4.2. List of primers used.	117
Table 4.3. ChIPseq peaks detected and predicted gene targets of the Pf4 RepC.	138
Table 4.4. ChIPseq peaks detected for Pf6 RepC.....	142

Table 4.5. ChIP peak detected and the gene predicted to be targets of the Pf6 RepC.	143
Table 4.6. Test for the effect of complementation on gene expression.....	145
Table 4.7. Permutation test for homogeneity of multivariate dispersion.....	146
Table 4.8. Top 20 genes downregulated in PAO1 Δ Pf4 biofilms relative to PAO1.	148
Table 4.9. Top 20 genes upregulated in PAO1 Δ Pf4_ pJN105_RepC biofilms relative to the PAO1 Δ Pf4 deletion mutant.	150
Table 4.10. Top 20 genes upregulated in PAO1 Δ Pf4_ pJN105_sdmRepC biofilms relative to PAO1 Δ Pf4.	152
Table 4.11. Common differentially expressed genes, relative to the Δ Pf4 mutant.	156

List of Figures

Figure 1.1. Mechanism of nutrient removal in EBPR.....	7
Figure 1.2. Biological elimination of ammonium.	8
Figure 1.3. Physiological differences of sludge during the process of granulation.	11
Figure 1.4. The developmental stages of a microbial granule (a) and its similarity to the biofilm life cycle (b).	15
Figure 1.5. The morphological groups of protozoa	18
Figure 1.6. Microscopic images of <i>Pseudomonas aeruginosa</i> biofilms and aerobic granules.	23
Figure 1.7. The bistable genetic switch of lambda phage.	27
Figure 2.1. Setup of sequencing batch reactors 1 to 4.....	34
Figure 2.2. Tangential filtration flow setup to concentrate the viral fraction.....	38
Figure 2.3. Changes in particle size and SVI ₅ (a) and settling time (b) during granulation and microscopic images of the granulation process (c).....	42
Figure 2.4. Changes in biomass for the four sequencing batch reactors.	44
Figure 2.5. Changes in microbial communities of SBRs throughout the process of granulation.....	46
Figure 2.6. Changes in community richness and diversity during granulation.....	47
Figure 2.7. Relative abundance of the dominant bacteria groups during the granulation process.	49
Figure 2.8. The relative abundance of viral families during granulation.	51
Figure 2.9. Viral families that affect bacterial community changes.....	52
Figure 3.1. Homologous recombination by Lambda Red.	68

Figure 3.2. Relative primer annealing sites to screen for successful Pf6 knockout mutant.....	72
Figure 3.3. Schematic representation of the continuous flow reactor tubing biofilm system used.....	74
Figure 3.4. Transcriptomic analysis pipeline.	77
Figure 3.5. Expected PCR products generated from screening Pf6 prophage knockouts.	79
Figure 3.6. Screening for Pf6 knockout generated with Lambda RED mediated recombination.....	80
Figure 3.7. Phage production from biofilms formed by wild type PAO1 (a, b) and the PAO1 Δ Pf4 mutant (c, d).....	82
Figure 3.8. Clustering of samples through principle component plot (a) and heatmap showing sample-to-sample distances (b).	85
Figure 3.9. Early stage biofilms were more differentially expressed compared to late stage biofilms.....	101
Figure 3.10. Alignment of the Pf4 and Pf6 prophage genomes.	104
Figure 4.1. Sequence of site-directed mutagenesis of the repressor <i>c</i> region.	120
Figure 4.2. SNPs upstream of repressor <i>c</i> gene for <i>repC</i> EMSA probe.....	121
Figure 4.3. Sites of sequential site directed mutagenesis of <i>repC</i>	130
Figure 4.4. EMSA of RepC protein gradients on target DNA.	132
Figure 4.5. Immunity assay of repressor C.	133
Figure 4.6. Peak coverage plot of Pf4 and Pf6 RepC across PAO1 genome.	135
Figure 4.7. Predicted Pf4 and Pf6 RepC binding motifs.....	136

Figure 4.8. Validation of Pf4 RepC binding to selected target regions.	140
Figure 4.9. The complemented strains were more similar to the wild type....	145
Figure 4.10. The top 20 differentially expressed genes relative to PAO1ΔPf4.	154
Figure 4.11. Euler diagrams of differentially expressed genes in all three comparisons.	155
Figure 4.12. Mutation located within the conserved region.	158
Figure 5.1. Model for the role of repressor <i>c</i>	172

List of Abbreviations

ACLAME	A CLAssification of Mobile genetic Elements
AHL	Acyl-homoserine lactone
APHA	American Public Health Association
AOB	Ammonium oxidising bacteria
BALO	<i>Bdellovibrio</i> -like organism
COD	Chemical oxygen demand
d	Day
dbRDA	Distance-based redundancy analysis
DPAO	Denitrifying PAO
DO	Dissolved oxygen
EBPR	Enhanced biological phosphorus removal
EHEC	Enterohemorrhagic <i>E. coli</i>
eDNA	Extracellular DNA
EPS	Extracellular polymeric substances
GAO	Glycogen accumulating organism
h	Hour
LEE	Locus of enterocyte effacement
LCA	Lowest common ancestor
LB	Luria-Bertani
MAG	Metagenome-assembled genome
min	Minute
MLSS	Mixed liquor suspended solids
MLVSS	Mixed liquor volatile suspended solids
TLA	Modified top layer agar
NOB	Nitrite oxidising bacteria
PFU	Plaque forming unit
PHA	Poly-b-hydroxylalkanoate
PCR	Polymerase chain reactions
PAO	Polyphosphate accumulating organism
PS	Polysaccharide

PLC	Programmable logic controller
PN	Protein
QS	Quorum sensing
rlog	Regularized logarithm
RF	Replicative form
s	Second
SBR	Sequencing batch reactor
SND	Simultaneous nitrification and denitrification
SNDPR	Simultaneous nitrification-denitrification and phosphorus removal
SDM	Site-directed mutagenesis
SCV	Small colony variants
SDS	Sodium dodecyl sulphate
SOC	Super optimal broth with catabolite repression
SUAV	Superficial upflow air velocity
SI	Superinfective
TFF	Tangential flow filter
T3SS	Type III secretion system
T6SS	Type VI secretion system
VST	Variance stabilizing transformations
VFAs	Volatile fatty acids
WT	Wild type

Abstract

Bacteriophages can impact bacterial communities, not just through predation, but also through the regulation of host genes that can improve biofilm development. Bacterial communities are essential in the function of most ecosystems including engineered systems such as activated floccular sludges of water reclamation plants. Activated floccular sludge is often used to cultivate aerobic granules and the formation of such granules is normally achieved by selection through physical parameters. However, these granules are unstable and have long initiation times. It has been hypothesized that bacteriophages are important in the formation of granules. In this study, a metagenomics approach was used to monitor the abundance of bacteriophage and bacterial communities of granulating sludge. Analysis suggested that Inoviridae and Microviridae influence the bacterial community change the most during initiation of granulation. Additionally, the changes in the abundance of bacteria, especially '*Candidatus Accumulibacter*', and physical factors such as settling time and biomass discharge are also important in this process. To better define the impacts of filamentous phage on a bacterial host, a model phage-host system was chosen, based on *Pseudomonas aeruginosa* and its endogenous filamentous phages, Pf4 and Pf6. Transcriptomics revealed that in the presence of Pf4, virulence factor associated genes were upregulated, in agreement with the previous observation that the wild type *P. aeruginosa* is more virulent than the phage deletion mutant. This is likely due to a DNA-binding protein, repressor C, which was also found to prevent reinfection by Pf4. Collectively, this supports the hypothesis that filamentous phages might play an important role in biofilm development and begins to shed information on the role of filamentous phage in a more complex community, as found in aerobic granules.

Chapter 1 – Literature Review

1.1. Introduction to bacterial biofilms

Bacteria generally exist in a single-celled planktonic population or as immobile multicellular aggregates known as biofilms. Biofilms are multicellular aggregates adhered onto surfaces and are encased in a self-produced matrix called extracellular polymeric substances (EPS). The growth of bacteria in a biofilm can be described as being a lifecycle that includes transitions from individual, planktonic cells into mature mushroom-shaped microcolonies, which subsequently undergo dispersal to release individual bacterial cells back into the planktonic phase (Hall-Stoodley *et al.*, 2004).

Briefly, during the attachment phase, free-swimming planktonic gram-negative bacteria reversibly attach themselves to a surface using surface adhesins such as flagella or pili while gram-positive bacteria produce various biopolymers such extracellular DNA, proteins and lipids to mediate attachment (Abee *et al.*, 2011). These cells transition to irreversible attachment when EPS is secreted by the attached bacteria. The bacteria within the biofilm propagate, maturing the biofilm and forming microcolonies. At the final stage, the mature biofilm undergoes dispersal, where bacterial cells are released so that they can colonise other surfaces, repeating this life cycle (Costerton *et al.*, 1999; Monroe, 2007).

The EPS also mediates adhesion of the bacteria to surfaces, immobilising them, which keeps them close to each other and allows for interactions such as cell-cell communication and gene transfer. The EPS is

made up of a variety of macromolecules including polysaccharides, glycolipids, glycoproteins, nucleic acids, humic substances and water (Flemming *et al.*, 2007; Flemming & Wingender, 2010). It is within this EPS matrix that bacteria continue to propagate and expand the biofilm before bacterial cells are released during the dispersal phase.

Biofilm dispersal can either be an active or passive process. Active dispersal is a highly regulated process, initiated by the bacteria and triggered by changes in a range of environmental factors including temperature, oxygen, nutrients and metabolite (cyclic di-GMP, nitric oxide, quorum sensing molecules) concentrations (McDougald *et al.*, 2011). Differential expression of genes also contributes to active dispersal; cell motility and polysaccharide degradation associated genes such as flagella biosynthesis and dispersin secretion are upregulated, while EPS production genes are downregulated (McDougald *et al.*, 2011; Toyofuku *et al.*, 2016). In contrast, passive dispersal is a process of biofilm detachment that involves physical forces like hydrodynamic shear, abrasion, or mechanical intervention (Kaplan, 2010; Matilla, 2017).

1.1.1. Ubiquity of biofilms

In terms of total biomass as well as the variety and extent of colonised habitats, bacteria are one of the most successful forms of life on earth. As biofilms represent the dominant mode of life for bacteria, it follows that they are widespread in nature (Costerton *et al.*, 1995). Biofilms were first observed as dental plaque in humans (Gibbons & Nygaard, 1970). Since then, they have been observed in almost all environments, including wastewater, drinking water, the human gastrointestinal tract and even showerheads (Elliott *et al.*, 2006;

Ledder *et al.*, 2008; Min & Rickard, 2009; Simoes *et al.*, 2008; Vornhagen *et al.*, 2013).

From a human perspective, the ability of bacteria to grow as a biofilm can be either beneficial or detrimental. In the human oral cavity, there is a taxonomically complex microbial community. In this environment, colonisation follows a specific sequence where primary colonisers, *Streptococcus* spp., *Veillonella* spp. and *Actinomyces* spp. adhere to the tooth pellicle. This is followed by secondary colonisation by anaerobic gram-negative bacteria such as *Fusobacterium* spp. and *Porphyromonas gingivalis* through complex interactions with the primary colonisers (Diaz *et al.*, 2006; Palmer *et al.*, 2003; Periasamy & Kolenbrander, 2010; Ritz, 1967; Zijnga *et al.*, 2010). Dysbiosis in the oral microbiota and interactions with the immune system leads to inflammation and the onset of periodontal disease (Kinane *et al.*, 2017).

Chronic wounds have also been described to harbour biofilms. From the isolated bacteria and community composition observed from sequencing, the most common pathogens were *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus epidermidis* and *Enterococcus faecalis* (Bjarnsholt *et al.*, 2013; Chong *et al.*, 2017; Dowd *et al.*, 2008; Giacometti *et al.*, 2000; Goh *et al.*, 2017; Rahim *et al.*, 2016). Chronic wounds are difficult to treat and growth within biofilms promotes the survival and persistence of these pathogens as they avoid the host immune response (Chong *et al.*, 2017; Hall-Stoodley & Stoodley, 2009).

In various industries, such as shipping, food and oil/gas industries, the presence of biofilms also has negative implications. Food spoilage and

transmission of pathogens are some of the consequences of biofilms in food processing environments (Carpentier & Cerf, 1993; Van Houdt & Michiels, 2010). Fouling by biofilms in shipping and oil/gas industries cause problems, where they increase fuel consumption as well as costs of ship maintenance due to the contribution of microorganisms to corrosion (Hunsucker *et al.*, 2018; Schultz *et al.*, 2011; Swain *et al.*, 2007). For pipelines, the leakage of material running through the corroded pipelines, i.e. oil, can lead to environmental pollution (Vigneron *et al.*, 2018).

While biofilms have been shown to be a problem both medically and industrially, there are also benefits of this mode of microbial growth. Biofilms can be used in bioremediation to reduce and degrade contaminants in the environment which are potential risks to both the environment and humans. Global nutrient cycling is mediated by biofilms in soil, sediments and water (Edwards & Kjellerup, 2013). When comparing bioremediation between biofilms and planktonic cells, biofilm mediated bioremediation is more effective as the cells are encased and protected by the extracellular biofilm matrix. For example, *Pseudomonas stutzeri* T102 was used to degrade natural petroleum contaminated soils (Shimada *et al.*, 2012). It was observed that only cells in the biofilm survived after 10 weeks and naphthalene degradation was also higher in the biofilm.

The improved performance of a surface attached biofilm is also seen in water reclamation systems where it is preferred over loosely floating, activated floccular sludge. The increased thickness of biofilms that form on biofilm carriers contributes to the increased tolerance of such biofilms to loading fluctuations and keeps the discharged effluent from exceeding set discharge levels (Wilderer

et al., 2000). These support biofilm carriers, which can be either suspended or fixed, also helps improve settling times while maintaining the nutrient removal efficiency of activated floccular sludge (Arnold *et al.*, 2000; Wilderer *et al.*, 2000). For example, the polyethylene biofilm carriers used in moving bed biofilm reactors (MBBRs) provide a large surface area for bacterial cell attachment leading to biofilm formation. Thus, suspended carriers also allow for a compact configuration of treatment facilities while achieving a highly active biomass with a low settling time for downstream separation of solids (Biswas & Turner, 2012). An example of a fixed carrier is the use of plastic nets in aeration tanks which serve as substrates for bacterial attachment and biofilm growth. Using nets increases the overall biomass in the tank and in turn improves the nutrient removal efficiency of the system. This combination was shown to have increased the removal of biochemical oxygen demand and reduced the sludge settling time in synthetic wastewater (Gebara, 1999).

1.2. The use of activated sludge in water reclamation

In biological wastewater reclamation systems, the activated sludge process is one that utilises microorganisms to remove both inorganic and organic matter. Typically, these microorganisms exist as suspended biofilm aggregates, called flocs, and are responsible for nutrient removal in wastewater. Flocs are separated from the effluent by gravitational settling and produce high quality effluents which are relatively free of suspended solids and organic matter. However, flocs have long settling times due to their poor settling characteristics. As such, they require large aeration tanks and secondary clarifiers to retain sufficient biomass to achieve the targeted efficiency of nutrient removal

(Sadrzadeh & Dulekgurgen, 2014). To accommodate these requirements, activated sludge systems have a large spatial footprint.

Water reclamation plants that utilise the activated sludge process are usually operated for the removal of nutrients such as phosphorus and nitrogen prior to effluent discharge. The removal of these nutrients is important to prevent eutrophication of aquatic systems, which is associated with algal blooms and the excessive growth of other photosynthetic microorganisms (Cloete & Bosch, 1994; Toerien *et al.*, 1975). Currently, the two common wastewater reclamation technologies that are utilised to remove biological nutrients are enhanced biological phosphorus removal and simultaneous nitrification-denitrification and phosphorus removal.

1.2.1. Enhanced biological phosphorus removal

Enhanced biological phosphorus removal (EBPR) is a configuration that is applied to activated sludge systems for the removal of carbon and phosphorus without the use of chemical additives. This is achieved through alternating between anaerobic and aerobic conditions. The removal of biological phosphorus largely relies on polyphosphate accumulating organisms (PAOs) which accumulate phosphate intracellularly. PAOs assimilate volatile fatty acids (VFAs) in the anaerobic phase as poly- β -hydroxylalkanoates (PHAs), and the energy required for this process is obtained through a concurrent degradation of polyphosphate to get ATP and glycogen which produces the reducing agent NADH_2 (Figure 1.1). During the aerobic phase, PAOs utilise the accumulated intracellular PHA to provide energy for assimilation of phosphate from the

wastewater into intracellular polyphosphate. This cycle enables phosphorus removal from wastewater (Figure 1.1) (Wu *et al.*, 2014).

In EBPR bioreactors, the most commonly identified PAOs are '*Candidatus Accumulibacter phosphatis*' (Accumulibacter) (He *et al.*, 2007; He & McMahon, 2011) followed by microorganisms that belong to the genera *Actinobacteria* (Kawakoshi *et al.*, 2012). *Tetrasphaera* spp., while having physiological differences to Accumulibacter, have also been shown to be associated with phosphorus removal (Kristiansen *et al.*, 2013; Nielsen *et al.*, 2012).

Figure 1.1. Mechanism of nutrient removal in EBPR.

The metabolism of PAOs during the anaerobic and aerobic phases in activated sludge utilising the EBPR process. During the anaerobic phase, VFAs are taken up by PAOs from the wastewater for accumulation as PHAs while poly-Ps are released as phosphate after degradation. During the aerobic phase, PHAs are hydrolysed which provides the energy used in phosphate assimilation and eventual storage as intracellular poly-P. Adapted from Wu *et al.* (2014).

1.2.2. Simultaneous nitrification and denitrification

Nitrogen removal from wastewater is achieved through a two-stage process of nitrification and denitrification. In this process, ammonium oxidising bacteria (AOB) and nitrite oxidising bacteria (NOB) convert ammonium to nitrite and nitrate respectively during nitrification. Nitrate is subsequently reduced to nitrogen gas by denitrifying bacteria (Figure 1.2a) (Helmer & Kunst, 1998). However, a process known as simultaneous nitrification and denitrification (SND) has been demonstrated to perform both nitrification and denitrification concurrently in a bioreactor operated under conditions of low dissolved oxygen (DO) (Figure 1.2b) (Bertanza, 1997; Pochana & Keller, 1999).

Figure 1.2. Biological elimination of ammonium.

AOBs first oxidise ammonium to nitrite before NOB further oxidise nitrite to nitrate. The reduction of nitrate during denitrification produces nitrogen gas (a). For SND, ammonium is partially oxidised to nitrate within the oxic layer before it undergoes denitrification in the anoxic layer to form nitrogen gas (b).

During this process of SND, ammonium undergoes partial oxidation to nitrite before it is denitrified to nitrogen gas (Yoo *et al.*, 1999). This process has 40% lower chemical oxygen demand (COD) as compared to the two-step biological elimination of nitrogen, making it more energy efficient (Turk &

Mavinic, 1986). Interestingly, a novel group of PAOs known as denitrifying PAOs (DPAOs) have been shown to mediate denitrification in systems that utilise the anaerobic and anoxic EBPR process (van Loosdrecht *et al.*, 1998). These organisms are capable of SND coupled with the uptake of phosphate using the same COD (Kern-Jespersen *et al.*, 1994). Nitrate is utilised by DPAOs instead of oxygen as an electron acceptor despite being less energy efficient. This results in a biomass yield that is 20% lower than when oxygen is the electron acceptor (Murnleitner *et al.*, 1997). Despite this inefficiency, bench-scale reactors enriched with DPAOs display sufficient removal of phosphorus to be considered for commercial use (Peng *et al.*, 2006).

1.2.3. Simultaneous nitrification-denitrification and phosphorus removal

Recently, a process called simultaneous nitrification-denitrification and phosphorus removal (SNDPR) has been described that integrates the EBPR and SND processes and is a cost-effective and an efficient alternative for nutrient removal. The feasibility of this integrated process has been shown in both laboratory-scale systems (Meyer *et al.*, 2005; Zeng *et al.*, 2003) and in a water reclamation plant (Yang *et al.*, 2016). During the anaerobic phase, acetate is consumed, followed by the release of phosphorus. The phosphorus is then assimilated during the aerobic phase while ammonium is converted to nitrogen gas without nitrite or nitrate accumulation (Lemaire *et al.*, 2006). However, without the formation of both aerobic and anoxic zones in the microbial aggregates, the SNDPR process is unstable, resulting in poor nitrification and denitrification (Meyer *et al.*, 2005). To overcome this, the use of aerobic granules was suggested to be beneficial (de Kreuk *et al.*, 2005).

1.2.4. Aerobic granulation in biological wastewater reclamation

Aerobic granules are high-density, compact aggregates of suspended biofilms typically comprised of bacteria and fungi (Liu & Tay, 2004; Weber *et al.*, 2007). This dense and compact structure limits the transfer of oxygen, facilitating the formation of both aerobic and anoxic zones that are required for reliable SNDPR (de Kreuk *et al.*, 2005; Lemaire *et al.*, 2006). Several studies have successfully implemented aerobic granular sludge in sequencing batch reactors (SBRs) that operate and maintain SNDPR (Bassin *et al.*, 2012; Lu *et al.*, 2006; Yilmaz *et al.*, 2008). SBRs are fill-and-draw activated sludge systems where wastewater is added in batches to be treated before discharge and are operated in a predetermined sequence of operations ("Biological Processes," 2006; Khan *et al.*, 2011).

The formation of aerobic granules is typically achieved without the need for carriers and the granules form under specific selective conditions that increase the aggregation of flocs (Weissbrodt *et al.*, 2012). While flocs have an average particle size of between 30 to 100 μm , granules are defined by their dense and compact structures (Figure 1.3), having an average particle size between 100 μm to 2 mm (Barr *et al.*, 2010a). The density of the granules contributes to their fast settling and low sludge volumetric index, greatly increasing the dewaterability of the system (Lemaire *et al.*, 2008a). These features also reduce the chance of the biomass being washed out of the system, resulting in high amounts of active biomass while generating effluent with a high clarity and quality.

The factors described above contribute to a smaller spatial footprint (25% of the conventional treatment plant) as compared to systems that utilise flocs (de Bruin *et al.*, 2004). Aerobic granular sludge offers additional advantages over flocs, such as high tolerance to toxic or inhibitory compounds and the opportunity to use high organic loading rates (Adav *et al.*, 2008). Therefore, granules have been used successfully in the treatment of contaminants like phenols, heavy metals and chloroanilines in wastewater (Maszenan *et al.*, 2011).

Figure 1.3. Physiological differences of sludge during the process of granulation.

Representative microscopy images of sludge during the aerobic granulation process over a period of 11 weeks. The reactor was seeded with floccular sludge (a) and the formation of compact aggregates were seen from week 4 (b). These aggregates slowly mature and expand into granules by week 8 of reactor operation (c). The images were taken at 40X magnification (red bar, 200 µm). The images were kindly provided by Dr Chan Siew Herng (SCELSE).

Despite the successes in establishing aerobic granular sludge at both the pilot (Isanta *et al.*, 2012) and full (Pronk *et al.*, 2015) scale bioreactors, this technology is not without limitations. Aerobic granular sludge can be unstable and spontaneously disintegrate, leading to failure of the system. Some

contributing mechanisms that lead to this instability include the overgrowth of filamentous bacteria (Aqeel *et al.*, 2016; Liu & Liu, 2006) and the breakdown of the anaerobic core (de Kreuk *et al.*, 2007; Zheng *et al.*, 2006). These factors eventually lead to poor settling, reducing the dewaterability of the system and the subsequent washout of the biomass that can collapse the performance of the reactors.

1.2.4.1. Non-biological and biological factors impacting the formation of aerobic granules

As mentioned previously, floccular sludge can be gradually converted to aerobic granules when the reactor is operated under specific operational conditions. These are physical means that have been studied and implemented in lab-scale SBRs to exert a strong selection pressure for the dense, fast settling granules. The first operational parameter is the time allowed for settling to select for sludge that has good settleability. Short settling times of 5 min and below have been shown to more quickly and completely convert the flocs to highly dense granules (McSwain *et al.*, 2004; Qin *et al.*, 2004). Another key operational parameter is the amount of hydrodynamic shear force that is provided by the upflow sparging of air and nitrogen during the reactor operation. This is measured as the superficial upflow air velocity (SUAV). The formation of granules was reported to occur in SBRs that were operated at a high SUAV of $1.2 \text{ cm}\cdot\text{s}^{-1}$ and above (Tay *et al.*, 2001a, 2001b).

The formation of aerobic granules is also affected by biological factors. The presence of hydrodynamic shear promotes the formation of a more tightly bound biofilm (Mohle *et al.*, 2007). Since they have apparent similarities in

formation and development, factors that drive biofilm development and dispersal may also apply to granules. The presence of hydrodynamic shear force stimulates the production of EPS. The EPS of aerobic granules consists of organic compounds like polysaccharides, proteins, lipids, humic-like substances and nucleic acids (Wang *et al.*, 2006). These substances mediate the adhesion between cells and promote the initial formation of microbial aggregates (Tay *et al.*, 2001c). Aerobic granulation was shown to be coupled to the ratio of polysaccharides (PS) to protein (PN). When the reactor was operated at low SUAV, there was also a low PS/PN ratio of 3, while a PS/PN ratio between 9 to 15 was observed when operated at high SUAV (Tay *et al.*, 2001a, 2001c). This suggests the role of EPS production in the process of aerobic granulation.

Another factor that has been established to affect the production of EPS in aerobic granules is quorum sensing (QS) (Ren *et al.*, 2010; Tan *et al.*, 2014; Wang *et al.*, 2017). QS in bacteria is due to small biochemical signalling molecules called autoinducers. Autoinducers are constitutively produced and released into the environment, and the concentration of these autoinducers are highly dependent on the density of the cells as well as the local diffusion parameters (Hense *et al.*, 2007). These autoinducers include autoinducing peptides produced by Gram-positive bacteria and acyl-homoserine lactones (AHLs) produced by Gram-negative bacteria. It has been demonstrated that EPS production and granulation was positively correlated with the concentration of AHLs (Tan *et al.*, 2014). Interestingly the PS/PN ratio of the sludge EPS was also positively correlated with the amount of AHLs present in the sludge. The exogenous addition of various AHLs increased the amount of PS by 14-36% and PN by 7-16% after 1 h of incubation as compared to before addback (Tan

et al., 2014). Additionally, more than half of the top 50 bacterial taxa, as revealed by total RNA sequencing of the sludge biomass, have a positive correlation with at least one of the AHLs. Some of the genera identified were *Nitrosomonas*, *Burkholderiales* and *Xanthomonadaceae*, and these have been demonstrated to be AHL producers (Tan *et al.*, 2014).

The development of aerobic granules has been suggested to be analogous to surface associated biofilms albeit with subtle differences (Ren *et al.*, 2010). The presence of EPS (Liu & Wang, 2007), cell-cell contact and adhesion (Liu & Tay, 2004) are examples of the similarities between surface associated biofilm and granules, as are dispersal and maturation. One of the proposed mechanisms for aerobic granulations states that post inoculation, fungi or filamentous bacteria dominate the system and form networks of hyphae that develop into mycelial pellets that act as a scaffold (Beun *et al.*, 1999) (Figure 1.4). The scaffold is a substratum for attachment and accumulation of other microorganisms that would otherwise be discharged from the system. The flocs eventually grow large enough that the inner regions become oxygen deprived, leading to the disintegration of the flocs and release of small aggregates or pre-granules (Beun *et al.*, 1999). Under high hydrodynamic shear in a sequencing batch reactor (SBR), these aggregates develop into mature granules (Liu & Tay, 2002; Liu *et al.*, 2005).

Figure 1.4. The developmental stages of a microbial granule (a) and its similarity to the biofilm life cycle (b).

a) Fungi with higher settling properties dominate the system after inoculation. The network of hyphae forms a mycelial pellet that acts as scaffolding for the population of bacterial colonies. Developing flocs eventually grow to a size where there is oxygen deprivation in the inner regions, eventually lysing the flocs to give rise to small aggregates that develop into compact and dense granules, which are mainly bacterial cells. The figure was adapted from Beun *et al.* (1999).

b) In the biofilm life cycle, the inoculated cells bind reversibly to the surface at the initial stage (b1), followed by secretion of exopolysaccharides into the surrounding environment causing them to bind irreversibly (b2). The young biofilms then undergo the early stage of biofilm development (b3); followed by the late stages of biofilm maturation into a defined, mushroom shape-like 3D structure (b4). Finally, the mature biofilm disperses, releasing free swimming cells to colonise new surfaces (b5). Adapted from Monroe (2007).

1.3. Predators of bacteria in the environment and in water reclamation systems

The presence of other microorganisms, such as bacteriophages and protozoa, can also affect the formation and stability of aerobic granules. Bacteriophages have been suggested to be responsible for the collapse of a granular EBPR reactor (Barr *et al.*, 2010b). Protozoans, particularly the sessile ciliates, have also been frequently observed in high abundance on mature granules (Lemaire *et al.*, 2008a). These ciliates have also been suggested as a substratum for the attachment of bacteria, which contributes to the initial formation of compact aggregates and the eventual increase in granular size (Weber *et al.*, 2007). Predatory bacteria have also been found to reduce microbial activity and the total biomass of aerobic granular sludge (Feng *et al.*, 2017). These suggests that the presence of predators could have an impact on granulation.

Predation is described as the biological interaction between predator and prey organisms. Bacterial predators are considered to be the primary drivers of mortality and changes in the bacterial community of all ecosystems as it exerts a strong selective pressure on the community. Predator and prey population dynamics affect both the phenotypic and taxonomic structure of these communities (Jürgens, 2007; Lima, 1998). Common bacterial predators in the environment include protozoans, predatory bacteria and bacteriophages.

1.3.1. Protozoa

Protozoa are unicellular, heterotrophic microorganisms of the kingdom Protista (Yaeger, 1996). Protists are ubiquitous in aquatic environments as well

as in soils and sediments. In these environments, protozoa control the abundance of bacteria by predation, contributing to nutrient cycling (Barcina *et al.*, 1991; Ronn *et al.*, 2002; Trap *et al.*, 2016).

Protozoa are generally classified by morphology as flagellates, ciliates or amoeba (Madoni, 2003). Flagellates are distinguished by the presence of one or more flagella which are used for movement, feeding and attachment to surfaces (Figure 1.5a). Many species of flagellates are heterotrophic and feed on bacteria or take up soluble organic molecules as sources of food. They are highly abundant and are the main consumers of bacteria in aquatic environments (Patterson, 2001). Amoebae have pseudopodia that extend out from the cell periphery (Anderson, 2010) (Figure 1.5b). The pseudopodia contracts and extends which allow the amoeba movement on surfaces as well as a means of engulfing prey. The ciliates are characterised by features such as sexual conjugation, nuclear dimorphism, the presence of cilia during certain stages of their life cycle (Figure 1.5c) and are the largest group of protozoa (Anderson, 2010; Lynn, 2012). Cilia that are close to the oral region assist in the capture of prey and are also responsible for drawing in water through the creation of feeding currents.

In activated floccular sludge, protozoa can constitute approximately 5% of the total biomass (Madoni, 2003). Grazing by protozoa can help clarify the effluent through the removal of planktonic bacteria and bacteria loosely attached on sludge particles (Madoni, 2011). Ciliates are particularly important as the increase in their numbers have been demonstrated to positively correlate with good effluent quality (Esteban *et al.*, 1991). In response to protozoan predation, bacteria employ several defence mechanisms like acquiring an oversized

morphology, increased motility, as well as the formation of biofilms (Matz & Kjelleberg, 2005). Despite the increase in grazing pressure due to the biofilm stationary mode of growth, better cell-cell communication is achievable, allowing more complex antipredator mechanisms like achieving a size that is beyond the protozoa's feasible prey size and QS-mediated upregulation of toxic factors like rhamnolipids and cytotoxins.

Figure 1.5. The morphological groups of protozoa.

A flagellate (a), amoeba (b), and ciliate (c). This figure is reproduced from Anderson (2010).

1.3.2. Predatory bacteria

Predatory bacteria are those that kill and consume other microbes. The best studied are the *Bdellovibrio* and the *Bdellovibrio*-like organisms (BALOs), which feed by penetrating the bacterial periplasm and reproducing intracellularly. The *Bdellovibrio* spp. are Gram-negative bacteria with high motility that target other Gram-negative bacteria (Stolp & Starr, 1963). When *Bdellovibrio* spp. collide with prey bacteria, they attach and penetrate the periplasm to form

bdelloplasts. Hydrolytic enzymes are then synthesised to degrade the host cellular contents to provide nutrients for the growth of the predator (Rendulic *et al.*, 2004; Sockett, 2009). The bacteria grow as filaments within the bdelloplasts and when the resources within the bdelloplasts are exhausted, they septate into several progeny cells. The progeny lyse the host and are released into the environment where the lifecycle repeats (Sockett, 2009).

In activated sludge, the strong predatory selective pressure that is exerted by these predatory bacteria can affect the bacterial community, possibly disrupting reactor performance. *Bdellovibrio bacterivorous* is a BALO strain that was isolated from activated sludge and was observed to predate on Gram-negative bacteria, including those of the phylum *Proteobacteria* and *Bacteroidetes* (Feng *et al.*, 2017). Ammonia-oxidising and phosphate-accumulating bacteria are two such groups in those phyla and are essential for the nutrient removal processes in wastewater reclamation.

1.3.3. Bacteriophages

Bacteriophages are the largest of all virus groups known to date (Ackermann, 2003) and have been found in the genomes of most sequenced bacteria (Canchaya *et al.*, 2003b). Bacteriophages are prokaryotic viruses that can be either double- or single-stranded DNA or RNA (dsDNA, ssDNA, dsRNA, ssRNA). They may be present as lysogenic prophages in hosts or as free phage particles in the environment. Lytic bacteriophages are virulent phages that inject their nucleic acid into the host cell and eventually cause death of the host by cell lysis, resulting in the release of phage progeny. Phages are known to alter competition among bacterial species (Joo *et al.*, 2006; Koskella *et al.*, 2012),

mediate horizontal gene transfer (Canchaya *et al.*, 2003a) and influence bacterial diversity (Buckling & Rainey, 2002; Rodriguez-Valera *et al.*, 2009).

The temperate and lytic lifestyles of phages are transmission strategies to exploit their prokaryotic host cells for their own reproduction and vary as a function of the bacterial density. When more bacteria are available, strategies that emphasise rapid infection of a host, such as short periods of latency, can be favoured. In contrast, when bacterial numbers are low, phage may survive as lysogens or by producing larger burst sizes during host cell lysis (Abedon, 2011). While phages can infect planktonic cells, they also infect bacteria in biofilms, which is not surprising given that most bacteria in the environment are present as biofilms and because biofilms typically contain high densities of bacterial cells (Hall-Stoodley *et al.*, 2004). Thus, biofilms are an ideal niche for the propagation of phage.

It has been shown that the relationship between phage and biofilm associated host bacteria is important for biofilm differentiation and also facilitates dispersal of the surviving cells (Webb *et al.*, 2004; Webb *et al.*, 2003). Even cryptic prophage can play a role in controlling the expression of genes that regulate the fitness of its host. For example, CP4-57 in *E. coli* biofilms excises during the early stages of biofilm formation and results in the development of a diversified population with specialised functions in terms of nutrient metabolism and motility (Wang *et al.*, 2009).

Bacteriophages appear to be active components of activated sludge systems and phage-mediated mortality has the potential to influence the treatment performance of the system by controlling the abundance of key

functional groups. For example, Wanner *et al.* (2000) attributed the failure of nitrification in a sludge system to a low bacterial concentration or the absence of nitrifying bacteria, as a consequence of the activity of lytic bacteriophage infection of those functional groups. An unexpected decline in phosphorus removal was also attributed to a phage infection of '*Candidatus Accumulibacter phosphatis*' (Barr *et al.*, 2010b). In the same study, the addition of supernatant from the failing reactor to a working reactor caused a decrease in the abundance of *Accumulibacter* in both granular and floccular activated sludges. Furthermore, after the phage attack, the numbers of *Accumulibacter* in the granular sludge failed to recover (Barr *et al.*, 2010b).

1.3.3.1. Bacteriophages and their potential roles in water reclamation systems

Given the abundance of bacteriophages in nature and in the genomes of most microbes as prophages, together with their role in modulating biofilm development (Rice *et al.*, 2009; Wang *et al.*, 2009; Webb *et al.*, 2004), it is likely that bacteriophages also play important roles in granulation. This can be either through selection for granule forming organisms or the evolution of phage infection resistant strains, where resistance is mediated by changes in surface proteins, which are the most common receptors for phage infection.

Bacteriophage and their host can be considered to be in a continuous evolutionary race where the bacteria host modifies itself to prevent attack from lytic phages, and the phages change to ensure their ability to infect a host to continue propagation. The dynamics between a phage and its host can be described as a "kill the winner" model where the rapid propagation of phages is

dependent on the host taxa becoming abundant in the environment (Suttle, 2007; Thingstad, 2000). The dominant bacteria in an environment is usually the ideal host for a phage that is enriched. Once the critical threshold to support replication of the phage is reached, rapid infection of the population occurs, resulting in host lysis and the numbers of the host rapidly declines. As a consequence of reduction of numbers of the dominant bacteria, other bacterial species can then exploit the now available resources and thus, will increase in number and hence, will dominate the niche. A study on *Nitrosospira multiformis*, a model AOB, showed that an increase in viral production due to exposure to stressors reduced bacterial counts and resulted in a failure to denitrify (Choi *et al.*, 2010).

In the model organism *P. aeruginosa* PAO1, the role of bacteriophage in the formation of biofilms has been intensively studied. PAO1 has been demonstrated to carry the filamentous phages Pf4 (Rice *et al.*, 2009; Webb *et al.*, 2004) and Pf6 (Tay, 2008) in their genome. The Pf4 phage has been shown to be responsible for cell death and dispersal, and the formation of morphotypic variants within the biofilm. Pf4 has also been demonstrated to stabilise the biofilm and confer resistance to surfactants such as sodium dodecyl sulphate (SDS) (Rice *et al.*, 2009).

In part, the phage may play a physical role in protecting the cells within the biofilm. For example, the PAO1 biofilm surface contains a high density of Pf4 phage (as determined with anti-Pf4 antibodies) that form a dense lattice (Webb *et al.*, 2004) (Figure 1.6a). Similar structures were also observed in aerobic granules using scanning electron microscopy (SEM) (Liu & Wang, 2007) (Figure 1.6c). More recently, it has been shown that the presence of Pf4,

together with biopolymers in the biofilm matrix, self-assemble into highly ordered liquid crystals (Figure 1.6d, e) that enhance biofilm adhesion and tolerance to desiccation (Secor *et al.*, 2015). Although the composition of the observed structures in the granules is largely unknown, there is the possibility that aggregates of filamentous phages are a part of the granule, as they are for the surface associated biofilms.

Figure 1.6. Microscopic images of *Pseudomonas aeruginosa* biofilms and aerobic granules.

Transmission electron microscopy images of *P. aeruginosa* cells labelled with anti-Pf4 immune-gold tagged antibodies specific for Pf4 filamentous phage (a, b), adapted from Webb *et al.* (2004). Scanning electron microscope image of a granule from a sequencing batch reactor (c), adapted from Liu and Wang (2007). Microscopy image of a *P. aeruginosa* biofilm as highly ordered liquid crystals in the presence of phage and microbial polymers (d, e), adapted from Secor *et al.* (2015).

1.4. *Pseudomonas aeruginosa* biofilms

The development of *P. aeruginosa* biofilms has been extensively studied (Hoiby *et al.*, 2001; Sauer *et al.*, 2004; Webb *et al.*, 2004; Webb *et al.*, 2003;

Whiteley *et al.*, 2001). This microorganism is a Gram-negative, opportunistic pathogen that is ubiquitous in the environment, making it an ecologically significant organism. *P. aeruginosa* is found in various settings such as water reclamation (Grobe *et al.*, 1995) and food preparation (Thanomsub *et al.*, 2006), homes (Finch *et al.*, 1978), swimming pools (Rice *et al.*, 2012) and plants and soil (Elrod & Braun, 1942). *P. aeruginosa* is also clinically important as it is responsible for biofilm-related infections and is best known for infection of the lungs of patients with cystic fibrosis (Bjarnsholt *et al.*, 2009; Sadikot *et al.*, 2005). Based on its environmental, industrial and clinical significance, *P. aeruginosa* is often used a model organism for the study of biofilms.

1.4.1. Role of bacteriophage in biofilm development, dispersal, and variant formation

The Pf4 filamentous phage is present as a prophage in *P. aeruginosa* PAO1 and continuously replicates without killing the host (Webb *et al.*, 2003). Furthermore, the PAO1 host is protected from reinfection by the phage due to the presence of the prophage. However, during biofilm development, the Pf4 phage has been shown to convert into a superinfective (SI) form where the phage becomes able to reinfect its host (Rice *et al.*, 2009). Interestingly, the appearance of superinfective Pf4 can be correlated to two key events during biofilm development; cell death and dispersal, and the appearance of genetic variants in the biofilm population.

Deletion of the Pf4 prophage from the host genome demonstrated that the mutant no longer underwent cell death and formed hollow colonies. The biofilms of the Pf4 knockout mutant were also less stable (being sensitive to

surfactant treatment) and less virulent than the PAO1 wild type (WT) when tested in an acute lung infection mouse model (Rice *et al.*, 2009). Furthermore, biofilms formed by the Pf4 mutant did not generate morphotypic variants, e.g. small colony variants (SCV), unlike PAO1 WT biofilms (Rice *et al.*, 2009). This demonstrates that prophages play an integral role in biofilm development of PAO1. While it is clear that phage play an important role in biofilm development, the specific mechanisms of how Pf4 mediates these effects are currently not known.

The contribution of bacteriophages to the virulence and other properties of the host is not limited to *P. aeruginosa*. Specific phage products have been shown to have an impact on their host (Canchaya *et al.*, 2004). In *E. coli*, the *bor* and *lom* genes encoded by a prophage confer resistance to the lysogen during growth (Barondess & Beckwith, 1990). In *E. coli* O157, the Shiga-like toxin carried by the lambdoid phage induces damage in the human gut during iron deficiency (Wagner *et al.*, 2002). Thus, phage mediated horizontal gene transfer can convert a non-pathogenic strain of *E. coli* into a highly virulent strain. Elucidation of how phage conversion happens and the molecular mechanisms of the contribution of Pf4 to biofilm development will enhance the understanding of its role in modifying host behaviours. The role of filamentous phage in the biology of *P. aeruginosa* is very complex as it was recently shown that this bacterium encodes a second filamentous phage (later designated as Pf6) that was found in a Pf4 knockout strain (Tay, 2008). This was unexpected since *P. aeruginosa* PAO1 was previously thought to harbour only one filamentous phage. It is also somewhat rare for a host to have two closely related

bacteriophages in their genome, which suggests that Pf6 (or its gene products) could make significant contributions to its host.

1.4.1.1. The bacteriophage lytic and lysogenic switch

Since Pf4 phage activity and superinfectivity contributes to PAO1 biofilm development and variant formation, it is important to understand the mechanisms of bacteriophage infection and replication, as well as how the bacterium gains immunity to re-infection. Bacteriophage infection has been shown to be either lytic or lysogenic. The lytic cycle is when the host cell lyses during phage release. In contrast, the lysogenic cycle is when the bacteriophage exists within a bacterial host as an extrachromosomal element or a prophage without resulting in the death of the host (Brussow *et al.*, 2004).

Phage lambda is one of the best studied phage and has been shown to exist as a lysogenic prophage in *E. coli* (Lesic & Rahme, 2008; Oppenheim *et al.*, 2005; Roberts & Roberts, 1975). The mechanism of the lysogenic to lytic switch involves a bistable genetic switch (Figure 1.7). During the lysogenic life cycle, the repressor C (encoded by the *cI* gene) protein binds to the operator site of the early promoters, P_R (*cro* promoter) and P_L , and regulates the *cI* repressor gene expression from the P_{RM} promoter through a positive autoregulatory loop (Nickels *et al.*, 2002). This represses the expression of the lytic phage *cro* genes and the prophage does not excise and reproduce. However, in response to DNA damage and activation of the SOS response, the repressor C protein is cleaved by RecA, inactivating repressor C and allowing transcription of the *cro* genes (Oppenheim *et al.*, 2005; Roberts & Roberts, 1975; Roberts *et al.*, 1978). Cro is an essential lytic regulator that represses *cI* from

the P_{RM} promoter and also indirectly lowering the levels of CII (which activates expression of *cI*) (Johnson *et al.*, 1978; Oppenheim *et al.*, 2005). Mutations in the *cI* promoter also leads to the activation of the lytic cycle, demonstrating the importance of the Cro protein in repressing the P_{RM} promoter. In other bacteriophage, for example those that infect *Xanthomonas campestris* pv. *citri* or *Lactobacillus*, the loss of the repressor protein can also lead to induction of the lytic cycle (Cheng *et al.*, 1999; Ladero *et al.*, 1998).

Figure 1.7. The bistable genetic switch of lambda phage.

During the lysogenic cycle (green arrows), the *cI* gene encodes for the repressor C protein (λ) that represses the transcription of the Cro protein (C) when it binds to the operator site O_{R1} and O_{R2} of the *cro* promoter (P_R). However, during the lytic cycle (red arrows), the *cro* gene encodes the Cro protein, which blocks the transcription of the repressor C protein (λ) through binding to the operator site O_{R3} of the *cI* promoter (P_{RM}). This figure was adopted from Hasty *et al.* (2001).

The repressor C protein also confers immunity to the host against infections by phage containing a closely related repressor. If a related phage

enters the lysogenised host, repressor C binds to the lytic operator site and suppresses expression of the lytic phage genes (Ptashne, 2004). Repressor C protein homologues have also been identified in other phages, e.g. the Pf4 phage of *P. aeruginosa* and P2 phage of *E. coli* (Ljungquist *et al.*, 1984; Rice *et al.*, 2009). Unlike the lambda phage, the Pf4 prophage continually produces phage particles without resulting in the death of the host. For Pf4, cell death is specifically associated with superinfection. In this way, the filamentous phage shares features of both temperate and lytic phage. However, the specific mechanism of killing is unknown and the role of repressor C in controlling the superinfection is unknown (Rice *et al.*, 2009). More recently, it was observed that the superinfective phages produced by two morphotypes of PAO1, have mutations involving the repressor C homolog. In a small colony variant, there was a 3,410 bp deletion in the 5' end of Pf4 that contained the repressor *c* gene; and in a spreader variant, 4 single nucleotide mutations were found within and upstream of the repressor *c* gene (Hui, 2013). The repressor C protein seems to mediate host immunity to the Pf4 filamentous phage, although it is likely to function differently from what has been described in lambda phage.

