ABSTRACT

Title of dissertation: THE ROLE OF GENOME ORGANIZATION AND FILAMENTOUS BACTERIOPHAGE ON GONOCOCCAL BIOLOGY AND PATHOGENICITY

Jessica Sarah Kopew, Doctor of Philosophy, 2024

Dissertation directed by: Professor Daniel C. Stein
Department of Cellular Biology and Molecular Genetics

Gonorrhea, caused by the bacterium *Neisseria gonorrhoeae* (GC), represents a significant global health concern as it is the second most common bacterial STI and has a rising rate of antimicrobial resistance. The first study of this thesis aims to elucidate the causes and consequences of gonococcal genome reorganization. Here I found that when looking at many GC strains' genomes, each GC strain had a unique genome organization including both inversions and translocation events. I also saw a consistent pattern of DNA high sequence similarity on both sides of the translocation or inversions, consistent with homologous recombination driven reorganization. PCR analysis of inversion events suggests that these large-scale reorganization events are both stable and rare. Growth curve analysis demonstrates a wide variability in growth rate between strains. Proteomic analysis suggests reorganization driven changes to replication termination location leads to upregulation of many kinds of proteins including energy metabolism and antimicrobial resistance associated genes. This study suggests that homologous recombination driven genome reorganization can have large impacts on gonococcal biology and pathogenicity. This study demonstrates the need for future gonococcal studies to use multiple GC strains from a diverse background to capture the wide variability in GC phenotypes. The second study of this thesis sought to uncover the role filamentous bacteriophage play in GC biology. I found that every GC strain currently in the NCBI database at the date of this study contains four filamentous bacteriophage

gene regions in the GC genome. I found that FA1090ΔfiI (a GC strain lacking all four filamentous bacteriophage gene regions) grew poorly at 37°C both in broth and on agar, as compared to wild type FA1090. However, there was no difference when the strains were grown at 34°C or when grown without shaking, demonstrating the condition dependent nature of this growth advantage. FA1090ΔfiI formed larger bacterial aggregates than FA1090 WT. When these strains were analyzed for their ability to produce biofilms, no differences were seen in the overall biofilm's biomass, yet the overall structure of the biofilms were different, with FA1090ΔfiI producing taller and rougher biofilms. Previous unpublished research in the Stein Lab demonstrates that filamentous phage derived proteins are capable of deteriorating the integrity of epithelial cell cultures and cervical tissue explants. The data from this chapter suggests that filamentous phage provide the gonococcus with a growth advantage, inhibit bacterial aggregation, alter the structure of the GC biofilm, and that phage proteins can lead to loss of the integrity of the epithelium. Taken en toto, these studies demonstrate that both alterations in bacterial genome organization and contributions from filamentous bacteriophage genomes can impact gonococcal biology and pathogenicity, which could be key to preventing and treating GC infections.

THE ROLE OF GENOME ORGANIZATION AND FILAMENTOUS BACTERIOPHAGE ON GONOCOCCAL BIOLOGY AND PATHOGENICITY

By

Jessica Sarah Kopew

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Daniel Stein, Ph.D., Chair
Assistant Professor Brantley Hall, Ph.D.
Associate Professor Jason Kahn, Ph.D.
Professor Kevin McIver, Ph.D.
Professor Wenxia Song, Ph.D.

© Copyright by
Jessica Sarah Kopew
2024

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Daniel C. Stein. My entire thesis and graduate career would not have been possible without his unwavering support, guidance, and criticism. His mentorship over these past five years has been invaluable in shaping me into the scientist I am today and for that I thank him. Additionally, I would like to thank Dr. Wenxia Song for being with me every step of the way through this journey. Her support during this process made me feel like an unofficial Song Lab member, which I deeply appreciate. I would also like to thank the rest of my dissertation committee members: Dr. Kevin McIver, Dr. Brantley Hall, and Dr. Jason Kahn. I appreciate their time, effort, and dedication to helping me shape this project into what it is today.

I would like to thank the past and present members of the Stein and Song lab for being with me during my time at The University of Maryland. I want to especially thank Dr. John (Chris) Vitucci, Salsawi Kebede, Dr. Sofia Di Benigno, Dr. Yiwai Dai, Dr. Margaret (Katie) Seeley-Fallen, Dr. Jurriaan van Haaren, and Dr. Anshuman Bhanja. Your listening ears, both inside and outside lab, were greatly appreciated and made this journey more enjoyable.

To my friends and colleagues in the Biological Sciences program at The University of Maryland, you have my deepest gratitude for your encouragement and companionship throughout our time here. I would like to specifically give my deepest thanks to my friends Dr. Maria Colt and Melissa Perry for their friendship, guidance, support, and inspiration. This process would not have been possible without you two keeping me sane with many fun nights together, and I am so glad I had you in my corner from day one.

