

Phenotypic characterization and genetic requirements of *Streptococcus pneumoniae* biofilms

Suyen Solange Espinoza Miranda

A dissertation for PhD
submitted to the Faculty of
the department of Biology
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

[December, 2023]

Phenotypic Characterization and Genetic Requirements of *Streptococcus pneumoniae* biofilms

Suyen Solange Espinoza Miranda

Advisors: Tim van Opijnen, Ph.D.

Abstract

Although bacteria are often studied as planktonic or free-living organisms, they frequently grow in complex surface-attached communities known as biofilms. Biofilms are communities of microorganisms attached to surfaces and embedded in a self-produced extracellular matrix. Biofilms are dynamic structures analogous to human settlements shaped by space and environment. These microbial communities fulfill critical roles in multiple infections in the human body. *Streptococcus pneumoniae* is a human pathogen that can cause biofilm-associated infections in various tissues and organs. This thesis offers a unique outlook for the study of *S. pneumoniae* biofilms by combining *in vitro*, genome-wide, and *in vivo* experiments to elucidate the complex population dynamics of *S. pneumoniae* biofilms. Existing methods to cultivate *S. pneumoniae* biofilms fail to fully capture the complexity of these communities, and most studies are limited to short periods of time. We developed a robust *in vitro* assay to grow *S. pneumoniae* biofilms. This assay can be maintained forever rather than days. We then use this robust assay to study their behavior *in vivo* and monitor disease outcomes. After establishing clear differences in biofilm and dispersal samples, we monitor population dynamics using genome-wide techniques (Tn-seq, RNA-seq and WGS) to provide some insights into this complex mode of growth. This work includes the first global identification of genetic requirements during biofilm establishment in two different *S. pneumoniae* strains using Tn-Seq. Coupled with our transcriptomic analysis, we found that genes involved in multiple pathways, such as capsule biosynthesis, nucleotide metabolism, and stress response, contributed to biofilm growth. Lastly, we studied the development of antibiotic resistance to three different types of antibiotics under *S. pneumoniae* biofilm conditions. We revealed common adaptive pathways to achieve biofilm growth and

antibiotic resistance (antibiotic target genes), as well as novel routes of adaptation to develop resistance. Our findings add to the growing body of knowledge in the field of bacterial genetics and antimicrobial resistance, paving the way for future research and therapeutic advancement.

TABLE OF CONTENTS

Table of Contents	iii
List of Tables.....	v
List of Figures.....	vi
List of Abbreviations	viii
Acknowledgements.....	x
Chapter 1. Introduction	1
1.1 Bacterial biofilm formation, structure, and function resemble those in human settlements.....	2
1.2 Biofilms are both phenotypically and genotypically diverse.....	5
1.3 The role of biofilms in pathogenesis.....	8
1.4 <i>Streptococcus pneumoniae</i> as a major pathogen and biofilm former.....	
1.4.1 Current knowledge and challenges in the study of <i>S. pneumoniae</i> biofilms.....	11
1.5 Overview of this thesis.....	13
1.6. References.....	15
Chapter 2. Development of a long-term robust and reproducible in vitro assay for <i>Streptococcus pneumoniae</i> biofilm growth.....	20
2.1 The incessant quest for the long-term in vitro biofilm assay.....	21
2.2 Results.....	
Development of a long-term <i>in vitro</i> assay for <i>Streptococcus pneumoniae</i> biofilm growth.....	25
Media composition optimization to support robust biofilm growth.....	27
Fluorescent microscopy visually confirms SDMM is the best choice of growth medium.....	29
Glucose concentration affects biofilm formation.....	30
Transferring the biofilms is necessary to promote long-term biofilm growth.....	32
Use of Confocal Microscopy to study biofilm population dynamics.....	33
Biofilm characterization of <i>S. pneumoniae</i> clinical isolates with different virulence backgrounds.....	38
2.3 Conclusions.....	41
2.4 Materials and Methods.....	43
2.5 References.....	46
Chapter 3 Phenotypic Characterization of TIGR4 <i>S. pneumoniae</i> biofilms.....	48
3.1 Background.....	49
3.2 Results.....	50
3.2.1 TIGR4 <i>S. pneumoniae</i> biofilms undergo spatial heterogeneity.....	50
3.2.2 <i>In vivo</i> characterization: Different growth states from the strain TIGR4 exhibit varying levels of virulence.....	49
3.2.3 Phenotypic antibiotic tolerance profile in <i>S. pneumoniae</i> biofilms.....	57
3.3 Conclusions.....	
3.4 Methods.....	65
3.5 References.....	69

Chapter 4. Deciphering genetic requirements of <i>Streptococcus pneumoniae</i> biofilms	72
4.1 The value of high functional genomic approaches in bacterial research.	73
4.2 Results	
4.2.1 Time-lapse Tn-Seq uncovers genetic requirements between two <i>S. pneumoniae</i> strains. grown as biofilms.	74
4.2.2 Genes contributing to biofilm formation are distributed across multiple functional categories.	83
4.2.3 Tn-seq validation	86
4.2.4 Capsule levels could explain some of the high-biofilm phenotypes of <i>Streptococcus pneumoniae</i> knockouts.	90
4.2.5 Cyclic di-AMP signaling in <i>Streptococcus pneumoniae</i> biofilms.	93
4.2.6 <i>Streptococcus pneumoniae</i> biofilm-forming factors unrelated to capsule.	97
4.2.7 <i>In vivo</i> phenotype of knockouts with different biofilm formation capacities.	99
4.3 Conclusions	102
4.4 Methods	104
4.5 References	107
Chapter 5. Development of antibiotic resistance in <i>S. pneumoniae</i> biofilms	111
5.1 Antimicrobial resistance in <i>S. pneumoniae</i>	112
5.2 The role of adaptive diversification in biofilms towards antibiotic resistance	112
5.3 Results	
5.3.1 Identification of genomic factors strongly favor biofilm selection in <i>S. pneumoniae</i>	115
5.3.2 Genes members of cell wall assembly, capsule metabolism, and carbohydrate transport are favored for biofilm growth.	117
5.3.3 Development of antibiotic resistance in <i>S. pneumoniae</i> biofilms.	123
5.4 Conclusions	131
5.5 Methods	133
5.6 Supplementary Information	135
5.7 References	136
Chapter 6. Discussion	141
6.1 Delving into the complexity of <i>S. pneumoniae</i> biofilms: A systems-level approach.	142
6.2 Disease outcome differs according to cell growth state and biofilm formation capacity.	144
6.3 Looking ahead: The promise of biofilm and dispersal research.	145
6.4 References	147
Appendix A. Transcriptome profiling of TIGR4 <i>S. pneumoniae</i> biofilms	150
1. Background	151
2. Results	152
3. Conclusions	169
4. Methods	171
5. Supplementary Information	174
6. References	177

LIST OF TABLES

Chapter 2. Development of a long-term robust and reproducible in vitro assay for *Streptococcus pneumoniae* biofilm growth.

Table 1. Methods used for growing and quantifying <i>S. pneumoniae</i> biofilms	p 23-24
Table 2. SDMM recipe	p 44-45
Table 3. List of <i>S. pneumoniae</i> strains used.	p 35

Chapter 4. Deciphering genetic requirements of *S. pneumoniae* biofilm growth

Table 1. Selected genes for in vitro validation	p 86
Table 2. Primers used in this study to generate knockouts.	p 105

Chapter 5. Development of antibiotic resistance in *S. pneumoniae* biofilms

Table 1. Mutations over 50% frequency in all population after 40 days of biofilm growth	p 135
--	-------

Appendix A. Transcriptome profiling of TIGR4 *S. pneumoniae* biofilms

Table 1. Sequence of adapters and primers used for RNA sequencing	p 172
Table 2. Average read count per sample	p 174
Table 3. Top 10 upregulated genes in biofilms	p 175
Table 4. Top 10 downregulated genes in biofilms	p 175
Table 5. Top 10 upregulated genes in dispersal	p 176
Table 6. Top 10 upregulated genes in dispersal	p 176

LIST OF FIGURES

Chapter 1. Introduction

- Figure 1.** Bacterial biofilms have structural and functional organization similar to a metropolitan area. p 3
- Figure 2.** The biofilm life cycle progresses through several stages p 4
- Figure 3.** Graphical abstract of this thesis. p 14

Chapter 2. Development of a long-term robust and reproducible in vitro assay for *Streptococcus pneumoniae* biofilm growth.

- Figure 1.** The biofilm life cycle in the context of pneumococcal endocarditis p 21
- Figure 2.** Development of a robust and consistent long term in vitro biofilm assay for *S.pneumoniae*. p 26
- Figure 3.** SDMM shows a higher cell viability recovery in biofilm during eight days of growth p 28
- Figure 4.** *S. pneumoniae* biofilm formation with different media. p 30
- Figure 5.** Effect of different glucose and initial bacterial optical density (OD₆₀₀) concentration on biofilm formation. p 31
- Figure 6.** Glucose concentration affects *S. pneumoniae* biofilm development p 32
- Figure 7.** Biofilm-growth assay can be carried out for weeks allowing to perform long-term experimental adaptation in biofilms. p 33
- Figure 8.** Quantification parameters used for biofilm characterization. p 34
- Figure 9.** Previously characterized strains validated the assay developed in this thesis. p 36
- Figure 10.** Two-dimensional principal component analysis separates low from high biofilm former strains. p 37
- Figure 11.** Biofilm characterization of seven *S. pneumoniae* clinical isolates. p 40

Chapter 3. Phenotypic Characterization of TIGR4 *S. pneumoniae* biofilms.

- Figure 1.** Biofilm and dispersal growth dynamics of the strain TIGR4. P 53
- Figure 2.** Different growth states from the strain TIGR4 exhibit varying levels of virulence. p 55
- Figure 3.** The virulence phenotype of dispersal cells is exacerbated when the inoculum is increased. p 56
- Figure 4.** Differentiating Resistance, Tolerance and Persistence p 58
- Figure 5.** Minimum Inhibitory Concentration (MIC) determination for vancomycin and ciprofloxacin p 60
- Figure 6.** *S. pneumoniae* biofilm and dispersal cells harbors tolerant cells upon addition of antibiotic. p 62
- Figure 7.** Dispersal cells are not a dormant population. p 63
- Figure 8.** Summary of tolerant assay for biofilm and dispersal cells samples. p 68

Chapter 4. Deciphering genetic requirements of *Streptococcus pneumoniae* biofilms.

Figure 1. TIGR4 and BHN97 forms distinct type of biofilms.	p 75
Figure 2. Tn-seq identifies genetic requirements for <i>S. pneumoniae</i> biofilm growth and maintenance.	p 80
Figure 3. Distribution of significant genes across all time points.	p 81
Figure 4. Venn diagrams show number shared genes across different time points from TIGR4.	p 82
Figure 5 Tn-seq reveals functional gene categories important for biofilm formation.	p 84
Figure 6. Detailed view of mutant genes and their contribution during biofilm formation.	p 85
Figure 7. <i>In vitro</i> validations of selected mutants.	p 89
Figure 8. Reduction of capsule increases biofilm formation.	p 92
Figure 9. Cyclic di-AMP levels across all mutants.	p 96
Figure 10. Low and high biofilm maker strains have different phenotypes <i>in vivo</i> .	p 100

Chapter 5. Development of antibiotic resistance in *S. pneumoniae* biofilms

Figure 1. Consequences of biofilm growth for population-genetic diversity.	p 114
Figure 2. Summary of variants detected during 40 days of biofilm growth.	p 116
Figure 3. Biofilm growth dynamics of adapted populations.	p 118
Figure 4. Evolutionary dynamics of bacterial population A (Pop A).	p 119
Figure 5. Evolutionary dynamics of bacterial population E (Pop E).	p 121
Figure 6. Genotype trajectories and abundance of adapted populations reflected in Muller plots.	p 122
Figure 7. Overview of methodology for identification of adapted genes under antibiotic conditions.	p 124
Figure 8. Frequency of trajectories for mutations over 50% at passage 10 in levofloxacin adapted populations.	p 126
Figure 9. Frequency of trajectories for mutations over 50% at passage 10 in vancomycin adapted populations.	p 127
Figure 10. Frequency of trajectories for mutations over 50% at passage 10 in rifampicin adapted populations.	p 129

Appendix A. Transcriptome profiling of TIGR4 *S. pneumoniae* biofilms

Figure 1. Optimization of RNA isolation from biofilm and dispersal cells	p 153
Figure 2. Transcriptional landscape of biofilms and dispersal cells.	p 154
Figure 3. Enrichment analysis of all DE genes significant under both biofilm and dispersal cells revealed a multitude of processes key for three-day old biofilms.	p 156
Figure 4. Multiple avenues for regulation during biofilm and dispersal states.	p 158
Figure 5. Different transcriptional responses among low and high biofilm strains.	p 162
Figure 6. Summary data visualization of DE genes in Δ SP_2205.	p 165
Figure 7. Summary data visualization of DE genes in Δ SP_1645.	p 167
Figure 8. Summary data visualization of DE genes in Δ SP_1739.	p 168

ABBREVIATIONS

EPS	Extracellular Polymeric Substance
ARG	Antibiotic Resistance Genes
HGT	Horizontal Gene Transfer
CF	Cystic Fibrosis
WHO	World Health Organization
BAI	Biofilm Associated Infections
CbpA	Choline Binding Protein A
LytA	Autolysin
CMV	Colony Morphology Variant
ARG	Antibiotic Resistance Genes
MIC	Minimum Inhibitory Concentration
SCV	Small Colony Variants
RA	Roughness Coefficient
OD	Optical Density
THY	Todd Hewitt + Yeast
SDMM	Semi-defined Minimal Media
CDM	Chemically Define Minimal Media
CFU	Colony Forming Units
CSLM	Confocal Scanning Laser Microscopy
PCA	Principal Component Analysis
PC1	Principal Component 1
PC2	Principal Component 2
ANOVA	Analysis of Variance
CPS	Capsule
PBS	Phosphate Buffered Saline
PG_	Pangenome
IAV	Influenza A virus
FIJI	Fiji is just Image J
H.p.i	Hours post infection
DE	Differentially Expressed
PTS	Phosphotransferase System
CSP	Competence Stimulated Peptide

CAMP	Cationic Antimicrobial Peptide
MDK	Minimum Duration of Killing
RIN	RNA Integrity
i.n	Intranasally
BA	Blood Agar
IACUC	Institutional Animal Care and Use Committee
TIS	Transposon Insertion. Sequencing
Tn-seq	Transposon Sequencing
TA	Thymine-adenine
FDR	False Discovery Rate
W	Fitness
WT	Wild Type
ELISA	Enzyme-linked Immunosorbent Assay
CDC	Center for Disease Control and Prevention
MDR	Multidrug-resistant
AMR	Antimicrobial Resistance
WPGS	Whole-Population Genome Sequencing

ACKNOWLEDGEMENTS

I have dreamt multiple times with this moment. The moment where I am finally writing the acknowledgements of my thesis. The road was not easy and the many times I wanted to quit were countless. But I was able to finish this Ph.D. thanks to the amazing support system I had. This thesis is dedicated to:

My dad, I know you are extremely proud of me, and I wouldn't have been able to finish this if it wasn't for your words "El mundo esta en tus manos" and I can only say this kept me moving after you left.

My mom, thanks for teaching that even in our lowest moments we can find beauty and comfort. For always believing I could make this happening.

To my brother and the rest of my family, thank you so much for cheering me up at every single step of the way.

To Hector, thanks for always believing in me, for doing this Ph.D. by my side, for constantly lifting me up, especially when that impostor syndrome hits hard and for bringing Chamu into my life.

To Max, thanks for always being there, for the laughs, the many audible credits you got me were the only way I could finish those animal experiments.

To Dr. Huete, my first mentor, thank you for encouraging me to apply for a Ph.D. and for believing I could finish this. For the countless advices while going through this and for being an inspiration.

To all my people, my tribe here in Boston: Irem, Bharathi, Ali, Defne, Tuan, Fati, doña Alixe, thank you for taking care of me, for being here for the day to the day craziness, for all the unconditional love and support you have given me for the past six years. I am grateful I ended up in Boston and got to meet each one of you.

To my friends in Nicaragua, Lu, Ana, Cata, Fania, Mary, Lynda, Adri you guys brought me joy and unforgettable moments anytime I came home for a quick vacation. Thank you for your words of encouragement and for reassuring every time I was capable of finishing this.

To current and old members of the lab, thank you for being welcoming, for all the great scientific discussions and most importantly for being part of the long hours in lab.

To all my mentors: Federico, thank you for all the guidance and for pushing me to finish this with the best possible outcome. Tim, thanks for all the support along these years and for giving me a chance to do my Ph.D in your lab. Michelle, thank you for being a role model in science, for all the mentorship and unconditional support. Vaughn, thanks for being an outstanding mentor and all the insightful conversations throughout my Ph.D. Babak, thank you for your unvaluable feedback and amazing mentorship, I am glad I started my Ph.D journey with a mentor like you.

And finally, this also goes to my country, to all those little girls who think their dreams are unattainable.

Dreams do come true!

Chapter 1

Introduction*

*Parts of this chapter are taken from: Espinoza-Miranda S., van Opijnen T. and Cooper V. : Causes and consequences of the evolution of genetic diversity in bacterial biofilms. 'In preparation' invited to *Nature Reviews in Microbiology*.

1.1 Bacterial biofilm formation, structure, and function resemble those in human settlements.

A human settlement is an organized permanent or temporary residence for people on a small or large area of land that contains all the necessary infrastructure for optimal living. Urban settlements are characterized by high population density, variety of economic activities and a complex social structure. Similar to humans, bacteria mostly grow and exist in nature as communities or physical aggregates [1] (settlements) and less commonly they can also be found as individuals (commuters). Free-swimming individual bacteria are referred to as planktonic, and the phenomenon of aggregation is termed a biofilm [2, 3].

Biofilms and human settlements are characterized by spatial heterogeneity, meaning that different regions have a variety of characteristics and functions [4, 5] (**Figure 1**). In biofilms, spatial diversity creates selective pressures and opportunities for genetic exchange, leading to genotypic and phenotypic diversification. Cells in biofilms can transition from uncommitted to committed by responding to various cues such as environmental stress, nutrient availability, and cell density. A possible analogy to this process is that commuters become urban residents. Commuters resemble planktonic cells that travel to and from the city but do not settle permanently. Urban residents are like biofilm cells that live in cities and contribute to their structure and function. Commuters can become urban residents by responding to different factors such as job opportunities, housing prices, social networks, and personal preferences. Some commuters may also have a family history of living in the city, which makes them more likely to become urban residents. Finally, migration is also possible; just as humans can relocate to find new opportunities, cells can detach from the biofilm and migrate to other areas to look for new sites to colonize and create new microcolonies.

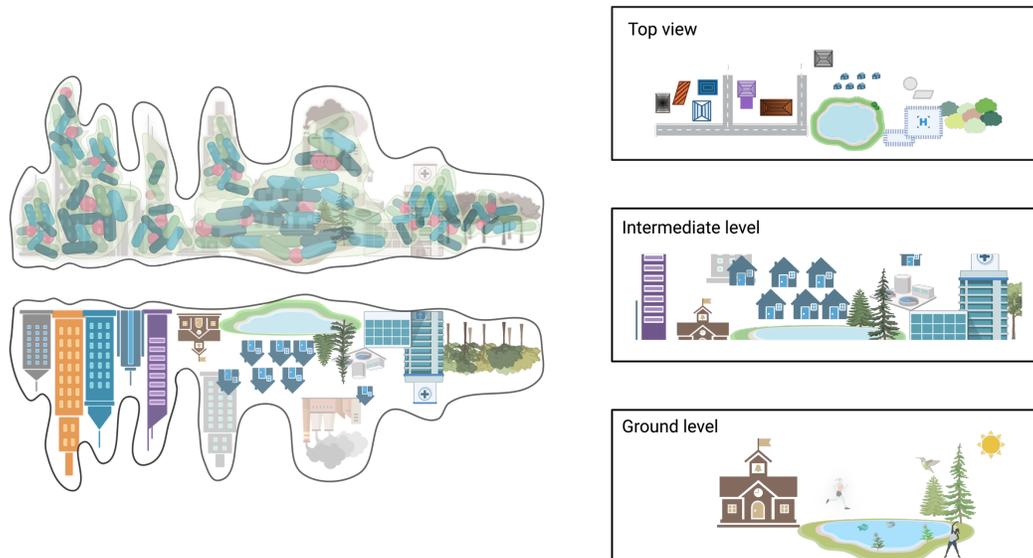


Figure 1. Bacterial biofilms have structural and functional organization like a metropolitan area. As bacteria attach and expand on a surface, the population is fragmented into distinct niches that respond differently to the environment. These distinct niches interact with each other to have a functional entity just as they do in a city. We visualize biofilm development from different angles: top view, intermediate level and ground level. At each level common genetic pathways play an important role in the establishment and structural framework of bacterial communities. A **top view** of this city-like metaphor is used to appreciate different emergent adaptive phenotypes such as colony morphology variants. As we descend into the city (**intermediate level**), more details of the composition of building architecture and social interactions can be appreciated. In the case of a biofilm, the intermediate level view could capture the composition of the Extracellular Polymeric Substance (EPS) matrix encasing biofilm cells and where cross-feeding between different cell aggregates takes place. Finally, at the **ground level**, we can explore the interactions and behaviors of individual bacteria within the biofilm, such as motility, quorum sensing, and detachment.

Some of the current understanding of biofilms is largely anchored by the view that biofilms represent a sessile developmental stage in the life of unicellular microbial organisms. Biofilm-forming organisms undergo a **life cycle (Figure 2)** that involves both sessile and motile stages: **planktonic, surface-attached, and dispersed cells**. During this life cycle, biofilm formation is initiated by the attachment of **planktonic** cells to the substrate. **Planktonic** cells are bacteria that do not associate with the biofilm matrix. These can be observed in the surrounding liquid medium. This cycle is followed by the

proliferation and recruitment of cells from the surrounding environment, generating the foundation for a biofilm structure (**Figure 2**- steps 3 &4). These **surface-attached** cells in the biofilm then form microcolonies, and as bacteria grow and divide in the biofilms, they release extracellular DNA, polysaccharides, lipids, and proteins, forming an extracellular polymeric (EPS) matrix [6]. Formation of the EPS is vital for the biofilm life cycle as it provides a defense mechanism that encases cells in a “slime” structure that protects bacteria from environmental stresses such as antibiotics and host immune responses. When a mature biofilm develops, bacteria detach releasing **dispersal** cells, that can serve as inocula for the initiation and development of new biofilms [2, 5]. These bacteria can disperse in aggregates of different sizes or as single cells. Dispersal cells are essential for the survival of biofilms, as they are responsible for maintaining the life cycle. Despite their clinical significance as they contribute to the progression of disease through the active release of cells [3, 7, 8], dispersal cell research has receive little attention. In chapter 3, we expand on dispersal cells and their relevance in causing disease.

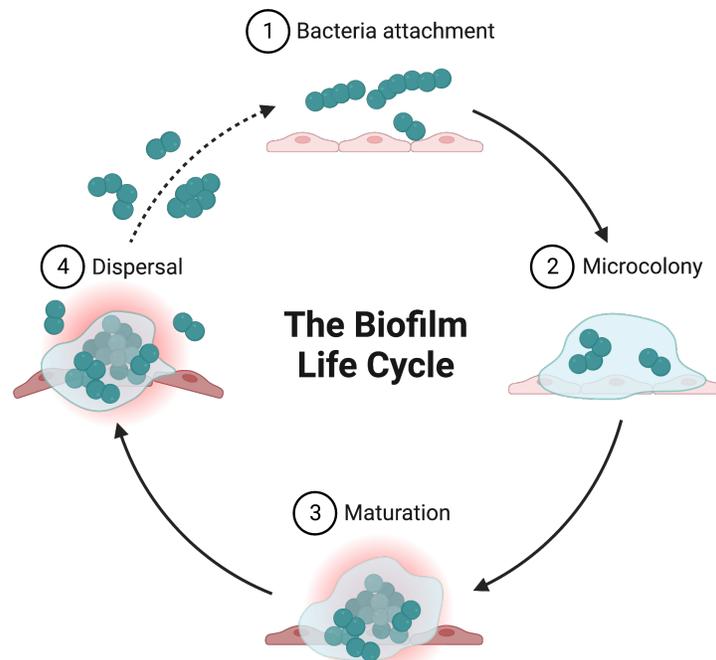


Figure 2. The biofilm life cycle progresses through several stages:1) attachment of free-swimming (planktonic) cells to a surface; 2) and 3) expansion and microcolony formation mediated by cell division and a self-produced extracellular polymeric substance (EPS) matrix that holds the biofilm

together; 4) once the biofilm is mature, cells actively escape through the disassemble of microcolonies releasing dispersal cells which then return to the cycle to colonize another surface.

Every stage of the biofilm life cycle is modulated by regulators that receive both internal and external signals [9]. Many of these master genetic regulators have been well-characterized and are shared among multiple species. One example is the small molecule c-di-GMP which at least partially controls the reprogramming of the bistable motile-to-sessile phenotype in biofilms and can produce distinct colony variants in *Pseudomonas* and *Bacilli* species [10, 11]. Rather than acting as an ON and OFF switch for biofilm development, recent work showed that c-di-GMP acts as an oscillator at the single-cell level that can directly impact cell lineages and requires a more elaborate and fine-tuned sensing response from the c-di-GMP regulatory machine [12]. Additionally, Armbruster *et al.* found that high c-di-GMP-producing cells can be seen as “biofilm founders” [13], which is in line with the **ground level** characterization in our city metaphor.

1.2 Biofilms are both phenotypically and genotypically diverse.

Bacterial populations growing in biofilms can change their physical properties and genetic makeup in response to both intrinsic and extrinsic stimuli [14]. Some of these physical properties include increased surface adherence, biomass production, and facilitation of bacterial persistence [15] which in turn can produce spatial and transcriptional heterogeneity independent of the selective pressure. For example, when nutrients do not reach all cells within a biofilm, bacteria can switch their metabolic pathways to survive and continue growing [16].

Biofilms possess the remarkable ability to alter their genetic content to adapt to their surroundings. Cell-to-cell interactions in biofilms create an ideal environment for horizontal gene transfer (HGT) for a variety of mobile elements [17]. Horizontal transfer rates are typically higher in biofilm communities than in planktonic states [18]. For example, after cultivating a combination of 25 staphylococcal strains commonly found on human skin in the form of biofilms, gene and plasmid transfers from one strain

were detected in another [19]. Not only do biofilms facilitate plasmid exchange but they also ensure the preservation of these mobile elements. Genetic diversity is more likely to be maintained in biofilms than planktonic growth. For instance, in *P. putida* plasmid loss occurs at a higher rate in planktonic cells but is maintained within the biofilms. This is done by quantifying the expression of an mCherry reporter within the plasmids [20]. This maintenance of mobile elements suggests that biofilms serve as reservoirs of genotypic diversity.

Biofilm structure. The spatial structure of biofilms is shaped by the production of **EPS components, nutrient availability, cell-to-cell interactions,** and **external factors.** The emergence of novel colony morphotypes is a trademark of phenotypic diversity in biofilms. Wrinkly, rugose, small, opaque, and transparent phenotypes have been linked to increased attachment and bacterial persistence. The emergence of colony variants traits often results from altered production of EPS, fimbriae, flagella, or other traits affecting aggregation at the cellular level. Such morphological colony variants have been observed in a wide range of species in natural environments, such as *Vibrio* species attached to aquatic plants and animals [21] and during chronic infections of the cystic fibrosis (CF) airway [22].

Nutrient availability. Since nutrients are not uniformly available within the biofilm, cells are subject to experience varying levels of nutrients. The outer layers of the biofilm are exposed to higher concentrations of nutrients and oxygen than the inner layers [23]. This gradient in nutrient availability creates different microenvironments within the biofilm that can support different phenotypic variants with varying metabolic activities. For example, anaerobic microorganisms are often found in the inner layers of biofilms, where oxygen concentrations are low. This can also give rise to slow-growth phenotypes [2, 24], which are variants characterized by restricted transcriptional and metabolic programs because of limited nutrients and/or oxygen [25, 26]. Nutrient availability in biofilms can thereby play a critical role in determining the growth rate and metabolic activity of bacterial cells within biofilms.

Cell-to-cell interactions. As aggregates form and grow, they may interact with neighboring aggregates, changing the environment and prevailing selective forces on one another. A classic example of the interaction between ecology and spatial arrangement comes from *P. fluorescens*. Experiments showed how isogenic populations diversified in multiple morphological phenotypes, which were maintained in the presence of environmental heterogeneity [27]. These diverse morphological phenotypes consume different resources facilitating total biofilm growth. For example, Poltak *et al.* observed that three distinct colony morphology variants, isolated from *Burkholderia cenocepacia* biofilm evolution, exhibited increased biomass exclusively when all three variants coexisted. This biomass expansion was the outcome of each morphotype's contribution to the biofilm architecture at different stages of biofilm establishment [28]. However, in cases where these variants compete for a common critical resource, antagonistic interactions may emerge. On the other hand, cells that generate valuable shared resources, such as polymers, can guide their progeny towards oxygen-rich parts of the biofilm; simultaneously gaining an edge over non-producer cells by utilizing available oxygen [29]. Facilitation of cell-to-cell interactions within biofilms leads to diverse morphological phenotypes that promote resource consumption and contribute to overall biomass growth.

Environmental interactions. In addition to nutrient availability and cell-cell interactions determining the spatial arrangement of biofilms, these microbial communities are also subjected to environmental interactions [30]. Environmental heterogeneity in biofilms can increase the evolution of antibiotic resistance even in the absence antibiotic pressure. For example, *E. coli* biofilms grown for 15 days without antibiotic treatment showed increased resistance to multiple classes of antibiotics compared to the initial inoculum [31]. Short- and long-term evolution experiments of *in vitro* biofilms have shown that accumulation of mutations intensifies when harsh environmental pressures are applied, resulting in antibiotic resistant phenotypes [32]. It has been concluded that an increase in the number of mutational changes can be linked to accelerated biofilm production [33, 34], changes in antibiotic

susceptibility, and the maintenance of genotypic diversity [35]. Bacterial resistance to last-resort antibiotics results in dire patient outcome. Exposing *P. aeruginosa* to 20 and 40 times the minimum inhibitory concentration (MIC) of imipenem in a cyclic manner resulted in multiple hyperbiofilm and hypervirulent phenotypes [36]. In summary, the environmental interactions and inherent biofilm heterogeneity play a significant role in driving the evolution of antibiotic resistance.

The intricate interplay between genetic makeup, environmental cues, and cooperative interactions shapes the structure and behavior of biofilms, leading to cell heterogeneity [37]. This diversity in phenotypes and genotypes enables biofilms to better withstand environmental pressures and survive for extended periods of time.

1.3 The role of biofilms in pathogenesis.

Biofilms can form on or in various tissues or 'surfaces' in the human body, including the skin, teeth, organs, epithelial cells and mucosa, as well as on medical devices such pacemakers, urinary tract catheters, prosthetic joints, breast implants among others [38-40]. Biofilms have been detected in a variety of bacterial infections in humans [41] and are usually linked to other afflictions. In fact, 60-80% of infections in hospitals are attributed to BAI (biofilm-associated infections)[42]. Examples of BAIs include pneumonia caused by *P. aeruginosa* in cystic fibrosis patients and catheter-related infections caused by *E. coli* [43]. Biofilm-related infections can also be caused by multiple species; for instance, *P. aeruginosa* and *S. aureus* have been found together respiratory infections in Cystic fibrosis (CF) patients and in diabetic foot infections and on prosthetics [44, 45].

The ability of cells to adhere to surfaces and form a protective polymeric matrix enables them to withstand environmental stresses including drugs and the immune system, leading to chronic and persistent infections. In addition to the **barrier effect** created by the presence of the extracellular matrix, which reduces the diffusion and penetration of antibiotics into biofilms, there are other

characteristics that contribute to pathogenesis [46, 47]. Biofilms harbor **persister cells**, which are a subpopulation of dormant or slow-growing bacteria that can survive high concentrations of antibiotics and resume growth when the antibiotic treatment is discontinued [48, 49]. It is worth noting that persister cells can be stochastically generated within the biofilm, as cells experience a variety of nutrient fluxes during the different growth stages [49]. A third reason is **genetic exchange** among bacterial species within the biofilm, which facilitates the horizontal transfer of antibiotic resistance genes (ARGs) through mechanisms such as plasmids and transposons. The causes mentioned above are closely linked to the fact that biofilms are spatially structured and environmentally heterogeneous, generating distinct ecological opportunities that select for both generalist and specialist genotypes.

Current knowledge of how biofilms may contribute to the pathogenesis of disease indicates several different mechanisms. For example, biofilms not only serve as a source of pathogenic bacteria that can detach and invade host tissues, but they can also modulate the host immune response either by evading immune detection or clearance or exacerbating inflammation [41]. *S. aureus* biofilm impairs macrophage polarization, suppressing bactericidal and pro-inflammatory activities by inhibiting the activation of NF- κ B, a pathway involved in innate immune responses [50, 51]. Likewise, *A. baumannii* and *P. aeruginosa* biofilms have been observed to cause delayed wound healing, inadequate vascularization, epidermis necrosis, and an augmented inflammatory response by the host in *in vivo* burn wound models [52]. Therefore, biofilms are not just passive entities during an infection but rather but are active contributors to the disease outcome.

Considering the significant threat that BAIs pose to healthcare, it is crucial to have advanced treatments and diagnostic methods in development to combat and eliminate these infections. Traditional bacterial diagnostic methods in the clinic can be labor intensive, unreliable and mostly developed using planktonic growth [53]. These methods may not be suitable for BAI detection due to several factors 1) the EPS can interfere with diagnostic reagents, 2) the prevalence of different physiological states,

such as persister cells, can lead to false-negative results and 3) capturing cells at specific stages of the dynamic biofilm life cycle in real time might not be possible [43]. Although emerging diagnostics promise a rapid and sensitive detection, there is lack of attention of bacteria growing as biofilms. Accurate diagnosis of BAIs is essential making better treatment decisions.

1.4 *Streptococcus pneumoniae* as a major pathogen and biofilm former.

A medically relevant biofilm-forming bacterium is the opportunistic pathogen *Streptococcus pneumoniae*, which colonizes the human respiratory tract. This gram-positive bacterium can migrate to sterile tissues and organs, causing acute and chronic infections [54, 55]. Children aged 0-5 years old are particularly susceptible to *S. pneumoniae* infections, with a prevalence of the bacterium in this age group often reaching 50-60% [56]. As a result, they are at a higher risk of developing associated diseases. The polysaccharide capsule of *S. pneumoniae* is the current target of pneumococcal vaccines. To date over 90 distinct capsular polysaccharides have been identified and they are used as a metric to differentiate strains serotypes [57]. Despite the development of vaccines, pneumococcal infections remain as one of the leading causes of pneumonia, meningitis, and middle ear infections with children and the elderly population at higher risks of developing any of these infections [58, 59].

In 2017, the World Health Organization (WHO) registered *S. pneumoniae* as one of the 12 serious threat pathogens due to the rising rates of antibiotic resistance found in *S. pneumoniae*-related infections in hospitals. Resistance to β -lactams and macrolide antibiotics is the most common type of resistance among *S. pneumoniae* isolates worldwide [60]. Clinical studies have shown that patients with an infection caused by an antibiotic-resistant isolate may experience additional symptoms and require ventilation and longer regimens of more than one antibiotic treatment [61].

Much of the research on *S. pneumoniae* has been devoted to exploring the molecular mechanisms of colonization and attachment to the nasopharynx, as well as the transmission from host to host [54]. *S.*

pneumoniae biofilms can be formed on abiotic or biotic surfaces such as plastic, glass, epithelial cells, or mucin layers[62]. *S. pneumoniae* biofilms are associated with asymptomatic colonization of the nasopharynx, which is a prerequisite for invasive disease. These biofilms are also involved in chronic and recurrent infections of the middle ear (otitis media), sinuses (sinusitis), and heart (pneumococcal endocarditis), which can be difficult to treat even with a combination of antibiotics. In the nasopharynx, *S. pneumoniae* biofilms are resistant to desiccation and influence the co-infection of other bacterial species recruited to the tissue [63]. Diseases such as otitis media are explained by the presence of biofilms, and over 80% of the cases are associated with a single strain pneumococcal isolate [56]. In addition, higher bacterial loads in the nasopharynx are correlated with the incidence of ear infections [56]. Considering the widespread occurrence of pneumococcal infections, additional research is required to minimize the impact of *S. pneumoniae* BAI.

1.4.1 Current knowledge and challenges in the study of *S. pneumoniae* biofilms

Evidence of *S. pneumoniae* to interact with epithelial cells is dated to 1992 and this interplay was initially described as “forming a thickened gelatinous layer”[64]. Subsequent reports have highlighted the role *S. pneumoniae* in causing diseases across various respiratory track tissues, it can also migrate to the heart, cause microlesions and evade the host response by quickly destroying macrophages [65]. In addition, nasopharyngeal carriage and colonization are essential precursors to invasive pneumococcal disease (IPD). These findings have sparked growing interested in biofilm research of this bacterial specie.

S. pneumoniae biofilms have been cultivated, but the current knowledge is limited to the isolated study of the attachment and expansion stages of the biofilm life cycle. This approach overlooks the holistic nature of biofilms. One of the main challenges in cultivating *S. pneumoniae* biofilms is that conventional techniques are unable to accurately replicate the extended duration of bacterial incubation that occurs during infections. Bacterial incubation within the host lasts for much longer periods than what

conventional methods can replicate[66]. For example, assays have been conducted for less than 3 days, while clinical evidence suggests that *S. pneumoniae* can persist in the nasopharynx of children for an average of 2 to 10 weeks [15].

From the attachment stage studies have revealed that genes related to cell adhesion and virulence, such as (i) adhesin choline binding protein A (CbpA), which helps bacteria attach to mucosal epithelial cells; (ii) polysaccharide capsule; and (iii) autolysin (LytA), which facilitates the release of pneumococcal DNA, are all necessary for the initial step of biofilm formation. Proteomic analysis of *S. pneumoniae* biofilms has revealed the presence of proteins associated with virulence, energy and nucleotide metabolism, capsule formation, cell adhesion, and resistance [62, 67]. Biofilms in *S. pneumoniae* also lead to the emergence of different colony variants which have been previously reported in other well-studied biofilm species[62, 68]. Colony morphology variants (CMV) and a hyper adherence phenotype are generated in *S. pneumoniae* biofilms shortly after attachment. This suggests the complexity of *S. pneumoniae* biofilms roots from phenotypic and genotypic heterogeneity. While this knowledge has founded the basis for understanding some of the most basic behavior of biofilm it disregards the hard evidence that biofilms progress beyond attachment and these known virulence factors might constitute only a minor subset of the genetic requirements for biofilm maintenance and cell dispersion.

Having a robust in vitro assay also allows for the study of the whole biofilm life cycle as an integrated process. In species such as *P. aeruginosa* looking at all the biofilms' stages have been provided information not just about bacterial attachment but also untangling the mechanism behind cell dispersion. For patients, dispersal cells can have severe implications, as it would allow bacteria to colonize other parts of the body such as the case of *Burkholderia* or *Pseudomonas species* migrating from CF lung of patients to cause systemic infections [69, 70]. The knowledge of *S. pneumoniae* dispersal cells is extremely limited, and these cells have just been investigated in the context of dispersion upon external conditions such as changes in temperature [71].

Given that pneumococcal colonization is frequent and infection cases are higher, disturbing this commensal bacterium might have a huge impact in decreasing the number of upper respiratory related diseases. Therefore, this thesis aims to provide a comprehensive understanding of the full complete *S. pneumoniae* biofilm life cycle. This study will contribute on the genetic factors of biofilm formation and possible targets to be used for biofilm disruption.

1.5 Overview of this thesis and its aims

Biofilms are highly dynamic and heterogeneous communities. To improve our understanding of *S. pneumoniae* biofilms beyond the well-studied 24-hour models, in Chapter 2 we propose and validate a reliable and consistent assay that enables the long-term study of *S. pneumoniae* biofilms. This assay separates low from high biofilm formers by combining microscopy visualization with quantification. Our assay serves as a platform for exploring the growth dynamics of different strains of biofilm in detail.

In Chapter 3, we conduct a comprehensive phenotypic analysis of biofilm dynamics *in vitro* and *in vivo* phenotypes by inoculating mice with three cells states: planktonic, biofilms, and dispersal cells to assess the capacity of these three different phenotypic populations to cause disease. Additionally, we demonstrate that the developed assay provides a new platform that enables isolation of antibiotic-tolerant cells from surface-attached and dispersed cells. In Chapter 4, we employ Tn-seq, as a genome wide approach to identify the key genetic elements involved in the formation of *S. pneumoniae* biofilms at different stages of development in two distinct biofilm-forming strains. As a result, we detected mutants with the ability to produce less and more biofilm and explored the contribution to disease outcome.

Finally, since biofilms are known to be inherent reservoirs of antibiotic recalcitrance (tolerance and resistance), the last chapter aims to understand the emergence of antibiotic resistance in biofilms. We discuss how the study of experimentally evolved biofilm conditions has enabled the discovery of novel routes of evolutionary dynamics that are key to understanding phenomena, such as the emergence of antibiotic resistance. We present a summary and overview of the genes that are altered during biofilm growth and those that acquire mutations under in the presence of three different types of antibiotics.

This work offers an improved method for cultivating *S. pneumoniae* biofilms in the laboratory that allows for the detailed study of in vitro biofilm population dynamics. Our method can generate cells in a ready-to-infect state that exacerbate disease outcome. We use cutting-edge omics techniques to pinpoint the essential genes for *S. pneumoniae* biofilm formation, the transcriptional differences between low and high biofilm formers, and the trajectory of these populations to acquire antibiotic resistance.

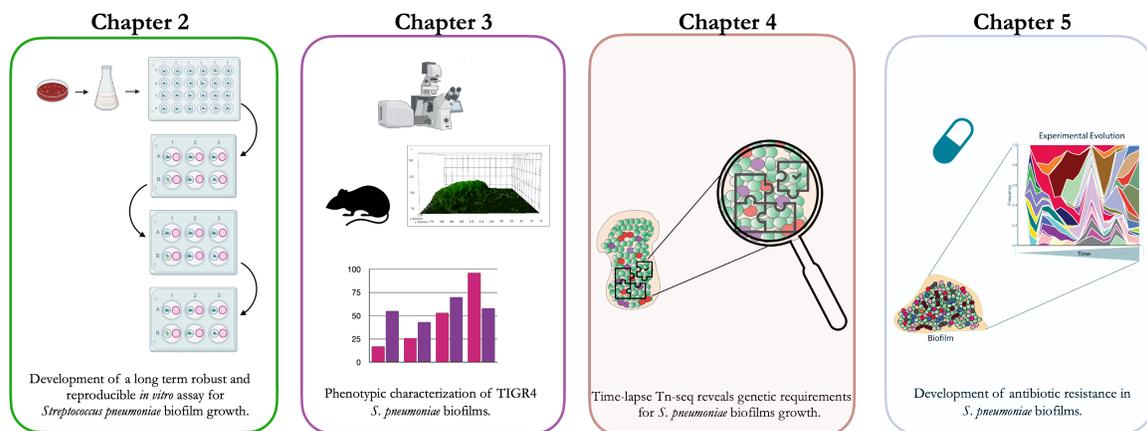


Figure 3. Graphical abstract of this thesis. **Chapter 2.** Development of a long-term robust and reproducible in vitro assay for *Streptococcus pneumoniae* biofilm growth. **Chapter 3.** In vitro and in vivo phenotypic characterization of the strain TIGR4. **Chapter 4.** Time-lapse Tn-seq revealed the genetic requirements for *S. pneumoniae* biofilm growth. **Chapter 5.** Development of antibiotic resistance in *S. pneumoniae* biofilms.

References

1. Cai, Y.M., *Non-surface Attached Bacterial Aggregates: A Ubiquitous Third Lifestyle*. Front Microbiol, 2020. **11**: p. 557035.
2. Penesyan, A., et al., *Three faces of biofilms: a microbial lifestyle, a nascent multicellular organism, and an incubator for diversity*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 80.
3. Sauer, K., et al., *The biofilm life cycle: expanding the conceptual model of biofilm formation*. Nat Rev Microbiol, 2022. **20**(10): p. 608-620.
4. Paula, A.J., G. Hwang, and H. Koo, *Dynamics of bacterial population growth in biofilms resemble spatial and structural aspects of urbanization*. Nat Commun, 2020. **11**(1): p. 1354.
5. Watkiss Paula, K.R., *Biofilm, City of Microbes*. Journal of bacteriology, 2000. **182**(10).
6. Flemming, H.C. and J. Wingender, *The biofilm matrix*. Nat Rev Microbiol, 2010. **8**(9): p. 623-33.
7. O'Rourke, D., et al., *There and back again: consequences of biofilm specialization under selection for dispersal*. Front Genet, 2015. **6**: p. 18.
8. Rumbaugh, K.P. and K. Sauer, *Biofilm dispersion*. Nat Rev Microbiol, 2020. **18**(10): p. 571-586.
9. Romling, U., M.Y. Galperin, and M. Gomelsky, *Cyclic di-GMP: the first 25 years of a universal bacterial second messenger*. Microbiol Mol Biol Rev, 2013. **77**(1): p. 1-52.
10. Flynn, K.M., et al., *Evolution of Ecological Diversity in Biofilms of Pseudomonas aeruginosa by Altered Cyclic Diguanylate Signaling*. J Bacteriol, 2016. **198**(19): p. 2608-18.
11. Conner, J.G., et al., *The ins and outs of cyclic di-GMP signaling in Vibrio cholerae*. Curr Opin Microbiol, 2017. **36**: p. 20-29.
12. Lee, C.K., et al., *Broadcasting of amplitude- and frequency-modulated c-di-GMP signals facilitates cooperative surface commitment in bacterial lineages*. Proc Natl Acad Sci U S A, 2022. **119**(4).
13. Armbruster, C.R., et al., *Heterogeneity in surface sensing suggests a division of labor in Pseudomonas aeruginosa populations*. Elife, 2019. **8**.
14. Lujan, A.M., et al., *Evolution and adaptation in Pseudomonas aeruginosa biofilms driven by mismatch repair system-deficient mutators*. PLoS One, 2011. **6**(11): p. e27842.
15. Tyerman, J.G., et al., *The evolution of antibiotic susceptibility and resistance during the formation of Escherichia colibiofilms in the absence of antibiotics*. BMC Evolutionary Biology, 2013. **13**(22).
16. Malviya, J., et al., *Metabolomic profiling of bacterial biofilm: trends, challenges, and an emerging antibiofilm target*. World J Microbiol Biotechnol, 2023. **39**(8): p. 212.
17. Madsen, J.S., et al., *The interconnection between biofilm formation and horizontal gene transfer*. FEMS Immunol Med Microbiol, 2012. **65**(2): p. 183-95.

18. Luo, A., et al., *Formation, Development, and Cross-Species Interactions in Biofilms*. Front Microbiol, 2021. **12**: p. 757327.
19. Aguila-Arcos, S., et al., *Biofilm-Forming Clinical Staphylococcus Isolates Harbor Horizontal Transfer and Antibiotic Resistance Genes*. Front Microbiol, 2017. **8**: p. 2018.
20. Roder, H.L., et al., *Biofilms can act as plasmid reserves in the absence of plasmid specific selection*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 78.
21. Ma, L., et al., *Analysis of Pseudomonas aeruginosa conditional psl variants reveals roles for the psl polysaccharide in adhesion and maintaining biofilm structure postattachment*. J Bacteriol, 2006. **188**(23): p. 8213-21.
22. Starkey, M., et al., *Pseudomonas aeruginosa rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung*. J Bacteriol, 2009. **191**(11): p. 3492-503.
23. Young, E., G. Melaugh, and R.J. Allen, *Active layer dynamics drives a transition to biofilm fingering*. NPJ Biofilms Microbiomes, 2023. **9**(1): p. 17.
24. Petroff, A.P., et al., *Reaction-diffusion model of nutrient uptake in a biofilm: theory and experiment*. J Theor Biol, 2011. **289**: p. 90-5.
25. Serra, D.O. and R. Hengge, *A c-di-GMP-Based Switch Controls Local Heterogeneity of Extracellular Matrix Synthesis which Is Crucial for Integrity and Morphogenesis of Escherichia coli Macrocolony Biofilms*. J Mol Biol, 2019. **431**(23): p. 4775-4793.
26. Cantillon, D., et al., *Three-dimensional low shear culture of Mycobacterium bovis BCG induces biofilm formation and antimicrobial drug tolerance*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 12.
27. Rainey, P. and M. Travisano, *Adaptive radiation in a heterogeneous environment*. Nature, 1998. **394**: p. 69-72.
28. Poltak, S.R. and V.S. Cooper, *Ecological succession in long-term experimentally evolved biofilms produces synergistic communities*. ISME J, 2011. **5**(3): p. 369-78.
29. Xavier, J.B. and K.R. Foster, *Cooperation and conflict in microbial biofilms*. Proc Natl Acad Sci U S A, 2007. **104**(3): p. 876-81.
30. DAVEY, M.E. and G.A. O'TOOLE, *Microbial Biofilms: from Ecology to Molecular Genetics*. MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, 2000. **64**(4): p. 20.
31. Kostakioti, M., M. Hadjifrangiskou, and S.J. Hultgren, *Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era*. Cold Spring Harb Perspect Med, 2013. **3**(4): p. a010306.
32. Winstanley, C., S. O'Brien, and M.A. Brockhurst, *Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections*. Trends Microbiol, 2016. **24**(5): p. 327-337.

33. Fauzia, K.A., et al., *Biofilm Formation and Antibiotic Resistance Phenotype of Helicobacter pylori Clinical Isolates*. Toxins (Basel), 2020. **12**(8).
34. Hall, C.W. and T.F. Mah, *Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria*. FEMS Microbiol Rev, 2017. **41**(3): p. 276-301.
35. Ahmed, M.N., et al., *The evolutionary trajectories of P. aeruginosa in biofilm and planktonic growth modes exposed to ciprofloxacin: beyond selection of antibiotic resistance*. NPJ Biofilms Microbiomes, 2020. **6**(1): p. 28.
36. Duan, X., et al., *rpoS-mutation variants are selected in Pseudomonas aeruginosa biofilms under imipenem pressure*. Cell Biosci, 2021. **11**(1): p. 138.
37. Stewart, P.S. and M.J. Franklin, *Physiological heterogeneity in biofilms*. Nat Rev Microbiol, 2008. **6**(3): p. 199-210.
38. Sanchez, C., et al., *Biofilm formation by clinical isolates and the implications in chronic infections*. BMC Infectious Diseases, 2013. **13**(47).
39. Khatoon, Z., et al., *Bacterial biofilm formation on implantable devices and approaches to its treatment and prevention*. Heliyon, 2018. **4**(12): p. e01067.
40. Yadav, M.K., J.E. Vidal, and J.-J. Song, *Microbial biofilms on medical indwelling devices*, in *New and Future Developments in Microbial Biotechnology and Bioengineering: Microbial Biofilms*. 2020. p. 15-28.
41. Vestby, L.K., et al., *Bacterial Biofilm and its Role in the Pathogenesis of Disease*. Antibiotics (Basel), 2020. **9**(2).
42. Zhao, A., J. Sun, and Y. Liu, *Understanding bacterial biofilms: From definition to treatment strategies*. Front Cell Infect Microbiol, 2023. **13**: p. 1137947.
43. Hall-Stoodley, L., et al., *Towards diagnostic guidelines for biofilm-associated infections*. FEMS Immunol Med Microbiol, 2012. **65**(2): p. 127-45.
44. Silva, V., et al., *Biofilm Formation of Multidrug-Resistant MRSA Strains Isolated from Different Types of Human Infections*. Pathogens, 2021. **10**(8).
45. Armbruster, C.R., et al., *Adaptation and genomic erosion in fragmented Pseudomonas aeruginosa populations in the sinuses of people with cystic fibrosis*. Cell Rep, 2021. **37**(3): p. 109829.
46. Sharma, D., L. Misba, and A.U. Khan, *Antibiotics versus biofilm: an emerging battleground in microbial communities*. Antimicrob Resist Infect Control, 2019. **8**: p. 76.
47. Uruen, C., et al., *Biofilms as Promoters of Bacterial Antibiotic Resistance and Tolerance*. Antibiotics (Basel), 2020. **10**(1).
48. Fisher, R.A., B. Gollan, and S. Helaine, *Persistent bacterial infections and persister cells*. Nat Rev Microbiol, 2017. **15**(8): p. 453-464.
49. Wood, T.K., S.J. Knabel, and B.W. Kwan, *Bacterial persister cell formation and dormancy*. Appl Environ Microbiol, 2013. **79**(23): p. 7116-21.