1.5. Thesis aims

Bacteriophages have a significant role in facilitation of horizontal gene transfer and controlling the abundance of bacteria through lysis in the environment. They can also have other effects on the host like controlling growth rate, biofilm formation and formation of variants. The contributions of bacteriophage are significant in microbial ecology and evolution, affecting the function of ecosystems where bacteria are the drivers of major biogeochemical cycles. In activated sludge, the presence of bacteriophage has been implicated

in the collapse of working bioreactors and phage-like structures have been observed in an electron micrograph of a granule (Barr *et al.*, 2010b; Liu & Wang, 2007). Therefore, we hypothesize that bacteriophages play similar roles in activated sludge, and that the changes in bacteriophage diversity influence the changes in bacterial communities of granulating reactors.

P. aeruginosa PAO1 and the Pf4 filamentous phage is a defined system for understanding host-phage interactions through the use of a genetically amenable host and a class of phage which can modulate its host gene expression. The presence of dense lattice structures in sludge, similar to ones observed in a *P. aeruginosa* biofilm, allows for a bottom up approach for studying the intricacies of the interaction between a host and filamentous phage (Liu & Wang, 2007; Webb *et al.*, 2004). In *P. aeruginosa* PAO1, it has been shown that the Pf4 filamentous phage has a significant contribution to biofilm development (Rice *et al.*, 2009). However, the role of the Pf6 prophage in PAO1 biofilm development has yet to be studied. Previous studies suggest that the superinfective form of the Pf6 phage also seems to be linked to essential stages of the biofilm development, including cell death (Tay, 2013). Therefore, it is essential to understand how the phage mediates these effects and to understand the mechanism that allows the Pf phage to become superinfective. Therefore, the specific aims of the thesis chapters were:

Chapter 2: To determine the role of phages in the establishment, maintenance and collapse of aerobic granules. Metagenomic sequencing was used to characterise the bacterial communities in floccular and granular sludge. At the same time, the dynamics of the virome present in the reactor effluent was also elucidated to correlate viro-types with granulation.

Chapter 3: To determine the role of the Pf filamentous phages in the development of a single species *P. aeruginosa* biofilm. Differential gene expression analysis was undertaken to determine if the Pf4 phage plays a role in the regulation of host genes.

Chapter 4: To study the role of the repressor C protein homolog in the Pf4 phage and its possible role in the control of superinfection and immunity to reinfection. ChIPseq was performed to determine the potential regulon of the repressor C. Additionally, differential expression analysis on the repressor C complemented Pf4 phage-deficient strain was undertaken to determine the potential role of the protein in controlling host gene expression.

Chapter 2 – The role of bacteriophage in the formation of aerobic granules

2.1. Introduction

Most current large-scale water reclamation plants are based on floccular sludge systems. To improve the economics and quality of treated water in water reclamation, there is a need to develop and better understand approaches that can perform waste and nutrient removal at a higher efficiency. Aerobic granular sludge is a complex, man-made ecosystem that consists of diverse microbial communities, phylogenetically and functionally, that are used for water remediation. The granules are densely packed biofilm aggregates that are developed from floccular sludge and have shown promise for efficient wastewater reclamation (Singh & Srivastava, 2010; Slater, 2006). The process of aerobic granule formation from aerobic flocs has been improved, mainly through an increased understanding of the impact of physiochemical factors such as hydrodynamic shear force, hydraulic retention time, settling time and discharge time (Liu, 2008; McSwain *et al.*, 2004; Qin *et al.*, 2004; Tay *et al.*, 2001a). However, despite this increased understanding of physical factors, many limitations remain due to slow start-up and granule instability in full-scale reactors (Wilén *et al.*, 2018). Given that biological processes are also likely to drive granulation, a better understanding these processes could improve the process of aerobic granulation.

The development of aerobic granules, a suspended form of biofilm, is analogous to surface associated biofilms (Ren *et al.*, 2010). As such, the biological factors that drive granulation share some common mechanisms as

observed for surface associated biofilms. For example, N-acyl-homoserine lactone (AHL)-mediated quorum sensing was positively correlated with extracellular polymeric substance (EPS) production of surface associated biofilms (Koutsoudis *et al.*, 2006; Molina *et al.*, 2005) and granules (Ren *et al.*, 2013; Tan *et al.*, 2014). In addition, pili facilitated adhesins also play a role in biofilm formation, where the interaction of fimbriae with specific ligands promotes adhesion to both abiotic and epithelial surfaces (Korea *et al.*, 2010; Korea *et al.*, 2011).

However, the role of bacteriophages, which have been implicated in biofilm formation has not been well studied in relation to granule formation. In *Streptococcus pneumoniae*, pneumococcal strains that carried lysogenic phage showed enhanced biofilm formation compared to non-lysogenised strains due to the release of extracellular DNA from the spontaneous activation of a prophage that causes cell lysis in a subpopulation of the biofilm (Carrolo *et al.*, 2010). In contrast, for *Streptococcus pyogenes* biofilms, the endolysin PlyC that is encoded in the bacteriophage C₁ genome, was shown to play a role in biofilm disruption through degradation of the biofilm matrix (Shen *et al.*, 2013). In sludge systems, an increased amount of viral-like particles was correlated with a reduction in phosphate removal in an '*Candidatus Accumulibacter*' enriched EBPR reactor (Barr *et al.*, 2010b). Thus, bacteriophages could potentially have a constructive or destructive role in both surface associated biofilms as well as in suspended biofilms such as granules. Despite this, there have been few reports to correlate the shift in bacterial communities with changes in the bacteriophage communities. More specifically, the dynamics of the different

bacteriophage families throughout the process of aerobic granulation have also not been studied.

Here, as part of an ongoing study, four sequencing batch reactors (SBRs) were seeded with activated floccular sludge and operated for simultaneous nitrification-denitrification and phosphorus removal (SNDPR) and under conditions that favour the formation of aerobic granules. Previously, the reactors were used to study granule formation till their collapse, and have correlated certain quorum sensing signals with the appearance of structured granules (Tan *et al.*, 2014). For this study, the aim was to take a more detailed study of the initiation of granulation. Using total metagenomic sequencing of the activated sludge and effluent collected over the course of granulation, both the bacterial and bacteriophage communities were characterized to investigate the potential role of bacteriophages during this process.

2.2. Materials and methods

2.2.1. Sequencing batch reactor operation

Four SBRs were inoculated with activated floccular sludge from the Ulu Pandan Water Reclamation Plant (Figure 2.1). All of the reactors were operated to perform simultaneous nitrification-denitrification and phosphorous removal (SNDPR) (Zhou *et al.*, 2010) and were operated for a total of 11 weeks. Each SBR had a final working volume of 2 L and was operated at 22°C. The operation was controlled using a programmable logic controller (PLC) and operated in a 6 h cycle comprising 2 phases: Phase 1 – feeding (8 min), anaerobic (60 min), aerobic (80 min at week 0 with gradual increase to 95 min by week 5), and

anoxic (40 min at week 0 with gradual increase to 50 min by Week 5); as well as Phase 2 – feeding (2 min), anaerobic (30 min), aerobic (40 min at week 0 with gradual increase to 70 min by week 5) followed by an anoxic stage (30 min). The cycles were completed with a settling stage (120 min at week 0 with a gradual decrease to 5 min at the end of week 6 onwards) followed by a discharging stage (10 min).

Figure 2.1. Setup of sequencing batch reactors 1 to 4.

The four sequencing batch reactors (1 to 4, from left to right) were seeded with floccular sludge from Ulu Pandan Water Reclamation Plant. The reactor operation was controlled and monitored by the programmable logic controller which also controls the feeding and discharge through the peristaltic pumps as seen in the setup above.

During the feeding stages, 1 L of synthetic wastewater (150 mg COD/L acetate and 50 mg COD/L propionate as carbon source, 10 mg/L orthophosphate, 20 mg/L ammonium and trace nutrient elements (Smolders *et*

al., 1994) was fed to each SBR, and 1 L of treated effluent was discharged at the end of each cycle. The synthetic wastewater was prepared as described by Zhou *et al.* (2010). The pH and dissolved oxygen (DO) levels were monitored using specific probes connected to the PLC. Throughout the cycles, nitrogen was sparged intermittently into the reactors at a flow rate of 1 L/min except the settling and discharging stage. To maintain the DO levels between 3 to 4 mg/L during the aerobic phases, air was sparged into the reactor. The pH of every reactor was maintained between 6.8 - 7.2 through PLC based monitoring and automatic dosing with 0.1 M HCl or 0.1 M NaOH as appropriate throughout the operation.

2.2.2. Sampling and performance studies of the sequencing batch reactors

A cycle study was conducted weekly to evaluate the reactor performance and the granulation process as part of an ongoing study. For the purposes of this study, the key operational parameter was that the system was moving towards granulation and to have a measure of biological function during operation. Therefore, the percentage removal of phosphorus and nitrogen was given here, and not the full biochemical measurements associated with SNDPR characterization. The mixed liquor suspended solids (MLSS) and mixed liquor volatile suspended solids (MLVSS) of the biomass, together with ammonia, nitrite, nitrate and orthophosphate concentrations were quantified using the American Public Health Association (APHA) standard engineering methods (Eaton *et al.*, 2005). Sludge density and compactness were measured by the sludge volumetric index at 5 min of settling (SVI₅) as described by Liu (2008). The particle size of the biomass was determined using a laser diffraction particle

size analyser (SALD-3101, Shimadzu, Japan) and the morphology was observed using a bench top microscope (Carl Zeiss, Germany).

Effluent from each reactor was discharged into their respective containers prior to transferring into 25 L carboys. This was done to collect the viral fraction to study the bacteriophage dynamics in the SBR. Samples were collected from weeks 1 to 10 of the study. Prior to the transfer, the biomass was allowed to settle for at least 16 h before passing through a 25 µm filter bag (Puridea, Singapore) into the carboy to remove any suspended biomass not separated by settling. The filtrate was stored at 4°C during collection and pooled over 4 d to obtain 20 L of filtrate, at which time, 2 mL of DNase I (200 U/mL) (Calbiochem, USA) was added to the samples to digest any extracellular DNA (eDNA).

2.2.3. Concentration of the viral fraction

The 20 L of filtered effluent was processed to remove bacteria by passing through a 0.2 µm tangential flow filter (TFF) before being concentrated using a 100 kDa TFF. The samples were filtered through a 0.2 µm Sartocon Slice Disposable TFF (Sartorius Stedim, Germany) using an automated system previously developed and optimised in-house (Tay, 2013) (Figure 2.2) and the filtrate was collected in a clean carboy as the permeate. Any particles larger than 0.2 µm were pumped back into the unfiltered reservoir as retentate, while the permeate was used in the subsequent concentration step (next paragraph below). Between samples, the system was treated with 0.1 M NaOH and then flushed with tap water to neutralise the system (Figure 2.2a) as part of the cleaning process to avoid cross-contamination.

The permeate was then used to concentrate the viral fraction using a 100 kDa Sartoclon Slice Disposable TFF unit (Sartorius Stedim, Germany) (Figure 2.2b). The fluid phase and particles smaller than 100 kDa were removed while retaining the viral fraction in the reservoir, concentrating the fraction as the volume decreased. To elute the viral fraction, SM buffer (100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-Cl at pH 7.5) was added to the phage reservoir. The volume of the viral fraction after elution was reduced from 20 L to 100 mL. To minimise cross-contamination between samples, the 100 kDa TFF was treated with 0.5 M NaOH for 1 h before flushing with 10 L of deionised water to neutralise the system.

The volume of the concentrated viral fraction was further reduced using Vivaspin 20 100,000 MWCO Centrifugal Concentrators (Sartorius Stedim, Germany). The samples were concentrated to a final volume between 2 to 3 mL after multiple rounds of centrifugation. For each round, 20 mL of samples was added to their respective column and centrifuged at 5000 *g* for 30 min at 4°C. The concentrate was then transferred to a 50 mL tube and the cycle was repeated until the entire sample has been concentrated. The concentrate was then aliquoted into 1.5 mL tubes in 200 µL volumes and stored at -80°C.

Figure 2.2. Tangential filtration flow setup to concentrate the viral fraction.

The viral concentration setup was optimised previously (Tay, 2013). The automated 0.2 µm tangential flow filtration (TFF) setup (a) and the setup for the viral fraction concentration using a 100 kDa TFF (b) used to process the effluent from the reactors. Prior to filtration and between changing carboys, the tubings were flushed using tap water to prevent cross contamination. The flow and sequence of the samples from the carboys were controlled using an Arduino UNO microcontroller of the valves (Arduino, USA). The transmembrane pressure for both systems was maintained between 1 to 2 kg/cm².

2.2.4. Bacteriophage nucleic acid extraction and multiple displacement amplification

Nucleic acids were extracted from 200 µL of concentrated viral fraction using the QIAamp MinElute Virus Spin kit (Qiagen, Germany) as per the manufacturers' instructions. The extracted viral DNA was used for whole genome multiple displacement amplification (MDA) using random hexamers with the illustra GenomiPhi V2 DNA Amplification kit as per the manufacturers' instructions. The amplified DNA was purified using the ethanol precipitation method (Green & Sambrook, 2016). Briefly, sodium acetate was added to a final concentration of 0.3 M at pH 5.2 was added to the amplified sample and mixed.

Two volumes of cold 100% molecular grade ethanol was added and incubated at -20°C overnight. The sample was then centrifuged at 15,000 *g* for 30 min and the supernatant was removed. One mL of 70% molecular grade ethanol was added and incubated at -20°C for 2 h before the DNA was pelleted by centrifugation at 15,000 *g* for 30 min. The supernatant was discarded and the pellet was air dried for 5 min before resuspension in 50 µL of sterile dH₂O.

2.2.5. Total genomic DNA extraction from aerobic granular sludge

One millilitre of suspended sludge was pelleted by centrifuging at 10,000 *g* for 10 min. The total genomic DNA was extracted using the pelleted sludge with the FastDNA™ SPIN Kit for Soil (MP Biomedical, USA) as per the manufacturers' instruction with one change. The homogenisation step was performed twice in the FastPrep® Instrument for 40 s at a speed setting of 6.0. To improve the quality of the extracted DNA, the genomic DNA was cleaned up using the Genomic DNA Clean & Concentrator (Zymo Research, USA) as per the manufacturers' instructions.

2.2.6. Total metagenome analysis of the aerobic granular sludge

The concentration of DNA samples was quantified on both the NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA) and the Qubit® 2.0 Fluorometer (Life Technologies, USA) using the dsDNA HS Assay kit. The amplified DNA was sequenced using the Illumina HiSeq platform as 250 bp paired end reads at the SCELSE sequencing facility. The quality of the metagenome reads was assessed using FastQC version 0.11.5 before it was adapter and quality trimmed using BBMap version 36.38 (Bushnell, 2015). Co-assembly of contigs was performed using MEGAHIT version 1.0.6-3 (Li *et al.*,

2016) with the meta-sensitive preset mode before ORF prediction was done using Prodigal version 2.6.3 (Hyatt *et al.*, 2010) with the meta mode. Using nucleic acid ORF sequences, sequence redundancy was reduced using cd-hit-est version 4.6.8 (Fu *et al.*, 2012) with the options for 95% sequence identity and word length of 10. The reduced nucleic acid ORF sequences were then used in a protein homology search using the Blastx program of DIAMOND version 0.8.22 (Buchfink *et al.*, 2015) against the NCBI nr database. Based on the Blastx output, the lowest common ancestor (LCA) annotation for the contigs were performed using MEGAN6 Community Edition version 6.8.12 (Huson *et al.*, 2016). The metagenome reads were then mapped to the co-assembled contigs using Bowtie 2 version 2.2.6 (Langmead & Salzberg, 2012) before the read coverage was obtained using the idxstats function of Samtools version 1.3.1 (Li *et al.*, 2009) to obtain a contig abundance table. The contig abundance and LCA-annotated contigs were then input into Phyloseq version 1.22.3 (McMurdie & Holmes, 2013) for downstream analyses.

2.2.7. Viral fraction analysis

The quality of the metavirome reads were assessed using FastQC version 0.11.5 before it was adapter and quality trimmed using BBMap version 36.38 (Bushnell, 2015). The co-assembly of contigs was performed using MEGAHIT version 1.0.6-3 (Li *et al.*, 2016) with the meta-sensitive preset mode before performing ORF prediction using Prodigal version 2.6.3 (Hyatt *et al.*, 2010) with the meta mode. The viral ORF amino acid sequences were then used in a protein homology search using the Blastp program of DIAMOND version 0.8.22 (Buchfink *et al.*, 2015) against the A CLAssification of Mobile genetic Elements (ACLAME) database. Based on the Blastp output, the lowest common

ancestor (LCA) annotation for the viral contigs were performed using MEGAN6 Community Edition version 6.8.12 (Huson *et al.*, 2016). The metavirome reads were then mapped to the co-assembled viral contigs using Bowtie 2 version 2.2.6 (Langmead & Salzberg, 2012) before the read coverage was obtained using idxstats function of Samtools version 1.3.1 (Li *et al.*, 2009) to obtain a contig abundance table. The LCA annotation of the viral contigs and their abundances were used for downstream analyses.

2.3. Results

2.3.1. Development of aerobic granular sludge from activated floccular sludge

Samples were collected from the reactors on a weekly basis to monitor the granulation of activated floccular sludge from Ulu Pandan that was used to seed the reactors to aerobic granules. Aerobic granules have been defined as compact aggregates, characterized by a minimum particle size of 100 μm with a SVI_5 of 50 mL/g or less (Barr *et al.*, 2010a). Granulation was defined here as occurring in 5 distinct phases: floccular, initiation, maturation, maintenance and dispersal (Tan *et al.*, 2014). In this study, the reactor was only operated to observe the floccular, initiation and maturation stages. During the floccular phase, the mean particle size of the sludge was 51.5 μm (50th percentile) (Figure 2.3). The initial reduction in the settling time, from 120 to 56 min, resulted in an average loss of 13% of the biomass (Figure 2.4).

Figure 2.3. Changes in particle size and SVI₅ (a) and settling time (b) during granulation and microscopic images of the granulation process (c).

The particle sizes of the sludge are represented as line plots of the 10th (---), 50th (—) and 90th (---) percentile (a), which represent the percentage of total particles below the corresponding size distribution. The compactness of the sludge particles was measured as SVI₅ (▼) (a) as the settling time (•) (b) was decreased throughout the process. Reactor 2 was taken as a representative for the weekly microscopic visualization of the four SBRS taken at 40X magnification (white bar, 200 μm) (c). The images, from left to right are from weeks 0 to 11, respectively. Sludge images for all reactors can be found as Appendix Figure 1, and plots for individual reactors can be found as Appendix Figure 2. The error bars represent standard deviation, n = 4.

By the end of week 3, the mean particle size was 96.2 μm (50th percentile) and small aggregates could be observed in the sludge (Figure 2.3a and c). Decreasing the settling time from 56 to 24 min (Figure 2.3b) did not further affect the amount of biomass (Figure 2.4) until week 4, when the sludge entered the

initiation phase, as determined by the presence of compact aggregates. The settling time was gradually reduced from 24 to 5 min from week 4 to the end of week 6 (Figure 2.3b) resulting in a further 44.24% decrease in biomass volume (SVI_5) (113 to 63 ml/g), a 29% average loss in MLVSS (3.44 to 2.45 g/L), and a mean particle size increase from 108.5 to 193 μm (50th percentile). The reduction in the settling time to 5 min ultimately selected for particles of > 482.5 μm and a significant decrease in SVI_5 to 16 mL/g (Figure 2.3a). It was apparent that the biomass was undergoing aerobic granulation as the biomass achieved the defined criteria of granules based on biomass particle size and SVI_5 (Barr *et al.*, 2010a).

The reactors were operated under conditions to achieve SNDPR. In terms of the reactor performance, nitrogen removal was stable within the range of 67% to 79% from week 1 to 11 (Appendix Figure 3a). In terms of phosphorus removal, by week 4, the phosphorous removal efficiency increased to 83.97% (Appendix Figure 3b) when compact aggregates were observed. Subsequently, phosphorous removal was maintained within the range of 50% to 76% during weeks 5 to 11. All four reactors formed mature granules by the end of the reactor operation. Throughout the reactor operation, the sludge and the viral fraction was collected to determine whether there was any correlation between the microbial composition and viral dynamics in the reactor and the granulation process.

Figure 2.4. Changes in biomass for the four sequencing batch reactors.

The mixed liquor suspended solids (MLSS) (•) and the mixed liquor volatile suspended solids (MLVSS) (■) along the reactor operation. The phase of granulation at time of sampling is indicated on the top of the graph. The error bars represent standard deviations. $n = 4$. Note errors bars given for all data points, although they may not be apparent for some due to the low error.

2.3.2. Microbial community composition of floccular and granular sludge

Traditionally, obtaining pure cultures of bacteria is the gold standard for experimentation and identification. It is increasingly understood that many microbes in nature often function as multi-species entities including for surface attached and suspended biofilms. Many such communities are physically and metabolically connected, and this may be why it is difficult to recover them as pure cultures under laboratory conditions (Schmeisser *et al.*, 2007). In contrast, metagenomic approaches offer a culture-independent way to probe the questions of “Who is there?” and “What are they doing?”, allowing the study of the microbial community composition, function and dynamics to provide new insights of different environments (Ercolini, 2013; Handelsman, 2004). For this study, the total genomic DNA was sequenced for the aerobic sludge samples

obtained from the four SBRs throughout the process of granulation. The metagenomic reads were assembled into contigs, annotated and the reads mapped to determine the relative abundance of the microorganisms in the system. This allows tracking of their diversity and changes in abundance as the granulation takes place over 11 weeks of reactor operation. The total number of reads for each week for all four SBRs ranged from 3.07×10^6 to 6.95×10^6 (Appendix Table 1), and therefore the data was transformed to give an even sampling depth.

In the early floccular stage (weeks 0 and 1), the clustering based on the relative abundances of the four SBRs suggested that the microbial communities were similar (Figure 2.5). However, from week 2, reactor 4 diverged from the other three reactors and by week 6 all four of the reactors showed slightly different community compositions. At week 11, there was > 60% dissimilarity between the communities of the mature granules (week 11) and the initial, floccular community (week 0) (Figure 2.5). PERMANOVA of the observed unique contigs in the reactors also showed that the microbial communities were highly dissimilar between the weeks ($P < 2e^{-16}$) and phases ($P < 2e^{-16}$) but similar between the reactors ($P = 0.954$).

Figure 2.5. Changes in microbial communities of SBRs throughout the process of granulation.

Principle coordinate analysis (PCoA) of the distance matrix of the reactor samples. The distance matrix was calculated using the Bray-Curtis dissimilarity. The phase of granulation is represented by the different symbols and the weeks are differentiated by colour.

To measure the richness and diversity of the aerobic granulation process, the Chao1 and the Shannon-Wiener indices were used, respectively. Bacterial communities were rich ($2,742,315 \pm 196,384$) (Figure 2.6a) and diverse (13.86 ± 0.045) (Figure 2.6a) in week 1, but by week 11 richness and diversity declined to $1,522,165 \pm 70,299$ and 12.82 ± 0.19 , respectively. The evenness of the community also decreased over time (Figure 2.6b). The decrease in the community diversity appeared to be correlated with the decrease of the evenness due to loss of richness of certain members of the community.

Figure 2.6. Changes in community richness and diversity during granulation.

Community richness was measured by the Chao 1 (Chao1) index (blue circles), and the community diversity was determined using the Shannon-Weiner index (red squares) (a). The community evenness was obtained by the formula:

$$\frac{\text{Shannon-Weiner index}}{\log_e \text{Chao1 index}} \quad (\text{b}).$$

The top 5 most abundant phyla across the weeks were *Proteobacteria*, *Nitrospirae*, *Bacteroidetes*, *Verrucomicrobia* and *Chloroflexi* (Appendix Figure 4b). The *Proteobacteria* steadily increased from 13.5% to 73.6% while the other members of the top 5 phyla gradually decreased in abundance. Of the *Proteobacteria*, the genus ‘*Candidatus Accumulibacter*’ (a PAO and denitrifier) was the most abundant, increasing from 5.76 to 56.34% by week 11 (Figure 2.7). The glycogen accumulating organisms (GAOs), ‘*Candidatus Competibacter*’ and ‘*Candidatus Contendobacter*’ did not change appreciably between, 1.45 to 4.65% and 2.08 to 5.7%, respectively (Figure 2.7). There was a peak in *Thauera* (a denitrifier) abundance at week 1 at 11.79%, followed by a significant reduction to 2.73% at week 11 (Figure 2.7). Nitrifiers, such as *Nitrospira*, progressively decreased from 33.8 to 4.26% over the course of the experiment. The other members of the top 15 genera generally had a lower

abundance between 0.02% and 3.96%, *Burkholderia* having the lowest abundance between 0.02% and 0.05%.

The biomass is considered to be granular when the sludge has a high particle size ($> 100\mu\text{m}$) and low SVI_5 (50 mL/g) (Barr *et al.*, 2010a). To determine which bacterial groups were correlated with granule formation, the relationship between the abundance of the top 15 genera and the particle size (50th percentile) (Appendix Table 2) and SVI_5 (Appendix Table 3) were determined using the Pearson correlation coefficient. The denitrifiers '*Candidatus Accumulibacter*' and *Dechloromonas* were positively correlated to the particle size of the granulating sludge. These two genera, within the family Rhodocyclaceae, are also important in phosphorus removal and the oxidation of ammonia (Tsuneda *et al.*, 2005). The genera *Burkholderia*, *Fibriimonas*, *Sulfuritalea*, *Terrimonas* and *Verrucomicrobium* were negatively correlated with granulation (Appendix Table 2).

Figure 2.7. Relative abundance of the dominant bacteria groups during the granulation process.

The abundance of the top 15 genera are coloured according to the colour gradient. The data are presented separately for each reactor. The contigs that could not be annotated at the genus level and those that are not in the top 15 removed from this plot. The plot with all genera included can be found as Appendix Figure 4. Note that for presentation purposes, the number of reads mapped have been log-transformed.

As there was an inverse relationship between particle size (increasing) and SVI_5 (decreasing) (Figure 2.3a), the abundance of the genera that contribute to SVI_5 reduction would have a negative correlation with SVI_5 . The genera '*Candidatus Accumulibacter*' and *Dechloromonas* were significantly

negatively correlated to SVI₅ and hence may directly contribute to the decrease in SVI₅ (Appendix Table 3). In contrast, the abundance of *Caldilinea*, *Flavobacterium*, *Nitrospira*, *Pseudomonas*, *Sulfuritalea* and *Verrucomicrobium* decreased with reduction of SVI₅.

2.3.3. Virome dynamics of a granulating sludge

Bacteriophages exert a complex influence over their microbial hosts and additionally may play a structural component of the matrix (Liu & Wang, 2007; Secor *et al.*, 2015; Webb *et al.*, 2004). At the same time, they may select for strains that are infection resistant, for example due to changes in their lipopolysaccharide structure, leading to increased surface hydrophobicity and in turn granulation (Wang *et al.*, 2006). Therefore, the relationship between phage community dynamics and granulation was investigated here.

Only DNA bacteriophages were targeted here as they represent the vast majority of phage. The sequences were assembled into viral contigs and the relative abundance for each viral family was obtained for weeks 1 to 10 of reactor operation (Figure 2.8). Throughout the reactor operation, the Microviridae were the most abundant and were present in all of the samples at 60.8 to 99.8%. At the end of the initiation phase (week 7), the abundances of the Podoviridae and Siphoviridae began to increase significantly and from week 9, they were the second and third most abundant viral families, respectively. At week 10, Podoviridae had an abundance of $11.15\% \pm 1.66\%$ and Siphoviridae had an abundance of $8.34\% \pm 0.96\%$. The abundance of Inoviridae also increased to $0.2\% \pm 0.05\%$ at week 5, when the sludge had developed into compact aggregates (Figure 2.3c) and peaked at week 9 at $1.4\% \pm 0.32\%$.

Siphoviridae, Microviridae and Myoviridae were also positively correlated with increased particle size (Table 2.1). Additionally, a distance based redundancy analysis (dbRDA) was done with the viral family abundance to identify covariates which have an effect on the changes in bacterial community (McArdle & Anderson, 2001). The changes in the bacterial community during the initiation (weeks 4 to 7) and maturation (weeks 8 to 9) phase seemed to be related to the Microviridae and Inoviridae viral families (Figure 2.9).

Figure 2.8. The relative abundance of viral families during granulation.

The relative abundance of the viral families across reactor operation for the individual reactors. The number of reads mapped to each family for the individual reactors can be found as Appendix Table 4. Note that for presentation purposes, the number of reads mapped have been log-transformed. Boxes that have been greyed would represent weeks where no counts were detected.

Table 2.1. Pearson's correlation between virus family and particle size.

Inoviridae	Podoviridae	Siphoviridae	Tectiviridae	Microviridae	Myoviridae
<u>Pearson r coefficient</u>					
0.4993	0.667	0.6038	0.4032	-0.6493	0.7093
<u>R squared</u>					
0.2493	0.4449	0.3646	0.1626	0.4216	0.5031
<u>P-value</u>					
0.1418	0.0351	0.0645	0.2479	0.0422	0.0216
ns	*	ns	ns	*	*

Figure 2.9. Viral families that affect bacterial community changes.

Distance-based redundancy analysis showing viral families that have the most influence on the bacterial community changes across the weeks of reactor operation. The points are coloured by the week of sampling.

2.4. Discussion

The effect of physical factors, such as shear force and settling time on granulation have been well studied and used to enhance the transition from floccular sludge into granules. However, the biological factors that contribute to the initiation of granulation, or the instability of the granule structure which often leads to granule disintegration after long periods of reactor operation is not as well studied. Therefore, it is important to improve our understanding of the factors that affect granulation. One under explored factor is the presence of bacteriophages in the reactor, which are one of the known factors that cause collapse of reactors (Barr *et al.*, 2010b) and have been suggested to be a structural component of aerobic granules (Liu & Wang, 2007). To study the microbial community change and role of bacteriophage in the granulation process, four SBRs were inoculated with floccular sludge and operated according to standard procedures.

2.4.1. The aerobic granulation process

Within the first four weeks of reactor operation, granules were successfully formed and this was in part due to the gradual reduction of the settling time (Figure 2.3). This was also observed in other studies which concluded that short settling times provided a strong selection pressure for compact and dense granules (Adav *et al.*, 2009; McSwain *et al.*, 2004). The reduction in settling time exerted a strong selection pressure on the floccular sludge and was associated with a 29% reduction in biomass from weeks 4 to 6, selecting for denser floccular sludge that eventually formed granules. As observed from weeks 6 to 8, the amount of biomass increased from a mean

MLVSS value of 2.705 to 4.435 g/L and particle size increased from 193 to 482.5 μm . The continual expansion of the granules could be attributed to microcolony outgrowth or aggregation. Outgrowth is characterized by the selection and enrichment of a particular group of bacteria in a floc, following which the microcolony outgrows to form larger and white granules. White granules were also observed in several EBPR systems where 97.5% of the granule biomass was composed of '*Candidatus Accumulibacter phosphatis*' (Barr *et al.*, 2010a). Alternatively, aggregation can involve the aggregation of numerous flocs, made up of diverse and distinct microbial populations, giving rise to yellow granules. In those granules, '*Candidatus Accumulibacter phosphatis*' only constituted 12.3% of the community, while the remaining 57.9% was comprised of '*Candidatus Competibacter phosphatis*' and the rest were other bacteria (Barr *et al.*, 2010a). The reduction in the settling time and the discharge of looser and lighter flocs were important processes that contributed to granulation.

2.4.2. Changes in the sludge microbial community

As the flocs developed into granules, there was a shift in bacterial communities in all of the SBRs (Figure 2.5 and Figure 2.7). The bacterial diversity was highest in the floccular phase (week 0 to 3) (Figure 2.6). This is supported by other studies showing a high diversity in activated sludge flocs (Ibarbalz *et al.*, 2013; Ju *et al.*, 2014; Yu & Zhang, 2012). In the four SBRs, the dominant phyla *Proteobacteria*, *Bacteroidetes*, *Verrucomicrobia* and *Nitrospirae* also dominated the bacterial communities of activated sludge (Appendix Figure 4b) (Yu & Zhang, 2012; Zhang *et al.*, 2012). As the sludge in the reactor granulated, the diversity of the bacterial community decreased (Figure 2.6b) and certain bacterial genera began to dominate the SBRs (Figure 2.7). This

decrease in diversity can be attributed to the short settling time and hydraulic retention time that caused an extensive discharge of the biomass, though these physical factors are necessary for conversion to aerobic granules (Weissbrodt *et al.*, 2012). In that study, the community was dominated by *Zoogloea* and *Dechloromonas* after the reactor was operated for 5 d (Weissbrodt *et al.*, 2012). Furthermore, *Proteobacteria*, *Bacteroidetes* and *Nitrospira* are well documented to be biofilm forming organisms (Chao *et al.*, 2015; Charnock & Nordlie, 2016; Martiny *et al.*, 2005; Martiny *et al.*, 2003; Rozej *et al.*, 2015) and their ability to readily form biofilms (as flocs) may have prevented their washout from the reactors. The short settling and discharge cycles were hypothesised to represent strong selection pressures to select for sludge that could settle within the stipulated time.

During the early stages of reactor operation, the *Proteobacteria* were not as abundant as *Acidobacteria*, *Bacteroidetes* and *Nitrospirae*. However, as the sludge granulated, the abundance of *Proteobacteria* increased and eventually dominated the reactor. This is due to the increased abundance of ‘*Candidatus Accumulibacter*’, ‘*Candidatus Competibacter*’ and ‘*Candidatus Contendobacter*’ (Figure 2.7), despite significant biomass washout during reduction of settling time. Additionally, the synthetic wastewater that consisted of acetate and propionate as carbon sources may also have played a part in the selection of the dominant bacteria. ‘*Candidatus Accumulibacter*’ are proficient in propionate uptake while ‘*Candidatus Competibacter*’ better utilise acetate (Oehmen *et al.*, 2005). The relatively higher abundance of ‘*Candidatus Accumulibacter*’, ‘*Candidatus Competibacter*’, ‘*Candidatus Contendobacter*’ and *Nitrospira*

throughout reactor operation suggests that these members were closely associated with denser flocs that settled faster and avoided washout.

In the maturation phase, the abundance of '*Candidatus Accumulibacter*' increased significantly relative to '*Candidatus Competibacter*'. This was despite the use of synthetic wastewater that contained a 3:1 ratio of acetate to propionate for use as carbon source. However, similar to previous studies of other SNDPR systems, '*Candidatus Accumulibacter*' and '*Candidatus Competibacter*' were the most abundant in the community (Barr *et al.*, 2010a; Lemaire *et al.*, 2008b). Aerobic granules generally are dominated by nitrifiers in the oxygen-rich outer layers, while the PAOs and denitrifiers are found in both the aerobic and anoxic zones (Winkler *et al.*, 2013). The GAOs, such as '*Candidatus Competibacter*', would usually dominate the anoxic zone to the core of the granule (Lemaire *et al.*, 2008b). The abundance of '*Candidatus Competibacter*' did not increase significantly and could have been due to the limited diffusion of nutrients into the inner anoxic zone. For example, granules with radii larger than 0.5 mm have been found to be diffusion limited for both acetate and DO (Li & Liu, 2005; Li *et al.*, 2008). The '*Candidatus Competibacter*' would likely become metabolically inactive or die in those deeper parts of the granules where access to a carbon source is limited. The data suggest that the Rhodocyclaceae like '*Candidatus Accumulibacter*' and *Dechloromonas* were important for granule formation. Other than nutrient availability and granular structure, the change in the microbial community might also be a result of other factors, such as bacteriophage predation (Shapiro *et al.*, 2010).

2.4.3. Population dynamics of the bacteriophage community during granulation

In addition to bacteria, other organisms such as fungi and protozoa are present in the reactors and these organisms have been demonstrated to also be hosts for viruses (Ghabrial *et al.*, 2015; Wang & Wang, 1991a, 1991b). In this study, the taxonomic assignment (based on protein homology against ACLAME and LCA clustering using MEGAN6) of the viral fraction assembled contigs showed that the majority of the contigs were of bacteriophage origin and reflected that the majority of the microbial population consisted of bacteria. The limitation of the method lies in the LCA clustering by MEGAN6 that utilised a database that is not prokaryotic virus-oriented. Many bacteriophages (especially those of the Inoviridae family) are often annotated not as viruses but instead as bacteria as these bacteriophages can exist in the host as a lysogen.

Microviridae were the most abundant members of the SBR viral community (Figure 2.8). Microviridae are single stranded DNA viruses that were abundant in the viral metagenomes of many aquatic environments and have the ability to lysogenise their host and facilitate horizontal gene transfer (Angly *et al.*, 2009; Lopez-Bueno *et al.*, 2009; Roux *et al.*, 2012). Myoviridae and Podoviridae are generally lytic bacteriophages that lack genes required to establish lysogeny (Lavigne *et al.*, 2009), while Siphoviridae are temperate phages found to infect a range of bacteria (Brussow & Desiere, 2001; Desiere *et al.*, 2000; Niu *et al.*, 2012). We hypothesize that their presence in the reactors is an indication that phage predation and lysis of specific bacteria might be required during granulation and their maturation, though this remains to be tested. Additionally, the change in its abundance is positively correlated with the

increase in sludge particle size (Table 2.1) and seemed to influence the bacterial community change that was observed (Figure 2.9). It would be confirmative if adback of specific or multiple viral species of the Microviridae family during reactor operation could enhance the formation or cause collapse of aerobic granules, but this would only be explored in future work. However, it has been demonstrated in a full-scale membrane bioreactor that the increase and decrease of plaque forming unit per millilitre counts of bacteriophages in the reactor effluent appeared to correlate with the abundance of their potential hosts (Shapiro *et al.*, 2010). For example, the lysis of bacterial genera that are associated with foaming, e.g. *Nostocoida limicola*, '*Candidatus* Miclothrix' and *Chloroflexi*, could improve the granulation process by selecting for faster settling sludge (Guo & Zhang, 2012; Nielsen *et al.*, 2009). Additionally, the lysis of bacteria releases cellular DNA into the sludge, and this eDNA may play a role in enhancing the biofilm matrix as a component of the macromolecular scaffold (Carrolo *et al.*, 2010; Jakubovics *et al.*, 2013).

Inoviridae bacteriophages were also detected and these phages are generally non-lytic (Ploss & Kuhn, 2010). Interestingly, there was an increase in abundance of Inoviridae detected at weeks 5, 7, 8 and 9 (Figure 2.8) and is consistent with the appearance of compact aggregates (Figure 2.3c) as well as the initiation of granulation as defined by Barr *et al.* (2010a). Additionally, like for Microviridae, it seemed that Inoviridae influenced the bacterial community change (Figure 2.9), more specifically for the weeks 4 to 8 when the sludge was undergoing initiation of granulation till the early maturation phase. In *Pseudomonas aeruginosa* PAO1, the Pf4 filamentous phage organises the biofilm matrix into a liquid crystalline structure that has increased adhesion and

viscosity (Secor *et al.*, 2015). It is possible that the role of the filamentous phage in the process of granulation is that of a structural component, as structures similar to that observed in a *P. aeruginosa* PAO1 biofilm where the Pf4 filamentous phage formed a dense lattice structure in the matrix, was observed in granules (Liu & Wang, 2007; Webb *et al.*, 2004). Electron microscopy of the granules from our reactors could be done to confirm if such structures can be observed.

Phages that are generally lytic, e.g. Microviridae, Myoviridae and Podoviridae, were positively correlated with increased particle size (Table 2.1). There is evidence that the lysogenic life cycle is favoured in nutrient limited environments (McDaniel & Paul, 2005; Williamson & Paul, 2006; Wilson & Mann, 1997) as was observed for the activated sludge process. This selection may in part be due to the loss of free phage particles that are washed out during the discharge stage. A prophage can be induced to become lytic due to sudden perturbation by temperature, UV exposure, or presence of toxic substance in the environment (Weinbauer, 2004). These sudden perturbations have led to deflocculation and decreased reactor performance (Alighardashi *et al.*, 2009; Morgan-Sagastume & Grant Allen, 2005), and might be due to the induction of prophages. In this study, the increase in Podoviridae and Siphoviridae from week 7 might have led to killing of organisms responsible for phosphorous removal, e.g. '*Candidatus Accumulibacter*', *Dechloromonas* and *Thauera* and hence could have contribution to the unstable phosphorus removal. Additionally, lytic bacteriophages may contribute to granulation by lysis of filamentous bacteria that cause foaming (light sludge that floats), and work have been done using Siphoviridae to control levels of *Gordonia*, *Nocardia* and *Rhodococcus*

(Khairnar *et al.*, 2016; Khairnar *et al.*, 2014; Liu *et al.*, 2015; Thomas *et al.*, 2002).

Although these studies used Siphoviridae, they highlight the potential and possibility of other bacteriophage families playing a role in granulation.

The relationship between bacteriophages and their hosts have been modelled. In a simple environment, if there is no cost for the resistance to phages, the population of functionally redundant phages and hosts oscillate with time and in the manner described in the “kill the winner” hypothesis (Bohannan & Lenski, 1997; Rodriguez-Brito *et al.*, 2010; Rodriguez-Valera *et al.*, 2009). However, if there is a cost, bacteriophages have been shown to drive host diversification (Middelboe *et al.*, 2009; Winter *et al.*, 2010). Diversification of bacteria may occur in the phage receptor region (Harvey *et al.*, 2018) or on CRISPR elements that quickly evolve to provide a defence system to the host (Held *et al.*, 2010). Another dynamic between hosts and bacteriophages are those for temperate phages where the phage encodes genes or regulate host gene expression that are useful for the host (Hernandez-Doria & Sperandio, 2018; Willner *et al.*, 2009). Cyanophage P-SSP7 of *Prochlorococcus* have been demonstrated to encode bacterial-like genes that play a role in photosynthesis and central metabolism (Lindell *et al.*, 2007). These auxiliary metabolic genes were thought to ensure a source of energy needed for phage replication in nutrient poor areas of the ocean (Breitbart *et al.*, 2007) as burst sizes were bigger for phages that carry photosynthesis genes (Bragg & Chisholm, 2008; Hellweger, 2009). In *Salmonella enterica* serotype Typhimurium, usually found in inflamed gut, the SopE Φ phage increased the production of inducible nitric oxide synthase that facilitated production of nitrate precursors and may be fitness-enhancing by extending pathways used in anaerobic respiration (Lopez

et al., 2012). The prevalence of phage-encoded biological functions in host cells and the recombination between the phage and host genes suggest that phage populations serve as gene reservoirs which contribute to niche partitioning of microbial species in various environments.

The results from this study led us to hypothesize that bacteriophages may play a role in the granulation of sludge in the reactors. While it remains to be tested, lytic bacteriophages may have selected for bacteria that made up denser sludge, and filamentous bacteriophage may be part of the granule structure. The process of aerobic granulation that provided a strong selective pressure through short settling times and discharge of light biomass resulted in a less diverse bacterial community. While the selection of bacteria that promote granule formation allow the sludge to be superior in terms of resistance to toxic substances (due to high density of cells) and improved dewaterability, the reduced diversity of a granule community may also explain its instability. This is important as bacterial diversity has been shown to promote community stability as well as resilience to perturbation.

Chapter 3 – The role of Pf filamentous phage in *Pseudomonas aeruginosa* PAO1 biofilm development

3.1. Introduction

Bacteriophages are prokaryotic viruses that are host specific and utilize the cellular machinery of the host to aid in their replication. Unlike the majority of phages where new progeny are released via host cell lysis, filamentous phages are constitutively produced and released from the cell without cell lysis. The Inoviridae family of filamentous phages have circular, ssDNA genomes that are non-enveloped and packed in a helical capsid which gives it the filamentous structure (Ackermann & Kropinski, 2007).

M13, that infects *Escherichia coli*, is one of the most comprehensively studied filamentous phage (Abedon, 2011; Ackermann, 2003; Ploss & Kuhn, 2010). The phage utilises its minor coat protein to recognise and bind to the F pilus of its host. Subsequently, the F pilus is retracted, which facilitates the interaction between the phage receptor protein to the receptors on the cell membrane, releasing the genome of the phage into the host (Jacobson, 1972). The complementary strand of the ssDNA is synthesized by host polymerases, forming the replicative form (RF) of the phage. This double stranded form of the phage generates new copies of the ssDNA genome via the rolling circle replication. The RF is similar to a plasmid and is responsible for expression of the encoded phage genes as well as the generation of new phage genomes for packaging. The newly synthesized ssDNA genome is subsequently coated with thousands of copies of the major coat protein as the phage is extruded from the host cell membrane (Ploss & Kuhn, 2010).

Infection by filamentous phages have also been described in a wide range of bacteria such as PH75 in *Thermus thermophilus* (Pederson *et al.*, 2001), Ypf phi in *Yersinia pestis* (Chouikha *et al.*, 2010) and f1 and fd in *E. coli* (Straus *et al.*, 2008). Filamentous phages, which include the Pf phages of *Pseudomonas aeruginosa*, share proteins that are probable homologues and a core genome (Mooij *et al.*, 2007; Park *et al.*, 2010; Webb *et al.*, 2004). For example, the coat proteins of Enterobacteria f1, *Xanthomonas* ΦLf and *Pseudomonas* Pf3 have 98.6%, 8.8% and 8.3% identity to *E. coli* Ff phage, respectively. The proteins encoded by these phages are not well conserved and their classification were based on the organisation of their genes as well as their morphology (Mai-Prochnow *et al.*, 2015).