To my friends and family, you have my deepest gratitude for sustaining me throughout my time here at Maryland. I'd like to thank my parents, Regina and Steve, for their support, encouragement, and love during my entire life. I'd also like to thank my siblings Frances and Aaron, as well as my brother-in-law Adam Ciesielski, for their help and laughter along the way. My family has supported me throughout my entire life, encouraging me to pursue my dreams and sustaining me throughout my journey; for that they will have my gratitude forever. To my friends back home, Emma, Becca, and Supriya, thank you for being my cheerleaders from afar. To all my friends I've made here in Maryland and DC, thank you for the countless Shabbat meals, kickball games, movie nights, rock climbing, game nights, birthday parties, rant

sessions, and so much more. And to the many feline friends who have been there along the way (Lily, Jasmine, Callie, Tofu, and Beans), I thank you for keeping sane.

You have all made my time here so rich, and this accomplishment would not have been possible without each and every one of you.

Table of Contents

Acknowledgements..... ii

Table of Contents…….. iv

List of Tables…….. vii

List of Figures…… viii

List of Abbreviations…….. x

Chapter 1: Introduction……. 1

1.1 Gonorrhea as a disease…….. 1

1.2 Gonorrhea infection progression within the female reproductive tract…… 1

1.3 Gonococcal genome organization.…….. 3

1.4 Filamentous bacteriophage within Neisseria gonorrhoeae..…..… 11

1.5 Significance and Goals…… 16

1.5.1 Aim 1: The role of genome organization in GC biology.…….……..... 16

1.5.2 Aim 2: The role of filamentous bacteriophage on GC biology.…… 16

Chapter 2: The role of Genome Organization in GC Biology……. 19

2.1 Introduction…… 19

2.2 Materials and Methods…… 21

2.2.1 Whole genome alignment and sequence comparison……. 21

2.2.2 Growth analysis of N. gonorrhoeae strains…… 21

2.2.3 PCR verification of small inversion event……... 21

2.3 Results…….. 22

2.3.1 Genomic pattern variation among reference strain's genomes…….. 22

2.3.2 Homologous recombination is the key driver of genome reorganization…….. 22

2.3.3 PCR verification of reorganization events existence and stability…….. 25

2.3.4 Growth differences between WHO gonococcal strains…… 29

2.3.6 Changes in gene expression of key genes not due to gene mutations……. 31

2.3.6 Changes in gene expression of key genes not due to gene mutations……. 32

2.4 Discussion……. 37

Chapter 3: The role of Filamentous Bacteriophage in GC Biology……... 42

3.1 Introduction…….. 42

3.2 Materials and Methods……... 44

3.2.1 Growth of N. gonorrhoeae strains…… 44

3.2.2 Creation of Filamentous Bacteriophage Mutants. …… 44

3.2.3 Biofilm Analysis of GC…….. 45

3.2.4 Aggregation analysis of GC…… 45

3.2.5 DNA sequencing and verification of GC……... 46

3.2.6 Statistical Analysis……. 46

3.3 Results…….. 46

3.3.1 Genetic Composition of FA1090 Wild Type and FA1090Δfil GC…….. 46

3.3.2 Filamentous Bacteriophage are Required for Growth…….. 47

3.3.3 All Four Regions of Filamentous Bacteriophage Genes not Required for Rescue of Growth……. 48

3.3.4 Filamentous Bacteriophage Promote Smoother and Shorter Biofilms…….. 55

3.3.5 Filamentous Bacteriophage inhibit gonococcal aggregation……... 55

3.4 Discussion…… 61

Chapter 4: General discussion and future work……... 69

4.1 General summary…… 69

4.2 Working Model…….. 70

4.3 Future Direction……... 72

4.3.1 Future directions of genome reorganization study…… 72

4.3.2 Future directions of filamentous bacteriophage study…….. 73

Appendices…… 76

Bibliography…….. 84

List of Tables

Table 1 Gonococcal Genomic Mobile Elements

Table 2 Relative location of various genes involved in antibiotic efflux

Table 3 List of PCR primers for filamentous phage detection

Table 4 List of PCR primers for filamentous phage creation

Table 5 Strain names and corresponding accession numbers

List of Figures

Figure 1 *Neisseria gonorrhoeae* infection progression model

Figure 2 DNA and RNA Polymerases' response to gene orientation

Figure 3 Homologous recombination of circular genome results in an inversion event

Figure 4 *Neisseria gonorrhoeae* filamentous bacteriophage and genome protein organization

Figure 5 Purified filamentous phage protein ORF9 impacts T84 cell monolayer and cervical explant integrity.