50. Yao, Y., X.H. Xu, and L. Jin, *Macrophage Polarization in Physiological and Pathological Pregnancy*. Front Immunol, 2019. **10**: p. 792.
51. Mirzaei, R., et al., *Immunometabolism in biofilm infection: lessons from cancer*. Mol Med, 2022. **28**(1): p. 10.
52. Maslova, E., et al., *Burns and biofilms: priority pathogens and in vivo models*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 73.
53. Peri, A.M., et al., *New Microbiological Techniques for the Diagnosis of Bacterial Infections and Sepsis in ICU Including Point of Care*. Curr Infect Dis Rep, 2021. **23**(8): p. 12.
54. Weiser, J.N., D.M. Ferreira, and J.C. Paton, *Streptococcus pneumoniae: transmission, colonization and invasion*. Nat Rev Microbiol, 2018. **16**(6): p. 355-367.
55. Henriques-Normark, B. and E.I. Tuomanen, *The pneumococcus: epidemiology, microbiology, and pathogenesis*. Cold Spring Harb Perspect Med, 2013. **3**(7).
56. Schilder, A.G., et al., *Otitis media*. Nat Rev Dis Primers, 2016. **2**(1): p. 16063.
57. Paton, J.C. and C. Trappetti, *Streptococcus pneumoniae Capsular Polysaccharide*. Microbiol Spectr, 2019. **7**(2).
58. Engholm, D.H., et al., *A visual review of the human pathogen Streptococcus pneumoniae*. FEMS Microbiol Rev, 2017. **41**(6): p. 854-879.
59. Reid, S.D., et al., *Streptococcus pneumoniae forms surface-attached communities in the middle ear of experimentally infected chinchillas*. J Infect Dis, 2009. **199**(6): p. 786-94.
60. Cillóniz, C., et al., *Antimicrobial Resistance Among Streptococcus pneumoniae*, in *Antimicrobial Resistance in the 21st Century*. 2018. p. 13-38.
61. Cilloniz, C., et al., *The Effect of Macrolide Resistance on the Presentation and Outcome of Patients Hospitalized for Streptococcus pneumoniae Pneumonia*. Am J Respir Crit Care Med, 2015. **191**(11): p. 1265-72.
62. Domenech, M., E. Garcia, and M. Moscoso, *Biofilm formation in Streptococcus pneumoniae*. Microb Biotechnol, 2012. **5**(4): p. 455-65.
63. Gilley, R.P. and C.J. Orihuela, *Pneumococci in biofilms are non-invasive: implications on nasopharyngeal colonization*. Front Cell Infect Microbiol, 2014. **4**: p. 163.
64. C Feldman, et al., *The interaction of Streptococcus pneumoniae with intact human respiratory mucosa in vitro*. European Respiratory Journal, 1992. **5**: p. 8.
65. Shenoy, A.T., et al., *Streptococcus pneumoniae in the heart subvert the host response through biofilm-mediated resident macrophage killing*. PLoS Pathog, 2017. **13**(8): p. e1006582.
66. Jourdain, S., et al., *Differences in nasopharyngeal bacterial carriage in preschool children from different socio-economic origins*. Clin Microbiol Infect, 2011. **17**(6): p. 907-14.

67. Allegrucci, M., et al., *Phenotypic characterization of Streptococcus pneumoniae biofilm development*. J Bacteriol, 2006. **188**(7): p. 2325-35.
68. Allegrucci, M. and K. Sauer, *Characterization of colony morphology variants isolated from Streptococcus pneumoniae biofilms*. J Bacteriol, 2007. **189**(5): p. 2030-8.
69. Coutinho, C.P., et al., *Long-term colonization of the cystic fibrosis lung by Burkholderia cepacia complex bacteria: epidemiology, clonal variation, and genome-wide expression alterations*. Front Cell Infect Microbiol, 2011. **1**: p. 12.
70. Blanchard, A.C. and V.J. Waters, *Microbiology of Cystic Fibrosis Airway Disease*. Semin Respir Crit Care Med, 2019. **40**(6): p. 727-736.
71. Chao, Y., et al., *The serine protease HtrA plays a key role in heat-induced dispersal of pneumococcal biofilms*. Sci Rep, 2020. **10**(1): p. 22455.

Chapter 2

Development of a long-term robust and reproducible in vitro assay for *Streptococcus pneumoniae* biofilm growth

The contents of this chapter were adapted from the first part of the following manuscript:

Espinoza-Miranda S, Schmiede E, Chu C, Chowdhury A, Rosconi F, van Opijnen T. Genome-wide identification of genetic requirements of *Streptococcus pneumoniae* biofilms. Manuscript Preparation.

Author contributions:

Suyen Espinoza-Miranda (SE), Tim van Opijnen (TvO) conceived and designed the study. SE, Emma Schmiede (ES), Cedrick Chu (CC) and Allison Chowdhury (AC) performed the experiments. Federico Rosconi (FR) and SE wrote the manuscript. Tim van Opijnen (TvO) edited and approved the manuscript.

2.1 The incessant quest for the long-term in vitro biofilm assay.

Although bacteria are most often studied as planktonic or free-living organisms, they frequently grow in complex surface-attached communities known as biofilms [1, 2]. Importantly, biofilms have been associated with the development of antibiotic resistance, increased virulence, and immune escape,[3] which highlights why an integrated understanding of how biofilms contribute to such events is critical. A microbial biofilm lifestyle follows a lifecycle that can be defined by at least three stages: 1. bacterial aggregation and attachment to a specific surface, 2. expansion/growth (biofilm maturation), and 3. dispersion (Figure 1) [4, 5]. During the expansion and growth stage, bacteria continuously release DNA, proteins and lipids which then forms the extracellular polymeric matrix (EPS). The matrix provides stability and integrity to the biofilm structure, as well as protection from the environment. While the life cycle for developing biofilms follows the same stages, biofilms vary in matrix composition, structure, and function, depending on the bacterial species [6].

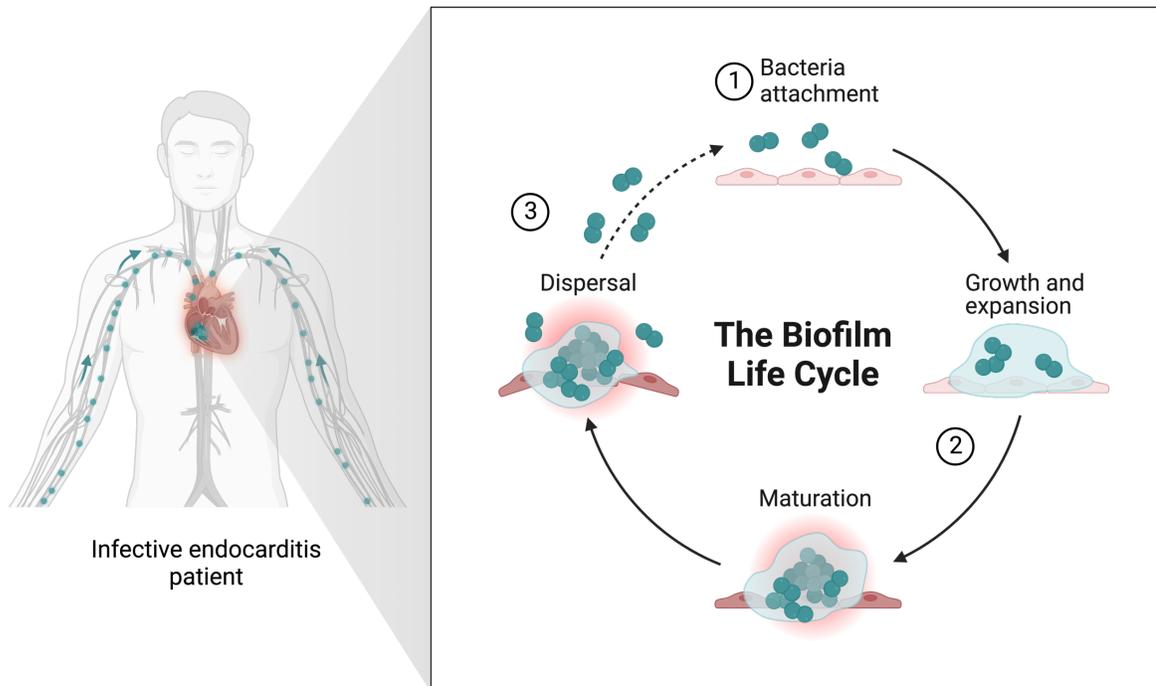


Figure 1. The biofilm life cycle in the context of pneumococcal endocarditis. 1. Bacteria attach to a surface, followed by 2. Maturation: referred to the continuous growth and expansion of a biofilm

and 3. Dispersal: cells detach once the biofilm reaches a maturation state and disperse to look for an empty surface to colonize.

Most biofilm research has focused on the attachment and maturation stage. Little is known about biofilm maintenance, defined as the progression from biofilm maturation to dispersion, the process by which microcolonies formed from the attached bacteria start expanding and growing to finalized in detaching from the biofilm. Biofilm maintenance is needed for the biofilm to continue its growth cycle [7, 8]. The importance of maintenance has often been overlooked as normally biofilm studies investigate growth in isolated stages rather than a full system [9]. The lack of standardized and reliable *in vitro* assays that mimic all the stages of biofilm growth is one of the main challenges in the study of biofilm communities as an integrated system [10]. The importance of a long-term *in vitro* assay lies in the fact that biofilm growth is a complex system that requires the coordination of multiple stages, and each stage relies on the previous one. Therefore, it is more effective to study the growth of the biofilm as a whole, rather than analyzing each stage in isolation.

Multiple methods have been developed to study biofilms, each comes with different advantages and limitations in terms of reproducibility, scalability, applicability, and validity [11]. All these methods are derived from two main approaches: 1. Static, which refers to methods that are characterized by having a limited supply of nutrients. For example, “the microtiter plate assay” is the most widely used approach and is defined by bacterial growth at the bottom of microtiter plates. This method is inexpensive, easy to implement, and allows high-throughput screening under multiple conditions when using 96 well plates [12]. However, this method fails to capture the biofilm architecture and dynamics that occur in this mode of growth during the maintenance and dispersal stage. In addition, this approach is extremely variable from well to well, leading to poor reproducibility [13]. 2. Dynamic, which are methods such as flow cell reactors that can provide continuous nutrient flow and enable real-time observations [14]. However, these approaches are in general expensive, low throughput, and are prone to contamination [10]. While development of an *in vitro* assay that can capture the individual biofilm stages as an

integrated system seems challenging, it is crucial to understand how biofilms contribute to the development of drug resistance, virulence and immune escape.

Streptococcus pneumoniae (*S. pneumoniae*) is a human respiratory pathogen capable of causing a range of serious diseases, including community-acquired pneumonia, meningitis, and sepsis [15]. The pathogen is known to form biofilms in the nose, the ears, the lungs and the heart [16]. Various approaches derived from static model systems have been used to study *S. pneumoniae* biofilm formation. Examples of some variation from static models include the use of 96 and 24 well plates, and falcon tubes, which depend on the attachment of bacteria to the bottom of the plate or tube [16]. These static approaches enable greater control when studying biofilms under different changing as it facilitates the manipulations of variation in nutrients, pH changes, and temperature [16, 17]. Additionally, methods used so far have been unable to accurately replicate the extended duration of bacterial incubation that occurs during infection. For instance, the longest an assay has been able to culture *S. pneumoniae* biofilms is 72 hours, while the average carriage time of *S. pneumoniae* in children varies between 13 and 65 days [18]. To further illustrate current methods, Table 1 summarizes methods used for cultivating *S. pneumoniae* biofilms, duration, major findings, and quantification approaches.

Table 1. Methods used for growing and quantifying *S. pneumoniae* biofilms.

Assay	<i>S. pneumoniae</i> strain used	Length of the assay	Quantification approach	Findings	Reference
96-well (flat bottom) polystyrene	R6	6 hours	Crystal violet	LytA amidase, LytC lysozyme, LytB glucosaminidase, CbpA adhesin, PcpA putative adhesin, and PspA (pneumococcal surface protein A) mutants had a decreased capacity to form biofilms,	[19]
microtiter plates	T4, 6A10, 6B8, 6A16, 6A18, 6B21	18 hours	Crystal violet, microscopy	Biofilm formation is strain dependent	[20]
96-well (flat-bottom) polystyrene	3 (AR 33118, FL 2812 and FL 5629), 2 serotype 6B (MJD 1225 and MJD	16 and 24 hours	Crystal violet	Human serum albumin (25000 microg/mL) reduced biofilm formation in 7 strains	[21]

	3032), 2 serotype 9V (AR 06016 and AR 09164), and 3 serotype 23F (FJD 60, AR 30118 and MJD 573)				
Six well plates	22F (ST433)	3 days	Microscopy	Enhanced biofilm development link to mutations in RpoE from 12 distinct SCV phenotypes	[22]
6-well polystyrene plates , 24 well plates	6A10, D39	48 hours	Crystal violet, Carbohydrate Detection Kit (BioVision)	Neuraminidase A and B-galactosidase are required for biofilm growth	[23]
96 well microtiter plates	30 clinical isolates	16 hours	Crystal violet and tetrazolium dyes	All nasal swabs were able to form some degree of biofilm	[24]
96 well microtiter plates	R6, D39	6 hours	Crystal violet	yefM-yoeB and relBE Toxin-Antitoxin Operons Participate in Oxidative Stress and Biofilm Formation	[25]
6-well tissue culture plates	BS68 (serotype 9V); BS69 (serotype 14); BS71 (serotype 3); BS72 (serotype 23F); BS73 (serotype 6A); and BS75 (serotype 19F)	1-3 days	Crystal violet, CFU, microscopy	Evidence of capsule downregulation in <i>S. pneumoniae</i> biofilms	[26]

For *S. pneumoniae* an approach that captures each stage of a biofilm and can study it as an integrated system is largely unavailable. Instead, early biofilms have been mostly studied, which have revealed key genes for the attachment stage. Knowledge on the later stages of *S. pneumoniae* biofilm growth are limited since in later stages of cultivation, pneumococci activate a hydrolytic mechanism that leads to the demise of noncompetent cells, which is a typical feature of *S. pneumoniae* regardless of type of growth [16]. Although Hall-Stoodley and others have been able to grow *S. pneumoniae* biofilms for up to 72 hours [26], their protocol is not suitable for longer incubation periods, which is likely, at least partially, due to the use of rich media, which decreases the yield of bacterial growth for longer periods of time. In addition, they fail to capture information from the dispersal stage (Table 1).

This chapter focuses on the development of a long-term *in vitro* assay for *S. pneumoniae* biofilm formation that captures each stage of the lifecycle and enables analysis of the biofilm as an integrated system. By testing multiple growth media and enabling biofilm dispersion we find that the *S. pneumoniae* biofilm lifecycle can be successfully maintained for 40 days. Moreover, by coupling this approach with confocal microscopy detailed characterization and quantification of biofilm structures are enabled across the entire lifecycle. Importantly, this approach is reproducible and high-throughput, and allows for the integrated study of all stages of biofilm establishment, growth, maintenance and dispersion.

2.2 Results

Development of a long-term *in vitro* assay for *Streptococcus pneumoniae* biofilm growth.

The process of sustaining and stabilizing growth to maintain a biofilm is continuous and influences the ability of a biofilm to survive in nature [9]. The absence of an assay for studying long-term *S. pneumoniae* biofilms has limited research to early life cycle stages, neglecting key survival aspects such as biofilm maintenance and dispersion. To address this limitation, an approach is needed that not only allows for longer growth of *S. pneumoniae* biofilms, but that also allows for studying the biofilm as an integrated system.

Multiple *in vitro* biofilm assays use microtiter plates, which enable bacteria to attach to the bottom of a well and form a biofilm. While these approaches constrain the ability of a biofilm to continue growing by the surface area of the well, by how many bacteria can divide inside the well and do not include dispersion, they are easy to manage and upscale. Because the advantages of a static system are large and the limitations might be overcome, we opted to establish a long-term biofilm growth assay based on a static *in vitro* approach. To achieve this glass coverslips were placed inside 24-well plates to serve as a surface for cell attachment. Similar to the biofilm life cycle, the experiment was initiated with the addition of planktonic bacteria to each well containing coverslips, which subsequently would attach,

grow, and divide to form biofilms on the coverslips (**Figure 2**). Coverslips are used for two reasons: 1. It allows for easy imaging by confocal microscopy which consequently enables microbial community quantification and 2. It enables the passage of bacteria attached to a transferable surface, thus enhancing the potential for long-term cultivation of biofilms.

The use of coverslips as a means to transfer biofilms is inspired by the bead transfer model which was developed specifically for studying biofilms in the context of experimental evolution [27]. In this approach, a bead is placed in a test tube with actively growing bacteria, bacteria attach to the bead, and after a certain amount of time, the bead-containing bacteria (old bead) is transferred to a tube with clean beads containing fresh media. In the new tube, bacteria detaching from the old bead can find a new surface to attach to the new bead allowing for the biofilm lifecycle to continue. This process can basically be repeated indefinitely enabling long-term biofilm cultivation. This method has been extensively used for species such as *P. aeruginosa*, *B. cenocepacia* and *A. baumannii* [1, 27]. Before proceeding to test whether *S. pneumoniae* biofilms can be efficiently transferred from one coverslip to another the right growth medium had to be found that would help promote biofilm formation without decreasing bacterial yield.

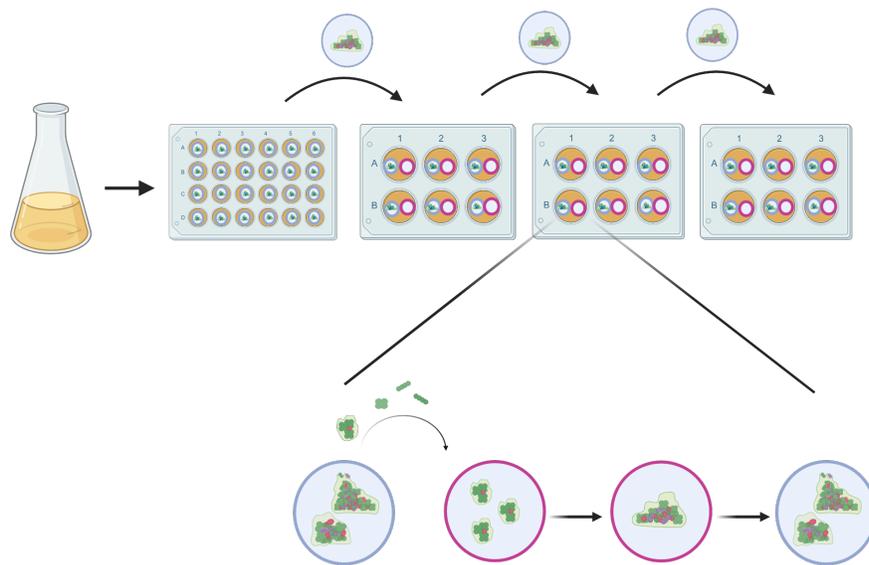


Figure 2. Development of a robust and consistent long-term in vitro biofilm assay for *S. pneumoniae*. Planktonic bacteria were harvested and diluted in 24-well plates containing clear round coverslips (light blue). The medium was changed every day to promote biofilm formation. After the first biofilm cycle (four days), the biofilm-containing cover slip is transferred to a new plate and paired with a new cover slip (pink), which allows for a new biofilm to be established on the clean coverslip. This process can then be repeated indefinitely.

Media composition optimization to support robust biofilms.

The choice of media can influence biofilm behavior and outcome[6] and for *in vitro S. pneumoniae* biofilms, different types of media and carbon sources have been utilized. In general, rich media seem to be preferred in biofilm studies [28] , likely because it is assumed that an abundance of nutrients would sustain a biofilm longer. However, some studies have used more minimal, chemically defined media [16, 29]. To cast a wide net three different media types were evaluated that include: 1. Todd Hewitt (THY) [30], a rich medium. 2. Semi-defined minimal medium (SDMM), a medium-rich medium [31], and 3. Chemically Defined Medium (CDM) [32] a minimal medium. To promote biofilm maintenance, media was changed every day. After every day bacteria were scraped from the coverslips, centrifuged, washed with 1XPBS and serially diluted and plated in blood agar (BA) plates to count number of viable cells within the biofilm. To ensure dissociation of cell aggregates, samples were separated we used vortex for 30 seconds at max speed.

To initially quantify the number of viable bacteria attached to coverslips over five days Colony Forming Units (CFUs) were enumerated from blood agar plates. While all three types of media can be used to form biofilms clear differences among the media can be observed: 1) The rich medium, THY, had an approximate 1000-fold drop in CFU between days 2 and 3 and never really recovered, indicating that this medium affects the yield of the recovered or even dispersed bacteria overtime. 2) CFU recovered from SDMM, the medium-rich medium, showed that the number of bacteria stayed relatively constant over the 5 days, with only small fluctuations. 3) The minimal medium, CDM, was not able to support the formation of a robust biofilm, indicated by the low number of CFUs recovered across the five days of growth (Table 2). Overall, there was a decreased in the number of viable cells at day 5. Therefore day 4 was selected to perform a transfer step.

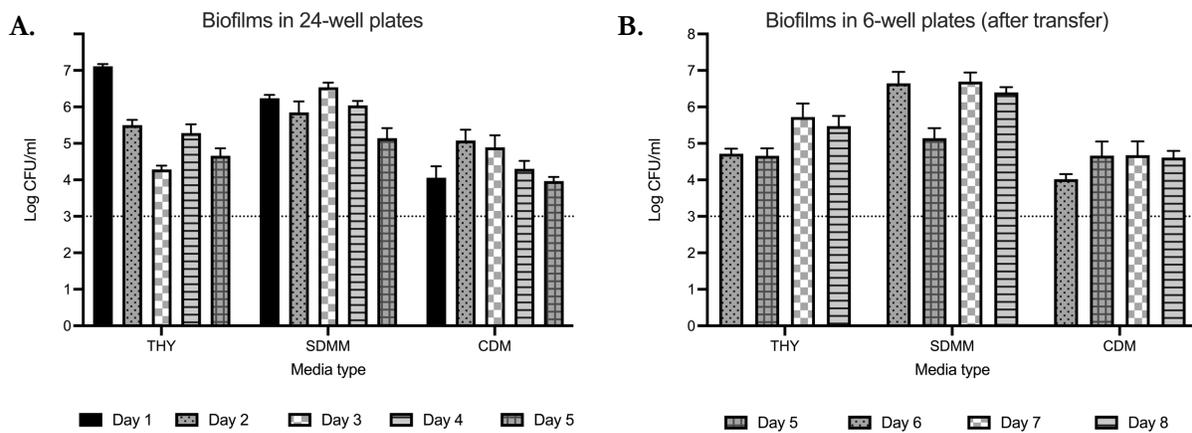


Figure 3. SDMM shows a higher cell viability recovery in biofilm during eight days of growth. Biofilms were grown on coverslips and media was change daily to promote growth. Daily samples were taken to calculate the number of viable cells withing the biofilm. Samples were scraped from the coverslips, centrifuged, wash and resuspended in 1XPBS followed by serial dilution and plating. **A.** Average of three independent experiments of cell viability quantification using different media during five days of biofilm grow (24-well plates). **B.** Average of three independent experiments of cell viability quantification using different media after transferring step (six-well plates).

In the process of selecting an appropriate growth medium for subsequent experiments, mature biofilms, aged four days, were carefully transferred to a fresh six-well plate and paired with clean coverslips. This approach was used to promote biofilm maintenance. The choice of four-day-old

biofilms was deliberate, aimed at preserving the bacterial population and preventing further reduction. Subsequently, CFU were determined for biofilms ranging from day 5 to day 8, cultivated in the new six-well plate. The obtained data indicates that the use of SDMM fosters a more consistent growth pattern for *Streptococcus pneumoniae* biofilms, as evidenced by the uniform CFU recovery rates illustrated in **Figure 3**.

Fluorescent microscopy visually confirms SDMM is the best choice of growth medium.

To further confirm suitability of SDMM as an optimal medium for nurturing robust and stable biofilm formations we visualize explored the ability of bacteria to grow as biofilms on the coverslips. Biofilms were stained using a live/dead cell viability kit (Invitrogen). This kit contains two dyes: 1. Syto 9, which stains DNA green of live cells with an intact cell wall, and 2. propidium iodide, which diffuses through a permeable membrane of dead cells and stains cells red. Stained biofilms were visualized under a fluorescent microscope (Olympus-IX83). **Figure 4** shows representative microscopy images of *S. pneumoniae* TIGR4 biofilms on day four (panels A, C, E) and day eight (which we call to the new re-established biofilms after the passage step) (panels B, D, F) when grown using three different media. Visualization was done on days 4 and 8, which represent the most mature biofilms that are grown in this assay, and which are at the time point right before the next 'biofilm passage' (See Figure 2). Figure 3 shows that each of the three media supported attachment of bacteria to glass coverslips, as indicated by the green-stained bacteria. After 4 and 8 days of growth, SDMM seemed to support the most dense biofilms, while CDM supported the lowest density biofilms, which confirms the number of CFUs recovered in Table 2 and 3. Importantly, this shows that the coverslip method enables the clear visualization of a biofilm, and possibly microscopy quantification, which we further explore below.

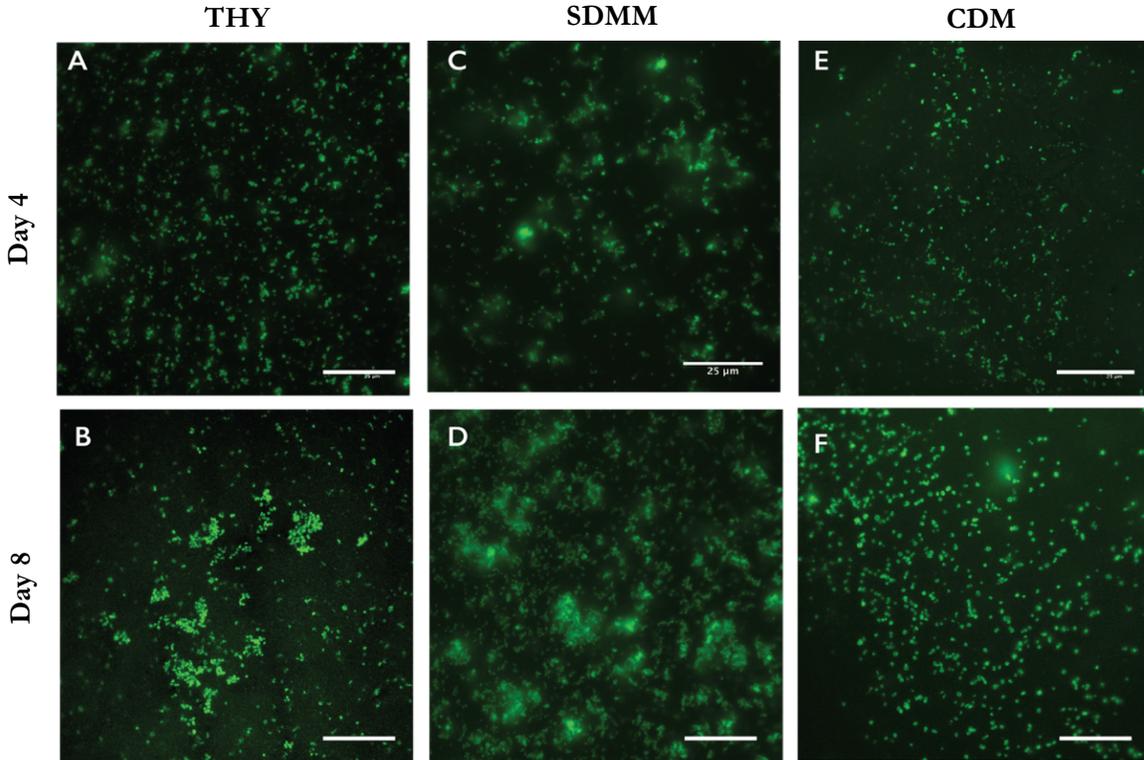


Figure 4. *S. pneumoniae* biofilm formation with different media. **A** and **B** grown with THY (rich medium). **C** and **D** with SDMM (semi-defined minimal medium). **E** and **F** grown with CDM (chemically defined medium). Top panels are biofilms after 4 days of growth and bottom panels after 8 days of growth. Biofilms are stained with BacLight live/dead staining and visualized with an Olympus IX83 fluorescence microscope using CellSense software. Magnification is at 60X, and the scale bar is 25µm for all the images.

Glucose concentration affects biofilm formation.

Bacteria in biofilms require a source of carbon for energy, growth and production of the EPS [33]. The impact of the carbon source concentration on biofilm formation varies among species. For example, *S. aureus* is more efficient at forming biofilms with higher glucose concentrations which corresponds to 1% more of glucose when compared to normal media conditions [34], whereas *Salmonella enterica* prefers lower glucose concentrations (0.25% of the total glucose normally added to the media) and its growth is inhibited at higher levels (1% more of glucose added to media) [35]. Depending on the

availability in the host environment, *S. pneumoniae* can make use of a variety of carbon sources, including glucose, galactose, fructose, sucrose, and lactose [36]. Nevertheless, glucose is the preferred carbon source, as it boosts the production of capsule precursor, and the presence of capsule facilitates adherence to surfaces which is a crucial step for the early stages of biofilm formation [36, 37].

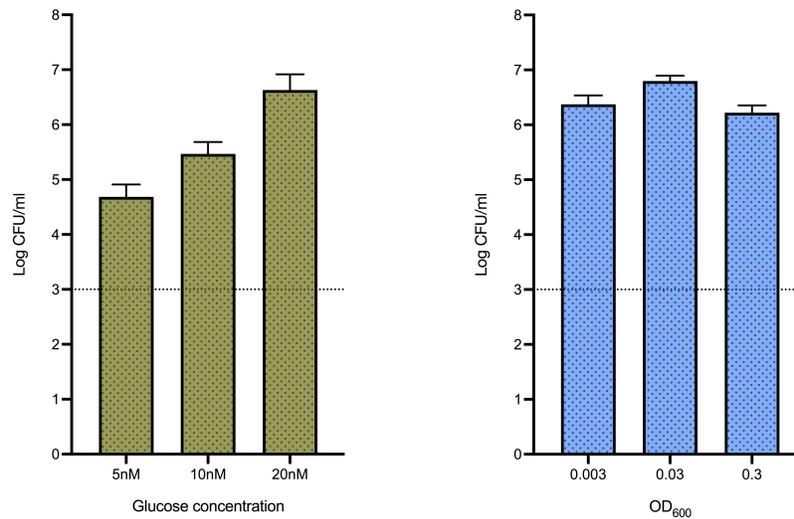


Figure 5. Effect of different glucose and initial bacterial optical density (OD₆₀₀) concentration on biofilm formation. Biofilms were grown using SDMM, media was changed every day and CFU was recorded after four days of growth. **A.** Three different glucose concentrations 5mM, 10mM and 20mM were tested. **B.** Three different initial OD₆₀₀ concentration were tested, revealing no significant different among them after 24 hours of growth.

Since SDMM with glucose as the carbon source provided the most robust and consistently sustained biofilms, three different glucose concentrations 5mM, 10mM, and 20mM (low, medium and high respectively) were explored to determine their effect on the biofilm lifecycle. Biofilms were visualized via fluorescence microscopy and after 4 days of biofilm growth 10mM and 20mM glucose led to a higher number of bacteria attached to the coverslip compared to 5mM (**Figure 5 and 6**). Moreover, cell counts indicate that 20mM leads to ~100x more viable cells in a coverslip-biofilm compared to 10mM glucose. Additionally, the size of the biofilm starting inoculum was evaluated by exploring three different starting optical densities OD₆₀₀: 0.3, 0.03, and 0.003, which did not lead to significant variation

in the number of viable cells adhered to the coverslip after 24 h, indicating that the initial inoculum size does not impact biofilm formation.

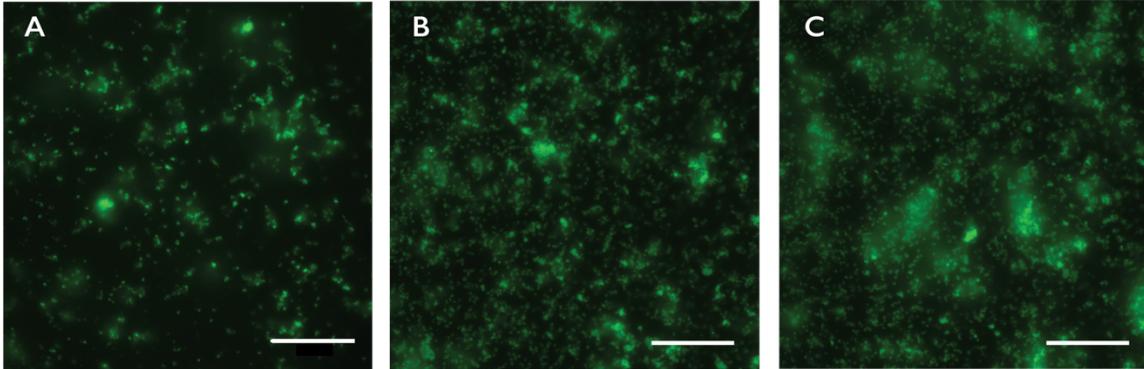


Figure 6. Glucose concentration affects *S. pneumoniae* biofilm development. Biofilms were grown using SDMM and three different glucose concentrations 5mM (A), 10mM (B) and 20mM (C). Biofilms were stained with BacLight live/dead stain and visualized using an Olympus IX83 fluorescence microscope using Cell Sense software. Magnification is at 60X, and scale bar is 25 μ m for all images.

Transferring the biofilms is necessary to promote long-term biofilm growth.

The bead-based biofilm approach from which this coverslip approach has taken its inspiration practically enables indefinite biofilm propagation and experimental evolution. To explore whether biofilms can be propagated for extended periods of time, and potentially indefinitely with the coverslip approach, a 4-day biofilm- containing coverslip (old – blue edge coverslip in **Figure 2**) was placed into a six well plate next to a clean coverslip (represented by the pink edge coverslip). We hypothesized that bacteria detaching from the old coverslip and biofilm would form a new biofilm on the new coverslip. Indeed, this approach enables biofilms to go through a complete lifecycle, including dispersion and reestablishment on a new surface. Transfer was repeated for four cycles, leading to 16 days of culture, which lead to a consistent and robust biofilm after every transfer (**Figure 7**). The coverslip approach

thus allows for the long-term maintenance of biofilm communities, which is the first report of a biofilm assay that enables long-term analysis of *S. pneumoniae* biofilms.

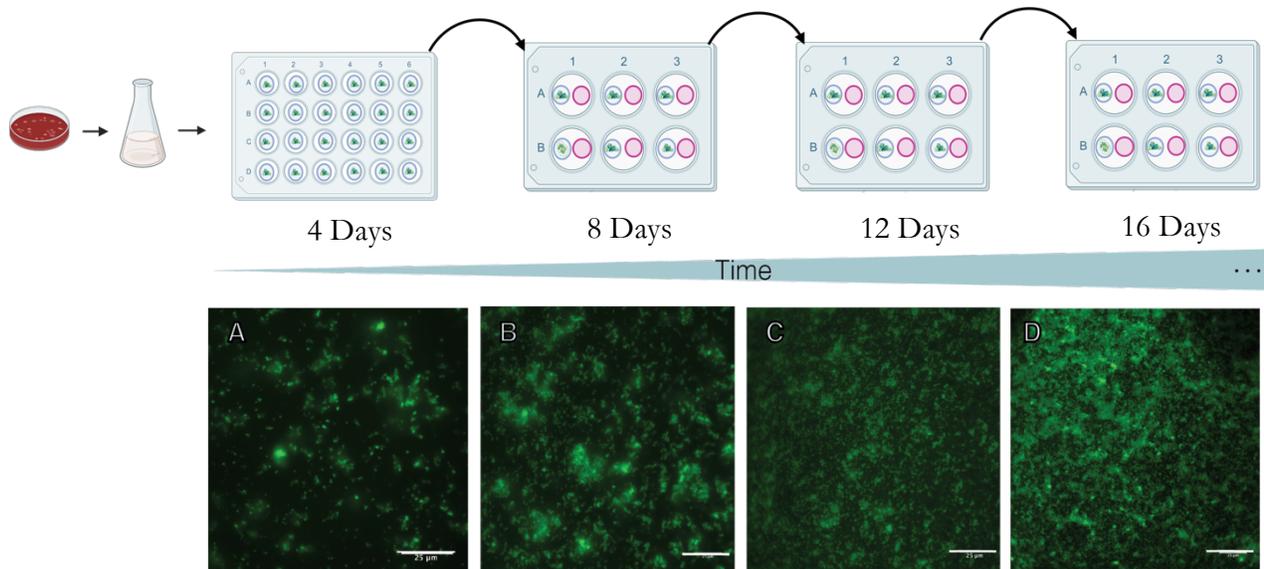


Figure 7. Biofilm-growth assay can be carried out for weeks allowing to perform long-term experimental adaptation in biofilms. Fluorescent images of four biofilm cycles are shown: A) 4, B) 8, C) 12 and D) 16 days. All scale bars are 25µm. Each biofilm was stained with BacLight and visualized using an Olympus IX83 fluorescence microscope using Cell Sense software. Magnification at 60X.

Use of Confocal Microscopy to study biofilm population dynamics.

To further understand how biofilms are developing overtime it is important to accurately quantify specific characteristics of a biofilm. The most common quantification method is crystal violet staining, which binds cellular components by ionic interactions [38]. A stained biofilm is dissolved in a solvent, and the absorbance of the solution can be measured by a spectrophotometer. More biofilm leads to more staining and increased absorbance. While this is a simple, low-cost, and high throughput method, it has multiple limitations, including the loss of biofilm during the washing and staining steps, especially for aerobic biofilm or early phase biofilm [12, 39]. Additionally, this method does not truly quantify biomass because it does not distinguish between live and dead cells, and more importantly, it does not

provide information on the spatial distribution, structure, or composition of the biofilm [40]. Alternatively, quantification through confocal microscopy is a powerful tool as it can provide information on the spatial distribution, structure, and composition of the biofilm[41], it can distinguish between living and dead cells and can visualize the biofilm three-dimensional architecture [42, 43]. Moreover, microscopy can be used to study the dynamics of formation and monitor the effects of antimicrobial agents [44].

To quantify biofilm features, confocal laser scanning microscopy (CSLM) images were analyzed with COMSTAT, a program for quantifying three-dimensional biofilm structures [45]. COMSTAT processes stacks of CSLM images and generates features for quantifying biofilm, including biomass, average thickness, roughness coefficient, and maximum thickness [46]. Biomass [$\mu\text{m}^3/\mu\text{m}^2$] is the amount of biologic material present in each area. Average thickness is the most common metric used in biofilms and provides a sense of the spatial organization at each step of the Z-stacks. Maximum thickness is the highest point of the biofilm at a given position. Roughness Coefficient (RA^*) is an indicator of biofilm heterogeneity and provides a measure of how much the thickness of the biofilm varies, e.g., the smaller the RA^* value, the higher uniformity among aggregates (**Figure 8**).

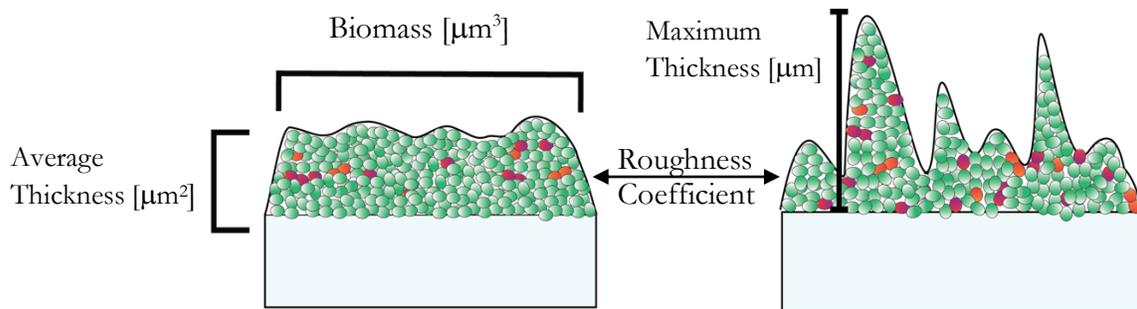
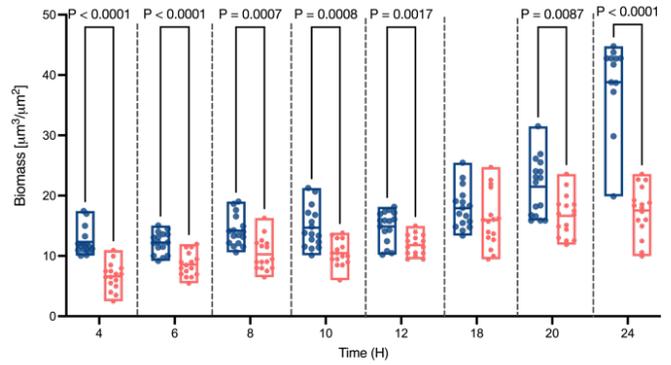
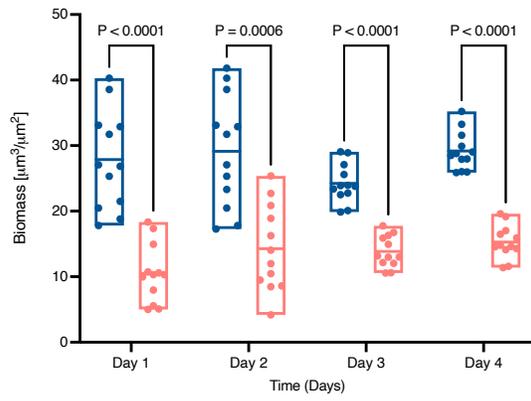


Figure 8. Quantification parameters used for biofilm characterization. Biomass is the 3D feature, average thickness examines the biofilm in 2D, maximum thickness measures the highest biofilm point in the Y-axis (1D). Roughness coefficient evaluates the overall composition; biofilm on the left depicts a smooth biofilm and, on the right, displays a rough biofilm.

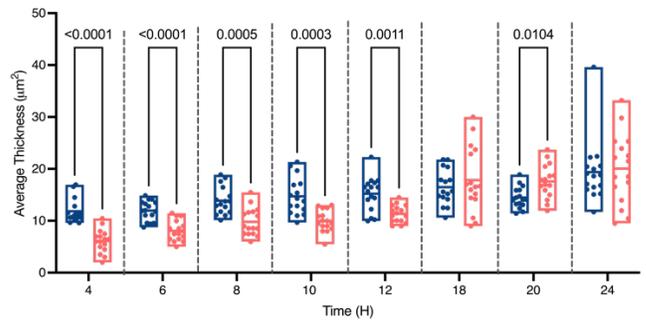
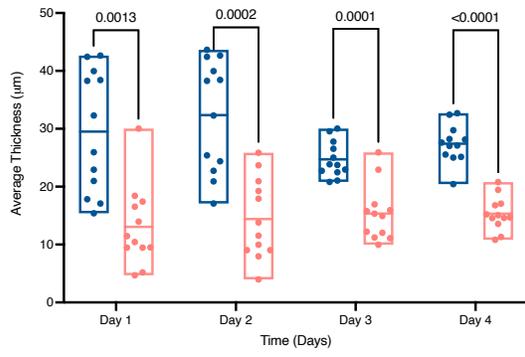
To explore the accuracy of the assay, two *S. pneumoniae* clinical isolates BS71 (serotype 3) and BS72 (serotype 23F), previously reported by Hall-Stoodley as low and high biofilm strains respectively, were tested. In this work authors defined a high biofilm strain as one with a biomass $>40\mu\text{m}^3$, and low with a biomass $<20\mu\text{m}^3$ [26]. Using the growth and quantification methods developed in this thesis, BS72 shows significantly more biomass across four days of biofilm growth than BS71, indicating that BS72 makes more abundant cell clusters (**Figure 9A-D**). Less variability among replicates for each parameter can be observed at day 3. Additionally, the average and maximum thickness of BS72, the high biofilm former, is higher than BS71, the low biofilm former. The high roughness coefficient [RA*] values suggest a larger variability in microcolony height for BS71, and a more uniform growth for BS72 (**Figure 9D**). While BS71 RA decreases over time, the biofilm roughness remains more variable compared to BS72. Taking a closer look at the roughness coefficient for BS71, it seems that the variation ranges from 0.4-0.6 and is constant across the four days.

Finally, the average and maximum thickness for BS71 is less than BS72 across 4 days of growth. These results demonstrate that the longer bacteria are cultivated in a biofilm, the less variability there is among replicates. This result was reinforced after looking at the growth dynamics of the first 24 hours of growth, suggesting that studying biofilms in such a short period of time masks the important architectural and dynamic differences that occur during biofilm growth (Right panels **Figure 9**).

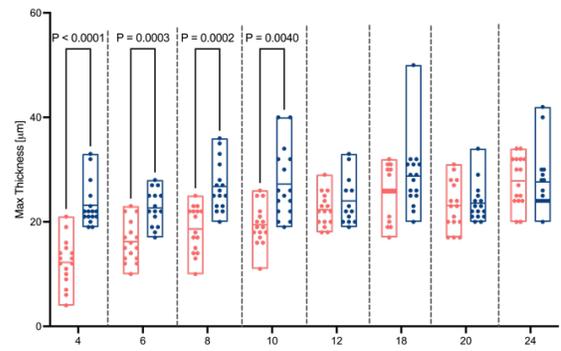
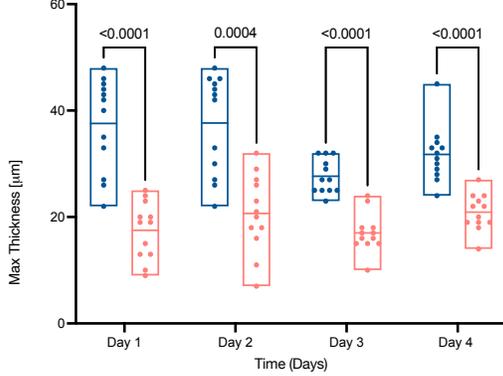
A.



B.



C.



D.

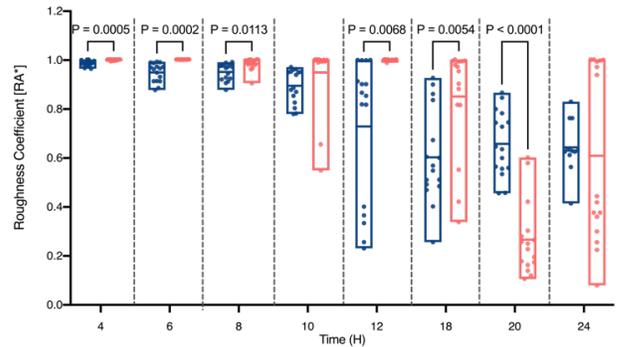
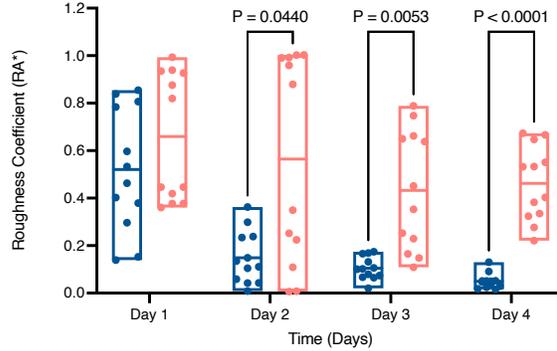


Figure 9. Previously characterized strains validate the assay developed in this thesis. Two previously reported strains **BS72: High biofilm strain** and **BS71: Low biofilm strain** were used to validate our assay. Biofilms were developed as previously described in figure1. Each biofilm was stained with BacLight. For each strain six Z-stacks from 4 biological replicates were acquired using 3i spinning disk confocal microscopy. **A.** Biomass, **B.** Maximum Thickness, **C.** Average Thickness and **D.** Roughness Coefficient. Fluorescence images were analyzed using COMSTAT followed by Two-way ANOVA followed Bonferroni post-test. Right panels show values for early.

In an attempt to clearly differentiate the strains from each other the dimensions of the dataset containing information on the four quantification parameters across four consecutive days of biofilm growth was reduced by performing a Principal Component Analysis (PCA). This analysis clearly separates the low and high-biofilm formers, with the first principal component (PC1) accounting for more than 60% of the variation in the dataset (**Figure 10A**), and the second component (PC2) explaining approximately 12% of the variation. For PC1, the main contributor of variance is biomass, while for PC2 it is roughness coefficient. These results suggest that changes in biofilm architecture can be illustrated mainly by these two parameters (**Figure 10B**).

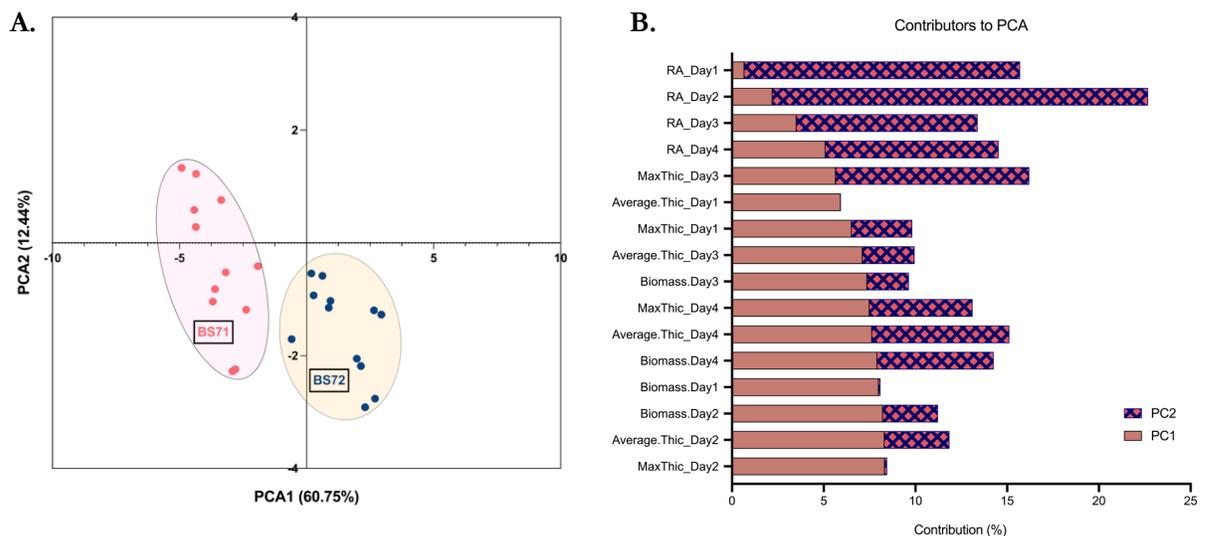


Figure 10. Two-dimensional principal component analysis separates low from high biofilm former strains. **A.** Biofilm former strains. Quantification data from all four parameters across four days were used to perform principal component analysis. **B.** Biomass and average thickness are the principal contributors to the first principal components, while Roughness coefficient is the primary

contributor to the second principal component. X- axis shows the percentage of contribution of variance of each variable for components 1 and 2.

In conclusion, using the developed assay, biofilm characterization of BS72 and BS71 confirms previously established biofilm differences with respect to high and low biomass. However, the new assay, along with the systematic biofilm quantification, reveals several additional differences in biofilm growth dynamics, increasing the resolution that can be obtained when charactering bacterial isolates.

Biofilm characterization of *S. pneumoniae* clinical isolates with different virulence

backgrounds.

S. pneumoniae has a high level of genomic diversity and adaptability [31]. This diversity is often investigated in the context of virulence and antibiotic resistance. Several research teams have studied biofilm growth in multiple *S. pneumoniae* strains, but only in the context of short in vitro experiments (12-24 hours). These studies concluded that all strains are capable of forming biofilms but to varying degrees. For example, Allegruci and others have showed that pneumococcal clinical isolates with different serotype background from the nose and ear form biofilms, but some are better than others. For instance, by measuring biomass using the CV staining method they showed that the isolates BS69 (serotype 14) and BS74 (serotype 18) form less biofilm compared to the clinical isolate BS71 (a serotype 3). Here, we investigated the capacity of seven *S. pneumoniae* strains from our laboratory pangenome collection, which includes strains with varying levels of virulence and different genomic backgrounds [31] (See materials and methods).

Figure 11 displays the biomass and roughness coefficient for each of the 7 strains tested on days 1 and 4 of biofilm formation. These two parameters were selected based on the PCA results, which indicate that they can mainly define biofilm formation differences between strains. As previously suggested all strains are able to produce biofilms. TIGR4, a virulent and common lab strain, produces lower and

rougher biofilms when compared to the 'high biofilm former' BS72. In contrast, the acapsular version (TIGR4-CPS) produces biofilms with greater biomass, but more roughness than BS72. This high biomass phenotype from the acapsular version is expected since non-encapsulated *S. pneumoniae* has been demonstrated to produce more robust biofilms, as strains deficient to produce capsule enhance cell attachment and transformation efficiency [19, 47]. Importantly, BHN97, a model strain for otitis media studies, forms even larger biofilms than BS72, which could explain its ability to cause ear infections. After 1 day of growth, there appears to be no difference in the number of bacteria attached to the coverslip for three of the tested strains (PG04, PG06 and PG16), but by day 4, distinct differences can be observed. PG04 is similar compared to BS72, while PG06 and PG16 exhibit lower biomass yields. Similarly, roughness coefficient for day 1 shows a significant difference for PG16, with a median lower RA number indicating a more even formation of the microcolonies in the biofilm. In contrast, the roughness coefficient on day 4 indicates that all strains are significantly different from BS72. Overall, this assay emphasizes the importance of examining biofilms for longer periods of times to assess a strain's true capacity to form biofilms and maintain the mode of growth. Importantly, the high biomass yield for BHN97 suggests there may be a link between otitis media and the ability of a strain to form high-biomass biofilms.