Usually, it would be expected that the host would lose or inactivate the prophage unless it confers a fitness advantage. Filamentous phage have been linked to resistance to reinfection by the same or similar phage, as for the SW1 filamentous phage of *Shewanella piezotolerans* (Wang *et al.*, 2007), cf filamentous phage of *Xanthomonas campestris* (Ahmad *et al.*, 2014; Cheng *et al.*, 1999), the CTXΦ phage of *Vibrio cholerae* (Davis *et al.*, 1999; Davis *et al.*, 2000) and the Pf4 filamentous phage of *P. aeruginosa* PAO1 (Rice *et al.*, 2009). In *P. aeruginosa* PAO1, it was shown that Pf4 contributed to biofilm development and stress resistance (Rice *et al.*, 2009; Secor *et al.*, 2015; Webb *et al.*, 2004). The Pf4 prophage knockout mutant no longer underwent cell death and nor did it form hollow colonies. The biofilms of the Pf4 knockout mutant were also less stable to surfactant exposure and were less virulent than the *P. aeruginosa* PAO1 wild type (WT) in an acute lung infection mouse model. Furthermore, biofilms formed by the Pf4 mutant also did not generate

morphotypic variants, e.g. small colony variants (SCV), unlike *P. aeruginosa* PAO1 WT biofilms (Rice *et al.*, 2009).

It has recently been shown that *P. aeruginosa* PAO1 encodes a second filamentous phage, designated as Pf6 (Tay, 2008). This is somewhat rare for a host to have two closely related bacteriophages in their genome, which suggests that Pf6 also makes significant contributions to its host, although this remains to be demonstrated. Therefore, the aim of this chapter was to generate a Pf6 deficient PAO1 strain and together with a Pf4 deficient strain, use transcriptomic analysis to measure the differential expression of genes within the PAO1 host genome in the presence or absence of the Pf filamentous phage in early and late stages of biofilm development. At the same time, production of the two phages during biofilm development was determined by measuring plaque formation as well as by quantitative PCR.

3.2. Materials and methods

3.2.1. Bacterial strains and culture conditions

All bacterial strains (Table 3.1) were maintained on Luria-Bertani (LB) medium, either in broth or on plates supplemented with 1.5% (w/v) agar when necessary. Bacteria were cultured in either LB broth or M9 minimal growth medium (48 mM Na₂HPO₄, 22 mM KH₂PO₄, 9 mM NaCl, 19 mM NH₄Cl, 2 mM MgSO₄, 0.1 mM CaCl₂, supplemented with 0.04% (w/v) glucose). For plasmid maintenance in *E. coli*, the medium was supplemented with either 100 µg/mL carbenicillin or gentamicin (Carb100 or Gm100). The *P. aeruginosa* Pf4 knockout (PAO1ΔPf4), putative Pf6 knockout (PAO1ΔPf6) and the Pf4 and Pf6

double knockout (PAO1 Δ Pf4 Δ Pf6) mutants were grown in medium supplemented with 100 μ g/mL gentamicin (LB Gm100 or M9 Gm100). The cultures were incubated overnight at 37°C, at 200 rpm in a rotary incubator.

3.2.2. Gel electrophoresis

Gel electrophoresis analyses were performed using either 1 or 2% (w/v) agarose dissolved in 1 $\times$ TAE buffer (39.95 mM Tris-base, 11.42% (v/v) glacial acetic acid, 127.29 μ M EDTA). One or two microliters of 6 $\times$ DNA loading dye (Fermentas Life Sciences, USA) was mixed with 10 μ L of DNA or 1 μ L of GeneRuler™ 1 Kbp or 100 bp plus ladder (Fermentas Life Sciences, USA), before being loaded into the wells. Electrophoresis was performed at 80 V using a Bio-Rad Power Pac 300 (Bio-Rad Laboratories, USA) for 80 min and subsequently submerged in a 0.5 μ g/mL ethidium bromide staining buffer for 30 min. All gel images were digitally recorded using a Molecular Imager Gel Doc System (Bio-Rad Laboratories, USA).

3.2.3. Genomic DNA extraction

Genomic DNA was extracted using QIAamp® DNA Mini Kit (Qiagen, Germany) as per the manufacturers' instructions. The concentration and purity was assessed using both the NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA) and Qubit® 2.0 Fluorometer (Life Technologies, USA).

3.2.4. Polymerase chain reaction

Polymerase chain reactions (PCR) were performed in a C1000™ Thermal Cycler (Bio-Rad, USA) using either Platinum™ Taq DNA polymerase (Invitrogen, USA) or KAPA HiFi HotStart DNA polymerase ReadyMix (Kapa Biosystems, USA) as per the manufacturers' recommendation. The cycling conditions were

set as follows: (1) initial denaturation at 95°C for 3 min; (2) 25 to 30 cycles of 95°C for 30 s, annealing at primer specific temperatures (Table 3.2) for 30 s and extension at 72°C for the time as specified in Table 3.2; and (3) final extension at 72°C for 5 min.

Table 3.1. Bacterial strains and plasmids.

Strain or Plasmid Host	Characteristics	Origin or Source
<i>Pseudomonas aeruginosa</i>		
PAO1	Wild type strain	Stover <i>et al.</i> (2000)
PAO1ΔPf4	ΔPf4::Gm ^r , Pf4 prophage deletion mutant	Tay (2013)
PAO1ΔPf4ΔPf6 (putative)	ΔPf4, ΔPf6::Gm ^r , Pf4 and Pf6 putative prophage deletion mutant	Tay (2013)
PAO1ΔPf6 (putative)	ΔPf6::Gm ^r , putative Pf6 prophage deletion mutant	This study
<i>Escherichia coli</i>		
DH5α	F-φ80/lacZΔM15 Δ(lacZYA-argF)U169 <i>recA1 endA1 hsdR17(r_k⁻, m_k⁺) phoA supE44 thi-1 gyrA96 relA1 tonA</i>	Taylor <i>et al.</i> (1993)
TOP10	F- <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80/lacZΔM15 Δ <i>lacX74 recA1 araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU galk rpsL (StrR) endA1 nupG</i>	Invitrogen
Plasmids		
pCR4 – TOPO	TA TOPO Plasmid for Sequencing. Amp ^r , Kan ^r .	Invitrogen

Table continued on next page

Table continued from previous page

pCR4_Pf4qPCRcontrol	TA TOPO Plasmid with Pf4_435_3050 amplified region inserted. Amp ^r , Kan ^r . 6571 bp.	This study
pCR4_Pf6qPCRcontrol	TA TOPO Plasmid with Pf6_8541_10448_2 amplified region inserted. Amp ^r , Kan ^r . 5863 bp.	This study
pDEST_Pf6KO	pDEST TM R4-R3 vector hosting the Pf6 knockout construct. Amp ^R , Gm ^R .	Tay (2013)
pUCP18-RedS	pUCP18 hosting the Red operon (<i>gam</i> , <i>bet</i> , <i>exo</i>) inducible by arabinose (<i>araC</i>), <i>sacB</i> . Gm ^R (<i>bla</i>).	Lesic and Rahme (2008)

3.2.5. Generation of Pf6 mutant via Lambda RED mediated recombination

The Lambda RED recombinase system (Invitrogen, USA) was used to generate the Pf6 knockout deletion mutant from PAO1. The system facilitates homologous recombination either between linear DNA or plasmid vector hosting the knockout construct (Figure 3.1) through the use of lambda phage recombinase and has been previously adapted for use in *P. aeruginosa* (Lesic & Rahme, 2008). For the linear DNA, the gentamicin resistance cassette flanked by nucleotides upstream and downstream of Pf6 was amplified via PCR from pDEST_Pf6KO (Table 3.1) using Pf6-UpF_KO and Pf6-DnR_KO (Table 3.2). The amplicon was purified and concentrated using PureLink PCR Purification Kit (Invitrogen, USA) as per the manufacturers' recommendation.

Figure 3.1. Homologous recombination by Lambda Red.

After the Red operon from the pUCP18-RedS in the host was expressed, the host was transformed with either a linear fragment (a) or a vector hosting the knockout construct (b). Expression of the Red genes prevented degradation of the linear fragment and facilitated the homologous recombination. Using the linear fragment, both recombination events have to occur for successful insertion of the gentamicin resistance cassette. Similarly for plasmids, two recombination events are also required for successful homologous recombination. The first event involved the integration of the vector either upstream or downstream of the target region, followed by the second event that looped out the integrated vector, leaving behind the gentamicin resistance cassette. However, for plasmids, a single recombination event is enough to survive selection as at this point the whole plasmid, including the gentamicin cassette, has been integrated into the chromosome.

The process was adapted from Lesic and Rahme (2008). Briefly, prior to the homologous recombination step, the host was transformed with pUCP18-RedS (Table 3.1) and grown in super optimal broth with catabolite repression (SOC) medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 20 mM MgSO₄, 19.9 mM glucose, 8.56 mM NaCl, 2.5 mM KCl) containing 200 µg/mL of

carbenicillin at 37°C to an OD₆₀₀ of 0.2. L-arabinose was then added to a final concentration of 0.2% (w/v) and the culture was allowed to grow for an additional 90 min which allows the expression of the Red operon. After induction, electrocompetent cells were prepared by incubation on ice for 30 min followed by 4 rounds of washing with deionised water containing 10% (v/v) glycerol. During the wash step, centrifugation was performed at 10,000 *g* for 10 min at 4°C. The washed cells were resuspended with deionised water containing 10% (v/v) glycerol to one-tenth its initial volume.

One hundred microlitres of electrocompetent cells were mixed with either 4 µg of purified amplicon or 2 µg of pDEST_Pf6KO plasmid in a 2 mm electrocuvette (BioRad, USA) for electroporation using the GenePulser (BioRad, USA) with the following settings: voltage at 2500 V, capacitance at 25 µF and resistance at 200 Ω. Following electroporation, 900 µL of SOC medium was immediately added to the cells and transferred to a 15 mL tube. The cells were allowed to recover for 2 h at 37°C in a rotary incubator. The transformants were then plated onto LB Gm100 plates for selection and allowed to grow for at least 16 h in a 37°C incubator. After incubation, any transformants that grew on the plate were streaked onto a NaCl-free LB Gm100 plate containing 10% sucrose to counter-select the pUCP18-RedS plasmid.

Table 3.2. Primer sequences.

Primer names	Sequences (5' → 3')	Annealing position ¹	Ta (°C)	Product size (bp)
Pf4_435_3050_F	CCCCGCTCCCTCGA TCTTTTACTC	785744 → 785767	57.8	2615
Pf4_435_3050_R	ACTCCCGCCTTACC CCTCAATCAC	788359 → 788382		
Pf6_8541_10448_2_F	ACCGGAGGAGCGT AGCGATGATTC	8541 → 10448 (RGP42)	57.8	1907
Pf6_8541_10448_2_R	GGGGGCGGTCTTC TGGTAGTGCT	8541 → 10448 (RGP42)		
Pf6_Tq_2_F1	AGACAACACCGACT GAAGTAACTCA	9097 → 9121 (RGP42)	57.8	79
Pf6_Tq_2_R	TGACGGCGCTTCAT TTTTTC	9157 → 9176 (RGP42)		
Pf4_Tq_2_F1	CCCCGCTCCCTCGA TCT	785744 → 785861	57.8	75
Pf4_Tq_2_R	AAGGCGTGATAGTC CCAGTAGAGA	785796 → 785819		
16S_27F ²	AGAGTTTGATCMTG GCTCAG	27 → 1492 (bacterial 16S)	55	1465
16S_1492R ²	ACGGTTACCTTGTT ACGACTT	27 → 1492 (bacterial 16S)		
Pf6-UpF_KO	GGTGAGGGCCTGG GCAACAA	5240993 → 5241012	60	3538
Pf6-DnR_KO	ACACCCCGTCGATC ATCACCAGTA	5242025 → 5242007		
Pf6FlankUp_F	CGGCGCTGGCGGG ATTCTT	5240351 → 5240369	60	3347
Pf6detectR	CGGGCGGGCCAAC ATAGCAT	5242672 → 5242653		

Table continued on next page

Table continued from previous page

Pf6detectF	TCATCCGCGCCGAA GTGG	5241702 5241719	→	60	3427
Pf6FlankDn_R	ACTTCCGGCGCGT GGTAGGTG	5244103 5244083	→		
repC_ORF-P_F	TAGAGAAATAATTA CCCTGG	788886 788867	→	56	133
repC_ORF-P_R	TTTTGGGGCCGATT TTC	788754 788770	→		
gapA_EMSA_F	ATACCAGCGTAGAC GAGGTCAA	3588190 3588211	→	60	100
gapA_EMSA_R	TTGAAATCCACCGA CACCAG	3588289 3588270	→		

¹ Annealing positions are with reference to the complete sequence of *P. aeruginosa* PAO1 unless otherwise stated.

² Universal primers for bacterial 16S gene.

3.2.6. Screening for Pf6 knockout mutants

Filamentous bacteriophage can be present in the host in 3 forms, the prophage that is integrated into the genome, the phage particle that has a ssDNA genome, and as a replicative form (RF) in the host, which is a dsDNA form of the phage that serves as a template for genome replication prior to the extrusion of phage particles. Primers were designed to test and discriminate between true knockout mutants (successful double-crossover recombination with entire prophage removed from the genome) and single crossover mutants that retain a functional copy of the Pf6 prophage. PCR primers were designed to discriminate between: (1) complete removal of Pf6 prophage in the genome using sequences flanking the homologous recombination region (Figure 3.2, blue and green arrows); (2) whether the vector integrated upstream or

downstream if it is a single crossover mutant (Figure 3.2, amplicon generated from blue or green arrows); and (3) a unique region in the Pf6 genome (Figure 3.2, purple arrows). The primers targeting the unique region (Table 3.2) determines if all forms of Pf6 have been removed from the host.

Figure 3.2. Relative primer annealing sites to screen for successful Pf6 knockout mutant.

Primers were designed to check if the prophage has been removed from the genome. The flanking primers target the outside of the region of homologous recombination. To check if the prophage was removed from the *P. aeruginosa* PAO1 genome, the amplicons from both the blue and green arrows were amplified by PCR. The purple arrows show the relative binding site for the primers that target the Pf6 unique sequence and was used to detect for both the replicative and prophage form of Pf6.

3.2.7. Construction of quantitative PCR quantification standards

Using primers Pf4_435_3050_F, Pf4_435_3050_R, Pf6_8541_10448_2_F and Pf6_8541_10448_2_R (Table 3.2), the unique regions of the Pf4 and Pf6 prophages were amplified. The amplified regions were purified using PureLink® PCR Purification Kit (Invitrogen, USA) as per the manufacturers' instructions before being used in a TOPO® Cloning reaction into a pCR™4-TOPO vector (Invitrogen, USA). The resulting plasmids were designated as pCR4_Pf4qPCRcontrol and pCR4_Pf6qPCRcontrol, respectively (Table 3.1). The plasmids were checked for the correct size using *Nde*I (New

England Biolabs, USA) and Sanger sequencing of the insert (AITBiotech, Singapore) to ensure that the qPCR primer binding sites do not contain point mutations. The sequencing data was analysed using SeqMan Pro™ version 11.1.0 (DNASTAR, Inc., USA).

3.2.8. Tubing biofilm system

Biofilms were allowed to develop in 15 cm long silicon tubing size 3.18 mm (Silastic® laboratory tubing, USA) in a continuous flow reactor setup under sterile conditions (Figure 3.3). The biofilm system was fed with M9 minimal salts medium and primed prior to inoculation.

During inoculation of the bacterial culture, the pump was switched off and the upstream tubing was clamped to prevent backflow. Using a 3 mL syringe with a 26 G × 1½" needle, 2 mL of 1×10^8 cells/mL bacterial culture was injected into the lumen of the tubing. The bacterial cells were allowed to attach to the tubing surface for 1 h at room temperature under static conditions. The tubing was unclamped and flow was restored at a flow rate of 9 mL/h. Biofilms were developed at room temperature for 2 or 7 d.

Figure 3.3. Schematic representation of the continuous flow reactor tubing biofilm system used.

The components of the system are connected by Silastic® silicon tubing size 1.57 mm, except for the tubing running through the Ismatec IPC High Precision Multichannel Dispenser Pump and connecting to the bubble trap, where Plat Silicone Tubing size 1.42 mm (Masterflex, USA) was used.

3.2.9. Quantification of phage titre in biofilm effluent

The phage titre in the biofilm effluent was determined using a modified top layer agar (TLA) method previously described by Webb *et al.* (2003). Briefly, TLA plates were made by seeding the LB top layer containing 0.75% (w/v) agar with 0.5 mL overnight culture for every 5 mL soft agar. The TLA plates were allowed to dry before proceeding. Phage particles were collected from PAO1 and PAO1ΔPf4 biofilm effluent daily and passed through a 0.2 µm Acrodisc® syringe filter (Pall Life Sciences, USA) to remove planktonic cells. The filtered effluent was serially diluted and spotted onto TLA plates, with a lawn of either PAO1 or PAO1ΔPf4. The plates were dried and incubated at 37°C overnight.

Phage induced plaques were observed and counted, and the number of plaque forming units (PFU) per mL for each sample was calculated from dilutions that yielded plaque counts between 10 to 100 plaques per spot.

3.2.10. Extraction of DNA from phage in biofilm effluent

As previously described, phage particles were collected daily and passed through a 0.2 µm syringe filter. In order to remove any contaminating extracellular DNA (eDNA), DNase I (Calbiochem, USA) was added to the filtered effluent at a final concentration of 100 U/mL of effluent and incubated at 37°C for 30 min. DNase was inactivated by heating at 80°C for 20 min. The extraction of phage DNA from the DNase I-treated samples was performed using the PureLink® Viral RNA/DNA Mini Kit (Invitrogen, USA) as per the manufacturers' instruction, with the exception that the elution buffer was heated to 60°C and the volume was fixed to 50 µL per sample.

3.2.11. Determination of phage population dynamics via qPCR

Phage DNA extracted from the *P. aeruginosa* PAO1 and PAO1ΔPf4 biofilm effluent was used to determine the bacteriophage population dynamics of Pf4 and Pf6. SyBr Green qPCR detection of Pf4 was done with primers Pf4_Tq_2_F and Pf4_Tq_2_R, and detection of Pf6 with primers Pf6_Tq_2_F and Pf6_Tq_2_R (Table 3.2). Amplicons from both primer sets were checked for primer specificity and the formation of primer dimers via gel electrophoresis using the respective qPCR quantification standards.

A standard curve was generated from the reactions carried out with serial dilutions of pCR4_Pf4qPCRcontrol and pCR4_Pf6qPCRcontrol, respectively (Table 3.1) with 0.05% Tween 20 (Sigma Aldrich, USA). Each qPCR reaction

mixture contained: 10 µL of 2X Fast SYBR® Green Master Mix (Applied Biosystems, USA) added to 0.2 µL of forward and reverse primers respectively and 2 µL of template DNA before topped up with 7.6 µL of dH₂O to a final volume of 20 µL.

All qPCR reactions were performed using the StepOnePlus™ Systems (Life Technologies). The cycling conditions were: pre-cycling at 95°C for 2 min, followed by 40 cycles of 95°C for 10 s and annealing and extension at 60°C for 30 s. Results were analysed using the StepOnePlus™ Software version 2.3 (Life Technologies), using the standard curve to quantify the amount of target phage in each sample.

3.2.12. Transcriptomic analysis of PAO1 biofilms

P. aeruginosa PAO1 and PAO1ΔPf4 biofilms were allowed to develop for 2 and 7 d before total RNA was extracted. The samples were treated with RNeasy Protect Bacteria Reagent (Qiagen, USA) before extraction. RNA was extracted using the RNeasy® Mini Kit (Qiagen, USA) inclusive of in-column DNA digestion. To ensure improved removal of contaminating DNA, the samples were further treated with TURBO™ DNase (Ambion®, USA), while DNase removal and RNA cleanup was performed using RNA Clean & Concentrator™ (Zymo Research, USA) as per the manufacturers' instruction. The RNA was sequenced as cDNA using the Illumina HiSeq as 100 bp paired end reads at the SCELSE sequencing facility. The raw total RNA reads were adapter and quality trimmed using BBMap version 36.38 (Bushnell, 2015) before the rRNA reads were depleted using SortMeRNA version 2.1 (Kopylova *et al.*, 2012) to get mRNA reads. The data analysis was then continued using the mRNA reads as

described in Figure 3.4. After differential expression analysis, gene set enrichment analysis was performed using the piano package (version 1.18.1).

Figure 3.4. Transcriptomic analysis pipeline.

The RNAseq reads were quality checked to removed failed reads and trimmed to remove any ambiguous or low quality bases. Reads were then mapped using Bowtie 2 against the *P. aeruginosa* PAO1 genome (NC_002516) modified to include the Pf6 prophage genome. HTSeq (Anders *et al.*, 2015) was used to count the number of reads in each feature with the modified genome as a reference. Differentially expressed genes that were statistically significant was identified using the DESeq2 package (Love *et al.*, 2014) in R and the datasets were tested using PERMANOVA to test for significant differences between the different conditions (strain and day). The log2 fold change values extracted at the end were based on the false discovery rate cutoff of 0.1.

3.3. Results

3.3.1. Generation and screening of Pf6 knockout mutants

To better understand how filamentous phage play a role in granulation or biofilm formation, a model phage-host system was used. The objective was to understand the mechanisms by which the filamentous phage control gene expression in their host, that might be linked to enhanced biofilm formation, a key requisite for granule formation. The *P. aeruginosa* genome contains two filamentous phages, Pf4 and Pf6. For detailed studies, it is ideal to have a deletion strain to better understand the impact of the two phages on the host. A Pf4 phage deletion strain was previously generated (Rice *et al.*, 2009) and therefore, deletion of the Pf6 prophage was attempted here. Using a plasmid vector hosting the knockout construct (pDEST_Pf6KO) to generate a knockout mutant has the disadvantage that the recombinant derived could be single or double recombinants. Single cross-over recombinants are clones that have the vector integrated either downstream or upstream of the target sequence (Figure 3.1b) and are not true knockout mutants since the target sequence is still present in the genome and possibly still functional. PCR was used to validate whether gentamicin resistant transformants were single or double recombinants. Using a linear PCR product of the knockout construct, only one recombination event (effectively a double cross-over) can occur in the process, unlike single or double recombinants in the case when using pDEST_Pf6KO. To increase the likelihood of success, both methods were used here (Figure 3.1).

A combination of external flanking primers that anneal outside the region of homologous recombination and primers that anneal within the region were

designed to differentiate the two kinds of recombinants. If the Pf6 prophage has been completely replaced by the gentamicin cassette, primers Pf6FlankUp_F and Pf6detectR would produce a 3,321 bp while Pf6detectF and Pf6FlankDn_R would produce a 3,401 bp PCR product (Figure 3.5). Using the DNA of *P. aeruginosa* PAO1, PAO1ΔPf4 or a single recombinant would not yield any of the expected PCR products since the annealing sites will be too far apart (approximately 15 Kbp) for complete amplification under the set conditions. Single recombinants would contain the primer sites for PCR_FlankUp and not PCR_FlankDown if pDEST_Pf6KO was integrated upstream of Pf6 (Figure 3.1b, left), since the Pf6 prophage would be between the sites recognised by the primer PCR_FlankDown.

Figure 3.5. Expected PCR products generated from screening Pf6 prophage knockouts.

The two primer sets depicted were designed to differentiate double and single recombinants. Successful Pf6 prophage knockouts are double recombinants and would be positive for both PCR_FlankUp and PCR_FlankDown.

After transformation using both linear DNA and pDEST_Pf6KO, only one transformant using pDEST_Pf6KO grew on LB10 supplemented with 100 µg/mL gentamicin and 10% (w/v) sucrose. As controls for the PCR screening, genomic DNA of *P. aeruginosa* PAO1 and PAO1ΔPf4 were used. At the same time, a

putative PAO1 Δ Pf4 Δ Pf6 (obtained from Tay (2013) generated using Lambda RED mediated recombination of linear DNA) was also screened. The transformant was positive only for PCR_FlankDown, indicating that it is a single recombinant with the vector integrated downstream of Pf6 (Figure 3.6).

Figure 3.6. Screening for Pf6 knockout generated with Lambda RED mediated recombination.

One putative PAO1 Δ Pf6 clone was recovered after homologous recombination. Amplicons were generated from PAO1, PAO1 Δ Pf4, and a putative PAO1 Δ Pf4 Δ Pf6 from a previous study (Tay, 2013) as controls for the primers used. The left most lanes for each gel contain the 1 Kbp DNA ladder, where sizes are given in bp. The other lanes contain the reaction products for PCR_FlankUp (U), PCR_FlankDown (D), the Pf4 unique region (4) and the Pf6 unique region (6).

As expected, PAO1 and PAO1 Δ Pf4 were positive for the Pf6 unique region and negative for the Pf6 screening primers. Interestingly, the putative PAO1 Δ Pf4 Δ Pf6 mutant was positive for both Pf6 screening primers and the Pf6 unique region. It should be noted that this transformant was made in the

PAO1 Δ Pf4 background and should only have Pf6 prophage present in the genome. Sequencing of the Pf6 unique region showed 100% identity to the unique region (data not shown), suggesting the presence of Pf6 sequences in putative PAO1 Δ Pf4 Δ Pf6 deletion strain despite being gentamicin resistant. Thus, a Pf6 deletion mutant was not obtained using any of the methods described here.

3.3.2. Phage population dynamics

Effluent from the biofilm collected over 7 d was used to determine the phage titre and also to extract the bacteriophage DNA to determine the Pf4 and Pf6 population dynamics during the *P. aeruginosa* PAO1 and PAO1 Δ Pf4 biofilm development. For the PAO1 biofilm, phage could be detected from day 1 and steadily increased as determined by the plaque formation on the PAO1 Δ Pf4 top layer agar (Figure 3.7a and c). The emergence of superinfective phage occurred on day 4 as seen by the formation of plaques on the wild type PAO1 top layer agar (Figure 3.7a). Since the detection of wild type Pf6 phage can only be determined using a host lacking the Pf6 prophage, total Pf6 titres could not be quantified. However, superinfective Pf6 could be quantified from the *P. aeruginosa* PAO1 top layer agar spotted with PAO1 Δ Pf4 biofilm effluent (Figure 3.7c) and was observed from day 6.

Figure 3.7. Phage production from biofilms formed by wild type PAO1 (a, b) and the PAO1ΔPf4 mutant (c, d).

The bars in black (■) represent the phage titre of the biofilm on a PAO1ΔPf4 top layer agar while the bars in grey (■) represent the biofilm phage titre on a PAO1 top layer agar (a, c). The points in blue (•) and the points in red (■) indicate the amounts of Pf4 or Pf6 DNA, respectively, detected in the biofilm effluent (c, d). The amounts of phage DNA are represented as the calculated copy numbers and the limit of detection set at 1×10^8 (indicated by the dashed line, ---). The biofilms were developed over 7 d. Statistical significance using t-test was also done between the Pf6 copy numbers of PAO1 (b) and PAO1ΔPf4 (d) and revealed a significant difference on day 7 with a P value of 0.0092. The error bars represent standard deviations. For all figures, there were 3 biological replicates with 3 technical replicates for each, i.e. $N = 3$ and $n = 3$.

To understand the dynamics of Pf4 and Pf6 production in the PAO1 and PAO1ΔPf4 biofilms, and as a complement to the top layer agar assay, qPCR was performed on the biofilm effluents over a period of 7 d. As a host lacking the Pf6 prophage was not obtained, plaque formation using a top layer agar assay cannot be used to determine Pf6 titres. Using qPCR, the emergence of Pf6 could be determined through the presence of Pf6 gDNA extracted from the biofilm effluent. All samples showed an increase in phage production over time (Figure 3.7b and d). For the wild type PAO1 biofilm, the Pf4 phage was present at or below the limit of detection (copy number 1×10^8) for days 1 to 2. From day 5 onward, there was a significant increase in the Pf4 phage titre before reaching a maximum on day 7 (copy number 1.03×10^{13}). The Pf6 phage was detected at lower amounts, peaking on day 7 with a copy number of 3.53×10^{10} . In contrast, while the timing of appearance was the same (day 3), the amount of Pf6 phage DNA detected in the PAO1ΔPf4 biofilm was higher, with a maximum copy number of 8.61×10^{10} on day 7.

3.3.3. Transcriptomic comparison of *P. aeruginosa* PAO1 and PAO1ΔPf4 biofilms

To investigate the effect of the Pf4 prophage on gene expression, transcriptomic analyses of the PAO1 and PAO1ΔPf4 biofilms were conducted. Since count values can be zero in certain conditions, pseudocounts were used and were obtained either by using a regularized log (rlog) function or variance stabilizing transformations (VST). The DESeq2 package in R utilizes these two methods to transform data on a log2 scale and the data are normalized with respect to library size. A standard deviation relative to the mean plot based on the various transformations allows a visual representation of the effect of the

transformation to the counts. Both the rlog and VST (Appendix Figure 5b and c) methods worked well. However, upon inspection of the transformed values, the rlog approach created negative values (data not shown) when the counts of specific RNA reads were zero and thus was not suitable for PERMANOVA analysis. The VST method was therefore chosen to perform downstream analyses as the experimental conditions in this study were expected to have zero values in the data, for example for phage gene expression data, when comparing between the wild type and phage deletion strains.

Before performing PERMANOVA, clustering of the data was performed and represented by both the PCA plot (Figure 3.8a) and as a heatmap (Figure 3.8b). These steps were performed as data quality assessment and controls in parallel to the differential expression analysis. The purpose was to test if any particular samples suffered from an abnormality would negatively influence the analysis. The heatmap showed that the samples and their replicates were clustered as per their experimental conditions (2 d PAO1, 2 d PAO1 Δ Pf4, 7 d PAO1, and 7 d PAO1 Δ Pf4) (Figure 3.8b). These observations were also reflected in the PCA plot showing that the replicates within conditions are similar to each other and are clustered well with no outliers (Figure 3.8a).

Figure 3.8. Clustering of samples through principle component plot (a) and heatmap showing sample-to-sample distances (b).

The samples labelled WT are *P. aeruginosa* PAO1 wild type biofilms and the samples labelled Pf4 are the *P. aeruginosa* PAO1ΔPf4 biofilms. PCA plot shown in the 2D plane spanned by their first two principal components to visualize the overall effect of experimental covariates and batch effects (a). The heatmap shows the Euclidean distances between the samples calculated from the variance stabilizing transformations (b).

PERMANOVA results based on the factors showed that there was interaction between strain and day ($P < 0.001$) and that there were also significant differences between the different strains ($P < 0.001$) and days ($P < 0.001$) (Table 3.3). As a test for the homogeneity of variance across the samples for strain and days, dispersions between PAO1 and PAO1ΔPf4, and between days 2 and 7 were tested (Table 3.4). The test showed no significant differences ($P > 0.1$) in dispersion. From the raw counts of the data, differential expression analysis was performed to generate log2 fold changes, p values, and adjusted p values. The function generating the results table automatically performs independent filtering based on the mean of normalized counts each

gene. This optimizes the number of genes, which have an adjusted p value below the set false discovery rate of 0.1. The datasets extracted were for the comparisons between: (1) 2 d PAO1 Δ Pf4 vs 2 d PAO1; and (2) 7 d PAO1 Δ Pf4 vs 7 d PAO1. The data was further filtered to show only the differentially expressed genes that have an adjusted p value < 0.05 .

Table 3.3. Test for the effect of sample conditions on gene expression.

	Df	Sum of Squares	Mean Squares	F. model	R ²	P value	
Strain	1	0.0008048	0.00080477	6.5767	0.25361	1e ⁻⁰⁴	***
Day	1	0.0008126	0.00081260	6.6407	0.25608	3e ⁻⁰⁴	***
Strain:Day	1	0.0005770	0.00057696	4.7150	0.18182	8e ⁻⁰⁴	***
Residuals	8	0.0009789	0.00012237		0.30850		
Total	11	0.0031733			1.00000		

The PERMANOVA was performed with using the following call in the Vegan package in R: `adonis (formula = df ~ type * day, permutations = 9999)`. P value: <0.001 (***).

Between the 2 d PAO1 Δ Pf4 and wild type biofilms, there were a total of 1,294 significantly and differentially expressed genes with a log2 fold change value more/less than ± 1.0 , of which 381 were hypothetical, unclassified or unknown based on PseudoCAP annotations (Winsor *et al.*, 2011). Based on PseudoCAP annotations, other than the hypothetical proteins, the top 20 downregulated genes in 2 d PAO1 Δ Pf4 relative to the wild type were primarily related to secreted factors associated with virulence, including the Rhl quorum sensing and pyochelin-related proteins (Table 3.5). A non-coding RNA, *phrD*, was also downregulated. Interestingly, many other virulence factor related genes were also downregulated in PAO1 Δ Pf4 relative to PAO1, such as four

genes associated with Las elastase and protease, genes associated with the Type III secretion system (T3SS) and sixteen pyoverdine-related genes (Table 3.6). The top 20 upregulated genes in 2 d PAO1ΔPf4 biofilms relative to wild type PAO1 include genes associated with CupA, which are motility and attachment related proteins as well as proteins associated with adaptation and protection, such as multidrug efflux and copper resistance genes (Table 3.7).

Table 3.4. Permutation test for homogeneity of multivariate dispersion.

	Df	Sum of Squares	Mean Squares	F	Permutations	P value
Strain	1	2.14e ⁻⁰⁵	2.14e ⁻⁰⁵	1.8703	9999	0.1605
Residuals	10	1.14e ⁻⁰⁴	1.14e ⁻⁰⁵			
Days	1	9.15e ⁻⁰⁸	9.15e ⁻⁰⁸	0.0319	9999	0.8423
Residuals	10	2.87e ⁻⁰⁵	2.87e ⁻⁰⁶			

Spatial median used in the analysis done using the vegan package in R. Significant differences, $p < 0.05$.

Between the 7 d PAO1ΔPf4 and wild type PAO1 biofilms, there were 324 significantly, differentially expressed genes (log2 fold change value more/less than ± 1.0), of which 220 were hypothetical, unclassified or unknown proteins. The top 20 genes downregulated in 7 d PAO1ΔPf4 biofilms included the Pf4 genes (data not shown) and these were excluded from the list shown here (Table 3.8). The top 20 downregulated genes were mostly hypothetical proteins and membrane proteins involved in the transport of small molecules (Table 3.8). Interestingly, in 7 d PAO1ΔPf4 biofilms (Table 3.9), the genes most highly induced were associated with mobile genetic elements such phage and transposons. Another interesting observation was the upregulation of genes

related to virulence factors in 7 d PAO1ΔPf4 biofilms relative to 7 d PAO1, which include two related to Type VI secretion system (T6SS), four pyoverdine, two pyochelin and three alkaline metalloprotease related genes (Table 3.10).

The distribution of differentially expressed genes based on KEGG top level categorisation were also examined (Figure 3.9). Of the total differentially expressed genes, 65.9% (day 2) and 81.5% (day 7) were not annotated in KEGG and are labelled as “Hypothetical” for PAO1ΔPf4 biofilms relative to PAO1 biofilms. The “Folding, sorting and degradation”, “Glycan biosynthesis and metabolism” and “Replication and repair” top level categories were differentially expressed in 2 d biofilms (Figure 3.9a) but not in 7d biofilms (Figure 3.9b). The genes are responsible for functions such as RNA degradation and protein export, lipopolysaccharide and peptidoglycan biosynthesis as well as the base excision and mismatch repair mechanisms. Gene set enrichment analysis was then used to determine a common “theme” in data of differentially expressed data where genes under the same pathway would be overrepresented in response to a set condition or perturbation (Hung *et al.*, 2012). A more detailed KEGG module annotation was used and revealed that 29 modules were significantly enriched in 2 d PAO1ΔPf4 biofilms relative to PAO1 biofilms (Appendix Table 7). Of these modules, the groups “Biosynthesis of siderophore group non-ribosomal peptides”, “Bacterial secretion system” and “Starch and sucrose metabolism” were distinctly downregulated for 2 d PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms (Appendix Table 7). In contrast, only 2 categories were significantly enriched in 7 d PAO1ΔPf4 biofilms relative to PAO1 biofilms. The categories “Biosynthesis of siderophore group non-ribosomal peptides” and “Bacterial chemotaxis” were significantly

upregulated and downregulated, respectively (Appendix Table 8). Given the observation that the siderophore biosynthesis seemed to be significantly differentially expressed and that 2 d biofilms had a larger number of differentially expressed genes, a set of pyoverdine proteins were used in validation of RNAseq data using qRT-PCR (Appendix Table 9). The data showed that the RNAseq and qRT-PCR data correlated well with each other.

Table 3.5. Top 20 genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 2.

Downregulated in PAO1ΔPf4					
Locus tag	Name	Description	log2	padj	PseudoCAP
PA3479	<i>rhIA</i>	Rhamnosyltransferase chain A	-8.44	1.35e ⁻⁴⁹	Secreted factors (toxins, enzymes, alginate)
PA3478	<i>rhIB</i>	Rhamnosyltransferase chain B	-6.26	6.15e ⁻⁵³	Secreted factors (toxins, enzymes, alginate)
PA1900	<i>phzB2</i>	Phenazine biosynthesis protein	-6.19	1.84e ⁻¹⁵	Secreted factors (toxins, enzymes, alginate)
PA4226	<i>pchE</i>	Dihydroaeruginoic acid synthetase	-6.14	1.65e ⁻³⁵	Secreted factors (toxins, enzymes, alginate); Transport of small molecules
PA1707	<i>pcrH</i>	Regulatory protein	-5.92	7.15e ⁻¹⁰	Secreted factors (toxins, enzymes, alginate); Protein secretion/export apparatus
PA1168	PA1168	Hypothetical protein	-5.88	1.49e ⁻⁶⁹	Hypothetical, unclassified, unknown
PA2069	PA2069	Carbamoyl transferase	-5.85	1.79e ⁻¹³	Putative enzymes
PA4570	PA4570	Hypothetical protein	-5.73	3.23e ⁻⁵⁰	Hypothetical, unclassified, unknown
PA4141	PA4141	Hypothetical protein	-5.62	7.03e ⁻⁹⁵	Hypothetical, unclassified, unknown
PA4225	<i>pchF</i>	Pyochelin synthetase	-5.55	2.13e ⁻²⁶	Secreted factors (toxins, enzymes, alginate); Transport of small molecules
PA4228	<i>pchD</i>	Pyochelin biosynthesis protein	-5.55	2.60e ⁻³⁶	Secreted factors (toxins, enzymes, alginate); Transport of small molecules
PA3332	PA3332	Hypothetical protein	-5.54	4.50e ⁻²⁸	Hypothetical, unclassified, unknown
PA3331	PA3331	Cytochrome P450	-5.47	3.27e ⁻⁴²	Adaptation, Protection; Carbon compound catabolism

Table continued on next page

Table continued from previous page

PA3417	PA3417	Pyruvate dehydrogenase E1 component subunit alpha	-5.37	6.07e ⁻¹³	Energy metabolism
PA4221	<i>fptA</i>	Fe(III)-pyochelin outer membrane receptor	-5.32	2.19e ⁻⁶⁵	Transport of small molecules
PA0714.1	<i>phrD</i>	ncRNA	-5.31	1.82e ⁻⁰⁸	Non-coding RNA gene
PA4222	PA4222	ABC transporter ATP-binding protein	-5.29	8.49e ⁻¹¹	Transport of small molecules
PA5383	PA5383	hypothetical protein	-5.251	9.06e ⁻³⁵	Hypothetical, unclassified, unknown
PA2146	PA2146	hypothetical protein	-5.226	2.37e ⁻⁵³	Hypothetical, unclassified, unknown
PA1183	<i>dctA</i>		-5.132	5.63e ⁻⁰⁹	Transport of small molecules

Table 3.6. Virulence associated genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 on day 2.

Downregulated in PAO1ΔPf4				
Locus tag	Name	Description	log2	padj
PA1871	<i>lasA</i>	LasA protease	-4.08	4.058e ⁻²⁵
PA3724	<i>lasB</i>	Elastase LasB	-4.66	4.350e ⁻⁷⁷
PA1432	<i>lasI</i>	Autoinducer synthesis protein	-1.14	6.589e ⁻¹³
PA1430	<i>lasR</i>	Transcriptional regulator	-1.99	8.525e ⁻¹⁷
PA1718	<i>pscE</i>	Type III export protein	-4.57	2.417e ⁻⁰⁵
PA1694	<i>pscQ</i>	Type III export protein	-2.21	3.852e ⁻⁰²
PA3841	<i>exoS</i>	Exoenzyme S	-4.78	7.213e ⁻³⁵
PA0044	<i>exoT</i>	Exoenzyme T	-3.85	1.519e ⁻¹⁹
PA2191	<i>exoY</i>	Adenylate cyclase	-3.60	2.518e ⁻⁰⁷
PA2386	<i>pvdA</i>	L-ornithine N5-oxygenase	-3.93	9.503e ⁻⁵⁴
PA2399	<i>pvdD</i>	Pyoverdine synthetase D	-3.54	4.025e ⁻³⁸
PA2397	<i>pvdE</i>	Pyoverdine biosynthesis protein	-3.54	1.033e ⁻³⁵
PA2396	<i>pvdF</i>	Pyoverdine synthetase F	-3.15	1.399e ⁻²⁹
PA2413	<i>pvdH</i>	Diaminobutyrate--2-oxoglutarate aminotransferase	-3.73	1.731e ⁻⁴¹
PA2400	<i>pvdJ</i>	Protein PvdJ	-3.40	2.965e ⁻⁴⁷
PA2424	<i>pvdL</i>	Peptide synthase	-3.39	6.419e ⁻⁷⁰

Table continued on next page

Table continued from previous page

PA2394	<i>pvdN</i>	Protein PvdN		-2.86	1.214e ⁻²¹
PA2395	<i>pvdO</i>	Protein PvdO		-3.06	2.552e ⁻¹⁷
PA2392	<i>pvdP</i>	Protein PvdP		-3.56	3.558e ⁻³⁰
PA2385	<i>pvdQ</i>	3-oxo-C12-homoserine lactone acylase		-2.94	3.656e ⁻¹⁹
PA2389	<i>pvdR</i>	Protein PvdR	Pyoverdine iron acquisition proteins	-2.55	2.644e ⁻²¹
PA2426	<i>pvdS</i>	Extracytoplasmic-function sigma-70 factor		-3.63	1.292e ⁻³⁹
PA2390	<i>pvdT</i>	Protein PvdT		-2.50	4.030e ⁻¹⁰
PA2398	<i>fpvA</i>	Ferripyoverdine receptor		-2.66	3.197e ⁻²⁶
PA4168	<i>fpvB</i>	Second ferric pyoverdine receptor		-2.28	7.344e ⁻⁰⁹

Table 3.7. Top 20 genes upregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 2.

Upregulated in PAO1ΔPf4					
Locus tag	Name	Description	log2	padj	PseudoCAP
PA2129	<i>cupA2</i>	Chaperone CupA2	8.34	8.251e ⁻²⁰	Motility & Attachment; Chaperones & heat shock proteins
PA2128	<i>cupA1</i>	Fimbrial subunit	8.10	1.154e ⁻¹¹⁴	Motility & Attachment
PA2494	<i>mexF</i>	Multidrug efflux transporter	6.92	7.289e ⁻⁵⁶	Transport of small molecules; Membrane proteins; Antibiotic resistance and susceptibility
PA2130	<i>cupA3</i>	Usher CupA3	6.91	1.303e ⁻²⁷	Motility & Attachment
PA2131	<i>cupA4</i>	Fimbrial subunit	6.80	1.338e ⁻¹⁷	Motility & Attachment
PA2065	<i>pcoA</i>	Copper resistance protein A	6.07	1.159e ⁻¹⁷	Adaptation, Protection
PA2493	<i>mexE</i>	Multidrug efflux membrane fusion protein	5.48	1.880e ⁻³⁶	Transport of small molecules; Antibiotic resistance and susceptibility
PA2133	PA2133	Cyclic-guanylate-specific phosphodiesterase	5.46	1.272e ⁻⁰⁶	Hypothetical, unclassified, unknown
PA2495	<i>oprN</i>	Multidrug efflux outer membrane protein	5.39	3.618e ⁻¹⁸	Membrane proteins; Antibiotic resistance and susceptibility; Transport of small molecules
PA5530	PA5530	MFS dicarboxylate transporter	5.13	2.738e ⁻³⁷	Membrane proteins; Transport of small molecules
PA2132	<i>cupA5</i>	Chaperone CupA5	4.97	9.823e ⁻⁰⁸	Motility & Attachment; Chaperones & heat shock proteins
PA2064	<i>pcoB</i>	Copper resistance protein B	4.89	3.237e ⁻⁰⁶	Adaptation, Protection

Table continued on next page

Table continued from previous page

PA4623	PA4623	Hypothetical protein	4.77	5.139e ⁻⁰⁸	Hypothetical, unclassified, unknown
PA1982	<i>exaA</i>	Quinoprotein ethanol dehydrogenase	4.59	2.515e ⁻⁰⁹	Carbon compound catabolism
PA2807	PA2807	Hypothetical protein	4.50	1.178e ⁻¹¹	Hypothetical, unclassified, unknown
PA0729.1	PA0729.1	tRNA-Gly	4.17	6.973e ⁻⁰⁶	Non-coding RNA gene
PA2077	PA2077	Oleate 10S-lipoxygenase	4.07	4.145e ⁻⁰⁵	Fatty acid and phospholipid metabolism
PA3532	PA3532	Hypothetical protein	3.98	4.287e ⁻⁰⁷	Membrane proteins
PA3519	PA3519	Hypothetical protein	3.62	7.258e ⁻⁰⁴	Hypothetical, unclassified, unknown
PA2285	PA2285	Hypothetical protein	3.55	6.807e ⁻⁰⁵	Hypothetical, unclassified, unknown

Table 3.8. Top 20 genes downregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 7.