Figure 6 Different genome patterns in different strains of GC

Figure 7 Mauve alignment of various *N. gonorrhoeae*

Figure 8 Homologous DNA sequence on both sides of large inversion

Figure 9 Homologous DNA sequence on both sides of small inversion

Figure 10 Confirmation of inversion event in database sequences via PCR detection

Figure 11 Selection pressure against selecting inversions that move replication terminus site

Figure 12 Possible inversions to convert FA1090 like genome organization to WHO L organization.

Figure 13 Variation in growth in both broth and agar

Figure 14 DNA replication machinery and genes near origin are over expressed in WHO strain with shifted terminal sequence

Figure 15 Multiple bacterial protein types upregulated in WHO L

Figure 16 Filamentous prophage regions in FA1090 genome

Figure 17 Construction of filamentous prophage region deletions in FA1090Δfil genome

Figure 18 Sequence analysis of FA1090 WT and FA1090Δfil

Figure 19 Filamentous phage required for growth at 37^{o}C

Figure 20 Growth defect in FA1090ΔfiI mutants rescued in static growth conditions

Figure 21 All four regions of filamentous phage regions are not required for WT levels of growth

Figure 22 Representative images of biofilm assays at 37^{o}C at 6 and 24 hours

Figure 23 Filamentous phage promote shorter and smoother biofilms at 37^{o}C

Figure 24 Filamentous phage promote shorter biofilms at 34^{o}C

Figure 25 Filamentous phage inhibit bacterial aggregation in piliated GC

Figure 26 Filamentous phage inhibit bacterial aggregation in non-piliated GC

Figure 27 Hypothesized model of filamentous phage's role in GC infection progression

Figure S1 Mauve Alignment of all GC strains analyzed

List of Abbreviations

AMR	Antimicrobial resistance
CEACAM	Carcinoembryonic antigen-related cell adhesion molecules
CREE	Corriea repeat enclosed element
GC	*Neisseria gonorrhoeae*
DGI	Disseminated gonococcal disease
DNA	Deoxyribonucleic acid
ECC	Ectocervix
ENC	Endocervix
FRT	Female reproductive tract
GC	*Neisseria gonorrhoeae*
HIV	Human immunodeficiency virus
HSPG	Heparan sulfate proteoglycans
LOS	Lipo-oligosaccharide
Opa	Opacity proteins
PCR	Polymerase Chain Reaction
Phage	Bacteriophage
PID	Pelvic inflammatory disease
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid

SNP	Single nucleotide polymorphisms
SSREE	Spencer-Smith repeat enclosed element
tRNA	Transfer ribonucleic acid
TZ	Transformation zone
WHO	World Health Organization
WT	Wild type
Zot	Zonula occludens toxin

Chapter 1: Introduction

1.1 Gonorrhea as a disease

Gonorrhea, caused by the bacterium *Neisseria gonorrhoeae* (GC), represents a significant global health concern, with an estimated 78 million individuals worldwide and over 600,000 domestically infected annually.[1,2] Increasing GC infection rates can be attributed to various factors, such as decreased condom usage, increased urbanization and international travel, false negative tests, and failing treatments.[1,3] With rising rates of GC infections, the World Health Organization (WHO) has become increasingly concerned by the increased prevalence of antimicrobial resistance (AMR) within GC populations, leading to increased efforts to survey and combat this public health threat.[4]

Within an individual, GC infection can lead to a wide range of symptoms.[1,5] While most female infections are asymptomatic, some women will develop more dangerous sequelae, which can range from localized infection within the female reproductive tract (FRT) such as pelvic inflammatory disease (PID) to disseminated gonococcal disease (DGI).[1] Both of these GC related diseases can have severe consequences on the individual, with prolonged lack of treatment potentially leading to infertility, arthritis, endocarditis, and even death.[1] While men have higher rates of symptomatic infections than women, men tend to have less dangerous symptoms, such as pain upon urination or penile discharge.[6] GC infection can also increase susceptibility to other infections, such as human immunodeficiency virus (HIV).[3] Given the multifaceted implications of GC infection on both the individual and global scale, a comprehensive understanding of its pathogenic mechanisms is paramount for effective disease treatment and prevention strategies.