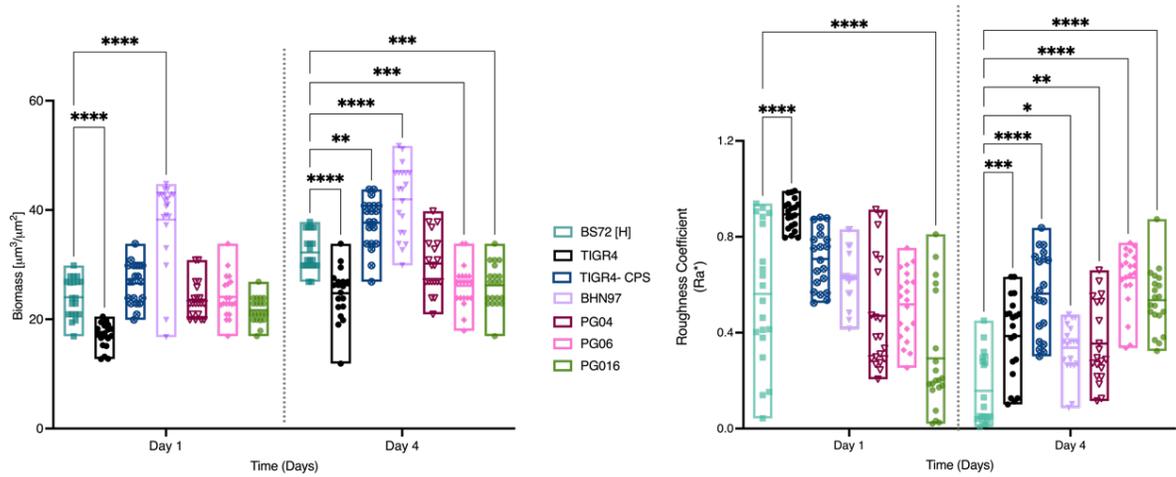


Figure 11. Biofilm characterization of seven *S. pneumoniae* clinical isolates. Biofilms for each strain were grown using the developed assay. Images from 12 biofilm replicates were taken for each strain using confocal microscopy followed by COMSTAT analysis. **A.** Biomass & **B.** Roughness Coefficient. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $p < 0.0001$ using Two-way ANOVA followed by Bonferroni post-test.

2.3 Conclusions

In nature, bacteria assemble as physical aggregates known as biofilms. The life cycle of a biofilm involves an initial attachment step followed by the expansion of attached microcolonies. Biofilm maintenance represents the crucial transition between the two last stages of the cycle: biofilm maturation and cell dispersion. Unfortunately, biofilm research has been hindered by the fragmented study of biofilm life cycle stages in isolation, rather than comprehensively understanding biofilm growth as a unified and continuous process. Currently, much of the knowledge about biofilm comes from studying the attachment and expansion stage, but there remains a significant gap in understanding biofilm maintenance and the underlying bacterial processes necessary for the survival of bacterial cells inside a biofilm.

Traditional techniques for cultivating *S. pneumoniae* biofilms in vitro typically involve polystyrene plates and are conducted for shorter periods of time of less than 24 hours [16, 48]. While these studies have shed light on some of the genetics in *S. pneumoniae* biofilms, little is known about the maintenance and cell dispersion stages. This gap in knowledge is caused by the lack of robust and long-term assays that allow the complete study of the life cycle of biofilms as an integrated system.

In this chapter, a reliable, reproducible, and long-term biofilm assay method was developed. Using glass coverslips as a surface for bacteria to attach to, a semi-defined minimal medium supplemented with glucose as the carbon source, and daily medium replenishment, biofilms were continuously grown for 96h. Moreover, by ‘passaging’ coverslips every four days biofilm growth was extended up to 16 days. While four passages and 16 days was the end-point of the experiment, coverslips can potentially be passaged indefinitely. The coverslip assay was successfully validated using two previously reported *S. pneumoniae* strains confirming their phenotypes as low and high biofilm formers. To successfully visualize and quantify bacteria growing in the biofilms, a live/dead cell viability staining was

implemented, coupled with microscopy and COMSTAT analysis. Parameters such as biomass, average and maximum thickness, along with roughness coefficient, characterize biofilm spatial growth dynamics.

In addition, this assay was used to explore biofilm growth dynamics for several *S. pneumoniae* strains with different virulence and serotype backgrounds. All strains tested were capable of forming biofilms, while some were quicker to achieve a higher biomass than others. The data presented in this work indicates there is considerable variability in the early stages of biofilm formation, followed by a more controlled growth as the biofilm matures. Indicating that biofilms need to be studied as a whole process instead of just unveiling single cycle stages.

This long-term biofilm assay is suitable for studying the detailed dynamics of biofilm formation and maintenance over long periods of time. This assay development opens the doors to new streptococcal biofilm research which we explore in the following chapters and include: . For instance: 1) research on dispersal cells detaching at different points of the biofilm life cycle is possible, 2) The discovery of the processes regulating the different stages of growth is possible with this new approach, 3) *S. pneumoniae* biofilms can now be used for experimental evolution studies to further assist the prediction of biofilm-related mutations which then can be used to design effective therapeutics to eradicate biofilms.

2.4 Materials and Methods

***S. pneumoniae* strain and biofilm assay growth conditions:** Strains used in this study are reported in **Table 3**. were initially grown overnight at 37C in 5% CO₂ on 5% defibrinated sheep blood agar plates. Biofilm assay: Bacteria were grown in liquid using semi-defined minimal media (SDMM- see full recipe on **table 2**) at PH 7.3 with 20mM glucose, 5ul/ml oxyrase, and 150 U/ml catalase (Worthington Bio Corp LS001896) for 3 generations (~2h) and diluted to an OD₆₀₀ of 0.03. 700ul of planktonic bacteria were seeded in a 24 well plate containing clean round coverslips and incubated at 37C in 5%CO₂. After 24 h, the dispersed cells were discarded or saved depending on the experiment, and each well was replenished with fresh SDMM to promote long-term biofilm formation. For extended growth periods, coverslips were transferred after 4 days to a new six-well plate containing a clean coverslip to provide a new surface space for dispersal cells from the old biofilm to attach to the new surface, this transferred process is repeated every four days.

Biofilm visualization and quantification: Biofilms were stained by adding 70ul of a solution containing Syto9 and propidium iodide (final concentrations of 2.5μM and 12.5μM, respectively) from the LIVE/DEAD™ BacLight™ (Molecular Probes, Life Technologies) and incubated from light for 15 min. The stain was washed with PBS, and the samples were placed on microscope slides. Z-stacks were acquired using a spinning disk confocal microscope (3i, Zeiss) with an APO 40X/1.10 oil immersion objective, and the focal planes were taken 0.3μm step size. Biomass (μm³/μm²), average thickness (μm²), maximum thickness (μm), and roughness coefficient were measured using COMSTAT2 software[45]. Six image stacks from four independent experiments were used for each analysis. In addition to imaging, cell viability was obtained from biofilms by scraping cells from coverslips with 1X PBS and from dispersal cells by centrifuging cells and resuspended in 1X PBS. Colony forming units (CFU) were recovered by serial dilution and plating on blood agar plates.

Statistical analysis: were performed using GraphPad Prism9. Imaging data were analyzed using one-way analysis of variance (ANOVA) with a post hoc Bonferroni test for multiple comparisons of measurements between *S. pneumoniae* strains. Principal component analysis was done using the PCA function under the multiple variable analysis tab in Prism9. The function generates PCA score diagram and contribution of variables table.

Table 2. SDMM recipe

Reagent	Stock (%)	1 L
Acid hydrolyzed Casein		5gr
Enzyme hydrolyzed Casein		1gr
L-cys.HCL	1	4ml
L-trp	0.57	1.05ml
L-asn	0.5	10ml
L-gln	0.1	10ml
Adenine	0.016	31.25ml
Ca-pantothenate	0.1	1.2ml
Nicotinic Acid	0.03	1ml
Pyridoxine.HCL	0.07	0.428ml
Thiamine.HCL	0.09	0.33ml
Riboflavine	0.002	7ml
Biotin	0.0035	0.017ml
K ₂ HPO ₄		8.5gr
NaOAc		2gr
NaHCO ₃		0.4gr
MgCl ₂ .6H ₂ O		0.5gr
CaCl ₂	0.6	1ml
CuSO ₄ .5H ₂ O	0.1	0.5ml
ZnSO ₄ .7H ₂ O	0.3	0.16ml
MnSO ₄ .4H ₂ O	0.07	0.286ml
Yeast Extract		0.5gr
Uracil	0.1	20ml
Adjust pH to 7.3 using 6N HCL		

Before inoculating cells add:

Choline.Cl	0.3% stock
FeSO4.7H2O	0.04% stock
Catalase	30 units/ml
Sugar	20 mM

Table 3. List of *S. pneumoniae* strains used.

Strain	Serotype	Virulence background
PG04	14	High
PG06	33A/F	Low
PG16	19A	High
TIGR4	4	High
TIGR4- CPS (2394)	4	Low
BHN97	19F	High
BS72	23F	Not known
BS71	3	Not Known

2.5 References

1. Steenackers, H.P., et al., *Experimental evolution in biofilm populations*. FEMS Microbiol Rev, 2016. **40**(3): p. 373-97.
2. Flemming, H.-C., et al., *Biofilms: an emergent form of bacterial life*. Nature Reviews Microbiology, 2016. **14**(9): p. 563-575.
3. Vestby, L.K., et al., *Bacterial Biofilm and its Role in the Pathogenesis of Disease*. Antibiotics (Basel), 2020. **9**(2).
4. Penesyan, A., et al., *Three faces of biofilms: a microbial lifestyle, a nascent multicellular organism, and an incubator for diversity*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 80.
5. Sauer, K., et al., *The biofilm life cycle: expanding the conceptual model of biofilm formation*. Nat Rev Microbiol, 2022. **20**(10): p. 608-620.
6. Yin, W., et al., *Biofilms: The Microbial "Protective Clothing" in Extreme Environments*. Int J Mol Sci, 2019. **20**(14).
7. Katharios-Lanwermyer, S., et al., *Pseudomonas aeruginosa Uses c-di-GMP Phosphodiesterases RmcA and MorA To Regulate Biofilm Maintenance*. mBio, 2021. **12**(1).
8. Liu, C., et al., *cAMP and c-di-GMP synergistically support biofilm maintenance through the direct interaction of their effectors*. Nat Commun, 2022. **13**(1): p. 1493.
9. Stefan Katharios-Lanwermyer, G.A.O.T., *Biofilm Maintenance as an Active Process: Evidence that Biofilms Work Hard to Stay Put*. Journal of bacteriology, 2022. **204**(4).
10. Guzman-Soto, I., et al., *Mimicking biofilm formation and development: Recent progress in in vitro and in vivo biofilm models*. iScience, 2021. **24**(5): p. 102443.
11. Azevedo, N.F., J. Allkja, and D.M. Goeres, *Biofilms vs. cities and humans vs. aliens - a tale of reproducibility in biofilms*. Trends Microbiol, 2021. **29**(12): p. 1062-1071.
12. O'Toole, G.A., *Microtiter dish biofilm formation assay*. J Vis Exp, 2011(47).
13. Kragh, K.N., et al., *Into the well-A close look at the complex structures of a microtiter biofilm and the crystal violet assay*. Biofilm, 2019. **1**: p. 100006.
14. Schwartz, K., et al., *The use of drip flow and rotating disk reactors for Staphylococcus aureus biofilm analysis*. J Vis Exp, 2010(46).
15. Henriques-Normark, B. and E.I. Tuomanen, *The pneumococcus: epidemiology, microbiology, and pathogenesis*. Cold Spring Harb Perspect Med, 2013. **3**(7).
16. Domenech, M., E. Garcia, and M. Moscoso, *Biofilm formation in Streptococcus pneumoniae*. Microb Biotechnol, 2012. **5**(4): p. 455-65.
17. Yadav, P., et al., *Deciphering Streptococcal Biofilms*. Microorganisms, 2020. **8**(11).

18. Brook, P.S., et al., *Different Methods for Culturing Biofilms In Vitro*, in *Biofilm Infections*, T. Bjarnsholt, Jensen, P., Moser, C., Høiby, N, Editor. 2010, Springer: New York. p. 251-266.
19. Moscoso, M., E. Garcia, and R. Lopez, *Biofilm formation by Streptococcus pneumoniae: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion*. J Bacteriol, 2006. **188**(22): p. 7785-95.
20. Blanchette, K.A., et al., *Neuraminidase A-Exposed Galactose Promotes Streptococcus pneumoniae Biofilm Formation during Colonization*. Infect Immun, 2016. **84**(10): p. 2922-32.
21. del Prado, G., et al., *Biofilm formation by Streptococcus pneumoniae strains and effects of human serum albumin, ibuprofen, N-acetyl-L-cysteine, amoxicillin, erythromycin, and levofloxacin*. Diagn Microbiol Infect Dis, 2010. **67**(4): p. 311-8.
22. Churton, N.W., et al., *Parallel Evolution in Streptococcus pneumoniae Biofilms*. Genome Biol Evol, 2016. **8**(5): p. 1316-26.
23. Blanchette-Cain, K., et al., *Streptococcus pneumoniae biofilm formation is strain dependent, multifactorial, and associated with reduced invasiveness and immunoreactivity during colonization*. mBio, 2013. **4**(5): p. e00745-13.
24. Wróblewska, J., et al., *Biofilm formation of Streptococcus pneumoniae from bronchial alveolar lavage and from nasal swab*. Medical Research Journal, 2016. **1**(3).
25. Chan, W.T., et al., *The Streptococcus pneumoniae yefM-yoeB and relBE Toxin-Antitoxin Operons Participate in Oxidative Stress and Biofilm Formation*. Toxins (Basel), 2018. **10**(9).
26. Hall-Stoodley, L., et al., *Characterization of biofilm matrix, degradation by DNase treatment and evidence of capsule downregulation in Streptococcus pneumoniae clinical isolates*. BMC Microbiol, 2008. **8**: p. 173.
27. Traverse, C.C., et al., *Tangled bank of experimentally evolved Burkholderia biofilms reflects selection during chronic infections*. Proc Natl Acad Sci U S A, 2013. **110**(3): p. E250-9.
28. Domenech, M., E. Garcia, and M. Moscoso, *Versatility of the capsular genes during biofilm formation by Streptococcus pneumoniae*. Environ Microbiol, 2009. **11**(10): p. 2542-55.
29. Yadav, M.K., S.W. Chae, and J.J. Song, *In Vitro Streptococcus pneumoniae Biofilm Formation and In Vivo Middle Ear Mucosal Biofilm in a Rat Model of Acute Otitis Induced by S. pneumoniae*. Clin Exp Otorhinolaryngol, 2012. **5**(3): p. 139-44.
30. Sanchez-Rosario, Y. and M.D.L. Johnson, *Media Matters, Examining Historical and Modern Streptococcus pneumoniae Growth Media and the Experiments They Affect*. Front Cell Infect Microbiol, 2021. **11**: p. 613623.
31. Rosconi, F., et al., *A bacterial pan-genome makes gene essentiality strain-dependent and evolvable*. Nat Microbiol, 2022. **7**(10): p. 1580-1592.
32. Jensen, P.A., Z. Zhu, and T. van Opijnen, *Antibiotics Disrupt Coordination between Transcriptional and Phenotypic Stress Responses in Pathogenic Bacteria*. Cell Rep, 2017. **20**(7): p. 1705-1716.

33. Limoli, D.H., C.J. Jones, and D.J. Wozniak, *Bacterial Extracellular Polysaccharides in Biofilm Formation and Function*. Microbiol Spectr, 2015. **3**(3).
34. Waldrop, R., et al., *Biofilm growth has a threshold response to glucose in vitro*. Clin Orthop Relat Res, 2014. **472**(11): p. 3305-10.
35. Roy, P.K., et al., *Effects of environmental conditions (temperature, pH, and glucose) on biofilm formation of Salmonella enterica serotype Kentucky and virulence gene expression*. Poult Sci, 2021. **100**(7): p. 101209.
36. Troxler, L.J., et al., *Carbon source regulates polysaccharide capsule biosynthesis in Streptococcus pneumoniae*. J Biol Chem, 2019. **294**(46): p. 17224-17238.
37. Werren, J.P., et al., *Carbon Source-Dependent Changes of the Structure of Streptococcus pneumoniae Capsular Polysaccharide with Serotype 6F*. Int J Mol Sci, 2021. **22**(9).
38. Wilson, C., et al., *Quantitative and Qualitative Assessment Methods for Biofilm Growth: A Mini-review*. Res Rev J Eng Technol., 2017. **6**(4).
39. Allkja, J., et al., *Interlaboratory study for the evaluation of three microtiter plate-based biofilm quantification methods*. Sci Rep, 2021. **11**(1): p. 13779.
40. Vandecandelaere, I., H. Van Acker, and T. Coenye, *A Microplate-Based System as In Vitro Model of Biofilm Growth and Quantification*. Methods Mol Biol, 2016. **1333**: p. 53-66.
41. C, W., et al., *Quantitative and Qualitative Assessment Methods for Biofilm Growth: A Mini-review*. Res Rev J Eng Technol., 2017. **6**(4).
42. Mountcastle, S.E., et al., *Biofilm viability checker: An open-source tool for automated biofilm viability analysis from confocal microscopy images*. NPJ Biofilms Microbiomes, 2021. **7**(1): p. 44.
43. Larimer, C., et al., *A method for rapid quantitative assessment of biofilms with biomolecular staining and image analysis*. Anal Bioanal Chem, 2016. **408**(3): p. 999-1008.
44. Giorgi, F., J.M. Curran, and E.A. Patterson, *Real-time monitoring of the dynamics and interactions of bacteria and the early-stage formation of biofilms*. Sci Rep, 2022. **12**(1): p. 18146.
45. Heydorn, A., et al., *Quantification of biofilm structures by the novel computer program COMSTAT*. Microbiology, 2000. **146**: p. 12.
46. Verotta, D., et al., *Mathematical Modeling of Biofilm Structures Using COMSTAT Data*. Comput Math Methods Med, 2017. **2017**: p. 7246286.
47. Camilli, R., A. Pantosti, and L. Baldassarri, *Contribution of serotype and genetic background to biofilm formation by Streptococcus pneumoniae*. Eur J Clin Microbiol Infect Dis, 2011. **30**(1): p. 97-102.
48. Valente, C., et al., *Intra-Species Interactions in Streptococcus pneumoniae Biofilms*. Front Cell Infect Microbiol, 2021. **11**: p. 803286.

Chapter 3

Phenotypic Characterization of TIGR4 *S. pneumoniae* biofilms. *

*The contents of this chapter were adapted from the first part of the following manuscript:

Espinoza-Miranda S, Schmiege E, Chu C, Chowdhury A, Bodrog S, Rosconi F, van Opijnen T. Genome-wide identification of genetic requirements of *Streptococcus pneumoniae* biofilms. Manuscript in preparation.

Author contributions:

Suyen Espinoza-Miranda (SE) and Tim van Opijnen (TvO) conceived and design the study. SE, Emma Schmiege (ES), Cedrick Chu (CC), Allison Chowhury (AC) and Sophie Bodrog (SB) performed the experiments. Federico Rosconi (FR) and SE wrote the manuscript. Tim van Opijnen (TvO) edited and approved the manuscript.

3.1 Background

Streptococcus pneumoniae (*S. pneumoniae*) usually referred to as the pneumococcus, is a Gram-positive bacterium recognized as an important colonizer of the upper respiratory tract in humans. *S. pneumoniae* can easily migrate to other tissues and organs, leading to both acute and chronic infections. Moreover, it can adhere to tissues in the form of biofilms [1, 2]. Given the frequency and extended time of pneumococcal colonization in the nasopharynx and other organs, it is crucial to gain a comprehensive understanding of biofilm development across all stages of its life cycle.

Previous studies have focused on specific genes related to cell attachment, such as the polysaccharide capsule, and essential virulence factors like pneumolysin, which are necessary for biofilm formation[3]. However, the biofilm lifecycle consists of different stages including attachment, maturation (division and expansion of attached cells) and dispersion, the last of which refers to the detachment of cells from the biofilms that will find a new surface to colonize. Studies focusing only on a small part of the lifecycle, such as attachment, thus tend to miss the complexity that make biofilms so important for many facets of bacterial survival and virulence. To comprehensively study *S. pneumoniae* biofilms on a genome-wide scale, it is therefore not only imperative to capture the entire biofilm lifecycle, but also understand their importance *in vivo*.

Biofilm dispersion, the process where cells detach from a biofilm and enter the environment (as illustrated in **Figure 1 - Chapter 2**), plays a vital role in the ecology and pathogenesis of bacteria. It enables colonization of new surfaces, transmission of infections, and evasion from environmental stress or host immune responses[4, 5]. Research on dispersal cells is limited and remains a topic of discussion, as these cells exhibit a unique phenotype, distinct from both planktonic and surface-attached cells[6]. For example, studies have shown that dispersal cells from *Pseudomonas aeruginosa* are highly virulent *in vivo* and display specific antibiotic susceptibility patterns [7, 8]. In contrast, limited

knowledge exists regarding *S. pneumoniae* dispersal cells. Only one group has explored dispersion after exposure to factors such as Influenza A virus (IAV) and changes in temperature[9, 10]. A previous study showed there is a 10-fold increase in the dispersal cell population upon addition of IAV when compared to their biofilm cell counterparts. In addition, they showed that induction of dispersion after viral treatment increases inflammation levels in a murine model. However, much remains unknown about *S. pneumoniae* dispersal cells, including their virulence. Therefore, a thorough phenotypic characterization of dispersal cells is key to achieving a comprehensive understanding of this critical biofilm stage.

In this chapter, a detailed phenotypic characterization of *S. pneumoniae* biofilms is conducted by tracking and quantifying their growth. Subsequently, biofilms and their dispersal cells are investigated in a murine model to explore severity of the disease of these populations against planktonic cell. And finally, we assess the survival of bacteria in biofilms, planktonic and dispersal conditions upon antibiotic treatment.

3.2 Results

3.2.1 TIGR4 *S. pneumoniae* biofilms undergo spatial heterogeneity.

Biofilm surface-attached cells characterization

To further explore the biofilm architecture and growth dynamics of *S. pneumoniae* biofilms, we used the clinical isolate TIGR4. This serotype 4 strain isolated from a 30-year-old patient with meningitis from Norway [11] has been used in multiple in vitro characterizations under planktonic conditions and is commonly used in pneumonia in vivo models[12]. TIGR4 biofilms were grown using the assay developed in **Chapter 2**. Briefly, glass coverslips are used as a surface for planktonic bacteria to attach, grow and divide. Glass coverslips are placed on 24-well-plates and media is changed every day to promote biofilm formation. Biofilms are stained, imaged and quantified as previously described

(Chapter 2, Methods sections). Representative biofilm images from each day are shown in **Figure 1F** (top panel).

Large variation in biomass, average and maximum thickness and roughness coefficient is observed among TIGR4 biological replicates. The formation of biofilm within the initial 24 hours of growth appears to be a variable process, particularly for TIGR4, as indicated by a decreasing variability over time. For example, biomass from day 1 ranges from ~10 to 25 [$\mu\text{m}^3/\mu\text{m}^2$], whereas at day 4 the range is reduced to ~22 to 30 [$\mu\text{m}^3/\mu\text{m}^2$] (**Figure 1A**).

It is important to notice that an increase in biomass does not necessarily indicate an expansion in the number of viable cells, as the ratio of live to dead cells at day 4 is around 50% of total biomass (**Figure 1C**). The parameter of total biomass takes into account live and dead cells in the images, really quantifying all the material attached, while CFU captures the numbers of live cells within the biofilm. Moreover, the roughness coefficient decreases over time, indicating a more homogeneous cell aggregate topography. Overall, the highest values for average and maximum thickness are reached on day 4 (**Figure 1B**).

Biofilm dispersal cells characterization

The number of viable cells was assessed for surface-attached cells and dispersal cells released from TIGR4 biofilms over a period of four days. Dispersal cells were obtained by carefully pipetting them out from the 24-well plates, centrifuging, and resuspending in 1XPBS. Cell viability was achieved by serially diluting bacteria and plating on blood agar plates. The initial observation was that the number of viable cells recovered from the dispersal mode of growth was always higher (~10X) than those attached to the coverslips throughout the four days (**Figure 1D**). After one day of growth, both dispersal and surface-attached cells exhibited opaque, transparent, and small colony variants (SCV), a

common phenotype that emerges shortly after attachment during biofilm development [13, 14]. Biofilm- derived *S. pneumoniae* variants are the result of varying levels of capsule production [15, 16]. To further characterize dispersal cells, they were stained with Syto 9 and propidium iodide to assess live and dead cells respectively. Twelve images were acquired using fluorescence microscopy after each day for four days (representative images are shown on **Figure 1F**- bottom panel). The size of cell aggregates was quantified utilizing the particle analysis function in FIJI. Following imaging and quantification, several observations were made: 1) these cells were different from planktonic as they detached in the form of both cell aggregates and individual cells, 2) upon examining hundreds of cell aggregates, it was determined that the average size of aggregates increased steadily each day, 3) larger aggregates ($< 15 \mu\text{m}^2$) were present on all days, indicating that dispersal cells also constituted a heterogeneous population throughout the growth period (**Figure 1D**).

Altogether, we observed that the *in vitro* characterization of biofilms indicates that biofilm establishment is highly variable in the early stages of growth. A more controlled growth with higher biomass production is observed as the biofilm develops. It was also demonstrated that dispersal cells consist of a mixture of single cells and cell aggregates.

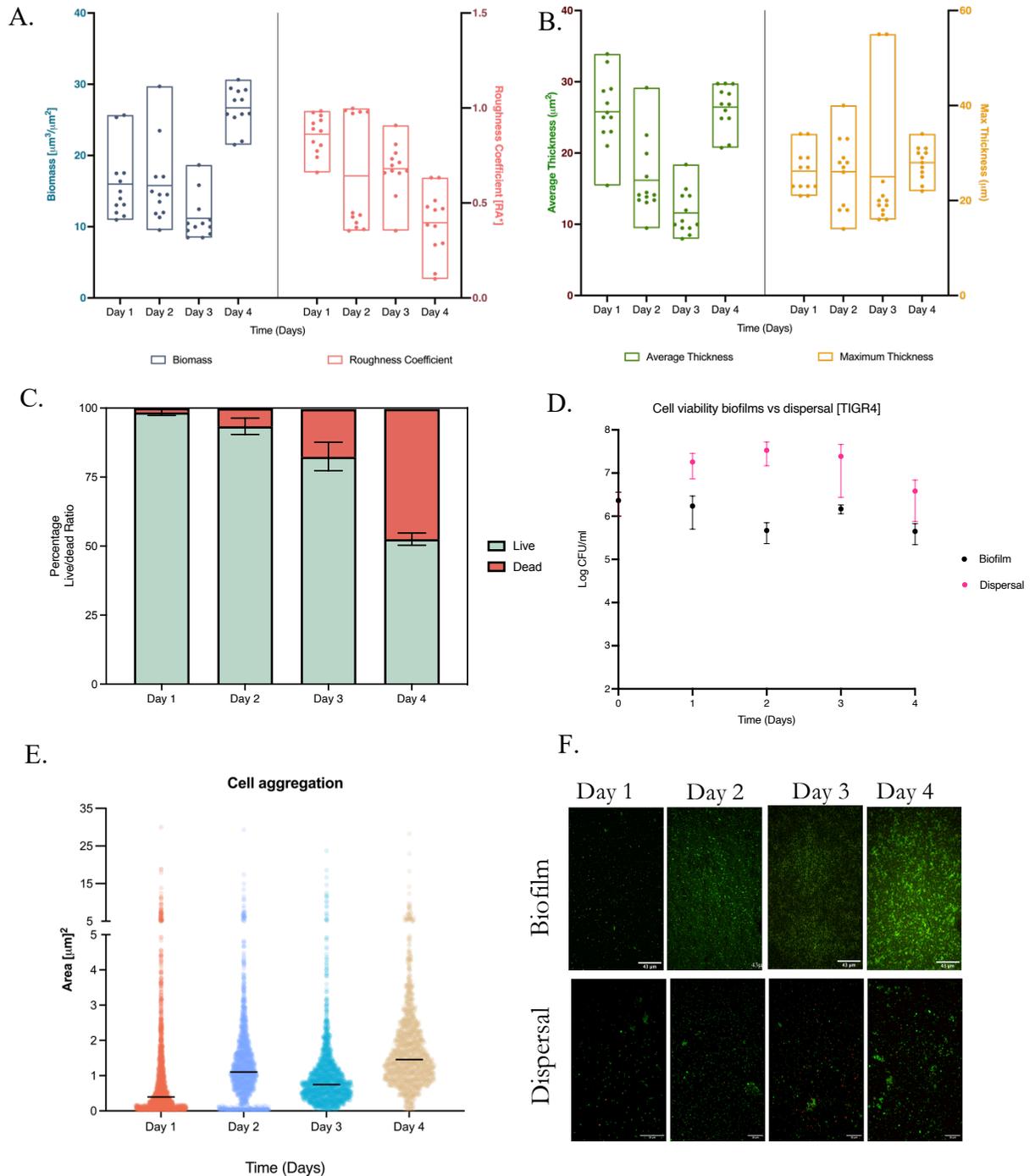


Figure 1. Biofilm and dispersal growth dynamics of the strain TIGR4. **A & B.** Biofilm quantification parameters for four days. Biofilms were stained using BacLight imaged after each day as described previously. COMSTAT was used to generate all measurements. **A.** Biomass (left Y axis) and roughness coefficient (right axis). **B.** Average thickness (left Y axis) and maximum thickness (right axis). **C.** Percentage of live and dead bacteria calculated from total biomass. **D.** Cell viability of biofilms and dispersal cells across all four days. **E.** Dispersal cell aggregation size. Dispersal cells were stained

using BacLight kit and images were taken using 3i spinning disk and cell aggregates were measure using FIJI. **F.** Representative images of biofilm growth over four days. Top panels show biofilms, and bottom panels dispersal cells. Biofilms were grown, stained, and imaged as previously described in Chapter 2.

3.2.1 *In vivo* characterization: Different growth states exhibit different levels of virulence.

Different organs in the human body, including the nasopharynx, ears, lungs, and the heart can be colonized by *S. pneumoniae* biofilms [17-19]. After characterizing the *in vitro* growth of surface-attached and dispersal cells, we focused on the assessment of the phenotypes of planktonic, biofilm, and dispersed populations *in vivo*. Mice were intranasally administered 2-4 X 10⁶ CFU of planktonic bacteria, three-day-old *in vitro* grown biofilms, or dispersal bacteria. To intranasally inoculate eight mice with 50 µl at a 10⁶ concentration the inocula preparation proceeded as followed: Biofilms were grown with daily media change, on the third day cells were scraped from 12 biofilm-containing coverslips, recovered, centrifuged and washed three times with 1X PBS and finally resuspended in 500ul of 1X PBS. The dispersal samples were treated as follow: a total from 36ml of dispersal cells from three-day old biofilms were pipetted out, centrifuged, washed three times with 1X PBS followed by a final resuspension of 1ml in 1X PBS. After 48 hours mice inoculated with biofilm and planktonic cells were sacrificed and bacterial counts were obtained from nasopharyngeal lavage, lungs, and blood. In addition, all mice were scored for signs of disease where a score of 0 indicates no disease and a score of 7 illustrates a lethargic animal [20].

Biofilm cells exhibit an attenuated phenotype.

Mice from the three sample groups (planktonic, biofilm and dispersal cells) were all able to colonize the nasopharynx, but based on the number of CFUs recovered, the dispersal cells were the most successful colonizers (**Figure 2B**). Mice inoculated with three-day-old biofilms were unable to cause

lung infection 48 hours after infection, whereas their dispersal counterparts were able to spread to the lungs at least 14 hours post-infection. Moreover, biofilms were unable to transition to the blood, while both planktonic and dispersal populations spread to the blood and reached similar CFU levels (**Figure 2B**). The avirulent phenotype of biofilms has been previously reported by Sanchez and colleagues [21] where they found that two-day-old biofilms from TIGR4 were unable to cause pneumonia or bacteremia, whereas mice infected with planktonic bacteria showed signs of invasive disease (after 48 hours post infection -h.p.i- at a 10^6 inoculum), defined based on the number of bacteria recovered from lungs and blood. Subsequently, they developed bacteremia and, ultimately, succumbed to the infection [21]. Overall, our results confirm that animals inoculated with biofilms are unable to cause severe disease but can effectively colonize the nasopharynx. This contrasts with mice inoculated with planktonic cells, which are characterized by severe clinical scores, and high bacterial loads in the lungs and blood (**Figure 2**).

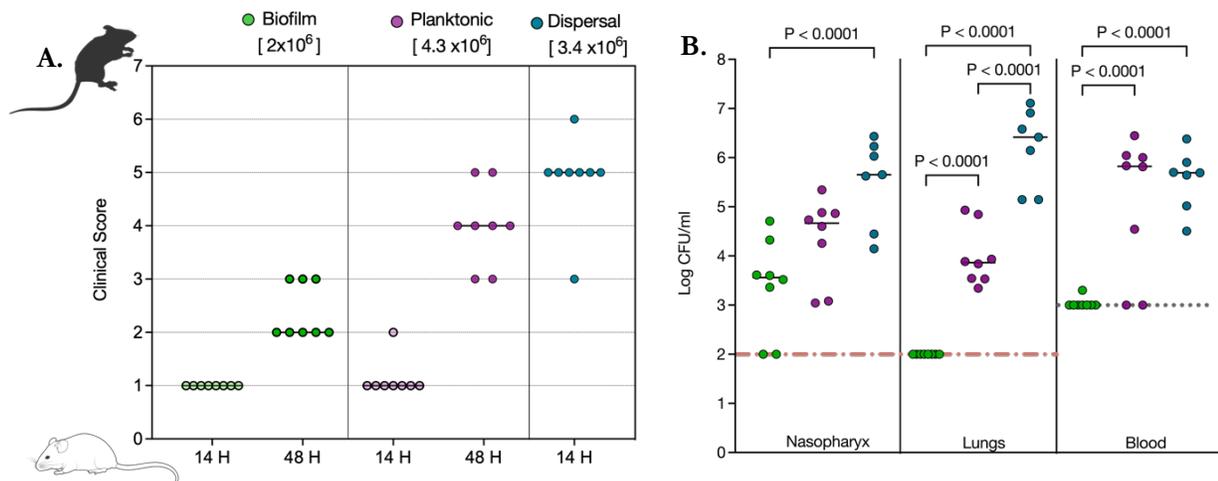


Figure 2. Different growth states from the strain TIGR4 exhibit varying levels of virulence. 6-week-old female Swiss Webster mice were intranasally inoculated using planktonic, three-day-old biofilm or dispersal cells. **A.** Clinical scores from mouse experiments at the time of sacrifice. A score of 1 indicates a lack of clinical signs, and a score of 6 denotes a slightly lethargic animal. **B.** Nasopharyngeal lavage, lungs, and blood were collected from each mouse. Each experiment has an $n=8$. Statistical analysis was performed using one-way ANOVA with no multiple comparison correction.

Dispersal cells are in a ready-to-infect state.

Mice inoculated with dispersal cells exhibited a 1000-fold higher level of colonization of the nasopharynx than mice inoculated with biofilm cells (**Figure 2B**). Moreover, dispersal cells exhibited a hypervirulent phenotype: 1) Dispersal cells caused signs of disease at 14 h.p.i (**Figure 2A**); and 2) Bacterial loads are 1000-fold higher in lungs compared to planktonic cells. Unlike the biofilm inoculum at the same dose, dispersal cells migrate to the blood. This outcome is in line with previous findings, which showed an increased virulence of dispersal-derived cells post-temperature and IAV treatment from strains D39 and EF3030 to cause infection in a mouse model [22]. In contrast, we show that dispersal cells do not need to be stressed (e.g., by temperature or IAV treatment). Instead, naturally released dispersal cells from biofilms exhibit higher virulence compared to both biofilms and planktonic cells.

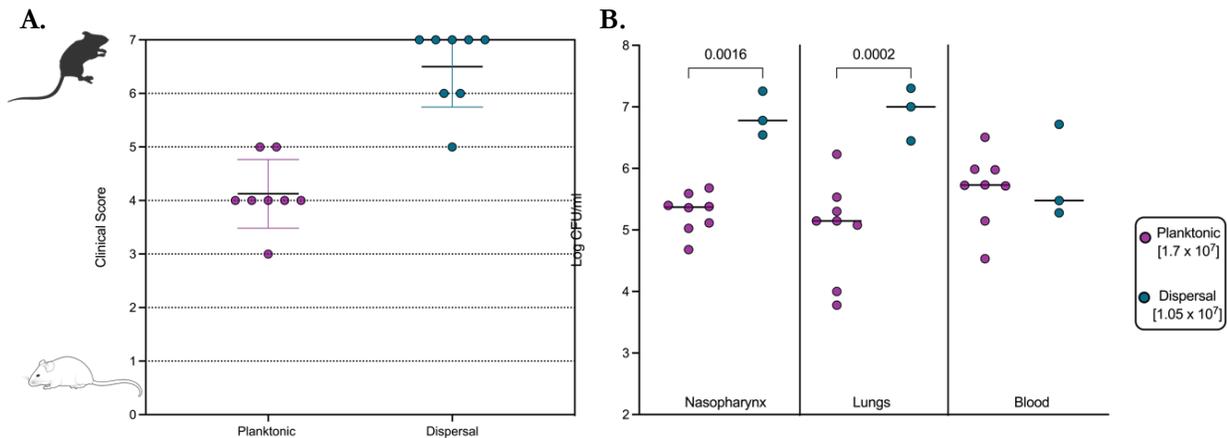


Figure 3. The virulence phenotype of dispersal cells is exacerbated when the inoculum is increased. 6-week-old female Swiss Webster mice were intranasally inoculated using planktonic or dispersal cells. **A.** Clinical scores from mouse experiments at the time of sacrifice. A score of 1 indicates a lack of clinical signs, and a score of 6 denotes a lethargic animal. **B.** Nasopharyngeal lavage, lungs, and blood were collected from each mouse. At 12 hours post infection, 5 out of 8 mice inoculated with dispersal cells were rigor mortis. Each experiment has an n=8. Statistical analysis was performed using one-way ANOVA.

Moreover, when a higher inoculum was used (1×10^7 cells) to infect mice either with dispersal or planktonic cells, of the 8 animals inoculated with dispersal cells, 5 died before 12 hours post-infection (clinical score of 7), and the other 3 were lethargic (clinical score of 5 and 6). In contrast, the same dose (1×10^7) of planktonic cells took 24 hours to reach symptoms and enter the bloodstream. Altogether, this shows that *S. pneumoniae* dispersal cells are more virulent than planktonic and biofilm populations, as evidenced by the higher clinical score and bacterial load in mice inoculated with dispersal cells at 14 hours.

3.2.3 Phenotypic antibiotic tolerance profile in *S. pneumoniae* biofilms.

One of the most significant characteristics of biofilms is the ability of bacteria to survive high doses of antibiotics [23, 24]. This type of survival in biofilm-associated antibiotic treatment is referred to as "tolerance," which distinguishes it from "resistance." Resistance involves the acquisition of genetic mutations that alter the target site of an antibiotic or confer the ability to degrade or efflux the antibiotic, resulting in an increase in minimum inhibitory concentration (MIC) [25]. In contrast, antibiotic tolerance refers to the ability of biofilm-associated bacteria to survive exposure to antibiotic concentrations that are lethal to planktonic cells without changing the MIC and without undergoing cell replication [26]. Factors contributing to tolerance include metabolic heterogeneity and reduced antibiotic penetration [27]. The combination of tolerance and resistance in biofilms makes them a major public health concern, contributing to the persistence of chronic infections, increased healthcare costs, and the emergence of multidrug-resistant bacteria.

Tolerance reflects the overall ability of a bacterial community to survive exposure to antibiotics [27]. On the other hand, the subpopulation of cells within the bacterial community that withstand antibiotic exposure is referred to as persistence [26, 28]. Persistent cells are typically characterized by a state of dormancy or low metabolism levels and normally constitute 10^{-2} to 10^{-5} of a population [29, 30] but

these numbers are sufficient to reseed the biofilm and cause a relapsed in the infection. The standard method of measuring tolerance and persistence is by performing a time-kill measurements. In this method bacteria are exposed to high concentrations of antibiotics and viable CFUs are plotted against time [31]. Tolerance is characterized by an increased in the minimum duration of killing (MDK) that takes the antibiotic to eradicate certain proportion of the bacterial population (**Figure 4**) [31]. MDK₉₉ is used for tolerance measuring tolerance which refers to the minimum duration of killing for the antibiotic to eradicate 99% of the population. Persistence can be measured by detecting higher MDKs, characteristic of a biphasic mode of killing [30] (**Figure 4**).

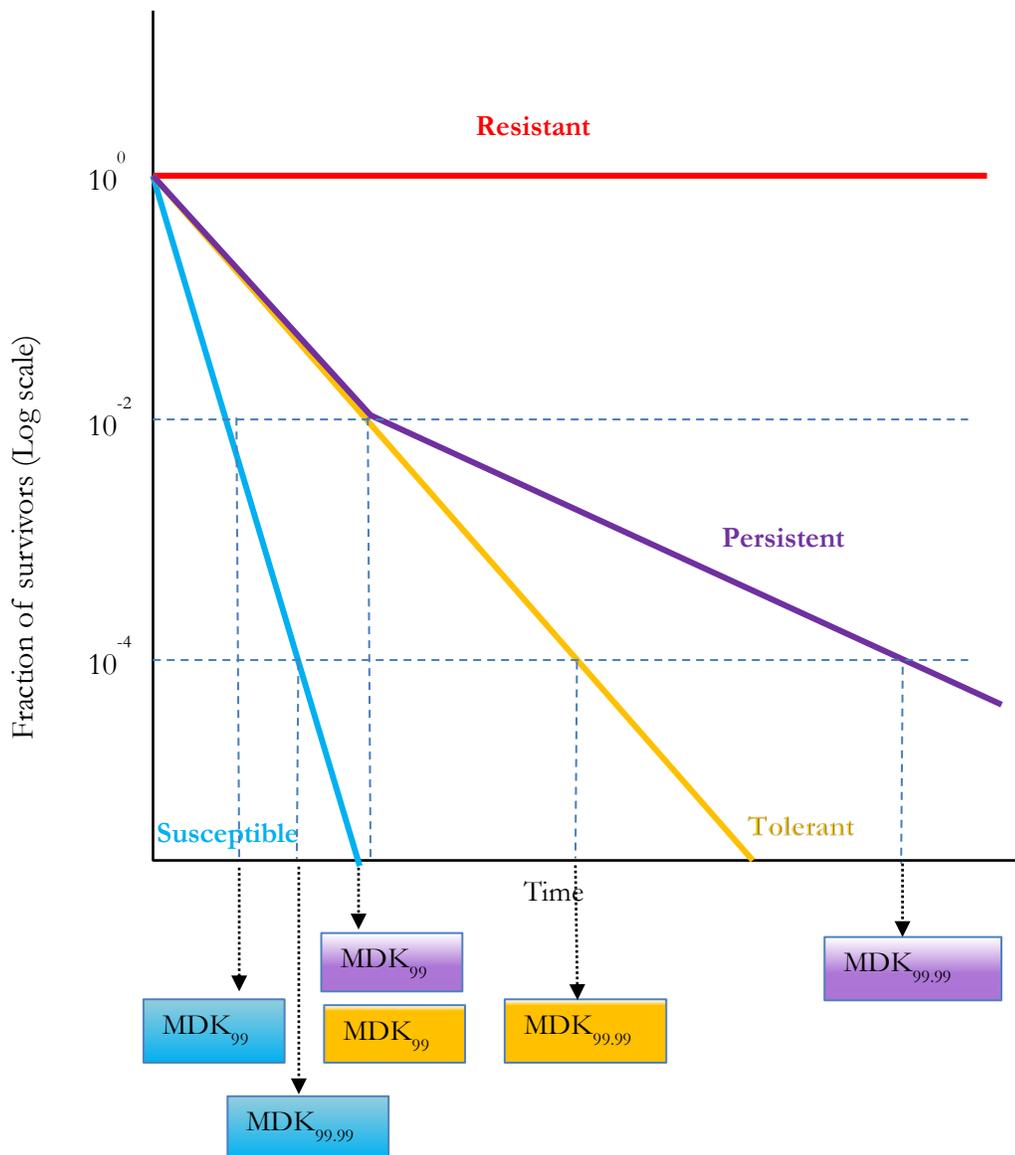


Figure 4. Differentiating Resistance, Tolerance and Persistence. Resistance is the ability of bacteria to withstand the effect of antibiotics through acquired genetic mutations. Tolerance is the capacity of bacterial population to survive high antibiotics concentrations. A tolerant population is characterized by a longer minimum duration of killing (MDK) when compared to susceptible populations. Lastly, persistence refers to a small subset of bacteria and is illustrated as a biphasic killing curve where the first phase represents rapid killing and the second is characterized by a prolonged killing therefore having a higher MDK. MDK₉₉ represents the minimum time of killing that will take for 99% of the bacterial population to be eradicated.

Understanding how biofilms behave under antibiotic exposure is crucial for developing effective strategies to combat biofilm-associated infections. Tolerance and persistence have been mostly studied in planktonic growth and just a few species such as *E. coli*, *S. aureus* and *P. aeruginosa* have been used as models to study persistence in biofilms. In this section we assessed the ability of *S. pneumoniae* biofilms to survive antibiotic treatment compared to planktonic cells.

Biofilm-derived dispersal cells are a source of antibiotic-tolerant populations.

When bacteria live in a repressed metabolism state, they lose sensitivity to antibiotics [27]. We hypothesize that biofilm and dispersal cells house tolerant populations of cells when exposed to antibiotics. It is presumed that biofilms harbor a high number of tolerant cells due to the nature of the biofilms to protect bacteria leaving those in the bottom layer of the biofilm at a minimum metabolic level which leave bacteria in a dormant state preventing them to get killed [30, 32]. To test this hypothesis, we subjected biofilms, planktonic, and dispersal populations to two different classes of antibiotics: the cell wall synthesis inhibitor vancomycin and the DNA synthesis inhibitor ciprofloxacin. First the minimum inhibitory concentration (MIC) was determined for both antibiotics in planktonic growth using SDMM and growth curves are shown in **Figure 5**.

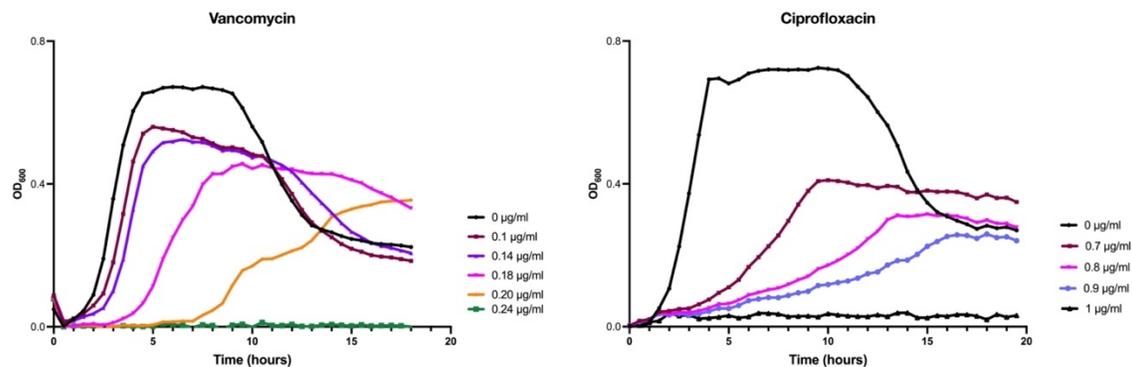


Figure 5. Minimum Inhibitory Concentration (MIC) determination for vancomycin and ciprofloxacin. 1 to 5×10^5 CFU of mid- exponential bacteria was diluted in fresh SDMM with a single antibiotic to achieve a final concentration gradient of vancomycin: 0 - $0.24 \mu\text{g}/\text{mL}$ and ciprofloxacin: 0 - $1 \mu\text{g}/\text{mL}$. Growth was monitored on BioSpa reader at 37°C for 20 hours. MIC is determined as the lowest concentration that abolishes bacterial growth. In this case MIC for vancomycin: $0.24 \mu\text{g}/\text{mL}$ and ciprofloxacin $1 \mu\text{g}/\text{mL}$.

After determining the minimum inhibitory concentration (MIC) of the antibiotic, we employed a concentration five to ten times higher than the MIC to assess the killing of planktonic, biofilm, and dispersed cells. The tolerance assay was conducted in the following manner for the three types of growth: for exponentially growing planktonic cells, the cells were exposed to the antibiotic, and samples were taken at various time points for bacterial counting [33]. For biofilm samples, two-day-old biofilms were treated with fresh SDMM medium containing or not containing antibiotics. Simultaneously, dispersal cells removed from two-days-old biofilms were spun down and resuspended in fresh SDMM medium with or without antibiotics. Samples from the three different groups were collected at different time points for 24 hours for bacterial counting. To calculate the rate of cell survival, the number of bacteria at each time point was divided by the initial population.

The addition of 5 - 10 times of vancomycin to two-day-old biofilms generated tolerant cells as depicted by the killing curve shown in **Figure 6**. A biphasic killing curve was generated to show killing dynamics among the planktonic, biofilm and dispersal samples. We were able to isolate viable bacteria in biofilms

and dispersal populations, even after 24 h post-antibiotic treatment. In contrast, planktonic bacteria were undetectable after 12 h post treatment (**Figure 6A**). Furthermore, we highlight the MDK₉₉, as a measure of tolerance for the three different populations. As tolerance is characterized by a longer MDK, we note that while it takes around ~6 hours for the antibiotic to kill 99% of the planktonic population, for biofilms, killing is delayed by 4 hours compared to planktonic. When analyzing dispersal cells, the MDK₉₉ extends approximately to 16- 18 hours post-antibiotic exposure, indicating that dispersal cells harbor a higher number of tolerant cells (**Figure 6A**). To the best of our knowledge, this is the first report on the identification of antibiotic-tolerant cells in *S. pneumoniae* biofilms and dispersal populations.

Tolerance can be influenced by the type of antibiotic used, for example for the strain D39 *S. pneumoniae*, higher rates of survival bacteria can be isolated with vancomycin than cefepime and amoxicillin [34]. To answer whether tolerance is dependent on the type of antibiotics for TIGR4 *S. pneumoniae* biofilms, we exposed biofilms, planktonic and dispersal cells to 10 times the MIC of Ciprofloxacin and Vancomycin for 8 hours. After repeating the experiment on three different days, we confirmed that tolerance was not influenced by the type of antibiotics used in planktonic and biofilm samples (**Figure 6B**). The proportion of surviving bacteria in dispersal cells was higher for ciprofloxacin than vancomycin after 8 hours of treatment. As dispersal cells from other species have been identified to have a different physiological profile than their planktonic and biofilm counterparts [7], it might be that *S. pneumoniae* dispersal cells exhibit an unique transcriptome profile which can be the cause for different antibiotic sensitivity from the sessile biofilm population.