Downregulated in PAO1ΔPf4					
Locus tag	Name	Description	log2	padj	PseudoCAP
PA1894	PA1894	Hypothetical protein	-3.41	2.007e ⁻⁰⁴	Hypothetical, unclassified, unknown
PA3359	PA3359	Hypothetical protein	-3.24	2.877e ⁻⁰²	Membrane proteins
PA5100	<i>hutU</i>	Urocanate hydratase	-3.23	2.466e ⁻²⁷	Amino acid biosynthesis and metabolism
PA1195	PA1195	Hypothetical protein	-2.93	4.756e ⁻⁰²	Hypothetical, unclassified, unknown
PA4306	<i>flp</i>	Type IVb pilin	-2.93	3.867e ⁻⁰⁵	Motility & Attachment
PA4200	PA4200	Hypothetical protein	-2.89	6.342e ⁻⁰⁹	Hypothetical, unclassified, unknown
PA2171	PA2171	Hypothetical protein	-2.81	3.894e ⁻²⁰	Hypothetical, unclassified, unknown
PA0918	PA0918	Cytochrome b561	-2.78	1.171e ⁻¹⁰	Energy metabolism
PA4352	PA4352	Hypothetical protein	-2.73	1.752e ⁻¹²	Hypothetical, unclassified, unknown
PA0514	<i>nirL</i>	Heme d1 biosynthesis protein	-2.70	2.347e ⁻⁰³	Hypothetical, unclassified, unknown; Energy metabolism; Biosynthesis of cofactors, prosthetic groups and carriers
PA5099	PA5099	Probable transporter	-2.64	6.429e ⁻⁰³	Membrane proteins; Transport of small molecules
PA0176	<i>aer2</i>	Aerotaxis transducer Aer2	-2.63	3.462e ⁻⁰⁶	Adaptation, Protection; Chemotaxis
PA1892	PA1892	Hypothetical protein	-2.62	4.801e ⁻⁰²	Hypothetical, unclassified, unknown
PA2121	PA2121	Probable transcriptional regulator	-2.58	7.077e ⁻⁰⁴	Transcriptional regulators

Table continued on next page

Table continued from previous page

PA3853	PA3853	Probable transferase	-2.58	2.230e ⁻⁰²	Putative enzymes
PA4836	PA4836	Hypothetical protein	-2.57	4.668e ⁻⁰³	Hypothetical, unclassified, unknown
PA5094	PA5094	Probable ATP-binding component of ABC transporter	-2.52	6.013e ⁻⁰³	Transport of small molecules
PA5383	PA5383	Hypothetical protein	-2.44	1.033e ⁻⁰⁸	Hypothetical, unclassified, unknown
PA5480	PA5480	Hypothetical protein	-2.37	2.952e ⁻⁰²	Hypothetical, unclassified, unknown
PA5092	<i>hutI</i>	Imidazolonepropionase	-2.34	3.729e ⁻⁰²	Amino acid biosynthesis and metabolism

Table 3.9. Top 20 genes upregulated in PAO1ΔPf4 biofilms relative to wild type PAO1 biofilms on day 7.

Upregulated in PAO1ΔPf4					
Locus tag	Name	Description	log2	padj	PseudoCAP
PA0434	PA0434	Hypothetical protein	7.21	2.095e ⁻¹³	Membrane proteins
PA2493	<i>mexE</i>	Multidrug efflux membrane fusion protein	5.79	1.271e ⁻²⁰	Transport of small molecules; Antibiotic resistance and susceptibility
PA3790	<i>oprC</i>	Copper transport outer membrane porin	5.40	1.465e ⁻¹⁷	Transport of small molecules
PA0433	PA0433	Hypothetical protein	4.65	5.474e ⁻⁰⁶	Hypothetical, unclassified, unknown
PA3785	PA3785	Hypothetical protein	4.43	1.659e ⁻¹⁴	Hypothetical, unclassified, unknown
PA1466	PA1466	Hypothetical protein	4.42	4.585e ⁻⁰⁴	Hypothetical, unclassified, unknown
PA3784	PA3784	Hypothetical protein	4.27	1.474e ⁻⁰⁶	Hypothetical, unclassified, unknown
PA0616	PA0616	Hypothetical protein	4.07	2.718e ⁻⁰⁵	Related to phage, transposon, or plasmid
PA0435	PA0435	Hypothetical protein	4.06	2.070e ⁻⁰³	Hypothetical, unclassified, unknown
PA2494	<i>mexF</i>	Multidrug efflux transporter	3.60	7.471e ⁻³⁷	Transport of small molecules; Membrane proteins; Antibiotic resistance and susceptibility
PA0247	<i>pobA</i>	4-hydroxybenzoate 3-monooxygenase	3.57	1.255e ⁻⁰²	Carbon compound catabolism
PA0635	PA0635	Hypothetical protein	3.39	9.310e ⁻⁰³	Related to phage, transposon, or plasmid

Table continued on next page

Table continued from previous page

PA1970	PA1970	Hypothetical protein	3.36	1.512e^{-03}	Hypothetical, unclassified, unknown
PA2009	<i>hmgA</i>	Homogentisate 1,2-dioxygenase	3.36	6.461e^{-04}	Carbon compound catabolism
PA4891	<i>ureE</i>	Urease accessory protein	3.31	2.230e^{-02}	Biosynthesis of cofactors, prosthetic groups and carriers
PA0640	PA0640	Probable bacteriophage protein	3.26	7.366e^{-04}	Related to phage, transposon, or plasmid
PA4500	PA4500	Probable binding protein component of ABC transporter	3.24	8.450e^{-16}	Transport of small molecules
PA0618	PA0618	Probable bacteriophage protein	3.16	5.192e^{-03}	Related to phage, transposon, or plasmid
PA3229	PA3229	Hypothetical protein	3.13	4.439e^{-24}	Hypothetical, unclassified, unknown
PA0617	PA0617	Probable bacteriophage protein	3.09	1.255e^{-02}	Related to phage, transposon, or plasmid

Table 3.10. Virulence associated genes upregulated in PAO1ΔPf4 biofilms relative to wild type biofilms on day 7.

Upregulated in PAO1ΔPf4				
Locus tag	Name	Description		log2 padj
PA0085	<i>hcp1</i>	protein Hcp1	Type VI secretion associated proteins	1.09 3.618e ⁻⁰⁵
PA3484	<i>tse3</i>	protein Tse3		1.57 2.917e ⁻⁰²
PA2386	<i>pvdA</i>	L-ornithine N5-oxygenase	Pyoverdine iron acquisition proteins	1.11 1.671e ⁻⁰⁴
PA2389	<i>pvdR</i>	protein PvdR		1.11 6.950e ⁻⁰⁴
PA2392	<i>pvdP</i>	protein PvdP		1.44 1.827e ⁻⁰⁵
PA2385	<i>pvdQ</i>	3-oxo-C12-homoserine lactone acylase		1.46 3.369e ⁻⁰⁴
PA4225	<i>pchF</i>	pyochelin synthetase	Pyochelin iron uptake proteins	1.23 4.319e ⁻⁰³
PA4229	<i>pchC</i>	pyochelin biosynthetic protein		2.18 8.361e ⁻⁰³
PA1249	<i>aprA</i>	alkaline metalloproteinase	Alkaline metalloproteinase proteins	1.39 6.255e ⁻¹¹
PA1246	<i>aprD</i>	alkaline protease secretion protein		1.74 4.781e ⁻⁰³
PA1247	<i>aprE</i>	alkaline protease secretion protein		2.33 4.335e ⁻⁰³

Figure 3.9. Early stage biofilms were more differentially expressed compared to late stage biofilms.

The differentially expressed genes were separated by KEGG top level categories for 2 d (a) and 7 d (b) PAO1ΔP4 biofilms relative to wild type PAO1 biofilms. Only the up- (•) and down-regulated (•) genes with a log2 fold change of > 1 are shown. The more detailed KEGG modules and the gene set enrichment analysis can be seen as Appendix Table 7 and Appendix Table 8.

3.4. Discussion

3.4.1. The Pf6 prophage may be essential

P. aeruginosa PAO1 has been shown to harbour the filamentous phages Pf4 (Rice *et al.*, 2009; Webb *et al.*, 2004) and Pf6 (Tay, 2013) and these two phages have been shown to be closely related (Figure 3.10). However, the impact of Pf6 on its host, and whether there are any interactions between Pf4 and Pf6 that can affect the host, is unclear. The presence of two filamentous phages in a bacterial host is not unprecedented as *Vibrio cholerae* also has two copies of the CTX filamentous phage (Davis *et al.*, 2000; Kim *et al.*, 2017). The CTX phage is present either separately on the two chromosomes (as in the classical strains), or as tandem repeats in the El Tor or O139 biotypes (Davis *et al.*, 2000). The CTX phages contribute to the host pathogenicity as the phage encode the cholera toxin and the phage can also transduce non-toxigenic strains and make them pathogenic. In order to explore the role of the Pf filamentous phages during *P. aeruginosa* PAO1 biofilm development, an effort to generate a Pf6 prophage deletion strain as well as a Pf4 - Pf6 double knockout strain was made. The use of Lambda Red mediated homologous recombination, with both the linear fragment and plasmid hosting the Pf6 knockout cassette, failed to generate any Pf6 prophage knockouts. One of the controls used for Pf6 screening was a putative double knockout mutant (PAO1 Δ Pf4 Δ Pf6) and this strain was generated in the background of PAO1 Δ Pf4 which should only have Pf6 present in the genome (Tay, 2013). Despite attempts to screen for a pure isolate of PAO1 Δ Pf4 Δ Pf6, PCR screening still showed the presence of Pf6 in the isolates but not in its original location (Tay, 2013), suggesting that Pf6 might have reinserted itself into the PAO1 genome.

An alternative method could be considered for future attempts at generating a Pf6 knockout. For example, it has been recently shown that the use of a suicide vector containing *sacB* was efficient in generating gene knockouts in PAO1 (Huang & Wilks, 2017). Alternatively, if generating a complete knockout of the Pf6 is not successful, generating a knockout in the integrase of Pf6 would also be of interest. It has been described that the knockout of the Pf4 integrase prevented the secretion of Pf4 phage particles in PAO1 (Secor *et al.*, 2015) and the same could likely be done for Pf6. Although this method would be limited from determining how Pf6 genes can affect host gene expression, it would allow us to probe if Pf6 phage particles are also able to contribute to the biofilm structure and confer any advantage to the PAO1 host.

Given that the same strategy was used to remove the Pf4 phage, which was observed at a relatively higher efficiency, the failure to delete the Pf6 phage suggests that it may confer some form of advantage to its PAO1 host. Alternatively, it is also possible that Pf filamentous phage integration is essential and that Pf6 is therefore maintained under a mixed culture condition. If this is true, it is likely that when the mixed culture is grown, the Pf6 positive strains would produce Pf6 and re-infect the double knockout mutants, eventually reintegrating with the chromosome. Due to the lack of either a Pf6 single knockout mutant or a Pf4 and Pf6 double knockout mutant, the Pf phage population dynamics was performed with wild type PAO1 and PAO1 Δ Pf4 strains only.

Figure 3.10. Alignment of the Pf4 and Pf6 prophage genomes.

The open reading frames for the Pf4 (right, ■) and Pf6 (left, ■) genomes are indicated. The homologous regions are shown as grey areas and the single nucleotide differences in the Pf6 genome relative to the Pf4 genome are depicted by the white hatched marks within the grey areas. Adapted from Tay (2013).

3.4.2. Pf phage population dynamics in PAO1 and PAO1ΔPf4 biofilms

The benefits of the presence of the Pf4 prophage in the host towards biofilm development and fitness have been established (Rice *et al.*, 2009; Webb *et al.*, 2004). Since PAO1 constitutively produces both filamentous phages, it

potentially incurs a high metabolic cost for the host. However, there was no difference in the growth rate of the wild type and Pf4 deficient mutant (data not shown). Pf4 phage was detected in the effluent at 1 d as compared to Pf6 at 3 d for wild type PAO1 biofilms (Figure 3.7b). Pf6 was found to be above the detectable limit at the same time for the wild type (copy number 4.75×10^8) (Figure 3.7b) and the Pf4 knockout (copy number 5.74×10^8) (Figure 3.7d) biofilms (day 3). The peak copy number of Pf6 detected on day 7, though marginally lower in the WT as compared to Pf4 knockout, was statistically significant ($P < 0.01$) (Appendix Table 5). This suggests that Pf4 possibly plays some sort of role in suppression on the production of Pf6 through an unknown mechanism. Here, the data suggests that the Pf6 phage was also detected in the superinfective form, as there was plaque formation on both the wild type PAO1 (5.29×10^3 PFU/mL) and PAO1 Δ Pf4 (7.16×10^3 PFU/mL) lawns from day 6 (day 6 copy number 1.42×10^9). It would be interesting to sequence superinfective Pf6 that is being expressed by PAO1 Δ Pf4 biofilms to see how the Pf6 becomes superinfective and will be considered in the future.

Both the Pf4 and Pf6 prophages encode an unannotated open reading frame that is homologous to the repressor C protein of phage P2 (Figure 3.10, in the intergenic region of PA4673.13 and PA4673.14 for Pf6, and PA0716 and PA0717 for Pf4) (Tay, 2013; Webb *et al.*, 2004). It is possible that this repressor plays a role in inhibition of Pf6 expression. If the repressor is functional, it may regulate more than just expression or inhibition of Pf6. Other phages, such as Mycobacteriophage L5 of *Mycobacterium smegmatis* (Donnelly-Wu *et al.*, 1993), lambda (Fogg *et al.*, 2010; Oppenheim *et al.*, 2005) and P2 phage (Ljungquist *et al.*, 1984) of *E. coli*, contain repressors that play a role in the host immunity

against reinfection by the same type of phage. There are also other filamentous phages that code for a repressor protein that plays similar roles. For example, the SW1 filamentous phage of *Shewanella piezotolerans* (Wang *et al.*, 2007) and the CTX Φ filamentous phage of *Vibrio cholerae* (Davis *et al.*, 1999) have been shown to also encode repressor genes which regulate the phage gene expression. Recently, it was also observed that in *P. aeruginosa* PAO1, superinfective Pf4 either has a deletion or mutation of the repressor C region (Hui, 2013), suggesting that repressor C plays some form of regulatory role. This protein and its role in superinfection and host phenotypes will be addressed in the following chapter.

3.4.3. Transcriptome of the wild type *P. aeruginosa* PAO1 and the Pf4 knockout biofilms

Biofilm development in *P. aeruginosa* and most other pathogens have been shown to benefit the host in terms of survival in the environment, through increased resistance to stressors such as antimicrobials and host immunity (Corona-Izquierdo & Membrillo-Hernandez, 2002; Davies *et al.*, 1998; Hufnagel *et al.*, 2004; Jefferson, 2004; Yoshida & Kuramitsu, 2002). The deletion of the Pf4 prophage has been shown to reduce the ability of PAO1 to form mature biofilms and increased sensitivity to SDS (Rice *et al.*, 2009). In the same study, using a nasal aspiration mouse model of acute pneumonia, presence of Pf4 was demonstrated to increase the virulence of PAO1. However, the mechanism which mediated this increased virulence is currently unknown. The transcriptomes of PAO1 and PAO1 Δ Pf4 were analysed here to identify phage controlled genes during biofilm development.

For 2 d PAO1ΔP_{f4} biofilms, relative to 2 d biofilms of wild type PAO1, it was observed that there was downregulation in genes associated with quorum sensing, iron acquisition proteins and other virulence associated genes (Table 3.5 and Table 3.6). For *P. aeruginosa*, it has been established that quorum sensing (i.e. cell-to-cell signalling) is involved in its biofilm development (Davies *et al.*, 1998; De Kievit *et al.*, 2001). A *lasI* mutant formed thin, undifferentiated biofilms that was also sensitive to SDS treatment, and supplementation of the signal molecule reverted the phenotype of the mutant biofilm to be more like the wild type (Davies *et al.*, 1998). The quorum sensing systems of *P. aeruginosa* also regulate the expression of other virulence genes (Glessner *et al.*, 1999; Pesci *et al.*, 1997), and in this study the T3SS-associated and iron acquisition-associated proteins (Table 3.6). The T3SS proteins in *P. aeruginosa* are used to inject effector proteins in breached epithelial barriers in airways. Once injected, the effector proteins retard wound healing which allows further exploitation of the breach (Hauser, 2009). Pyoverdine is one of the major siderophores of *P. aeruginosa* (Braud *et al.*, 2010) and pyoverdine-deficient mutants have been shown to be avirulent when injected into a burnt mouse infection model (Meyer *et al.*, 1996). Additionally, the genes that are part of the “Biosynthesis of siderophore group nonribosomal peptides” and “Bacterial secretion system” KEGG modules were significantly enriched and downregulated (Appendix Table 7). The role of prophage in increasing virulence of their host has also been demonstrated in other bacteria. For example, the presence of the phage MDAΦ in *Neisseria meningitidis* increased colonisation of epithelial cells (Bille *et al.*, 2017), and the Sp5 and Sp15 prophages that encode Shiga toxin in *E. coli* O157:H7 which damages the eukaryotic host by

inhibiting protein synthesis when an adenine residue from the 28S RNA of the ribosome 60S subunit is cleaved (Hayashi *et al.*, 2001; Ohnishi *et al.*, 2001). The upregulation of virulence associated genes when the Pf4 prophage is present in the host genome suggests a possible role of Pf4, or one of the genes of Pf4, in the regulation of virulence of *P. aeruginosa* PAO1. The observations may also explain why the *P. aeruginosa* PAO1 wild type is more virulent than the Pf4 knockout mutant (Rice *et al.*, 2009).

In the absence of the Pf4 prophage, 2 d PAO1 Δ Pf4 biofilms relative to the wild type showed upregulation of genes associated with motility and attachment, and also adaptation and protection proteins (Table 3.7). The CupA1-A5 proteins are involved in the assembly of fimbrial subunits and the presence of these adhesins, in addition to other motility and adhesin proteins such as flagellar (Flg) proteins, suggests that PAO1 possess a wide range of factors that function in biofilm formation (Vallet *et al.*, 2001). Many KEGG modules related to cell metabolism like “Metabolic pathways” and “Carbon metabolism”, as well as the “Mismatch repair” mechanism were also found to be significantly enriched and upregulated (Appendix Table 7). The results suggest that these genes were upregulated to support biofilm development in the absence of the Pf4. In *E. coli* biofilms, it was observed that excision of the cryptic prophages, CP4-57 and DLP12, increased biofilm formation by up to 17 fold and transcriptomic analyses revealed upregulation of the motility genes *flg*, *flh* and *fli* (Wang *et al.*, 2009). Interestingly, in the 7 d biofilms of the PAO1 Δ Pf4 mutant relative to the wild type, there were more virulence-associated genes that were upregulated, including the T6SS, iron acquisition-related proteins as well as alkaline metalloproteinase proteins (Table 3.10). The “Biosynthesis of

siderophore group nonribosomal proteins” KEGG module was also significantly enriched and upregulated in 7 d PAO1ΔPf4 biofilms (Appendix Table 8). This agrees with the observation that extended incubation of streaked agar plates and liquid culture of Pf4 deficient mutant was greener as it produced more pyoverdine (data not shown). Furthermore, the overall transcriptome was more similar to the 2 d and 7 d PAO1 transcriptome relative to that of the Pf4 knockout strain and clustered together (Figure 3.8b). The data suggests that Pf4 deficient mutants would be less virulent due to the delayed expression of virulence factors and that Pf4-deficient mutants could potentially be as virulent as wild type PAO1 once the biofilm matures (e.g. in the scenario of a chronic infection).

In conclusion, the data in this study suggest that the two filamentous phages play important roles in *P. aeruginosa* PAO1, regulating host gene expression and function, and may explain why these filamentous bacteriophages are not eliminated by their host. There seemed to be maintenance of a population with the Pf6 prophage, as approaches that were previously successful in generating the Pf4 knockout was not readily successful here. There was also a slight, but significant, suppression of Pf6 phage production when the Pf4 prophage was present. Transcriptomic analyses suggested the role of the Pf4 in regulating the expression of virulence factor-associated genes that explain why the wild type *P. aeruginosa* PAO1 is more virulent than the Pf4 knockout strain. These data also suggest that the phage may directly exert its effects on gene expression at the transcriptional level. The mechanism of how the phage controls gene expression in the wild type was explored in Chapter 4.

Chapter 4 – The role of the repressor C of the *Pseudomonas aeruginosa* PAO1 filamentous phages in global gene regulation

4.1. Introduction

Heritable traits are acted on by natural selection, where the frequency of beneficial alleles increase in response to the given environment while deleterious alleles decline. This process is known as adaptive evolution (Boundless, 2017). In adaptive evolution, stabilizing, directional and diversifying selection are several selection processes that can affect population variation. Diversifying selection in a bacterial population supports a range of ecological and clinical phenomena such as antibiotic resistance (Osorio *et al.*, 2013) and niche adaptation (Koh *et al.*, 2012; Shea *et al.*, 2011). In this process, single populations explore multiple fitness peaks through genetic variation, favouring extreme values rather than intermediate ones, resulting in multiple subpopulations with different adaptive mutations (Boundless, 2017; Maynard Smith, 1962).

Clinical isolates and dispersal populations of lab-grown bacteria produce phenotypic variants at high frequency (Bernardy *et al.*, 2016; Drenkard & Ausubel, 2002; Hashem *et al.*, 2017; Hoffmaster *et al.*, 2008; Ramazanzadeh *et al.*, 2013; Valle *et al.*, 2007; Woo *et al.*, 2012). In biofilms, genetic variants are commonly observed during biofilm development (Boles *et al.*, 2004; Lee *et al.*, 2015) and it is possible that diversifying selection causes the diversification of biofilms. For example, changes in physiological status due to nutrient availability in different regions of the biofilm, elevated mutation rates in the

centres of biofilm microcolonies and DNA damage may promote diversification (Conibear *et al.*, 2009; Oliver & Mena, 2010). Bacteriophages have also been increasingly appreciated to contribute to the diversification of bacterial species. In *Salmonella*, it was demonstrated that a prophage-encoded superoxide dismutase was the primary oxidative stress defence mechanism for the bacterium (Farrant *et al.*, 1997). Additionally, in *Escherichia coli* O157 and *Corynebacterium diphtheriae*, prophage have been shown harbour genes that encode toxins and virulence factors, which enhance the pathogenicity of the bacterial host (Huang *et al.*, 1987; Pappenheimer, 1977; Wagner *et al.*, 2001). Thus, it has been suggested that prophage contribute to bacterial fitness and evolution (Brussow & Hendrix, 2002).

In *Pseudomonas aeruginosa* PAO1, the Pf4 prophage has been shown to play a crucial role in the development of biofilm where it influences biofilm maturation, cell death, dispersal and variant formation. Furthermore, these effects are seen in biofilms where the Pf4 phage acquires a superinfective phenotype (Rice *et al.*, 2009). Recently, in PAO1, a second closely related filamentous phage was found in a Pf4 deficient strain and designated as Pf6 (Tay, 2008). Unlike phage lambda, the Pf4 and Pf6 prophage continually produces phage particles without resulting in the death of the host (Tay, 2013). For Pf4, cell death is specifically associated with superinfection. Recently, it was observed that the superinfective phages produced by two morphotypes of *P. aeruginosa* PAO1 have mutations involving a gene that has homology to an immunity protein of the P2 and lambda phage, called the repressor C protein. In a small colony variant, there was a 3,410 bp deletion in the 5' end of the Pf4 prophage that encompassed the repressor *c* gene and in a spreader variant,

four single nucleotide mutations were found within, and upstream of, the repressor *c* gene (Hui, 2013). The repressor C protein was hypothesised to mediate host immunity to the Pf4 filamentous phage, although it is likely to function differently from what has been described in lambda and P2 phage. However, the specific mechanism by which the repressor C controls superinfection remains unknown (Rice *et al.*, 2009). Bacteriophage infection has been shown to be separated into lytic or lysogenic cycles and this process is regulated by a bistable genetic switch. In the *E. coli* phage lambda, the repressor C protein represses the expression of the Cro protein, an activator of the lytic cycle (Oppenheim *et al.*, 2005; Roberts & Roberts, 1975). The repressor C protein also confers immunity to the host against reinfection by the same phage. If a related phage enters the lysogenised host, repressor C binds to the lytic operator site and suppresses expression of the lytic phage genes (Ptashne, 2004).

Additionally, it has also been recently shown that phage transcription factors can exert some regulation over the host gene expression. In Enterohemorrhagic *E. coli* (EHEC), the deletion of certain bacteriophage regulatory genes decreased the expression of the locus of enterocyte effacement (LEE)-encoded type III secretion system. This led to the identification of the bacteriophage-encoded Cro which serves as a transcriptional activator of LEE as well as other host virulence genes responsible for fimbria and flagellar expression (Hernandez-Doria & Sperandio, 2018). Similarly, deletion of Pf4 decreased the virulence of *P. aeruginosa* PAO1 (Rice *et al.*, 2009) and it is possible that the repressor C homolog of Pf4 also plays a role in the expression of host genes.

The mechanism by which the repressor C controls superinfection was investigated here. Using a combination of mutants, DNA binding assays and ChIPseq analysis, it was observed that repressor C was responsible for immunity against reinfection by binding to its putative promoter region and that it can control host gene expression by binding to promoters of other host genes. The effect of the RepC protein on controlling gene expression more generally in *P. aeruginosa* PAO1 was examined by transcriptomic analysis and validated the role of RepC in regulating expression of certain host genes observed through ChIPseq.

4.2. Materials and methods

4.2.1. Bacterial strains and culture conditions

All bacterial strains (Table 4.1) were maintained on Luria-Bertani (LB) medium, either in broth or on plates supplemented with 1.5% (w/v) agar when necessary. Bacteria were cultured in either LB broth or M9 minimal growth medium (48 mM Na₂HPO₄, 22 mM KH₂PO₄, 9 mM NaCl, 19 mM NH₄Cl, 2 mM MgSO₄ and 0.1 mM CaCl₂, supplemented with 0.04% (w/v) glucose). For plasmid maintenance in *E. coli*, the medium was supplemented with either 100 µg/mL carbenicillin or gentamicin (Carb100 or Gm100). *P. aeruginosa* PAO1 strains carrying the pJN105 vector were grown in medium supplemented with 100 µg/mL gentamicin (LB Gm100 or M9 Gm100). The cultures were incubated overnight at 37°C, at 200 rpm.

Table 4.1. Bacterial strains and plasmids used.

Strain/Plasmid Host	Characteristics	Source
<i>Pseudomonas aeruginosa</i>		
PAO1	Wild type ATCC strain	Stover <i>et al.</i> (2000)
PAO1ΔPf4::Gm ^r	ΔPf4::Gm ^r , Pf4 prophage deletion mutant	Tay (2013)
PAO1ΔPf4	Pf4 prophage deletion mutant	This study
PAO1ΔPf4_ pJN105	Pf4 prophage deletion mutant carrying pJN105 plasmid. Gm ^r	This study
PAO1ΔPf4_ pJN105_RepC	Pf4 prophage deletion mutant carrying pJN105_RepC_CT6His plasmid. Gm ^r	This study
PAO1ΔPf4_ pJN105_sdmRepC	Pf4 prophage deletion mutant carrying pJN105_sdmRepC_CT6His plasmid. Gm ^r	This study
PAO1_pJN105	Wild type PAO1 carrying the pJN105 plasmid. Gm ^r	This study
PAO1_ pJN105_RepC	Wild type PAO1 carrying the pJN105_RepC_CT6His plasmid. Gm ^r	This study
PAO1_pJN105_ sdmRepC	Wild type PAO1 carrying the pJN105_RepC_CT6His plasmid. Gm ^r	This study
PAO1_pJN105_ Pf6RepC	Wild type PAO1 carrying the pJN105_RepC_CT6His plasmid. Gm ^r	This study
<i>Escherichia coli</i>		
NEB5α	<i>fhuA2</i> Δ(<i>argF-lacZ</i>)U169 <i>phoA</i> <i>glnV44</i> Φ80 Δ(<i>lacZ</i>)M15 <i>gyrA96</i> <i>recA1</i> <i>relA1</i> <i>endA1</i> <i>thi-1</i> <i>hsdR17</i>	New England Biolab

Table continued on next page

Table continued from previous page

TOP10	F ⁻ <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi80/lacZ\Delta M15$ $\Delta lacX74$ <i>recA1</i> <i>araD139</i> $\Delta(ara-leu)$ 7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i> λ -	Invitrogen
-------	--	------------

Plasmids

pFLP2	FLP recombinase. Amp ^r , <i>sacB</i>	Hoang <i>et al.</i> (1998)
pCR4_TOPO	TA TOPO Plasmid for Sequencing. Amp ^r , Kan ^r	Invitrogen
pCR4_RepC	TA TOPO containing RepC insert. Amp ^r , Kan ^r	This study
pCR4_RepC_SDM04	TA TOPO containing RepC insert with SDM04 SNP. Amp ^r , Kan ^r	This study
pCR4_RepC_SDM03	TA TOPO containing RepC insert with SDM04-SDM03 SNP. Amp ^r , Kan ^r	This study
pCR4_RepC_SDM02	TA TOPO containing RepC insert with SDM04-SDM02 SNP. Amp ^r , Kan ^r	This study
pCR4_RepC_SDM01	TA TOPO containing RepC insert with all four SNPs. Amp ^r , Kan ^r	This study
pCR4_motifSNP	TA TOPO containing RepC insert with 788826G>T SNP. Amp ^r , Kan ^r	This study
pUCP18	Escherichia-Pseudomonas shuttle vector. Amp ^r	Schweizer (1991)
pUCP18_RepC_CT6His	pUCP containing RepC with C-terminal hexahistidine tag fusion. Amp ^r	This study
pUCP18_RepC_CT6His_ $\Delta lacZ$	pUCP containing RepC with C-terminal hexahistidine tag fusion and partial <i>lacZ</i> deletion. Amp ^r	This study

Table continued on next page

Table continued from previous page			
pBAD_yhjH	Escherichia-Pseudomonas shuttle vector. Gm ^r	Christensen <i>et al.</i> (2013)	
pJN105_RepC	pJN105 containing RepC with C-terminal hexahistidine tag fusion. Gm ^r	This study	
pJN105_sdmRepC	pJN105 containing mutant RepC with C-terminal hexahistidine tag fusion. Gm ^r	This study	
pJN105_Pf6RepC	pJN105 containing RepC insert with C-terminal hexahistidine tag fusion. Gm ^r	This study	

4.2.2. Genomic DNA and plasmid DNA extraction

Genomic DNA was extracted using the QIAamp® DNA Mini Kit (Qiagen, Germany) while plasmid DNA was extracted with the FavorPrep plasmid extraction kit (Favorgen, Taiwan) as per the manufacturers' instructions. The concentration was assessed using both a NanoDrop 2000 Spectrophotometer (Thermo Scientific, USA) and Qubit® 2.0 Fluorometer (Life Technologies, USA).

4.2.3. Polymerase chain reaction

Polymerase chain reactions (PCR) were performed in a Mastercycler pro (Eppendorf, USA) using either Platinum™ Taq DNA polymerase (Invitrogen, USA) or Q5 DNA polymerase (New England Biolabs, USA) as per the manufacturers' recommendation. The cycling conditions were: (1) initial denaturation at 95°C for 3 min; (2) 25 to 30 cycles of 95°C for 30 s, annealing at temperatures as specified (Table 4.2 or Appendix Table 10) for 30 s, and extension at 72°C for the time as specified (Table 4.2 or Appendix Table 10)

and (3) a final extension at 72°C for 5 (Platinum™ Taq DNA polymerase) or 2 min (Q5 DNA polymerase).

Table 4.2. List of primers used.

Primer	Sequences (5'-3')	T _m (°C)	Product size (bp)
repC_EMSA_F	[Btn]TAGAGAAATAATTACCCTGG	56	133
repC_EMSA_R	[Btn]TTTTGGGGCCGATTTTC	56	133
repC_ORF_FrontU p_F	TAGAGAAATAATTACCCTGG	56	407
repC_P_ORF_R	GCTTCATTTCTTGCCTTCCATCCG	56	407
gapA_EMSA_F	[Btn]ATACCAGCGTAGACGAGGTCAA	60	100
gapA_EMSA_R	[Btn]TTGAAATCCACCGACACCAG	60	100
Pf6repC_EMSA_F	[Btn]TTTCAGTGAAGTTCAGGCGAT AA	65	133
Pf6repC_EMSA_R	[Btn]GGTCTATGCCAGCCTTATCTATT AA	65	133
repC_F_KpnI	ACCAGGTACCTAATGAGCACGTCAG CC	60	312
repC_F_PstI	ACCACTGCAGTAATGAGCACGTCAG CC	60	312
repC_SDM01_F_P stI	ACCACTGCAGTAATGAGCACGCCAG CC	60	312
repC_R_CT6His_ XbaI	TTGTTCTAGATATTAGTGGTGATGGT GATGATGAGCTCCCGCGTTTTGATT G	60	312
Pf6repC_F_PstI	ACCACTGCAGTAATGGAATCAATACA GAGCAG	60	315
Pf6repC_R_ CT6His_XbaI	TTGTTCTAGATATTAGTGGTGATGGT GATGATGAGCTCCCGCGTTTTGTCC GGCCA	60	315
repC_SDM01_C- A_F	GGACAAGTTTaGATTGGCTTC	58	407

Table continued on next page

Table continued from previous page			
repC_SDM01_C-A_R	AATCAAAACGCGGGATAG	58	407
repC_SDM02_A-G_F	CTATCGGCTGgCGTGCTCATA	63	407
repC_SDM02_A-G_R	AGCAAGACTATTAATCAAGAAAATCG	63	407
repC_SDM03_G-T_F	TTTATTCCGGtGAAATAGTTGCC	59	407
repC_SDM03_G-T_R	TTTCTATATGAGCACGTCAG	59	407
repC_SDM04_A-G_F	GCTTTGGGGAgATAATTTCTCCAG	62	407
repC_SDM04_A-G_R	CACGTCAGGGCAACTATTTTC	62	407
pUCP18-RC_deleteLacZ_F	GGTACCTAATGAGCACGTC	57	4814
pUCP18-RC_deleteLacZ_R	AGCTGTTTCCTGTGTGAAATTG	57	4814
repC_qPCR_F	GACTTCTGGGGCGATGGAAC	68	84
repC_qPCR_R	AGCACGGAAGAAATCGACGTC	68	84

[Btn] : 5' biotinylated primers

4.2.4. Gel electrophoresis

Gel electrophoresis analyses were performed using either 1 or 2% (w/v) agarose dissolved in 1× TAE buffer (39.95 mM Tris-base, 11.42% (v/v) glacial acetic acid, 127.29 μM EDTA). One or two microliters of 6× DNA loading dye (Fermentas Life Sciences, USA) was mixed with 10 μL of DNA or 1 μL of GeneRuler™ 1 Kbp or 100 bp DNA ladder (Fermentas Life Sciences, USA), before being loaded into the wells. Electrophoresis was performed at 80 V using

a Bio-Rad Power Pac 300 (Bio-Rad Laboratories, USA) for 80 min and subsequently submerged in a 0.5 µg/mL ethidium bromide staining buffer for 30 min. All gel images were digitally recorded using a Molecular Imager Gel Doc System (Bio-Rad Laboratories, USA).

4.2.5. Site-directed mutagenesis (SDM) of repressor *c*

The wild type (WT) *repC* sequence was amplified with *repC_ORF_FrontUp_F* and *repC_P_ORF_R* primers (Table 4.2) using Platinum *Taq* DNA polymerase (Invitrogen, USA) as it introduces an adenine nucleotide overhang. The overhang is required for accurate insertion of *repC* into the pCR4 plasmid of the TOPO TA Cloning Kit (Invitrogen, USA) before it was transformed into TOP10 cells as per the manufacturers' instructions. Site-directed mutagenesis (SDM) of the four *repC* nucleotide mutations on the resultant plasmid, pCR4_RepC, was carried out using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs, USA) according to manufacturers' instructions. All plasmids were transformed into NEB5α cells through heat shock and grown overnight with constant shaking before being extracted. Briefly, *repC_SDM04_A-G_F* and *repC_SDM04_A-G_R* primers (Table 4.2 and Figure 4.1) first introduced an A to G (788857A>G) nucleotide substitution in pCR4_RepC plasmid. Primers *repC_SDM03_G-T_F* and *repC_SDM03_G-T_R* (Table 4.2 and Figure 4.1) introduced the second substitution (788826G>T) into the plasmid pCR4_RepC_SDM04. Primers *repC_SDM02_A-G_F* and *repC_SDM02_A-G_R* (Table 4.2 and Figure 4.1) were then used for the third substitution (788799A>G) in plasmid pCR4_RepC_SMD03. Finally, primers *repC_SDM01_C-A_F* and *repC_SDM01_C-A_R* (Table 4.2 and Figure 4.1) were used to introduce the last substitution (788570C>A) into

pCR4_RepC_SDM02 to form pCR4_repC_SDM01 containing all four SNPs. A plasmid which only has the SNP in the 788826G>T was also generated using the primers repC_SDM03_G-T_F and repC_SDM03_G-T_R (Table 4.2) to modify this region from pCR4_RepC to generating the plasmid pCR4_motifSNP.

Figure 4.1. Sequence of site-directed mutagenesis of the repressor *c* region.

The direction of mutagenesis from right to left is indicated and performed sequentially using the respective primers (diagonal text) to generate the construct of interest. Single nucleotide polymorphisms are indicated by the red vertical line and the nucleotide substitution. Plasmids with different combinations of the mutations indicated above are listed in Table 4.1 and in Figure 4.3.

4.2.6. Electrophoretic mobility shift assay (EMSA) of the RepC

The *repC*, mutated *repC* and Pf6 *repC* sequences were sent as targets to the Protein Production Platform (PPP) (Biopolis, Singapore) for expression and purification of C-terminal hexahistidine tagged (CT6His) proteins to be used for electrophoretic mobility shift assay (EMSA). Prior to EMSA, the proteins were diluted in 1X TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). All primers used for amplification of EMSA targets were 5' biotinylated (IDT, Singapore) (Table

4.2). The *repC* target was amplified with repC_EMSA_F and repC_EMSA_R while the non-target was amplified with gapA_EMSA_F and gapA_EMSA_R using PAO1 genomic DNA as the template. The *repC** probe, which contained the two SNPs in the promoter region and the motifSNP probe which contained only the SNP in the SDM03 region (788826G>T) (Figure 4.1 and Figure 4.2), were also amplified with the primers repC_EMSA_F and repC_EMSA_R, from the plasmids pCR4_RepC_SDM03 and pCR4_motifSNP, respectively. The Pf6 *repC* probe was amplified with the primers Pf6repC_EMSA_F and Pf6repC_EMSA_R from PAO1 genomic DNA. The target probes used for validation of ChIPseq peaks and potential regulon were also amplified using 5' biotinylated primers (IDT, Singapore) (Appendix Table 10).

Figure 4.2. SNPs upstream of repressor *c* gene for *repC* EMSA probe.

Red lines indicate the location of SNPs in the promoter region of repressor *c*. The direction of arrow represents the transcription orientation in the genome. The repressor *c* is located in the intergenic region between PA0716 and PA0717.

Protein-DNA complexes were separated from free DNA on 6% polyacrylamide DNA retardation gels (Invitrogen, USA). All binding and chemiluminescence reactions were carried out with the Pierce LightShift Chemiluminescent EMSA kit (Thermo Scientific, USA) according to manufacturers' instruction. Binding reactions were performed with protein concentrations in lanes 1 to 10 at either 0 to 90 nM, or 0 to 20480 nM, and 10

fmol of target probe for 30 min at 30°C in RepC binding buffer (10 mM Tris, 1 mM DTT, 2.5% glycerol, 5 µg/µL BSA, 0.05% NP-40, 80 mM KCl, 5 mM MgCl₂ and 2 mM EDTA). After incubation, 5 µL of 5X Novex Hi-Density TBE Sample Buffer (Thermo Scientific, USA) was mixed with the binding reaction. The protein-DNA complex was then electrophoresed for 120 min at 110 V in 0.5X TBE (40 mM Tris-Cl pH 8.3, 45 mM boric acid and 1 mM EDTA) buffer through a 6% polyacrylamide gel at 4°C.

4.2.7. Construction and expression of C-terminal hexahistidine-tagged RepC

The Pf4 *repC* was amplified using the primers repC_F_KpnI and repC_R_CT6His_XbaI (Table 4.2) and the PCR product was then ligated into the *KpnI*-*XbaI* site of pUCP18 to generate pUCP18_RepC. The partial *lacZα* gene upstream of the inserted *repC* in pUCP18_RepC was deleted using the pUCP18-RC_deleteLacZ_F and pUCP18-RC_deleteLacZ_R primers (Table 4.2) to prevent expression of a truncated RepC along with the LacZ fusion. The deletion was carried out using a Q5 SDM Kit (New England Biolabs, USA) according to the manufacturers' instructions, resulting in pUCP18_RepC_CT6His_ΔlacZ. Recombinant plasmid constructs were transformed by heat shock into NEB5α cells and grown overnight. The plasmid constructs were then extracted and electroporated into PAO1 using the GenePulser (BioRad, USA) with the following settings: voltage, 2500 V; capacitance, 25 µF; and resistance, 200 Ω.

The Pf4 *repC*, mutated *repC* and Pf6 *repC* genes were amplified with repC_F_PstI and repC_R_CT6His_XbaI, repC_SDM01_F_PstI and

repC_R_CT6His_XbaI, and Pf6repC_F_PstI and Pf6repC_R_CT6His_XbaI (Table 4.2) primers respectively and ligated into the *PstI*-*XbaI* site of pBAD_yhjH, which has a pJN105 plasmid backbone, to generate the pJN105_RepC, pJN105_sdmRepC and pJN105_Pf6RepC plasmid respectively. The *yhjH* gene was excised with the digestion of pBAD-yhjH using *PstI* and *XbaI* restriction endonucleases. The pJN105_RepC, pJN105_sdmRepC and pJN105_Pf6RepC plasmids were transformed into *E. coli* TOP10 cells by heat shock, grown overnight and extracted. The pJN105 plasmid backbone had a gentamicin resistance cassette. Therefore, to prevent selection of false positive clones, the gentamicin resistance cassette in PAO1ΔPf4::GmR was excised. Briefly, the pFLP2 plasmid was electroporated into PAO1ΔPf4::Gm^r and spread on LB Carb100 plates after recovery. Transformants were then plated on LB agar with 10% sucrose (w/v) to counter select for the plasmid pFLP2. To ensure that the gentamicin resistance cassette was excised, transformants were then spread onto LB Gm100 plates to test for resistance. Transformants that could not grow on LB Gm100 plates were labelled as PAO1ΔPf4 for the insertion of extracted pJN105_RepC, pJN105_sdmRepC and pJN105_Pf6RepC plasmids, generating PAO1ΔPf4_pJN105_RepC, PAO1ΔPf4_pJN105_sdmRepC and PAO1ΔPf4_pJN105_Pf6RepC respectively. The empty pJN105 plasmid was transformed into *P. aeruginosa* PAO1 as a control to produce PAO1ΔPf4_pJN105.

4.2.8. *In vitro* immunity assay of repressor C protein

PAO1ΔPf4_pJN105_RepC was cultured in M9 minimal growth medium (48 mM Na₂HPO₄, 22 mM KH₂PO₄, 9 mM NaCl, 19 mM NH₄Cl, 2 mM MgSO₄, 0.1 mM CaCl₂, 0.4% (w/v) glucose) with 0.25% (w/v) arabinose overnight on an

orbital shaker at room temperature. The culture was passed through a 0.2 µm Acrodisc syringe filter (Pall Life Sciences, USA) to obtain phage particles. Filtered phage effluent was similarly obtained from overnight cultures of PAO1, PAO1ΔPf4, PAO1ΔPf4_ pJN105 and uninduced PAO1ΔPf4_ pJN105_RepC. Each filtered phage effluent was serially diluted for determination of plaque forming units (PFU) using a modified top layer agar (TLA) method previously described by Webb *et al.* (2003). TLA plates were made by seeding the LB agar top layer with 5 mL LB soft agar (0.75% w/v) containing 750 µL of PAO1, PAO1ΔPf4, PAO1ΔPf4_pJN105 or PAO1ΔPf4_ pJN105_RepC M9 medium overnight culture. Ten microliters of each effluent dilution was spotted in triplicate onto TLA plates with PAO1, PAO1ΔPF4, PAO1ΔPf4_ pJN105 or PAO1ΔPf4_ pJN105_RepC lawns. The plates were dried and incubated overnight at 37°C. The PFU per mL for each sample was obtained from dilutions yielding plaque counts between 10 to 100 plaques per spot.

4.2.9. Quantification of *repC* expression by qRT-PCR

The levels of gene expression of *repC* in PAO1, PAO1ΔPF4, PAO1ΔPf4_ pJN105 and PAO1ΔPf4_ pJN105_RepC was determined using qRT-PCR and the primers *repC*_qPCR_F and *repC*_qPCR_R (Table 4.2). A standard curve was generated from the reactions carried out with serial dilutions of pCR4_RepC (Table 4.1). Each qRT-PCR reaction mixture contained: 10 µL of 2X Fast SYBR® Green Master Mix (Applied Biosystems, USA) added to 0.2 µL of forward and reverse primers respectively, and 1 µL of template DNA before being topped up to a final volume of 20 µL with deionised water. All qRT-PCR runs were performed using the StepOnePlus™ System (Life Technologies). The cycling conditions were: pre-cycling at 95°C for 2 min, followed by 40 cycles of

95°C for 10 s and annealing and extension at 60°C for 30 s. Results were analysed using the StepOnePlus™ Software version 2.3 (Life Technologies), using the standard curve to quantify the amount of *repC* transcript in each sample.

4.2.10. Chromatin immunoprecipitation sequencing (ChIPseq) sample preparation

The ChIPseq protocol used here was modified from Bonocora and Wade (2015). Briefly, PAO1_pJN105, PAO1_pJN105_RepC, PAO1_pJN105_sdmRepC and PAO1_pJN105_Pf6RepC was inoculated in LB and incubated overnight in a 37°C, at 200 rpm, rotary incubator. The following day, the overnight culture was diluted to a final OD₆₀₀ value of 0.1 and a volume of 42 mL in a sterile flask. The diluted culture was incubated in a 37°C rotary incubator with shaking at 200 rpm to an OD₆₀₀ of 0.3 before the addition of L-arabinose to a final concentration of 0.25% (w/v) with further incubation for 1 h. Molecular grade formaldehyde (Pierce Protein Biology, USA) was then added to a final concentration of 1% and incubated at room temperature to crosslink any protein-DNA complexes. After incubation, glycine was added to a final concentration of 0.5 M to neutralise any free formaldehyde in the medium. The culture was then pelleted and washed 3 times with 1X TBS (50 mM Tris-Cl, pH 7.5, 150 mM NaCl) before finally being pelleted in a 1.5 mL tube. The pellet was resuspended in 1 mL FA lysis buffer (50 mM HEPES-KOH, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 1% Triton-X100 and 150 mM NaCl) containing 4 mg/mL lysozyme and incubated for 30 min at 37°C followed by 5 min on ice.

To completely lyse the cell pellet and shear the genomic DNA, the samples were sonicated using a VCX 750 probe sonicator (Sonics & Materials Inc., USA) with the program of a 30 s on and off cycle for 30 min at 40% amplitude. After sonication, the lysate was transferred into a 1.5 mL tube on ice before being centrifuged for 5 min at 21,130 g at 4°C to pellet cell debris. The supernatant was transferred to a 15 mL tube containing 2.5 mL prechilled FA lysis buffer and mixed by inversion.