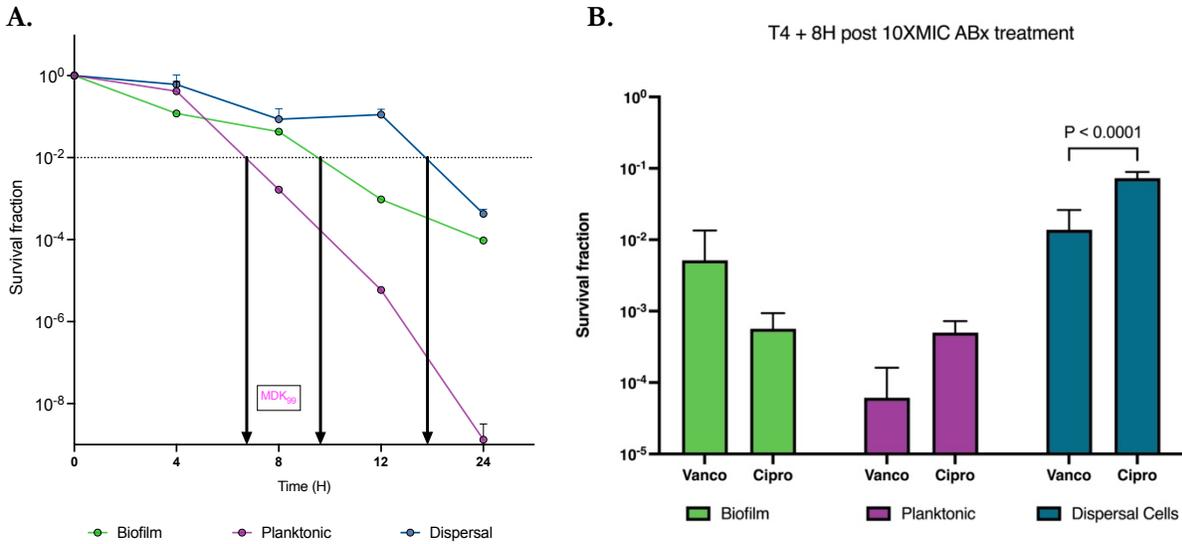


Figure 6. *S. pneumoniae* biofilm and dispersal cells harbors tolerant cells upon addition of antibiotic. **A.** Tolerant cells were isolated from planktonic, four-day-old biofilms and dispersal populations. 5X MIC of vancomycin was used, and samples were taken at 4-, 8-, 12-, and 24-hours post-treatment and plated in blood agar plates for enumeration. **B.** Survival fraction of recovered tolerant cells after 8 h post-antibiotic exposure with 10X MIC of vancomycin or ciprofloxacin for the three different states of growth (planktonic, biofilm and dispersal). Statistical analyses were performed using Two-way Anova comparing each treatment within sample groups. Followed by a Bonferroni correction was performed to show significance.

By definition, tolerance happens in populations where there is no change in the MIC [35]. We tested if dispersal cells which is the sample group with higher amount of survival cells upon antibiotic treatment changed their MIC. To test that dispersal cells are tolerant but not resistant they were cultivated in the presence of 1X MIC vancomycin, and their optical density was recorded every 30 minutes. After 20 hours, three biological replicates of dispersal cells exhibited a lack of growth under 1X MIC vancomycin conditions in contrast to the no antibiotic control which exhibited normal grow (Figure 7A). These results confirm that dispersal cells are tolerant and not resistant to the antibiotic treatment applied.

Since tolerance is the global phenotype of the population that survives the antibiotic treatment, we wondered if dispersal cells were in a dormant state which would be characteristic of persister cells. To

test is these cells are dormant we measure the growth rate and doubling time of dispersal cells over four days. To achieve this, we remove dispersal cells from the well, washed them with 1X PBS and grew them for 20 hours. TIGR4 in planktonic growth was used as a control for each day. After calculating growth rate and doubling time, we conclude there is not significant differences between the growth of dispersal cells and planktonic TIGR4 indicating that these cells are not in a dormant state (**Figure 7B**).

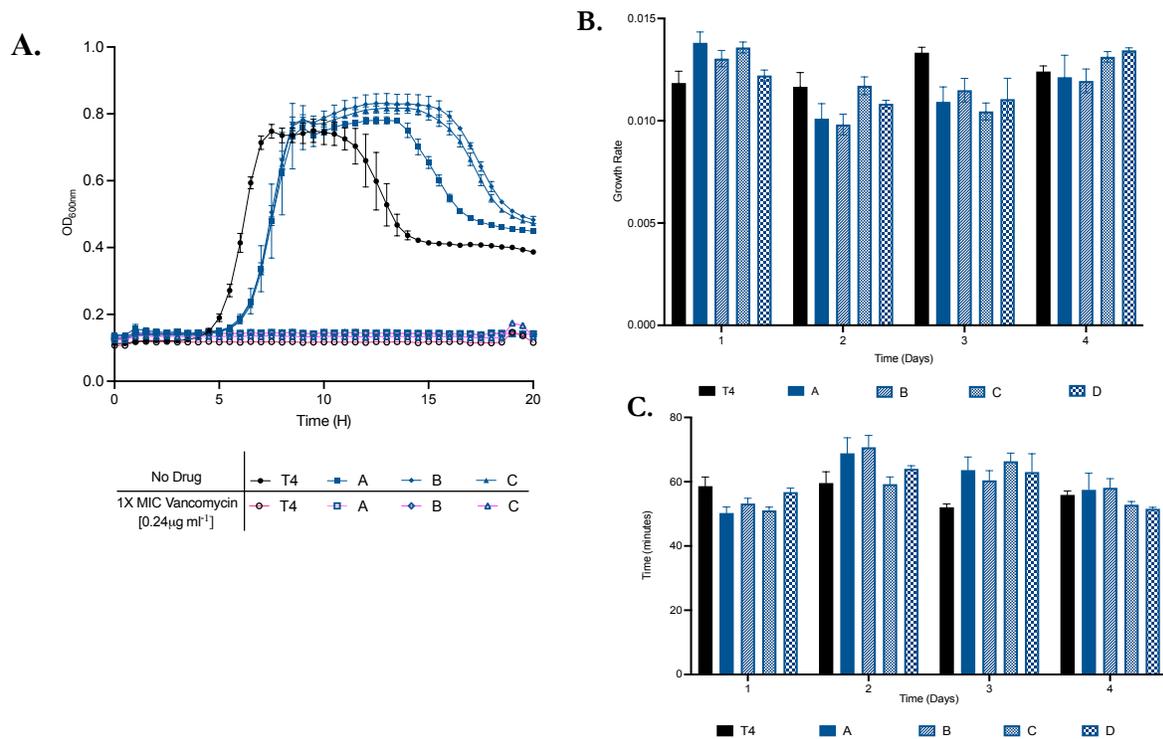


Figure 7. Dispersal cells are not a dormant population. **A.** Growth curves of dispersal cells using SDMM show no change in growth rate when compared to planktonic growth. Dispersal cells are tolerant and not resistant as they fail to grow under the exposure of 1X MIC of the antibiotic. **B & C.** Growth rate and doubling time from dispersal cell over four days are not different when compared to planktonic growth. Dispersal cells were recovered after each day and were subjected to growth compared to planktonic cells in SDMM.

3.3 Conclusions

The biofilm life cycle starts with planktonic cells adhering to a surface followed by the division and expansion of the formed microcolonies. This process culminates with the release of dispersal cells from mature biofilms, enabling them to search for new surfaces and continue the cycle. Among these stages, the attachment phase is the most studied. In *S. pneumoniae* genes associated with colonization and epithelial cell attachment to epithelial have been identified as crucial for biofilm formation [17]. Little is known about the dispersal phase and how these bacteria detach from the biofilms. Utilizing the assay developed in Chapter 2, we explore the phenotype of biofilm and dispersal cells *in vitro* and *in vivo* and upon exposure of antibiotic treatment.

This chapter summarizes a detailed phenotypic characterization of the three states of growth in the clinical strain TIGR4: planktonic, biofilm, and dispersal cells. *In vitro*, biofilm growth measurements confirm dynamics in the architecture of how bacteria form biofilms over four days. *In vitro*, dispersal cells detach from the biofilm as single cells and cell aggregates of various sizes. Indicating a mixture of phenotypes is characteristic of dispersal cells *in vitro*. Mice were inoculated with the same dose of planktonic, biofilm and dispersal cells to assess virulence and disease severity *in vivo*. Our results showed that all three environments successfully colonize the nasopharynx. However, samples inoculated with biofilm cells exhibit diminished virulence in the lungs and bloodstream. Additionally, it was observed that mice inoculated with dispersal cells became moribund as early as 14 hours post-inoculation, compared to 48 hours for planktonic cells. In conclusion, dispersal cells are in a ready-to-infect state, while biofilms are unable to cause disease.

After establishing clear *in vitro* and *in vivo* distinctions between biofilm and dispersal cells. We explored whether *S. pneumoniae* biofilms harbored tolerant cells. We subjected planktonic, biofilm and dispersal cells to high concentrations of antibiotics, confirming that *S. pneumoniae* biofilms harbor a larger

population of tolerant cells than their planktonic counterparts, evident in the higher proportion of live cells following antibiotic exposure. The biofilm lifestyle endows bacteria with properties to survive harsh perturbations. For example, the architecture and production of EPS act as a mechanical barrier for antibiotics to penetrate and reach bacteria at the bottom of biofilms. In multiple species such as *E. coli*, *P. aeruginosa* and *S. aureus* biofilms can withstand high concentrations of antibiotics [23, 36]. Tolerant cells are known for being the reason of relapsed infections in humans and are even considered to become the first step towards antibiotic resistance [31, 37] . Our study revealed that biofilms and dispersal cells have a high number of tolerant cells when compared to planktonic cells. This observation was supported by the prolonged duration of killing under antibiotic treatment.

Our work has contributed to a more comprehensive understanding of the characteristics of *S. pneumoniae* biofilms and dispersed cells. However, there are still several unanswered questions. For instance, is there a strain-specific tolerance phenotype? Are there differences in the number of tolerant biofilms and dispersed cells in response to vancomycin and ciprofloxacin treatment? What are the mechanisms underlying tolerance, and are these mechanisms consistent across planktonic, biofilm, and dispersal cells? We hope that our findings will serve as a foundation for the isolation of tolerant cells from *S. pneumoniae* biofilms, ultimately leading to a better understanding of the treatment of biofilm-associated infections, particularly those that result in relapse due to antibiotic tolerance.

3.4 Methods

***S. pneumoniae* biofilm growth, visualization, and quantification**

Biofilms were grown, visualized, and quantified as previously described in Chapter 2 (p. 32).

Dispersal cell size measurement was done using FIJI along with the plug-in particle size. At least 200 cells were analyzed per image.

Murine lung infection model and sample collection

Groups of eight 6–8-week-old Swiss Webster mice (Charles River) were anesthetized by isoflurane inhalation and challenged intranasally (i. n.) with 50ul of $\sim 2\text{-}5 \times 10^6$ CFU bacterial suspension in 1X PBS. For infections of mice with biofilms and dispersal samples, the bacterial population was recovered by scraping cells (biofilms) and pipetting out (dispersal cells), followed by centrifugation and three washes in 1XPBS to remove any dead cells. The inoculum was confirmed by serial dilution and plating on blood agar (BA). Mice were inspected every 12 hours for illness and morbidity by monitoring the presence of hunching, starry coat, and activity levels. Mice were euthanized by CO₂ asphyxiation at 12-14 hours post infections (h.p.i) for dispersal cell samples inoculated mice or 48 h.p.i for planktonic and biofilm for inoculated mice. Blood obtained by cardiac puncture, nasopharynx lavage, and total homogenized lungs were collected from each animal. Bacterial burden was determined by serial dilution and plating on BA. Mice were housed under a 12 h/ 12 dark/light cycle. The room temperature setpoint was 71 °F (± 22 °), and the humidity setpoint was 40%.

Ethics approval

Animal work was conducted in compliance with Boston College's animal core facility and the guidelines of the Institutional Animal Care and Use Committee (IACUC). Mouse experiments were performed under the approved protocol #2022-008-01 from the IACUC of Boston College.

Determination of relative minimal inhibitory concentration (MIC)

1 to 5 x 10⁵ CFU of mid-exponential bacteria in 100uL was diluted with 100uL of fresh SDMM medium with a single antibiotic to achieve a final concentration gradient of vancomycin (0.1-0.5 µg/mL) and ciprofloxacin (0.125-4.0 µg/mL), and in 96-well plates. Each concentration was tested in triplicate. Growth was monitored on a BioSpa plate reader at 37° C for 16 hours. MIC is determined as the lowest concentration that abolishes bacterial growth.

Tolerance assay

Planktonic tolerance experiments were done as previously described [33]. Mature biofilms were developed, and dispersal cells were recovered by centrifugation in 1.5ml tubes, and washed with 1XPBS. Fresh SDMM media with (+ABX) or without (-ABX) antibiotic was added to wells containing a developed biofilm and to harvested dispersal cells (Supplementary Figure 5). Samples were taken at different time points over a 24-hour period, including T₀ (starting population), washed with 1XPBS, and plated on BA for enumeration and survival killing curve construction. The number of surviving bacteria at each timepoint is divided by the T₀ population to determine the survival proportion at each timepoint.

$$\textit{Survival Fraction} = \frac{\textit{Number of Surviving Cells}}{\textit{Total number of cells before antibiotic treatment}} \times 100$$

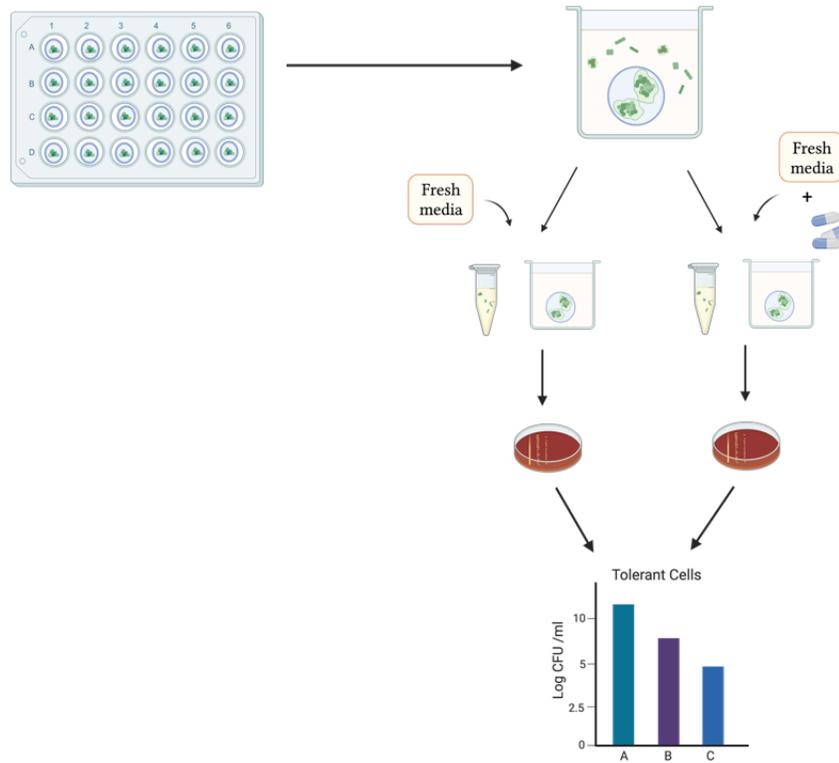


Figure 8. Summary of tolerant assay for biofilm and dispersal cells samples. In brief, matured biofilms and dispersal cells were either exposed with fresh media (control sample) or with fresh media + antibiotic (treatment samples). Samples were taken at different time points for bacterial enumeration and further tolerance analysis.

3.5 References

1. Joanna, W., B. Agata, and K. Emilia, *Biofilm formation of Streptococcus pneumoniae from bronchial alveolar lavage and from nasal swab*. Medical research journal, 2016. **1**(3): p. 5.
2. Gilley, R.P. and C.J. Orihuela, *Pneumococci in biofilms are non-invasive: implications on nasopharyngeal colonization*. Front Cell Infect Microbiol, 2014. **4**: p. 163.
3. Moscoso, M., E. Garcia, and R. Lopez, *Biofilm formation by Streptococcus pneumoniae: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion*. J Bacteriol, 2006. **188**(22): p. 7785-95.
4. Guilhen, C., C. Forestier, and D. Balestrino, *Biofilm dispersal: multiple elaborate strategies for dissemination of bacteria with unique properties*. Mol Microbiol, 2017. **105**(2): p. 188-210.
5. Barraud, N., S. Kjelleberg, and S.A. Rice, *Dispersal from Microbial Biofilms*. Microbiol Spectr, 2015. **3**(6).
6. Rumbaugh, K.P. and K. Sauer, *Biofilm dispersion*. Nat Rev Microbiol, 2020. **18**(10): p. 571-586.
7. Chua, S.L., et al., *Dispersed cells represent a distinct stage in the transition from bacterial biofilm to planktonic lifestyles*. Nat Commun, 2014. **5**: p. 4462.
8. Chua, S.L., et al., *In vitro and in vivo generation and characterization of Pseudomonas aeruginosa biofilm-dispersed cells via c-di-GMP manipulation*. Nat Protoc, 2015. **10**(8): p. 1165-80.
9. Chao, Y., et al., *Streptococcus pneumoniae biofilm formation and dispersion during colonization and disease*. Front Cell Infect Microbiol, 2014. **4**: p. 194.
10. Pettigrew, M.M., et al., *Dynamic changes in the Streptococcus pneumoniae transcriptome during transition from biofilm formation to invasive disease upon influenza A virus infection*. Infect Immun, 2014. **82**(11): p. 4607-19.
11. Tettelin Hervé , et al., *Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae*. Science, 2001. **293**(5529): p. 8.
12. Engholm, D.H., et al., *A visual review of the human pathogen Streptococcus pneumoniae*. FEMS Microbiol Rev, 2017. **41**(6): p. 854-879.
13. Li, J. and J.R. Zhang, *Phase Variation of Streptococcus pneumoniae*. Microbiol Spectr, 2019. **7**(1).
14. Chai, M.H., et al., *Proteomic comparisons of opaque and transparent variants of Streptococcus pneumoniae by two dimensional-differential gel electrophoresis*. Sci Rep, 2017. **7**(1): p. 2453.
15. Allegrucci, M. and K. Sauer, *Characterization of colony morphology variants isolated from Streptococcus pneumoniae biofilms*. J Bacteriol, 2007. **189**(5): p. 2030-8.
16. Allegrucci, M. and K. Sauer, *Formation of Streptococcus pneumoniae non-phase-variable colony variants is due to increased mutation frequency present under biofilm growth conditions*. J Bacteriol, 2008. **190**(19): p. 6330-9.

17. Domenech, M., E. Garcia, and M. Moscoso, *Biofilm formation in Streptococcus pneumoniae*. Microb Biotechnol, 2012. **5**(4): p. 455-65.
18. Munoz-Elias, E.J., J. Marciano, and A. Camilli, *Isolation of Streptococcus pneumoniae biofilm mutants and their characterization during nasopharyngeal colonization*. Infect Immun, 2008. **76**(11): p. 5049-61.
19. Shenoy, A.T., et al., *Streptococcus pneumoniae in the heart subvert the host response through biofilm-mediated resident macrophage killing*. PLoS Pathog, 2017. **13**(8): p. e1006582.
20. Jonczyk, M.S., et al., *Variation in Inflammatory Response during Pneumococcal Infection Is Influenced by Host-Pathogen Interactions but Associated with Animal Survival*. Infect Immun, 2016. **84**(4): p. 894-905.
21. Sanchez, C., et al., *Biofilm formation by clinical isolates and the implications in chronic infections*. BMC Infectious Diseases, 2013. **13**(47).
22. Marks, L.R., et al., *Interkingdom signaling induces Streptococcus pneumoniae biofilm dispersion and transition from asymptomatic colonization to disease*. mBio, 2013. **4**(4).
23. Hoiby, N., et al., *Antibiotic resistance of bacterial biofilms*. Int J Antimicrob Agents, 2010. **35**(4): p. 322-32.
24. Flemming, H.-C., et al., *Biofilms: an emergent form of bacterial life*. Nature Reviews Microbiology, 2016. **14**(9): p. 563-575.
25. Schrader, S.M., J. Vaubourgeix, and C. Nathan, *Biology of antimicrobial resistance and approaches to combat it*. Sci Transl Med, 2020. **12**(549).
26. Ciofu, O., et al., *Tolerance and resistance of microbial biofilms*. Nat Rev Microbiol, 2022. **20**(10): p. 621-635.
27. Stokes, J.M., et al., *Bacterial Metabolism and Antibiotic Efficacy*. Cell Metab, 2019. **30**(2): p. 251-259.
28. Uruen, C., et al., *Biofilms as Promoters of Bacterial Antibiotic Resistance and Tolerance*. Antibiotics (Basel), 2020. **10**(1).
29. Bakkeren, E., M. Diard, and W.D. Hardt, *Evolutionary causes and consequences of bacterial antibiotic persistence*. Nat Rev Microbiol, 2020. **18**(9): p. 479-490.
30. Yan, J. and B.L. Bassler, *Surviving as a Community: Antibiotic Tolerance and Persistence in Bacterial Biofilms*. Cell Host Microbe, 2019. **26**(1): p. 15-21.
31. Brauner, A., et al., *An Experimental Framework for Quantifying Bacterial Tolerance*. Biophys J, 2017. **112**(12): p. 2664-2671.
32. Bowler, P., C. Murphy, and R. Wolcott, *Biofilm exacerbates antibiotic resistance: Is this a current oversight in antimicrobial stewardship?* Antimicrob Resist Infect Control, 2020. **9**(1): p. 162.
33. Leshchiner, D., et al., *A genome-wide atlas of antibiotic susceptibility targets and pathways to tolerance*. Nat Commun, 2022. **13**(1): p. 3165.

34. Geerts, N., et al., *Antibiotic Tolerance Indicative of Persistence Is Pervasive among Clinical Streptococcus pneumoniae Isolates and Shows Strong Condition Dependence*. Microbiology Spectrum, 2022. **10**(6).
35. Wood, T.K., S.J. Knabel, and B.W. Kwan, *Bacterial persister cell formation and dormancy*. Appl Environ Microbiol, 2013. **79**(23): p. 7116-21.
36. Sharma, D., L. Misba, and A.U. Khan, *Antibiotics versus biofilm: an emerging battleground in microbial communities*. Antimicrob Resist Infect Control, 2019. **8**: p. 76.
37. Liu, J., et al., *Effect of tolerance on the evolution of antibiotic resistance under drug combinations*. Science, 2020. **367**: p. 5.

Chapter 4

Deciphering genetic requirements of *Streptococcus pneumoniae* biofilms.

The content of this chapter is adapted from the following manuscript:

Espinoza-Miranda S, Schmiege E, Chu C, Bodrog S, Rosconi F, van Opijnen T. Genome-wide identification of genetic requirements of *Streptococcus pneumoniae* biofilms. Manuscript in preparation.

Author contributions:

Suyen Espinoza-Miranda (SE), Federico Rosconi (FR) and Tim van Opijnen (TvO) conceived and design the study. SE, Emma Schmiege (ES) and Cedrick Chu (CC), Sophie Bodrog (SB) performed experiments. FR and SE wrote the manuscript. Tim van Opijnen (TvO) edited and approved the manuscript.

4.1 The value of high functional genomic approaches in bacterial research

In the past decade, high-throughput screening approaches have emerged as powerful tools for exploring bacterial behavior under different stresses, including exposure to antibiotics and their interaction with the host in a comprehensive manner[1]. These approaches can be combined with targeted experiments on a smaller scale to validate and deepen our understanding of the biological significance of these findings. Transposon-insertion sequencing (TIS) evaluates the genotype-phenotype relationship by combining transposon mutagenesis with high throughput DNA sequencing. Transposons are introduced randomly into bacterial genomes resulting in loss-of-function mutation via gene disruption. The collection of mutants is then propagated under the desired conditions. Genomic DNA is extracted and subjected to high-throughput sequencing, precisely identifying the location of transposon insertions in the genome. This method allows for the quantification of final insertions frequency compared to the frequency of the initial population [1, 2].

Transposon sequencing (Tn-seq), one of multiple published TIS approaches, has significantly impacted the field of bacterial genomics. Tn-seq uses the *mariner* transposon to create insertions it into thymine-adenine (TA) sites across the genome, followed by DNA extraction and high throughput sequencing. Tn-seq measures the contribution on non-essential genes under specific conditions, as cells with disrupted essential genes cannot survive [3]. Most of Tn-Seq studies have been carried out under planktonic conditions, and though they provide valuable information, they are insufficient to address bacterial growth under biofilm conditions. Recently, the first report on using Tn-seq to identify the determinants of motility during biofilm growth was conducted on *P. aeruginosa* [4]. Apart from this study, only a handful of other species have used Tn-seq to uncover genes important during biofilm formation, including *Enterococcus faecalis* [5], *Escherichia coli* [6], *Yersinia pestis* [7], *Staphylococcus aureus* [8] and *Mycobacterium tuberculosis* [9].

In the context of streptococcal biology, Tn-seq has revealed numerous novel genetic requirements and genetic interactions on a global and single-cell scale under multiple conditions [3, 10], including the addition of antibiotics [11, 12] and within *in vivo* infection models [3, 13]. More recently, it has provided groundbreaking insights into universally essential, strain-specific essential, and accessory essential genes in a large-scale pangenome analysis [14]. While individual genes crucial for *S. pneumoniae* biofilm formation have been identified through single-gene knockout experiments, a comprehensive evaluation of genes essential for both biofilm establishment and maintenance is lacking in the existing literature. This chapter fills this gap by presenting the first global identification of the genetic requirements for *S. pneumoniae* biofilm formation.

4.2 Time-lapse Tn-Seq uncovers genetic requirements among two *S. pneumoniae* strains grown as biofilms.

Two different strains, TIGR4 and BHN97, were selected to determine the genetic requirements of *S. pneumoniae* biofilms. Chapter 3 includes a comprehensive characterization of TIGR4 strain biofilm growth. BHN97 is a serotype 19F strain commonly used for otitis media *in vivo* models [15]. Both strains have similar rates of growth during planktonic growth (**Figure 1A**). However, striking differences were observed when biofilms were grown using the assay developed in this study. Compared to TIGR4, BHN97 produced more biomass, as well as thicker and taller aggregates than TIGR4 (**Figure 1B**). BHN97 biofilms grew more smoothly on the glass surface, as indicated by their low roughness coefficient (RA*) values (**Figure 1C**). PCA analysis using the information extracted from biomass, roughness coefficient, average and maximum thickness quantified from the imaging data showed that neither TIGR4 nor BHN97 could be grouped as low (BS71) or high biofilm (BS72) formers; instead, the replicates of each strain group together, and form their own subset separated from the low and high biofilm-former strains used in Chapter 2 (**Figure 1D**). This demonstrates that

the ability to grow as a biofilm varies among *S. pneumoniae* strains and that the high and low phenotype might not be sufficient to characterize biofilm formation capacities.

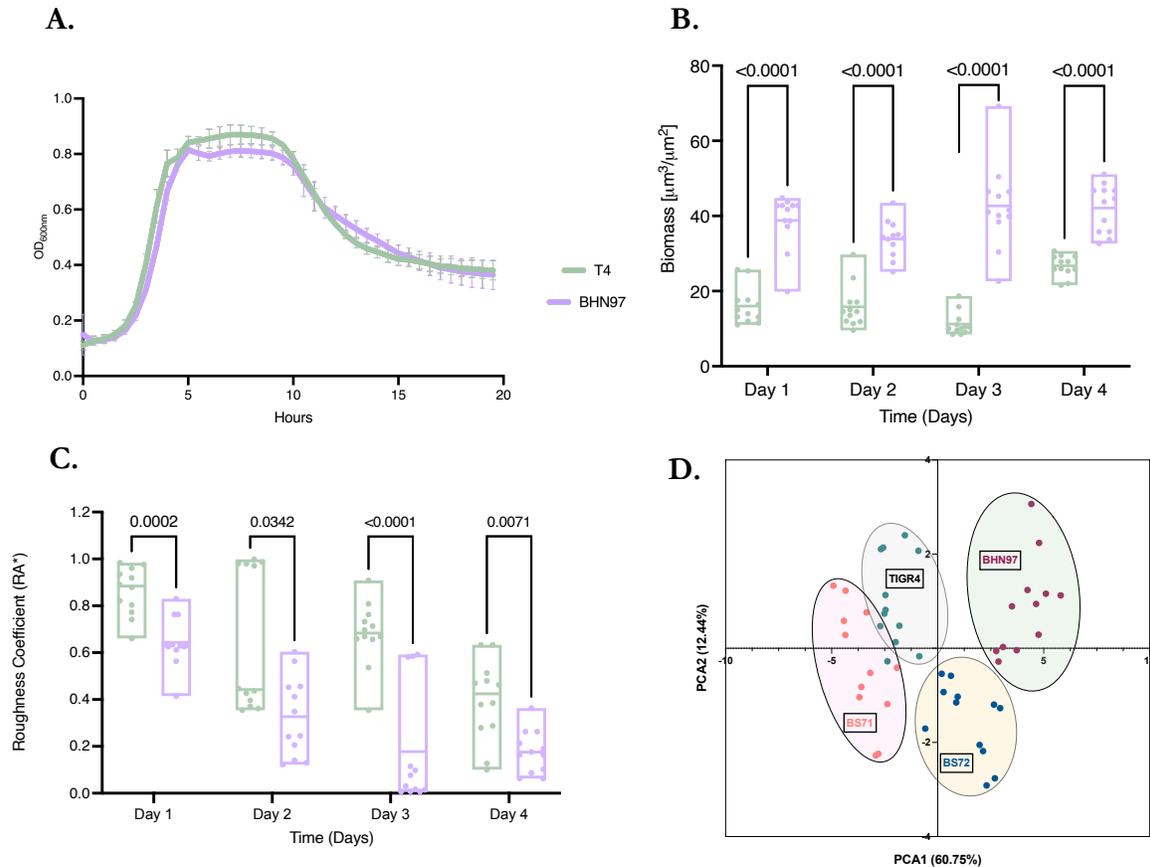


Figure 1. TIGR4 and BHN97 forms distinct type of biofilms. **A** Growth in liquid media shows no evident difference between the strains. Error bars: standard error of three biological replicates. **TIGR4** (light green symbols) and **BHN97** (lilac symbols) **B.** BHN97 produces more biomass when compared to TIGR4 over 4 days of growth & **C.** Roughness Coefficient for BHN97 is always lower for BHN97, indicating less heterogeneity in the biofilm architecture. Statistical analyses were done using two-way ANOVA followed Bonferroni post-test. **D.** PCA analysis of quantification parameters from BHN97, TIGR4, and the low (BS71) and high (BS72) strains used for validation of the methodology shows that imaging data is able to group biofilm replicates of the same strains by their ability to form biofilms.

To obtain a genome-wide view of genetic determinants important under biofilm formation conditions, Tn-seq was performed on growing biofilms using four independent highly saturated transposon mutant libraries of each strain TIGR4 and BHN97 constructed in our lab [14]. The inoculum used to seed the plates for the assay was used and served as the time zero (t1) for further analysis. Biofilms and dispersal cells were harvested at 4, 6, 12, 24, 48, 72, and 96 h and each sample was used as an independent t2 for further analysis (**Figure 2A**). The genomic DNA was sequenced and analyzed to determine the contribution of non-essential genes of *S. pneumoniae* to fitness during biofilm growth at each time point in (**Figure 2B** – TIGR4 :green, BHN97: lilac).

To determine which genes were either beneficial or deleterious for the formation or survival in biofilms, Tn-Seq fitness values were calculated as previously described [10, 16, 17] for each t2 independently. In this calculation, if the average fitness of all transposon mutants in a gene is close to 1, the inactivation of this gene indicates a neutral phenotype in relation to the wild type (WT), implying that the gene does not play a role in biofilm growth (**Figure 2A** grey dots). When the average fitness value is below 1, the lack of this gene results in reduced biofilm formation (pink dots), while an average fitness value higher than 1 indicates the mutant is a better biofilm former (teal dots). **Figure 2B** displays the average fitness values for all non-essential genes for biofilms across all time points for both strains.

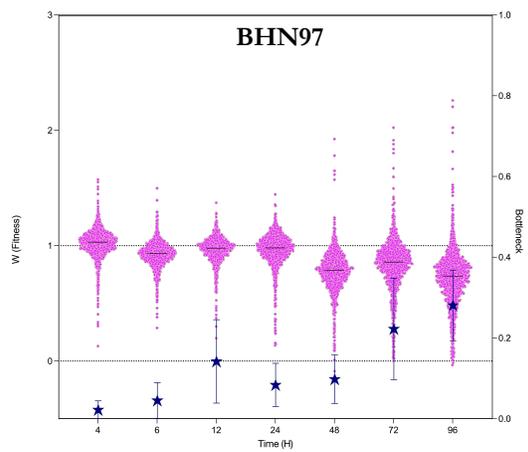
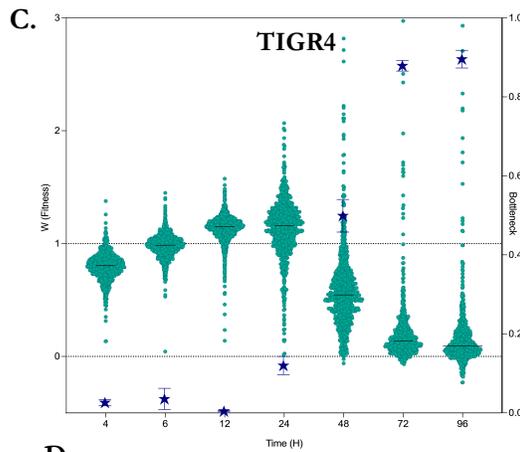
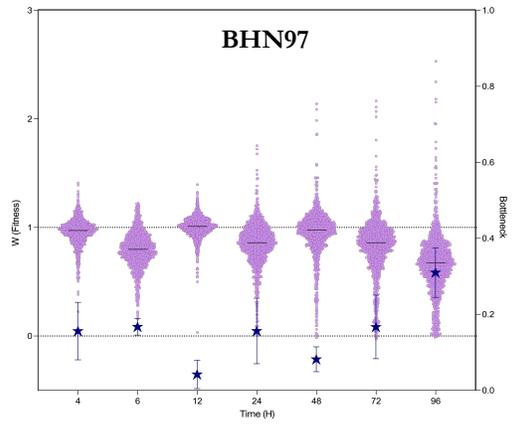
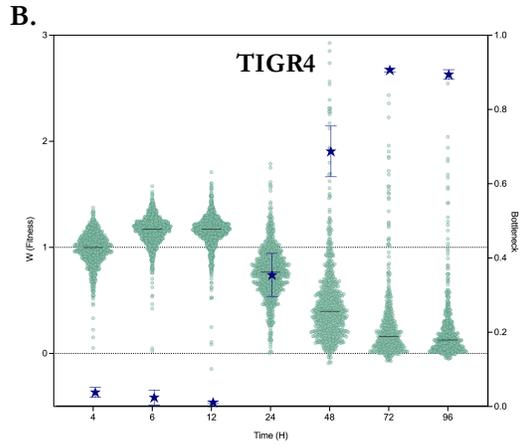
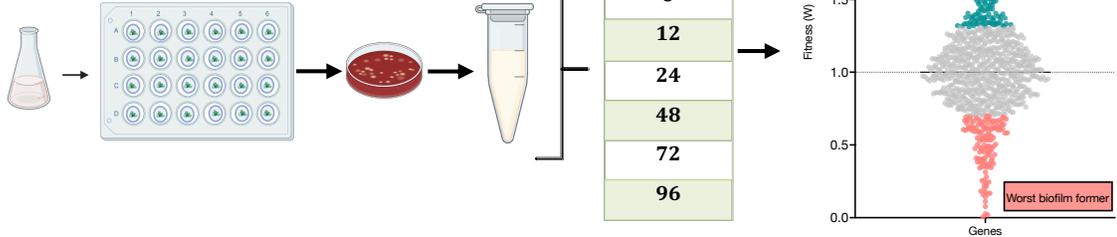
Biofilms can experience population bottlenecks during different stages of establishment as the dynamic attachment and dispersion of cells naturally changes the population size [18]. In nature, a sharp reduction in the number of viable cells in a population generates a drift in the genetic content [19]. This reduction is commonly referred to as “the bottleneck effect” [20]. In Tn-Seq experiments, bottlenecks randomly reduce the proportion of the transposon library, resulting in data variability and potential significant gene identification with misleading biological significance [21]. To overcome this limitation of the Tn-Seq tool, we estimate the strength of the bottleneck effect and correct the results accordingly.

To estimate the bottleneck effect in our Tn-seq dataset, we computed a bottleneck value for each t2, defined as the proportion of TA insertion sites with a positive count at t1 but 0 counts at t2. In other words, we calculate the proportion of mutants that disappeared at time 2 [22]. In TIGR4, the proportion of mutants lost at t2 differed from each time point; for 4, 6, and 12 h, the proportion of loss mutants was less than 10% and then increased drastically after 24 h, reaching a maximum of nearly 90% reduction in diversity at 72 h (Bottleneck values are depicted as star symbols- **Figure 2B & C**). In contrast, the fraction of lost mutants in BHN97 never exceeded 30% of the library diversity at any time point. The same trend in bottleneck values was observed for the dispersal counterparts, indicating that the random loss of mutants is not exclusive of surface-attached cells (**Figure 2C**). Assessing the number of viable cells after each day of biofilm growth for both strains showed a particular reduction of live cells for strain TIGR4 (**Figure 2D**). This reduction in the proportion of live/dead cells could explain the more dramatic changes in bottleneck values from Tn-Seq results in the strain TIGR4.

Our approach to mitigate the bottleneck effect relies on the following observations: 1) when bottleneck values are close to zero (no bottleneck), the median value of the distribution of average fitness for all genes is close to 1, and 2) this median decreases when the bottleneck effect increases (**Figure 2B & C**). Hence, we assume that a gene's average fitness close to the median indicates a neutral phenotype, lower than the median a deleterious phenotype, and higher than the median an advantage (**Figure 2B**). To identify those genes whose average fitness significantly deviates from a neutral-phenotype baseline, we compare the observed average fitness W of each gene to the overall median fitness. This comparison is grounded in the assumption that the median fitness serves as the most suitable representation of an expected value for neutral fitness. Genes that deviate significantly from the expected neutral value (median) are tagged as true or false discoveries by a one-sample t-test analysis and post hoc correction for multiple comparisons (false discovery rate of 5% using the method of

Benjamini & Hochberg). Genes tagged as true discoveries and with ΔW (W observed – median) with an absolute value higher than 0.3 were considered to be important for biofilms.

A.
Mutant library t1



D.

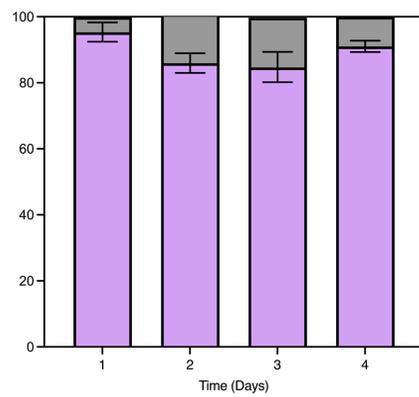
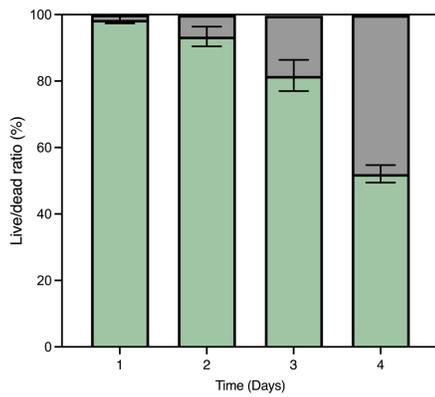


Figure 2. Tn-seq identifies genetic requirements for *S. pneumoniae* biofilm growth and maintenance. **A.** Overview of the experimental setup: transposon (tn) mutant libraries for **TIGR4 (green)** and **BHN97 (lilac)** were grown as biofilm using our assay. t1 represents the initial Tn library used to set up biofilm assay, and multiple t2 timepoints were taken for early and late stages of biofilm formation. Right panel shows how higher and a low biofilm former mutants look in our data set: teal: better biofilm former mutant ($\Delta W > 0.3$), pink: worst biofilm mutant ($\Delta W < -0.3$), and grey: mutants with no significant phenotype under biofilm growth ($|\Delta W| < 0.3$). **B.** Contribution of non-essential genes of *S. pneumoniae* to overall fitness during biofilm growth for TIGR4 and BHN97. Fitness for each timepoint is plotted in the left Y-axis, and the median is shown as a black bar. Right Y-axis displays calculated bottleneck values for each t2 (blue stars \pm SD). **C.** Same trends in fitness (W) and bottleneck distributions are observed for dispersal cells **D.** Percentage of live to dead bacteria for TIGR4 and BHN97 calculated from total biomass. Grey indicates dead cells.

Following the evaluation of fitness values for each t2, a total number of 186 genes were identified as important for biofilm growth in TIGR4. Among them, 126 genes are shared in more than two time points. **Figure 3** shows the distribution of positive and negative differences from significant genes across all time points in biofilm and dispersal cells for both strains. The loss of significant genes with negative impacts on for biofilm formation in TIGR4 at 72 and 96 hours might be caused by the large bottleneck value at these timepoints (**Figure 3A**). Large bottleneck values at these timepoints might be a result of the innate loss of viable cells in TIGR4 biofilms as reflected in **Figure 2D**.

A total of 218 genes were found to be significant under dispersal conditions for TIGR4, and 80 of these genes appeared at single time points. In BHN97, 378 genes were significant under biofilm conditions, but only 56 genes appeared at single time points. In addition, 271 genes were identified under dispersal conditions, with only 70 of the genes appearing at single time points. **Figure 4** shows the number of significant genes shared and unique among the different timepoints. For visual simplicity we show Venn diagrams for the first four timepoints (2, 6, 12 and 24h) and later timepoints (48, 72 and 96h) are depicted in a separate Venn diagram. These results show that there are multiple genetic requirements shared across the different stages of biofilm formation and between both strains.

However, the number of genes required at single time points suggest that forming biofilms might also be a complex dynamic process with multiple players at different stages of establishment.

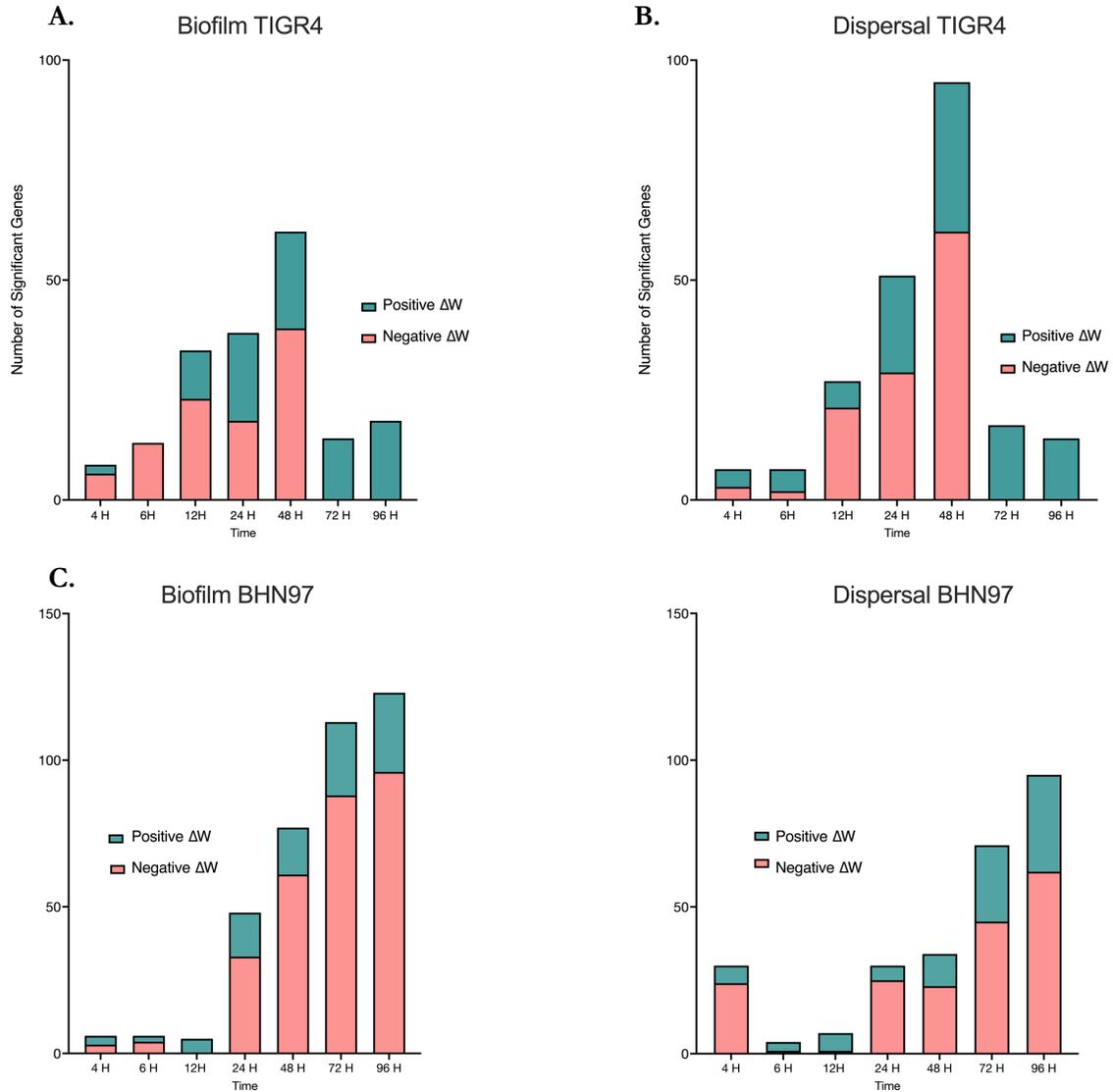


Figure 3. Distribution of significant genes across all time points. Each graph shows genes enriched under each condition indicated by a positive difference in fitness (teal) and genes with a detrimental effect showed by a negative difference in fitness (pink). Top panels show values for TIGR4 (A & B) and bottom panels for BHN97 (C & D).

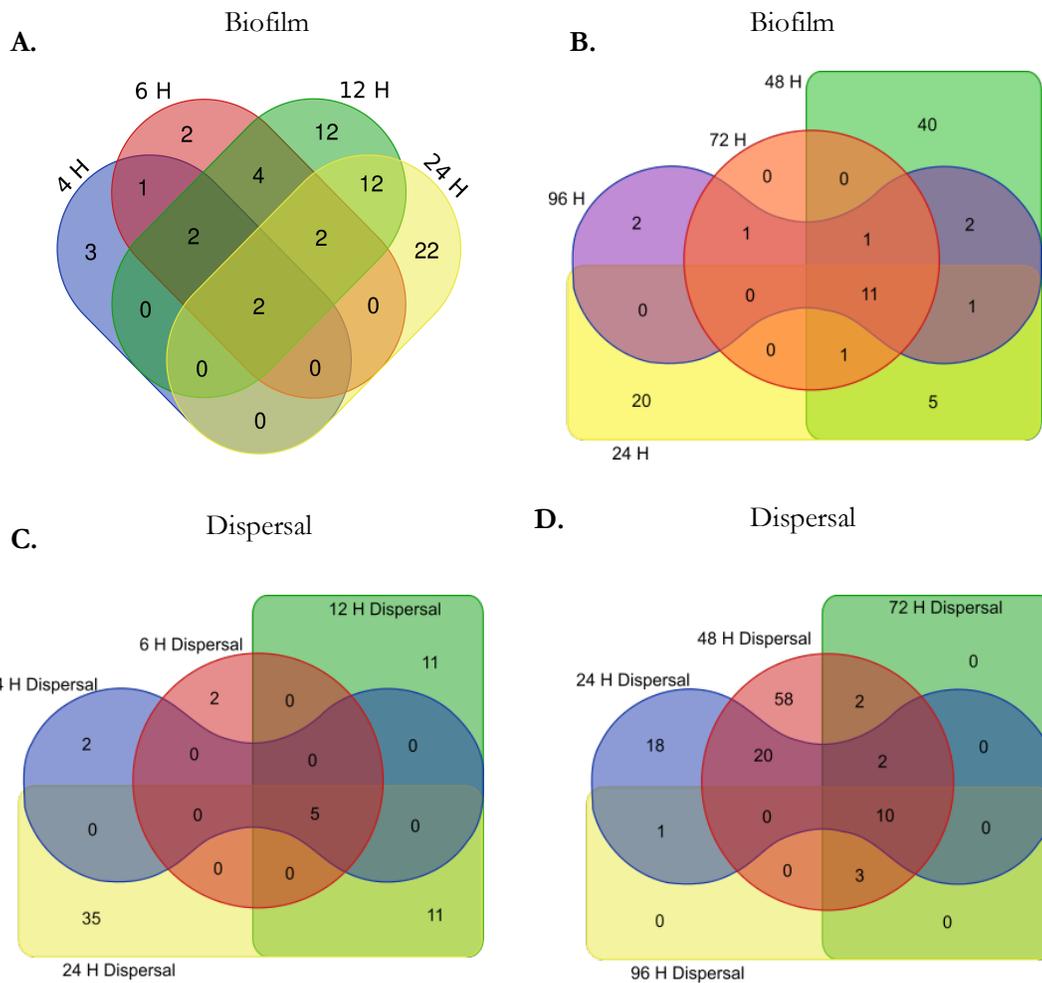


Figure 4. Venn diagrams show number shared genes across different time points from TIGR4. For simplicity the dataset was separated in early **A & C** (4,6,12 & 24 H) and late **B & D** (48, 72 and 96 H). Top panels show biofilm data and bottom panels display dispersal cells data. There are multiple significant genes that overlap across multiple timepoints. Tables 1 and 2 contain detailed information for the genes.

4.2.2 Genes contributing to biofilm formation are distributed across multiple functional categories.

To determine whether a particular process or pathway is specifically involved in contributing to biofilm growth at each t₂, the distribution of both beneficial and detrimental significant genes within each process was compared to the overall genomic distribution of the pathways. An adjusted p-value was calculated for each process, and processes p < 0.05 were considered statistically enriched. Tn-seq analysis identified that significant biofilm-related genes were distributed across multiple functional categories. Functional enrichment analysis revealed that a variety of pathways, such as cell wall metabolism and fructose and mannose metabolism, were consistently important for both strains during biofilm formation. Simultaneously, the dataset showed that capsule metabolism genes were critical for both strains during different stages of biofilm formation, confirming that impairment of the polysaccharide capsule enhances biofilm formation. Furthermore, this enrichment analysis also highlighted some strain dependencies on the genetic requirements during biofilm formation, such as the absence of genes involved in pyruvate metabolism pathways from any point in biofilm formation for BHN97 (**Figure 5**). The data also revealed the temporal dependencies of pathways. For instance, genes involved in purine metabolism are crucial for TIGR4 biofilms only during the first 24 h of growth, whereas in BHN97, they become relevant after 48 h.

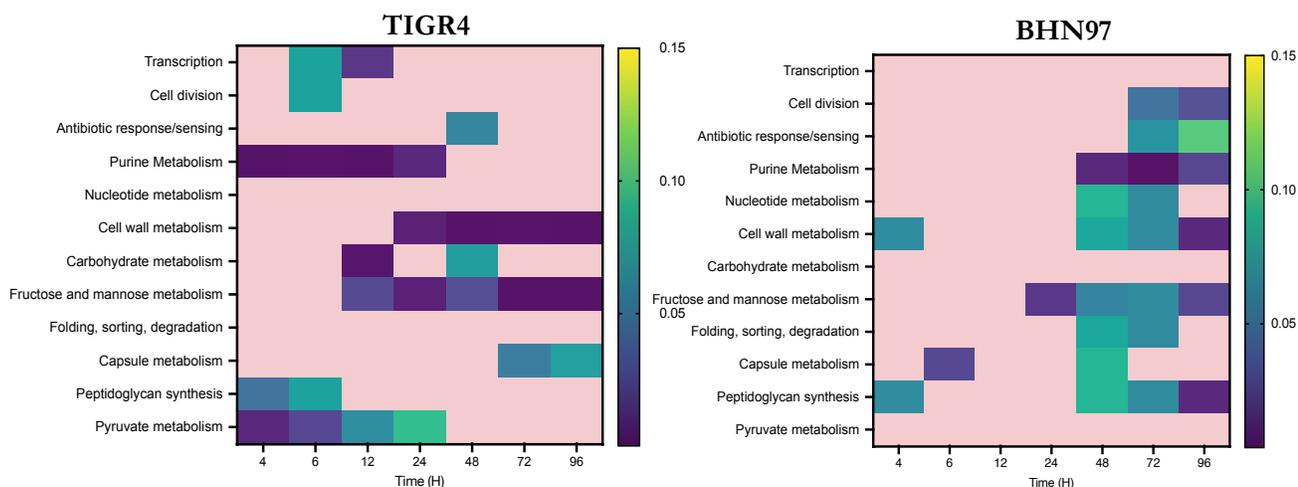


Figure 5. Tn-seq reveals functional gene categories important for biofilm formation. Enriched functional categories in biofilm individual datasets are presented as a heatmap with the corresponding p-values of false discovery rates. p-value scale is ranging from 0 to 0.15 shown in the figure key.

A closer examination of the data showed that the absence of certain genes in pathways such as peptidoglycan metabolism, cell division, and translation had a detrimental impact on biofilm formation by TIGR4, and this effect was observed as early as 4 h and persisted over time as expected since every t2 was compared to t1 independently. A clear example of these recurring genes over multiple t2 is the penicillin-binding protein A PbpA (coded by SP_0369) involved in peptidoglycan assembly in *S. pneumoniae* and ribonuclease Y RnaseY (coded by SP_1739), which is responsible for mRNA stability in the cell (**Figure 6**). In contrast, teichoic acid biosynthesis genes were enriched under biofilm conditions only for the strain TIGR4. Tables 1 and 2 in the supplementary information contain a summary of the shared and unique genes across time points.

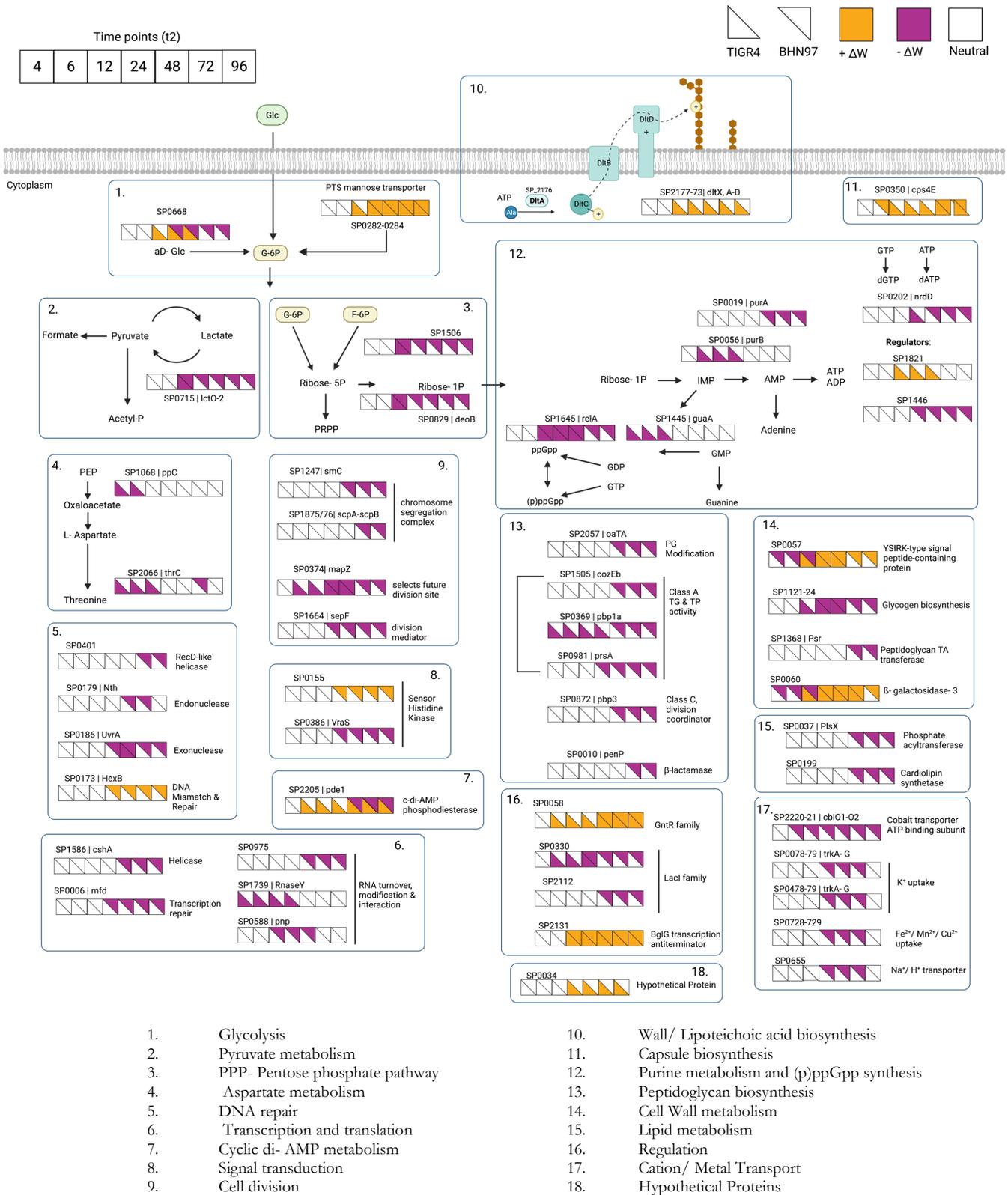


Figure 6. Detailed view of mutant genes and their contribution during biofilm formation. Each square represents a time point. The bottom half of the square indicates TIGR4, top half BHN97. White means no contribution (a mutant with neutral phenotype); purple means a negative difference

in fitness (low biofilm-former mutants), and yellow indicates a positive difference in fitness (high biofilm-former mutants). Gene designations for both BHN97 and TIGR4 are indicated.

Table 1. Selected genes for *in vitro* validation

Mutant	Strain	Gene product	Pathway	Biofilm validation	Resistant cassette
SP_0034	T4	Hypothetical	Hypothetical	More	CAT
SP_0058	T4	GntR family transcriptional regulator	Regulation	More	SPECT
SP_0060	T4	Beta-galactosidase 3	Carbohydrate	More	CAT
SP_0284	T4	PTS system, mannose-specific IIAB components	Membrane transport	More	CAT
SP_0350	T4	Sugar transferase	Capsule biosynthesis	More	CAT
SP_1645	T4	Putative guanosine-3',5'-bis(diphosphate)	Nucleotide metabolism	Less	SPECT
SP_1739	T4	Ribonuclease Y	Translation	Less	ERY
SP_2205	T4	DHH subfamily 1 protein	Ci-d- AMP	More	SEPC
SP_2176	T4	D-alanine--poly(phosphoribitol) ligase subunit 1	Wall teichoic acid	More	SPECT
SP_0350	BHN97	Sugar transferase	Capsule biosynthesis	More	CAT
SP_0975	BHN97	Ribonuclease R	Translation	Less	SPECT
SP_1645	BHN97	Putative guanosine-3',5'-bis(diphosphate)	Nucleotide metabolism	Less	SPECT
SP_2205	BHN97	DHH subfamily 1 protein	Ci-d- AMP	Less	CAT

4.2.3 Tn-seq validation

Biofilm growth progresses through different stages: attachment, cell growth/expansion (maintenance) and finally dispersion [23]. Therefore, we hypothesize that the bacterium requires a diverse set of genes to effectively form and maintain a biofilm. Thus, from the pool of significant genes extracted from Tn-seq dataset, we strategically selected candidates spanning from multiple cellular pathways for further investigation. After carefully examining the data, a total of 12 genes were selected for validation: nine genes in TIGR4 and four in BHN97 (Table 1). Genes selected were part of capsule (SP_0350, SP_0058) and carbohydrate metabolism (SP_0060), purine metabolism (SP_1645), cyclic di-AMP production (SP_2205), wall teichoic acid metabolism (SP_2176), membrane transport (SP_0284), translation (SP_1739) and one gene listed as hypothetical (SP_0034). These selected genes were shared across various time points, as well as among the surface-attached and dispersal cells and their expected fitness effect on biofilms was shared among positive and negative significance.

Following this, single gene knockout strains were constructed, and the biofilm-forming capacity of each mutant was assessed. Throughout the study, gene designations strictly followed TIGR4 locus names, and BHN97 genes were represented as homologous counterparts in TIGR4. In our experiment the significance of a mutant implies it could either enhance or impede *S. pneumoniae* biofilm formation. Table 1 has detail information of expected biofilm formation capacities for each selected gene. Consequently, to validate the anticipated phenotypes from Tn-seq, we grew each mutant strain in our previously develop biofilm assay, biofilms were stained, and images were acquired at day 1 and 4, as these two time points were found to contribute the better separation of the strain's ability to form biomass (Chapter 2, Figure 7).

A reliable indicator of biofilm production is the measurement of biomass. [24]. After measuring biomass on both day one and four for all selected mutants we observed that quantifying low biofilm

former strains on day 4 yielded more accurate results with respect to the mutant expected phenotype. For instance, the predicted low biofilm-former phenotypes for Δ SP_1645 and Δ SP_1739 were only validated after 4 days, as they exhibited biofilm levels comparable to the wild type (WT) after 1 day of growth. Specifically focusing on these two mutants (Δ SP_1645 and Δ SP_1739), it appears that their biomass remains constant over time without an increase (**Figure 7**). We successfully confirmed expected phenotypes predicted by Tn-seq, validating them in at least one of the two data points in TIGR4. For BHN97, the Δ SP_0350 mutant responsible for the initial glucose phosphate transferase in the sugar linkage to a lipid carrier for the capsule operon [25], did not exhibit a significant difference when compared to WT BHN97. We speculate that this lack of difference might be attributed to the existing proficiency of BHN97 as a biofilm former. While the absence of a capsule in other *S. pneumoniae* isolates contributes to enhanced biofilm formation, the relationship between capsule serotypes and biofilm formation remains incompletely understood. The other three mutant strains constructed in BHN97 were able to validate for their expected Tn-seq phenotype illustrated by their significant change in biomass when compared to WT (**Figure 7B**).

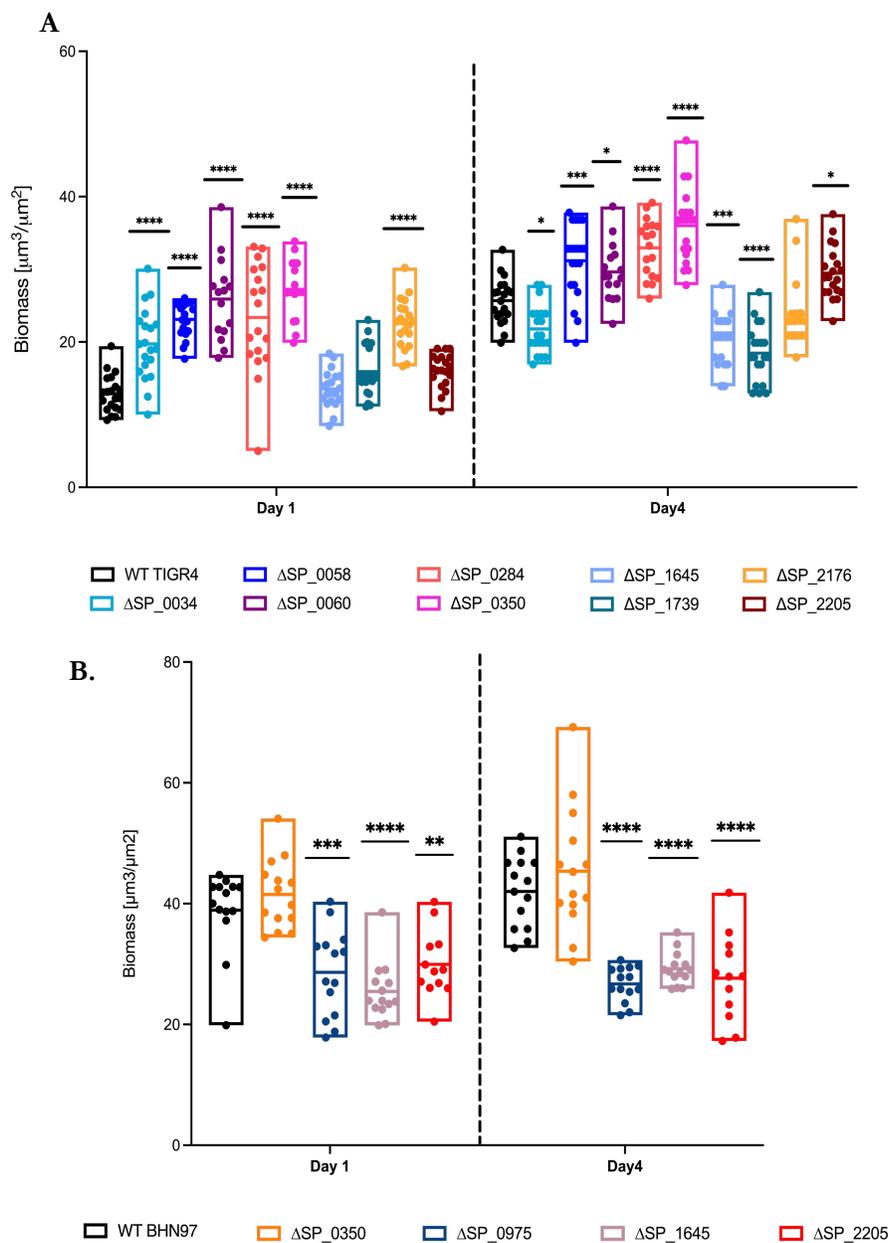


Figure 7. *In vitro* validations of selected mutants. Single knockouts were constructed, and biofilms grown using the methodology developed in Chapter 2. Biofilms were stained and quantified after 1 and 4 days of biofilm growth. **A.** WT TIGR4 and its respective mutants. **B.** WT BHN97 and its respective mutants. Biomass from each knockout is being compared against the biomass from TIGR4 wildtype. Statistical analyses were done using two-way ANOVA followed Bonferroni post-test. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $p < 0.0001$.

4.2.4 Capsule levels could explain some of the high-biofilm phenotypes of *Streptococcus pneumoniae* knockouts.

The polysaccharide capsule in *S. pneumoniae* is a major virulence factor, shielding the bacterium from host immune defense [26]. The absence of capsule facilitates adhesion to various surfaces including epithelial. A crucial inverse relationship exists, where the quantity of produced capsule is inversely proportional to the bacterium's capacity for biofilm formation [27, 28]. In essence, reduced capsule production corresponds to an increased biomass in the formed biofilm. In a previous screen low throughput transposon mutant screen, transposon insertions mapped to *cpsE* (a gene that encodes a glycosyl-phosphotransferase that transfers glucose-1-phosphate units which initiates the first cascade of capsule biosynthesis) were found and confirmed to be hyperbiofilm formers in the strain TIGR4 [29]. In agreement with this, our high-throughput Tn-seq data showed enhanced biofilm formation capacity when capsule biosynthesis genes were interrupted. After constructing a single knockout strain for the *cpsE* gene (Δ SP_0350), biofilms were grown and quantified on days 1 and 4. The Δ SP_0350 mutant produced more biomass than WT TIGR4 (**Figure 8**). In addition, the amount of biomass produced by Δ SP_0350 mutant was comparable to that produced by the unencapsulated version of the WT TIGR4-AC2394 [30] (**Figure 8A**). Suggesting that the disruption of one single gene in the capsule operon is enough to trigger the high biofilm phenotype. We consider the Δ SP_0350 confirmation of a high biofilm former as a positive control in our dataset.

Since capsule reduction increases for biofilm formation, we explored whether lower capsule levels could explain the observed high-biofilm phenotype of the selected genes with a beneficial fitness difference in biofilm formation. Capsule was quantified for all selected mutants by immunodot blot assay. Bacteria were grown to mid-exponential in SDMM, diluted to normalize cell number (same optimal density) and subsequently transferred to a nitrocellulose membrane. Blots were developed using unconjugated rabbit anti-serotype 4 and 19F serum for TIGR4 and BHN97 respectively,

followed by using a Cy5-conjugated goat anti-rabbit as a secondary antibody (**Figure 8B**). The amount of capsule produced by the low-capsule predicted strains, such as Δ SP_0350 and T4 Δ CPS (the TIGR4 unencapsulated version- used as a positive control), possessed approximately 85 % less capsule than WT TIGR4 (**Figure 8C**). The strain lacking the gene SP_0058, confirmed to have a high biofilm phenotype is annotated as a GntR family transcriptional regulator. Dot blot result show that Δ SP_0058 produces approximately 70% less capsule when compared to WT TIGR4 (**Figure 8C**). In D39, the product of SP_0058 is referred to as the cps locus repressor, CpsR, and can repress the capsule locus in the presence of glucose [31]. The observed reduction in capsule for Δ SP_0058 in TIGR4 suggests it may function similarly to *cpsR* in D39.

A second knockout revealing a significant decrease in capsule production compared to wild type was Δ SP_0284. This gene, part of the *ManLMN* operon, functions as a component of a phosphoenolpyruvate phosphotransferase system (PTS) responsible for transporting carbohydrates into the cell [32]. The PTS system operates as a complex cascade, orchestrating the transport and phosphorylation of the specified sugars [33]. Tn-Seq results showed transposon mutants *ManL*, coded by SP_0284, are enriched during biofilm growth at multiple time points. The Δ SP_0284 knockout presented increased biomass production per the *in vitro* validation. Immunodotblot assay showed a reduction in capsule production in Δ SP_0284 when compared to the wild type. While the contribution of this operon in *S. pneumoniae* biofilm is not previously established, in related species such as *S. mutants*, Δ Man-PTS is unable to form biofilms [34]. In contrast, in *Vibrio cholerae* and *E.coli* impairment of any component of Δ Man-PTS enhances biofilm formation [35]. We hypothesized that the high biofilm phenotype of Δ SP_0284 in our study might be also linked to a reduction in capsule levels.

When we measure capsule levels in BHN97 and its mutants, we confirmed that for Δ SP_0350 capsule levels are lower as expected. Similarly, lower capsule production was observed for Δ SP_0975, a mutant that lacks Rnase R, this understudied family of genes in *S. pneumoniae* have not been reported to exhibit

a relationship with capsule biosynthesis. Interestingly, this knockout is a low-biofilm former. The final =knockout displaying a significant reduction in capsule levels in TIGR4 was Δ ASP_2205, a gene involved in the regulation of cyclic di-AMP levels in the cell. There is no previous relationship studied between capsule production and cyclic di-AMP regulation in *S. pneumoniae*. However, cyclic di-AMP has been previously linked with biofilms [36, 37] a relationship we will explore further in the next section.

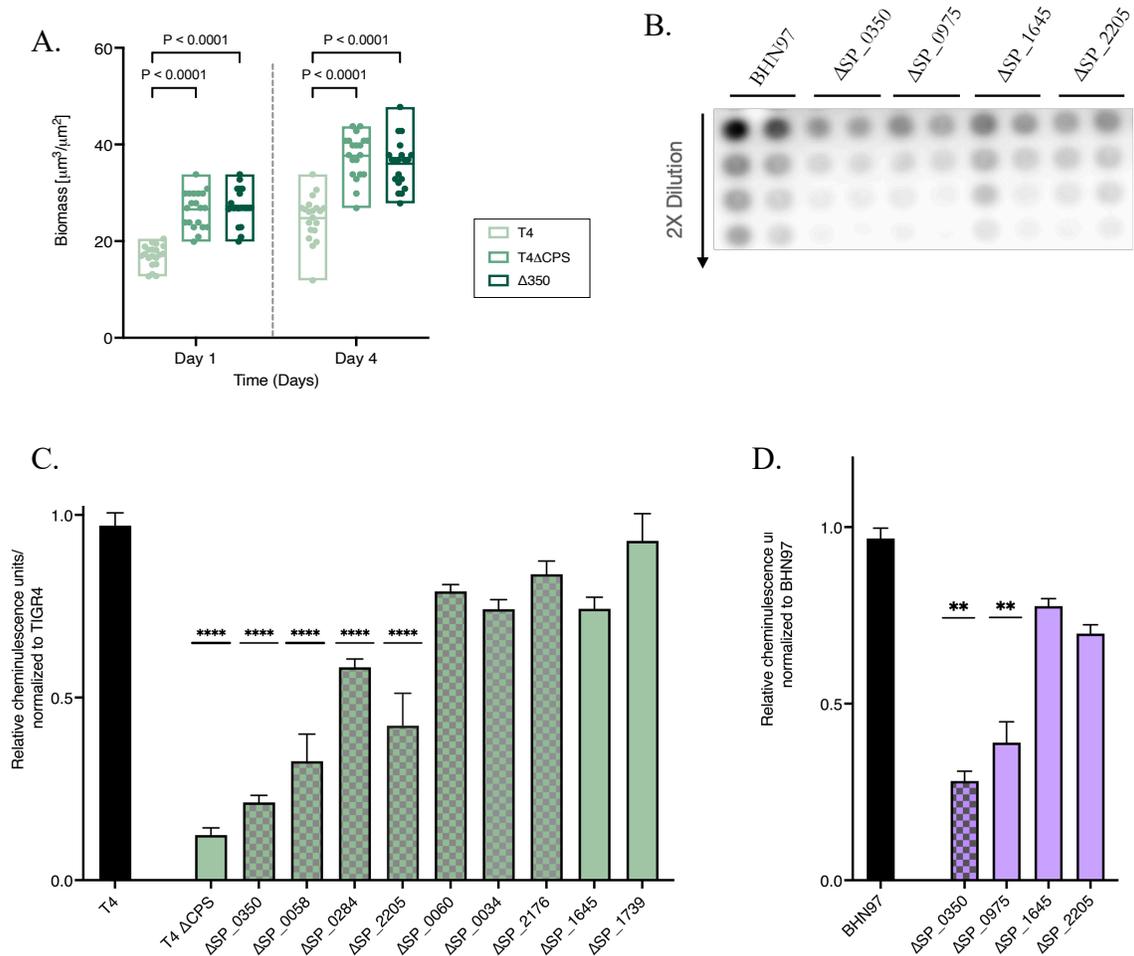


Figure 8. Reduction of capsule increases biofilm formation **A.** Biomass from capsule deficient strains and wild type. Capsule quantification via immunodotblotting. Capsule was isolated from OD₆₀₀-matched, mid-log phase and normalized by CFU. For each strain, 25 μ L spots of a 2-fold serial dilution were applied to a nitrocellulose membrane, which was developed using an unconjugated rabbit anti-serotype specific to the strain serotype, HRP-conjugated goat anti-rabbit antibody and ECL Blotting Substrate **B)** Data shown are from one blot representative of 3 independent experiments. **C. & D.** Capsule levels for TIGR4 and BHN97 strains. Data presented are the means \pm SD from 3 independent

experiments. Filled bars indicate knockouts with high biofilm-forming phenotype predicted by Tn-Seq. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $p < 0.0001$.

4.2.5 Cyclic di-AMP signaling in biofilms in *Streptococcus pneumoniae*.

While the association between the capsule and *Streptococcus pneumoniae* biofilm formation is well recognized, our Tn-seq approach successfully identified genes exhibiting heightened biofilm formation capabilities that are not necessarily linked to reduced capsule production. These include TIGR4-derived knockouts of SP_2176, SP_0034 and SP_0060. These mutants suggest that other important players might be key to the life cycle of biofilms. In the following section we explore some of the undescribed and novel ways of *S. pneumoniae* to form and maintain biofilm growth. More specifically we focus on potential role of secondary messenger molecule related genes and a global mRNA regulator.

Multiple signaling pathways in bacteria including biofilm formation and maintenance are regulated through cyclic dinucleotide second messengers[38]. Cyclic di-GMP (c-di-GMP) is the primary signal and is known to signal the transition between free-swimming growth and cell aggregation [39, 40]. While strong evidence has been provided on the key involvement of c-di-GMP molecules in multiple stages of biofilm formation, recent studies have indicated that c-di-AMP also has the ability to control biofilm growth [41]. In *B. subtilis* and *M. smegmatic* cyclic diadenosine monophosphate (c-di-AMP) plays a role in virulence by regulating the expression of genes involved in colonization [41]. Research performed in various model organisms suggests that c-di-AMP levels must remain within a certain range to avoid causing harm to the cell/colony, although some exceptions exist [42]. The impact of high or low c-di-AMP concentrations on biofilm formation varies by organism. In *Streptococcus mutans*, higher levels of c-di-AMP promote biofilm formation[42]. In *Streptococcus pyogenes*, biofilm formation also increases in knockouts with high levels of c-di-AMP [43], however no connection between ci-di-AMP and biofilm behavior has been reported in *S. pneumoniae*.

In *S. pneumoniae*, cyclic di-AMP is produced solely by the deadenylate cyclase CdaA (coded by SP_1096) and is degraded by the two phosphodiesterases Pde1 and Pde2 (coded by SP_2205 and SP_1298 respectively) [44-46]. Deletion of any of the genes involved in degradation results in increased c-di-AMP levels in the cell [44]. c-di-AMP influences bacterial growth, chain length, stops K⁺ uptake upon binding to the K⁺ transporter (Trk) family protein, and is essential for disease development in mouse models [44, 45, 47]. Our Tn-seq data revealed that when gene SP_2205 encoded by *pde1* is interrupted, biofilm biomass increases (**Figure 7A**). A Δ SP_2205 knockout validated the high biofilm former phenotype as well as exhibiting low capsule levels (**Figure 8C**). Our work indicates that lack of *pde1* in the cell might lead to reduce capsular polysaccharide levels.

To test the relationship between cyclic di-AMP levels and the ability of a mutant to form biofilms, intracellular levels of c-di-AMP were measured in the selected mutants and WT TIGR4 and BHN97 by performing a competitive ELISA in cell lysates derived from planktonic growth (**Figure 9A**). As expected, Δ SP_2205 accumulated more c-di-AMP than the WT, which is in agreement with previous reports [48, 49]. For mutants Δ SP_0058 and Δ SP_0034 (hypothetical protein) both form a better biofilm than Δ SP_2205 does, but they have lower c-di-AMP levels in the cells. C-di-AMP levels from BHN97 WT and mutants showed a different result (**Figure 9B**). Contrary to the high c-di-AMP accumulation levels from the TIGR4 Δ SP_2205, BHN97 Δ SP_2205 does not accumulate this secondary molecule in the cell. On the other hand, the capsule mutant (SP_0350) accumulates c-di-AMP substantially higher than WT. While, these experiments were performed using bacterial lysates from planktonic growth rather than biofilm or dispersal cells, we did not observe any clear correlation between the ability to form biofilm and total c-di-AMP accumulation in our collection of mutants.

To assess the relationship of c-di-AMP and biofilm behavior in *S. pneumoniae* more directly, we measured the levels of planktonic (at stationary phase) and three-day old biofilm and their respective

dispersal cells to identify any pattern in the production of this secondary molecule across the three modes of growth in WT TIGR4 and the $\Delta pde1$ (ΔSP_{2205}). The levels of cyclic di-AMP in biofilms and dispersal cells after three days of growth were almost five times higher than those in the planktonic population for the WT. In the case of ΔSP_{2205} , c-di-AMP levels in biofilms are 20 times higher than in planktonic cells. However, in dispersal cells, the levels of the secondary messenger did not change when comparing WT and ΔSP_{2205} (**Figure 9C**). This indicates that c-di-AMP accumulation occurs at higher rate in dispersal cells, and this accumulation is independent of SP₂₂₀₅ activity. Variation in intracellular c-di-AMP levels was noted across different growth modes, as evidenced by Wooten et al. In their study, the ΔSP_{2205} mutant exhibited levels below the limit of detection when grown on blood agar plates when compared to planktonic cells, indicating a likely mode-dependent fluctuation in c-di-AMP degradation[50].

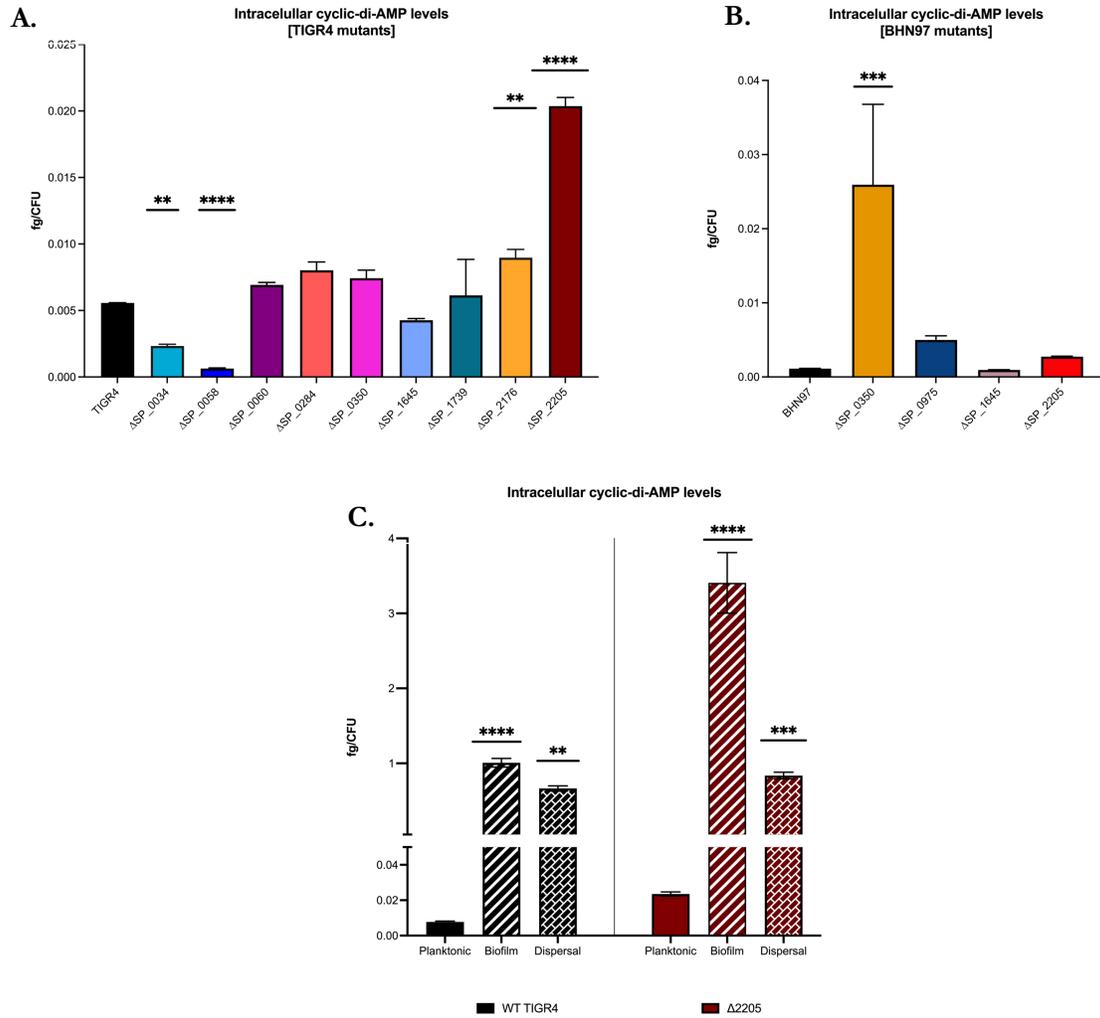


Figure 9. Cyclic di-AMP levels across all mutants. A competitive ELISA assay was performed to measure intracellular c-di-AMP levels in **A.** c-di-AMP measurements for all mutants in TIGR4 planktonic growth and **B.** BHN97 and its respective mutants. **C.** Three-day-old biofilms and dispersal for WT TIGR4 and Δ2205 (*pde1*). N=3 biological replicates and normalized to colony forming unit (CFU). Statistical analysis was performed using one-way Anova, and all mutants were compared to WT TIGR4 as a control group. * P≤0.05, ** P≤0.01, *** P≤0.001, **** p<0.0001.

4.2.6 *Streptococcus pneumoniae* biofilm-forming factors unrelated to reduced capsule levels.

Our genome-wide approach identified genes essential and detrimental under biofilm growth. In the previous section we explored the effect of the polysaccharide capsule in many of the selected mutants. In this section we mention some of the mutants are unrelated to capsule biosynthesis. One of the selected genes for validation is a member of the *dltABCD* (SP_2176-2173) operon in *S. pneumoniae*. This operon is responsible for encoding proteins involve in d-alanylation of teichoic acids, which leads to alterations of the cell surface charge [51]. Deletion of one member of the SP_2176-2173 genes enhanced bacterial attachment *in vivo* [52]. We showed that deletion of SP_2176 increased biofilm formation after 1 day of growth but showed no difference after 4 days (**Figure 7A**), indicating this operon might be important for the establishment of biofilms.

Cells growing within the complex architecture of biofilms must have the ability to adapt to limited nutrient and oxygen availability as well as mechanical stress [53]. Therefore, it is not surprising that genes involved in the stringent response pathway, cell wall integrity, and nutrient homeostasis appeared to be significant in the Tn-seq dataset. SP_1645 (*relA*) is responsible for the hydrolysis and synthesis of the (p)ppGpp molecule [54]. (p)ppGpp is a nucleotide that acts as an alarmone involved in the stringent response in bacteria. It is involved in regulating virulence gene expression and helps bacteria cope with nutrient-related stress [55]. When there is a shortage of amino acids, (p)ppGpp inhibits RNA synthesis and decreases translation in the cell, conserving the amino acids available in the environment [56]. Evidence suggests that bacteria devoid of (p)ppGpp are unable to swim, form biofilms, invade host cells, or resist innate immunity [57, 58]. A mutant strain lacking the *relA* gene was constructed, biofilms were grown, and biomass was quantified. Results showed that biomass at day 4 was less than that of WT TIGR4 (**Figure 7A**). This agrees with several previous studies that showed that lack of (p)ppGpp resulted in decreased biofilm formation in *Pseudomonas*, *Enterococcus faecalis*, and *S. aureus*

species [59, 60]. However, this is the first report of *relA*'s relationship with *S. pneumoniae* biofilm formation.

Another unexpected finding from our Tn-seq screen was SP_1739 (*rmv*). This protein is one of the conserved RNases found in *S. pneumoniae*. In bacteria, ribonucleases regulate transcript abundance, mRNA stability and degradation [61]. The role of *rmv* in biofilms have been studied in multiple species, in *P. aeruginosa*, RNA turnover regulates the expression of genes involved in EPS production and motility, which are essential for biofilm formation [62]. In *E. coli* biofilms, RNA turnover controls quorum sensing, a coordinated behavior that allows bacteria to communicate[63]. While the role of RNases in *S. pneumoniae* has recently attracted attention, there remains a gap in understanding their potential association with biofilm formation in this pathogen. It was recently shown that *rmv* in *S. pneumoniae* is essential for pathogenesis, as the deletion of the gene showed strong attenuation of virulence in invasive pneumoniae *in a vivo* model in the strain D39 [64]. SP_1739 is one of the few genes in our dataset that appears in the early time points and stays through the latter time points. We observed a low biomass production when growing this mutant in the biofilm assay (**Figure 7A**). In addition, our lab previously showed that deletion of SP_1739 is responsible for an increase of antibiotic tolerance [12].

To gain further insights into the distinctions between low and high biofilm formers, we analyzed the *in vivo* profiles of three specific knockouts. Among these, one knockout (SP_2205) displayed a high biofilm phenotype and was associated with decreased capsule production, and two other knockouts (SP_1645 and SP_1739), which we identified as low biofilm formers.

4.2.7 *In vivo* phenotype of knockouts with different biofilm formation capacities.

In Chapter 3, variations in *in vivo* infection outcomes were demonstrated depending on the source of inoculum used (planktonic, surface attached, dispersal). It was observed that disease induction is unattainable by surface-attached cells, while dispersal cells are found in a state ready to cause infection. With a focus on the three genes examined in the preceding sections—SP_1645, SP_1739, and SP_2205— and their respective knockouts showcasing diverse capacities for biofilm formation, an investigation into the disease-causing potential of these strains, characterized by low and high biofilm-forming abilities, was further pursued. Groups of 8 mice were intranasally inoculated with either planktonic, three-day- old biofilm, or three-day-old dispersal inoculum (**Figure 10**). Biofilm and planktonic inoculated animals were sacrificed 48 hours post-infection (h.p.i) while dispersal inoculated mice were sacrificed at 12 h.p.i., two hours less than the characterization done in Chapter 3, due to the high mortality rate observed with the Δ SP_1739 knockout. Infection was monitored at three sites, the nasopharynx, lungs, and blood (**Figure 10**), to examine aspects of colonization and invasive disease.

When mice were inoculated with dispersal cells from the low biofilm former SP_1739 knockout, bacterial titers were higher in the nasopharynx, while all cell states showed hypervirulence in the lungs. This suggests that this knockout generates cells highly adapted for lung infection, which can be confirmed by observing the increased bacterial titer even from the biofilm-derived inoculum. However, blood invasiveness phenotype is not different than WT. This phenotype observed from Δ SP_1739 in TIGR4 differs from the role in pathogenesis in D39 where lack of *maseY* attenuated pneumococcal growth in mice [64]. Mice inoculated with planktonic cells also exhibited a hypervirulent phenotype and higher titers of bacteria observed after sacrifice. This suggests that changes leading to a hypervirulent phenotype extend beyond the biofilm life cycle. In conclusion, the Δ SP_1739 strain is a hypervirulent strain that forms weak biofilms, and its hypervirulent-based mechanisms are lung-

restricted. This hypervirulent phenotype can be exacerbated when the infection develops from a dispersal inoculum.

Deletion of ΔSP_2205 , a high biofilm former strain, resulted in an attenuated phenotype in all the cell states, which is consistent with previous reports that SP_2205 is required to cause pneumococcal disease [65]. We attribute this phenotype to lower capsule reductions (**Figure 8**). For instance, in nasopharynx colonization where the capsule is not relevant since the immune response is not active, CFU numbers are similar to WT [66]. Surprisingly, the in animals infected with dispersal cells, the virulent phenotype is restored to WT. This phenotype restoration resembled the unchanged levels of c-di-AMP of WT versus ΔSP_2205 in dispersal cells (**Figure 9**).

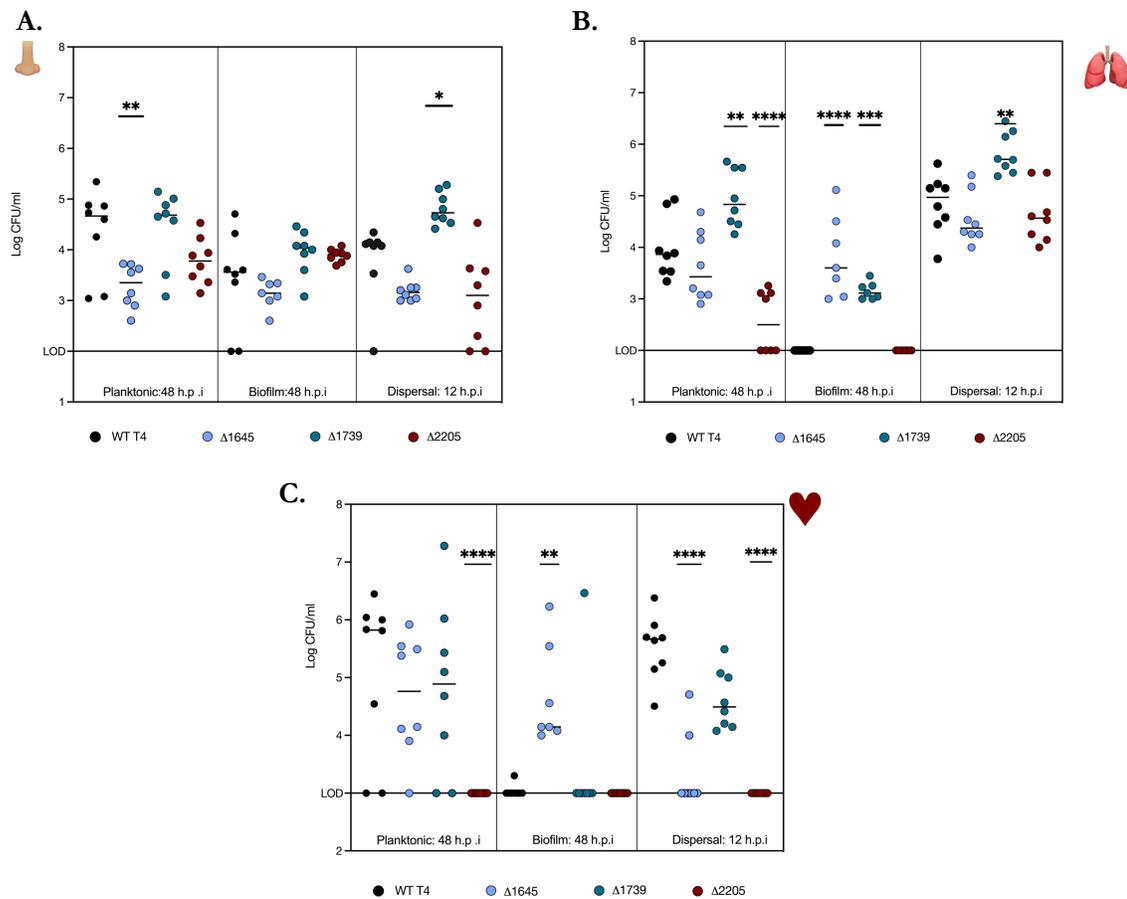


Figure 10. Low and high biofilm maker strains have different phenotypes *in vivo*. 6–8-week-old Swiss Webster mice were inoculated intranasally with either: planktonic bacteria ($\sim 5 \times 10^6$) and sacrificed 48

h.p.i, three-day old biofilms grow in vitro ($2-4 \times 10^6$) and sacrificed 48 h.p.i., three-day old dispersal cells ($2-4 \times 10^6$) and sacrificed 12 h.p.i. **A.** Nasopharyngeal lavage, **B.** lungs and **C.** blood were harvested for all mice experiments. One-way Anova was performed. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $p < 0.0001$.

Although mice inoculated with the low biofilm former ΔSP_{1645} in a planktonic state showed a decreased ability to colonize the nasopharynx compared to the ability of WT to colonize the nasopharynx using the same dose, this knockout colonizes the nasopharynx bacteria at the same rate across planktonic, biofilm and dispersal cells. Similarly, during lung invasion, we observed that biofilm-derived cells were as virulent as the planktonic inoculum. However, it was found that the dispersal inoculum from the ΔSP_{1645} presented a severely compromised blood invasion phenotype. These results confirm that dispersal cells rely on different mechanisms to achieve invasion compared to planktonic or biofilm derived cells.

4.3 Conclusions

We investigated at a genome-wide scale the genetic elements that influence biofilm formation, using Tn-seq on two strains with different biofilm formation capacities: TIGR4, a bacteremia model strain, and BHN97, an otitis media model strain. The transposon mutant libraries were grown in the assay previously developed on chapter 2 and sampled temporally to obtain a detailed resolution of the genetic requirements needed at different stages of biofilm growth and dispersion. This method proved to be highly effective in identifying both the functional categories that were shared and unique among the two strains, as well as the pathways that were enriched in a temporal manner.

A set twelve of genes was selected from various functional categories and their expected phenotype from Tn-seq was validated. Single knockouts were created, and biofilms were grown from each mutant strain using our previously developed biofilm assay described in Chapter 2. After measuring biomass on days 1 and 4, it was confirmed that each mutant strain displayed the expected Tn-seq phenotype at least once between the two time points. Since the reduction of capsule is a crucial factor in biofilm formation, capsule levels were measured in each strain and the results showed that not all strains had a deficiency in capsule production, suggesting that the capsule is not solely responsible for enhanced biofilm formation.

In our discussion, we highlight certain genes unrelated to capsule production. Notably, our research unveils, the cyclic di-AMP oscillation observed across various stages of biofilm growth. Furthermore, we delve into an exploration of the roles played by low biofilm formers Δ SP_1645 and Δ SP_1739 in biofilm formation. SP_1645, encoding for *relA*, emerges as a key player, as evidenced by its significance in various species like *P. aeruginosa*, where it is crucial for the interplay of biofilms with the environment [60]. It's worth noting that bacteria generate secondary metabolites acting as signaling molecules,

enabling communication and regulation of diverse physiological behaviors at the multicellular level, a phenomenon that includes biofilm formation.

We evaluated the influence of two low (Δ SP_1645 and Δ SP_1739) and high biofilm-forming (Δ SP_2205) strains on their ability to cause disease, when the inoculum was sourced from either surface-attached or dispersed cells. The inoculum from surface-attached low biofilm-forming strains showed improved migration to the lungs. We hypothesize that low biofilm formers might be less tightly packed in compared to high biofilm former strains resulting in better rates of dispersion which could explain the signs of disease in lungs when compared to TIGR4 and the high biofilm former strain. In addition, dispersal cells still exhibited hypervirulence, regardless of the presence of low- or high-biofilm strains when compared to samples from the planktonic inoculum. Findings from this work also confirmed that intracellular cyclic di-AMP regulation is necessary for disease to occur in mice irrespective of the type of inoculum.

In essence, the methodology explained in this chapter facilitates the identification of both established and previously undiscovered genes governing biofilm growth in *S. pneumoniae*, contributing to a more thorough comprehension of the intricate molecular network at play. The dataset we've generated stands as a valuable asset for future researchers, offering a foundation to delve into the genetic intricacies of biofilm development. Notably, the identified biofilm determinants hold promise as potential targets for the development of therapeutics focused on combating biofilm formation or dispersion.

4.4 Methods:

Tn-seq experiment and fitness analysis: TIGR4 and BHN97 transposon libraries previously constructed by Rosconi et al. were used for this study. In brief, libraries containing 10000–20000 insertions were constructed using Magellan6 [14]. Libraries were grown in our biofilm assay, and biofilm and dispersed bacteria were harvested at 4, 6, 12, 24, 48, 72, and 96 h. gDNA was isolated using the Qiagen DNeasy kit with a previous step of resuspending cells in lysis buffer (Tris 20mM pH 7.9; EDTA2mM; Triton 1.2%; deoxycholate 1.5%) with a 10-minute incubation at 37C. Post-gDNA processing and library preparation were performed as described previously. Sequencing analysis and fitness calculations were performed using Aerobio v2.3. To determine whether the fitness effects of a single mutant (W_i) are statistically different, four requirements had to be fulfilled[14]: (1) W_i was calculated from at least five data points, (2) median average fitness across all genes was calculated, (3) the observed W and expected W were compared by a one-sample t-test, and obtained P values were corrected by an FDR of 5%, (4) genes identified as true discovery and with a DW (W observed - W expected) higher than 0.3 were considered required and consequentially as high biofilm formers ($DW < -0.3$) or disadvantageous ($DW > 0.3$) low biofilm formers.

Mutant Generation and growth curves: Gene knockouts were created by replacing the coding sequence with either chloramphenicol (4 μ g/ml), spectinomycin (200 μ g/ml), or erythromycin (1 μ g/ml) resistance cassette, as previously described[14]. Primers used and knockout constructed are described in **Table 2**. Randomly selected colonies were selected, and gene knockout was confirmed via Sanger sequencing. To assess normal growth in mutant strains, growth was recorded for three biological replicates at OD₆₀₀ using 96-well plates on a Biospa 8 plate reader (Biotek) every 30 minutes for ~22 h period. To validate Tn-seq results, mutant strains were grown in our biofilm assay and quantification parameters via imaging as well as CFU were obtained as explained previously.

Table 2. Primers used in this study to generate knockouts.

Strain	Gene	Resistance cassette	Primer 1	Primer 2
TIGR4	SP_0034	CAT	AGTACTGGTGGAC TCTTGATT _{aac}	CATCAAGCTTATCGATACCGA GAAATGACATAAAAAACCTCC
TIGR4	SP_0284	CAT	ACTCATTATTGTC ATTCCCTCC	CATCAAGCTTATCGATACCGC TIGCGAATACTGTACAAC _{Tg}
TIGR4	SP_0060	CAT	TAGGACAGTAGG GAATTGTT	CATCAAGCTTATCGATACCGA TCTCGTATCTCAAATCGTG
TIGR4	SP_0058	SPECT	TAGGTGTAATCA GTGTAGC	CGGTATCGATAAGCTTGATT ACTAAGAAGAAAACCTG
TIGR4	SP_1645	SPECT	TCTCCAAT _{TCTTC} CTACCAT	CGGTATCGATAAGCTTGATT GTTGATCT _{TCTCGATTGTC}
TIGR4	SP_2176	SPECT	GATGCAT _{TAAAG} AACAGTCG	CGGTATCGATAAGCTTGATT GGTTTATTTGACACAATAGG
BHN97	EQH44_0486 5 (SP_0975)	SPECT	GGTTTGATTGTG TCCAGTAT	CGGTATCGATAAGCTTGATAC CTTCCCTTGCTTGTGTA
BHN97	EQH44_0794 5(SP1645)	SPECT	GCATCTCTACTCT CCAATTC	CGGTATCGATAAGCTTGAT GCTTTTCGCCTTCTAGATAA
BHN97	EQH44_1104 5(SP2205)	CAT	ATAATACCCCTCA AAAAGCG	CATCAAGCTTATCGATACCGG TIGTAAT _{TCCCCTAAGTCC}
BHN97	EQH44_0188 5 (350)	SPECT	CGTATIGTCTCTG TTTCAGT	CGGTATCGATAAGCTTGATTT TTCAATCCTTTTTCATCCA

Capsule quantification using immunodot blot assay: strains were grown in 3 ml of THY plus oxyrase (5ul/ml) to mid-log (OD600 of ~ 0.5–0.7), bacteria were pelleted, washed with 1X PBS, and stored at -20°C until use. Samples were normalized based on colony forming units (CFU), resuspended in cell wall lysis buffer (50mM Tris-HCl pH 7.5, 1 mg/ml lysozyme, 30U/ml mutanolysin), and incubated samples at 37°C for 30 min. Immunodotblot was done as previously described[66] with some modifications. Samples were sonicated (Branson, Inc) for 30s ON 15s OFF while on ice. Samples were serially diluted in Tris-buffered saline (TBS), and 20 µL was spotted on a 0.2µm-pore-size nitrocellulose membrane (Invitrogen, Inc.) with suction. Membranes were blocked with 20ml of 5% milk in TBS plus Tween-20 overnight. The membrane was then probed with rabbit anti-type 4 serum or anti-type 19F serum (Statens Serum Institut, 1:1000) in 10 ml of 2.5% milk in 1X TBST for 1h, at room temperature with shaking, followed by a second incubation with anti-rabbit HRP secondary antibody in 10 ml of 2.5% in 1X TBST for 1h at room temperature and incubated in the dark. The

membrane was washed three times with TBS, developed with EHCL substrate (MilliporeSigma™), visualized using a Typhoon FLA 9500 biomolecular imager (GE Healthcare), and quantified using ImageQuant TL.

Murine lung infection model and sample collection

Groups of eight 6–8-week-old were infected, and samples were collected as described in Chapter 3 (p.59)

c-di-AMP measurements by ELISA

Strains were grown in SDMM at 37C in a 5% CO₂ atmosphere up to 0.4 OD₆₀₀. Cells were pelleted and resuspended in 1 ml of 50mM Tris-HCl PH 8.0), sonicated for 40s (20s ON with 15s OFF), boiled for 5 min (95°C), and centrifuged for 5 min at high speed at 4°C. Lysates were centrifuged at 15 x000g for 5 min at 4°C, and clear supernatants were used for the cyclic-di-AMP assay following the manufacturer's instructions and a cyclic-di-AMP ELISA assay kit (catalog number 501960) from Cayman Chemicals. Cyclic-di-AMP levels of the strains were calculated based on the standard curve, normalized to the culture OD, and finally presented as rmole/OD.

4.5 References

1. van Opijnen, T. and H.L. Levin, *Transposon Insertion Sequencing, a Global Measure of Gene Function*. *Annu Rev Genet*, 2020. **54**: p. 337-365.
2. Burby, P.E., et al., *Implementation and Data Analysis of Tn-seq, Whole-Genome Resequencing, and Single-Molecule Real-Time Sequencing for Bacterial Genetics*. *J Bacteriol*, 2017. **199**(1).
3. van Opijnen, T. and A. Camilli, *A fine scale phenotype-genotype virulence map of a bacterial pathogen*. *Genome Res*, 2012. **22**(12): p. 2541-51.
4. Nolan, L.M., et al., *A global genomic approach uncovers novel components for twitching motility-mediated biofilm expansion in Pseudomonas aeruginosa*. *Microb Genom*, 2018. **4**(11).
5. Willett, J.L.E., et al., *Comparative Biofilm Assays Using Enterococcus faecalis OG1RF Identify New Determinants of Biofilm Formation*. *mBio*, 2021. **12**(3).
6. Holden, E.R., et al., *Massively parallel transposon mutagenesis identifies temporally essential genes for biofilm formation in Escherichia coli*. *Microb Genom*, 2021. **7**(11).
7. Kara R. Eichelberger, et al., *Tn-Seq Analysis Identifies Genes Important for Yersinia pestis Adherence during Primary Pneumonic Plague*. *mSphere*, 2020. **5**(4).
8. DeFrancesco, A.S., et al., *Genome-wide screen for genes involved in eDNA release during biofilm formation by Staphylococcus aureus*. *Proc Natl Acad Sci U S A*, 2017. **114**(29): p. E5969-E5978.
9. Jacob P. Richards, et al., *Adaptation of Mycobacterium tuberculosis to Biofilm Growth Is Genetically Linked to Drug Tolerance*. *Antimicrobial Agents and Chemotherapy*, 2019. **63**(11).
10. Thibault, D., et al., *Droplet Tn-Seq combines microfluidics with Tn-Seq for identifying complex single-cell phenotypes*. *Nat Commun*, 2019. **10**(1): p. 5729.
11. Jensen, P.A., Z. Zhu, and T. van Opijnen, *Antibiotics Disrupt Coordination between Transcriptional and Phenotypic Stress Responses in Pathogenic Bacteria*. *Cell Rep*, 2017. **20**(7): p. 1705-1716.
12. Leshchiner, D., et al., *A genome-wide atlas of antibiotic susceptibility targets and pathways to tolerance*. *Nat Commun*, 2022. **13**(1): p. 3165.
13. Rowe, H.M., et al., *Bacterial Factors Required for Transmission of Streptococcus pneumoniae in Mammalian Hosts*. *Cell Host Microbe*, 2019. **25**(6): p. 884-891 e6.
14. Rosconi, F., et al., *A bacterial pan-genome makes gene essentiality strain-dependent and evolvable*. *Nat Microbiol*, 2022. **7**(10): p. 1580-1592.
15. Rosch, J.W., et al., *A live-attenuated pneumococcal vaccine elicits CD4+ T-cell dependent class switching and provides serotype independent protection against acute otitis media*. *EMBO Mol Med*, 2014. **6**(1): p. 141-54.
16. van Opijnen, T. and A. Camilli, *Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms*. *Nat Rev Microbiol*, 2013. **11**(7): p. 435-42.