The supernatant, containing the DNA-protein complexes, was then divided into three, 800 µL aliquots to which 20 µL of ChIP-Grade Protein A/G Magnetic beads (Pierce Protein Biology, USA) and 20 µL of anti-His-tag Rabbit mAb (Cell Signalling Technology, USA) were added. This mixture was incubated overnight on a Labquake™ rotisserie (Thermo Fisher Scientific, USA) at 4°C. Protein A/G Magnetic beads and associated protein-DNA complexes were recovered by placing tubes on a magnetic rack for 2 min or until the supernatant had cleared. The supernatant was discarded and 700 µL FA lysis buffer was added and incubated for 3 min at room temperature on the Labquake™ rotisserie. This washing procedure was repeated with 2 rounds of 750 µL FA lysis buffer followed by 2 rounds of 750 µL 10 mM Tris-HCl, pH 7.5.

To elute the cross-linked protein-DNA complexes, the beads were recovered using the magnetic rack before addition of 100 µL ChIP elution buffer (50 mM Tris-HCl, pH 7.5, 1% SDS and 10 mM EDTA). The mixture was incubated at 65°C for 10 min with shaking at 500 rpm in a ThermoMixer® C (Eppendorf, Germany). The magnetic beads were recovered and the supernatant was transferred to a clean 1.5 mL tube before the sample was

boiled in a ThermoMixer® C for 10 min. After boiling, the tubes were spun briefly to collect any condensate.

To recover the DNA regions that were bound by the protein, the samples were purified using the Zymo ChIP DNA Clean & Concentrator (Zymo Research, USA) as per the manufacturers' instructions. The DNA was quantified with a Qubit Fluorometer (Thermo Fisher Scientific, USA) and the quality was assessed with a High Sensitivity D1000 ScreenTape (Agilent, USA) before the DNA was sent for sequencing as 75 bp paired end reads on an Illumina MiSeq platform.

4.2.11. ChIPseq analysis

The quality of the ChIPseq reads were assessed using FastQC version 0.11.5 before it was adapter and quality trimmed using BBDuk version 36.38 (Bushnell, 2015). The data analysis was continued using the trimmed reads in R through the systemPipeR (version 1.12.0) package for pipeline control. The reads were aligned to the PAO1 genome using Bowtie2 version 2.2.6 and the output SAM file was converted to an indexed and sorted BAM file. In order to perform the peak calling, the MACS2 (version 2.1.1.20160309) (Zhang *et al.*, 2008) 'callpeak' function was used with the reads from either the PAO1_pJN105_RepC or PAO1_pJN105_Pf6RepC as the treatment file and the PAO1_pJN105 reads as the control. The identified peaks were annotated using the ChIPseeker package (version 1.14.2) and the sequences of the peaks were extracted using the GenomicRanges package (version 1.30.3) to facilitate motif discovery. The motif discovery was performed using the online tool MEME-ChIP (Machanick & Bailey, 2011) and the options used were set for an expected motif

site distribution of ≥ 1 occurrence. Finally, the top scoring motif was used to screen the PAO1 genome to identify genes that may putatively be regulated by RepC using the online tool Virtual Footprint (version 3.0) (Munch *et al.*, 2005).

4.2.12. Tubing biofilm system

Biofilms were allowed to develop for 2 d at room temperature in 15 cm long silicon tubing size 3.18 mm (Silastic® laboratory tubing, USA) in a continuous flow reactor setup under sterile conditions. The biofilm system was fed with M9 minimal salts medium. During inoculation of the bacterial culture, the pump was switched off and the upstream tubing was clamped to prevent backflow. Using a 3 mL syringe with a 26 G $\times$ 1½" needle, 2 mL of 1×10^8 cells/mL bacterial culture was injected into the lumen of the tubing. The bacterial cells were allowed to attach to the tubing surface for 1 h at room temperature under static conditions. The tubing was then unclamped and flow was restored at a flow rate of 9 mL/h.

4.2.13. Transcriptomic analysis of *repC* complemented biofilms

PAO1, PAO1 Δ Pf4, PAO1 Δ Pf4_pJN105_RepC and PAO1 Δ Pf4_pJN105_sdmRepC biofilms were allowed to develop for 2 d before total RNA was extracted. As there is no Pf6 deficient strain, Pf6 RepC was not included in the transcriptome study. The samples were treated with RNAprotect Bacteria Reagent (Qiagen, USA) before extraction. RNA was extracted using the RNeasy® Mini Kit (Qiagen, USA) inclusive of in-column DNA digestion. To ensure improved removal of contaminating DNA, the samples were further treated with TURBO™ DNase (Ambion®, USA) and DNase removal and RNA cleanup was performed using RNA Clean & Concentrator™ (Zymo Research,

USA) as per the manufacturers' instruction. The RNA was converted into cDNA and sequenced as 100 bp paired end reads using the Illumina HiSeq platform.

The total raw RNA reads were adapter and quality trimmed using BBDMap version 36.38 (Bushnell, 2015) before the rRNA reads were depleted using SortMeRNA version 2.1 (Kopylova *et al.*, 2012) to obtain mRNA reads. The mRNA reads were mapped using Bowtie2 against the *P. aeruginosa* PAO1 genome (NC_002516) modified to include the Pf6 prophage genome. HTSeq (Anders *et al.*, 2015) was used to count the number of reads in each feature with the modified genome as a reference. Differentially expressed genes that were statistically significant was identified using the DESeq2 package (Love *et al.*, 2014) in R and the datasets were tested using PERMANOVA to test for significant differences between the different conditions (strain). The log2 fold change values extracted at the end were based on the false discovery rate cutoff of 0.1.

4.3. Results

4.3.1. Plasmid construction

To obtain target sequences with the mutations in *repC* as observed by Hui (2013), sequential site directed mutagenesis was performed using the pCR4 plasmid containing *repC* using the appropriate primers designed to introduce a single SNP. The first SNP was introduced into the putative promoter region and termed SDM04, followed by SDM03, while SDM02 and SDM01 were SNPs introduced in the coding region of repressor *c* (Figure 4.3).

Figure 4.3. Sites of sequential site directed mutagenesis of *repC*.

The *repC* was inserted into a pCR4 plasmid backbone containing an ampicillin resistance cassette as means for selection as depicted by the plasmid map. Red lines represent the location of the singular SNPs made within and upstream of the *repC* gene.

For complementation of RepC in a Pf4 deficient host, the wild type *repC* gene was amplified from *P. aeruginosa* using primers that introduced a 6His tag in the C-terminal and inserted into the *lacZα* gene region of pUCP18 and downstream of the constitutive *lac* promoter. The plasmid, pUCP18_RepC_CT6His, was introduced into *E. coli* TOP10 and *P. aeruginosa* PAO1. To ensure that RepC was expressed, the partial *lacZα* located upstream of the *repC* was deleted to generate pUCP18_RepC_CT6His_Δ*lacZ*. The transformation of this plasmid produced colonies in *E. coli* TOP10 but not in *P. aeruginosa* PAO1. Therefore, the *repC* gene was inserted to the pJN105 plasmid backbone. This plasmid has a low copy number and a controllable

expression of the gene of interest through the *araBAD* promoter. This construct was successfully introduced into *P. aeruginosa* PAO1.

4.3.2. Electrophoretic mobility shift assay of the repressor C proteins

To test whether the Pf4 and Pf6 RepC proteins bind to their putative promoter regions located upstream of the respective *repC* genes, *in vitro* binding assays were performed. The *gapA* gene served as a negative control to determine the specificity of RepC to non-target regions (Figure 4.4). The concentration of the protein gradients used in this assay were from 10 to 90 nM. The Pf4 *repC* promoter showed clear retardation in migration after incubation with Pf4 RepC (Figure 4.4f). Further, binding was observed for RepC at 20 nM and showed two distinct bands that suggested both single (C1) and putative dimer or multimer protein-DNA complexes (C2). At 80 nM of Pf4 RepC, only the dimer or multimer protein-DNA complex was observed.

Superinfective phage variants have mutations in the putative promoter site for RepC, suggesting that the binding site may be altered. Therefore, to determine if Pf4 RepC could bind the mutated promoter, the *repC** and motifSNP probes were used. The motifSNP probe is the Pf4 putative promoter region with a single SNP in the SDM03 region (788826G>T), while the *repC** probe contained the two SNPs (in both the SDM04 and SDM03 region) in the putative promoter region (Figure 4.3). Compared to the WT Pf4 putative promoter sequence (Figure 4.4f), the motifSNP probe showed the formation of a protein-DNA complex only at ≥ 70 nM (Figure 4.4h), while the *repC** probe showed no binding to the Pf4 RepC protein even at 90 nM (Figure 4.4j).

Figure 4.4. EMSA of RepC protein gradients on target DNA.

The binding of RepC to the selected target regions was tested using an increasing protein gradient. In all assays, protein concentrations used in lanes 1 – 10 were 0, 10, 20, 30, 40, 50, 60, 70, 80 and 90 nM respectively. The target probes were the regions upstream of the respective *repC* (133 bp) and were also the putative promoter region, while the non-target *gapA* (100 bp) was randomly chosen as the negative control for the assays. All of the RepC protein variants were synthesized by the Protein Production Platform (Biopolis, Singapore). The notations indicate free DNA (D), protein-DNA complex (C1) and protein dimer-DNA complex (C2).

As Pf6 is highly similar to Pf4 (Tay, 2013), the Pf6 RepC was predicted to behave similarly to the Pf4 RepC. However, the Pf6 RepC showed no sign of binding to its own putative promoter region even at 90 nM, or to any of the other targets that were used here (Figure 4.4, a - e).

4.3.3. The repressor *c* confers immunity to superinfection

To verify whether the RepC protein confers immunity to PAO1 against infection by the filamentous Pf4 phage, plaque assays were conducted for PAO1, PAO1ΔPf4, PAO1ΔPf4_ pJN105 and PAO1ΔPf4_ pJN105_RepC. PAO1ΔPf4, PAO1ΔPf4_ pJN105 and PAO1ΔPf4_ pJN105_RepC do not have

Pf4 in their genomes and thus, it was unsurprising that they did not produce plaques on any of the lawns. As expected, PAO1 effluent containing Pf4 phage formed plaques on PAO1ΔPf4 (5×10^6 PFU/mL) and PAO1ΔPf4_pJN105 lawns (4.86×10^6 PFU/mL), while no plaque formation was observed when PAO1 effluent was spotted on the PAO1 lawn (Figure 4.5a).

Figure 4.5. Immunity assay of repressor C.

Phage effluent from PAO1 was spotted onto the different bacterial lawns and the PFU was enumerated (a). The clear zones in the representative images of the lawns indicated plaque formation. Expression of *repC* was also determined by qRT-PCR; the blue dashed line represents the detection threshold (b). Error bars represent standard deviations. For (a), N = 3 and for (b), N = 3 and n = 3.

The Pf4 deficient mutant was complemented with the Pf4 RepC (PAO1ΔPf4_pJN105_RepC) and the effluent from wild type PAO1 no longer produced plaques on this strain. This suggests that the RepC protein is sufficient to provide immunity against phage infection. Unexpectedly, no plaques were observed in the absence of arabinose induction. Therefore, the expression of *repC* was determined by qRT-PCR (Figure 4.5b). The copy number of *repC*

detected in wild type PAO1 was 4.77×10^4 , while it was 7.22×10^7 for arabinose induced PAO1 Δ Pf4_pJN105_RepC. In non-induced PAO1 Δ Pf4_pJN105_RepC samples (i.e. no arabinose), the copy number was 1.54×10^5 . This explains the lack of plaques detected when PAO1 effluent was spotted onto arabinose negative PAO1 Δ Pf4_pJN105_RepC lawns. This therefore suggests that the *araBAD* system is leaky in *P. aeruginosa* PAO1. As expected, no expression of *repC* was detected in the uncomplemented Pf4 deletion strain.

4.3.4. ChIPseq analysis of the Pf4 and Pf6 RepC

The results above show that the repressor C is able to bind to its upstream putative promoter region and that complementation with the RepC protein in a Pf4 deficient mutant can confer immunity to infection by Pf4 in an otherwise susceptible host. To determine if RepC can bind and regulate genes outside of the prophage genome, ChIPseq was performed. While this was successful for the wild-type Pf4 and Pf6 RepC, no detectable DNA was obtained for the mutant RepC, sdmRepC. The analysis of the wild type RepC revealed 46 peaks where Pf4 RepC (Figure 4.6a) potentially binds and 466 peaks for Pf6 RepC (Figure 4.6b).

Figure 4.6. Peak coverage plot of Pf4 and Pf6 RepC across PAO1 genome. The relative positions of the detected peaks bound by Pf4 (a) and Pf6 (b) RepC in the PAO1 genome. The peaks heights were the $-\log_{10}(p \text{ value})$ obtained from MACS2 and represents the probability of the peak being a true match.

The sequences associated with the peaks above and a motif analysis was performed using MEME-ChIP to identify a putative binding motif. Four possible motifs were obtained for each RepC protein, ranked by their E-value. In the subsequent analysis, only the top motif for each protein was used. For the Pf4 RepC protein, the top scoring motif, SSRGGGCAAYADYTTCTSGH, was found in 32 of the 46 regions identified in the ChIPseq dataset (Figure 4.7a1). The top scoring motif for Pf6 RepC, MDGGGCTADHMCCCRB TVKSBCSSCACTW, was found in 50 of 466 regions (Figure 4.7b1).

(a) Possible Pf4 RepC binding motifs

Sites	E-value
32	$4.5e^{-23}$
6	$2.9e^{-15}$
6	$4.0e^{-7}$
19	$4.0e^{-7}$

(b) Possible Pf6 RepC binding motifs

Sites	E-value
50	$1.3e^{-306}$
40	$2.2e^{-239}$
27	$1.1e^{-218}$
15	$1.1e^{-150}$

Figure 4.7. Predicted Pf4 and Pf6 RepC binding motifs.

For every motif (a and b), the measurement of conservation is represented as a bit score (y-axis) for every nucleotide in their respective position (x-axis). The occurrence of the motif in the submitted sequences are represented as “Sites”.

4.3.4.1. The Pf4 RepC binding targets

For Pf4 RepC, not including the 46 peaks detected by ChIPseq, an additional 853 potential binding regions were detected by Virtual Footprint using the top scoring motif. Of those, 80 were located in intergenic regions, 766 in coding regions and 7 in non-coding region between genes. One site was the putative promoter region for *repC* (TCAGGGCAACTATTTCCCCGG), which was shown to bind RepC (Figure 4.2 and Figure 4.4). Interestingly, the Pf4 RepC ChIPseq data showed an interaction of Pf4 RepC with the Pf6 *repC* putative promoter region (annotated as the intergenic region of PA4673.13) (Table 4.3). When the Pf4 RepC binding motif sequence was compared against the Pf6 genome, a binding region in the putative promoter of Pf6 *repC* (data not shown) was also found. However, the EMSA performed on the Pf6 *repC* putative promoter, containing this site, did not show any detectable binding by the Pf4 RepC (Figure 4.4).

Other potential targets for regulation by the Pf4 RepC include several virulence-factor related genes such as *phzA1* (a probable phenazine biosynthesis protein), *pilQ* (a type IV fimbrial biogenesis protein), *pldB* (a T6SS effector), *aprA* (a precursor of alkaline metalloproteinase), *fptA* (a precursor of the pyochelin outer membrane receptor) and *cdrA* (an adhesion of a two-partner secretion pathway) (Table 4.3). Other potential Pf4 RepC targets were associated with energy and central metabolism such as *fixC* (an oxidoreductase) and *puuR* (an ABC transporter).

Table 4.3. ChIPseq peaks detected and predicted gene targets of the Pf4 RepC.

Locus tag ^a	Gene ^a	ChIPseq (MACS2)		Binding motif (Virtual Footprint)				Gene product ^a	
		Fold enrichment	-log10 <i>p</i> value	Conserved motif ^b	Strand ^b	Distance from PWM ^b gene start ^b	Location ^b		
PA0717		6.37126	179.659	TCAGGGCAACTATTTCCCCGG	-	302	10.92 Intergenic	Hypothetical protein of bacteriophage Pf1	
PA4210	<i>phzA1</i>	5.29093	85.064	CCGGAGAACTTTTCCCTCGC	-	32	9.76 Intergenic	Probable phenazine biosynthesis protein	
PA5040	<i>pilQ</i>	3.86006	46.604	GCGGGTCAAGCAGTTCCTGGA	-	295	9.84 Upstream ORF PA5041 (pilP)	Type 4 fimbrial biogenesis protein	
PA4673		3.40552	38.710	CAATGGAAATAGTTTCAGAGA	+	16	10.03 Intergenic	Conserved hypothetical protein of Pf6	
.13									
PA2953	<i>fixC</i>	3.49856	34.689	CGAGGGCAACTATATCATCTC	+	-	10.70 ORF	Electron transfer flavoprotein-ubiquinone oxidoreductase	
PA5089	<i>pldB</i>	2.82994	20.510	GCGGGTGAATATCTGCCCCCA	+	-	9.82 ORF	H3-T6SS-dependent phospholipase D effector	
PA1249	<i>aprA</i>	2.83107	18.793	CCAGGCGAACAATTGCCCGCA	-	324	10.39 Intergenic	Alkaline metalloproteinase precursor	

Table continued on next page

Table continued from previous page

PA5103	<i>puuR</i>	2.26708	10.928	GCAAGGAAATCGCTACCCCGT	+	273	10.44	Upstream PA5104	ORF ATP-binding transporters	cassette
PA0140	<i>ahpF</i>			CCGGGGATACGTTTGCCCGAA	+	18	9.54	intergenic	alkyl hydroperoxide reductase subunit F	
PA4221	<i>fptA</i>	Not found in ChIPseq but through Virtual Footprint		CGGAGACGATTTCTTCATAAT	-	45	9.53	intergenic	Fe(III)-pyochelin membrane precursor	outer receptor
PA4625	<i>cdrA</i>			CAATGTGAATTGCTTCTCAGA	-	71	9.63	intergenic	cyclic diguanylate- regulated TPS partner A	

^a ChIP-enriched RepC-binding regions were shown by the PA number of the associated genes; the respective gene names and gene annotations were extracted from <http://www.pseudomonas.com>.

^b Conserved motifs, distance to ATG and the PMW scores were all obtained by Virtual Footprint for the entire genome of *P. aeruginosa* PAO1.

^c Full tables for Pf4 RepC ChIPseq and Pf4 RepC Virtual Footprint are as Appendix Table 11 and Appendix Table 12 respectively.

Seven targets were chosen to validate Pf4 RepC binding as they all contained the predicted conserved motif (Table 4.3). The known binding target Pf4 *repC* putative promoter was used as the positive binding control and the previously used *gapA* was used as the negative control. For this binding assay, a high concentration of protein (80 – 20,480 nM) was used to assay for possible weak interactions. All targets showed binding to Pf4 RepC, although the degree of binding was low for *phzA1* and *fixC* even at extremely high concentrations (up to 20,480 nM of RepC) (Figure 4.8), while *pldB* showed evidence of binding, even at 1,280 nM (Figure 4.8).

Figure 4.8. Validation of Pf4 RepC binding to selected target regions.

The promoter regions were chosen based on Table 4.3. The concentration of Pf4 RepC used (left to right) (nM): 0, 80, 160, 320, 640, 1280, 2560, 5120, 10240 and 20480. The negative control target was *gapA* and the Pf4 *repC* putative promoter region was used as a positive control. Free DNA (D) and protein-DNA complex (C) are indicated.

4.3.4.2. The Pf6 RepC binding targets

Only one target, *hisF2* (a imidazoleglycerol-phosphate synthase, cyclase subunit), was observed in both the ChIPseq and Virtual Footprint data sets (Table 4.5). The other coding regions that were predicted binding sites for Pf6 RepC were PA2225 and PA1371, which are hypothetical proteins, while the rest of the predicted genes included PA0668.4, PA4280.2, PA4690.2 and PA5369.2, which are all components of the 23S ribosomal RNA complex (Appendix Table 14). Given that binding of the Pf6 RepC to its putative target promoter could not be detected and due to time limitations, further validation of these binding sites was not pursued, but should be in future studies.

Table 4.4. ChIPseq peaks detected for Pf6 RepC.

Locus tag	Gene	-log10(p value)	Fold enrichment	-log10(qvalue)	Annotation	Strand	Distance to TSS	Gene product
PA0001	<i>dnaA</i>	14.276	2.242	12.626	Promoter (≤1kb)	+	0	chromosomal replication initiator protein DnaA
PA0004	<i>gyrB</i>	8.95543	1.83173	7.40421	Promoter (≤1kb)	+	0	DNA gyrase subunit B
PA0576	<i>rpoD</i>	5.05791	1.601	3.62893	Promoter (≤1kb)	-	-18	sigma factor RpoD
PA0762	<i>algU</i>	16.64587	2.20102	14.96382	Promoter (≤1kb)	+	0	sigma factor AlgU
PA0913	<i>mgtE</i>	3.57801	1.46752	2.22749	Promoter (≤1kb)	+	-42	Magnesium transporter MgtE
PA0992	<i>cupC1</i>	44.35762	3.34101	42.36456	Promoter (≤1kb)	+	0	fimbrial subunit CupC1
PA1898	<i>qscR</i>	16.52517	2.3598	14.84418	Promoter (≤1kb)	+	0	quorum-sensing control repressor
PA2426	<i>pvdS</i>	81.35161	4.68027	79.09652	Promoter (≤1kb)	+	0	sigma factor PvdS
PA3622	<i>rpoS</i>	8.65425	1.90751	7.10961	Promoter (≤1kb)	-	0	sigma factor RpoS

Table continued on next page

Table continued from previous page

PA3841	<i>exoS</i>	7.68215	1.80304	6.16379	Promoter (≤ 1 kb)	-	0	exoenzyme S		
PA4210	<i>phzA1</i>	61.09945	3.88179	58.9892	Promoter (≤ 1 kb)	+	0	probable protein	phenazine	biosynthesis

^a Full table for Pf6 RepC ChIPseq can be seen as Appendix Table 13.

Table 4.5. ChIP peak detected and the gene predicted to be targets of the Pf6 RepC.

Locus tag	Gene	ChIPseq (MACS2)		Binding motif (Virtual Footprint)				Gene product	
		Fold enrichment	-log10 <i>p</i> value	Conserved motif ^a	Strand	Distance from gene start	Distance from PMW		Location
PA3151	<i>hisF2</i>	6.56812	154.0932	ATGGGCTGGAATCAAGTGTCC CCGCAATT CTCGTTTATACCGCTAAGTATA GGATGCTC ATAACTATGAACAAAATAGAAT CGGCTTTG	- + +	226 195 165	30.66 32.43 60	Coding region	Imidazoleglycerol-phosphate synthase, cyclase subunit

^a The conserved motifs shown here are for Pf6 RepC predicted motifs 1, 2 and 4 respectively.

^b Full table for Pf6 RepC Virtual Footprint can be seen as Appendix Table 14.

4.3.5. Differential gene expression by complementation of Pf4 RepC and sdmRepC

The ChIPseq analyses for the Pf4 and Pf6 RepC proteins suggest that both exert some form of control on the gene expression in *P. aeruginosa* PAO1. In the Pf4 prophage deficient mutant, virulence factor associated genes were downregulated (Chapter 3) and the PAO1ΔPf4 mutant was less virulent than its wild type counterpart (Rice *et al.*, 2009). Given that RepC appears to be a regulatory protein, capable of binding the promoters of different genes it is possible that RepC is sufficient to explain those differences in gene expression when the Pf4 phage is deleted. Therefore, the effect of RepC on global gene expression was determined. Biofilms of PAO1, PAO1ΔPf4, PAO1ΔPf4_pJN105_RepC and PAO1ΔPf4_pJN105_sdmRepC were allowed to develop for 2 d in a tubing biofilm system before the total RNA was extracted for sequencing. Since a Pf6 prophage deficient PAO1 could not be generated, complementation of Pf6 *repC* was not performed.

As expected, the samples for the wild type and PAO1ΔPf4 mutant clustered separately, while PAO1ΔPf4_pJN105_RepC and PAO1ΔPf4_pJN105_sdmRepC also clustered away from the PAO1ΔPf4 mutant (Figure 4.9). This showed that complementation with RepC altered gene expression and was more similar to the wild type than the PAO1ΔPf4 deletion strain. Thus, the complementation with *repC* was sufficient to restore wild type expression profiles in the phage deletion mutant.

PERMANOVA showed a significant difference between the different strains ($P < 0.001$) (Table 4.6). No significant difference ($P > 0.1$) in the

dispersions within sample groups (Table 4.7) were observed when the homogeneity of variance for the samples was tested. Furthermore, as the effect of complementing RepC and sdmRepC to PAO1ΔPf4 was the primary interest, only the comparisons of PAO1, PAO1ΔPf4_pJN105_RepC and PAO1ΔPf4_pJN105_sdmRepC to PAO1ΔPf4 where the adjusted *p* value of less than 0.1 were analysed.

Figure 4.9. The complemented strains were more similar to the wild type.

The samples PAO1 (wt, purple), PAO1ΔPf4 (p4, red), PAO1ΔPf4_pJN105_RepC (rC, green) and PAO1ΔPf4_pJN105_sdmRepC (sdm.rC, cyan) were ordinated based on their first two principal components to visualise the overall effect of experimental covariates and batch effects.

Table 4.6. Test for the effect of complementation on gene expression.

	Df	Sum of Squares	Mean Squares	F. model	R ²	P value	
Strain	3	0.00026797	8.9322e-05	2.1749	0.44921	2e-04	***
Residuals	8	0.00032856	4.1070e-05		0.55079		
Total	11	0.00059653			1.00000		

PERMANOVA was performed with using the following settings in the Vegan package of R: `adonis(formula = df ~ type * day, permutations = 9999)`. P value: <0.001 (***).

Table 4.7. Permutation test for homogeneity of multivariate dispersion.

	Df	Sum of Squares	Mean Squares	F	Permutations	P value
Strain	3	2.44e-07	8.15e-08	0.7385	9999	0.1605
Residuals	8	8.82e-07	1.10e-07			

Spatial median used in the analysis done using the vegan package in R. Significant differences, $p < 0.05$.

Between *P. aeruginosa* PAO1ΔPf4 and PAO1, there were 58 differentially expressed genes, of which 17 are hypothetical, unclassified or unknown based on PseudoCAP annotations (Winsor *et al.*, 2011). In this comparison, Pf4 genes were excluded from the table because they are absent in the deletion mutant. The top 20 downregulated genes (by log2 fold change) of known function included those involved in cell metabolism, such as PA3331 (cytochrome P450), *fabH2* and *hutU*, as well as virulence-factor related genes including those from the *rhl* and *las* quorum sensing systems, e.g. *rhlA*, *rhlB*, *rhlC* and *lasB* (Table 4.8).

When PAO1ΔPf4_pJN105_RepC and PAO1ΔPf4 were compared, 182 genes were found to be differentially expressed. Of those, 58 were hypothetical, unclassified or unknown. A significant portion of the genes are annotated as membrane proteins and are suggested to be involved in the transport of small molecules. The majority of the top 20 most upregulated genes were putative enzymes such as PA1202, *wbpH* and *wbpI*, as well as proteins that are part of the cell wall, LPS and/or capsule production, e.g. *cupE1* and *pslC* (Table 4.9).

Between PAO1ΔPf4_ pJN105_sdmRepC and PAO1ΔPf4, there were 162 differentially expressed genes, of which 52 are annotated as hypothetical, unclassified or unknown. These were mainly hypothetical proteins such as PA2699, PA2731 and *vgrG3* (Table 4.10). The other proteins were annotated as cell wall, LPS and capsule proteins such as *cupE1*, *wbpL*, *wbpG*, *wbpJ*, *hsiC3* and *hsiB3* (Table 4.10).

The top 20 most differentially expressed genes for each of the comparisons, regardless whether they were up- or downregulated, were plotted as a volcano plot (Figure 4.10). The common differentially expressed genes when considering all differentially expressed genes (Figure 4.11a) and the top 20 differentially expressed genes (Figure 4.11b) for the sample group comparisons to PAO1ΔPf4 were identified and plotted as an Euler plot. Eight genes were shared between all three comparisons, six that were shared between PAO1 vs PAO1ΔPf4 and sdmRepC vs PAO1ΔPf4, five that were shared between PAO1 vs PAO1ΔPf4 and RepC vs PAO1ΔPf4 and 102 that were shared between sdmRepC vs PAO1ΔPf4 and RepC vs PAO1ΔPf4.

Of the common differentially expressed genes (irrespective of whether they were induced or repressed) (Table 4.11), several were upregulated across all of the comparisons, including PA2698 and PA1202 which are probable hydrolases, hypothetical protein PA2699 and *cdrA* (a cyclic diguanylate-regulated TPS partner). Interestingly, there were also genes that were common with the putative regulon of Pf4 RepC (Appendix Table 15). The genes *aprA*, PA2323, *mexF* and *cdrA* are examples of differentially expressed genes that were also observed and detected by ChIPseq and Virtual Footprint analyses (Table 4.11).

Table 4.8. Top 20 genes downregulated in PAO1ΔPf4 biofilms relative to PAO1.

Locus tag	Name	Description	log2 FC	padj	PseudoCAP
PA4200	PA4200	Hypothetical protein	-2.662	1.92e ⁻²³	Hypothetical, unclassified, unknown
PA2698	PA2698	Probable hydrolase	-2.568	1.68e ⁻²⁸	Putative enzymes
PA3479	<i>rhlA</i>	Rhamnosyltransferase chain A	-2.444	1.00e ⁻¹²	Secreted Factors (toxins, enzymes, alginate)
PA3478	<i>rhlB</i>	Rhamnosyltransferase chain B	-2.410	9.71e ⁻²⁶	Secreted Factors (toxins, enzymes, alginate)
PA0714.1	<i>phrD</i>	PhrD	-2.257	8.83e ⁻¹¹	Non-coding RNA gene
PA3331	PA3331	Cytochrome P450	-2.123	1.00e ⁻¹¹	Adaptation, Protection; Carbon compound catabolism
PA3328	PA3328	FAD-dependent monooxygenase	-1.946	2.05e ⁻⁰⁷	Putative enzymes
PA1202	PA1202	Probable hydrolase	-1.939	2.68e ⁻³⁹	Putative enzymes
PA3329	PA3329	Hypothetical protein	-1.892	5.37e ⁻⁰⁸	Hypothetical, unclassified, unknown
PA2696	PA2696	Probable transcriptional regulator	-1.849	2.11e ⁻⁰⁹	Transcriptional regulators
PA3333	<i>fabH2</i>	3-oxoacyl-[acyl-carrier-protein] synthase III	-1.848	1.33e ⁻⁰⁸	Fatty acid and phospholipid metabolism

Table continued on next page

Table continued from previous page

PA3330	PA3330	Short chain dehydrogenase	-1.741	8.92e ⁻⁰⁶	Putative enzymes
PA3327	PA3327	Non-ribosomal peptide synthetase	-1.736	8.38e ⁻⁰⁷	Adaptation, Protection
PA2699	PA2699	Hypothetical protein	-1.720	2.56e ⁻¹²	Hypothetical, unclassified, unknown
PA0122	<i>rahU</i>	RahU	-1.588	8.83e ⁻¹¹	Adaptation, Protection
PA5100	<i>hutU</i>	Urocanase	-1.557	1.92e ⁻⁰⁵	Amino acid biosynthesis and metabolism
PA1130	<i>rhIC</i>	Rhamnosyltransferase 2	-1.504	6.08e ⁻⁰⁶	Cell wall / LPS / capsule; Adaptation, Protection
PA1168	PA1168	Hypothetical protein	-1.445	6.19e ⁻⁰⁵	Hypothetical, unclassified, unknown
PA3724	<i>lasB</i>	Elastase LasB	-1.319	1.55e ⁻⁰⁴	Secreted Factors (toxins, enzymes, alginate
PA1131	PA1131	Probable major facilitator superfamily (MFS) transporter	-1.275	7.87e ⁻⁰³	Membrane proteins; Antibiotic resistance and susceptibility; Transport of small molecules

Table 4.9. Top 20 genes upregulated in PAO1ΔPf4_ pJN105_RepC biofilms relative to the PAO1ΔPf4 deletion mutant.

Locus tag	Name	Description	log2 FC	padj	PseudoCAP
PA2331	PA2331	Hypothetical protein	2.340	3.80e ⁻¹²	Membrane proteins
PA2698	PA2698	Probable hydrolase	1.477	1.65e ⁻⁰⁸	Putative enzymes
PA1874	PA1874	Hypothetical protein	1.281	6.05e ⁻⁰⁸	Antibiotic resistance and susceptibility
PA4130	PA4130	Sulfite or nitrite reductase	1.095	1.10e ⁻⁰²	Central intermediary metabolism; Adaptation, Protection
PA1202	PA1202	Probable hydrolase	1.043	2.59e ⁻¹⁰	Putative enzymes
PA4648	<i>cupE1</i>	Pilin subunit	1.019	5.44e ⁻⁰³	Cell wall / LPS / capsule; Motility & Attachment
PA2667	<i>mvaU</i>	Heteromeric transcriptional activator	0.989	8.90e ⁻⁰⁷	Transcriptional regulators
PA3150	<i>wbpG</i>	LPS biosynthesis protein	0.967	1.50e ⁻⁰⁶	Cell wall / LPS / capsule
PA2367	<i>hcp3</i>	Hcp3	0.922	4.64e ⁻⁰⁴	Hypothetical, unclassified, unknown
PA3149	<i>wbpH</i>	Probable glycosyltransferase	0.907	2.04e ⁻⁰⁵	Putative enzymes; Cell wall / LPS / capsule
PA0572	PA0572	Hypothetical protein	0.805	7.49e ⁻⁰²	Hypothetical, unclassified, unknown

Table continued on next page

Table continued from previous page

PA3148	<i>wbpI</i>	UDP-N-acetylglucosamine 2-epimerase	0.801	3.31e^{-04}	Putative enzymes; Cell wall / LPS / capsule
PA1318	<i>cyoB</i>	Cytochrome o ubiquinol oxidase subunit I	0.788	2.50e^{-06}	Energy metabolism
PA4306	<i>flp</i>	Type IVb pilin	0.787	1.47e^{-02}	Motility & Attachment
PA2788	PA2788	Probable chemotaxis transducer	0.775	4.74e^{-02}	Adaptation, Protection; Chemotaxis
PA2371	PA2371	ClpV3	0.745	6.21e^{-02}	Translation, post-translational modification, degradation
PA1708	<i>popB</i>	Translocator protein	0.744	5.19e^{-02}	Protein secretion/export apparatus
PA3155	<i>wbpE</i>	UDP-2-acetamido-2-dideoxy-d-ribo-hex-3-uluronic acid transaminase	0.723	1.84e^{-04}	Putative enzymes; Cell wall / LPS / capsule
PA2699	PA2699	Hypothetical protein	0.723	3.26e^{-02}	Hypothetical, unclassified, unknown
PA2233	<i>psIC</i>	PsIC	0.718	9.20e^{-02}	Putative enzymes; Cell wall / LPS / capsule

Table 4.10. Top 20 genes upregulated in PAO1ΔPf4_ pJN105_sdmRepC biofilms relative to PAO1ΔPf4.

Locus tag	Name	Description	log2 FC	padj	PseudoCAP
PA2331	PA2331	Hypothetical protein	3.379	7.52e ⁻²⁷	Membrane proteins
PA4130	PA4130	Sulfite or nitrite reductase	2.800	4.85e ⁻¹⁸	Central intermediary metabolism; Adaptation, Protection
PA2698	PA2698	Probable hydrolase	1.511	5.60e ⁻⁰⁹	Putative enzymes
PA1874	PA1874	Hypothetical protein	1.202	5.62e ⁻⁰⁷	Antibiotic resistance and susceptibility
PA4306	<i>flp</i>	Type IVb pilin	1.015	4.69e ⁻⁰⁴	Motility & Attachment
PA2265	PA2265	Gluconate dehydrogenase	0.977	5.35e ⁻⁰⁴	Carbon compound catabolism
PA2699	PA2699	Hypothetical protein	0.961	1.14e ⁻⁰³	Hypothetical, unclassified, unknown
PA4625	<i>cdrA</i>	Cyclic diguanylate-regulated TPS partner A	0.891	5.54e ⁻⁰⁶	Secreted Factors (toxins, enzymes, alginate)
PA4648	<i>cupE1</i>	Pilin subunit	0.883	2.57e ⁻⁰²	Cell wall / LPS / capsule; Motility & Attachment
PA4660	<i>phr</i>	Deoxyribodipyrimidine photolyase	0.840	6.31e ⁻⁰²	DNA replication, recombination, modification and repair

Table continued on next page

Table continued from previous page

PA3584	<i>glpD</i>	Glycerol-3-phosphate dehydrogenase	0.814	3.02e^{-02}	Central intermediary metabolism; Energy metabolism
PA2366	<i>hsiC3</i>	HsiC3	0.807	2.67e^{-06}	Hypothetical, unclassified, unknown
PA2731	PA2731	Uncharacterized protein	0.791	4.49e^{-02}	Hypothetical, unclassified, unknown
PA0140	<i>ahpF</i>	Alkyl hydroperoxide reductase subunit F	0.781	8.87e^{-02}	Adaptation, Protection
PA2365	<i>hsiB3</i>	HsiB3	0.756	1.00e^{-03}	Hypothetical, unclassified, unknown
PA2373	<i>vgrG3</i>	VgrG3	0.739	4.72e^{-02}	Hypothetical, unclassified, unknown
PA3145	<i>wbpL</i>	Glycosyltransferase	0.727	5.15e^{-03}	Membrane proteins; Cell wall / LPS / capsule
PA3150	<i>wbpG</i>	LPS biosynthesis protein	0.725	1.17e^{-03}	Cell wall / LPS / capsule
PA2371	PA2371	ClpV3	0.721	8.73e^{-02}	Translation, post-translational modification, degradation
PA3147	<i>wbpJ</i>	Probable glycosyl transferase	0.716	1.07e^{-02}	Putative enzymes; Cell wall / LPS / capsule

Figure 4.10. The top 20 differentially expressed genes relative to PAO1ΔP_{f4}.

The genes that had a FDR value (or adjusted p value) of less than 0.05 were labelled red. Genes that were downregulated and upregulated will scatter to the left and right respectively. The top 20 most differentially expressed genes were labelled as text besides their corresponding point.

Figure 4.11. Euler diagrams of differentially expressed genes in all three comparisons.

The differentially expressed genes were plot into Euler diagrams using the Eulerr package (version 4.1.0) to help to identify common genes that were differentially expressed. The plots show all differentially expressed genes (a) and also for those that were only part of the top 20 most differentially expressed genes (b). The stress and diagError values were to see the accuracy of the area distribution for the Euler diagram.

Table 4.11. Common differentially expressed genes, relative to the Δ Pf4 mutant.

Locus tag	Name	Description	Log2 fold change relative to the Δ Pf4 mutant		
			WT	RepC	sdmRepC
PA2698		Probably hydrolase	2.57	1.48	1.51
PA1202		Probable hydrolase	1.94	1.04	0.64
PA2699		Hypothetical protein	1.72	0.72	0.96
PA3326		Hypothetical protein	1.25	-0.69	-0.77
PA4625	<i>cdrA</i>	Cyclic diguanylate-regulated TPS partner A	0.71	0.44	0.89
PA4225	<i>pchF</i>	Pyochelin synthetase	0.44	-0.49	-0.42
PA4226	<i>pchE</i>	Dihydroaeruginosic acid synthetase	0.42	-0.48	-0.37
PA4218	<i>ampP</i>		0.41	-0.33	-0.38
PA4224	<i>pchG</i>	Pyochelin biosynthetic protein	0.49	0.49	
PA2323		Probable glyceraldehyde	-0.56	-0.56	
PA2495	<i>oprN</i>	Multidrug efflux outer membrane protein precursor	-2.98	-2.98	
PA2493	<i>mexE</i>	Multidrug efflux outer membrane protein precursor	-3.78	-3.78	
PA2494	<i>mexF</i>	Multidrug efflux transporter	-4.09	-4.09	
PA3333	<i>fabH2</i>	3-oxoacyl-[acyl-carrier-protein] synthase III	1.85		-1.07
PA3327		Probable non-ribosomal peptide synthetase	1.74		-1.11

Table continued on next page

Table continued from previous page

PA4739		Conserved hypothetical protein	0.59	-0.40
PA4738		Conserved hypothetical protein	0.57	-0.58
PA1249	<i>aprA</i>	Alkaline metalloproteinase precursor	0.53	0.45
PA5429	<i>aspA</i>	Aspartate ammonia-lyase	-0.63	0.63

Cells that are not populated mean that they were not common differentially expressed genes for those comparisons. The log2 fold change for genes that were upregulated are coloured green while downregulated genes are red, based on the comparison.

4.4. Discussion

4.4.1. The role of RepC in immunity and phage conversion to a superinfective state

EMSA assays showed that the wild type RepC protein could bind to its putative promoter site, while the mutant protein could not. Since RepC shares homology with the repressor C protein from *E. coli* phage P2 (Webb *et al.*, 2004) and given that the P2 and λ repressor C play a role in the lytic and lysogenic switch (Berngruber *et al.*, 2010; Oppenheim *et al.*, 2005), it is possible that Pf RepC plays similar roles in *P. aeruginosa*.

The data presented here suggest that the upstream region of the Pf4 *repC* is likely to be its promoter since the protein is able to bind the target probe. Mutations within this region of the Pf4 *repC* promoter were observed by Hui (2013) and McElroy *et al.* (2014) and it was demonstrated here that binding of

the Pf4 RepC to that mutated promoter region was altered. It was also observed that there was a mutation in the highly conserved guanine residue of the conserved motif (Figure 4.12c) (Hui, 2013; McElroy *et al.*, 2014) and the experimental data in this chapter shows that single mutation can inhibit RepC binding. Furthermore, by complementing Pf4 RepC into a Pf4 deficient strain, immunity to Pf4 infection was conferred, shown by the lack of plaque formation when Pf4 was spotted onto lawns of Pf4 RepC complemented PAO1ΔPf4 (Figure 4.5a).

Figure 4.12. Mutation located within the conserved region.

The mutations within the *repC* and its putative promoter region is depicted in the table adapted from Hui (2013) and McElroy *et al.* (2014) (a). Based on the motif and conserved region (b), it is observed that there was a mutation of the highly conserved guanine residue of the motif (c, red arrow).

It is not clear why superinfection leads to host cell death since it has been reported that Ff phage of *E. coli*, which replicate in an uncontrolled fashion, achieving titres of 10¹³ per millilitre do not kill their host (Rakonjac, 2012). However, these data and previous studies (Hui, 2013) provide an insight into

the possible mechanism of how the Pf4 phage is able to cause superinfection (SI) in *P. aeruginosa* PAO1 that is otherwise immune to Pf4 reinfection due to the presence of the prophage within its genome. While the normally expressed and released filamentous phage do not cause cell lysis, it is possible that uncontrolled phage production overtaxes the membrane as the phage particles were released, resulting in lysis. This mechanism is similar to the *E. coli* λ phage where mutations within the *cl* (or repressor *c*) promoter disrupts repressor binding and the induction of the lytic life cycle of the phage (Oppenheim *et al.*, 2005; Ptashne, 2004). Additionally, the Pf4 encoded toxin-antitoxin system could also be responsible in host cell death. For example, if there was a mutation in the antitoxin, the toxin could accumulate and kill the host. However, while mutations in the phage genome have been reported in the RepC ORF and promoter region (Hui, 2013; McElroy *et al.*, 2014), no mutations in any of the other Pf4 genes were observed (McElroy *et al.*, 2014). Alternatively, metabolic stress encountered by the host cell during uncontrolled phage production could activate the toxin-antitoxin system and cause cell death, which has also been described for toxin-antitoxin systems such as CcdB/CCdA and ParE/ParD in *E. coli* (Buts *et al.*, 2005; Jiang *et al.*, 2002; Van Melderren & Saavedra De Bast, 2009).

There are also other filamentous phages that have transcriptional repressors within their genomes. The cf filamentous phage of *Xanthomonas campestris* pv. *citri* has a hypothetical repressor protein that can interfere with DNA-protein and RNA-RNA interactions that regulate lysogeny (Cheng *et al.*, 1999). In the CTX Φ phage of *Vibrio cholerae*, the repressor protein RstR

represses the *rstA* promoter by binding as a dimer-of-dimers and regulates phage replication (Kimsey & Waldor, 2004; Quinones *et al.*, 2005).

However, it is surprising that Pf4 RepC is unable to bind to the putative promoter region of the Pf6 *repC* since a putative binding site was detected in that region. It was also surprising that the sdmRepC was also unable to bind to the mutated promoter regions, given that the promoter mutation and mutations in RepC co-occur. Furthermore, the natural variant from which sdmRepC was identified, is protected from superinfection by its own phage, suggesting an intact immune function (Hui, 2013). Based on phage production profiles (Chapter 3), it was hypothesised that the Pf4 phage can repress Pf6 phage production. However, given that the Pf4 RepC does not bind the Pf6 promoter, it is not clear how the Pf4 phage can control Pf6 phage production. While it is possible, the mechanism is probably not as simple as Pf4 RepC binding to the Pf6 *repC* putative promoter and preventing expression of the Pf6 genes. *V. cholerae* El Tor and O139 strains also carry two copies of the CTX Φ phage (Davis *et al.*, 2000; Faruque & Mekalanos, 2012) which are present as tandem repeats, and usually have relatively high phage production. Deletion of one of the repeats results in low or no phage production (Davis *et al.*, 2000). As observed for the CTX Φ phage, it is possible that production of the Pf4 and Pf6 phage might not be specifically due to expression from their individual clusters but as a combination of the two, although this has to be tested in the future. Similarly, for sdmRepC, it was hypothesized that PAO1 morphotypes that produce SI Pf4 are immune to reinfection by SI Pf4 (Hui, 2013) due to sdmRepC binding to the mutated putative promoter and prevent the expression of SI Pf4 genes. It is also possible that immunity to SI Pf4 reinfection is mediated by a

different mechanism. For example, it has been shown that glycosylation of the *P. aeruginosa* type IV pili and mutations in *pilA* change the host susceptibility to phage infection (Harvey *et al.*, 2018). Based on the accumulated data to date, it is hypothesised that additional regulatory factors are also important in the control of superinfection.