17. Tim, v.O., B. Kip, and C. Andrew, *Tn-seq; high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms*. Nature Methods, 2009. **6**: p. 6.
18. Pedroso, M., *The Impact of Population Bottlenecks on the Social Lives of Microbes*. Biological Theory, 2018. **13**(3): p. 190-198.
19. Li, H. and M.J. Roossinck, *Genetic bottlenecks reduce population variation in an experimental RNA virus population*. J Virol, 2004. **78**(19): p. 10582-7.
20. Chakraborty, P. and M. Kimmel, *Bottleneck Effect*, in *Encyclopedia of Genetics*, Sydney Brenner and Jefferey H. Miller, Editors. 2001, Academic Press., p. 233-235.
21. Cain, A.K., et al., *A decade of advances in transposon-insertion sequencing*. Nat Rev Genet, 2020. **21**(9): p. 526-540.
22. Surujon, D., *Tn-seq data processing scripts*. 2019.
23. Sauer, K., et al., *The biofilm life cycle: expanding the conceptual model of biofilm formation*. Nat Rev Microbiol, 2022. **20**(10): p. 608-620.
24. Corte, L., et al., *Biofilm Specific Activity: A Measure to Quantify Microbial Biofilm*. Microorganisms, 2019. **7**(3).
25. Bentley, S.D., et al., *Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes*. PLoS Genet, 2006. **2**(3): p. e31.
26. Wen, Z., et al., *Allelic Variation of the Capsule Promoter Diversifies Encapsulation and Virulence In Streptococcus pneumoniae*. Sci Rep, 2016. **6**: p. 30176.
27. Domenech, M., E. Garcia, and M. Moscoso, *Versatility of the capsular genes during biofilm formation by Streptococcus pneumoniae*. Environ Microbiol, 2009. **11**(10): p. 2542-55.
28. Domenech, M., E. Garcia, and M. Moscoso, *Biofilm formation in Streptococcus pneumoniae*. Microb Biotechnol, 2012. **5**(4): p. 455-65.
29. Munoz-Elias, E.J., J. Marciano, and A. Camilli, *Isolation of Streptococcus pneumoniae biofilm mutants and their characterization during nasopharyngeal colonization*. Infect Immun, 2008. **76**(11): p. 5049-61.
30. Price, K.E., N.G. Greene, and A. Camilli, *Export requirements of pneumolysin in Streptococcus pneumoniae*. J Bacteriol, 2012. **194**(14): p. 3651-60.
31. Wu, K., et al., *CpsR, a GntR family regulator, transcriptionally regulates capsular polysaccharide biosynthesis and governs bacterial virulence in Streptococcus pneumoniae*. Sci Rep, 2016. **6**: p. 29255.
32. Bidossi, A., et al., *A functional genomics approach to establish the complement of carbohydrate transporters in Streptococcus pneumoniae*. PLoS One, 2012. **7**(3): p. e33320.
33. Houot, L., et al., *The phosphoenolpyruvate phosphotransferase system regulates Vibrio cholerae biofilm formation through multiple independent pathways*. J Bacteriol, 2010. **192**(12): p. 3055-67.
34. Abranches, J., et al., *Different roles of EILABMan and EIIGlc in regulation of energy metabolism, biofilm development, and competence in Streptococcus mutans*. J Bacteriol, 2006. **188**(11): p. 3748-56.

35. Lazazzera, B.A., *The phosphoenolpyruvate phosphotransferase system: as important for biofilm formation by Vibrio cholerae as it is for metabolism in Escherichia coli.* J Bacteriol, 2010. **192**(16): p. 4083-5.
36. Rorvik, G.H., et al., *The c-di-AMP signaling system influences stress tolerance and biofilm formation of Streptococcus mitis.* Microbiologyopen, 2021. **10**(4): p. e1203.
37. Liu, C., et al., *cAMP and c-di-GMP synergistically support biofilm maintenance through the direct interaction of their effectors.* Nat Commun, 2022. **13**(1): p. 1493.
38. Valentini, M. and A. Filloux, *Biofilms and Cyclic di-GMP (c-di-GMP) Signaling: Lessons from Pseudomonas aeruginosa and Other Bacteria.* J Biol Chem, 2016. **291**(24): p. 12547-12555.
39. Dahlstrom, K.M. and G.A. O'Toole, *A Symphony of Cyclases: Specificity in Diguanylate Cyclase Signaling.* Annu Rev Microbiol, 2017. **71**: p. 179-195.
40. Jenal, U., A. Reinders, and C. Lori, *Cyclic di-GMP: second messenger extraordinaire.* Nat Rev Microbiol, 2017. **15**(5): p. 271-284.
41. Fahmi, T., G.C. Port, and K.H. Cho, *c-di-AMP: An Essential Molecule in the Signaling Pathways that Regulate the Viability and Virulence of Gram-Positive Bacteria.* Genes (Basel), 2017. **8**(8).
42. Peng, X., et al., *Cyclic di-AMP mediates biofilm formation.* Mol Microbiol, 2016. **99**(5): p. 945-59.
43. Tazin Fahmi, et al., *The Second Messenger c-di-AMP Regulates Diverse Cellular Pathways Involved in Stress Response, Biofilm Formation, Cell Wall Homeostasis, SpeB Expression, and Virulence in Streptococcus pyogenes.* Infection and Immunity, 2019. **87**(6).
44. Zarrella, T.M. and G. Bai, *Cyclic di-AMP Signaling in Streptococcus pneumoniae*, in *Microbial Cyclic Di-Nucleotide Signaling*, S. Chou, et al., Editors. 2020, Springer, Cham.
45. Bai, Y., et al., *Cyclic di-AMP impairs potassium uptake mediated by a cyclic di-AMP binding protein in Streptococcus pneumoniae.* J Bacteriol, 2014. **196**(3): p. 614-23.
46. Zarella, T., et al., *Bacterial Second Messenger Cyclic di-AMP Modulates the Competence State in Streptococcus pneumoniae.* Journal of Bacteriology, 2020. **202**(4).
47. Li, Z., et al., *Signal Transduction of Streptococci by Cyclic Dinucleotide Second Messengers.* Curr Issues Mol Biol, 2019. **32**: p. 87-122.
48. Jana, B., et al., *CRISPRi-TnSeq: A genome-wide high-throughput tool for bacterial essential-nonessential genetic interaction mapping.* bioRxiv, 2023.
49. Bai, Y., et al., *Two DHH subfamily 1 proteins in Streptococcus pneumoniae possess cyclic di-AMP phosphodiesterase activity and affect bacterial growth and virulence.* J Bacteriol, 2013. **195**(22): p. 5123-32.
50. Wooten, A.K., et al., *Unique Roles for Streptococcus pneumoniae Phosphodiesterase 2 in Cyclic di-AMP Catabolism and Macrophage Responses.* Front Immunol, 2020. **11**: p. 554.
51. Kiriukhin, M.Y. and F.C. Neuhaus, *D-alanylation of lipoteichoic acid: role of the D-alanyl carrier protein in acylation.* J Bacteriol, 2001. **183**(6): p. 2051-8.

52. Cooper, V.S., et al., *Experimental Evolution In Vivo To Identify Selective Pressures during Pneumococcal Colonization*. mSystems, 2020. **5**(3).
53. Marks, L.R., R.M. Reddinger, and A.P. Hakansson, *Biofilm formation enhances fomite survival of Streptococcus pneumoniae and Streptococcus pyogenes*. Infect Immun, 2014. **82**(3): p. 1141-6.
54. Kazmierczak, K.M., et al., *Roles of rel(Spn) in stringent response, global regulation and virulence of serotype 2 Streptococcus pneumoniae D39*. Mol Microbiol, 2009. **72**(3): p. 590-611.
55. Chau, N.Y.E., et al., *Emerging and divergent roles of pyrophosphorylated nucleotides in bacterial physiology and pathogenesis*. PLoS Pathog, 2021. **17**(5): p. e1009532.
56. Wang, B., R.A. Grant, and M.T. Laub, *ppGpp Coordinates Nucleotide and Amino-Acid Synthesis in E. coli During Starvation*. Mol Cell, 2020. **80**(1): p. 29-42 e10.
57. Haurlyiuk, V., et al., *Recent functional insights into the role of (p)ppGpp in bacterial physiology*. Nat Rev Microbiol, 2015. **13**(5): p. 298-309.
58. Irving, S.E., N.R. Choudhury, and R.M. Corrigan, *The stringent response and physiological roles of (pp)pGpp in bacteria*. Nat Rev Microbiol, 2021. **19**(4): p. 256-271.
59. Liu, H., et al., *Influence of (p)ppGpp on biofilm regulation in Pseudomonas putida KT2440*. Microbiol Res, 2017. **204**: p. 1-8.
60. Steinchen, W., et al., *Dual role of a (p)ppGpp- and (p)ppApp-degrading enzyme in biofilm formation and interbacterial antagonism*. Mol Microbiol, 2021. **115**(6): p. 1339-1356.
61. Barria, C., S. Domingues, and C.M. Arraiano, *Pneumococcal RNase R globally impacts protein synthesis by regulating the amount of actively translating ribosomes*. RNA Biol, 2019. **16**(2): p. 211-219.
62. Holmqvist, E. and J. Vogel, *RNA-binding proteins in bacteria*. Nat Rev Microbiol, 2018. **16**(10): p. 601-615.
63. Parker, A., et al., *Alternative pathways for Escherichia coli biofilm formation revealed by sRNA overproduction*. Mol Microbiol, 2017. **105**(2): p. 309-325.
64. Sinha, D., et al., *Pivotal Roles for Ribonucleases in Streptococcus pneumoniae Pathogenesis*. mBio, 2021. **12**(5): p. e0238521.
65. Cron, L.E., et al., *Two DHH subfamily 1 proteins contribute to pneumococcal virulence and confer protection against pneumococcal disease*. Infect Immun, 2011. **79**(9): p. 3697-710.
66. Shainheit, M.G., M. Mule, and A. Camilli, *The core promoter of the capsule operon of Streptococcus pneumoniae is necessary for colonization and invasive disease*. Infect Immun, 2014. **82**(2): p. 694-705.

Chapter 5

Development of antibiotic resistance in *S. pneumoniae* biofilms.

*Some contents of this chapter are taken from: Espinoza-Miranda S., van Opijnen T. and Cooper V. Causes and consequences of the evolution of genetic diversity in bacterial biofilms. 'In preparation' to be submitted to *Nature Reviews in Microbiology*.

5.1 Antimicrobial resistance in *Streptococcus pneumoniae*

S. pneumoniae is the main pathogen responsible for community-acquired pneumonia, meningitis, otitis media and bacteremia is associated with high mortality [1]. Antibiotics have been used to treat *S. pneumoniae* infections since the 1900s, but the overuse of these drugs led to the emergence of penicillin-resistant *S. pneumoniae* isolates. The emergence of antimicrobial resistant (AMR) *S. pneumoniae* isolates has become a global concern, leading to increased treatment failures, prolonged hospital stays, and is classified as a serious threat by the Center for Disease control and Prevention(CDC) [2]. According to the CDC and meta-analysis studies at least 40% of infections caused by *S. pneumoniae* were resistant to at least one antibiotic. And the most common resistance found isolates were to: penicillin, macrolides, fluoroquinolones and tetracyclines[3].

S. pneumoniae isolates commonly exhibit antibiotic resistance to β -lactams, macrolides, and fluoroquinolones. Multidrug-resistant (MDR) isolates resistant to three or more classes of antibiotics are increasing [4], highlighting the need to understand how these resistant strains emerge and discover alternative treatments for pneumococcus-related infections. Moreover, the prevalence of biofilms as the predominant mode of bacterial growth, their high resistance to antibiotics, and their ability to generate phenotypic and genotypic diversity complicate the issue of antibiotic resistance. Consequently, relying solely on finding alternative treatments for pneumococcus-related infections and developing novel antibiotics may not be sufficient to tackle this challenge. Instead, we suggest investigating *S. pneumoniae* biofilms in the context of antibiotic resistance as a more effective approach.

5.2 The role of adaptive diversification in biofilms towards antibiotic resistance

Many pathogens that cause chronic infections use biofilms to evade host defenses and antimicrobial compounds[5-8]. Bacterial biofilms are responsible for more than 80% of tissue and medical-device infections in hospitals[9]. Secreted polymers form a protective layer around cells and reduce the diffusion rates of antibiotics to sub-inhibitory concentrations [10]. This unique mode of bacterial growth has been categorized as a serious global health threat because of the extra protection they offer to bacteria against antibiotic treatments [11]. The spread of antibiotic resistance and the number of biofilm-associated infections have steadily increased the challenge of successful bacterial clearance in clinical settings [12-14].

Multiple genetic pathways and phenotypes that cause antibiotic resistance have been identified in planktonic growth conditions. However, such studies are inadequate to explain the evolution of resistance in surface-associated or aggregated microbes [15, 16]. Studying biofilm-associated populations *in vivo* is challenging for several reasons: these studies are low-throughput, expensive, there is high variation among subjects, and assessing biofilm formation inside the host is still not well established [17]. The discovery that mutants selected in laboratory models of evolving biofilms frequently resembled those commonly discovered in clinical isolates instead suggests that these models are suitable for identifying biofilm adaptations and perhaps also pathways to antimicrobial resistance. This has been confirmed by studies that have found similar patterns of resistance in both laboratory and clinical settings [18, 19]. Moreover, experimental evolution is a valuable approach for identifying new drug targets and estimating the frequencies of different mutations under selection. These new targets and relative rates enable the prediction of newly arising: antibiotic resistance, tolerance, and heteroresistance (referring to a bacterial population exhibiting a wide range of susceptibilities to antibiotics)[20] within the host [21-23].

The perks of experimental evolution

Experimental evolution offers the opportunity to track evolutionary changes in real time under controlled conditions. The sequencing of evolving populations can be performed at a low cost with the development of whole-genome sequencing [24]. Bacteria growing in biofilms can display higher mutations compared to planktonic growth [25]. Experimental biofilm evolution has been successfully achieved for species such as *E. coli*, *Pseudomonas fluorescens*, *Burkholderia cenocepacia*, *Acinetobacter baumannii*, and concluded that adaptation to biofilm results in the occurrence of multiple mutations that lead to longer fixation times and the maintenance of diversity in the environment (**Figure 1**) [19, 26]. There is evidence of repeatability and phenotypic convergence between in vitro and in vivo biofilm evolution experiments, suggesting that biofilms show similar strong selective pressures for development inside a host or in the laboratory [27]. Biofilm evolution studies can provide key insights for predicting evolutionary outcomes, identifying different paths to adaptation to antibiotics, biofilm maintenance and persistence, and designing therapies to successfully disrupt biofilms [28].

Studying the evolution of antibiotic resistance in biofilms can provide insights into new mechanisms of resistance by which bacteria can survive chronic infections. For example, experimental evolution of *Acinetobacter baumannii* biofilms to ciprofloxacin showed more beneficial mutations for the bacteria and exhibit cross resistance to other types of antibiotics [29]. This chapter seeks to uncover the mechanisms behind the emergence of genetic diversity during *S. pneumoniae* biofilm growth and the evolution of antibiotic adaptation in our biofilm model.

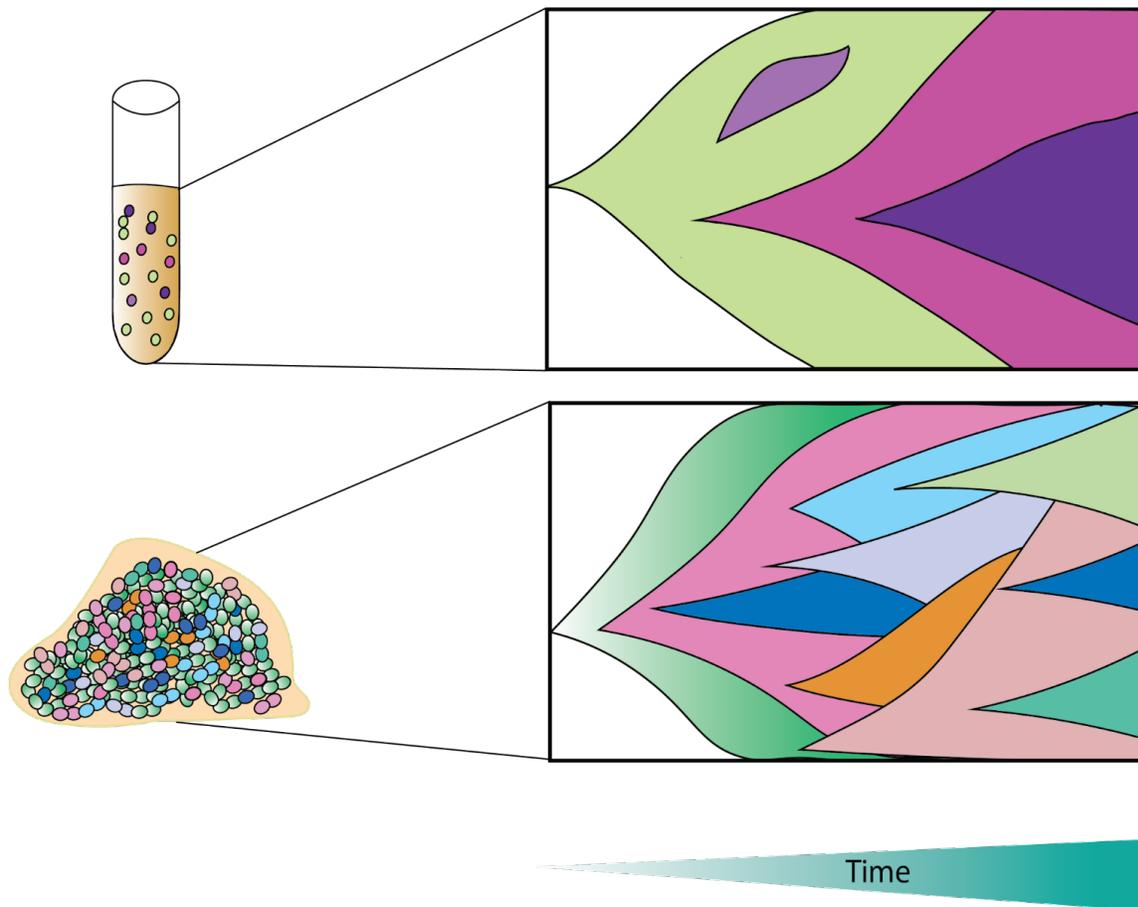


Figure 1. Consequences of biofilm growth for population-genetic diversity. Planktonic growth involves mass-action competition, in which few mutants contend for fixation, while the structured environment of biofilms gives rise to many contenders, resulting in sustained diversity. Each color indicates an emerging unique genotype in the population, and the width reflects its frequency.

5.3 Results

5.3.1 Identification of genomic factors strongly favor biofilm selection in *S. pneumoniae*.

In this chapter, we aim to expand our knowledge of the genetic factors that are most likely to be favored in *S. pneumoniae* biofilms over 40 days. Experimental evolution of bacterial biofilms requires the development of a proper assay that allows biofilms to grow for longer periods of time. We grew and passaged five independent biofilm populations of strain TIGR4 utilizing the assay described in Chapter 2 for 40 days. At every cycle passage (every four days), we scraped cells and isolated genomic

DNA, followed by Whole-Population Genome Sequencing (WPGS). We sequenced 10 data points for each population to create a detailed timeline of the emergence and fates of mutations. By comparing the sequenced populations with the genome sequence of the ancestor strain TIGR4, we identified variants that emerged during the process of biofilm growth using the breseq pipeline[30].

A total of 257 mutations were detected across all populations, with frequencies ranging from 5-100%. In this dataset, evolved mutations are defined as those that meet two criteria: 1) a frequency higher than 10%, and 2) filtering out background mutations, which are pre-existing in the parental strain or are media dependent. To accomplish this, the parental strain grown in SDMM was used as a control. As a first step to identify the genetic changes that are most likely to drive adaptive phenotypic changes, we focused on mutations that occurred at a frequency of over 50% and were present only in the final populations. Most of the variants detected were single nucleotide polymorphisms (SNPs, as shown in **Figure 2A**), although deletions ranging from 20 to 372 base pairs (bp) were also identified. Additionally, single nucleotide insertions were found and mostly located in intergenic regions. A total of 30 mutations over 50% frequency were observed across all populations, and Table 1 in the supplementary information provides a list of all mutations for the five adapted populations for biofilm growth.

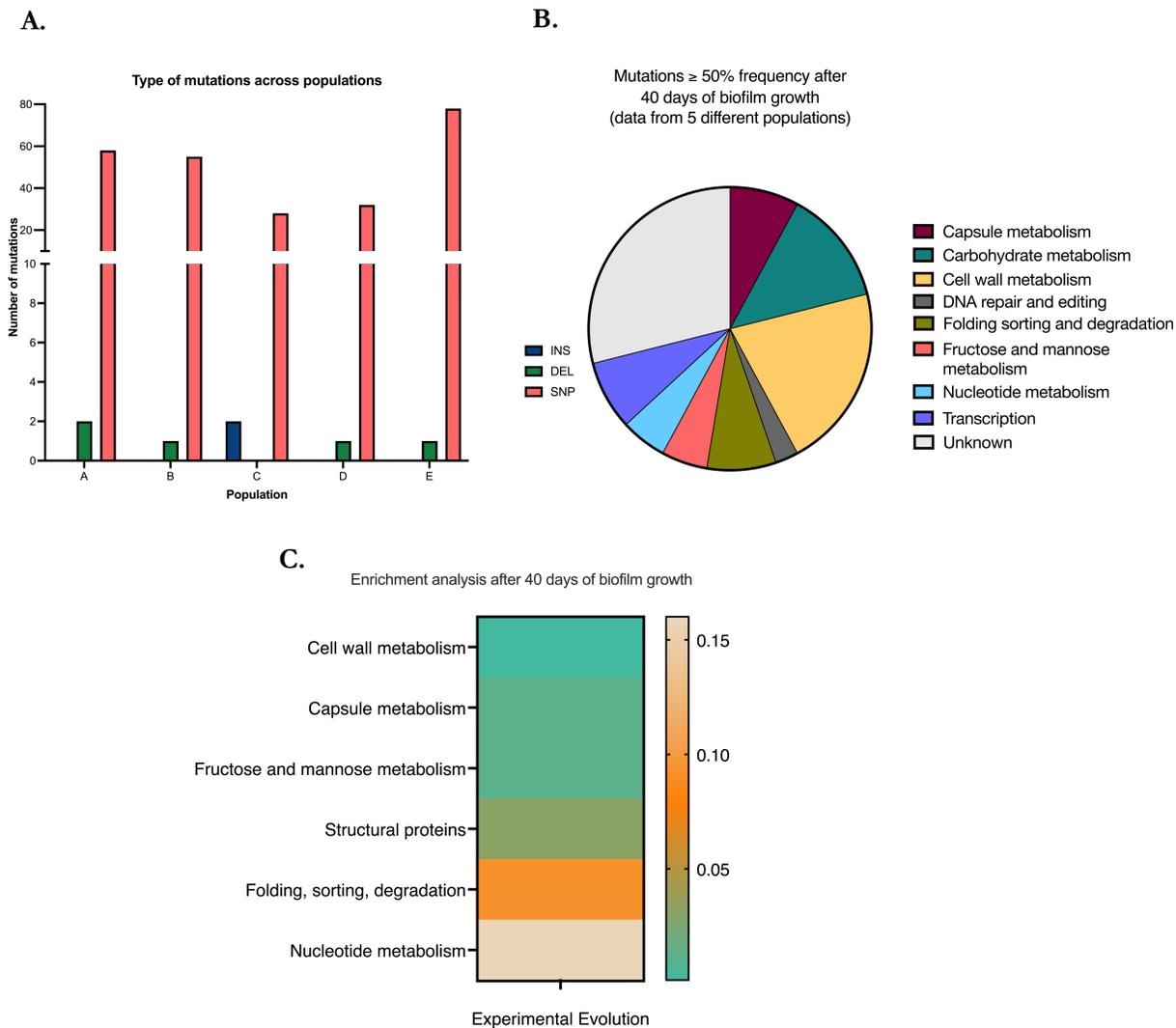


Figure 2. Summary of variants detected during 40 days of biofilm growth. **A.** Number and type of mutations detected per population (this includes all sequenced time points). **B.** Pie chart indicates the functional category of mutated genes over 50% at day 40. **C.** Enrichment analysis displays cellular pathways significantly enriched under biofilm growth.

Enrichment analysis of variant data demonstrated that genes with roles in capsule, carbohydrate and cell wall metabolism are predominant (**Figure 2C**). Additionally, genes associated with transcription, nucleotide metabolism and DNA repair and editing constitute a percentage of the total mutated genes (**Figure 2B**). It is also worth noting that we found variants in genes with unknown functions. To

determine if the five adapted populations (Pop A-E) were better at forming biofilms than WT, we tested their ability and grew each population in the assay developed in Chapter 2. We then quantified the biofilms using microscopy and compared the results after one and four days of growth. Pop E was found to be a superior biofilm former, as its biomass significantly increased compared to the other populations and WT after one day of growth. Additionally, after four days of growth, Pop A displayed a higher production of biomass, similar to Pop E. The rest of the populations were found to have biomass expansion similar to WT (**Figure 3A**). When we switch our attention to looking at roughness coefficient, we see that all populations have a lower roughness coefficient indicating that there is less variability in the thickness of these biofilms (**Figure 3B**). It is important to note that SNPs, deletions and stop mutations were detected in all 5 different populations, therefore Pop A Pop E were not exclusively one type. **Figure 6** shows the genotype trajectories of the other three populations (B,C, and D).

To determine whether the high biofilm phenotype displayed by Pop A and Pop E was a population-wide phenomenon, we selected two clones from each of the two population by randomly choosing single colonies post growing the population on blood agar plates. All the clones were successfully grown in our biofilm assay, and their biomass was compared to that of the parental strain TIGR4 and the populations they originated from. We were able to determine that individual clones displayed the same high biomass phenotype as their respective population. No significant differences were observed among clones isolated from the same population (**Figure 3 C& D**). Clones were also sequenced to attempt to establish a connection between the high biomass production phenotype and the genotype. Clones from adapted populations have the advantage of being isogenic populations, where all mutations are fixed.

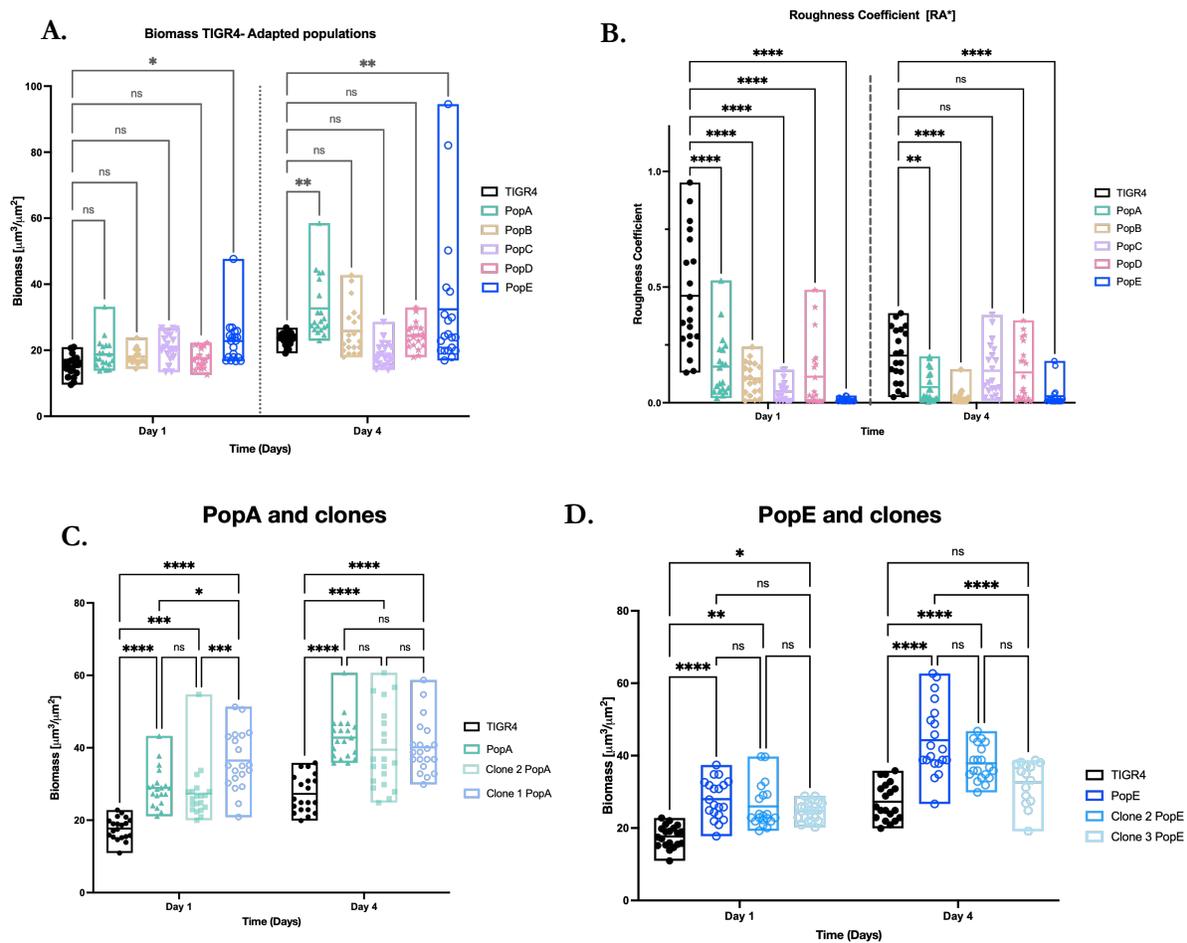


Figure 3. Biofilm growth dynamics of adapted populations. All populations were grown in our biofilms assay and quantified for day 1 and 4 of growth. **A.** Biomass of all five adapted populations. **B.** Roughness coefficient indicates all biofilms from populations exhibit less variability in their thickness. **C & D.** Biomass from selected clones confirm high biofilm production phenotype. Twenty Z-stacks were analyzed for each sample. For all graphs “ns” denotes non-significant comparisons, **** $p < 0.0001$ using Two-way ANOVA followed by Bonferroni post-test.

of evolution. **C.** Occurrence of fixed mutations in clones at day 40 and their presence on the population.

The mutation in *dltC* (T65I) in Pop A, also present in the two clones, quickly overtakes during adaptation and proceeds to fixation, as shown in the Muller plot (**Figure 4 A& B**) depicting genotype frequencies[31]. This SNP was also found at 65% and 87% frequency in Populations B and D, respectively. This suggests that there is some degree of genetic parallelism in *S. pneumoniae* biofilms where genes or identical mutations are present in independent populations. Parallelism is not limited to biofilm evolution, as it can also be observed in planktonic adaptation experiments[28]. For biofilms, parallelism has been reported for multiple species such as *P. aeruginosa*, *B. cenocepacia*, *S. aureus*, *A. baumannii* among others[27, 29, 32].

Population E has a 372 bp deletion in SP_2176 (*dltA*), a member of the *dltABCD* operon, which functions by reducing the overall negative charge of the cell envelope through the addition of d-alanine, and it is fixed at day 40. This particular deletion in *dltA* is located removes amino acids 71 to 194, overlapping the ATP binding pocket required to activate the D-alanyl-adenylate. Activated D-alanine is transferred through the D-alanyl carrier protein ligase (*dltC*) through the transmembrane *dltB* to finally add the molecule to the teichoic acid layer [33, 34]. Cooper et al showed *dltB* is strongly selected under in vivo nasal colonization of *S. pneumoniae* and confirmed that impairment of this gene enhanced bacterial attachment[35]. While these results are from a 19F serotype, we hypothesized that deletion of members of this operon would have the same enhanced attachment effect on TIGR4 since our previous Tn-seq data validation confirmed that $\Delta dltA$ increased biofilm formation after one day of growth but showed no difference after day 4, indicating the importance of this operon might be exclusive to the early stage of attachment in biofilms.

We observed strong selection on genes member of the PTS mannose II component system. PTS systems are responsible for the uptake of carbohydrates in bacteria. Unlike other transport systems, PTS use phosphoenolpyruvate (PEP) instead of ATP as phosphate donor. These systems typically consist of two components: 1) an embedded membrane permease EII (enzyme II) which drives sugar transport and 2) EI (enzyme I) complex, which regenerates the permease activity in the system. Multiple SNPs were detected in the PTS-mannose system *manLMN* (SP_0284-0282) as early as day 8 and frequencies increased over time until reaching fixation in multiple populations. *manLMN* in TIGR4 was confirmed to be essential for growth and acts as a high affinity glucose transporter[40], Δ *manL* has also shown to be detrimental to nasopharynx colonization and lung invasiveness in mice[41]. Likewise, we show in Chapter 4 that deletion of *manL* results in higher biofilm formation. We also identified hypothetical gene variants, such as SP_0034, which has 99% homology to spr0034 in the strain R6 and is thought to be a transmembrane protein. The identified mutations varied from single SNPs to 24AA deletion and stop mutations. Although no function has been assigned to this gene, it is worth mention that this gene was present in the Tn-seq data set from the previous chapter, and its deletion resulted in increased biomass production. This suggests that mutations that lead to a loss of function of the gene, such as stop and deletions, may be directly linked to biofilm production.

5.3.3 Development of antibiotic resistance in *S. pneumoniae* biofilms.

After identifying the variants that arose during biofilm growth, we exposed *S. pneumoniae* biofilms to increasing concentrations of antibiotics for 40 days. Here, we aim to understand how biofilm-associated bacteria adapt to antibiotics, the experimental design is illustrated in **Figure 7A**. We exposed three independent population replicates to three classes of antibiotics levofloxacin (target DNA synthesis), vancomycin (target cell wall synthesis), and rifampicin (target RNA synthesis), respectively (**Figure 7B**). Prior to exposing the biofilms to antibiotics, we determined the minimum inhibitory

concentration (MIC) of each antibiotic using the parental strain in planktonic growth (see Methods for details on the concentrations used). We first grew biofilms for a full cycle (4 days), and at the moment of the first passage, we added $0.5 \times$ of the predetermined MIC to the new media. Every day, the medium was changed along with the appropriate concentrations of antibiotics. The experiment consisted of increasing the concentration of antibiotics in 0.5X increments every four days for a total of 10 passages (transfers of coverslips, each transfer is a 4 day exposure to antibiotics), except for the vancomycin experiment, where increments were reduced to $0.25 \times$ MIC due to the harsh selective pressure. At least 4 time points per population were then sequenced and filtered as previously described.

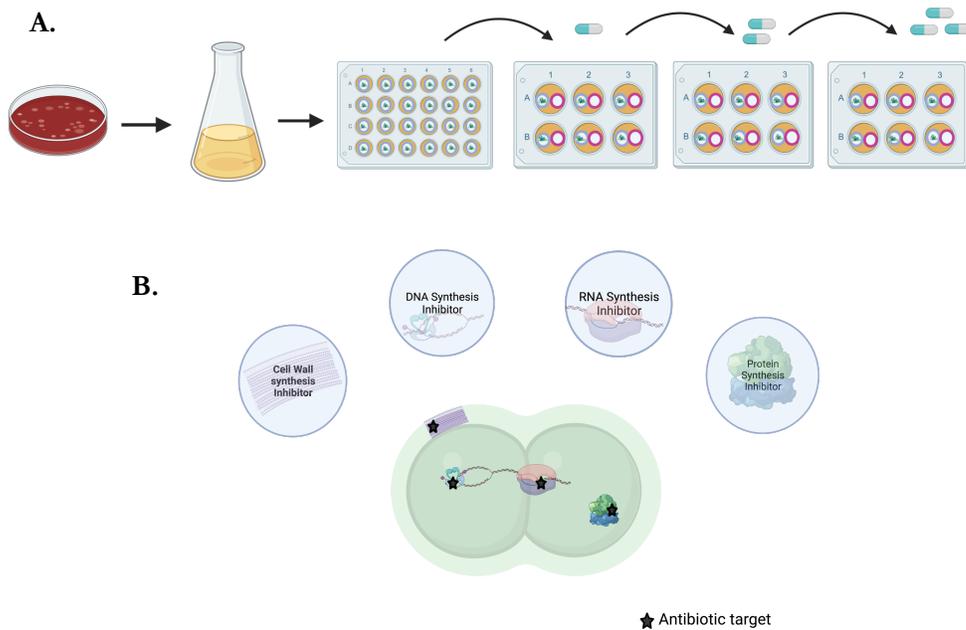


Figure 7. Overview of methodology for identification of adapted genes under antibiotic conditions. **A.** Experimental evolution is done in our biofilm assay, every four days biofilms are transferred, and antibiotic concentration is increased. **B.** Representation of the four classes of antibiotic used during experimental evolution.

Antibiotic pressure selects for antibiotic gene targets.

Levofloxacin is a fluoroquinolone antibiotic that targets two crucial enzymes: DNA gyrase, and topoisomerase IV, which collaborate to alleviate DNA overwinding prior to replication[42]. Mutations in the genes encoding these enzymes are well-known mechanisms of resistance to levofloxacin[43]. Our three adapted populations to levofloxacin harbored a combination of mutations in both classes of genes encoding *gyrA*, *gyrB*, *parC* and *parE*. One of the most studied variants associated with fluoroquinolone resistance is the mutation serine to phenylalanine (S81F) in *gyrA*, which has been confirmed in both vitro planktonic experiments and isolates from patients resistant to levofloxacin[44, 45]. This mutation reduces the binding affinity of the fluoroquinolone for the DNA gyrase enzyme, leading to decreased antibiotic susceptibility.

At least one of the mutations in *gyrA*, *gyrB*, *parC* and *parE* was fixed at passage 10 in all three populations (**Figure 8**). This suggests that when biofilms are exposed to antibiotics, they are more likely to select for a target mutation but still coexist with the other genotypes generated. Fluoroquinolone resistance has been previously reported in *S. pneumoniae* clinical isolates[46, 47]. Mutations in *parC* and *parE* have been detected to confer low-level resistance to DNA synthesis inhibitors and the addition of mutations in *gyrA* can exacerbate the resistance phenotype[46, 48]. While mutations in *gyrB* from levofloxacin resistance isolates are less common, they still contribute to resistant phenotype. We tested the MIC of our levofloxacin adapted populations and confirmed that the MIC increased by 9-18 times more of the parental strain MIC.

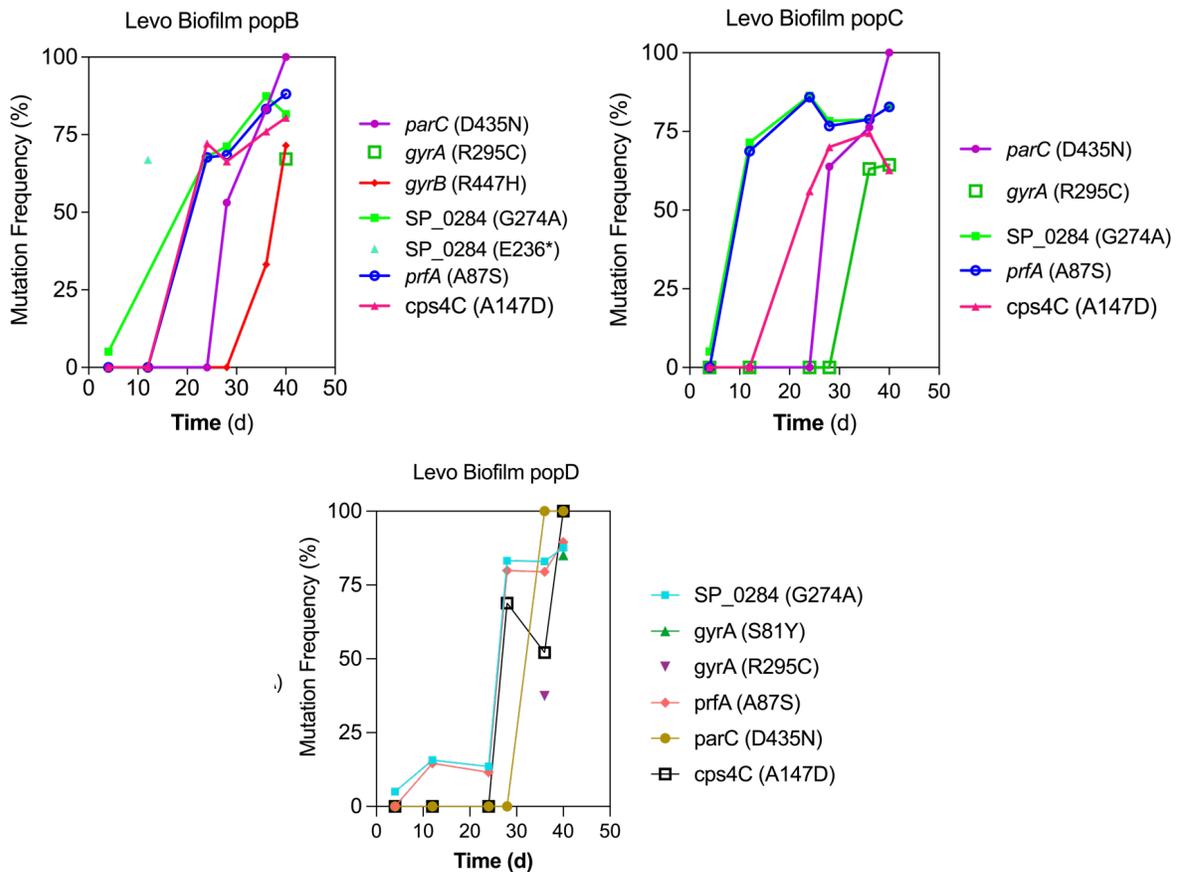


Figure 8. Frequency of trajectories for mutations over 50% at passage 10 in levofloxacin adapted populations. Biofilm populations were adapted to 2.5X MIC of levofloxacin.

Vancomycin inhibits bacterial cell wall synthesis by targeting peptidoglycan components[49]. This antibiotic is generally used the last line of defense use against many gram-positive bacteria[49]. Our approach identified variants in multiple genes related to cell wall metabolism and cell division. For instance, the *ciaH* (SP_0799) is a sensor histidine kinase member of the two-component system (TCS) CiaRH. TCS are diverse regulatory systems in bacteria that consists of two main components, 1) the histidine kinase (HK), which senses signals from the environment and undergoes autophosphorylation which then activates 2) the response regulator (RR) leading to a conformational change that controls a specific cellular function such as, competence, virulence and even antibiotic resistance[50, 51]. TIGR4 has 12 TCS and are widely distributed across the genome. The *CiaRH*, also known as TCS05, was

initially studied as an alternative TCS involved in the negative regulation of competence to counterbalance the activity of *ComDE*[50, 52]. Among dental streptococci the *CiaRH* operon play a substantial role in biofilm formation. In *S. pneumoniae*, *CiaRH* controls the expression of over 70 genes including *manLMN* and *dltABCD*[53]. Loss of the *CiaRH* operon results in decreased biofilm formation and major fitness defect in vivo[54]. Mutation on the *ciaH* (C306) resulted in cefotaxime resistance and some associations between vancomycin tolerance and *ciaRH* has been proposed; however, the exact mechanism has not been elucidated due to the large number of genes under the regulation of *ciaRH*[55].

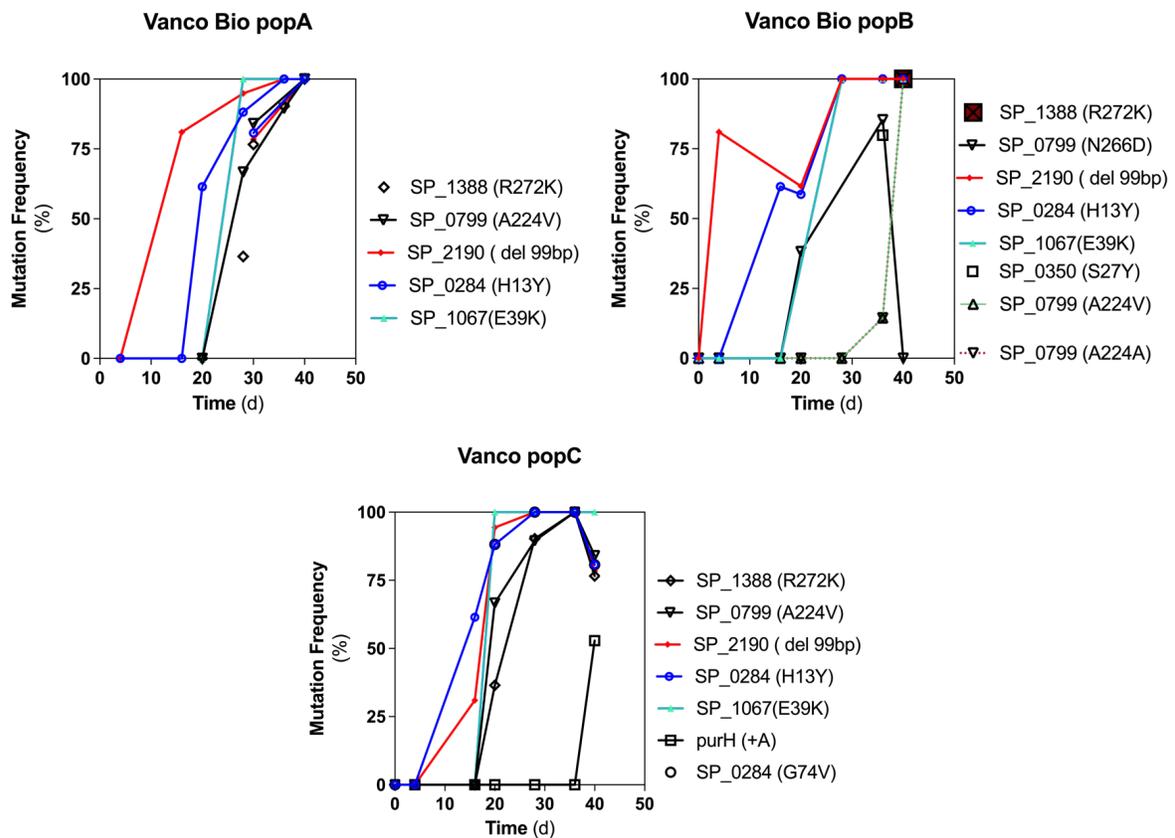


Figure 9. Frequency of trajectories for mutations over 50% at passage 10 in vancomycin adapted populations. Biofilm populations were adapted to 1.75X MIC of vancomycin.

The three populations adapted to vancomycin have mutations in the *SP_1067(fisW)* gene, which is part of the peptidoglycan(PG) complex and works with the Class B penicillin-binding protein (*pbp2X*) to

carry out septal separation during cell division in *S. pneumoniae*[56, 57]. Mutations in *ftsW* cause the blocking of the septal PG synthesis, resulting in elongated chains. We hypothesize that the mutation generated in our biofilm data set can lead to resistance to vancomycin by stopping PG synthesis at the moment of cell division.

Rifampicin inhibits the bacterial RNA polymerase from synthesizing all types of RNA. Rifampicin resistance can occur due to mutations in the β -subunit of the bacterial RNA polymerase, which prevents elongation.[58]. Rifampicin is commonly used to treat MDR *S. pneumoniae* isolates and has been effective when used as an adjuvant therapy with vancomycin. However, resistance to rifampicin has developed. Mutations in the *rpoB* gene have been localized to specific clusters within the well-conserved regions of the gene, which are divided into four main clusters: Cluster N (145-150 AA), I (478-510 AA), II (535-548 AA), and III (585-700 AA).[58]. Previous studies of *S. pneumoniae* rifampicin-resistant isolates have found that mutations in the *rpoB* gene most commonly occur at either cluster I or II, with the most frequently observed mutation being H499T[59, 60]. Our experiments have identified mutations in biofilms that are located within cluster I and III, although these specific mutations (H486N, A670V, E428D, D46N) have not been reported in clinical isolates. We predict that these mutations can enhance resistance to rifampicin by impairing *rpoB* binding, as rifampicin binding is directly related to the degree of impairment of RNA polymerase activity.

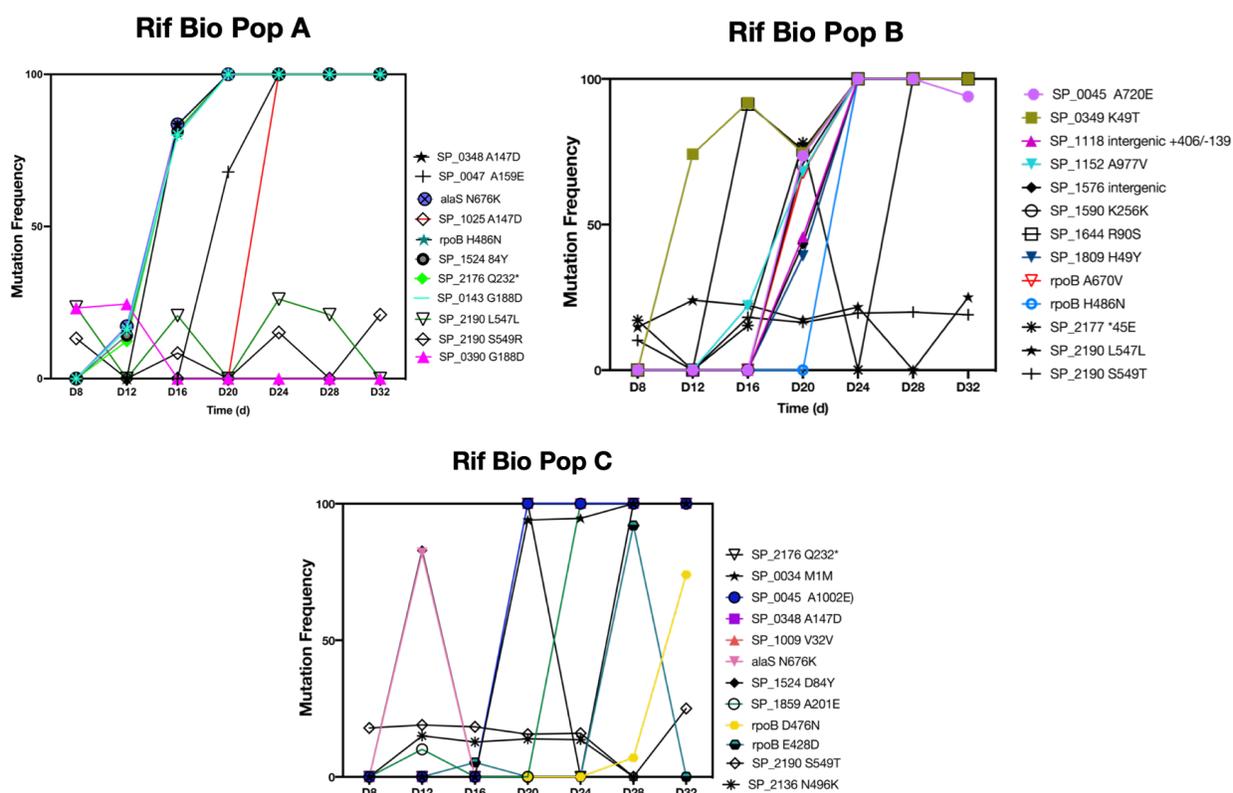


Figure 10. Frequency of trajectories for mutations over 50% at passage 10 in rifampicin adapted populations. Biofilm populations were adapted to 2. 5X MIC of rifampicin.

Antibiotic pressure on *S. pneumoniae* biofilms reveals non-antibiotic associated target mutations.

We observed that the emergence of antibiotic resistance in biofilms selects for mutations different from those related to antibiotic targets or biofilm-related adaptive genes. Specifically, we identified the following genes: peptide release factor 1 *prfA* (SP_1020) in the levofloxacin adaptation, adhesin A *cbpA* (SP_2190) and putrescine import ABC *potB* (SP_1388), both in the vancomycin adaptation, and two putative proteins with unknown functions (SP_0045, SP_1859) in the rifampicin adaptation experiment. Mutations on these genes were present only under antibiotic conditions. Little is known

about the function or contribution of these genes in antibiotic resistance or even contribution to biofilm formation.

Choline binding protein A, encoded by the gene *cbpA*, is one of the 15 proteins interact with choline residues in the teichoic acid (TA)[61]. *cbpA*, usually referred to as *PspA*, is one of the major choline-binding proteins with multiple functions in *S. pneumoniae*[62] including: 1) facilitating bacteria attachment to epithelial cells[63], 2) binding to component of the complement systems and limiting opsonophagocytic killing of pneumococci by immune cell [64], 3) being implicated in biofilm formation[65]. The 99bp deletion on *cbpA* gene in the vancomycin experiment resulted in fixation, we propose that a method to achieve vancomycin resistance is by effectively impeding the incorporation of choline into the bacterial TA layer. In addition to the changes in *cbpA*, we discovered variants in other choline-binding proteins (Cbp) that were present across all adaptation experiments at frequencies lower than 25%. Our results suggest that further investigation of cbp in the context of antibiotic resistance is warranted, as some Cbps are currently being explored by others as promising drug targets[66].

5.4 Conclusion

Experimental evolution is used to study evolutionary processes in real-time by subjecting organisms to controlled laboratory conditions and observing changes in their traits over multiple generations. This technique is particularly valuable in the context of antibiotic resistance in biofilms, as it provides insights into how bacteria adapt and develop resistance to antimicrobial agents within the complex and structured environments of biofilm communities. We show that evolution experiments conducted in multiple replicate biofilm populations in our previously developed biofilm assay, facilitate the sustained selection of enhanced biofilm formers. These experiments systematically monitor mutations contributing to improved biofilm formation over an extended period. In contrast to Tn-seq, which exclusively considers 'loss of function' mutations akin to stop codons and deletions in experimental evolution, the latter method allows for the identification of single amino acid substitutions. Such substitutions may prove sufficient to yield specific phenotypes.

Our research showed that as *S. pneumoniae* biofilms grow over time, they select genetic factors that promote biofilm formation. These genes are primarily related to cell wall, capsule, and carbohydrate metabolism. Additionally, we found that biofilms serve as a reservoir for genotype generation and maintenance, as they are less likely to experience competition for resources due to their spatial arrangement and provide a suitable environment for the coexistence of multiple haplotypes. We identified several mutations in capsule genes, including stop mutations suggestive of a loss-of-function phenotype. The observed reduction in polysaccharide capsule levels aligns with prior findings, indicating that lower capsule levels are associated with increased biomass production. Furthermore, our experiments selected for mutations in PTS-mannose and d-alanine-related genes. We speculate that these mutations contribute to the enhanced selection of superior biofilm formers. This speculation is supported by the observed loss of function in these genes in strains with significantly increased biofilm formation observed via the Tn-seq approach in the previous chapter.

Long term exposure of *S. pneumoniae* biofilms to increasing antibiotic concentrations revealed that genetic variants related to the antibiotic target are commonly found our data set and are likely to enhance antibiotic resistance in the population. Our analysis revealed multiple hypothetical variants across all experiments, suggesting that further research on biofilms and antibiotic resistance could help to clarify the function of these unknown genes. Studying the emergence of antibiotic resistance in these biofilms may provide a valuable opportunity to explore previously unexamined genes for their role in antibiotic resistance. Our dataset can be used for further predictions on bacterial adaptive outcomes which can result in a great prognostic of the emergence of antibiotic resistance in the clinic.

4.5 Methods

***S. pneumoniae* biofilm growth, visualization, quantification, and statistical analysis.**

Biofilms were grown, visualized and quantified as previously described in Chapter 2 (p. 32).

Experimental evolution, whole-genome sequencing and analysis

TIGR4 was used as the parental strain in biofilm and antibiotic evolution experiment. For antibiotic adaptation antibiotic increasing concentration of levofloxacin, vancomycin and rifampicin were added at every passage in 0.5X MIC steps with exception of vancomycin, which was added in 0.25X MIC steps. Genomic DNA was isolated from adapted populations and single strains using a DNase Blood and Tissue kit (Qiagen), concentrations of genomic DNA were measured on a Qubit 3.0 fluorometer (Invitrogen) and diluted to 5ng/uL for library preparation using a Nextera kit (Illumina). Libraries were sequenced on an Illumina NextSeq500 and reads were mapped to their corresponding reference genomes. Mutations were identified using the Breseq pipeline [30] we developed on our streaming Aerobio analyses platform with polymorphism mode for populations and consensus mode for adapted clones [30, 67]. Post processing is done using the following pipeline [68]. In brief adaptive mutations in each experiment are determined based on the following criteria: 1) the frequency of a mutation is greater than 10% in at least one replicate population; 2) the mutation is not present in any background adapted populations, and; 3) the mutation is a nonsense or missense mutation. To predict genotypes and lineages based on trajectory of mutations Muller plots were generated using the lollipop package v 0.9.0 [69].

Determination of minimum inhibitory concentration (MIC).

MICs were determined as previously described [70]. In short, $\sim 1 \times 10^5$ CFU of mid-exponential bacteria are cultured in 200 μ L in 96-well plates in fresh medium containing a single antibiotic at the following concentration gradients and increments: Levofloxacin gradient 0.4–1.2 μ g/mL with 0.1 μ g/mL increments; vancomycin 0.12–0.32 μ g/mL with 0.04 μ g/mL increments; rifampicin 0.001–0.005 with 0.0012 μ g/mL increments. MICs were determined in triplicate and monitored on BioSpa 8

(BioTek) at 37 °C for 20 h. MIC is determined as the lowest concentration that abolishes bacterial growth.

Antibiotic	MIC ($\mu\text{g/ml}$)
Levofloxacin	0.8
Vancomycin	0.24
Rifampicin	0.0032

4.6 Supplementary information

Table 1. Mutations over 50% frequency in all population after 40 days of biofilm growth

Gene	Annotation	Description	Pop A	Pop B	Pop C	Pop D	Pop E
SP_0034 ←	Δ76 bp	hypothetical protein	87%				
SP_0034 ←	E210* (GAG→TAG)	hypothetical protein					82%
SP_0114 ←	R63H (CGT→CAT)	hypothetical protein					100%
<i>purH</i> → / → SP_0051	intergenic (+111/-11)	bifunctional phosphoribosylaminoimidazole carboxamide formyltransferase		74%			
<i>mutL</i> →	S627* (TCG→TAG)	DNA mismatch repair protein		69%			
<i>secY</i> →	R369C (CGT→TGT)	preprotein translocase subunit SecY	86%				
SP_0250 → / → SP_0251	intergenic (+144/-7)	PTS system, IIC component/formate acetyltransferase, putative				91%	
SP_0282 ←	H221Y (CAT→TAT)	PTS system, mannose-specific IID component	85%				
SP_0284 ←	D70G (GAC→GGC)	PTS system, mannose-specific IIAB components					100%
SP_0348 →	A147D (GCT→GAT)	capsular polysaccharide biosynthesis protein Cps4C					100%
SP_0350 →	M121I (ATG→ATA)	capsular polysaccharide biosynthesis protein Cps4E	88%				
SP_0350 →	E92* (GAG→TAG)	capsular polysaccharide biosynthesis protein Cps4E		65%			
SP_0395 →	R136* (CGA→TGA)	transcriptional regulator, putative		68%			
SP_0797 →	E452K (GAA→AAA)	aminopeptidase N					70%
SP_0840 →	A18S (GCT→TCT)	hypothetical protein		57%			
SP_0897 →	A225E (GCA→GAA)	pyruvate kinase			95%		
SP_0875 →	T180T (ACG→ACA)	lactose phosphotransferase system repressor	87%				
SP_1118 →	G341D (GGT→GAT)	pullulanase, putative			100%		
SP_1118 →	Q345* (CAA→TAA)	pullulanase, putative				90%	
SP_0925 →	L158I (CIT→AIT)	hypothetical protein		65%			
SP_1060 →	A2A (GCC→GCI)	hypothetical protein	92%				
SP_1120 ← / → SP_1121	intergenic (-8/-22)	hypothetical protein/glycogen branching enzyme			100%	91%	
SP_1967 ←	G254D (GGT→GAT)	hypothetical protein	83%				
SP_2106 ←	W415* (TGG→TGA)	glycogen phosphorylase family protein					100%
SP_2143 ←	S532N (AGT→AAT)	hypothetical protein	88%				
SP_2174 ←	T65I (ACA→ATA)	D-alanine-poly(phosphoribitol) ligase subunit 2	100%	65%		87%	
SP_2176 ←	Δ372 bp	D-alanine-D-alanyl carrier protein ligase			90%		100%

4.7 References

1. Gierke, R., A.P. Wodi, and M. Kobayash, *Pneumococcal Disease*, in *Epidemiology and Prevention of Vaccine-Preventable Diseases*, C.f.D.C.a. Prevention, Editor. 2021, Public Health Foundation: Washington D.C.
2. CDC, *ANTIBIOTIC RESISTANCE THREATS IN THE UNITED STATES*. U.S. Department of Health and Human Services, 2019.
3. Mohanty, S., et al., *A Multicenter Evaluation of Trends in Antimicrobial Resistance Among Streptococcus pneumoniae Isolates From Adults in the United States*. Open Forum Infect Dis, 2022. **9**(9): p. ofac420.
4. Pallares, R., O. Capdevila, and I. Grau, *Treatment options for resistant pneumococcal infections*. Clin Microbiol Infect, 1999. **5 Suppl 4**: p. S3-S11.
5. Bjarnsholt, T., *The role of bacterial biofilms in chronic infections*. APMIS Suppl, 2013(136): p. 1-51.
6. Ribet, D. and P. Cossart, *How bacterial pathogens colonize their hosts and invade deeper tissues*. Microbes Infect, 2015. **17**(3): p. 173-83.
7. Baishya, J. and C.A. Wakeman, *Selective pressures during chronic infection drive microbial competition and cooperation*. NPJ Biofilms Microbiomes, 2019. **5**(1): p. 16.
8. Bisht, K. and C.A. Wakeman, *Discovery and Therapeutic Targeting of Differentiated Biofilm Subpopulations*. Front Microbiol, 2019. **10**: p. 1908.
9. Vestby, L.K., et al., *Bacterial Biofilm and its Role in the Pathogenesis of Disease*. Antibiotics (Basel), 2020. **9**(2).
10. Rybtke, M., et al., *Pseudomonas aeruginosa Biofilm Infections: Community Structure, Antimicrobial Tolerance and Immune Response*. J Mol Biol, 2015. **427**(23): p. 3628-45.
11. Sharma, D., L. Misba, and A.U. Khan, *Antibiotics versus biofilm: an emerging battleground in microbial communities*. Antimicrob Resist Infect Control, 2019. **8**: p. 76.
12. Khan, J., et al., *Challenges of antibiotic resistance biofilms and potential combating strategies: a review*. 3 Biotech, 2021. **11**(4): p. 169.
13. Uruen, C., et al., *Biofilms as Promoters of Bacterial Antibiotic Resistance and Tolerance*. Antibiotics (Basel), 2020. **10**(1).
14. Bowler, P., C. Murphy, and R. Wolcott, *Biofilm exacerbates antibiotic resistance: Is this a current oversight in antimicrobial stewardship?* Antimicrob Resist Infect Control, 2020. **9**(1): p. 162.
15. Olivares, E., et al., *Clinical Impact of Antibiotics for the Treatment of Pseudomonas aeruginosa Biofilm Infections*. Front Microbiol, 2019. **10**: p. 2894.
16. Schrader, S.M., J. Vaubourgeix, and C. Nathan, *Biology of antimicrobial resistance and approaches to combat it*. Sci Transl Med, 2020. **12**(549).

17. Guzman-Soto, I., et al., *Mimicking biofilm formation and development: Recent progress in in vitro and in vivo biofilm models*. iScience, 2021. **24**(5): p. 102443.
18. Gloag, E.S., et al., *Pseudomonas aeruginosa Interstrain Dynamics and Selection of Hyperbiofilm Mutants during a Chronic Infection*. mBio, 2019. **10**(4).
19. Traverse, C.C., et al., *Tangled bank of experimentally evolved Burkholderia biofilms reflects selection during chronic infections*. Proc Natl Acad Sci U S A, 2013. **110**(3): p. E250-9.
20. El-Halfawy, O.M. and M.A. Valvano, *Antimicrobial heteroresistance: an emerging field in need of clarity*. Clin Microbiol Rev, 2015. **28**(1): p. 191-207.
21. McDonald, M.J., *Microbial Experimental Evolution - a proving ground for evolutionary theory and a tool for discovery*. EMBO Rep, 2019. **20**(8): p. e46992.
22. Kostakioti, M., M. Hadjifrangiskou, and S.J. Hultgren, *Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era*. Cold Spring Harb Perspect Med, 2013. **3**(4): p. a010306.
23. Mehta, H.H., A.G. Prater, and Y. Shamoo, *Using experimental evolution to identify druggable targets that could inhibit the evolution of antimicrobial resistance*. J Antibiot (Tokyo), 2018. **71**(2): p. 279-286.
24. Lenski, R.E., *Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations*. ISME J, 2017. **11**(10): p. 2181-2194.
25. Coenye, T., M. Bove, and T. Bjarnsholt, *Biofilm antimicrobial susceptibility through an experimental evolutionary lens*. NPJ Biofilms Microbiomes, 2022. **8**(1): p. 82.
26. Flynn, K.M., et al., *Evolution of Ecological Diversity in Biofilms of Pseudomonas aeruginosa by Altered Cyclic Diguanylate Signaling*. J Bacteriol, 2016. **198**(19): p. 2608-18.
27. Turner, C.B., C.W. Marshall, and V.S. Cooper, *Parallel genetic adaptation across environments differing in mode of growth or resource availability*. Evol Lett, 2018. **2**(4): p. 355-367.
28. Steenackers, H.P., et al., *Experimental evolution in biofilm populations*. FEMS Microbiol Rev, 2016. **40**(3): p. 373-97.
29. Penesyan, A., et al., *Rapid microevolution of biofilm cells in response to antibiotics*. NPJ Biofilms Microbiomes, 2019. **5**(1): p. 34.
30. Deatherage, D.E. and J.E. Barrick, *Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq*. Methods Mol Biol, 2014. **1151**: p. 165-88.
31. Scribner, M.R., et al., *Parallel Evolution of Tobramycin Resistance across Species and Environments*. mBio, 2020. **11**(3).
32. Koza, A., et al., *Adaptive radiation of Pseudomonas fluorescens SBW25 in experimental microcosms provides an understanding of the evolutionary ecology and molecular biology of A-L interface biofilm formation*. FEMS Microbiol Lett, 2017. **364**(12).

33. Kiriukhin, M.Y. and F.C. Neuhaus, *D-alanylation of lipoteichoic acid: role of the D-alanyl carrier protein in acylation*. J Bacteriol, 2001. **183**(6): p. 2051-8.
34. Engholm, D.H., et al., *A visual review of the human pathogen Streptococcus pneumoniae*. FEMS Microbiol Rev, 2017. **41**(6): p. 854-879.
35. Cooper, V.S., et al., *Experimental Evolution In Vivo To Identify Selective Pressures during Pneumococcal Colonization*. mSystems, 2020. **5**(3).
36. Shainheit, M.G., M. Mule, and A. Camilli, *The core promoter of the capsule operon of Streptococcus pneumoniae is necessary for colonization and invasive disease*. Infect Immun, 2014. **82**(2): p. 694-705.
37. Paton, J.C. and C. Trappetti, *Streptococcus pneumoniae Capsular Polysaccharide*. Microbiol Spectr, 2019. **7**(2).
38. Shainheit, M.G., et al., *Mutations in pneumococcal cpsE generated via in vitro serial passaging reveal a potential mechanism of reduced encapsulation utilized by a conjunctival isolate*. J Bacteriol, 2015. **197**(10): p. 1781-91.
39. Schaeffer, C.R., et al., *Accumulation-associated protein enhances Staphylococcus epidermidis biofilm formation under dynamic conditions and is required for infection in a rat catheter model*. Infect Immun, 2015. **83**(1): p. 214-26.
40. Fleming, E. and A. Camilli, *ManLMN is a glucose transporter and central metabolic regulator in Streptococcus pneumoniae*. Molecular Microbiology, 2016. **102**(3): p. 20.
41. van Opijnen, T. and A. Camilli, *A fine scale phenotype-genotype virulence map of a bacterial pathogen*. Genome Res, 2012. **22**(12): p. 2541-51.
42. Levine, C., H. Hiasa, and K.J. Marians, *DNA gyrase and topoisomerase IV: biochemical activities, physiological roles during chromosome replication, and drug sensitivities*. Biochimica et Biophysica Acta, 1998.
43. MUNˆOZ, R. and A.G.D.L. CAMPA, *ParC Subunit of DNA Topoisomerase IV of Streptococcus pneumoniae Is a Primary Target of Fluoroquinolones and Cooperates with DNA Gyrase A Subunit in Forming Resistance Phenotype*. ANTIMICROBIAL AGENTS AND CHEMOTHERAPY,, 1996. **40**(10): p. 5.
44. Ruiz, J., *Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection*. J Antimicrob Chemother, 2003. **51**(5): p. 1109-17.
45. Park, M., et al., *Novel Levofloxacin-Resistant Multidrug-Resistant Streptococcus pneumoniae Serotype 11A Isolates, South Korea*. Emerg Infect Dis, 2016. **22**(11): p. 1978-1980.
46. IRENE GONZA´LEZ, et al., *Fluoroquinolone Resistance Mutations in the parC, parE, and gyrA Genes of Clinical Isolates of Viridans Group Streptococci*. ANTIMICROBIAL AGENTS AND CHEMOTHERAPY,, 1998. **42**(11): p. 6.

47. Davies, T.A., et al., *Prevalence of single mutations in topoisomerase type II genes among levofloxacin-susceptible clinical strains of Streptococcus pneumoniae isolated in the United States in 1992 to 1996 and 1999 to 2000*. Antimicrob Agents Chemother, 2002. **46**(1): p. 119-24.
48. Hooper, D.C. and G.A. Jacoby, *Topoisomerase Inhibitors: Fluoroquinolone Mechanisms of Action and Resistance*. Cold Spring Harb Perspect Med, 2016. **6**(9).
49. Stogios, P.J. and A. Savchenko, *Molecular mechanisms of vancomycin resistance*. Protein Sci, 2020. **29**(3): p. 654-669.
50. Bhate, M.P., et al., *Signal transduction in histidine kinases: insights from new structures*. Structure, 2015. **23**(6): p. 981-94.
51. van Opijnen, T. and A. Camilli, *Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms*. Nat Rev Microbiol, 2013. **11**(7): p. 435-42.
52. Moscoso, M., M. Domenech, and E. Garcia, *Vancomycin tolerance in Gram-positive cocci*. Environ Microbiol Rep, 2011. **3**(6): p. 640-50.
53. Halfmann, A., et al., *Identification of the genes directly controlled by the response regulator CiaR in Streptococcus pneumoniae: five out of 15 promoters drive expression of small non-coding RNAs*. Mol Microbiol, 2007. **66**(1): p. 110-26.
54. Blanchette-Cain, K., et al., *Streptococcus pneumoniae biofilm formation is strain dependent, multifactorial, and associated with reduced invasiveness and immunoreactivity during colonization*. mBio, 2013. **4**(5): p. e00745-13.
55. He, L.Y., et al., *The Role and Regulatory Network of the CiaRH Two-Component System in Streptococcal Species*. Front Microbiol, 2021. **12**: p. 693858.
56. Briggs, N.S., et al., *The Pneumococcal Divisome: Dynamic Control of Streptococcus pneumoniae Cell Division*. Front Microbiol, 2021. **12**: p. 737396.
57. Nakamoto, R., et al., *The divisome but not the elongasome organizes capsule synthesis in Streptococcus pneumoniae*. Nat Commun, 2023. **14**(1): p. 3170.
58. Goldstein, B.P., *Resistance to rifampicin: a review*. J Antibiot (Tokyo), 2014. **67**(9): p. 625-30.
59. Chen, J.Y., et al., *Mutations of the rpoB gene in rifampicin-resistant Streptococcus pneumoniae in Taiwan*. J Antimicrob Chemother, 2004. **53**(2): p. 375-8.
60. Ferrandiz, M.J., et al., *New mutations and horizontal transfer of rpoB among rifampin-resistant Streptococcus pneumoniae from four Spanish hospitals*. Antimicrob Agents Chemother, 2005. **49**(6): p. 2237-45.
61. Luo, R., et al., *Solution structure of choline binding protein A, the major adhesin of Streptococcus pneumoniae*. EMBO J, 2005. **24**(1): p. 34-43.
62. Lane, J.R., et al., *A Jack of All Trades: The Role of Pneumococcal Surface Protein A in the Pathogenesis of Streptococcus pneumoniae*. Front Cell Infect Microbiol, 2022. **12**: p. 826264.

63. Shenoy, A.T. and C.J. Orihuela, *Anatomical site-specific contributions of pneumococcal virulence determinants*. Pneumonia (Nathan), 2016. **8**.
64. Kerr, A.R., et al., *The contribution of PspC to pneumococcal virulence varies between strains and is accomplished by both complement evasion and complement-independent mechanisms*. Infect Immun, 2006. **74**(9): p. 5319-24.
65. Moscoso, M., E. Garcia, and R. Lopez, *Biofilm formation by Streptococcus pneumoniae: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion*. J Bacteriol, 2006. **188**(22): p. 7785-95.
66. Maestro, B. and J.M. Sanz, *Choline Binding Proteins from Streptococcus pneumoniae: A Dual Role as Enzybiotics and Targets for the Design of New Antimicrobials*. Antibiotics (Basel), 2016. **5**(2).
67. Anthony, J.S.v.O., T, *A DAG computation server for fully integrated and automated massively parallel sequencing analyses*. 2019.
68. Surujon, D., *Adaptation*. 2019.
69. Deitrick, C., *Lolipop*. 2019. p. A set of scripts to cluster mutational trajectories into genotypes and cluster genotypes by background.
70. Leshchiner, D., et al., *A genome-wide atlas of antibiotic susceptibility targets and pathways to tolerance*. Nat Commun, 2022. **13**(1): p. 3165.

Chapter 6

Discussion

6.1 Delving into the complexity of *S. pneumoniae* biofilms: A systems-level approach.

Bacteria are often studied as planktonic or free-living organisms, and frequently grow in complex surface-attached communities known as biofilms[1, 2]. The life cycle of a biofilm is not linear. As bacteria attach to and colonize a surface, they have to surrender their autonomy to form part of a community. These bacterial assemblies mirror the complexities of human settlements, and not only is their architecture unique, but growth and cellular function during the biofilm life cycle require coordinated actions and intricate communication among all its players (Chapter 1).

This thesis offers a unique outlook for the study of *Streptococcus pneumoniae* biofilms. We combined *in vitro*, genome-wide, and *in vivo* experiments to elucidate the complexity of *S. pneumoniae* biofilms. Conventional methods to grow *S. pneumoniae* biofilms *in vitro* use polystyrene plates and are mostly carried out for periods no longer than 24 h [3, 4]. However, these methods are unable to study the maintenance of these communities over time and fail to capture the complexity of these microbial engines. We developed a long-term *in vitro* assay for *S. pneumoniae* that captures the complexity of the biofilm growth dynamics (Chapter 2). We applied a genome-wide functional approach and for the first time measured the fitness of individual mutants during the growth of *S. pneumoniae* biofilms (Chapter 4). We complemented our data by recording transcriptional response changes in these biofilm populations. Moreover, by exploring the genetic basis of antibiotic resistance in biofilm-associated *S. pneumoniae*, we contributed to a broader understanding of antimicrobial resistance dynamics (Chapter 5). These insights can guide the development of innovative strategies to combat resistance, including novel drug targets and combination therapies that can circumvent or mitigate the impact of resistance mechanisms.

Planktonic, biofilm, and dispersal stages have profound phenotypic differences.

After developing a long-term reproducible in vitro assay to grow *S. pneumoniae* biofilms, we confirmed that different strains exhibited different abilities to form biofilms (Chapter 2). We verified that dispersal cells actively released from *S. pneumoniae* biofilms were phenotypically different from their planktonic and aggregated counterparts. Our transcriptome study from the strain TIGR4 in Appendix A revealed that biofilms and dispersal cells undergo a multitude of changes in mRNA abundance when compared to planktonic growth. Suggesting that the transition from planktonic to biofilms (surface-attached) to dispersal cells and the maintenance of each of these cell states is an expensive energetic process. This dramatic change in surface-attached cells might seem counterintuitive, as these cells remain in a sessile state; however, secreting the polymers, extracellular DNA, and enzymes that form the extracellular polymeric matrix (EPS) are laborious tasks and most likely result in a higher global transcriptomic response than that experienced by planktonic cells. Overall, there are dramatic physiological changes and regulation rearrangement occurring in biofilm and dispersal cells.

Identification of the key functions and pathways involved in biofilm formation.

Our genome-wide approach, Tn-Seq, showed that key participants in biofilm development are spread across the genome. By exploring mutant libraries of two strains with different capacities to form biofilms, we obtained a dynamic view of the main pathways with roles at different stages during 96 h of biofilm establishment. For instance, we identified genes that appear to be important across multiple time points, but we also located genes that were required at single timepoints. To complement our findings about genetic factors favored under biofilm growth, we decided to explore biofilm growth for 40 days. We serially passaged our biofilms for over 40 days and found that the genetic determinants favored under long-term biofilm maintenance overlapped with some of the genes identified in the Tn-seq screens. Besides the seemingly chaotic nature of these bacterial communities, the multifaceted

approach taken in this study gives a higher resolution of bacterial population dynamics for *S. pneumoniae* biofilms.

Cells growing inside biofilms must be able to adapt to limiting nutrient availability and oxygen; therefore, it is not surprising that our Tn-seq and RNA-seq approaches were able to identify genes involved in the stringent response pathway (SP_1645), cell wall modifications (SP_2176), carbohydrate metabolism (SP_0284), and mRNA turnover (SP_1739). As expected, we confirmed that capsule downregulation, or lack thereof, was correlated with higher biomass production. We suspect that the pathways mentioned above work coordinated to grow and disperse microcolonies successfully.

Adapting *S. pneumoniae* biofilms to three different types of antibiotics revealed common adaptive pathways to achieve biofilm growth and antibiotic resistance (antibiotic target genes), as well as novel routes of adaptation to develop resistance. The implications of this study extend beyond *S. pneumoniae*, as the principles uncovered can be relevant to other bacterial pathogens that form biofilms and exhibit antibiotic resistance. Our findings add to the growing body of knowledge in the field of bacterial genetics and antimicrobial resistance, paving the way for future research and therapeutic advancement.

6.2 Disease outcome differs according to cell growth state and biofilm formation capacity.

S. pneumoniae forms biofilms in the nose, lungs, and ears and has recently been discovered to grow as cell aggregates in the heart [5-7]. Despite the development of pneumococcal vaccines, this facultative anaerobe remains a subject of study because of the emergence of antibiotic resistance and increased bacterial phenotypic heterogeneity in the clinic [8, 9]. Because biofilms play a critical role during different stages of infection, it is critical to study bacteria within the context of a biofilm. In Chapter 3, we explore the disease outcome phenotype of three cell growth states: planktonic, biofilm, and dispersed cells from the same strain TIGR4 in a murine lung infection model. Biofilm samples colonized the nasopharynx but were unable to migrate to the lungs and blood; however, planktonic

and dispersal populations spread to the blood at approximately the same magnitude (Chapter 3). Surprisingly, our data revealed that dispersal cells were in a ready to infect state, as mice inoculated with these samples showed symptoms of aggressive pneumonia as early as 14 h post-infection with the same inoculum size used for planktonic infections. In addition, bacterial titers in the lung were 10 times higher in mice inoculated with dispersal cells when compared to those with a planktonic inoculum. Altogether, we showed that dispersal cells isolated from our *in vitro* assay were more virulent than planktonic and biofilm populations. Previous *in vivo* Tn-Seq experiments have identified several genes required for infection in the TIGR4 strain[10] and in a collection of strains representative of the pangenome (Rosconi, unpublished data), but these experiments were conducted using Tn-Seq libraries in the planktonic stage. Given the significant differences in transcriptomics between planktonic and dispersal cells, it is possible that the genetic requirements for infection in the dispersal stage overlap with those identified in the planktonic stage, but also have unique requirements that are specific to the dispersal stage.

The disease outcome changed when we explored strains with different capacities to form biofilms *in vitro*. When mice were infected with biofilm inoculum from two low-biofilm mutant strains, Δ SP_1645 and Δ SP_1739, we observed that they were able to cause lung infection, unlike WT biofilm samples. We suspect that this phenotype is a result of the low biofilm formation capacity, as cells in these biofilms might not be held together as strongly within the EPS and can be primed to more detachment, releasing faster dispersal cells, which we confirmed are inherently more virulent. We observed opposite results with a high biofilm mutant strain, Δ SP_2205, in which biofilms are unable to cause disease. This indicates that the disease outcome in animals infected with Δ SP_2205 resembles that of WT TIGR4. We supported this through the analysis of the transcriptome of this high biofilm former, where we noticed a less chaotic response of the three states of growth when compared to the two low biofilm formers. Our laboratory previously introduced the concept of entropy as a chaotic response generated in the bacterial transcriptome upon exposure to stress [11]. We extend this entropy concept to the

overall mRNA disorder observed in biofilm and dispersal samples from the Δ SP_1645 and Δ SP_1739 as almost 65% of the expressed genes were differentially expressed under these conditions.

6.3 Looking ahead: The promise of biofilm and dispersal research.

The implementation of genome-wide approaches to complex systems, such as bacterial biofilms, allows the determination of genes and regulatory mechanisms important during biofilm and dispersal development. The contents of this thesis proved that it is extremely important to study not only biofilms but also dispersal cells. We propose that biofilm therapeutics, e.g., drugs or biologicals that disrupt biofilms, should also be targeted towards the creation of anti-dispersal therapies, as dispersal cells exhibit a hypervirulent phenotype. When exploring biofilm and dispersal cells, it is important to understand the role of these cells in host-to-host dissemination. Dispersion should also be explored in mixed-species biofilm to determine whether the hypervirulence phenotype remains the same when they come from multiple species. For example, a secondary infection with *S. pneumoniae* after influenza virus (H1N1) infection resulted in a more severe disease outcome [12].

The complexity of biofilms can be better dissected by performing even more sensible genome-wide techniques at the single-cell (sc) level. Spatial single-cell transcriptomics can be used to capture phenotypic heterogeneity in different biofilm layers [13]. Par-seqFish combines single-cell RNA sequencing with fluorescence in situ hybridization (FISH) to analyze the expression patterns of single cells. This approach used *P. aeruginosa* biofilms as a proof-of-concept and was able to determine unique metabolic changes in single cells in different parts of the biofilm [14]. With the development of new optimized single-cell spatial transcriptomic approaches, we have entered a new era of biofilm research. Although still a lot is left to know about *S. pneumoniae* biofilms, this thesis provides a robust data set with potential biofilm and dispersal diagnostic markers to be explored by future scientists.

6.4 References

1. Steenackers, H.P., et al., *Experimental evolution in biofilm populations*. FEMS Microbiol Rev, 2016. **40**(3): p. 373-97.
2. Flemming, H.-C., et al., *Biofilms: an emergent form of bacterial life*. Nature Reviews Microbiology, 2016. **14**(9): p. 563-575.
3. Valente, C., et al., *Intra-Species Interactions in Streptococcus pneumoniae Biofilms*. Front Cell Infect Microbiol, 2021. **11**: p. 803286.
4. Domenech, M., E. Garcia, and M. Moscoso, *Biofilm formation in Streptococcus pneumoniae*. Microb Biotechnol, 2012. **5**(4): p. 455-65.
5. Chao, Y., et al., *Streptococcus pneumoniae biofilm formation and dispersion during colonization and disease*. Front Cell Infect Microbiol, 2014. **4**: p. 194.
6. Shenoy, A.T., et al., *Streptococcus pneumoniae in the heart subvert the host response through biofilm-mediated resident macrophage killing*. PLoS Pathog, 2017. **13**(8): p. e1006582.
7. Wróblewska, J., et al., *Biofilm formation of Streptococcus pneumoniae from bronchial alveolar lavage and from nasal swab*. Medical Research Journal, 2016. **1**(3).
8. Obolski, U., et al., *Vaccination can drive an increase in frequencies of antibiotic resistance among nonvaccine serotypes of Streptococcus pneumoniae*. Proc Natl Acad Sci U S A, 2018. **115**(12): p. 3102-3107.
9. Cillóniz, C., et al., *Antimicrobial Resistance Among Streptococcus pneumoniae*, in *Antimicrobial Resistance in the 21st Century*. 2018. p. 13-38.
10. van Opijnen, T. and A. Camilli, *A fine scale phenotype-genotype virulence map of a bacterial pathogen*. Genome Res, 2012. **22**(12): p. 2541-51.
11. Zhu, Z., et al., *Entropy of a bacterial stress response is a generalizable predictor for fitness and antibiotic sensitivity*. Nat Commun, 2020. **11**(1): p. 4365.
12. Mueller Brown, K., et al., *Secondary infection with Streptococcus pneumoniae decreases influenza virus replication and is linked to severe disease*. FEMS Microbes, 2022. **3**: p. xtac007.
13. Homberger, C., L. Barquist, and J. Vogel, *Ushering in a new era of single-cell transcriptomics in bacteria*. MicroLife, 2022. **3**: p. uqac020.
14. Dar, D., et al., *Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution*. Science, 2021. **373**(6556).

Appendix A

Transcriptome profiling of TIGR4 *S.
*pneumoniae** biofilms.

1. Background

Bacteria grow in nature as a mixture of aggregates commonly referred to as biofilms [1]. Forming a biofilm requires of multiple stages including attachment, expansion, and maintenance, and lastly dispersal [2]. During the stages of biofilm establishment bacteria have to regulate their transition from a motile to sessile lifestyle [3]. For example, in multiple species such as *E. coli* and *P. aeruginosa* flagellar and curli genes in charge of EPS assembly are derepressed upon the attachment stage [4]. The switch from biofilms to dispersal cells have been found mostly to be regulated through the transitions in secondary molecule regulators which are able to oscillate in their expression and activate a cascade of changes in the overall transcriptome depending on the state bacteria is in (sessile or motile) [5]. Biofilm phenotypic adaptation have been shown to be associated with changes in gene expressions in response to their environment [6]. For example the different layers of nutrient gradients generated inside the biofilm create chemical heterogeneity resulting in changes in the transcriptome which eventually aid to the emergence of phenotypes like antibiotic tolerant or dormant cells [3].

Streptococcus pneumoniae (*S. pneumoniae*) naturally colonizes the upper respiratory tract and can cause invasive disease. *S. pneumoniae* changes its transcriptome profile depending on the colonizing tissue [7]. Previous transcriptomic analysis on 19F serotype *S. pneumoniae* biofilms after 24 hours of growth showed a complex transcriptional rearrange in the initial stages of biofilm formation. Throughout this study, the authors identified gene categories such as carbohydrate metabolism and virulence to be enriched under biofilm growth. In addition, they found that dispersed bacteria upon temperature or virus treatment yielded over 60% of the coding genes to have a significant change in expression when compared to planktonic growth [8]. While this study offered some insights into the transcriptional differences between biofilm and dispersal cells still does not fully answers what happens to *S. pneumoniae* biofilms older than 24 hours.

Throughout this thesis we have shown biofilms, planktonic and dispersal from *S. pneumoniae* cells are phenotypically different from each other *in vitro*. In addition, we have shown that disease severity changes depending on the type of inoculum (biofilm, planktonic and dispersal cells). The purpose of this appendix was to compare gene expression differences of three-day old biofilms and dispersal cells against planktonic growth to find possible explanations why these three states of growth might differ from each other. We find changes in expression common to biofilm and dispersal but also unique to the sessile (biofilm) and motile (dispersal) state. In the second part of the appendix, we briefly discuss how low and high biofilm former strains may change their transcriptome with respect to planktonic growth.

2. Results

Considering the distinct *in vitro* and *in vivo* phenotypes of the three different states of the lifecycle (planktonic, biofilm, and dispersal cells), we hypothesized that the observed differences could be explained by distinct gene expression profiles for each state. To test this hypothesis and gain a better understanding of the phenotypic alterations associated with biofilm and dispersal cells we use TIGR4 for all transcriptomic experiments. Four biological replicates for each mode of growth were collected using the *in vitro* model. Planktonic cells were in exponential phase 0.4-0.5 OD₆₀₀, and biofilm and dispersal cells were collected at day 3.

RNA isolation was done utilizing a lab-protocol adapted the RNeasy Mini kit from Qiagen. While this protocol was successful in producing high-quality RNA for planktonic cells; it failed to produce high quality RNA for biofilm and dispersal samples (**Figure 1** – pink dots). We reasoned that improvement on the quality of isolated RNA from biofilm and dispersal samples is critical for downstream applications. In an attempt to improve RNA integrity, we tested a low-output RNA protocol (Micro kit from Qiagen with some modifications). The RNA integrity is assessed by RNA integrity (RIN) score on TapeStation (Agilent). RNA isolation using the Micro kit protocol generated the best quality

RNA for biofilm as dispersal samples, with an average RIN score of 7 comparing to a RIN score of 2-4 from the Mini-kit protocol.

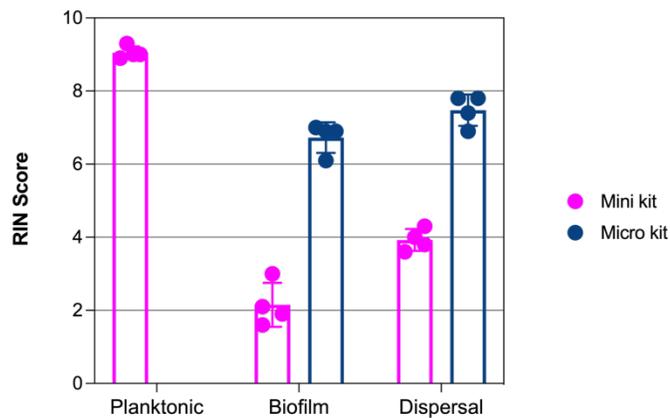


Figure 1. Optimization of RNA isolation from biofilm and dispersal cells. RNA integrity (RIN) scores determined by TapeStation 4200 (Agilent). Each dot represents an independent sample for RNA isolation. Mini kit on the left was unable to yield good quality RNA from biofilm and dispersal samples (Pink dots). The micro kit protocol was used for biofilm and dispersal samples downstream applications (Blue dots).

Differentially expressed genes under biofilm and dispersal conditions

Sequencing reads were aligned to the *S. pneumoniae* T4 genome (NC_003028), as described in Material and Methods. A total of 51-70 million reads were obtained for each experimental condition (Supplementary Table 1). Read counts were determined using feature Counts[9] and used for differential gene expression analysis using DEseq2 [10]. To identify differentially expressed genes during biofilm and dispersal growth, the mRNA- specific read counts are compared from the planktonic samples. In total, of the 1415 expressed genes during biofilm growth, 878 (62%) were significantly and differentially expressed (442 downregulated and 436 upregulated), and these were distributed across various functional categories (Figure 2A). Conversely, 912 genes were significantly and differentially expressed in the dispersal population (417 upregulated and 495 downregulated – Figure 2B). 734 genes were shared between biofilm and dispersal cells, while biofilm has 137 and dispersal cells have 144 specifically differentially expressed cells. This shows that while most of the

transcriptional alterations are common to both conditions, each can be defined by a specific set of transcriptional changes (**Figure 2C**).

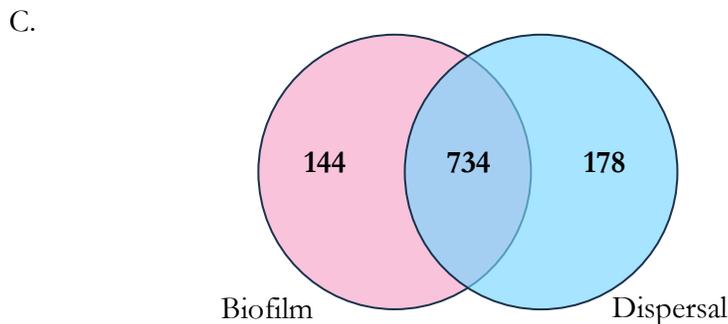
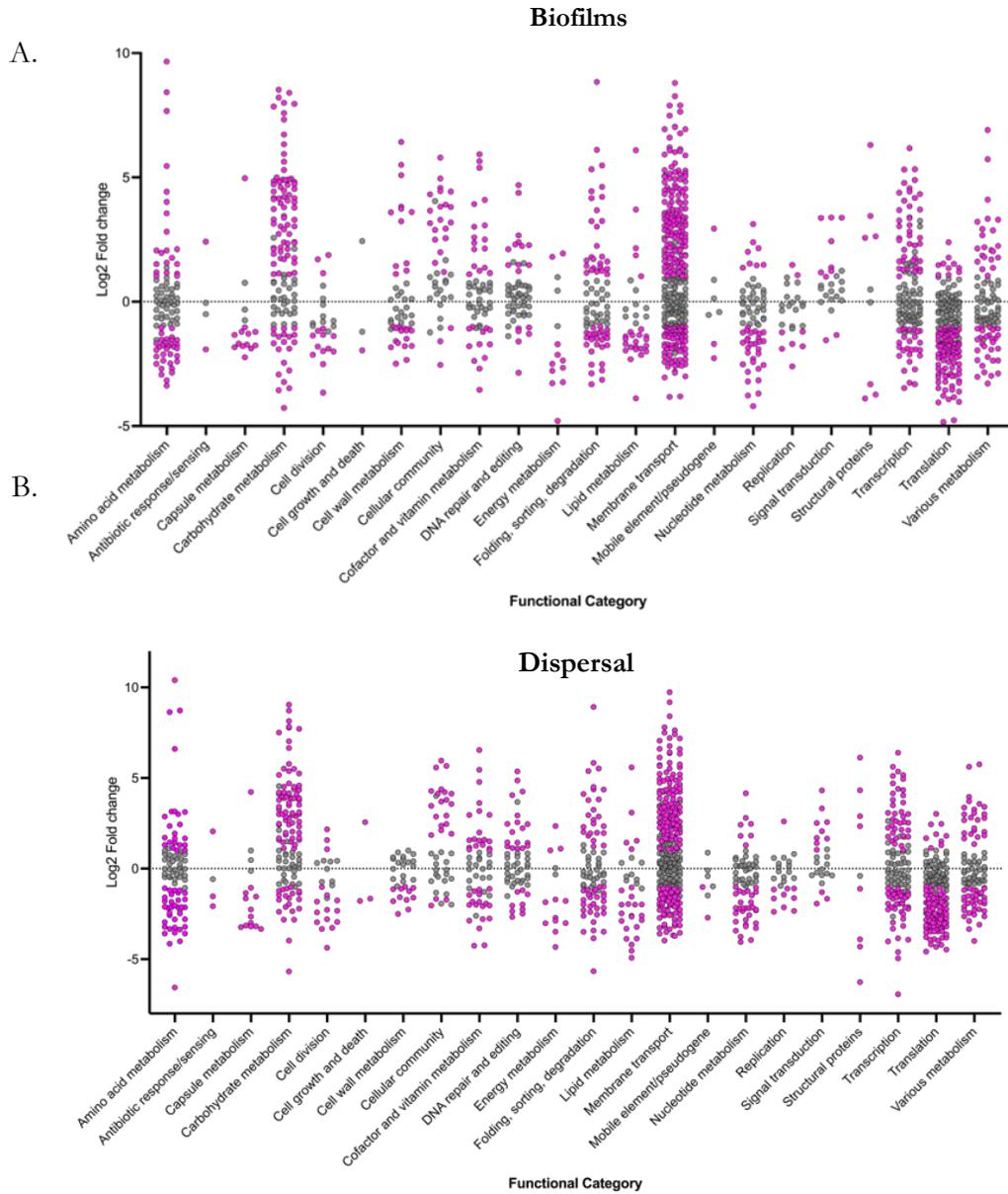


Figure 2. Transcriptional landscape of biofilms and dispersal cells. Distribution of gene functional category of differentially expressed (DE) genes **A.** Biofilms **B.** Dispersal cells. **C.** Venn diagram indicating the number of DE genes shared between conditions and unique for biofilms and dispersal cells.

To gain deeper insights on what where the differentially expressed (DE) genes, one approach we used was looking at the gene function and pathway enrichment. In this analysis we compared the number of DE genes compared to the overall genomic distribution. An adjusted p-value was calculated for each process, and processes $p < 0.05$ were considered statistically enriched. Functional enrichment analysis identified processes like carbohydrate and capsule metabolism to be enriched in biofilm samples. Dispersal samples were enriched in replication, peptidoglycan biosynthesis and nucleotide metabolism (**Figure 3**). In the coming sections we describe in detail some of the transcriptional shared and unique patterns among biofilm and dispersal cells.

Of the shared transcriptional patterns, some have already been extensively studied for their role in biofilm formation, such as capsule metabolism. Microarray, single knockout, metabolomics, proteomics, and RNA sequencing studies have revealed that *S. pneumoniae* cells reduce capsule production to form biofilms [8, 11-14]. In agreement with this, our data shows that all genes expressed in the TIGR4 *cps* locus are downregulated in both biofilms and dispersal cells (**Figure 4A**). Capsule residues are mostly carbohydrates. It is then likely that the regulation of other genes of the carbohydrate metabolism functional category can directly or indirectly influence the biosynthesis of the polysaccharide capsule. For instance, Cools and colleagues found that the absence of the highly conserved UTP-glucose-1-phosphate uridylyltransferase (*galU*) gene, which is involved in glucose and galactose metabolism, has a direct effect on capsule production, and consequently increases bacterial adhesion to epithelial cells and promotes biofilm formation [15]. The *galU* (SP_2092) gene in biofilm and dispersal cells showed a 4-fold decrease in expression which is thus in agreement with previous findings and confirms the association with biofilm formation.

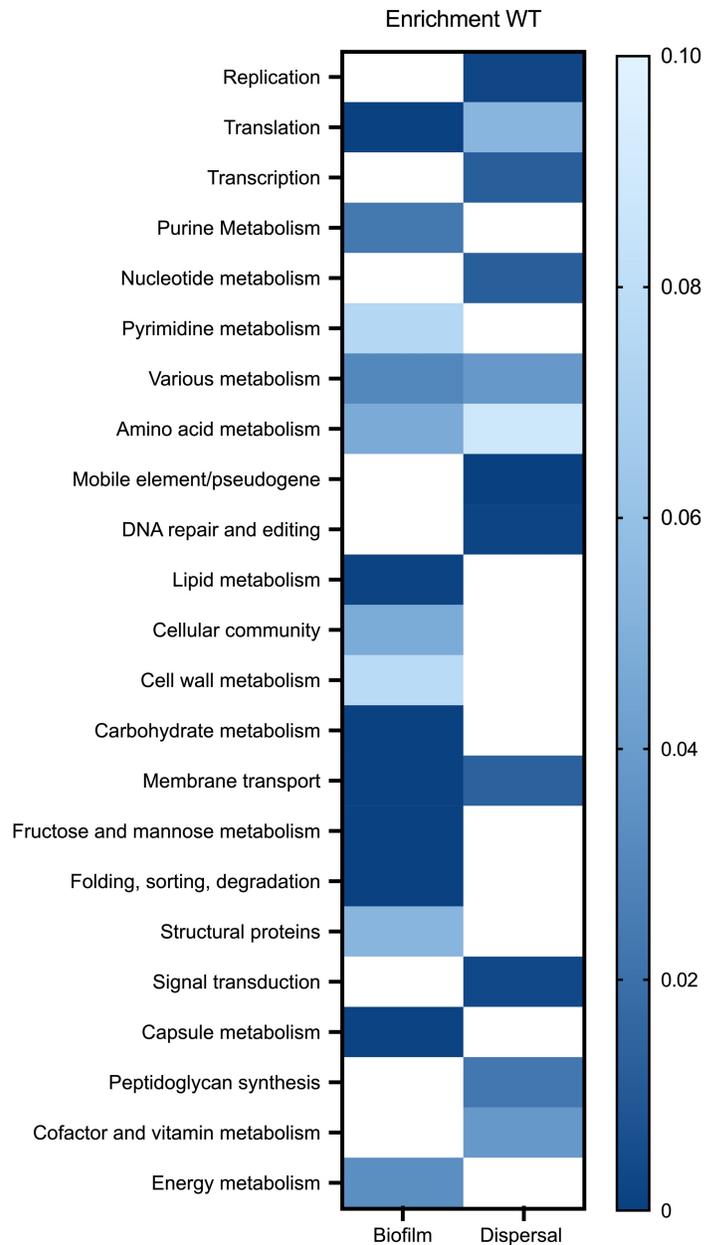


Figure 3. Enrichment analysis of all DE genes significant under both biofilm and dispersal cells revealed a multitude of processes key for three-day old biofilms. Enriched functional categories in biofilm individual datasets are presented as a heatmap with the corresponding p-values of false discovery rates. p-value scale is ranging from 0 to 0.10 shown in the figure key.

Streptococci regulate carbohydrate metabolism through carbon catabolite repression (CCR), a process that entails transcriptional control by the catabolite control protein CcpA and the uptake of

carbohydrates [16]. Around 75% of the expressed genes in the carbohydrate metabolism functional category underwent significant changes under biofilm conditions, resulting in one of the top enriched categories in our functional category enrichment analysis (**Figure 3**). The genes coding for β -galactosidase *bgaA* (SP_0648) and neuraminidase B *nanB* (SP_1687) are 16-18 times upregulated in biofilms and dispersal cells, respectively (**Figure 4C**). In *S. pneumoniae*, these surface glycosidases work together to harvest host glycans and release carbohydrates that are used as a carbon source, which is associated with enhanced virulence [17-19]. The anchored neuraminidase A coding gene *nanA* (SP_1688) and *bgaA* have been experimentally confirmed to promote biofilm formation[20], while the role of *nanB* has yet to be confirmed [17]. However, our work suggests that *nanB* might also contribute to biofilm formation indicated by the exacerbated upregulation in our biofilm samples.

S. pneumoniae harbors a significant number of transporters, including phosphotransferase (PTS) systems, which allows the bacterium to utilize various carbohydrates for metabolic purposes [21]. TIGR4 has 19 PTS systems, and at least one gene of each PTS system is upregulated in biofilms and dispersal cells (**Figure 4B**), except for the Mannose Phosphotransferase System (Man-PTS), which is considered an important entry route for the preferred carbon source, glucose [22]. We previously identified the Man-PTS operon as playing a fundamental role in biofilm formation as depicted by our Tn-seq dataset and experimental evolution. Our previous studies have identified the Man-PTS operon as playing a critical role in biofilm formation, as highlighted by our Tn-seq dataset and experimental evolution. Our findings suggest that a deficiency in Man-PTS gene members is essential for enhanced biofilm formation. Consequently, the downregulation of this PTS system observed in our transcriptomic data indicates that reduced production of the Man-PTS operon is beneficial for promoting biofilm growth.