4.4.2. The potential regulon of Pf4 RepC

To assay for the global regulon affected by the Pf4 RepC protein, ChIPseq, RNAseq and protein-DNA binding assays were performed. The Pf4 RepC appears to be able to bind to and regulate expression of certain virulence genes, including *phzA1*, *pldB* and *aprA* (Table 4.3, Figure 4.8 and Appendix Table 15). PhzA1 is a probable phenazine biosynthesis protein and is part of one of two *Pseudomonas* phenazines which are toxic to many other organisms and cell types (Recinos *et al.*, 2012). PldB is a phospholipase D enzyme that is injected by T6SS and degrades phosphatidylethanolamine of bacterial cells (Russell *et al.*, 2013; Sana *et al.*, 2016) as well as facilitates eukaryotic cell intracellular invasion by activation of the PI3K/Akt pathway (Jiang *et al.*, 2014).

Other than directly regulating expression of virulence related genes, an indirect control through the regulation of other transcriptional regulators might also be possible. The bistable expression regulator, BexR, is a LysR-type transcriptional regulator that controls a diverse set of genes such as *aprA* which codes for a secreted virulence factor alkaline metalloproteinase (Turner *et al.*, 2009). Other BexR controlled genes include PA1202, PA2698 and the *mexEF-oprN* operon, which were also differentially expressed in PAO1, RepC- and sdmRepC-complemented PAO1ΔPf4. It is possible that Pf4 RepC can control

the expression of *bexR*. Alternatively, both LysR-type transcriptional regulators and RepC contain a helix-turn-helix (HTH) domain (Koentjoro *et al.*, 2018). It is possible that the presence of RepC and BexR in the cell affect the positive feedback loop which controls the virulence gene expression in *P. aeruginosa* (Turner *et al.*, 2009), and that in turn contributed to the increased virulence of the wild type PAO1 compared to the Pf4 knockout strain in the acute lung infection model (Rice *et al.*, 2009).

4.4.3. The potential regulon of Pf6 RepC

Due to the lack of a Pf6 deficient PAO1 strain, RNAseq was not undertaken for the Pf6 *repC*. There was a large number of peaks detected (466 peaks, Appendix Table 13) for Pf6 RepC and despite that, a motif that could represent and explain for the obtained peaks was not obtained. It is likely that Pf6 RepC binds and behaves differently from Pf4 RepC and that it interacts with other transcriptional regulators. The predicted regulon of Pf6 RepC includes *dnaA* (a chromosomal replication initiator protein), *phzH* (a phenazine-modifying enzyme), 23S ribosomal RNA sequences (PA4280.2), *pilY1* (a type IV fimbrial biogenesis protein) and PA5503 (an ATP-binding component of ABC transporter). Given the high number of putative targets observed and failure to identify a clear consensus binding site may alternatively suggest that many of those targets are false positives. Therefore, further validation work is needed to clarify the role of the Pf6 RepC in regulating gene expression in the *P. aeruginosa* host.

4.4.4. Conclusions

The data presented in this study suggested that the Pf4 RepC is not only important in the control of superinfection, but may also be an important regulator of virulence factor-related genes. The Pf4 RepC binds its putative promoter and regulates the expression of Pf4 genes and has been determined as the minimal factor for prevention of reinfection by Pf4. The Pf4 RepC appeared to both directly and indirectly regulate gene expression of virulence genes and may explain the increased virulence of PAO1. In contrast, the function of Pf6 RepC could not be determined. This work highlights that simple genetic elements, such as the filamentous phage, can play a significant, albeit complex, role in the lifestyle and function of the host bacterium.

Chapter 5 – General discussion and conclusions

5.1. General discussion

Bacteriophages play important roles in the ecology and biology of microorganisms in many environments. For example, by driving the restructuring of microbial community species composition that may ultimately influence the community function. Given that bacteria preferentially exist within biofilms (Costerton *et al.*, 1995), understanding phage interactions with bacteria of biofilm communities and how they can drive changes in community composition and function is essential. Such studies have implications in various fields such as microbial ecology, infectious disease and industrial microbiology. In industrial processes, bacteria are responsible for many important biochemical and biotransformation processes (Barr *et al.*, 2010a; Leroy & De Vuyst, 2004; Sengun & Karabiyikli, 2011; Zhang *et al.*, 2011), or even as biological control agents, i.e. in aquaculture (Gomez-Gil *et al.*, 2000; Verschuere *et al.*, 2000). Biofilms have also been suggested to be responsible for more than 60% of bacterial infections (Chen & Wen, 2011). There are an increasing number of reports for the use of bacteriophages in these industries, though most of them focus on the use of lytic phages to kill and remove unwanted biofilms (Levin & Bull, 2004; Lu & Collins, 2007; Oliveira *et al.*, 2012; Spricigo *et al.*, 2013). The data presented in this thesis studies the dynamics of the bacterial community of granulating sludge and bacteriophages in a laboratory scale reactor operated for SNDPR, a type of nutrient removal process, and used a laboratory model to understand bacteriophage dynamics and mechanisms.

5.1.1. The role of bacteriophages in granulation

In water reclamation, bacteriophages have been known to exist in high abundance (Shapiro *et al.*, 2010; Wu & Liu, 2009) and their detection in high titres are hypothesised to result in deterioration of reactor performance. '*Candidatus Accumulibacter*' are often enriched in enhanced biological phosphorus removal (EBPR) systems as they play a key role in removal of phosphorus (Garcia-Martin *et al.*, 2006; Lu *et al.*, 2006). However, enrichments of '*Candidatus Accumulibacter*' are susceptible to phage predation, and infected floccular and granular sludge rapidly declined in '*Candidatus Accumulibacter*' populations and subsequently phosphorus removal (Barr *et al.*, 2010b). Furthermore, the presence of bacteriophages in sludge structures also suggest their contribution to biofilm formation or that they may be involved in selection of functional communities (Baloch *et al.*, 2008).

Here, the diversity and abundance of bacteria and bacteriophages throughout the aerobic granulation process were characterized using total genomic sequencing of the sludge and reactor effluent (Chapter 2). The data showed that '*Candidatus Accumulibacter*' and '*Candidatus Competibacter*' were the top 2 most abundant genera in the granulating reactor. However, correlation studies showed that only '*Candidatus Accumulibacter*' was positively correlated with increased particle size and decreased SVI₅, which are indicators of granulation. The structure of the aerobic granule has been investigated (Lemaire *et al.*, 2008a; Lemaire *et al.*, 2008b) and it was hypothesized that the outer aerobic layer would be dominated by '*Candidatus Accumulibacter*' while the inner anoxic and anaerobic granular core would be dominated by the '*Candidatus Competibacter*'. It was indeed observed in this study that there is a

high abundance of '*Candidatus Accumulibacter*' and '*Candidatus Competibacter*' (Chapter 2) although the spatial distribution of these two genera were not explored here. The dense and compact structure of granules have been shown to be more tolerant to toxic compounds (Carucci *et al.*, 2010) and also hypothesized to provide increased protection against predation. The biofilm matrix have been shown to reduce phage susceptibility (Darch *et al.*, 2017; May *et al.*, 2011) and biofilms of *Pseudomonas fluorescens* hyperproduce the biofilm matrix in response to phage attack (Scanlan & Buckling, 2012). Predation may significantly impact both floccular and granular sludge. The phenotypic changes observed are likely due to changes in bacteria host gene expression. Therefore, understanding the effect of bacteriophages during biofilm development would potentially provide insights to improve the process of aerobic granulation.

The data presented here also showed that there were changes in abundance of DNA phages as the sludge in the reactor started to granulate (Chapter 2). There was an increase in the amount of Inoviridae detected in the effluent, which suggested that they might be actively produced by the bacterial community during the process of granulation. As Inoviridae are extruded out of the cell, the high densities of the filamentous phage on the cell surface may play a direct role in enhancing biofilm formation through the interaction of the filaments with other components of the matrix (Ghigo, 2001; Reisner *et al.*, 2003; Webb *et al.*, 2004). It has been observed using scanning electron microscopy that aerobic granules have dense lattice of phage-like filaments (Liu & Wang, 2007) similar to that of a *P. aeruginosa* PAO1 biofilm (Webb *et al.*, 2004). However, only the lytic Podoviridae, Microviridae and Myoviridae were positively correlated with the post-initiation stages of granulation (Chapter 2). This may

suggest that while the Inoviridae could facilitate nucleation of granules, granule expansion and maturation may be driven by phage induced bacterial lysis. There might have been lysis of bacteria that are prone to cause reactor bulking as these bacteria could have made the sludge light and fail to form dense granules (Khairnar *et al.*, 2016; Liu *et al.*, 2015) and that the lysis of such organisms released extracellular DNA (eDNA) into the matrix to improve the granule structure. It has been demonstrated in *Streptococcus pneumoniae* and *Shewanella oneidensis* that lysis due to spontaneous phage induction released eDNA into the biofilm matrix and enhanced the biofilm development (Carrolo *et al.*, 2010; Gödeke *et al.*, 2010). Furthermore, phages may also play a role in regulating host gene expression that drives granule formation. The excision of the cryptic prophage CP4-57 in *E. coli* upregulated flagella genes (Wang *et al.*, 2009) which are suggested to be important in good biofilm formation (i.e. thickness of biofilm) (Wood *et al.*, 2006). In *P. aeruginosa* biofilms, filamentous phage genes are highly expressed (Whiteley *et al.*, 2001) and it is possible that phage-encoded genes are important in biofilm formation. It has also been shown that in the absence of phage, *P. aeruginosa* biofilms are less stable than the wild type (Rice *et al.*, 2009). Therefore, the molecular mechanisms of how phage can affect the bacterial host need to be better understood and is also addressed here (Chapter 3 and 4).

Overall, the data indicated that '*Candidatus Accumulibacter*' and physical factors such as the reduction of settling time and hydrodynamic shear force were important for the formation of granules in sequencing batch reactors operated for simultaneous nitrification denitrification and phosphorus removal. Bacteriophages in the system play a role by predation of bacteria that may be

detrimental to formation of granules or by enhancing the structure through cohesion of filament-like structures in the granular matrix.

5.1.2. The role of the Pf4 filamentous phage and its repressor C in biofilm development

Traditionally, studies on the role of bacteriophage have been performed on single species biofilms. In these systems, temperate and non-lytic bacteriophages have been shown to influence host gene expression (Lim *et al.*, 2016; Wang *et al.*, 2009). However, activated sludge systems are complex engineered ecosystems that are made up of highly diverse microbial and viral communities. To better define how bacteriophages can affect biofilm development, the interactions between *Pseudomonas aeruginosa* PAO1 and the Pf filamentous phages were studied here (Chapters 3 and 4). Transcriptomic analysis revealed that the Pf4 prophage was linked to the upregulation of genes involved in virulence, such as the type III secretion system, and the *las* and *rhl* quorum sensing systems (Chapter 3). Furthermore, the two quorum sensing systems are also involved in the regulation of genes that affect biofilm formation (O'Loughlin *et al.*, 2013; Sakuragi & Kolter, 2007), and it was previously shown that the deletion of the Pf4 prophage reduced the ability of the *P. aeruginosa* biofilm to withstand antibiotics and surfactants (Rice *et al.*, 2009). These could suggest that the Pf4 mediates some biofilm control effects through modulation of the QS response.

To better understand the mechanism by which the Pf4 phage can modulate host gene expression, the role of the Pf4 repressor *c* was explored here (Chapter 4). The repressor *c* was previously hypothesized to play a role in

the conversion of the Pf4 into a superinfective state, where either deletion or mutations of the repressor coding region were found in morphotypes of *P. aeruginosa* PAO1 that constitutively produce superinfective Pf4 (Hui, 2013; McElroy *et al.*, 2014). As the emergence of superinfective phages were also associated with key stages of the PAO1 biofilm development (Rice *et al.*, 2009; Webb *et al.*, 2004; Webb *et al.*, 2003), the role of repressor C as a transcriptional regulator was explored. The Pf4 repressor C was found to be the immunity factor that prevented reinfection by Pf4. In support of its roles as a regulatory element, RepC was shown to bind to the promoter region of PAO1 genes and regulate their expression, which can either directly affect the host phenotype or indirectly for genes that are upstream of various pathways. Here, it was shown that in the absence of the Pf4 prophage, virulence factor-related genes were downregulated while genes that supported biofilm development were upregulated (Chapters 3 and 4). Upon complementation of RepC in Pf4 deficient hosts, the gene expression was more similar to the wild type, with the expression of virulence factor-related genes (Chapter 4). It was recently described that a phage encoded transcription factor, Cro, regulated virulence gene expression in enterohemorrhagic *Escherichia coli* (Hernandez-Doria & Sperandio, 2018). Mutations in the promoter region of repressor *c* would explain why the wild type PAO1 is susceptible to superinfective Pf4, while mutations in the coding region probably altered the repressor binding affinity and affected RNA-RNA and DNA-protein interactions. Mutations in a repressor protein and the altered binding affinities has been previously described to contribute to the control of lysogenic and lytic conversions (Cheng *et al.*, 1999; Simons & Kleckner, 1988). However, mutations in the promoter region of repressor *c* does

not seem to be a compensatory mutation to the control of superinfective Pf4 phage expression and is possible that immunity to superinfective Pf4 is due to another mechanism. The role of Pf6 in PAO1 biofilm development remains unclear, although the data suggest that it may be essential since a Pf6 deletion could not be made (Chapter 3).

5.1.2.1. Proposed model

Based on the observations here and elsewhere, mutations in the repressor *c* gene of Pf4 is stimulated by changes in environmental conditions. Moreover, the superinfective form of the Pf4 phage is linked to the establishment of morphotypic variants as the biofilm matures. Here, a model is proposed for the role of repressor C in PAO1 (Figure 5.1). The repressor C in the cell is constitutively expressed and plays a role in the control of host gene expression (Figure 5.1a) and at the same time regulate Pf4 production (Figure 5.1b). The repressor C also confers immunity to Pf4 reinfection (Figure 5.1c). However, for a superinfective Pf4, which has mutations in the repressor *c* promoter region or lacking the region entirely, the regulation of phage production by repressor C is lost. Such mutants subsequently over produce the superinfective phage that leads to host cell lysis and the cell death phenotype observed in wild type PAO1 biofilms (Figure 5.1d). The mutated repressor C of a superinfective Pf4 producing PAO1 has an altered binding affinity but is still able to regulate host gene expression (Figure 5.1e) and the immunity against superinfective Pf4 may be due to a different mechanism, such as the glycosylation of type IV pili where the phage is prevented from host entry (Harvey *et al.*, 2018) (Figure 5.1f). We are currently exploring differences in crystal structures for the Pf4, mutant Pf4 and Pf6 RepC proteins to better understand how differences in amino acid

sequences may alter DNA binding domains. Such information will help to better elucidate the mechanisms of host gene regulation by these phage-encoded regulatory proteins.

The mechanism of how superinfective phage can lead to cell death is still unclear though it was hypothesized here that the unregulated production and extrusion of the phage may cause cell lysis due to stress exerted on the cell wall. Alternatively, the Pf4 encoded toxin-antitoxin system may also play a role where uncontrolled production of the phage caused metabolic stress to the cell. It has been described in other systems that elevated stress activates the toxin-antitoxin system as a mechanism of programmed cell death (Buts *et al.*, 2005; Couturier *et al.*, 1998; Jiang *et al.*, 2002; Van Melderen & Saavedra De Bast, 2009). It has also been proposed that an imbalance in the levels of the replicative form and ssDNA of that phage may influence the control of superinfective events (Russel, 1995) and is crucial for preventing cell death. When superinfective phage infects the cell, host enzymes are actively recruited to convert the ssDNA into the replicative form to produce phage-encoded proteins and to produce new ssDNA. This also puts the cell into a state of metabolic stress and leads to host cell death.

Figure 5.1. Model for the role of repressor c.

The mechanism for *repC/repC** in wild type *P. aeruginosa* PAO1 (WT, top) and PAO1 morphotype constitutively producing superinfective (SI) Pf4 (SI, bottom). In a normal state, *repC* in the Pf4 prophage produces RepC and regulates host and also Pf4 gene expression, contributing to PAO1 virulence (a) and the stable infection of Pf4 (b). When infecting Pf4 comes into WT PAO1, intracellular RepC binds to the *repC* promoter region of the infecting Pf4 and regulate the expression of the infecting phage genes, maintaining stable infection (c). A morphotype of PAO1 constitutively produces SI Pf4. When SI Pf4 infects WT PAO1, intracellular RepC is unable to bind to the *repC** promoter region of the infecting SI Pf4, leading to uncontrolled expression and extrusion of SI Pf4, causing lysis of the infected WT PAO1 (d). In SI Pf4 producing PAO1, the mutated RepC* can also control the host gene expression (e). Since RepC* is unable to bind to the mutated promoter, immunity to both wild type Pf4 and SI Pf4 is likely due to other mechanisms, for example, through glycosylation of type IV pili (f).

5.2. Limitations of this study

The work presented in Chapter 2 reported the dynamics of the virome composition in a lab scale reactor inoculated with sludge from the Ulu Pandan Wastewater Treatment Plant. The reactors were operated to achieve granulation and operated for SNDPR. The findings showed the changes in phage and bacterial compositions and how certain bacteriophage families influenced the bacterial community changes. The data presented were mainly based on a bioinformatics analysis of the sludge metagenome and metavirome. To further study the role of bacteriophages in this system, the reactors could be set up to test different operational conditions as well as to perform bacteriophage addback experiments to observe their effect on bacterial community change and granulation. Multivariate analysis can then be used to tease out the factors that affect granulation. To complement the phage addback, an attempt at identifying phage-host partners could be done through a modified single cell viral tagging technique approach by Deng *et al.* (2014). Bacteriophages in cell-free reactor effluent could be fluorescently stained (e.g. SYBR Gold) and added back to sludge homogenised to single cells (Hugenholtz & Džunková, 2019). As these single cell hosts are not fluorescent, cells that are infected or have phages adsorbed would now be fluorescent and can be separated from non-infected cells by fluorescence-activated cell sorting at single cell resolution. Sequencing these infected single cells and getting metagenome-assembled genomes (MAGs) would allow us to phage-host partners (MAGs identified as host or phages accordingly). Additionally, both transmission and scanning electron microscopy can be employed to better visualise the structure of the granules. Observing the presence of any bacteriophage-like structures in

the granules would shed light if bacteriophages, especially Inoviridae, play a role in granulation as a physical component.

In Chapters 3 and 4, *Pseudomonas aeruginosa* PAO1 was used as a model organism to investigate the role of filamentous phages in biofilm dynamics. The data showed that presence of Pf4 in the host genome contributed to differences in virulence factor expression. As a Pf6 deficient mutant could not be generated, a gap in understanding the roles of filamentous bacteriophage in *P. aeruginosa* biofilm development remains. Moving forward, instead of generating a complete knockout of Pf6 in the wild type and Pf4 deficient mutant, an integrase knockout could first be done. This would generate a mutant strain that would be unable to secrete Pf6 phage particles and first study the role of Pf6 in *P. aeruginosa* during biofilm development. Sequential knockout of Pf6 genes could then be attempted to eventually obtain a complete knockout. Additionally, Chapter 4 expanded on the role of repressor C in immunity to superinfection through a combination of protein-DNA binding assays, gene complementation and sequencing based approaches. A crystal structure of repressor C could be elucidated to model how the protein binds to the consensus sequence and modulate immunity against Pf4 reinfection. It would also be of interest to identify if repressor C interacts with other proteins in *P. aeruginosa*, either on its own or dependent on binding to a consensus sequence, which could further explain the transcriptional differences that was observed in chapters 3 and 4.

5.3. Conclusions

Predation by phages have been demonstrated to be a factor in the development of aerobic granules from activated floccular sludge. However, while the role of Inoviridae as a structural component is yet to be elucidated, there is an increase in Inoviridae abundance as the sludge becomes compact aggregates. Furthermore, distance based redundancy analysis suggested that Inoviridae was one of the two bacteriophage families that was associated with the bacterial community change during the initiation phase. The role of bacteriophages in granulation in this system is mainly inconclusive and requires further testing, though the data suggests their role in bacteria community change. While bacteria from the genera '*Candidatus Accumulibacter*' and other major bacterial groups were positively correlated to the increase in particle size, the reactor operating conditions are likely the main drivers of the granulation process. The findings here further the understanding of the role that bacteriophages could play in driving the initiation and maintenance of granules.

The role of filamentous phages in the development of biofilms was also evaluated using a less complex single-species biofilm. The results demonstrated that the presence of the Pf4 prophage conferred an advantage to *P. aeruginosa* PAO1 in terms of virulence and biofilm development through a phage-encoded transcriptional regulator. More work is needed to validate this control by repressor C. Overall, the study of bacteriophages and how they drive biofilm formation would be useful in enhancing granule stability for longer operations in a full-scale water reclamation plant.

Contributions

Chapter 2

Assoc. Prof Scott A. Rice, Dr Tan Chuan Hao, Dr Martin Tay Qi Xiang, Dr Chan Siew Herng and I initiated the granulating reactor project to study the initiation of granulation. Dr Tan Chuan Hao, Dr Chan Siew Herng and I operated and maintained the reactors. I and Dr Chan Siew Herng sampled the reactor for the nutrient analysis, particle size measurements and DNA extraction from the sludge. Dr Chan Siew Herng performed the microscopy to obtain the images of the sludge. I sampled the effluent for the reactor and extracted DNA from the viral fraction. I performed the metagenome analysis of the sludge and viral fractions.

Chapter 3

I and Dr Martin Tay Qi Xiang initiated the study of the project. The tubing biofilm setup was optimised by Dr Martin Tay and I. I collected the effluent from the setup daily and performed phage enumeration and extracted the phage DNA for qPCR. I performed the qPCR and the analysis. Dr Kelvin Lee Kai Wei assisted with the qPCR analysis. I extracted the RNA from the biofilm and also analysed the RNAseq data with initial guidance from Dr Martin Tay and Dr Kelvin Lee Kai Wei.

Chapter 4

I initiated the study as a follow up of the work that was done by Dr Janice Hui. The Nanyang Technological University's Protein Production Platform synthesized the proteins used in this study. I and Goh Yu Fen performed the

DNA-binding assays, designed and constructed the plasmids that were used in this study, and the *in vitro* immunity assays. I extracted the RNA, performed qPCR and analysed the data. I performed the ChIPseq experiment and analysed the data. I setup the tubing biofilm experiment, RNA extraction from the biofilm for RNAseq and performed the gene expression analysis.

Permissions

Figures that were adapted from publication were either those from an open access article distributed under the terms of Creative Commons Attribution – Non-Commercial License or with permission obtained from RightsLink. The publications and their respective permission license number are:

License number:

4501760174874. Beun *et al.* (1999). Aerobic granulation in a sequencing batch reactor. Water Res.

4503450147734. Hasty *et al.* (2001). Computational studies of gene regulatory networks: in numero molecular biology. Nat Rev Genet.

4503450571409. Secor *et al.* (2015). Filamentous Bacteriophage Promote Biofilm Assembly and Function. Cell Host Microbe.

4503450717611. Anderson, O. R. (2010). Protozoan Ecology eLS.

4500631117069. Wu *et al.* (2014). Simultaneous nitrogen and phosphorus removal in the sulfur cycle-associated Enhanced Biological Phosphorus Removal (EBPR) process. Water Res.

References

- Abedon, S. T. (2011). Bacteriophages and Biofilms. In W. C. Bailey (Ed.), *Biofilms: Formation, Development and Properties* (pp. 58): Nova Science Publishers, Inc.
- Abee, T., Kovács, Á. T., Kuipers, O. P., & van der Veen, S. (2011). Biofilm formation and dispersal in Gram-positive bacteria. *Current Opinion in Biotechnology*, 22(2), 172-179.
- Ackermann, H. W. (2003). Bacteriophage observations and evolution. *Res Microbiol*, 154(4), 245-251. doi:10.1016/s0923-2508(03)00067-6
- Ackermann, H. W., & Kropinski, A. M. (2007). Curated list of prokaryote viruses with fully sequenced genomes. *Res Microbiol*, 158(7), 555-566. doi:10.1016/j.resmic.2007.07.006
- Adav, S. S., Lee, D.-J., & Lai, J.-Y. (2009). Aerobic granulation in sequencing batch reactors at different settling times. *Bioresour Technol*, 100(21), 5359-5361. doi:10.1016/j.biortech.2009.05.058
- Adav, S. S., Lee, D.-J., Show, K.-Y., & Tay, J.-H. (2008). Aerobic granular sludge: Recent advances. *Biotechnol Adv*, 26(5), 411-423. doi:10.1016/j.biotechadv.2008.05.002
- Ahmad, A. A., Askora, A., Kawasaki, T., Fujie, M., & Yamada, T. (2014). The filamentous phage XacF1 causes loss of virulence in *Xanthomonas axonopodis* pv. *citri*, the causative agent of citrus canker disease. *Front Microbiol*, 5, 321. doi:10.3389/fmicb.2014.00321
- Alighardashi, A., Pandolfi, D., Potier, O., & Pons, M. N. (2009). Acute sensitivity of activated sludge bacteria to erythromycin. *J Hazard Mater*, 172(2-3), 685-692. doi:10.1016/j.jhazmat.2009.07.051
- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166-169. doi:10.1093/bioinformatics/btu638
- Anderson, O. R. (2010). Protozoan Ecology eLS.
- Angly, F. E., Willner, D., Prieto-Davo, A., Edwards, R. A., Schmieder, R., Vega-Thurber, R., Antonopoulos, D. A., Barott, K., Cottrell, M. T., Desnues, C., Dinsdale, E. A., Furlan, M., Haynes, M., Henn, M. R., Hu, Y., Kirchman, D. L., McDole, T., McPherson, J. D., Meyer, F., Miller, R. M., Mundt, E., Naviaux, R. K., Rodriguez-Mueller, B., Stevens, R., Wegley, L., Zhang, L., Zhu, B., & Rohwer, F. (2009). The GAAS metagenomic tool and its estimations of viral and microbial average genome size in four major biomes. *PLoS Comput Biol*, 5(12), e1000593. doi:10.1371/journal.pcbi.1000593

Aqeel, H., Basuvaraj, M., Hall, M., Neufeld, J. D., & Liss, S. N. (2016). Microbial dynamics and properties of aerobic granules developed in a laboratory-scale sequencing batch reactor with an intermediate filamentous bulking stage. *Appl Microbiol Biotechnol*, 100(1), 447-460. doi:10.1007/s00253-015-6981-7

Arnold, E., Böhm, B., & Wilderer, P. (2000). Application of activated sludge and biofilm sequencing batch reactor technology to treat reject water from sludge dewatering systems: A comparison. *Water Sci Technol*, 41(1), 115-122. doi:10.2166/wst.2000.0019

Baloch, M. I., Akunna, J. C., Kierans, M., & Collier, P. J. (2008). Structural analysis of anaerobic granules in a phase separated reactor by electron microscopy. *Bioresour Technol*, 99(5), 922-929. doi:10.1016/j.biortech.2007.03.016

Barcina, I., Ayo, B., Muela, A., Egea, L., & Iriberry, J. (1991). Predation rates of flagellate and ciliated protozoa on bacterioplankton in a river. *FEMS Microbiology Letters*, 85(2), 141-149.

Barondess, J. J., & Beckwith, J. (1990). A bacterial virulence determinant encoded by lysogenic coliphage lambda. *Nature*, 346(6287), 871-874. doi:10.1038/346871a0

Barr, J. J., Cook, A. E., & Bond, P. L. (2010a). Granule formation mechanisms within an aerobic wastewater system for phosphorus removal. *Appl Environ Microbiol*, 76(22), 7588-7597. doi:10.1128/aem.00864-10

Barr, J. J., Slater, F. R., Fukushima, T., & Bond, P. L. (2010b). Evidence for bacteriophage activity causing community and performance changes in a phosphorus-removal activated sludge. *FEMS Microbiol Ecol*, 74(3), 631-642. doi:10.1111/j.1574-6941.2010.00967.x

Bassin, J. P., Kleerebezem, R., Dezotti, M., & van Loosdrecht, M. C. (2012). Simultaneous nitrogen and phosphate removal in aerobic granular sludge reactors operated at different temperatures. *Water Res*, 46(12), 3805-3816. doi:10.1016/j.watres.2012.04.015

Bernardy, E. E., Turnsek, M. A., Wilson, S. K., Tarr, C. L., & Hammer, B. K. (2016). Diversity of Clinical and Environmental Isolates of *Vibrio cholerae* in Natural Transformation and Contact-Dependent Bacterial Killing Indicative of Type VI Secretion System Activity. *Appl Environ Microbiol*, 82(9), 2833-2842. doi:10.1128/aem.00351-16

Berngruber, T. W., Weissing, F. J., & Gandon, S. (2010). Inhibition of superinfection and the evolution of viral latency. *J Virol*, 84(19), 10200-10208. doi:10.1128/jvi.00865-10

Bertanza, G. (1997). Simultaneous nitrification-denitrification process in extended aeration plants: Pilot and real scale experiences. *Water Science and Technology*, 35(6), 53-61. doi:10.1016/S0273-1223(97)00095-4

Beun, J. J., Hendriks, A., van Loosdrecht, M. C. M., Morgenroth, E., Wilderer, P. A., & Heijnen, J. J. (1999). Aerobic granulation in a sequencing batch reactor. *Water Res*, 33(10), 2283-2290. doi:10.1016/S0043-1354(98)00463-1

Bille, E., Meyer, J., Jamet, A., Euphrasie, D., Barnier, J. P., Brissac, T., Larsen, A., Pelissier, P., & Nassif, X. (2017). A virulence-associated filamentous bacteriophage of *Neisseria meningitidis* increases host-cell colonisation. *PLoS Pathog*, 13(7), e1006495. doi:10.1371/journal.ppat.1006495

. Biological Processes. (2006). In M. Niaounakis & C. P. Halvadakis (Eds.), *Waste Management Series* (Vol. 5, pp. 171-218): Elsevier.

Biswas, K., & Turner, S. J. (2012). Microbial community composition and dynamics of moving bed biofilm reactor systems treating municipal sewage. *Appl Environ Microbiol*, 78(3), 855-864. doi:10.1128/aem.06570-11

Bjarnsholt, T., Alhede, M., Alhede, M., Eickhardt-Sorensen, S. R., Moser, C., Kuhl, M., Jensen, P. O., & Hoiby, N. (2013). The in vivo biofilm. *Trends Microbiol*, 21(9), 466-474. doi:10.1016/j.tim.2013.06.002

Bjarnsholt, T., Jensen, P. O., Fiandaca, M. J., Pedersen, J., Hansen, C. R., Andersen, C. B., Pressler, T., Givskov, M., & Hoiby, N. (2009). *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatr Pulmonol*, 44(6), 547-558. doi:10.1002/ppul.21011

Bohannan, B. J. M., & Lenski, R. E. (1997). Effect of Resource Enrichment on a Chemostat Community of Bacteria and Bacteriophage. *Ecology*, 78(8), 2303-2315. doi:10.1890/0012-9658(1997)078[2303:EOREOA]2.0.CO;2

Boles, B. R., Thoendel, M., & Singh, P. K. (2004). Self-generated diversity produces "insurance effects" in biofilm communities. *Proc Natl Acad Sci U S A*, 101(47), 16630.

Bonocora, R. P., & Wade, J. T. (2015). ChIP-seq for genome-scale analysis of bacterial DNA-binding proteins. *Methods Mol Biol*, 1276, 327-340. doi:10.1007/978-1-4939-2392-2_20

Boundless. (2017). The Evolution of Populations: Adaptive Evolution Boundless Biology. Boston, Massachusetts: Boundless. Retrieved from <https://courses.lumenlearning.com/boundless-biology/chapter/adaptive-evolution/>.

Bragg, J. G., & Chisholm, S. W. (2008). Modeling the fitness consequences of a cyanophage-encoded photosynthesis gene. *PLoS One*, 3(10), e3550. doi:10.1371/journal.pone.0003550

Braud, A., Geoffroy, V., Hoegy, F., Mislin, G. L., & Schalk, I. J. (2010). Presence of the siderophores pyoverdine and pyochelin in the extracellular medium reduces toxic metal accumulation in *Pseudomonas aeruginosa* and increases bacterial metal tolerance. *Environ Microbiol Rep*, 2(3), 419-425. doi:10.1111/j.1758-2229.2009.00126.x

Breitbart, M., Thompson, L. R., Suttle, C. A., & Sullivan, M. B. (2007). Exploring the vast diversity of marine viruses. *Oceanography*, 20(2), 135–139. doi:10.5670/oceanog.2007.58

Brussow, H., Canchaya, C., & Hardt, W. D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev*, 68(3), 560-602. doi:10.1128/mmr.68.3.560-602.2004

Brussow, H., & Desiere, F. (2001). Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. *Mol Microbiol*, 39(2), 213-222.

Brussow, H., & Hendrix, R. W. (2002). Phage genomics: small is beautiful. *Cell*, 108(1), 13-16.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat Methods*, 12(1), 59-60. doi:10.1038/nmeth.3176

Buckling, A., & Rainey, P. B. (2002). The role of parasites in sympatric and allopatric host diversification. *Nature*, 420(6915), 496-499. doi:10.1038/nature01164

Bushnell, B. (2015). *BBMap short-read aligner, and other bioinformatics tools*.

Buts, L., Lah, J., Dao-Thi, M. H., Wyns, L., & Loris, R. (2005). Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends Biochem Sci*, 30(12), 672-679. doi:10.1016/j.tibs.2005.10.004

Canchaya, C., Fournous, G., & Brussow, H. (2004). The impact of prophages on bacterial chromosomes. *Mol Microbiol*, 53(1), 9-18. doi:10.1111/j.1365-2958.2004.04113.x

Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M. L., & Brussow, H. (2003a). Phage as agents of lateral gene transfer. *Curr Opin Microbiol*, 6(4), 417-424.

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., & Brüssow, H. (2003b). Prophage Genomics. *Microbiol Mol Biol Rev*, 67(2), 238-276. doi:10.1128/MMBR.67.2.238-276.2003

Carpentier, B., & Cerf, O. (1993). Biofilms and their consequences, with particular reference to hygiene in the food industry. *J Appl Bacteriol*, 75(6), 499-511.

Carrolo, M., Frias, M. J., Pinto, F. R., Melo-Cristino, J., & Ramirez, M. (2010). Prophage spontaneous activation promotes DNA release enhancing biofilm formation in *Streptococcus pneumoniae*. *PLoS One*, 5(12), e15678. doi:10.1371/journal.pone.0015678

Carucci, A., Milia, S., Cappai, G., & Muntoni, A. (2010). A direct comparison amongst different technologies (aerobic granular sludge, SBR and MBR) for the

treatment of wastewater contaminated by 4-chlorophenol. *J Hazard Mater*, 177(1-3), 1119-1125. doi:10.1016/j.jhazmat.2010.01.037

Chao, Y., Mao, Y., Wang, Z., & Zhang, T. (2015). Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. *Sci Rep*, 5, 10044. doi:10.1038/srep10044

Charnock, C., & Nordlie, A.-L. (2016). Proteobacteria, extremophiles and unassigned species dominate in a tape-like showerhead biofilm. *Braz J Microbiol*, 47(2), 345-351. doi:10.1016/j.bjm.2016.01.018

Chen, L., & Wen, Y. M. (2011). The role of bacterial biofilm in persistent infections and control strategies. *Int J Oral Sci*, 3(2), 66-73. doi:10.4248/ijos11022

Cheng, C. M., Wang, H. J., Bau, H. J., & Kuo, T. T. (1999). The primary immunity determinant in modulating the lysogenic immunity of the filamentous bacteriophage cf. *J Mol Biol*, 287(5), 867-876. doi:10.1006/jmbi.1999.2651

Choi, J., Kotay, S. M., & Goel, R. (2010). Various physico-chemical stress factors cause prophage induction in *Nitrosospira multififormis* 25196--an ammonia oxidizing bacteria. *Water Res*, 44(15), 4550-4558. doi:10.1016/j.watres.2010.04.040

Chong, K. K. L., Tay, W. H., Janela, B., Yong, A. M. H., Liew, T. H., Madden, L., Keogh, D., Barkham, T. M. S., Ginhoux, F., Becker, D. L., & Kline, K. A. (2017). *Enterococcus faecalis* Modulates Immune Activation and Slows Healing During Wound Infection. *J Infect Dis*, 216(12), 1644-1654. doi:10.1093/infdis/jix541

Chouikha, I., Charrier, L., Filali, S., Derbise, A., & Carniel, E. (2010). Insights into the infective properties of YpfPhi, the *Yersinia pestis* filamentous phage. *Virology*, 407(1), 43-52. doi:10.1016/j.virol.2010.07.048

Christensen, L. D., van Gennip, M., Rybtke, M. T., Wu, H., Chiang, W. C., Alhede, M., Hoiby, N., Nielsen, T. E., Givskov, M., & Tolker-Nielsen, T. (2013). Clearance of *Pseudomonas aeruginosa* foreign-body biofilm infections through reduction of the cyclic Di-GMP level in the bacteria. *Infect Immun*, 81(8), 2705-2713. doi:10.1128/iai.00332-13

Cloete, T. E., & Bosch, M. (1994). Acinetobacter cell biomass, growth stage and phosphorus uptake from activated sludge mixed liquor. *Water Sci Technol*, 30(11), 219.

Conibear, T. C., Collins, S. L., & Webb, J. S. (2009). Role of mutation in *Pseudomonas aeruginosa* biofilm development. *PLoS One*, 4(7), e6289. doi:10.1371/journal.pone.0006289

Corona-Izquierdo, F. P., & Membrillo-Hernandez, J. (2002). A mutation in rpoS enhances biofilm formation in *Escherichia coli* during exponential phase of growth. *FEMS Microbiol Lett*, 211(1), 105-110.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., & Lappin-Scott, H. M. (1995). Microbial Biofilms. *Annu Rev Microbiol*, 49(1), 711-745. doi:10.1146/annurev.mi.49.100195.003431

Costerton, J. W., Stewart, P. S., & Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science*, 284(5418), 1318-1322.

Couturier, M., Bahassi el, M., & Van Melderren, L. (1998). Bacterial death by DNA gyrase poisoning. *Trends Microbiol*, 6(7), 269-275.

Darch, S. E., Kragh, K. N., Abbott, E. A., Bjarnsholt, T., Bull, J. J., & Whiteley, M. (2017). Phage Inhibit Pathogen Dissemination by Targeting Bacterial Migrants in a Chronic Infection Model. *MBio*, 8(2). doi:10.1128/mBio.00240-17

Davies, D. G., Parsek, M. R., Pearson, J. P., Iglewski, B. H., Costerton, J. W., & Greenberg, E. P. (1998). The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science*, 280(5361), 295-298.

Davis, B. M., Kimsey, H. H., Chang, W., & Waldor, M. K. (1999). *The Vibrio cholerae* O139 Calcutta bacteriophage CTXphi is infectious and encodes a novel repressor. *J Bacteriol*, 181(21), 6779-6787.

Davis, B. M., Moyer, K. E., Boyd, E. F., & Waldor, M. K. (2000). CTX prophages in classical biotype *Vibrio cholerae*: functional phage genes but dysfunctional phage genomes. *J Bacteriol*, 182(24), 6992-6998.

de Bruin, L. M., de Kreuk, M. K., van der Roest, H. F., Uijterlinde, C., & van Loosdrecht, M. C. (2004). Aerobic granular sludge technology: an alternative to activated sludge? *Water Sci Technol*, 49(11-12), 1-7.

De Kievit, T. R., Gillis, R., Marx, S., Brown, C., & Iglewski, B. H. (2001). Quorum-sensing genes in *Pseudomonas aeruginosa* biofilms: their role and expression patterns. *Appl Environ Microbiol*, 67(4), 1865-1873. doi:10.1128/aem.67.4.1865-1873.2001

de Kreuk, M. K., Heijnen, J. J., & van Loosdrecht, M. C. (2005). Simultaneous COD, nitrogen, and phosphate removal by aerobic granular sludge. *Biotechnol Bioeng*, 90(6), 761-769. doi:10.1002/bit.20470

de Kreuk, M. K., Picioreanu, C., Hosseini, M., Xavier, J. B., & van Loosdrecht, M. C. (2007). Kinetic model of a granular sludge SBR: influences on nutrient removal. *Biotechnol Bioeng*, 97(4), 801-815. doi:10.1002/bit.21196

Deng, L., Ignacio-Espinoza, J. C., Gregory, A. C., Poulos, B. T., Weitz, J. S., Hugenholtz, P., & Sullivan, M. B. (2014). Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature*, 513, 242. doi:10.1038/nature13459

Desiere, F., Pridmore, R. D., & Brussow, H. (2000). Comparative genomics of the late gene cluster from *Lactobacillus* phages. *Virology*, 275(2), 294-305. doi:10.1006/viro.2000.0498

- Diaz, P. I., Chalmers, N. I., Rickard, A. H., Kong, C., Milburn, C. L., Palmer, R. J., Jr., & Kolenbrander, P. E. (2006). Molecular characterization of subject-specific oral microflora during initial colonization of enamel. *Appl Environ Microbiol*, 72(4), 2837-2848. doi:10.1128/aem.72.4.2837-2848.2006
- Donnelly-Wu, M. K., Jacobs, W. R., Jr., & Hatfull, G. F. (1993). Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. *Mol Microbiol*, 7(3), 407-417.
- Dowd, S. E., Sun, Y., Secor, P. R., Rhoads, D. D., Wolcott, B. M., James, G. A., & Wolcott, R. D. (2008). Survey of bacterial diversity in chronic wounds using pyrosequencing, DGGE, and full ribosome shotgun sequencing. *BMC Microbiol*, 8, 43. doi:10.1186/1471-2180-8-43
- Drenkard, E., & Ausubel, F. M. (2002). *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature*, 416(6882), 740-743. doi:10.1038/416740a
- Eaton, A. D., Franson, M. A. H., Association, A. P. H., Association, A. W. W., & Federation, W. E. (2005). *Standard Methods for the Examination of Water and Wastewater*. Washington DC, USA: American Public Health Association.
- Edwards, S. J., & Kjellerup, B. V. (2013). Applications of biofilms in bioremediation and biotransformation of persistent organic pollutants, pharmaceuticals/personal care products, and heavy metals. *Appl Microbiol Biotechnol*, 97(23), 9909-9921. doi:10.1007/s00253-013-5216-z
- Elliott, D. R., Wilson, M., Buckley, C. M., & Spratt, D. A. (2006). Aggregative behavior of bacteria isolated from canine dental plaque. *Appl Environ Microbiol*, 72(8), 5211-5217. doi:10.1128/aem.01060-05
- Elrod, R. P., & Braun, A. C. (1942). *Pseudomonas aeruginosa*: Its Role as a Plant Pathogen. *J Bacteriol*, 44(6), 633-645.
- Ercolini, D. (2013). High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. *Appl Environ Microbiol*, 79(10), 3148-3155. doi:10.1128/aem.00256-13
- Esteban, G., Téllez, C., & Bautista, L. M. (1991). Dynamics of ciliated protozoa communities in activated-sludge process. *Water Res*, 25(8), 967-972. doi:10.1016/0043-1354(91)90145-G
- Farrant, J. L., Sansone, A., Canvin, J. R., Pallen, M. J., Langford, P. R., Wallis, T. S., Dougan, G., & Kroll, J. S. (1997). Bacterial copper- and zinc-cofactored superoxide dismutase contributes to the pathogenesis of systemic salmonellosis. *Mol Microbiol*, 25(4), 785-796.
- Faruque, S. M., & Mekalanos, J. J. (2012). Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence*, 3(7), 556-565. doi:10.4161/viru.22351

Feng, S., Tan, C. H., Constancias, F., Kohli, G. S., Cohen, Y., & Rice, S. A. (2017). Predation by *Bdellovibrio bacteriovorus* significantly reduces viability and alters the microbial community composition of activated sludge flocs and granules. *FEMS Microbiol Ecol*, 93(4). doi:10.1093/femsec/fix020

Finch, J. E., Prince, J., & Hawksworth, M. (1978). A bacteriological survey of the domestic environment. *J Appl Bacteriol*, 45(3), 357-364.

Flemming, H. C., Neu, T. R., & Wozniak, D. J. (2007). The EPS matrix: the "house of biofilm cells". *J Bacteriol*, 189(22), 7945-7947. doi:10.1128/jb.00858-07

Flemming, H. C., & Wingender, J. (2010). The biofilm matrix. *Nat Rev Microbiol*, 8(9), 623-633. doi:10.1038/nrmicro2415

Fogg, P. C., Allison, H. E., Saunders, J. R., & McCarthy, A. J. (2010). Bacteriophage lambda: a paradigm revisited. *J Virol*, 84(13), 6876-6879. doi:10.1128/jvi.02177-09

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150-3152. doi:10.1093/bioinformatics/bts565

Garcia-Martin, H., Ivanova, N., Kunin, V., Warnecke, F., Barry, K. W., McHardy, A. C., Yeates, C., He, S., Salamov, A. A., Szeto, E., Dalin, E., Putnam, N. H., Shapiro, H. J., Pangilinan, J. L., Rigoutsos, I., Kyrpides, N. C., Blackall, L. L., McMahon, K. D., & Hugenholtz, P. (2006). Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nat Biotechnol*, 24(10), 1263-1269. doi:10.1038/nbt1247

Gebara, F. (1999). *Activated sludge biofilm wastewater treatment system* (Vol. 33).

Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., & Suzuki, N. (2015). 50-plus years of fungal viruses. *Virology*, 479-480, 356-368. doi:10.1016/j.virol.2015.02.034

Ghigo, J.-M. (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature*, 412, 442. doi:10.1038/35086581

Giacometti, A., Cirioni, O., Schimizzi, A. M., Del Prete, M. S., Barchiesi, F., D'Errico, M. M., Petrelli, E., & Scalise, G. (2000). Epidemiology and microbiology of surgical wound infections. *J Clin Microbiol*, 38(2), 918-922.

Gibbons, R. J., & Nygaard, M. (1970). Interbacterial aggregation of plaque bacteria. *Arch Oral Biol*, 15(12), 1397-1400.

Glessner, A., Smith, R. S., Iglewski, B. H., & Robinson, J. B. (1999). Roles of *Pseudomonas aeruginosa las* and *rhl* Quorum-Sensing Systems in Control of Twitching Motility. *J Bacteriol*, 181(5), 1623-1629.