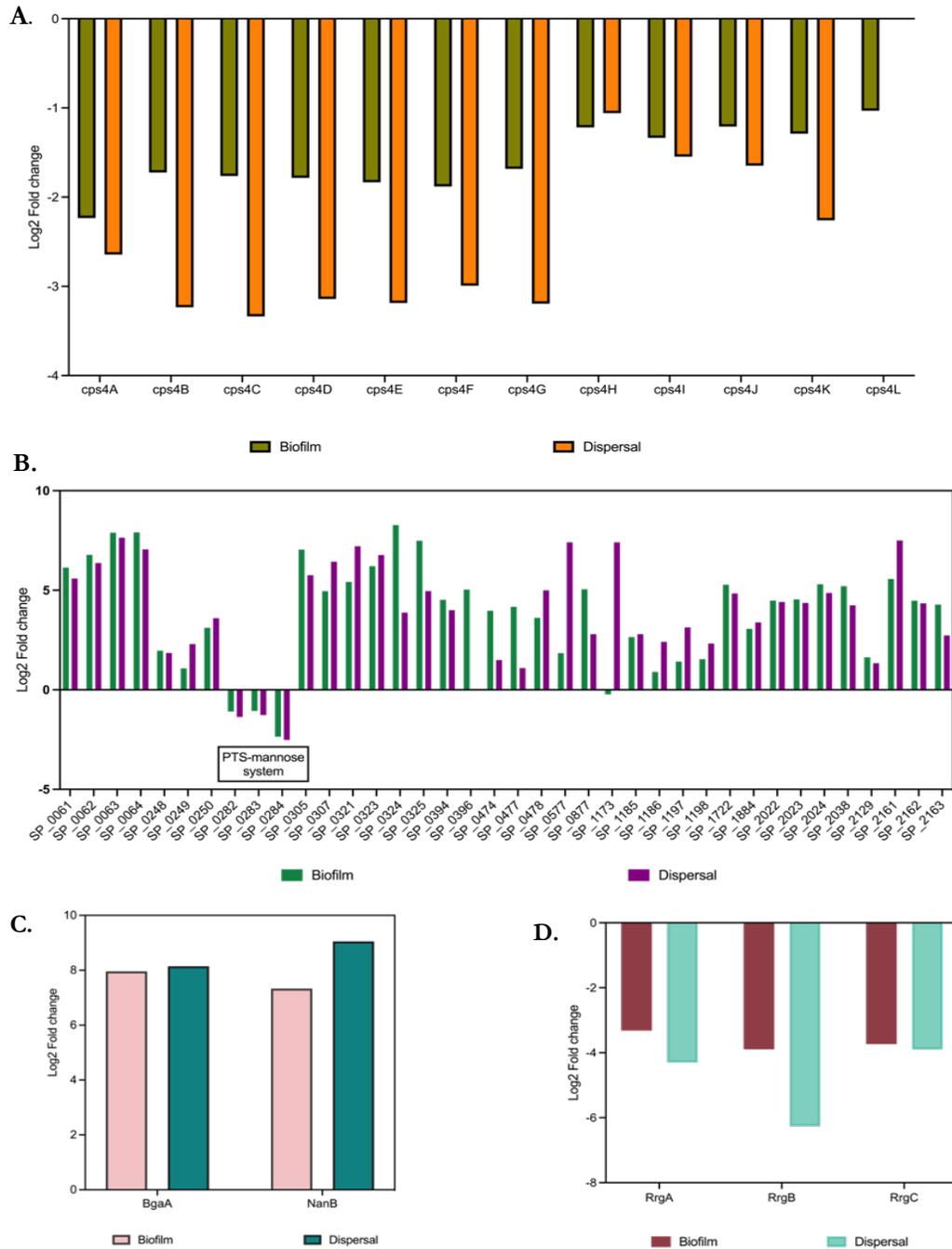


Figure 4. Multiple avenues for regulation during biofilm and dispersal states. **A.** Capsule operon is downregulated during both states of growth. This agrees with previous published work on *S. pneumoniae* having to produce less capsule to become a better biofilm former and colonizer. **B.** *S. pneumoniae* overexpresses gene members of all PTS transporter systems, probably to optimize the uptake of the available carbon sources. **C.** Upregulation of the virulence factors *bgaA* and *nanB* is essential for the life cycle of biofilms. **D.** Genes essential for the assembly of the type 1 pilus are downregulated in biofilm and dispersal cells.

Next, we identified that structural proteins are enriched in biofilms and dispersal cells. TIGR4 has a type 1 pilus which is assembled by the structural proteins encoded by *mgA* (SP_0462), *mgB* (SP_0463), *mgC* (SP_0464), the Rof-A- like transcriptional regulator *r/rA* (SP_0461) and three sortases [23, 24]. All three rrg structural proteins are under-expressed in biofilm and dispersal cells (**Figure 4D**). Pili plays a crucial role in interacting with various host cell types which ultimately has an effect on virulence. Among the assembly genes, only *RrgA* has been associated with biofilm formation [25]. Expression of pilus 1 oscillates in vivo, peaking in the early infection stages (around day 1) and declining during the later stages of colonization (on day 7). This suggests bacteria regulate pili production dynamically, enhancing colonization initially, and reducing it later, possibly to evade the host immune response once infection is established [26].

Unique transcriptional signatures in biofilms and dispersal cells.

The unique differentially expressed genes in each state were examined with the goal of uncovering distinct features of growth for each state (biofilm and dispersal cells) and potentially finding connections with the hypervirulence phenotype of dispersal cells showed in Chapter 3. There are 144 genes, distributed among all functional categories, that are unique for biofilm growth. Among the group of genes specific for biofilm growth we found those related to the competence pathway. Competence in *S. pneumoniae* is essential for the survival of the bacterium and is the primary route of Horizontal Gene Transfer (HGT). In biofilms, horizontal gene transfer plays a pivotal role in shaping microbial communities by facilitating the exchange of genetic material between different species and it is thought to be crucial for maintaining diversity [27, 28].

In bacteria, HGT primarily occurs through three major routes: direct cell-to-cell contact (conjugation), transfer facilitated by bacteriophages (transduction), or the uptake of DNA from competent cells (transformation) [27]. In *S. pneumoniae* competent cells can readily take up extracellular DNA and

safeguard the bacterium from the fratricide mechanism induced during this state [29-31]. Our results showed that *comFC* (SP_2207) and *comGF* (SP_2048) genes, members of the late competence machinery in *S. pneumoniae*, are eight-fold higher expressed in biofilm samples. Competence is induced by the secreted Competence stimulated peptide (CSP) *ComM* (part of the early competence pathway)[32]. When the bacterium detects CSP, it initiates an activation cascade of early competence genes, followed by activation of the late competence genes involved in the formation of a type IV pilin. This pilus then binds to dsDNA and mechanically introduces the dsDNA into the cell through the transmembrane partners of the ComGF and ComFC complex [33]. The upregulation of the competence genes suggests that biofilms prefer environments for transferring DNA from the cell to the environment.

Cellular processes such as transcription, nucleotide metabolism and peptidoglycan biosynthesis are enriched in dispersal cells (**Figure 3**). One of the distinct changes in dispersal cells is the upregulation of the *dltA* gene, which plays a crucial role in the incorporation of D-alanine into the teichoic acid layer, being 6 2fold higher when compared to planktonic growth. In *S. pneumoniae*, a functional *dlt* operon is necessary to neutralize the negative charge of the cell wall by incorporating positive residues [34]. These changes in surface charge could act as a defense mechanism against antimicrobial peptides (CAMPs). In a previous study it was found that D39 strain upregulates the *dltABCD* operon in the presence of CAMPs [35, 36]. In mutants of *S. pneumoniae*, defects in this operon lead to reduced bacterial adherence, and during mature biofilm formation, it alters the structure of the EPS, which in turn affects the structure [37]. We hypothesize that bacteria require a higher degree of d-alanylation on their teichoic cell walls after being released from biofilms. This conformation is outside the scope of our study, but our data provides a solid foundation for further experiments.

Transcriptomic differences between low and high biofilm formers.

The first part of this Appendix described the major changes that occur at the transcriptomic level to cells when forming biofilms. In chapter 4 we showed how low and high biofilm formers from a same background have a difference in disease outcome. For instance, the low biofilm forming strains (Δ SP_1645 and Δ SP_1739) were able to migrate to the lungs unlike WT TIGR4 and the high biofilm former strain (Δ SP_2205). In addition all mice infected with dispersal cells exhibited a hypervirulence phenotype in the lungs reflecting no difference compared to WT. To elucidate how a high biofilm strain differs from a low biofilm former, RNA-sequencing was performed on the high biofilm former strain Δ SP_2205 and the two low biofilm formers Δ SP_1645 and Δ SP_1739. Total RNA from biofilm, planktonic, and dispersal samples were used to generate RNA-seq reads and mapped to the TIGR4 genome. For pairwise differential expression analyses, each strain was compared within each state of growth and against the WT control. For example, all planktonic samples from mutants were compared against TIGR4

Figure 5A shows Log2Fold Change of all three mutants in the three different states of growth. Overall biofilms and dispersal samples have a higher number of differentially expressed (DE) genes than planktonic samples. We also observed that the high biofilm former Δ SP_2205 has the least amount of DE genes under the three different growth states when compared to the low biofilm formers. In order to identify systems-level patterns in the whole transcriptome, a Principal component analysis (PCA) was performed using the messenger RNA abundances (feature counts normalized by TPM) for each sample (**Figure 5B**). The first component clearly separates planktonic from biofilms and dispersal cells. This is expected considering the large transcriptome changes between planktonic and their counterparts as observed in the first part of this Appendix (**Figure 2**). We also noticed that Δ SP_1645 biofilms and dispersal replicates are closer to the planktonic samples in the PC1. This may explain why animals inoculated with biofilm samples mostly behaved as those inoculated with planktonic cells.

Principal Component 2 (PC2) separated Δ SP_1739 samples from the rest. This separation indicates a high number of changes on the transcriptome from this mutant, probably because Rnase Y participates in RNA turnover and maintenance, having a major effect on mRNA abundances. The Y-axis (PC2) also separated WT dispersal samples from Δ SP_1645 and Δ SP_2205, locating them closer to the dispersal and biofilm samples from Δ SP_1739.

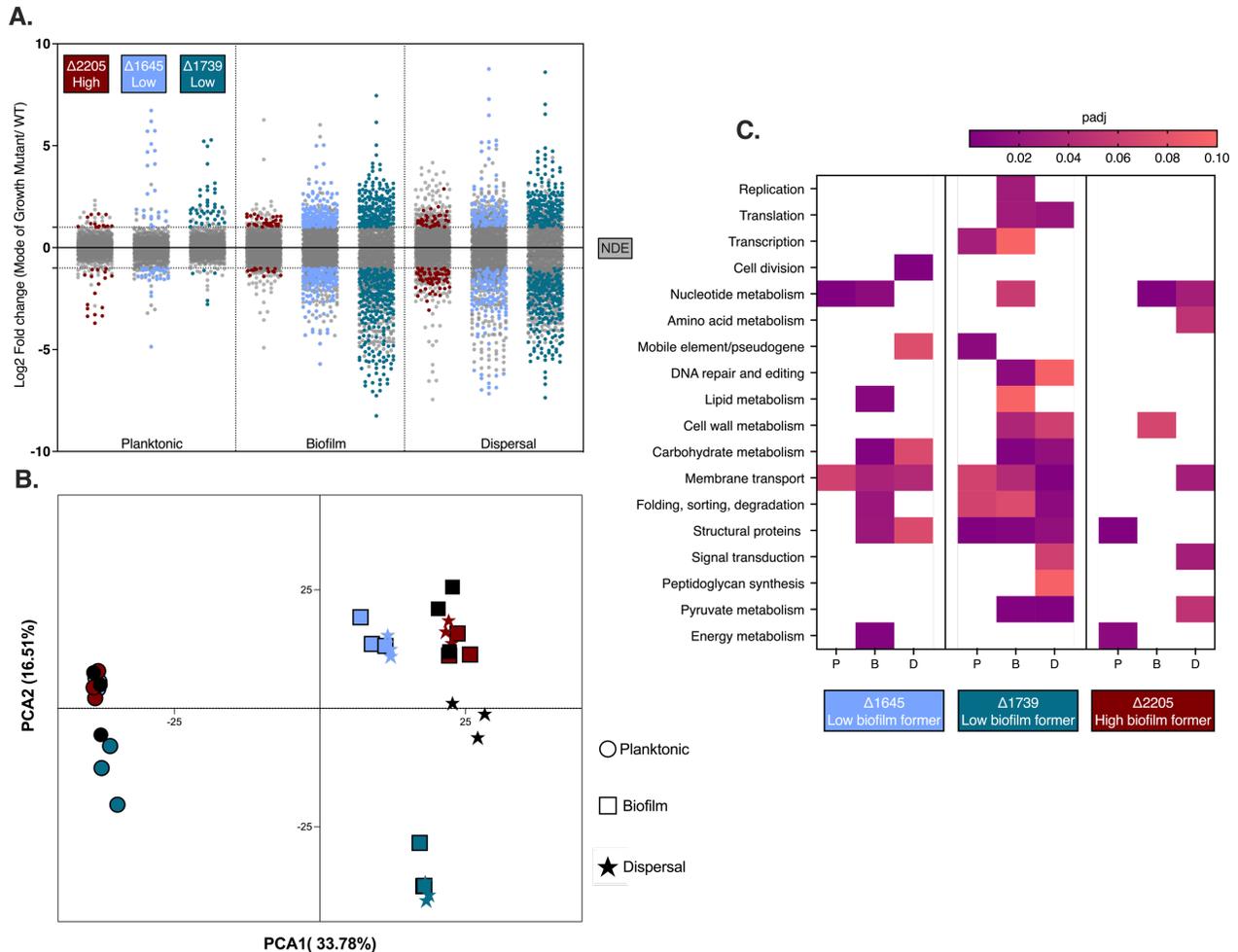


Figure 5. Different transcriptional responses among low and high biofilm strains. A. Distribution of Log₂Fold Change for all mutants in three states of growth. **B.** Two-dimensional (PC1-PC2) principal component analysis on transcriptome of planktonic(dots), biofilm (squares) and dispersal (stars) of all replicate samples from WT T4 (black), Δ SP_1645 (lilac), Δ SP_1739 (teal) and Δ SP_2205 (maroon). PC1 separates biofilm and dispersal conditions from planktonic samples. PC2 separates Δ SP_1739 from the rest of the samples. **C.** Enrichment analysis of DE genes for each mutant under each state of growth. Enriched functional categories are presented as a heatmap with the

corresponding q-values of false discovery rates. q-value scale is ranging from 0 to 0.10 shown in figure key.

We then identified the top 10 contributors for each PC component. The 10 contributor genes that mostly separated planktonic from biofilm and dispersal cells (PC1) belong to membrane transport, carbohydrate metabolism, membrane transport or have unknown functions (Table 1). This confirms the role of carbohydrate metabolism as a key signature in biofilm samples.

Table 1. Top 20 contributors for PC1, Log2Fold Change of comparison Biofilm vs planktonic is reflected. Yellow means upregulated and gray means downregulated.

Locus	Gene Description	Category	T4_Bio	Δ 1645 Bio	Δ 1739 Bio	Δ 2205 Bio
SP_0062	PTS system, IIC component	Membrane transport	6.77559819	7.032203024	7.86181612	7.288141551
SP_0065	sugar isomerase domain protein AgaS	Carbohydrate metabolism	8.00013915	7.430419963	9.23656977	7.514027491
SP_0066	aldose 1-epimerase	Carbohydrate metabolism	8.52846125	7.927177282	8.92680149	8.980647659
SP_0067	hypothetical protein	NA	8.52846125	7.927177282	9.15483286	8.40960868
SP_0068	hypothetical protein	NA	8.57506524	8.847161422	9.03081696	8.034773792
SP_0069	choline binding protein I	NA	7.96623082	7.406743758	9.29741565	8.363849267
SP_0110	hypothetical protein	Cellular community	4.582146736	3.391012985	4.2533033	5.161788079
SP_0338	putative ATP-dependent Clp protease, ATP-binding subunit	Folding, sorting, degradation	8.843328465	9.230469696	8.53087135	9.007152849
SP_0498	endo-beta-N-acetylglucosaminidase	Carbohydrate metabolism	6.331695002	3.564999268	5.86950777	6.406331531
SP_0648	beta-galactosidase	Carbohydrate metabolism	7.960431842	7.60246618	8.047735	7.934518503

An examination of the functional categories impacted under each condition was conducted through enrichment analysis, as previously described in the first section of this appendix. The results of the enrichment analysis revealed the presence of unique and shared pathways that are critical for each strain background, and which may vary depending on the state of growth (**Figure 5C**). Among the three mutans, Δ SP_1739 exhibited the highest number of enriched categories, suggesting that

significant changes occurred across functional categories. The large number of changes may be attributable to the function of this gene, mRNA turnover, as the deletion may trigger an exacerbated response in the overall transcriptome of the samples. This could potentially explain why we observed replication, translation, and transcription categories enriched in this particular mutant.

Transcriptional responses in a high biofilm former strain (Δ SP_2205)

The high biofilm former Δ SP_2205 has the lower number of significant changes across all three states of growth (**Figure 6**). A total of 45 significant genes were observed on the Δ SP_2205 mutant under biofilm conditions (14 downregulated and 31 upregulated). In planktonic growth there is seems to be a relationship between increased c-di-AMP levels an reduce pili production as all *rygA-B* genes are downregulated. In the Δ SP_2205 background, the SP_2176 (*dltA*) gene, responsible for initiating the cascade of alanine decoration of the teichoic acid layer [38] and part of our validation knockout panel, was observed to be upregulated in biofilms but downregulated in the dispersal state. For dispersal cells, multiple membrane transport genes, including ABC transporter permease (SP_1306, SP_2222), copper efflux (SP_0729), and potassium transport (SP_0078 *trkG*), appeared to be upregulated. The latter, SP_0078 has been shown to be key for potassium uptake under c-di-AMP depletion[39]. Interestingly, there are no DE genes shared among the three states in the Δ SP_2205. This lack of a common response suggests that the role of cAMP varies among different cell states.

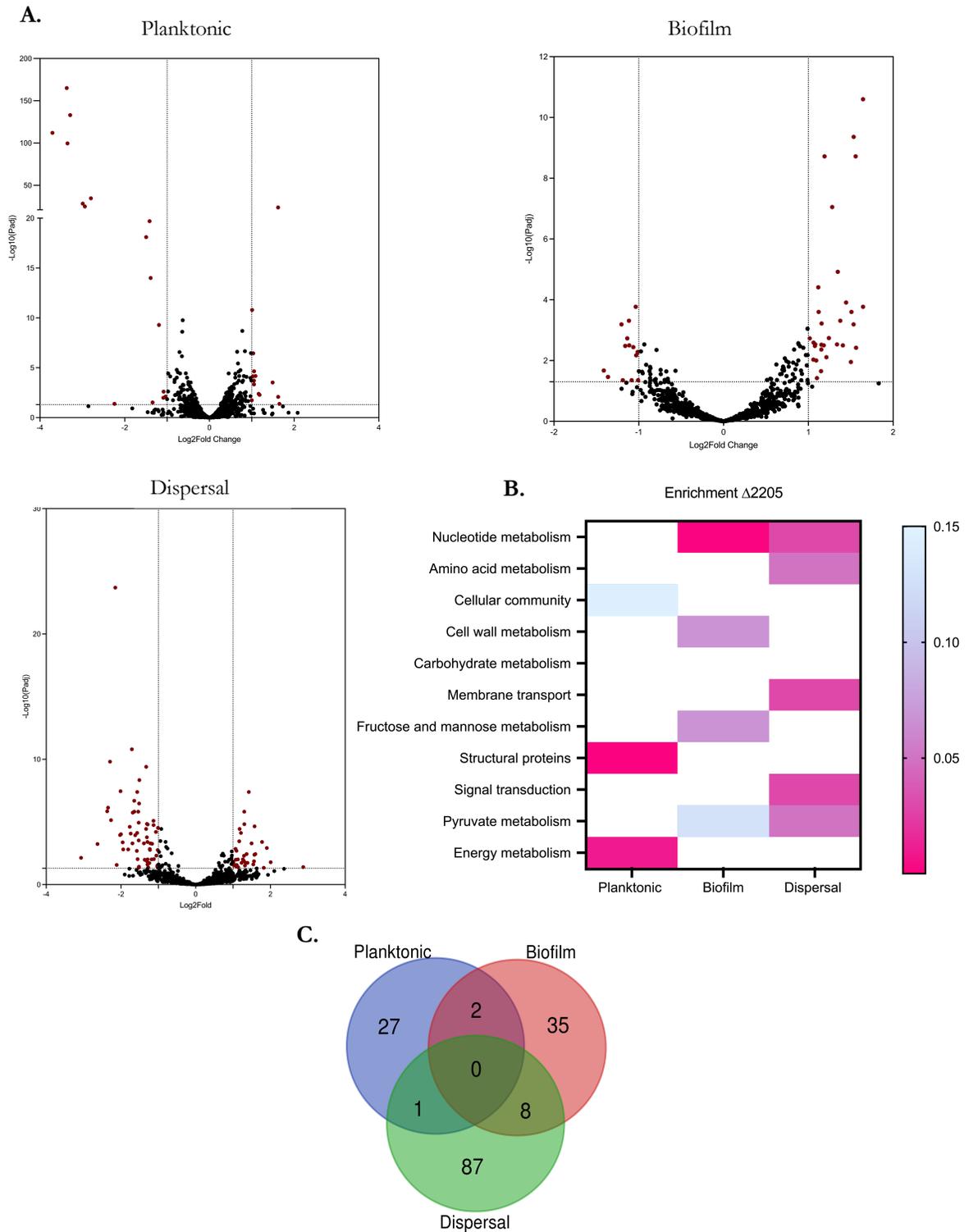


Figure 6. Summary data visualization of DE genes in Δ SP_2205. A. Volcano plots for each state of growth (DE genes are colored in dark red). **B.** Enrichment analysis shows pathways important under each state of growth. **C.** Venn diagram shows number of shared and unique genes among states of growth.

A glimpse into the transcriptional responses in two low biofilm former strains (Δ SP_1645 and Δ SP_1739)

For the Δ SP_1645 mutant, in planktonic growth a total of 52 DE genes were observed (30 downregulated and 22 upregulated genes). In biofilms, there were 321 genes (156 downregulated and 165 upregulated) while 234 significant genes showed in the dispersal condition (131 downregulated and upregulated). All three states of growth in the Δ SP_1645 background showed downregulation of the purine operon and upregulation of genes annotated as magnesium transport genes, indicating these changes in genes related to (p)ppGpp metabolism are independent from the cell state (**Figure 7**).

The other low biofilm former Δ SP_1739 had the highest number of transcriptome changes in the three states of growth which is expected as the absence of this RNase impacts the transcriptome, growth, and virulence in *S. pneumoniae*[40]. The homolog Δ SP_1739 in D39 results in abnormal cell morphology, which agrees with the RNA-seq results obtained here that show that genes members of the cell division pathways such as *ftsE*, *ftsX*, *ftsK* are downregulated in biofilm and dispersal growth. Pili upregulation is observed in the three states of growth for the Δ *rly* mutants, but the expression is even higher on the biofilm and dispersal samples (**Figure 8**).

Further exploration of this data set can open avenues to select gene expression markers characteristic of biofilm or dispersal cell stages. For instance, the main contributors to the principal components could serve as genes whose expression levels define if a cell is in a biofilm-like or dispersal-like transcriptome. Having data from strains with different biofilm-forming capacities enables a more general outlook for the possible selection of genes relevant under biofilm-formation conditions

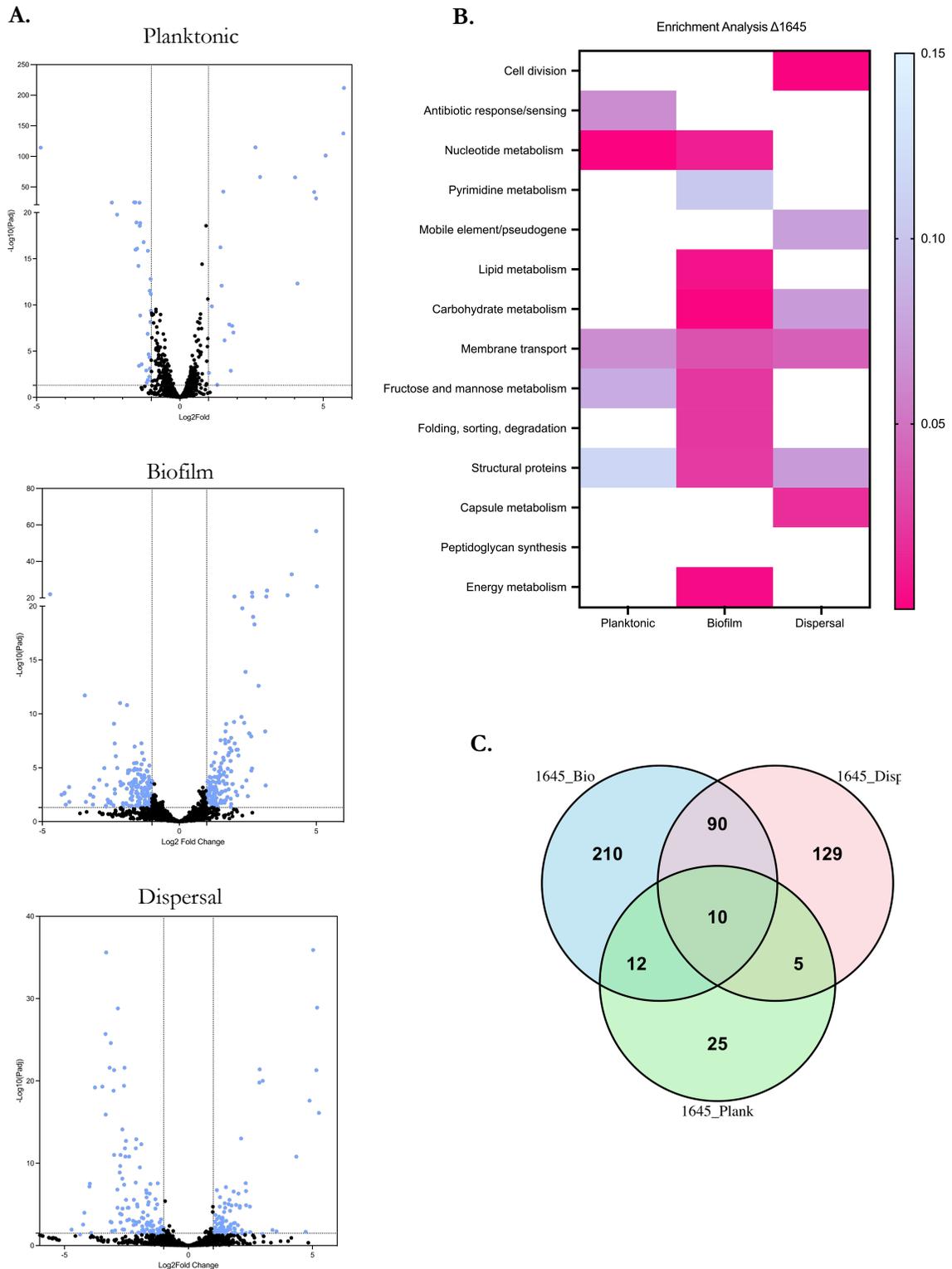


Figure 7. Summary data visualization of DE genes in ΔSP_{1645} **A.** Volcano plots for each state of growth (DE genes are colored in lilac). **B.** Enrichment analysis shows pathways important under each state of growth. **C.** Venn diagram shows number of shared and unique genes among states of growth.

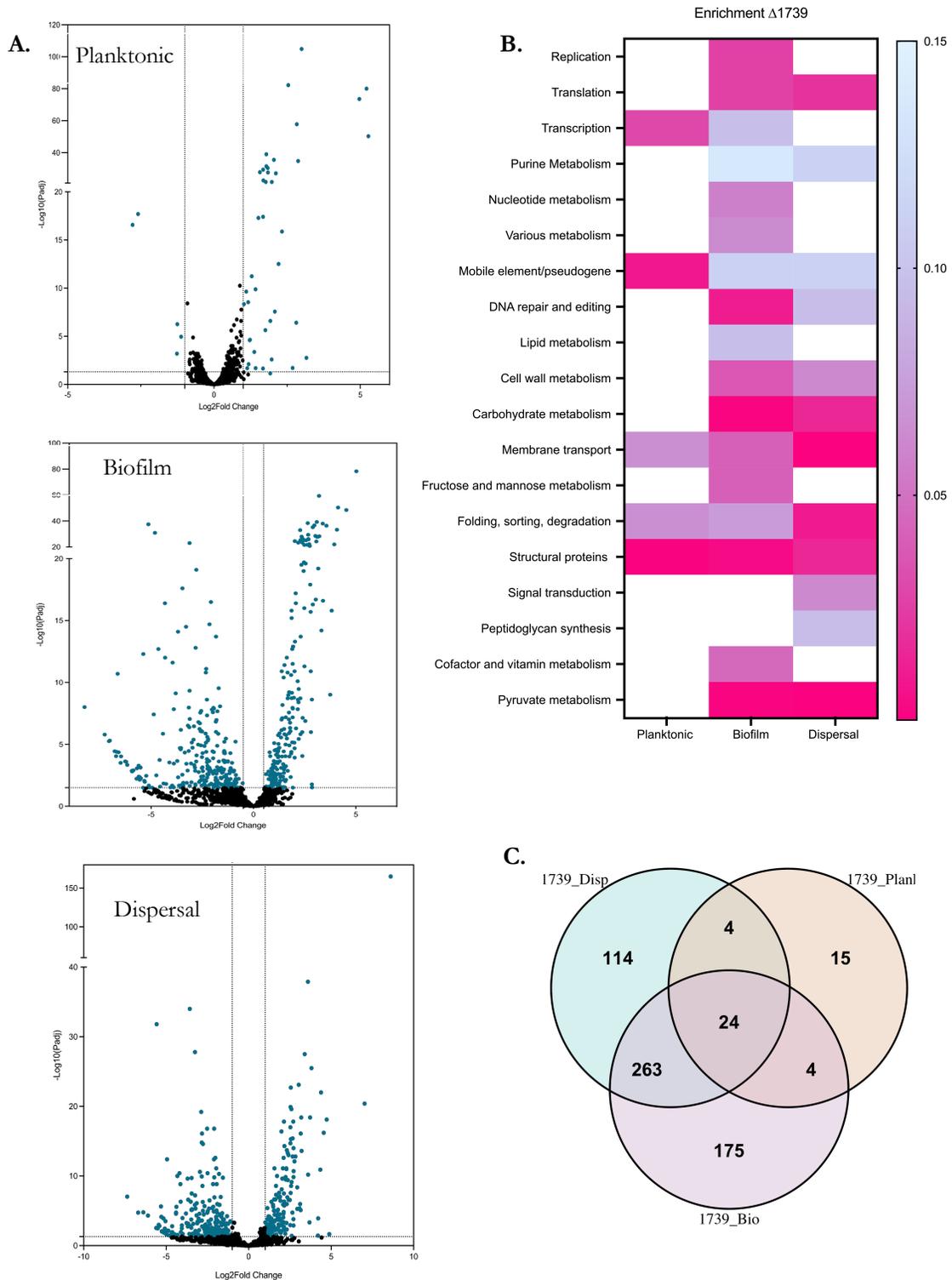


Figure 4. Summary data visualization of DE genes in ΔSP_{1739} . **A.** Volcano plots for each state of growth (DE genes are colored in teal). **B.** Enrichment analysis shows pathways important under each state of growth. **C.** Venn diagram shows number of shared and unique genes among states of growth.

Conclusions

In this thesis, we have demonstrated the phenotypic distinctions among planktonic, biofilm, and dispersal cells *in vitro* (Chapter 3). Furthermore, utilizing these distinct cell states independently as inocula in *in vivo* experiments revealed variations in disease severity. We also showed that strains with different biofilm formation capacities exhibited different phenotypes *in vivo* (Chapter 4). To explore the underlying factors that are changed in biofilm and dispersal cell, we hypothesized that scrutinizing the transcriptome of these populations could unveil some of these differences. The primary objective of this Appendix Chapter is to offer some insights into the transcriptomic responses of *S. pneumoniae* biofilms and dispersal cells in comparison to planktonic growth.

RNA sequencing revealed a plethora of genes that were either upregulated or downregulated in biofilm and dispersal cells. Approximately 75% of the differentially expressed (DE) genes were common to both stages, indicating the highly dynamic nature of the biofilm life cycle. A multitude of processes in biofilm and dispersal cells undergo changes in expression to maintain the mode of growth. For instance, decreasing the expression of capsule, pili proteins, and key carbohydrate metabolic players is crucial for the establishment of surface-attached (biofilm) cells. Among the numerous differences in transcriptome responses between biofilm and dispersal cells, we highlight how the late competence system is upregulated in biofilms, suggesting a higher rate of DNA uptake and exchange facilitation among cells within the biofilm.

When switching our attention to high and low biofilm formers, we found that lower biofilm formers changed their transcriptome at a much bigger scale when compare to WT TIGR4. We recognize the data in this chapter can be further analyzed and used for future experiments. The comprehensive dataset generated by our research serves as a valuable resource, providing a wealth of information for future investigators interested in exploring the genetic complexities of biofilm development in *S. pneumoniae*. It would be beneficial to understand at a mechanistic level why a low biofilm former strain

is able to cause disease in the lung, or perhaps exploring why dispersal cells are highly virulent in mice. Over time, extensive transcriptomic studies in biofilms can provide avenues for the design and discovery of antibiofilm or anti-dispersal therapeutics, contributing to the ongoing efforts to develop more effective strategies against biofilm-related infections.

Methods

RNA extraction and purification

Planktonic samples: Four biological replicates were grown in a 5ml SDMM to an early/mid log-phase, and were then harvested by centrifugation at 4 °C. The samples were then snap-frozen and stored at -80 °C for RNA isolation. Total RNA was isolated using a mechanical disruption method on a BioSpec Mini-beadbeater-16, followed by the use of a RNeasy Mini kit (Qiagen).

Biofilm and dispersal samples: Four biological replicates of a 3-day-old biofilm and dispersal cells were grown, and cells were collected using RNALater (Invitrogen) instead of 1XPBS to prevent RNA degradation. RNA was extracted on the same day of sample collection by mechanically disrupting cell pellets using PowerBead Pro and subjecting them to a low input RNA isolation using RNeasy Micro kit following the manufacturer's instructions. The eluted RNA was treated with diethyl pyrocarbonate (DEPC)-treated water (Qiagen) and 1µL of murine RNase inhibitor was added immediately after elution. The RNA integrity and concentration were assessed using TapeStation (Agilent).

Library preparation and RNA sequencing: 400ng of total RNA from each sample was used to generate cDNA libraries in accordance with the RNAtag-Seq protocol[41] with modifications described in the following section. Prior to ribosomal RNA depletion, each sample (in 26µL) was concentrated to a final volume of 16µL using Zymo RNA clean and concentrator -5 (Zymo Research) and added with 1µL of murine RNase inhibitor immediately after elution. rRNA was depleted using a custom a RnaseH approach[42] consisting of oligo probes complementary to rRNA unique to *S. pneumoniae* TIGR4. rRNA-depleted samples were cleaned up using SPRI beads (Agencourt) and subjected to reverse transcription, Illumina-compatible adapter ligation, end repair, PCR enrichment, and indexing using oligos following the manufacturer's instructions. Twelve enriched and indexed

libraries were pooled and paired-end sequenced on a NextSeq500 (Illumina) with 150-bp reads to ensure sufficient sequencing depth.

Bioinformatic analysis and data visualization: RNA-Seq data was processed and analyzed using Aerobio, an in-house developed analysis pipeline. In brief, raw reads were demultiplexed, trimmed to 150-bp and quality filtered (96% sequence quality > Q14). Filtered reads are mapped to both *S. pneumoniae* T4 reference genome (NC_003028.3) using bowtie2. Mapped reads were aggregated by Feature Count[9] and differential expression (DE) is calculated with DESeq2[10]. DE comparisons are made against planktonic growth. All DE comparisons are filtered based on two criteria: $|\log_2\text{FoldChange}| > 1$ and adjusted p-value (p_{adj}) < 0.05. Data visualization was done using ShinyOmics[43] and GraphPad Prism v.10.

Table 1. Sequences of adapters and primers

RNABC01-12

Number	Sequence	6nt BarCode+T
RNABC01	/5Phos/rArUrU rGrCrU rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	AAGCAAT
RNABC02	/5Phos/rArUrG rArArU rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	AATTCAT
RNABC04	/5Phos/rArGrG rCrUrG rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ACAGCCT
RNABC05	/5Phos/rArUrC rArGrG rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ACCTGAT
RNABC06	/5Phos/rArArU rUrArG rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ACTAATT
RNABC07	/5Phos/rArUrU rGrArG rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ACTCAAT
RNABC08	/5Phos/rArUrG rGrUrC rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	AGACCAT
RNABC09	/5Phos/rArGrU rUrUrA rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ATAAACT
RNABC10	/5Phos/rArArU rGrUrA rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ATACATT
RNABC11	/5Phos/rArArU rUrGrA rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ATCAATT
RNABC12	/5Phos/rArGrU rGrGrA rUrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	ATCCACT
RNABC19	/5Phos/rArUrU rCrGrU rGrArG rArUrC rGrGrA rArGrA rGrCrG rUrCrG rUrGrU rArG/3SpC3/	CACGAAT

RTS_Enr_P5_ RNAtag	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGGCTCTTCCGAT CT
RTS_Enr_P7_ BC_01	CAAGCAGAAGACGGCATAACGAGATTCGGTGTGC GTGACTGGAGTTCAGACGTGTG CTCTCCGATCT
RTS_Enr_P7_ BC_02	CAAGCAGAAGACGGCATAACGAGATTCGCCAGA GTGACTGGAGTTCAGACGTGTG CTCTCCGATCT
RTS_Enr_P7_ BC_04	CAAGCAGAAGACGGCATAACGAGATGGCTCCTC GTGACTGGAGTTCAGACGTGTG CTCTCCGATCT
RTS_Enr_P7_ BC_06	CAAGCAGAAGACGGCATAACGAGATAACATAAT GTGACTGGAGTTCAGACGTGTG CTCTCCGATCT

Blue indicates P7 specificity

Red indicates BC0_ primer

Primers for sequencing:

3Tr3 adapter: 5'-AGATCGGAAGAGCACACGTCTG/3SpC3-3'

RTS_AR2 : TACACGACGCTCTTCCGAT

3. Supplementary information:

Table 2. Average read count per sample WT TIGR4 planktonic, biofilm and dispersal samples.

Sample Name	Total # 4 lanes	Read count	Average Read count
316-Plank-a-AAGCAAT-R1.fastq.gz	17474336	4368584	
316-Plank-b-AATTCAT-R1.fastq.gz	16779540	4194885	
316-Plank-c-ACAGCCT-R1.fastq.gz	9327496	2331874	3631781.00
316-1645Plank-a-ACTAATT-R1.fastq.gz	13583384	3395846	
316-1645Plank-b-ACTCAAT-R1.fastq.gz	21261400	5315350	
316-1645Plank-c-AGACCAT-R1.fastq.gz	18695480	4673870	4461688.67
316-1739Plank-a-ATACATT-R1.fastq.gz	15197984	3799496	
316-1739Plank-c-ATCCACT-R1.fastq.gz	11350668	2837667	
316-1739Plank-d-CACGAAT-R1.fastq.gz	67514624	16878656	7838606.33
316-2205Plank-b-AATTCAT-R1.fastq.gz	41559852	10389963	
316-2205Plank-c-ACAGCCT-R1.fastq.gz	23321028	5830257	
316-2205Plank-d-ACCTGAT-R1.fastq.gz	37100572	9275143	8498454.33
316-Disp-a-ACTAATT-R1.fastq.gz	1383956	345989	
316-Disp-b-ACTCAAT-R1.fastq.gz	1973436	493359	
316-Disp-d-ATAAACT-R1.fastq.gz	1747168	436792	425380.00
316-1739Disp-a-ATACATT-R1.fastq.gz	2290764	572691	
316-1739Disp-b-ATCAATT-R1.fastq.gz	2258596	564649	
316-1739Disp-c-ATCCACT-R1.fastq.gz	2012240	503060	546800.00
316-1645Disp-b-AATTCAT-R1.fastq.gz	2728452	682113	
316-1645Disp-c-ACAGCCT-R1.fastq.gz	3550472	887618	
316-1645Disp-d-ACCTGAT-R1.fastq.gz	2092968	523242	697657.67
316-2205Disp-b-ACTCAAT-R1.fastq.gz	5558280	1389570	
316-2205Disp-c-AGACCAT-R1.fastq.gz	8032476	2008119	
316-2205Disp-d-ATAAACT-R1.fastq.gz	9950216	2487554	1961747.67
316-Bio-a-ATACATT-R1.fastq.gz	10097500	2524375	
316-Bio-c-ATCCACT-R1.fastq.gz	11577080	2894270	
316-Bio-d-CACGAAT-R1.fastq.gz	34678556	8669639	4696094.67
316-1645Bio-a-ACTAATT-R1.fastq.gz	4456592	1114148	
316-1645Bio-b-ACTCAAT-R1.fastq.gz	7486164	1871541	
316-1645Bio-d-ATAAACT-R1.fastq.gz	6156108	1539027	1508238.67
316-1739Bio-b-AATTCAT-R1.fastq.gz	5507044	1376761	
316-1739Bio-c-ACAGCCT-R1.fastq.gz	4334544	1083636	
316-1739Bio-d-ACCTGAT-R1.fastq.gz	4605772	1151443	1203946.67
316-2205Bio-b-ATCAATT-R1.fastq.gz	9361492	2340373	
316-2205Bio-c-ATCCACT-R1.fastq.gz	12827048	3206762	
316-2205Bio-d-CACGAAT-R1.fastq.gz	32456776	8114194	4553776.33

Table 3 Top 10 upregulated genes in TIGR4 WT biofilms.

Locus Tag	Log2 FC	Log10(padj)	Category	Gene Name	Gene description
SP_2150	9.65924531	105	Amino acid metabolism	<i>arcB</i>	ornithine carbamoyltransferase
SP_0338	8.84332847	287	Folding, sorting, degradation		ATP-dependent Clp protease ATP-binding subunit
SP_0326	8.8009136	52	Membrane transport	<i>yajC</i>	preprotein translocase subunit YajC
SP_0066	8.52846125	125	Carbohydrate metabolism	<i>galM</i>	aldose 1-epimerase
SP_2151	8.42405752	83.3	Amino acid metabolism	<i>arcC</i>	Carbamate kinase
SP_0318	8.40239061	49.7	Carbohydrate metabolism		Carbohydrate kinase
SP_0324	8.26924041	44.9	Membrane transport		PTS system transporter subunit IIC
SP_0317	8.21361573	58.5	Carbohydrate metabolism		keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase
SP_0065	8.00013915	197	Carbohydrate metabolism		sugar isomerase
SP_0648	7.96043184	263	Carbohydrate metabolism	<i>bgaA</i>	beta-galactosidase

Locus Tag	Log2 FC	-Log10(padj)	Category	Gene.Name	Gene description
SP_1414	-4.8450315	27.5	Translation	<i>rpsU</i>	30S ribosomal protein S21
SP_1514	-4.7978941	16.2	Energy metabolism	<i>atpE</i>	ATP synthase F0, C subunit
SP_0838	-4.7668871	30.3	Translation	<i>rpsT</i>	30S ribosomal protein S20
SP_0962	-4.2762502	33.3	Carbohydrate metabolism	<i>gloA</i>	Lactoylglutathione lyase
SP_1277	-4.2052885	33.7	Nucleotide metabolism	<i>pyrB</i>	Aspartate carbamoyltransferase
SP_1626	-4.0553041	49.8	Translation	<i>rpsO</i>	30S ribosomal protein S15
SP_2042	-4.0306089	10	Translation	<i>mpA</i>	ribonuclease P
SP_1299	-3.9260183	15	Translation	<i>rpmE2</i>	50S ribosomal protein L31 type B
SP_0463	-3.8935267	33.6	Structural proteins		cell wall surface anchor family protein
SP_0418	-3.8863595	263	Carbohydrate metabolism	<i>bgaA</i>	beta-galactosidase

Table 4 Top 10 downregulated genes in TIGR4 WT biofilms.

Table 5 Top 10 upregulated genes in dispersal TIGR4 WT.

Locus Tag	Log2 FC	-Log10(padj)	Category	Gene.Name	Gene description
SP_2150	10.3968738	92.75	Amino acid metabolism	<i>arcB</i>	ornithine carbamoyltransferase
SP_1688	9.73347159	60.75	Membrane transport		Putative ABC transporter, premease component
SP_1689	9.18060848	93.30	Membrane transport		Putative ABC transporter, premease component
SP_1687	9.05153874	130.63	Carbohydrate metabolism	<i>nanB</i>	Putative neuraminidase B
SP_0338	8.92114601	#NUM!	Folding, sorting, degradation		ATP-dependent Clp protease ATP-binding subunit
SP_2151	8.72288079	62.30	Amino acid metabolism	<i>arcC</i>	Carbamate kinase
SP_0066	8.71297189	132.42	Carbohydrate metabolism	<i>galM</i>	aldose 1-epimerase
SP_2148	8.64027797	92.96	Amino acid metabolism	<i>arcA</i>	arginine deiminase
SP_0326	8.40530805	86.46	Membrane transport	<i>yajC</i>	preprotein translocase subunit YajC
SP_0648	8.14676887	#NUM!	Carbohydrate metabolism	<i>bgaA</i>	beta-galactosidase

Table 6 Top 10 downregulated genes in dispersal TIGR4 WT.

Locus Tag	Log2 FC	-Log10(padj)	Category	Gene.Name	Gene description
SP_1278	-6.937351	7.71	Transcription	<i>pyrR</i>	Bifunctional pyrimidine regulatory protein/uracil phosphoribosyltransferase
SP_0856	-6.5679178	7.01	Amino acid metabolism	<i>ibvE</i>	branched-chain amino acid aminotransferase
SP_0463	-6.2683324	19.12	Structural proteins	<i>rrgB</i>	cell wall surface anchor family protein
SP_0962	-5.6845354	10.03	Carbohydrate metabolism	<i>gloA</i>	Lactoylglutathione lyase
SP_1427	-5.6656526	5.20	Folding, sorting, degradation		U32 family peptidase
SP_0461	-4.951729	12.65	Transcription		transcriptional regulator
SP_0418	-4.9236054	38.88	Lipid metabolism	<i>acpP</i>	acyl carrier protein
SP_1433	-4.6068651	3.46	Transcription		Transcriptional regulator, AraC family
SP_0580	-4.5810204	3.57	Translation		Ribosomal protein S18 acetylase RimI and related acetyltransferases
SP_1699	-4.5371427	3.45	Lipid metabolism	<i>acpS</i>	4'-phosphopantetheinyl transferase

References

1. Kostakioti, M., M. Hadjifrangiskou, and S.J. Hultgren, *Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era*. Cold Spring Harb Perspect Med, 2013. **3**(4): p. a010306.
2. Sauer, K., et al., *The biofilm life cycle: expanding the conceptual model of biofilm formation*. Nat Rev Microbiol, 2022. **20**(10): p. 608-620.
3. Young, E., G. Melaugh, and R.J. Allen, *Active layer dynamics drives a transition to biofilm fingering*. NPJ Biofilms Microbiomes, 2023. **9**(1): p. 17.
4. Nolan, L.M., et al., *A global genomic approach uncovers novel components for twitching motility-mediated biofilm expansion in Pseudomonas aeruginosa*. Microb Genom, 2018. **4**(11).
5. Lee, C.K., et al., *Broadcasting of amplitude- and frequency-modulated c-di-GMP signals facilitates cooperative surface commitment in bacterial lineages*. Proc Natl Acad Sci U S A, 2022. **119**(4).
6. Stewart, P.S. and M.J. Franklin, *Physiological heterogeneity in biofilms*. Nat Rev Microbiol, 2008. **6**(3): p. 199-210.
7. Ogunniyi, A.D., et al., *Identification of genes that contribute to the pathogenesis of invasive pneumococcal disease by in vivo transcriptomic analysis*. Infect Immun, 2012. **80**(9): p. 3268-78.
8. Pettigrew, M.M., et al., *Dynamic changes in the Streptococcus pneumoniae transcriptome during transition from biofilm formation to invasive disease upon influenza A virus infection*. Infect Immun, 2014. **82**(11): p. 4607-19.
9. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features*. Bioinformatics, 2014. **30**(7): p. 923-30.
10. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
11. Moscoso, M., E. Garcia, and R. Lopez, *Biofilm formation by Streptococcus pneumoniae: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion*. J Bacteriol, 2006. **188**(22): p. 7785-95.
12. Sanchez, C.J., et al., *Streptococcus pneumoniae in biofilms are unable to cause invasive disease due to altered virulence determinant production*. PLoS One, 2011. **6**(12): p. e28738.
13. Lu, H., et al., *Metabolomics Deciphered Metabolic Reprogramming Required for Biofilm Formation*. Sci Rep, 2019. **9**(1): p. 13160.
14. Altay, O., et al., *Revealing the Metabolic Alterations during Biofilm Development of Burkholderia cenocepacia Based on Genome-Scale Metabolic Modeling*. Metabolites, 2021. **11**(4).
15. Cools, F., et al., *Streptococcus pneumoniae galU gene mutation has a direct effect on biofilm growth, adherence and phagocytosis in vitro and pathogenicity in vivo*. Pathog Dis, 2018. **76**(7).

16. Carvalho, S.M., et al., *CcpA ensures optimal metabolic fitness of Streptococcus pneumoniae*. PLoS One, 2011. **6**(10): p. e26707.
17. Manco, S., et al., *Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis*. Infect Immun, 2006. **74**(7): p. 4014-20.
18. Mathew, B.J., et al., *Role of Streptococcus pneumoniae extracellular glycosidases in immune evasion*. Front Cell Infect Microbiol, 2023. **13**: p. 1109449.
19. Jennings, M.P., C.J. Day, and J.M. Attack, *How bacteria utilize sialic acid during interactions with the host: snip, snatch, dispatch, match and attach*. Microbiology (Reading), 2022. **168**(3).
20. Blanchette, K.A., et al., *Neuraminidase A-Exposed Galactose Promotes Streptococcus pneumoniae Biofilm Formation during Colonization*. Infect Immun, 2016. **84**(10): p. 2922-32.
21. Bidossi, A., et al., *A functional genomics approach to establish the complement of carbohydrate transporters in Streptococcus pneumoniae*. PLoS One, 2012. **7**(3): p. e33320.
22. Magoch, M., et al., *Crystal Structure of Mannose Specific IIA Subunit of Phosphotransferase System from Streptococcus pneumoniae*. Molecules, 2020. **25**(20).
23. Ness, S. and M. Hilleringmann, *Streptococcus pneumoniae Type 1 Pilus - A Multifunctional Tool for Optimized Host Interaction*. Front Microbiol, 2021. **12**: p. 615924.
24. Hilleringmann, M., et al., *Molecular architecture of Streptococcus pneumoniae TIGR4 pili*. EMBO J, 2009. **28**(24): p. 3921-30.
25. Munoz-Elias, E.J., J. Marcano, and A. Camilli, *Isolation of Streptococcus pneumoniae biofilm mutants and their characterization during nasopharyngeal colonization*. Infect Immun, 2008. **76**(11): p. 5049-61.
26. Pancotto, L., et al., *Expression of the Streptococcus pneumoniae pilus-1 undergoes on and off switching during colonization in mice*. Sci Rep, 2013. **3**: p. 2040.
27. Madsen, J.S., et al., *The interconnection between biofilm formation and horizontal gene transfer*. FEMS Immunol Med Microbiol, 2012. **65**(2): p. 183-95.
28. Abe, K., N. Nomura, and S. Suzuki, *Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism*. FEMS Microbiol Ecol, 2020. **96**(5).
29. Salvadori, G., et al., *Competence in Streptococcus pneumoniae and Close Commensal Relatives: Mechanisms and Implications*. Front Cell Infect Microbiol, 2019. **9**: p. 94.
30. Abranches, J., et al., *Different roles of EILABMan and EIIGlc in regulation of energy metabolism, biofilm development, and competence in Streptococcus mutans*. J Bacteriol, 2006. **188**(11): p. 3748-56.
31. Steinmoen, H., Eivind Knutsen, and L.S. Håvarstein, *Induction of natural competence in Streptococcus pneumoniae triggers lysis and DNA release from a subfraction of the cell population*. PNAS, 2002. **99**(11): p. 6.
32. Santoro, F., F. Iannelli, and G. Pozzi, *Genomics and Genetics of Streptococcus pneumoniae*. Microbiology Spectrum, 2019. **7**(3).

33. Muschiol, S., et al., *Gram-Positive Type IV Pili and Competence*. Microbiol Spectr, 2019. **7**(1).
34. Kovacs, M., et al., *A functional dlt operon, encoding proteins required for incorporation of d-alanine in teichoic acids in gram-positive bacteria, confers resistance to cationic antimicrobial peptides in Streptococcus pneumoniae*. J Bacteriol, 2006. **188**(16): p. 5797-805.
35. Denapaite, D., et al., *Biosynthesis of teichoic acids in Streptococcus pneumoniae and closely related species: lessons from genomes*. Microb Drug Resist, 2012. **18**(3): p. 344-58.
36. Majchrzykiewicz, J.A., O.P. Kuipers, and J.J. Bijlsma, *Generic and specific adaptive responses of Streptococcus pneumoniae to challenge with three distinct antimicrobial peptides, bacitracin, LL-37, and nisin*. Antimicrob Agents Chemother, 2010. **54**(1): p. 440-51.
37. Du, J., et al., *Dlt operon regulates physiological function and cariogenic virulence in Streptococcus mutans*. Future Microbiology, 2023. **18**(4).
38. Spengler, C., et al., *Using Knock-Out Mutants to Investigate the Adhesion of Staphylococcus aureus to Abiotic Surfaces*. Int J Mol Sci, 2021. **22**(21).
39. Bai, Y., et al., *Cyclic di-AMP impairs potassium uptake mediated by a cyclic di-AMP binding protein in Streptococcus pneumoniae*. J Bacteriol, 2014. **196**(3): p. 614-23.
40. Sinha, D., et al., *Pivotal Roles for Ribonucleases in Streptococcus pneumoniae Pathogenesis*. mBio, 2021. **12**(5): p. e0238521.
41. Shishkin, A.A., et al., *Simultaneous generation of many RNA-seq libraries in a single reaction*. Nature Methods, 2015. **12**(4).
42. Kraus, A.J., B.G. Brink, and T.N. Siegel, *Efficient and specific oligo-based depletion of rRNA*. Sci Rep, 2019. **9**(1): p. 12281.
43. Surujon, D. and T. van Opijnen, *ShinyOmics: collaborative exploration of omics-data*. BMC Bioinformatics, 2020. **21**(1): p. 22.