- Gödeke, J., Paul, K., Lassak, J., & Thormann, K. M. (2010). Phage-induced lysis enhances biofilm formation in *Shewanella oneidensis* MR-1. *ISME J*, 5, 613. doi:10.1038/ismej.2010.153
- Goh, H. M. S., Yong, M. H. A., Chong, K. K. L., & Kline, K. A. (2017). Model systems for the study of Enterococcal colonization and infection. *Virulence*, 8(8), 1525-1562. doi:10.1080/21505594.2017.1279766
- Gomez-Gil, B., Roque, A., & Turnbull, J. F. (2000). The use and selection of probiotic bacteria for use in the culture of larval aquatic organisms. *Aquaculture*, 191(1), 259-270. doi:10.1016/S0044-8486(00)00431-2
- Green, M. R., & Sambrook, J. (2016). Precipitation of DNA with Ethanol. *Cold Spring Harb Protoc*, 2016(12). doi:10.1101/pdb.prot093377
- Grobe, S., Wingender, J., & Truper, H. G. (1995). Characterization of mucoid *Pseudomonas aeruginosa* strains isolated from technical water systems. *J Appl Bacteriol*, 79(1), 94-102.
- Guo, F., & Zhang, T. (2012). Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing. *Water Res*, 46(8), 2772-2782. doi:10.1016/j.watres.2012.02.039
- Hall-Stoodley, L., Costerton, J. W., & Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. *Nat Rev Micro*, 2(2), 95-108.
- Hall-Stoodley, L., & Stoodley, P. (2009). Evolving concepts in biofilm infections. *Cell Microbiol*, 11(7), 1034-1043. doi:10.1111/j.1462-5822.2009.01323.x
- Handelsman, J. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol Mol Biol Rev*, 68(4), 669-685. doi:10.1128/MMBR.68.4.669-685.2004
- Harvey, H., Bondy-Denomy, J., Marquis, H., Sztanko, K. M., Davidson, A. R., & Burrows, L. L. (2018). *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. *Nature Microbiology*, 3(1), 47-52. doi:10.1038/s41564-017-0061-y
- Hashem, Y. A., Amin, H. M., Essam, T. M., Yassin, A. S., & Aziz, R. K. (2017). Biofilm formation in enterococci: genotype-phenotype correlations and inhibition by vancomycin. *Sci Rep*, 7(1), 5733. doi:10.1038/s41598-017-05901-0
- Hasty, J., McMillen, D., Isaacs, F., & Collins, J. J. (2001). Computational studies of gene regulatory networks: in numero molecular biology. *Nat Rev Genet*, 2(4), 268-279. doi:10.1038/35066056
- Hauser, A. R. (2009). The type III secretion system of *Pseudomonas aeruginosa*: infection by injection. *Nat Rev Microbiol*, 7(9), 654-665. doi:10.1038/nrmicro2199

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C. G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M., & Shinagawa, H. (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res*, 8(1), 11-22.

He, S., Gall, D. L., & McMahon, K. D. (2007). "*Candidatus Accumolibacter*" population structure in enhanced biological phosphorus removal sludges as revealed by polyphosphate kinase genes. *Appl Environ Microbiol*, 73(18), 5865-5874. doi:10.1128/aem.01207-07

He, S., & McMahon, K. D. (2011). Microbiology of '*Candidatus Accumolibacter*' in activated sludge. *Microb Biotechnol*, 4(5), 603-619. doi:10.1111/j.1751-7915.2011.00248.x

Held, N. L., Herrera, A., Cadillo-Quiroz, H., & Whitaker, R. J. (2010). CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLoS One*, 5(9). doi:10.1371/journal.pone.0012988

Hellweger, F. L. (2009). Carrying photosynthesis genes increases ecological fitness of cyanophage in silico. *Environ Microbiol*, 11(6), 1386-1394. doi:10.1111/j.1462-2920.2009.01866.x

Helmer, C., & Kunst, S. (1998). Simultaneous nitrification/denitrification in an aerobic biofilm system. *Water Sci Technol*, 37(4), 183-187. doi:10.1016/S0273-1223(98)00102-4

Hense, B. A., Kuttler, C., Muller, J., Rothballer, M., Hartmann, A., & Kreft, J. U. (2007). Does efficiency sensing unify diffusion and quorum sensing? *Nat Rev Microbiol*, 5(3), 230-239. doi:10.1038/nrmicro1600

Hernandez-Doria, J. D., & Sperandio, V. (2018). Bacteriophage Transcription Factor Cro Regulates Virulence Gene Expression in Enterohemorrhagic *Escherichia coli*. *Cell Host Microbe*, 23(5), 607-617.e606. doi:10.1016/j.chom.2018.04.007

Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., & Schweizer, H. P. (1998). A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene*, 212(1), 77-86.

Hoffmaster, A. R., Novak, R. T., Marston, C. K., Gee, J. E., Helsel, L., Pruckler, J. M., & Wilkins, P. P. (2008). Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing. *BMC Microbiol*, 8(1), 191. doi:10.1186/1471-2180-8-191

Hoiby, N., Krogh Johansen, H., Moser, C., Song, Z., Ciofu, O., & Kharazmi, A. (2001). *Pseudomonas aeruginosa* and the in vitro and in vivo biofilm mode of growth. *Microbes Infect*, 3(1), 23-35.

- Huang, A., Friesen, J., & Brunton, J. L. (1987). Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in *Escherichia coli*. *J Bacteriol*, 169(9), 4308-4312.
- Huang, W., & Wilks, A. (2017). A rapid seamless method for gene knockout in *Pseudomonas aeruginosa*. *BMC Microbiol*, 17, 199. doi:10.1186/s12866-017-1112-5
- Hufnagel, M., Koch, S., Creti, R., Baldassarri, L., & Huebner, J. (2004). A putative sugar-binding transcriptional regulator in a novel gene locus in *Enterococcus faecalis* contributes to production of biofilm and prolonged bacteremia in mice. *J Infect Dis*, 189(3), 420-430. doi:10.1086/381150
- Hugenholtz, P., & Džunková, M. (2019). *Single cell viral tagging*. Paper presented at the Australian Microbial Ecology Conference, Perth, Australia.
- Hui, G. K. J. (2013). *The genes and factors that drive the conversion of the Pseudomonas aeruginosa Pf4 prophage into the superinfective form*. (Doctor of Philosophy), The University of New South Wales.
- Hung, J.-H., Yang, T.-H., Hu, Z., Weng, Z., & DeLisi, C. (2012). Gene set enrichment analysis: performance evaluation and usage guidelines. *Brief Bioinform*, 13(3), 281-291. doi:10.1093/bib/bbr049
- Hunsucker, K. Z., Vora, G. J., Hunsucker, J. T., Gardner, H., Leary, D. H., Kim, S., Lin, B., & Swain, G. (2018). Biofilm community structure and the associated drag penalties of a groomed fouling release ship hull coating. *Biofouling*, 34(2), 162-172. doi:10.1080/08927014.2017.1417395
- Huson, D. H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H. J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol*, 12(6), e1004957. doi:10.1371/journal.pcbi.1004957
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119-119. doi:10.1186/1471-2105-11-119
- Ibarbalz, F. M., Figuerola, E. L., & Erijman, L. (2013). Industrial activated sludge exhibit unique bacterial community composition at high taxonomic ranks. *Water Res*, 47(11), 3854-3864. doi:10.1016/j.watres.2013.04.010
- Isanta, E., Suárez-Ojeda, M. E., Val del Río, Á., Morales, N., Pérez, J., & Carrera, J. (2012). Long term operation of a granular sequencing batch reactor at pilot scale treating a low-strength wastewater. *Chem Eng J*, 198-199, 163-170. doi:10.1016/j.cej.2012.05.066
- Jacobson, A. (1972). Role of F pili in the penetration of bacteriophage fl. *J Virol*, 10(4), 835-843.

Jakubovics, N. S., Shields, R. C., Rajarajan, N., & Burgess, J. G. (2013). Life after death: the critical role of extracellular DNA in microbial biofilms. *Lett Appl Microbiol*, 57(6), 467-475. doi:10.1111/lam.12134

Jefferson, K. K. (2004). What drives bacteria to produce a biofilm? *FEMS Microbiol Lett*, 236(2), 163-173. doi:10.1111/j.1574-6968.2004.tb09643.x

Jiang, F., Waterfield, N. R., Yang, J., Yang, G., & Jin, Q. (2014). A *Pseudomonas aeruginosa* type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic cells. *Cell Host Microbe*, 15(5), 600-610. doi:10.1016/j.chom.2014.04.010

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Mol Microbiol*, 44(4), 971-979.

Johnson, A., Meyer, B. J., & Ptashne, M. (1978). Mechanism of action of the cro protein of bacteriophage lambda. *Proc Natl Acad Sci U S A*, 75(4), 1783-1787.

Joo, J., Gunny, M., Cases, M., Hudson, P., Albert, R., & Harvill, E. (2006). Bacteriophage-mediated competition in *Bordetella* bacteria. *Proc Biol Sci*, 273(1595), 1843-1848. doi:10.1098/rspb.2006.3512

Ju, F., Guo, F., Ye, L., Xia, Y., & Zhang, T. (2014). Metagenomic analysis on seasonal microbial variations of activated sludge from a full-scale wastewater treatment plant over 4 years. *Environ Microbiol Rep*, 6(1), 80-89. doi:10.1111/1758-2229.12110

Jürgens, K. (2007). Predation on Bacteria and Bacterial Resistance Mechanisms: Comparative Aspects Among Different Predator Groups in Aquatic Systems. In E. Jurkevitch (Ed.), *Predatory Prokaryotes: Biology, Ecology and Evolution* (pp. 57-92). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kaplan, J. B. (2010). Biofilm dispersal: mechanisms, clinical implications, and potential therapeutic uses. *J Dent Res*, 89(3), 205-218. doi:10.1177/0022034509359403

Kawakoshi, A., Nakazawa, H., Fukada, J., Sasagawa, M., Katano, Y., Nakamura, S., Hosoyama, A., Sasaki, H., Ichikawa, N., Hanada, S., Kamagata, Y., Nakamura, K., Yamazaki, S., & Fujita, N. (2012). Deciphering the genome of polyphosphate accumulating actinobacterium *Microlunatus phosphovorus*. *DNA Res*, 19(5), 383-394. doi:10.1093/dnares/dss020

Kern-Jespersen, J. P., Henze, M., & Strube, R. (1994). Biological phosphorus release and uptake under alternating anaerobic and anoxic conditions in a fixed-film reactor. *Water Res*, 28(5), 1253-1255. doi:10.1016/0043-1354(94)90215-1

Khairnar, K., Chandekar, R., Nair, A., Pal, P., & Paunekar, W. N. (2016). Novel application of bacteriophage for controlling foaming in wastewater treatment plant- an eco-friendly approach. *Bioengineered*, 7(1), 46-49. doi:10.1080/21655979.2015.1134066

Khairnar, K., Pal, P., Chandekar, R. H., & Paunikar, W. N. (2014). Isolation and characterization of bacteriophages infecting nocardioforms in wastewater treatment plant. *Biotechnol Res Int*, 2014, 151952. doi:10.1155/2014/151952

Khan, A. A., Gaur, R. Z., Tyagi, V. K., Khursheed, A., Lew, B., Mehrotra, I., & Kazmi, A. A. (2011). Sustainable options of post treatment of UASB effluent treating sewage: A review. *Resources, Conservation and Recycling*, 55(12), 1232-1251.

Kim, E. J., Yu, H. J., Lee, J. H., Kim, J.-O., Han, S. H., Yun, C.-H., Chun, J., Nair, G. B., & Kim, D. W. (2017). Replication of *Vibrio cholerae* classical CTX phage. *Proc Natl Acad Sci U S A*, 114(9), 2343-2348. doi:10.1073/pnas.1701335114

Kimsey, H. H., & Waldor, M. K. (2004). The CTXphi repressor RstR binds DNA cooperatively to form tetrameric repressor-operator complexes. *J Biol Chem*, 279(4), 2640-2647. doi:10.1074/jbc.M311109200

Kinane, D. F., Stathopoulou, P. G., & Papapanou, P. N. (2017). Periodontal diseases. *Nat Rev Dis Primers*, 3, 17038. doi:10.1038/nrdp.2017.38

Koentjoro, M. P., Adachi, N., Senda, M., Ogawa, N., & Senda, T. (2018). Crystal structure of the DNA-binding domain of the LysR-type transcriptional regulator CbnR in complex with a DNA fragment of the recognition-binding site in the promoter region. *Febs j*, 285(5), 977-989. doi:10.1111/febs.14380

Koh, K. S., Matz, C., Tan, C. H., Le, H. L., Rice, S. A., Marshall, D. J., Steinberg, P. D., & Kjelleberg, S. (2012). Minimal increase in genetic diversity enhances predation resistance. *Mol Ecol*, 21(7), 1741-1753. doi:10.1111/j.1365-294X.2011.05415.x

Kopylova, E., Noe, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, 28(24), 3211-3217. doi:10.1093/bioinformatics/bts611

Korea, C. G., Badouraly, R., Prevost, M. C., Ghigo, J. M., & Beloin, C. (2010). *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ Microbiol*, 12(7), 1957-1977. doi:10.1111/j.1462-2920.2010.02202.x

Korea, C. G., Ghigo, J. M., & Beloin, C. (2011). The sweet connection: Solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli*: Multiple *E. coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tissue tropism. *Bioessays*, 33(4), 300-311. doi:10.1002/bies.201000121

Koskella, B., Lin, D. M., Buckling, A., & Thompson, J. N. (2012). The costs of evolving resistance in heterogeneous parasite environments. *Proc Biol Sci*, 279(1735), 1896-1903. doi:10.1098/rspb.2011.2259

Koutsoudis, M. D., Tsaltas, D., Minogue, T. D., & von Bodman, S. B. (2006). Quorum-sensing regulation governs bacterial adhesion, biofilm development,

and host colonization in *Pantoea stewartii* subspecies *stewartii*. *Proc Natl Acad Sci U S A*, 103(15), 5983-5988. doi:10.1073/pnas.0509860103

Kristiansen, R., Nguyen, H. T., Saunders, A. M., Nielsen, J. L., Wimmer, R., Le, V. Q., McIlroy, S. J., Petrovski, S., Seviour, R. J., Calteau, A., Nielsen, K. L., & Nielsen, P. H. (2013). A metabolic model for members of the genus *Tetrasphaera* involved in enhanced biological phosphorus removal. *ISME J*, 7(3), 543-554. doi:10.1038/ismej.2012.136

Ladero, V., Garcia, P., Bascaran, V., Herrero, M., Alvarez, M. A., & Suarez, J. E. (1998). Identification of the repressor-encoding gene of the *Lactobacillus* bacteriophage A2. *J Bacteriol*, 180(13), 3474-3476.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9, 357. doi:10.1038/nmeth.1923

Lavigne, R., Darius, P., Summer, E. J., Seto, D., Mahadevan, P., Nilsson, A. S., Ackermann, H. W., & Kropinski, A. M. (2009). Classification of Myoviridae bacteriophages using protein sequence similarity. *BMC Microbiol*, 9, 224. doi:10.1186/1471-2180-9-224

Ledder, R. G., Timperley, A. S., Friswell, M. K., Macfarlane, S., & McBain, A. J. (2008). Coaggregation between and among human intestinal and oral bacteria. *FEMS Microbiol Ecol*, 66(3), 630-636. doi:10.1111/j.1574-6941.2008.00525.x

Lee, K. W. K., Yam, J. K. H., Mukherjee, M., Periasamy, S., Steinberg, P. D., Kjelleberg, S., & Rice, S. A. (2015). Interspecific diversity reduces and functionally substitutes for intraspecific variation in biofilm communities. *ISME J*, 10, 846. doi:10.1038/ismej.2015.159

Lemaire, R., Meyer, R., Taske, A., Crocetti, G. R., Keller, J., & Yuan, Z. (2006). Identifying causes for N₂O accumulation in a lab-scale sequencing batch reactor performing simultaneous nitrification, denitrification and phosphorus removal. *J Biotechnol*, 122(1), 62-72. doi:10.1016/j.jbiotec.2005.08.024

Lemaire, R., Webb, R. I., & Yuan, Z. (2008a). Micro-scale observations of the structure of aerobic microbial granules used for the treatment of nutrient-rich industrial wastewater. *ISME J*, 2(5), 528-541. doi:10.1038/ismej.2008.12

Lemaire, R., Yuan, Z., Blackall, L. L., & Crocetti, G. R. (2008b). Microbial distribution of *Accumulibacter* spp. and *Competibacter* spp. in aerobic granules from a lab-scale biological nutrient removal system. *Environ Microbiol*, 10(2), 354-363. doi:10.1111/j.1462-2920.2007.01456.x

Leroy, F., & De Vuyst, L. (2004). Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci Technol*, 15(2), 67-78. doi:10.1016/j.tifs.2003.09.004

Lesic, B., & Rahme, L. G. (2008). Use of the lambda Red recombinase system to rapidly generate mutants in *Pseudomonas aeruginosa*. *BMC Mol Biol*, 9, 20. doi:10.1186/1471-2199-9-20

- Levin, B. R., & Bull, J. J. (2004). Population and evolutionary dynamics of phage therapy. *Nat Rev Microbiol*, 2, 166. doi:10.1038/nrmicro822
- Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., Yamashita, H., & Lam, T. W. (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3-11. doi:10.1016/j.ymeth.2016.02.020
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352
- Li, Y., & Liu, Y. (2005). Diffusion of substrate and oxygen in aerobic granule. *Biochem Eng J*, 27(1), 45-52. doi:10.1016/j.bej.2005.06.012
- Li, Y., Liu, Y., Shen, L., & Chen, F. (2008). DO diffusion profile in aerobic granule and its microbiological implications. *Enzyme Microb Technol*, 43(4), 349-354. doi:10.1016/j.enzmictec.2008.04.005
- Lim, W. S., Phang, K. K. S., Tan, A. H.-M., Li, S. F.-Y., & Ow, D. S.-W. (2016). Small Colony Variants and Single Nucleotide Variations in Pf1 Region of PB1 Phage-Resistant *Pseudomonas aeruginosa*. *Front Microbiol*, 7(282). doi:10.3389/fmicb.2016.00282
- Lima, S. L. (1998). Nonlethal Effects in the Ecology of Predator-Prey Interactions What are the ecological effects of anti-predator decision-making? *BioScience*, 48(1), 25-34. doi:10.2307/1313225
- Lindell, D., Jaffe, J. D., Coleman, M. L., Futschik, M. E., Axmann, I. M., Rector, T., Kettler, G., Sullivan, M. B., Steen, R., Hess, W. R., Church, G. M., & Chisholm, S. W. (2007). Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature*, 449(7158), 83-86. doi:10.1038/nature06130
- Liu, M., Gill, J. J., Young, R., & Summer, E. J. (2015). Bacteriophages of wastewater foaming-associated filamentous *Gordonia* reduce host levels in raw activated sludge. *Sci Rep*, 5, 13754. doi:10.1038/srep13754
- Liu, Y. (2008). *Wastewater Purification: Aerobic Granulation in Sequencing Batch Reactors*: CRC Press: Taylor & Francis.
- Liu, Y., & Liu, Q.-S. (2006). Causes and control of filamentous growth in aerobic granular sludge sequencing batch reactors. *Biotechnol Adv*, 24(1), 115-127. doi:10.1016/j.biotechadv.2005.08.001
- Liu, Y., & Tay, J.-H. (2002). The essential role of hydrodynamic shear force in the formation of biofilm and granular sludge. *Water Res*, 36(7), 1653-1665.
- Liu, Y., & Tay, J.-H. (2004). State of the art of biogranulation technology for wastewater treatment. *Biotechnol Adv*, 22(7), 533-563. doi:10.1016/j.biotechadv.2004.05.001

Liu, Y., & Wang, Z.-W. (2007). Essential Roles of Extracellular Polymeric Substances in Aerobic Granulation. In Y. Liu (Ed.), *Wastewater Purification* (pp. 181-194): CRC Press.

Liu, Y., Wang, Z. W., Qin, L., Liu, Y. Q., & Tay, J. H. (2005). Selection pressure-driven aerobic granulation in a sequencing batch reactor. *Appl Microbiol Biotechnol*, 67(1), 26-32. doi:10.1007/s00253-004-1820-2

Ljungquist, E., Kockum, K., & Bertani, L. E. (1984). DNA sequences of the repressor gene and operator region of bacteriophage P2. *Proc Natl Acad Sci U S A*, 81(13), 3988-3992.

Lopez, C. A., Winter, S. E., Rivera-Chavez, F., Xavier, M. N., Poon, V., Nuccio, S. P., Tsois, R. M., & Baumler, A. J. (2012). Phage-mediated acquisition of a type III secreted effector protein boosts growth of salmonella by nitrate respiration. *MBio*, 3(3). doi:10.1128/mBio.00143-12

Lopez-Bueno, A., Tamames, J., Velazquez, D., Moya, A., Quesada, A., & Alcamí, A. (2009). High diversity of the viral community from an Antarctic lake. *Science*, 326(5954), 858-861. doi:10.1126/science.1179287

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15(12), 550. doi:10.1186/s13059-014-0550-8

Lu, H., Oehmen, A., Virdis, B., Keller, J., & Yuan, Z. (2006). Obtaining highly enriched cultures of *Candidatus Accumulibacter phosphatus* through alternating carbon sources. *Water Res*, 40(20), 3838-3848. doi:10.1016/j.watres.2006.09.004

Lu, T. K., & Collins, J. J. (2007). Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci U S A*, 104(27), 11197-11202. doi:10.1073/pnas.0704624104

Lynn, D. H. (2012). *Ciliophora eLS*.

Machanick, P., & Bailey, T. L. (2011). MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, 27(12), 1696-1697. doi:10.1093/bioinformatics/btr189

Madoni, P. (2003). Protozoa in Activated Sludge. In G. Bittton (Ed.), *Encyclopedia of Environmental Microbiology*.

Madoni, P. (2011). Protozoa in wastewater treatment processes: A minireview. *Ital J Zool*, 78(1), 3-11. doi:10.1080/11250000903373797

Mai-Prochnow, A., Hui, J. G. K., Kjelleberg, S., Rakonjac, J., McDougald, D., & Rice, S. A. (2015). 'Big things in small packages: the genetics of filamentous phage and effects on fitness of their host'. *FEMS Microbiol Rev*, 39(4), 465-487. doi:10.1093/femsre/fuu007

Martiny, A. C., Albrechtsen, H. J., Arvin, E., & Molin, S. (2005). Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking

water distribution system: detection of a large nitrite-oxidizing population associated with *Nitrospira* spp. *Appl Environ Microbiol*, 71(12), 8611-8617. doi:10.1128/aem.71.12.8611-8617.2005

Martiny, A. C., Jorgensen, T. M., Albrechtsen, H. J., Arvin, E., & Molin, S. (2003). Long-term succession of structure and diversity of a biofilm formed in a model drinking water distribution system. *Appl Environ Microbiol*, 69(11), 6899-6907.

Maszenan, A. M., Liu, Y., & Ng, W. J. (2011). Bioremediation of wastewaters with recalcitrant organic compounds and metals by aerobic granules. *Biotechnol Adv*, 29(1), 111-123. doi:10.1016/j.biotechadv.2010.09.004

Matilla, M. A. (2017). A technology for the investigation of biofilm transmission under shearing pressures. *Microb Biotechnol*, 10(6), 1451-1453. doi:10.1111/1751-7915.12848

Matz, C., & Kjelleberg, S. (2005). Off the hook--how bacteria survive protozoan grazing. *Trends Microbiol*, 13(7), 302-307. doi:10.1016/j.tim.2005.05.009

May, T., Tsuruta, K., & Okabe, S. (2011). Exposure of conjugative plasmid carrying *Escherichia coli* biofilms to male-specific bacteriophages. *ISME J*, 5(4), 771-775. doi:10.1038/ismej.2010.158

Maynard Smith, J. (1962). Disruptive Selection, Polymorphism and Sympatric Speciation. *Nature*, 195, 60. doi:10.1038/195060a0

McArdle, B. H., & Anderson, M. J. (2001). Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology*, 82(1), 290-297. doi:10.1890/0012-9658(2001)082[0290:FMMTCD]2.0.CO;2

McDaniel, L., & Paul, J. H. (2005). Effect of Nutrient Addition and Environmental Factors on Prophage Induction in Natural Populations of Marine *Synechococcus* Species. *Appl Environ Microbiol*, 71(2), 842-850. doi:10.1128/AEM.71.2.842-850.2005

McDougald, D., Rice, S. A., Barraud, N., Steinberg, P. D., & Kjelleberg, S. (2011). Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nat Rev Microbiol*, 10(1), 39-50. doi:10.1038/nrmicro2695

McElroy, K. E., Hui, J. G. K., Woo, J. K. K., Luk, A. W. S., Webb, J. S., Kjelleberg, S., Rice, S. A., & Thomas, T. (2014). Strain-specific parallel evolution drives short-term diversification during *Pseudomonas aeruginosa* biofilm formation. *Proc Natl Acad Sci U S A*, 111(14), E1419.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8(4), e61217. doi:10.1371/journal.pone.0061217

McSwain, B. S., Irvine, R. L., & Wilderer, P. A. (2004). The influence of settling time on the formation of aerobic granules. *Water Sci Technol*, 50(10), 195-202.

Meyer, J. M., Neely, A., Stintzi, A., Georges, C., & Holder, I. A. (1996). Pyoverdine is essential for virulence of *Pseudomonas aeruginosa*. *Infect Immun*, 64(2), 518-523.

Meyer, R. L., Zeng, R. J., Giugliano, V., & Blackall, L. L. (2005). Challenges for simultaneous nitrification, denitrification, and phosphorus removal in microbial aggregates: mass transfer limitation and nitrous oxide production. *FEMS Microbiol Ecol*, 52(3), 329-338. doi:10.1016/j.femsec.2004.11.011

Middelboe, M., Holmfeldt, K., Riemann, L., Nybroe, O., & Haaber, J. (2009). Bacteriophages drive strain diversification in a marine *Flavobacterium*: implications for phage resistance and physiological properties. *Environ Microbiol*, 11(8), 1971-1982. doi:10.1111/j.1462-2920.2009.01920.x

Min, K. R., & Rickard, A. H. (2009). Coaggregation by the freshwater bacterium *Sphingomonas natatoria* alters dual-species biofilm formation. *Appl Environ Microbiol*, 75(12), 3987-3997. doi:10.1128/aem.02843-08

Mohle, R. B., Langemann, T., Haesner, M., Augustin, W., Scholl, S., Neu, T. R., Hempel, D. C., & Horn, H. (2007). Structure and shear strength of microbial biofilms as determined with confocal laser scanning microscopy and fluid dynamic gauging using a novel rotating disc biofilm reactor. *Biotechnol Bioeng*, 98(4), 747-755. doi:10.1002/bit.21448

Molina, L., Rezzonico, F., Defago, G., & Duffy, B. (2005). Autoinduction in *Erwinia amylovora*: evidence of an acyl-homoserine lactone signal in the fire blight pathogen. *J Bacteriol*, 187(9), 3206-3213. doi:10.1128/jb.187.9.3206-3213.2005

Monroe, D. (2007). Looking for chinks in the armor of bacterial biofilms. *PLoS Biol*, 5(11), e307. doi:10.1371/journal.pbio.0050307

Mooij, M. J., Drenkard, E., Llamas, M. A., Vandenbroucke-Grauls, C. M., Savelkoul, P. H., Ausubel, F. M., & Bitter, W. (2007). Characterization of the integrated filamentous phage Pf5 and its involvement in small-colony formation. *Microbiology*, 153(Pt 6), 1790-1798. doi:10.1099/mic.0.2006/003533-0

Morgan-Sagastume, F., & Grant Allen, D. (2005). Activated sludge deflocculation under temperature upshifts from 30 to 45 degrees C. *Water Res*, 39(6), 1061-1074. doi:10.1016/j.watres.2004.12.027

Munch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., & Jahn, D. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*, 21(22), 4187-4189. doi:10.1093/bioinformatics/bti635

Murnleitner, E., Kuba, T., van Loosdrecht, M. C., & Heijnen, J. J. (1997). An integrated metabolic model for the aerobic and denitrifying biological phosphorus removal. *Biotechnol Bioeng*, 54(5), 434-450. doi:10.1002/(sici)1097-0290(19970605)54:5<434::Aid-bit4>3.0.Co;2-f

- Nickels, B. E., Dove, S. L., Murakami, K. S., Darst, S. A., & Hochschild, A. (2002). Protein–Protein and Protein–DNA Interactions of $\sigma 70$ Region 4 Involved in Transcription Activation by λ cl. *J Mol Biol*, 324(1), 17-34. doi:[https://doi.org/10.1016/S0022-2836\(02\)01043-4](https://doi.org/10.1016/S0022-2836(02)01043-4)
- Nielsen, J. L., Nguyen, H., Meyer, R. L., & Nielsen, P. H. (2012). Identification of glucose-fermenting bacteria in a full-scale enhanced biological phosphorus removal plant by stable isotope probing. *Microbiology*, 158(Pt 7), 1818-1825. doi:10.1099/mic.0.058818-0
- Nielsen, P. H., Kragelund, C., Seviour, R. J., & Nielsen, J. L. (2009). Identity and ecophysiology of filamentous bacteria in activated sludge. *FEMS Microbiol Rev*, 33(6), 969-998. doi:10.1111/j.1574-6976.2009.00186.x
- Niu, Y. D., Stanford, K., Kropinski, A. M., Ackermann, H. W., Johnson, R. P., She, Y. M., Ahmed, R., Villegas, A., & McAllister, T. A. (2012). Genomic, proteomic and physiological characterization of a T5-like bacteriophage for control of Shiga toxin-producing *Escherichia coli* O157:H7. *PLoS One*, 7(4), e34585. doi:10.1371/journal.pone.0034585
- O'Loughlin, C. T., Miller, L. C., Siryaporn, A., Drescher, K., Semmelhack, M. F., & Bassler, B. L. (2013). A quorum-sensing inhibitor blocks *Pseudomonas aeruginosa* virulence and biofilm formation. *Proc Natl Acad Sci U S A*, 110(44), 17981-17986. doi:10.1073/pnas.1316981110
- Oehmen, A., Yuan, Z., Blackall, L. L., & Keller, J. (2005). Comparison of acetate and propionate uptake by polyphosphate accumulating organisms and glycogen accumulating organisms. *Biotechnol Bioeng*, 91(2), 162-168. doi:10.1002/bit.20500
- Ohnishi, M., Kurokawa, K., & Hayashi, T. (2001). Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors? *Trends Microbiol*, 9(10), 481-485.
- Oliveira, J., Castilho, F., Cunha, A., & Pereira, M. J. (2012). Bacteriophage therapy as a bacterial control strategy in aquaculture. *Aquaculture International*, 20(5), 879-910. doi:10.1007/s10499-012-9515-7
- Oliver, A., & Mena, A. (2010). Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin Microbiol Infect*, 16(7), 798-808. doi:10.1111/j.1469-0691.2010.03250.x
- Oppenheim, A. B., Kobilier, O., Stavans, J., Court, D. L., & Adhya, S. (2005). Switches in bacteriophage lambda development. *Annu Rev Genet*, 39, 409-429. doi:10.1146/annurev.genet.39.073003.113656
- Osorio, N. S., Rodrigues, F., Gagneux, S., Pedrosa, J., Pinto-Carbo, M., Castro, A. G., Young, D., Comas, I., & Saraiva, M. (2013). Evidence for diversifying selection in a set of *Mycobacterium tuberculosis* genes in response to antibiotic- and nonantibiotic-related pressure. *Mol Biol Evol*, 30(6), 1326-1336. doi:10.1093/molbev/mst038

Palmer, R. J., Jr., Gordon, S. M., Cisar, J. O., & Kolenbrander, P. E. (2003). Coaggregation-mediated interactions of streptococci and actinomyces detected in initial human dental plaque. *J Bacteriol*, 185(11), 3400-3409.

Pappenheimer, A. M., Jr. (1977). Diphtheria toxin. *Annu Rev Biochem*, 46, 69-94. doi:10.1146/annurev.bi.46.070177.000441

Park, S. H., Marassi, F. M., Black, D., & Opella, S. J. (2010). Structure and dynamics of the membrane-bound form of Pf1 coat protein: implications of structural rearrangement for virus assembly. *Biophys J*, 99(5), 1465-1474. doi:10.1016/j.bpj.2010.06.009

Patterson, D. J. (2001). Protozoan Diversity and Biogeography eLS.

Pederson, D. M., Welsh, L. C., Marvin, D. A., Sampson, M., Perham, R. N., Yu, M., & Slater, M. R. (2001). The protein capsid of filamentous bacteriophage PH75 from *Thermus thermophilus*. *J Mol Biol*, 309(2), 401-421. doi:10.1006/jmbi.2001.4685

Peng, Y.-z., Wang, X.-l., & Li, B.-k. (2006). Anoxic biological phosphorus uptake and the effect of excessive aeration on biological phosphorus removal in the A2O process. *Desalination*, 189(1), 155-164. doi:10.1016/j.desal.2005.06.023

Periasamy, S., & Kolenbrander, P. E. (2010). Central role of the early colonizer *Veillonella* sp. in establishing multispecies biofilm communities with initial, middle, and late colonizers of enamel. *J Bacteriol*, 192(12), 2965-2972. doi:10.1128/jb.01631-09

Pesci, E. C., Pearson, J. P., Seed, P. C., & Iglewski, B. H. (1997). Regulation of las and rhl quorum sensing in *Pseudomonas aeruginosa*. *J Bacteriol*, 179(10), 3127-3132.

Ploss, M., & Kuhn, A. (2010). Kinetics of filamentous phage assembly. *Phys Biol*, 7(4), 045002. doi:10.1088/1478-3975/7/4/045002

Pochana, K., & Keller, J. (1999). Study of factors affecting simultaneous nitrification and denitrification (SND). *Water Sci Technol*, 39(6), 61-68. doi:10.1016/S0273-1223(99)00123-7

Pronk, M., de Kreuk, M. K., de Bruin, B., Kamminga, P., Kleerebezem, R., & van Loosdrecht, M. C. M. (2015). Full scale performance of the aerobic granular sludge process for sewage treatment. *Water Res*, 84, 207-217. doi:10.1016/j.watres.2015.07.011

Ptashne, M. (2004). *A genetic switch*. New York: Cold Spring Harbor Laboratory Press.

Qin, L., Liu, Y., & Tay, J.-H. (2004). Effect of settling time on aerobic granulation in sequencing batch reactor. *Biochem Eng J*, 21(1), 47-52. doi:10.1016/j.bej.2004.03.005

Quinones, M., Kimsey, H. H., & Waldor, M. K. (2005). LexA cleavage is required for CTX prophage induction. *Mol Cell*, 17(2), 291-300. doi:10.1016/j.molcel.2004.11.046

Rahim, K., Qasim, M., Rahman, H., Khan, T. A., Ahmad, I., Khan, N., Ullah, A., Basit, A., & Saleha, S. (2016). Antimicrobial resistance among aerobic biofilm producing bacteria isolated from chronic wounds in the tertiary care hospitals of Peshawar, Pakistan. *J Wound Care*, 25(8), 480-486. doi:10.12968/jowc.2016.25.8.480

Rakonjac, J. (2012). Filamentous Bacteriophages: Biology and Applications eLS.

Ramazanzadeh, R., Zamani, S., & Zamani, S. (2013). Genetic diversity in clinical isolates of *Escherichia coli* by enterobacterial repetitive intergenic consensus (ERIC)-PCR technique in Sanandaj hospitals. *Iran J Microbiol*, 5(2), 126-131.

Recinos, D. A., Sekedat, M. D., Hernandez, A., Cohen, T. S., Sakhtah, H., Prince, A. S., Price-Whelan, A., & Dietrich, L. E. P. (2012). Redundant phenazine operons in *Pseudomonas aeruginosa* exhibit environment-dependent expression and differential roles in pathogenicity. *Proc Natl Acad Sci U S A*, 109(47), 19420.

Reisner, A., Haagenzen, J. A., Schembri, M. A., Zechner, E. L., & Molin, S. (2003). Development and maturation of *Escherichia coli* K-12 biofilms. *Mol Microbiol*, 48(4), 933-946.

Ren, T. T., Li, X. Y., & Yu, H. Q. (2013). Effect of N-acyl-L-homoserine lactones-like molecules from aerobic granules on biofilm formation by *Escherichia coli* K12. *Bioresour Technol*, 129, 655-658. doi:10.1016/j.biortech.2012.12.043

Ren, T. T., Yu, H. Q., & Li, X. Y. (2010). The quorum-sensing effect of aerobic granules on bacterial adhesion, biofilm formation, and sludge granulation. *Appl Microbiol Biotechnol*, 88(3), 789-797. doi:10.1007/s00253-010-2796-8

Rendulic, S., Jagtap, P., Rosinus, A., Eppinger, M., Baar, C., Lanz, C., Keller, H., Lambert, C., Evans, K. J., Goesmann, A., Meyer, F., Sockett, R. E., & Schuster, S. C. (2004). A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a genomic perspective. *Science*, 303(5658), 689-692. doi:10.1126/science.1093027

Rice, S. A., Tan, C. H., Mikkelsen, P. J., Kung, V., Woo, J., Tay, M., Hauser, A., McDougald, D., Webb, J. S., & Kjelleberg, S. (2009). The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. *ISME J*, 3(3), 271-282. doi:10.1038/ismej.2008.109

Rice, S. A., van den Akker, B., Pomati, F., & Roser, D. (2012). A risk assessment of *Pseudomonas aeruginosa* in swimming pools: a review. *J Water Health*, 10(2), 181-196. doi:10.2166/wh.2012.020

Ritz, H. L. (1967). Microbial population shifts in developing human dental plaque. *Arch Oral Biol*, 12(12), 1561-1568.

Roberts, J. W., & Roberts, C. W. (1975). Proteolytic cleavage of bacteriophage lambda repressor in induction. *Proc Natl Acad Sci U S A*, 72(1), 147-151.

Roberts, J. W., Roberts, C. W., & Craig, N. L. (1978). *Escherichia coli* recA gene product inactivates phage lambda repressor. *Proc Natl Acad Sci U S A*, 75(10), 4714-4718.

Rodriguez-Brito, B., Li, L., Wegley, L., Furlan, M., Angly, F., Breitbart, M., Buchanan, J., Desnues, C., Dinsdale, E., Edwards, R., Felts, B., Haynes, M., Liu, H., Lipson, D., Mahaffy, J., Martin-Cuadrado, A. B., Mira, A., Nulton, J., Pasic, L., Rayhawk, S., Rodriguez-Mueller, J., Rodriguez-Valera, F., Salamon, P., Srinagesh, S., Thingstad, T. F., Tran, T., Thurber, R. V., Willner, D., Youle, M., & Rohwer, F. (2010). Viral and microbial community dynamics in four aquatic environments. *ISME J*, 4(6), 739-751. doi:10.1038/ismej.2010.1

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasic, L., Thingstad, T. F., Rohwer, F., & Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nat Rev Micro*, 7(11), 828-836. doi:10.1038/nrmicro2235

Ronn, R., McCaig, A. E., Griffiths, B. S., & Prosser, J. I. (2002). Impact of protozoan grazing on bacterial community structure in soil microcosms. *Appl Environ Microbiol*, 68(12), 6094-6105.

Roux, S., Krupovic, M., Poulet, A., Debroas, D., & Enault, F. (2012). Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads. *PLoS One*, 7(7), e40418. doi:10.1371/journal.pone.0040418

Rozej, A., Cydzik-Kwiatkowska, A., Kowalska, B., & Kowalski, D. (2015). Structure and microbial diversity of biofilms on different pipe materials of a model drinking water distribution systems. *World J Microbiol Biotechnol*, 31, 37-47. doi:10.1007/s11274-014-1761-6

Russel, M. (1995). Moving through the membrane with filamentous phages. *Trends Microbiol*, 3(6), 223-228.

Russell, A. B., LeRoux, M., Hathazi, K., Agnello, D. M., Ishikawa, T., Wiggins, P. A., Wai, S. N., & Mougous, J. D. (2013). Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*, 496(7446), 508-512. doi:10.1038/nature12074

Sadikot, R. T., Blackwell, T. S., Christman, J. W., & Prince, A. S. (2005). Pathogen-host interactions in *Pseudomonas aeruginosa* pneumonia. *Am J Respir Crit Care Med*, 171(11), 1209-1223. doi:10.1164/rccm.200408-1044SO

Sadrzadeh, F., & Dulekgurgen, E. (2014). Improving the settling properties of activated sludge by gradually decreasing the settling time. *Desalination and Water Treatment*, 52(13-15), 2465-2471. doi:10.1080/19443994.2013.795874

- Sakuragi, Y., & Kolter, R. (2007). Quorum-sensing regulation of the biofilm matrix genes (*pel*) of *Pseudomonas aeruginosa*. *J Bacteriol*, 189(14), 5383-5386. doi:10.1128/jb.00137-07
- Sana, T. G., Berni, B., & Bleves, S. (2016). The T6SSs of *Pseudomonas aeruginosa* Strain PAO1 and Their Effectors: Beyond Bacterial-Cell Targeting. *Front Cell Infect Microbiol*, 6, 61. doi:10.3389/fcimb.2016.00061
- Sauer, K., Cullen, M. C., Rickard, A. H., Zeef, L. A., Davies, D. G., & Gilbert, P. (2004). Characterization of nutrient-induced dispersion in *Pseudomonas aeruginosa* PAO1 biofilm. *J Bacteriol*, 186(21), 7312-7326. doi:10.1128/jb.186.21.7312-7326.2004
- Scanlan, P. D., & Buckling, A. (2012). Co-evolution with lytic phage selects for the mucoid phenotype of *Pseudomonas fluorescens* SBW25. *ISME J*, 6(6), 1148-1158. doi:10.1038/ismej.2011.174
- Schmeisser, C., Steele, H., & Streit, W. R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl Microbiol Biotechnol*, 75(5), 955-962. doi:10.1007/s00253-007-0945-5
- Schultz, M. P., Bendick, J. A., Holm, E. R., & Hertel, W. M. (2011). Economic impact of biofouling on a naval surface ship. *Biofouling*, 27(1), 87-98. doi:10.1080/08927014.2010.542809
- Schweizer, H. P. (1991). *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19. *Gene*, 97(1), 109-121.
- Secor, P. R., Sweere, J. M., Michaels, L. A., Malkovskiy, A. V., Lazzareschi, D., Katznelson, E., Rajadas, J., Birnbaum, M. E., Arrigoni, A., Braun, K. R., Evanko, S. P., Stevens, D. A., Kaminsky, W., Singh, P. K., Parks, W. C., & Bollyky, P. L. (2015). Filamentous Bacteriophage Promote Biofilm Assembly and Function. *Cell Host Microbe*, 18(5), 549-559. doi:10.1016/j.chom.2015.10.013
- Sengun, I. Y., & Karabiyikli, S. (2011). Importance of acetic acid bacteria in food industry. *Food Control*, 22(5), 647-656. doi:10.1016/j.foodcont.2010.11.008
- Shapiro, O. H., Kushmaro, A., & Brenner, A. (2010). Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *ISME J*, 4(3), 327-336. doi:10.1038/ismej.2009.118
- Shea, P. R., Beres, S. B., Flores, A. R., Ewbank, A. L., Gonzalez-Lugo, J. H., Martagon-Rosado, A. J., Martinez-Gutierrez, J. C., Rehman, H. A., Serrano-Gonzalez, M., Fittipaldi, N., Ayers, S. D., Webb, P., Willey, B. M., Low, D. E., & Musser, J. M. (2011). Distinct signatures of diversifying selection revealed by genome analysis of respiratory tract and invasive bacterial populations. *Proc Natl Acad Sci U S A*, 108(12), 5039-5044. doi:10.1073/pnas.1016282108
- Shen, Y., Koller, T., Kreikemeyer, B., & Nelson, D. C. (2013). Rapid degradation of *Streptococcus pyogenes* biofilms by PlyC, a bacteriophage-encoded endolysin. *J Antimicrob Chemother*, 68(8), 1818-1824. doi:10.1093/jac/dkt104

Shimada, K., Itoh, Y., Washio, K., & Morikawa, M. (2012). Efficacy of forming biofilms by naphthalene degrading *Pseudomonas stutzeri* T102 toward bioremediation technology and its molecular mechanisms. *Chemosphere*, 87(3), 226-233. doi:10.1016/j.chemosphere.2011.12.078

Simoes, L. C., Simoes, M., & Vieira, M. J. (2008). Intergeneric coaggregation among drinking water bacteria: evidence of a role for *Acinetobacter calcoaceticus* as a bridging bacterium. *Appl Environ Microbiol*, 74(4), 1259-1263. doi:10.1128/aem.01747-07

Simons, R. W., & Kleckner, N. (1988). Biological regulation by antisense RNA in prokaryotes. *Annu Rev Genet*, 22, 567-600. doi:10.1146/annurev.ge.22.120188.003031

Singh, M., & Srivastava, R. K. (2010). Sequencing batch reactor technology for biological wastewater treatment: a review. *Asia Pac J Chem Eng*, 6(1), 3-13. doi:10.1002/apj.490

Slater, N. (2006). *Sequencing batch reactors: cost effective wastewater treatment*. Paper presented at the 32nd Annual Operators Seminar, Banff, Alberta.

Smolders, G. J. F., van der Meij, J., van Loosdrecht, M. C. M., & Heijnen, J. J. (1994). Model of the anaerobic metabolism of the biological phosphorus removal process: Stoichiometry and pH influence. *Biotechnol Bioeng*, 43(6), 461-470. doi:10.1002/bit.260430605

Sockett, R. E. (2009). Predatory lifestyle of *Bdellovibrio bacteriovorus*. *Annu Rev Microbiol*, 63, 523-539. doi:10.1146/annurev.micro.091208.073346

Spricigo, D. A., Bardina, C., Cortés, P., & Llagostera, M. (2013). Use of a bacteriophage cocktail to control *Salmonella* in food and the food industry. *Int J Food Microbiol*, 165(2), 169-174. doi:10.1016/j.ijfoodmicro.2013.05.009

Stolp, H., & Starr, M. P. (1963). *Bdellovibrio bacteriovorus* gen. et sp. n., a predatory, ectoparasitic, and bacteriolytic microorganism. *Antonie van Leeuwenhoek*, 29(1), 217-248. doi:10.1007/BF02046064

Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltry, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., Paulsen, I. T., Reizer, J., Saier, M. H., Hancock, R. E., Lory, S., & Olson, M. V. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature*, 406(6799), 959-964. doi:10.1038/35023079

Straus, S. K., Scott, W. R., Symmons, M. F., & Marvin, D. A. (2008). On the structures of filamentous bacteriophage Ff (fd, f1, M13). *Eur Biophys J*, 37(4), 521-527. doi:10.1007/s00249-007-0222-7

- Suttle, C. A. (2007). Marine viruses--major players in the global ecosystem. *Nat Rev Microbiol*, 5(10), 801-812. doi:10.1038/nrmicro1750
- Swain, G., Kovach, B., Touzot, A., Cassé, F., & Kavanagh, C. (2007). *Measuring the Performance of Today's Antifouling Coatings* (Vol. 23).
- Tan, C. H., Koh, K. S., Xie, C., Tay, M., Zhou, Y., Williams, R., Ng, W. J., Rice, S. A., & Kjelleberg, S. (2014). The role of quorum sensing signalling in EPS production and the assembly of a sludge community into aerobic granules. *ISME J*, 8(6), 1186-1197. doi:10.1038/ismej.2013.240
- Tay, J. H., Liu, Q. S., & Liu, Y. (2001a). The effects of shear force on the formation, structure and metabolism of aerobic granules. *Appl Microbiol Biotechnol*, 57(1), 227-233. doi:10.1007/s002530100766
- Tay, J. H., Liu, Q. S., & Liu, Y. (2001b). Microscopic observation of aerobic granulation in sequential aerobic sludge blanket reactor. *J Appl Microbiol*, 91(1), 168-175.
- Tay, J. H., Liu, Q. S., & Liu, Y. (2001c). The role of cellular polysaccharides in the formation and stability of aerobic granules. *Lett Appl Microbiol*, 33(3), 222-226.
- Tay, M. (2008). *Identification of the secondary phage in Pseudomonas aeruginosa and determination of its role in biofilm development*. University of New South Wales.
- Tay, M. (2013). *The Role of Bacteriophage in Granulation*. (Doctor of Philosophy (SBS)), Nanyang Technological University.
- Taylor, R. G., Walker, D. C., & McInnes, R. R. (1993). E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res*, 21(7), 1677-1678.
- Thanomsub, B., Pumeechockchai, W., Limtrakul, A., Arunrattiyakorn, P., Petchleelaha, W., Nitoda, T., & Kanzaki, H. (2006). Chemical structures and biological activities of rhamnolipids produced by *Pseudomonas aeruginosa* B189 isolated from milk factory waste. *Bioresour Technol*, 97(18), 2457-2461. doi:10.1016/j.biortech.2005.10.029
- Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr*, 45(6), 1320-1328. doi:10.4319/lo.2000.45.6.1320
- Thomas, J. A., Soddell, J. A., & Kurtboke, D. I. (2002). Fighting foam with phages? *Water Sci Technol*, 46(1-2), 511-518.
- Toerien, D., Hyman, K. L., & Bruwer, M. J. (1975). A preliminary trophic status classification of some South African impoundments. *water SA*, 1(1), 15.

Toyofuku, M., Inaba, T., Kiyokawa, T., Obana, N., Yawata, Y., & Nomura, N. (2016). Environmental factors that shape biofilm formation. *Biosci Biotechnol Biochem*, 80(1), 7-12. doi:10.1080/09168451.2015.1058701

Trap, J., Bonkowski, M., Plassard, C., Villenave, C., & Blanchart, E. (2016). Ecological importance of soil bacterivores for ecosystem functions. *Plant and Soil*, 398(1), 1-24. doi:10.1007/s11104-015-2671-6

Tsuneda, S., Miyauchi, R., Ohno, T., & Hirata, A. (2005). Characterization of denitrifying polyphosphate-accumulating organisms in activated sludge based on nitrite reductase gene. *J Biosci Bioeng*, 99(4), 403-407. doi:10.1263/jbb.99.403

Turk, O., & Mavinic, D. S. (1986). Preliminary assessment of a shortcut in nitrogen removal from wastewater. *Can J Civil Eng*, 13(6), 600-605. doi:10.1139/l86-094

Turner, K. H., Vallet-Gely, I., & Dove, S. L. (2009). Epigenetic control of virulence gene expression in *Pseudomonas aeruginosa* by a LysR-type transcription regulator. *PLoS Genet*, 5(12), e1000779. doi:10.1371/journal.pgen.1000779

Valle, J., Vergara-Irigaray, M., Merino, N., Penadés, J. R., & Lasa, I. (2007). $\sigma(B)$ Regulates IS256-Mediated *Staphylococcus aureus* Biofilm Phenotypic Variation. *J Bacteriol*, 189(7), 2886-2896. doi:10.1128/JB.01767-06

Vallet, I., Olson, J. W., Lory, S., Lazdunski, A., & Filloux, A. (2001). The chaperone/usher pathways of *Pseudomonas aeruginosa*: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. *Proc Natl Acad Sci U S A*, 98(12), 6911-6916. doi:10.1073/pnas.111551898

Van Houdt, R., & Michiels, C. W. (2010). Biofilm formation and the food industry, a focus on the bacterial outer surface. *J Appl Microbiol*, 109(4), 1117-1131. doi:10.1111/j.1365-2672.2010.04756.x

van Loosdrecht, M. C. M., Brandse, F. A., & de Vries, A. C. (1998). Upgrading of waste water treatment processes for integrated nutrient removal — The BCFS® process. *Water Sci Technol*, 37(9), 209-217. doi:10.1016/S0273-1223(98)00290-X

Van Melderén, L., & Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet*, 5(3), e1000437. doi:10.1371/journal.pgen.1000437

Verschuere, L., Rombaut, G., Sorgeloos, P., & Verstraete, W. (2000). Probiotic bacteria as biological control agents in aquaculture. *Microbiol Mol Biol Rev*, 64(4), 655-671.

Vigneron, A., Head, I. M., & Tsesmetzis, N. (2018). Damage to offshore production facilities by corrosive microbial biofilms. *Appl Microbiol Biotechnol*, 102(6), 2525-2533. doi:10.1007/s00253-018-8808-9

- Vornhagen, J., Stevens, M., McCormick, D. W., Dowd, S. E., Eisenberg, J. N., Boles, B. R., & Rickard, A. H. (2013). Coaggregation occurs amongst bacteria within and between biofilms in domestic showerheads. *Biofouling*, 29(1), 53-68. doi:10.1080/08927014.2012.744395
- Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W., Friedman, D. I., & Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol Microbiol*, 44(4), 957-970.
- Wagner, P. L., Neely, M. N., Zhang, X., Acheson, D. W., Waldor, M. K., & Friedman, D. I. (2001). Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *J Bacteriol*, 183(6), 2081-2085. doi:10.1128/jb.183.6.2081-2085.2001
- Wang, A. L., & Wang, C. C. (1991a). Viruses of parasitic protozoa. *Parasitol Today*, 7(4), 76-80.
- Wang, A. L., & Wang, C. C. (1991b). Viruses of the Protozoa. *Annu Rev Microbiol*, 45(1), 251-263. doi:10.1146/annurev.mi.45.100191.001343
- Wang, F., Wang, F., Li, Q., & Xiao, X. (2007). A novel filamentous phage from the deep-sea bacterium *Shewanella piezotolerans* WP3 is induced at low temperature. *J Bacteriol*, 189(19), 7151-7153. doi:10.1128/jb.00569-07
- Wang, S., Shi, W., Tang, T., Wang, Y., Zhi, L., Lv, J., & Li, J. (2017). Function of quorum sensing and cell signaling in the formation of aerobic granular sludge. *Rev Environ Sci Biotechnol*, 16(1), 1-13. doi:10.1007/s11157-017-9420-7
- Wang, X., Kim, Y., & Wood, T. K. (2009). Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. *ISME J*, 3(10), 1164-1179. doi:10.1038/ismej.2009.59
- Wang, Z., Liu, L., Yao, J., & Cai, W. (2006). Effects of extracellular polymeric substances on aerobic granulation in sequencing batch reactors. *Chemosphere*, 63(10), 1728-1735. doi:10.1016/j.chemosphere.2005.09.018
- Wanner, J., Ruzickova, I., Krhutkova, O., & Pribyl, M. (2000). Activated sludge population dynamics and wastewater treatment plant design and operation. *Water Sci Technol*, 41, 217-225.
- Webb, J. S., Lau, M., & Kjelleberg, S. (2004). Bacteriophage and phenotypic variation in *Pseudomonas aeruginosa* biofilm development. *J Bacteriol*, 186(23), 8066-8073. doi:10.1128/jb.186.23.8066-8073.2004
- Webb, J. S., Thompson, L. S., James, S., Charlton, T., Tolker-Nielsen, T., Koch, B., Givskov, M., & Kjelleberg, S. (2003). Cell death in *Pseudomonas aeruginosa* biofilm development. *J Bacteriol*, 185(15), 4585-4592.
- Weber, S. D., Ludwig, W., Schleifer, K. H., & Fried, J. (2007). Microbial composition and structure of aerobic granular sewage biofilms. *Appl Environ Microbiol*, 73(19), 6233-6240. doi:10.1128/aem.01002-07

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol Rev*, 28(2), 127-181. doi:10.1016/j.femsre.2003.08.001

Weissbrodt, D. G., Lochmatter, S., Ebrahimi, S., Rossi, P., Maillard, J., & Holliger, C. (2012). Bacterial Selection during the Formation of Early-Stage Aerobic Granules in Wastewater Treatment Systems Operated Under Wash-Out Dynamics. *Front Microbiol*, 3, 332. doi:10.3389/fmicb.2012.00332

Whiteley, M., Banger, M. G., Bumgarner, R. E., Parsek, M. R., Teitzel, G. M., Lory, S., & Greenberg, E. P. (2001). Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature*, 413(6858), 860-864. doi:10.1038/35101627

Wilderer, P. A., Arnz, P., & Arnold, E. (2000). Application of Biofilms and Biofilm Support Materials As A Temporary Sink and Source. *Water Air Soil Pollut*, 123(1), 147-158. doi:10.1023/a:1005237617334

Wilén, B.-M., Liébana, R., Persson, F., Modin, O., & Hermansson, M. (2018). The mechanisms of granulation of activated sludge in wastewater treatment, its optimization, and impact on effluent quality. *Appl Microbiol Biotechnol*, 102(12), 5005-5020. doi:10.1007/s00253-018-8990-9

Williamson, S. J., & Paul, J. H. (2006). Environmental Factors That Influence the Transition from Lysogenic to Lytic Existence in the *Listonella pelagia* Marine Phage: Host System. *Microb Ecol*, 52(2), 217-225.

Willner, D., Furlan, M., Haynes, M., Schmieder, R., Angly, F. E., Silva, J., Tammadoni, S., Nosrat, B., Conrad, D., & Rohwer, F. (2009). Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLoS One*, 4(10), e7370. doi:10.1371/journal.pone.0007370

Wilson, W., & Mann, N. (1997). *Lysogenic and lytic production in marine microbial communities* (Vol. 13).

Winkler, M. K. H., Kleerebezem, R., de Bruin, L. M. M., Verheijen, P. J. T., Abbas, B., Habermacher, J., & van Loosdrecht, M. C. M. (2013). Microbial diversity differences within aerobic granular sludge and activated sludge flocs. *Appl Microbiol Biotechnol*, 97(16), 7447-7458. doi:10.1007/s00253-012-4472-7

Winsor, G. L., Lam, D. K., Fleming, L., Lo, R., Whiteside, M. D., Yu, N. Y., Hancock, R. E., & Brinkman, F. S. (2011). *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Res*, 39(Database issue), D596-600. doi:10.1093/nar/gkq869

Winter, C., Bouvier, T., Weinbauer, M. G., & Thingstad, T. F. (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiol Mol Biol Rev*, 74(1), 42-57. doi:10.1128/mmbr.00034-09

Woo, J. K., Webb, J. S., Kirov, S. M., Kjelleberg, S., & Rice, S. A. (2012). Biofilm dispersal cells of a cystic fibrosis *Pseudomonas aeruginosa* isolate exhibit

variability in functional traits likely to contribute to persistent infection. *FEMS Immunol Med Microbiol*, 66(2), 251-264. doi:10.1111/j.1574-695X.2012.01006.x

Wood, T. K., Gonzalez Barrios, A. F., Herzberg, M., & Lee, J. (2006). Motility influences biofilm architecture in *Escherichia coli*. *Appl Microbiol Biotechnol*, 72(2), 361-367. doi:10.1007/s00253-005-0263-8

Wu, D., Ekama, G. A., Wang, H.-G., Wei, L., Lu, H., Chui, H.-K., Liu, W.-T., Brdjanovic, D., van Loosdrecht, M. C. M., & Chen, G.-H. (2014). Simultaneous nitrogen and phosphorus removal in the sulfur cycle-associated Enhanced Biological Phosphorus Removal (EBPR) process. *Water Res*, 49, 251-264. doi:10.1016/j.watres.2013.11.029

Wu, Q., & Liu, W. T. (2009). Determination of virus abundance, diversity and distribution in a municipal wastewater treatment plant. *Water Res*, 43(4), 1101-1109. doi:10.1016/j.watres.2008.11.039

Yaeger, R. G. (1996). Protozoa: Structure, Classification, Growth, and Development. In S. Baron (Ed.), *Medical Microbiology*. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston.

Yang, Q., Shen, N., Lee, Z. M. P., Xu, G., Cao, Y., Kwok, B., Lay, W., Liu, Y., & Zhou, Y. (2016). Simultaneous nitrification, denitrification and phosphorus removal (SNDPR) in a full-scale water reclamation plant located in warm climate. *Water Sci Technol*.

Yilmaz, G., Lemaire, R., Keller, J., & Yuan, Z. (2008). Simultaneous nitrification, denitrification, and phosphorus removal from nutrient-rich industrial wastewater using granular sludge. *Biotechnol Bioeng*, 100(3), 529-541. doi:10.1002/bit.21774

Yoo, H., Ahn, K.-H., Lee, H.-J., Lee, K.-H., Kwak, Y.-J., & Song, K.-G. (1999). Nitrogen removal from synthetic wastewater by simultaneous nitrification and denitrification (SND) via nitrite in an intermittently-aerated reactor. *Water Res*, 33(1), 145-154. doi:10.1016/S0043-1354(98)00159-6

Yoshida, A., & Kuramitsu, H. K. (2002). Multiple *Streptococcus mutans* Genes Are Involved in Biofilm Formation. *Appl Environ Microbiol*, 68(12), 6283-6291.

Yu, K., & Zhang, T. (2012). Metagenomic and metatranscriptomic analysis of microbial community structure and gene expression of activated sludge. *PLoS One*, 7(5), e38183. doi:10.1371/journal.pone.0038183

Zeng, R. J., Lemaire, R., Yuan, Z., & Keller, J. (2003). Simultaneous nitrification, denitrification, and phosphorus removal in a lab-scale sequencing batch reactor. *Biotechnol Bioeng*, 84(2), 170-178. doi:10.1002/bit.10744

Zhang, B., Ji, M., Qiu, Z., Liu, H., Wang, J., & Li, J. (2011). Microbial population dynamics during sludge granulation in an anaerobic-aerobic biological phosphorus removal system. *Bioresour Technol*, 102(3), 2474-2480. doi:10.1016/j.biortech.2010.11.017

Zhang, T., Shao, M. F., & Ye, L. (2012). 454 pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME J*, 6(6), 1137-1147. doi:10.1038/ismej.2011.188

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, 9(9), R137. doi:10.1186/gb-2008-9-9-r137

Zheng, Y.-M., Yu, H.-Q., Liu, S.-J., & Liu, X.-Z. (2006). Formation and instability of aerobic granules under high organic loading conditions. *Chemosphere*, 63(10), 1791-1800. doi:10.1016/j.chemosphere.2005.08.055

Zhou, Y., Ganda, L., Lim, M., Yuan, Z., Kjelleberg, S., & Ng, W. J. (2010). Free nitrous acid (FNA) inhibition on denitrifying poly-phosphate accumulating organisms (DPAOs). *Appl Microbiol Biotechnol*, 88(1), 359-369. doi:10.1007/s00253-010-2780-3

Zijnga, V., van Leeuwen, M. B., Degener, J. E., Abbas, F., Thurnheer, T., Gmur, R., & Harmsen, H. J. (2010). Oral biofilm architecture on natural teeth. *PLoS One*, 5(2), e9321. doi:10.1371/journal.pone.0009321

Appendix Figures and Tables

Appendix Figure 1. Weekly microscopic image of aerobic granular sludge development in reactors 1 to 4 (R1 – R4).

From left to right are images from week 0 to week 11 respectively, with the granule development stage indicated at the top. Floccular sludge undergoes aerobic granulation and are mature aerobic granules by week 11. Magnification x 40 (red bar, 200 μ m). The images were kindly taken by Dr Chan Siew Herng (SCELSE).

Appendix Figure 2. The particle size, SVI₅, and settling time for the individual reactors.

Appendix Figure 3. Nutrient removal profile of SBR.

The mean total ammonia removal (a) and total phosphorus removal (b) of four SBRS over a period of 11 weeks. Error bars represent standard deviation. $n = 4$.

Appendix Table 1. Percentage of sequencing reads across four SBRs.

Domain (%)					
Week	Total reads	Archaea	Bacteria	Eukarya	Unknown
0	3070427.5	0.732 ±	94.041 ±	0.042 ±	5.185 ±
		0.079	12.022	0.005	0.655
1	4417103.3	0.203 ±	93.090 ±	0.035 ±	6.672 ±
		0.019	8.071	0.003	0.581
2	5836281.55	0.111 ±	91.029 ±	0.035 ±	8.825 ±
		0.007	4.419	0.003	0.428
3	5927262.55	0.099 ±	90.059 ±	0.039 ±	9.803 ±
		0.007	7.556	0.006	1.485
4	6511833.55	0.076 ±	89.433 ±	0.038 ±	10.453 ±
		0.009	8.318	0.006	0.723
5	6948507.05	0.077 ±	88.742 ±	0.040 ±	11.141 ±
		0.010	6.069	0.008	1.140
6	6789836.8	0.069 ±	88.637 ±	0.040 ±	11.255 ±
		0.004	5.297	0.003	0.403
7	6804589.25	0.057 ±	87.621 ±	0.033 ±	12.289 ±
		0.004	9.163	0.004	1.657
8	6001128.75	0.054 ±	87.288 ±	0.032 ±	12.626 ±
		0.006	3.437	0.005	0.612
9	6229380.8	0.053 ±	87.322 ±	0.031 ±	12.594 ±
		0.007	5.188	0.004	0.989
10	6766966.55	0.053 ±	87.299 ±	0.030 ±	12.618 ±
		0.013	4.621	0.007	0.290
11	6197420.05	0.049 ±	87.244 ±	0.029 ±	12.677 ±
		0.009	6.307	0.005	1.282

Appendix Figure 4. Relative abundance plots at the genus (a) and phylum (b) level.

The colours represent the respective genera or phyla. The composition plot for the genera includes the contigs “Not annotated at Genus level” while the plot for the phyla does not include contigs “Not annotated at Phylum level”.

Appendix Table 2. Pearson's correlation coefficient between genera and particle size (50th percentile).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Pearson r coefficient															
-0.6108	-0.5305	0.8349	0.1263	-0.0910	0.6573	-0.7382	-0.5174	-0.1102	-0.5484	-0.5092	-0.6455	-0.6089	-0.0972	-0.7843	-0.4175
R squared															
0.3731	0.2814	0.6971	0.0159	0.0083	0.4320	0.5450	0.2677	0.0121	0.3007	0.2593	0.4167	0.3708	0.0094	0.6151	0.1743
P value (two-tailed)															
0.0349	0.076	0.0007	0.6956	0.7786	0.0202	0.0061	0.0849	0.7332	0.0649	0.0909	0.0234	0.0356	0.7639	0.0025	0.1769
*	ns	***	ns	ns	*	**	ns	ns	ns	ns	*	*	ns	**	ns

A : Burkholderia | B: Caldilinea | C: '*Candidatus* Accumulibacter' | D: '*Candidatus* Competibacter' | E: '*Candidatus* Contendobacter' | F: Dechloromonas | G: Fimbriimonas | H: Flavobacterium | I: Micavibrio | J: Nitrospira | K: Pseudomonas | L: Sulfuritalea | M: Terrimonas | N: Thauera | O: Verrucomicrobium | P: Zavarzinella

Appendix Table 3. Pearson's correlation coefficient between genera and SVI₅.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Pearson r coefficient															
0.5488	0.6845	-0.9384	-0.5285	-0.2895	-0.7336	0.4809	0.6955	-0.3065	0.7484	0.6205	0.8829	0.3362	0.3615	0.9363	0.415
R squared															
0.3012	0.4685	0.8805	0.2793	0.0837	0.5381	0.2313	0.4837	0.0939	0.5601	0.385	0.7795	0.1131	0.1307	0.8767	0.1722
P value (two-tailed)															
0.0646	0.0141	<0.0001	0.0773	0.3614	0.0066	0.1135	0.012	0.3326	0.0051	0.0313	0.0001	0.2852	0.2483	<0.0001	0.1797
ns	*	****	ns	ns	**	ns	*	ns	**	*	***	ns	ns	****	ns

A : Burkholderia | B: Caldilinea | C: '*Candidatus* Accumulibacter' | D: '*Candidatus* Competibacter' | E: '*Candidatus* Contendobacter' |

F: Dechloromonas | G: Fimbriimonas | H: Flavobacterium | I: Micavibrio | J: Nitrospira | K: Pseudomonas | L: Sulfuritalea |

M: Terrimonas | N: Thauera | O: Verrucomicrobium | P: Zavarzinella

Appendix Table 4. Number of reads mapped to individual Family.

SBR 1						
Week	Inoviridae	Podoviridae	Siphoviridae	Tectiviridae	Microviridae	Myoviridae
1	75	3136	3527	7	1312003	1755
2	27	517	899	4	861892	553
3	27	1295	3134	0	2021252	883
4	65	2232	1853	10	2474859	1682
5	9276	2655	2367	6	4601546	2999
6	0	377	320	1	454394	128
7	566	63000	22694	14249	1057590	15748
8	7175	57361	21854	124	719323	18595
9	180	99013	242706	2178	77107	31600
10	35	12000	19201	8	86709	7701
SBR 2						
Week	Inoviridae	Podoviridae	Siphoviridae	Tectiviridae	Microviridae	Myoviridae
1	56	4620	6019	143	1416262	2843
2	89	2590	4234	7	1296211	2921
3	12	1344	1540	14	1232880	1013
4	67	2947	2732	15	2226824	1995
5	11266	2265	2178	25	3841460	1967
6	16	811	420	2	1724858	317
7	4280	19483	13519	261	1758830	8168
8	3916	10899	7866	71	951457	3467
9	14821	116083	107314	170	1169214	28801
10	25	14572	7779	17	96812	3695
Table continued on next page						

Table continued from previous page

SBR 3						
Week	Inoviridae	Podoviridae	Siphoviridae	Tectiviridae	Microviridae	Myoviridae
1	300	8420	23027	2228	1697460	5724
2	35	932	1487	7	742211	1084
3	5	1158	1643	38	990727	967
4	30	1726	1575	14	3151509	1410
5	664	1142	936	27	1959618	1732
6	48	537	330	6	2123520	196
7	2997	31679	13833	2340	3553499	15726
8	8555	18258	9169	33	2567225	9417
9	9600	336482	98785	51506	156503	52721
10	128	34846	22327	42	189251	13555
SBR 4						
Week	Inoviridae	Podoviridae	Siphoviridae	Tectiviridae	Microviridae	Myoviridae
1	4080	8951	9291	6	1050044	4457
2	33	2151	2954	72	965617	2320
3	78	2709	4841	224	1050398	2042
4	724	7146	7425	17	872739	7206
5	1620	18806	9591	343	756761	10992
6	63	3357	2452	32	1315463	1795
7	15301	44857	20376	51268	719102	17111
8	67	51585	17739	113	78577	13296
9	26856	75912	42594	115	736959	40624
10	524	44337	29779	48	350045	14819

Appendix Table 5. Multiple t tests comparing Pf6 levels in biofilm effluent over 7 days.

Significant?	P value	PAO1 Mean1	PAO1ΔPf4 Mean2	Diff	SE of diff	t ratio	df
	0.998019	7.592474e+007	2.907666e+007	4.684808e+007	1.869854e+010	0.00250544	28
	0.997199	1.090934e+008	4.286224e+007	6.623121e+007	1.869854e+010	0.00354205	28
	0.995816	4.750430e+008	5.739832e+008	-9.894013e+007	1.869854e+010	0.00529133	28
	0.958773	3.538813e+008	1.329166e+009	-9.752842e+008	1.869854e+010	0.0521583	28
	0.968547	4.582508e+008	1.202177e+009	-7.439261e+008	1.869854e+010	0.0397853	28
	0.842088	5.179485e+009	1.419547e+009	3.759938e+009	1.869854e+010	0.201082	28
*	0.00925154	3.524783e+010	8.752241e+010	-5.227459e+010	1.869854e+010	2.79565	28

Appendix Figure 5. Per-gene standard deviation across samples against the rank of the mean.

The transformation methods used were the shifted logarithm $\log_2(n+1)$ (a), regularized log (b), and variance stabilizing transformation (c).

Appendix Table 6. All differentially expressed genes for Chapter 3.

Can be seen at <https://goo.gl/cycKxz>

Appendix Table 7. Gene set enrichment analysis by KEGG categories for 2 d PAO1ΔPf4 biofilms relative to PAO1 biofilms.

Name	N	Distinct direction				Non-directional		Mixed direction					
		Up		Down				Down			Up		
		p	padj	p	padj	p	padj	N	p	padj	N	p	padj
Citrate cycle (TCA cycle)	32	0.0035	0.012	1	1	0	0	11	0.0077	0.19	21	0	0
Oxidative phosphorylation	53	0	0	1	1	0	0	18	0.64	0.99	35	0	0
Purine metabolism	73	0	0	1	1	0	0	18	0.43	0.99	55	0	0
Pyrimidine metabolism	41	0	0	1	1	0	0	5	0.38	0.99	36	0	0
Valine and leucine and isoleucine biosynthesis	19	0.0028	0.0099	1	1	0	0	5	0.014	0.23	14	0.0003	0.0024
Aminoacyl-tRNA biosynthesis	49	0	0	1	1	0	0	4	0.81	0.99	45	0	0
Biosynthesis of siderophore group non-ribosomal peptides	6	1	1	0	0	0	0	6	0	0	0	NA	NA
Metabolic pathways	717	0	0	1	1	0	0	240	0.75	0.99	480	0	0
Biosynthesis of secondary metabolites	318	0	0	1	1	0	0	110	0.48	0.99	210	0	0
Carbon metabolism	123	0	0	1	1	0	0	43	0.26	0.99	80	0	0
2-Oxocarboxylic acid metabolism	30	0.0009	0.0038	1	1	0	0	9	0.18	0.99	21	0.0001	0.00087
Biosynthesis of amino acids	134	0	0	1	1	0	0	32	0.31	0.99	100	0	0
Ribosome	56	0	0	1	1	0	0	2	0.0062	0.19	54	0	0
Bacterial secretion system	67	1	1	0.0001	0.0055	0.0002	0.0016	47	0.0021	0.11	20	0.094	0.22
Histidine metabolism	24	0.13	0.26	0.87	1	0.0004	0.0028	10	0.019	0.27	14	0.0024	0.011
RNA polymerase	4	0	0	1	1	0.0004	0.0028	0	NA	NA	4	0.0004	0.0028
C5-Branched dibasic acid metabolism	13	0.082	0.18	0.92	1	0.0007	0.0045	5	0.1	0.92	8	0.0021	0.01
Mismatch repair	16	0.0001	0.00052	1	1	0.0034	0.021	2	0.3	0.99	14	0.0017	0.0088

Table continued on next page

Table continued from previous page

Protein export	18	0	0	1	1	0.0045	0.026	0	NA	NA	18	0.0037	0.015
Cysteine and methionine metabolism	39	0.0043	0.014	1	1	0.0077	0.041	14	0.43	0.99	25	0.0014	0.0077
Phenylalanine and tyrosine and tryptophan biosynthesis	26	0.0068	0.02	0.99	1	0.0081	0.041	8	0.23	0.99	18	0.0072	0.024
DNA replication	14	0	0	1	1	0.008	0.041	0	NA	NA	14	0.005	0.018
Pantothenate and CoA biosynthesis	18	0.0013	0.0051	1	1	0.0086	0.041	5	0.58	0.99	13	0.0009	0.0055
Pyruvate metabolism	54	0.41	0.64	0.59	1	0.011	0.051	30	0.39	0.99	24	0.0006	0.0039
Glycolysis or Gluconeogenesis	33	0.0087	0.024	0.99	1	0.012	0.053	13	0.6	0.99	20	0.0004	0.0028
Riboflavin metabolism	7	0.0079	0.022	0.99	1	0.013	0.053	1	0.32	0.99	6	0.018	0.055
Starch and sucrose metabolism	17	1	1	0.0027	0.074	0.013	0.054	13	0.0097	0.19	4	0.41	0.65
Microbial metabolism in diverse environments	217	0	0	1	1	0.019	0.076	89	0.81	0.99	130	0.0001	0.00087
Fatty acid biosynthesis	22	0.0002	0.001	1	1	0.026	0.098	5	0.76	0.99	17	0.0028	0.012

N: Number of genes; p : p value; padj: False discovery rate

The gene set enrichment is done by:

- 1) Selecting categories with p value < 0.05 and padj < 0.1 for the non-directional class.
- 2) If the p value is significant in either the up or down classes, then the gene set is regulated in a distinct direction
- 3) If the p value is not significant for either the up or down classes, then the gene set might have a mixed direction of regulation

The table cells are highlighted blue if the gene set is upregulated or pink if the gene set is downregulated.

Appendix Table 8. Gene set enrichment analysis by KEGG categories for 7 d PAO1ΔPf4 biofilms relative to PAO1 biofilms.

Name	N	Distinct direction				Non-directional		Mixed direction					
		Up		Down				Down			Up		
		p	padj	p	padj	p	padj	N	p	padj	N	p	padj
Biosynthesis of siderophore group non-ribosomal peptides	6	0	0	1	1	0	0	0	NA	NA	6	0	0
Bacterial chemotaxis	42	1	1	0	0	0.0032	0.18	31	0.001	0.1	11	0.66	0.98

N: Number of genes; p: *p* value; padj: False discovery rate

The gene set enrichment is done by:

- 1) Selecting categories with *p* value < 0.05 and padj < 0.1 for the non-directional class.
- 2) If the *p* value is significant in either the up or down classes, then the gene set is regulated in a distinct direction
- 3) If the *p* value is not significant for either the up or down classes, then the gene set might have a mixed direction of regulation

The table cells are highlighted blue if the gene set is upregulated or pink if the gene set is downregulated.

Appendix Table 9. Comparison of RNAseq and qPCR data.

		RNAseq	qPCR
Locus tag	Gene name	Log2 fold change	Log2 fold change
PA2386	<i>pvdA</i>	3.927	4.192
PA2426	<i>pvdS</i>	3.625	4.628
PA2392	<i>pvdP</i>	3.555	4.278
PA2397	<i>pvdE</i>	3.539	3.571
PA2395	<i>pvdO</i>	3.064	3.621
PA2385	<i>pvdQ</i>	2.941	4.382
PA2394	<i>pvdN</i>	2.863	3.562
PA2390	<i>pvdT</i>	2.495	2.925

Appendix Table 10. ChIPseq validation primers.

Primer	Sequences (5'-3')	Tm (°C)	Product size (bp)
phzA1_EMSA_F	[Btn]CCTTACCGAGCGACGA ACT	68	121
phzA1_EMSA_R	[Btn]CGAGAGGGCTCTCCA GGTAT	68	121
pilQup_EMSA_F	[Btn]CAGAAGGGCAACAAG GTGAT	64	80
pilQup_EMSA_R	[Btn]TCGAAGGTCTCGATGT TGAA	64	80
fixC_EMSA_F	[Btn]TACCAAGGTGCCCAAC TTCT	67	77
fixC_EMSA_R	[Btn]AGAGGTTGCCCAAGG AGAT	67	77
pldB_EMSA_F	[Btn]GGAAGTCGACGTTCTG CAAC	67	77
pldB_EMSA_R	[Btn]GATCCCTGGGGGCAG ATATT	67	77
aprA_EMSA_F	[Btn]ACCGATACGGAGGAAA ACC	65	119
aprA_EMSA_R	[Btn]TGCGTATTGTTCGACC TCTG	65	119
ahpF_EMSA_F	[Btn]CCTGGTCGGCAAGATC TAAG	62	150
ahpF_EMSA_R	[Btn]GAAATTCGGGCAAACG TATC	62	150
fptA_EMSA_F	[Btn]CCTCGGGGGATATAC GATTT	64	98
fptA_EMSA_R	[Btn]ACGGTCGTGGATCGAT AGAG	64	98

Appendix Table 11. Peaks detected by MACS2 for Pf4 RepC.

Locus tag	Name	Peak	Distance to TSS	-log10(pvalue)	Fold Enrichment	Description	PseudoCAP
PA0717	PA0717	peak_5a	0	179.659	6.37126	hypothetical protein of bacteriophage Pf1	Hypothetical, unclassified, unknown; Related to phage, transposon, or plasmid
PA5205	PA5205	peak_42	0	113.396	6.52198	conserved hypothetical protein	Membrane proteins
PA4021	PA4021	peak_30	380	85.618	5.40382	probable transcriptional regulator	Transcriptional regulators
PA4210	<i>phzA1</i>	peak_32a	0	85.064	5.29093	probable phenazine biosynthesis protein	Secreted Factors (toxins, enzymes, alginate)
PA5040	<i>pilQ</i>	peak_39	697	46.604	3.86006	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor	Motility & Attachment
PA2035	PA2035	peak_12	0	40.042	3.95772	probable decarboxylase	Putative enzymes
PA4673.13	PA4673.13	peak_35	-179	38.709	3.40552	Pf6	
PA2715	PA2715	peak_19	0	35.242	3.72661	probable ferredoxin	Energy metabolism
PA2953	PA2953	peak_22	160	34.689	3.49856	electron transfer flavoprotein-ubiquinone oxidoreductase	Energy metabolism
PA1477	<i>ccmC</i>	peak_9	0	33.072	3.58171	heme exporter protein CcmC	Membrane proteins; Transport of small molecules
Table continued on next page							

Table continued from previous page							
PA4979	PA4979	peak_38	0	30.098	3.46662	probable acyl-CoA dehydrogenase	Putative enzymes
PA0299	<i>spuC</i>	peak_2	0	27.260	3.20044	Polyamine:pyruvate transaminase	Putative enzymes; Carbon compound catabolism
PA0715	PA0715	peak_4a	0	24.504	3.05095	hypothetical protein	Related to phage, transposon, or plasmid
PA3305	PA3305	peak_24	0	22.682	3.06218	hypothetical protein	Membrane proteins
PA3016	PA3016	peak_23	0	21.914	2.8916	hypothetical protein	Hypothetical, unclassified, unknown
PA5089	<i>pldB</i>	peak_40	465	20.510	2.82994	PldB	Secreted Factors (toxins, enzymes, alginate)
PA3810	<i>hscA</i>	peak_27	154	20.323	2.89323	heat shock protein HscA	Chaperones & heat shock proteins
PA1249	<i>aprA</i>	peak_6	-110	18.793	2.83107	alkaline metalloproteinase precursor	Secreted Factors (toxins, enzymes, alginate)
PA3939	PA3939	peak_29	219	18.495	2.8021	hypothetical protein	Hypothetical, unclassified, unknown
PA0154	<i>pcaG</i>	peak_1	-113	15.460	2.50081	protocatechuate 3,4-dioxygenase, alpha subunit	Carbon compound catabolism
PA2327	PA2327	peak_15	-19	14.753	2.57107	probable permease of ABC transporter	Membrane proteins; Transport of small molecules
Table continued on next page							

Table continued from previous page							
PA2612	<i>serS</i>	peak_18	-99	14.753	2.57107	seryl-tRNA synthetase	Amino acid biosynthesis and metabolism; Translation, post-translational modification, degradation
PA4576	PA4576	peak_34	594	13.791	2.45922	probable ATP-dependent protease	Putative enzymes
PA1538	PA1538	peak_10	84	11.486	2.33997	probable flavin-containing monooxygenase	Putative enzymes
PA5103	<i>puuR</i>	peak_41	-79	10.928	2.26708	PuuR	Hypothetical, unclassified, unknown
PA5492	PA5492	peak_44	519	10.794	2.16447	conserved hypothetical protein	Hypothetical, unclassified, unknown
PA1806	<i>fabI</i>	peak_11	-127	10.720	2.28219	NADH-dependent enoyl-ACP reductase	Fatty acid and phospholipid metabolism
PA2345	PA2345	peak_16	0	10.720	2.28219	conserved hypothetical protein	Hypothetical, unclassified, unknown
PA2275	PA2275	peak_13	0	10.680	2.24062	probable alcohol dehydrogenase (Zn-dependent)	Putative enzymes
PA2475	PA2475	peak_17	423	10.559	2.26436	probable cytochrome P450	Adaptation, Protection; Putative enzymes
PA2937	PA2937	peak_21	314	9.276	1.96309	hypothetical protein	Hypothetical, unclassified, unknown
Table continued on next page							

Table continued from previous page							
PA1284	PA1284	peak_7	18	9.253	2.16664	probable acyl-CoA dehydrogenase	Putative enzymes
PA4293	<i>pprA</i>	peak_33	1754	8.562	2.07667	two-component sensor PprA	Two-component regulatory systems
PA5369.1	PA5369.1	peak_43	-166	7.933	1.9429	5S ribosomal RNA	Non-coding RNA gene
PA4156	<i>fvbA</i>	peak_31	0	7.875	2.05108	FvbA	Transport of small molecules
PA0717	PA0717	peak_5b	0	7.871	1.9989	hypothetical protein of bacteriophage Pf1	Hypothetical, unclassified, unknown; Related to phage, transposon, or plasmid
PA0715	PA0715	peak_4b	0	7.502	1.85463	hypothetical protein	Related to phage, transposon, or plasmid
PA2818	<i>arr</i>	peak_20	218	7.222	1.99331	aminoglycoside response regulator	Nucleotide biosynthesis and metabolism; Antibiotic resistance and susceptibility
PA4210	<i>phzA1</i>	peak_32b	0	6.990	1.9649	probable phenazine biosynthesis protein	Secreted Factors (toxins, enzymes, alginate)
PA5517	PA5517	peak_45	488	6.663	1.79795	conserved hypothetical protein	Hypothetical, unclassified, unknown
PA3316	PA3316	peak_25	-29	6.345	1.83515	probable permease of ABC transporter	Membrane proteins; Transport of small molecules
PA0717	PA0717	peak_5c	0	6.052	1.84109	hypothetical protein of bacteriophage Pf1	Hypothetical, unclassified, unknown; Related to phage, transposon, or plasmid
Table continued on next page							

Table continued from previous page							
PA3826	PA3826	peak_28	0	5.784	1.79297	hypothetical protein	Membrane proteins
PA3766	PA3766	peak_26	-244	5.480	1.80342	probable aromatic amino acid transporter	Transport of small molecules
PA4704	<i>cbpA</i>	peak_37	0	5.438	1.78182	cAMP-binding protein A	Hypothetical, unclassified, unknown
PA4690.2	PA4690.2	peak_36	0	5.091	1.67931	23S ribosomal RNA	Non-coding RNA gene
PA5523	PA5523	peak_46	0	4.901	1.58971	probable aminotransferase	Putative enzymes
PA1450	PA1450	peak_8b	-205	4.735	1.69336	conserved hypothetical protein	Hypothetical, unclassified, unknown
PA0668.4	PA0668.4	peak_3	42	4.412	1.67201	23S ribosomal RNA	Non-coding RNA gene
PA2306	<i>ambA</i>	peak_14	0	3.926	1.63948	AmbA	Membrane proteins; Secreted Factors (toxins, enzymes, alginate)
PA1450	PA1450	peak_8a	-205	3.782	1.59375	conserved hypothetical protein	Hypothetical, unclassified, unknown

Appendix Table 12. Pf4 RepC binding regions detected by Virtual Footprint.

Can be seen at <https://goo.gl/cycKxz>

Appendix Table 13. Peaks of Pf6 RepC identified by MACS2.

Can be seen at <https://goo.gl/cycKxz>

Appendix Table 14. Pf6 RepC binding regions for all motifs.

Can be seen at <https://goo.gl/cycKxz>

Appendix Table 15. Genes common in RNAseq and Virtual Footprint.

Locus tag	Name	Description	PseudoCAP
PA0020	<i>tsaP</i>	T4P secretin-associated protein TsaP	Hypothetical, unclassified, unknown
PA0036	<i>trpB</i>	tryptophan synthase beta chain	Amino acid biosynthesis and metabolism
PA0041	PA0041	probable hemagglutinin	Secreted Factors (toxins, enzymes, alginate)
PA0084	<i>tssC1</i>	TssC1	Protein secretion/export apparatus
PA0140	<i>ahpF</i>	alkyl hydroperoxide reductase subunit F	Adaptation, Protection
PA0459	PA0459	probable ClpA/B protease ATP binding subunit	Translation, post-translational modification, degradation
PA0546	<i>metK</i>	methionine adenosyltransferase	Amino acid biosynthesis and metabolism; Central intermediary metabolism
PA0572	PA0572	hypothetical protein	Hypothetical, unclassified, unknown

Table continued on next page

Table continued from previous page					
PA0581	PA0581	conserved protein	hypothetical	Hypothetical, unknown	unclassified,
PA0888	<i>aotJ</i>	arginine/ornithine protein AotJ	binding	Transport of small molecules	
PA0958	<i>oprD</i>	Basic amino acid, basic peptide and imipenem outer membrane porin OprD precursor		Transport of small molecules	
PA0976.1	PA0976.1	tRNA-Lys		Non-coding RNA gene	
PA1115	PA1115	hypothetical protein		Membrane proteins	
PA1131	PA1131	probable major facilitator superfamily transporter	(MFS)	Membrane proteins; Antibiotic resistance and susceptibility; Transport of small molecules	
PA1159	PA1159	probable cold-shock protein		Transcriptional regulators; Adaptation, Protection	
PA1249	<i>aprA</i>	alkaline metalloproteinase precursor		Secreted Factors (toxins, enzymes, alginate)	
PA1324	PA1324	hypothetical protein		Hypothetical, unknown	unclassified,
PA1337	<i>ansB</i>	glutaminase-asparaginase		Amino acid biosynthesis and metabolism	
PA1372	PA1372	hypothetical protein		Hypothetical, unknown	unclassified,
PA1592	PA1592	hypothetical protein		Hypothetical, unknown	unclassified,
PA2144	<i>glgP</i>	glycogen phosphorylase		Cell wall / LPS / capsule	
PA2147	<i>katE</i>	catalase HP11		Adaptation, Protection	
PA2151	PA2151	conserved protein	hypothetical	Hypothetical, unknown	unclassified,
PA2165	PA2165	probable glycogen synthase		Energy metabolism	
PA2167	PA2167	hypothetical protein		Hypothetical, unknown	unclassified,
Table continued on next page					

Table continued from previous page				
PA2323	PA2323	probable glyceraldehyde-3-phosphate dehydrogenase	Carbon catabolism	compound
PA2373	<i>vgrG3</i>	VgrG3	Hypothetical, unknown	unclassified,
PA2494	<i>mexF</i>	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter MexF	Transport of small molecules; Membrane proteins; Antibiotic resistance and susceptibility	
PA2642	<i>nuoG</i>	NADH dehydrogenase I chain G	Energy metabolism	
PA2751	PA2751	conserved hypothetical protein	Membrane proteins	
PA2777	PA2777	conserved hypothetical protein	Membrane proteins	
PA2862	<i>lipA</i>	lactonizing lipase precursor	Carbon catabolism; Secreted Factors (toxins, enzymes, alginate)	compound
PA3009	PA3009	hypothetical protein	Hypothetical, unknown	unclassified,
PA3145	<i>wbpL</i>	glycosyltransferase WbpL	Membrane proteins; Cell wall / LPS / capsule	
PA3151	<i>hisF2</i>	imidazoleglycerol-phosphate synthase, cyclase subunit	Amino acid biosynthesis and metabolism	
PA3157	PA3157	probable acetyltransferase	Putative enzymes; Cell wall / LPS / capsule	
PA3164	PA3164	still frameshift 3-phosphoshikimate 1-carboxyvinyltransferase prephenate dehydrogenase	pseudogene	
PA3329	PA3329	hypothetical protein	Hypothetical, unknown	unclassified,
PA3716	PA3716	hypothetical protein	Hypothetical, unknown	unclassified,
Table continued on next page				

Table continued from previous page				
PA4221	<i>fptA</i>	Fe(III)-pyochelin membrane precursor	outer receptor	Transport of small molecules
PA4296	<i>pprB</i>	two-component regulator, PprB	response	Transcriptional regulators; Two-component regulatory systems
PA4315	<i>mvaT</i>	transcriptional MvaT, P16 subunit	regulator	Transcriptional regulators
PA4370	<i>icmP</i>	Insulin-cleaving metalloproteinase membrane protein precursor	outer	Membrane proteins
PA4625	<i>cdrA</i>	cyclic diguanylate-regulated TPS partner A, CdrA		Secreted Factors (toxins, enzymes, alginate); Cell wall / LPS / capsule
PA4670	<i>prs</i>	ribose-phosphate pyrophosphokinase		Carbon catabolism; Nucleotide biosynthesis and metabolism
PA4877	PA4877	hypothetical protein		Hypothetical, unclassified, unknown
PA4879	PA4879	conserved protein	hypothetical	Hypothetical, unclassified, unknown
PA4933	PA4933	hypothetical protein		Membrane proteins
PA5040	<i>pilQ</i>	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor		Motility & Attachment
PA5111	<i>gloA3</i>	lactoylglutathione lyase		Central intermediary metabolism
PA5178	PA5178	conserved protein	hypothetical	Hypothetical, unclassified, unknown

The rows that are in **boldface** are genes that were found in both the Pf4 RepC ChIPseq and Virtual Footprint datasets and that were differentially expressed in PAO1 WT vs PAO1ΔPf4.