1.2 Gonorrhea infection progression within the female reproductive tract

In order for a woman to acquire a gonococcal infection, GC must undergo an infection progression where GC first adheres to the FRT and then colonizes the ectocervix (ECC), leading to an asymptomatic localized infection, or penetrates the transformation zone (TZ) and endocervix (ENC) to disseminate into deeper tissue (Figure 1).[7-9] GC is a gram-negative diplococcus whose key surface components during infection are opacity (Opa) proteins, lipooligosaccharides (LOS), and type IV pili,

which are all phase- and antigenically variable.[10-13] This high level of phase and antigenic variation in the GC surface proteins allows for high levels of immune evasion.[14] First, type IV pili contact host cells and retract to bring GC closer to the host surface, where LOS and Opa proteins facilitate more intimate interactions between GC and host cells.[10-15] Pili are absolutely required for GC adherence to the host and allow for natural transformation.[11] Pili are not only important for initial adherence to the host epithelium, but also play a key role in promoting bacterial aggregation to form microcolonies and, eventually, biofilms in the mucosal layer above the host epithelium.[14] While there are 11 different Opa proteins that can be all expressed simultaneously, none expressed, or some combination of any of the Opa proteins, different Opa proteins can lead to different infection outcomes.[11] Certain Opa proteins can bind to either Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) or Heparan sulfate proteoglycans (HSPGs) on the host cells. Current research suggests that Opa_{CEACAM}-expressing GC initiate localized asymptomatic ECC colonization, whereas Opa_{HSPG}-expressing GC leads to symptomatic penetration of the ENC.[11]

In the context of a female infection, the FRT can be split into two halves: the upper and lower FRT. The cervix serves as a gateway between these halves, with the unsterile lower FRT below it and the sterile upper FRT above. To prevent GC from penetrating deeper into tissue, the first layer of defense is tightly bound together epithelial cells, providing a physical barrier between the mucosal layer on its apical side and deeper tissue on its basal side. The cervix has many cell types, consisting of the multilayered squamous epithelial cells of the ECC, which faces down towards the vagina; the single layered polarized epithelial cells of the ENC, which faces up towards the uterus; and the TZ between them, which has a morphology between the ENC and ECC (Figure 1).[16] Tight junctions at the apical side of the polarized cells and at the parabasal level in nonpolarized cells help keep pathogens from passing between cells.[16,17] Calcium ions (Ca^{2+}) play a crucial role in keeping tight junctions regulated, as extracellular Ca^{2+} is necessary for tight junction formation.[18] It has been seen that GC can overcome this epithelial barrier in the ENC and TZ by causing Ca^{2+} influx and activation of non-muscle myosin II (NMII) leading to dissociation of junctions between epithelial cells and epithelial cell shedding, allowing for transmigration of GC past the now unstable epithelial cell barrier.[7-9] It has also been seen that other filamentous phage / phage proteins are involved in the dissociation of tight junctions in other pathogens.[19-21] Conversely,

increased integrin production in the squamous epithelial cells of the ECC decreases host cell shedding, allowing for high levels of GC colonization in the ECC.[8] While it is known that GC can form biofilms on abiotic surfaces and on cells in vitro, the role of biofilms in the context of ECC colonization is still not well understood.[14,22] Biofilm production in the FRT could lead to immune evasion, increased fitness against environmental stressors, and persistence post-antibiotic treatment, all leading to long-term consequences for the host.[22] Studies in *Neisseria meningitidis* have shown that *Neisseria* biofilms can have differential gene expression, which can aid biofilm defenses on both the surface of the biofilm and the bottom of the biofilm contacting the surface of the host cells.[23,24]

1.3 Gonococcal genome organization

Most bacterial genomes are composed of a circular chromosome and plasmids. When bacteria divide, the genome is bidirectionally replicated by deoxyribonucleic acid (DNA) polymerases starting at the origin of replication and proceeding in opposite directions from each other until they meet at the terminal sequence resulting in two complete copies of the genome.[25] These origin and terminal sequences are highly conserved within bacterial species. Variation in the speed of the DNA polymerase and genome size leads to diverse growth rates among bacterial species as each round of replication must produce a complete copy of the genome. This method of replication is highly conserved between bacterial species and allows for nearly perfect inheritance of the genomes from one generation to the next. However, selective pressure, imperfect genome replication, horizontal gene transfer, and mobile genomic elements creates a genome that is not perfectly stable.[26-30] This level of flexibility within the genome allows for bacteria to adapt to stressful and new environments and obtain new genes that can allow them to acquire higher fitness, which, in the context of pathogenicity, can result in higher AMR or immune evasion. When comparing different bacterial species, there is a wide range in the variation of genome stability.[25-28,30,31] Cyanobacteria and other free-living bacterial species tend to have lower genomic stability than pathogenic bacterial species, while endomutualist bacteria tend to have the most stable genomes.[30]

Overall genome organization has a large impact on bacterial biology and conserved genome organizational patterns can be observed across species. During replication, transcription and translation of proteins occur concurrently. This means that as DNA polymerases are replicating the genome, they will

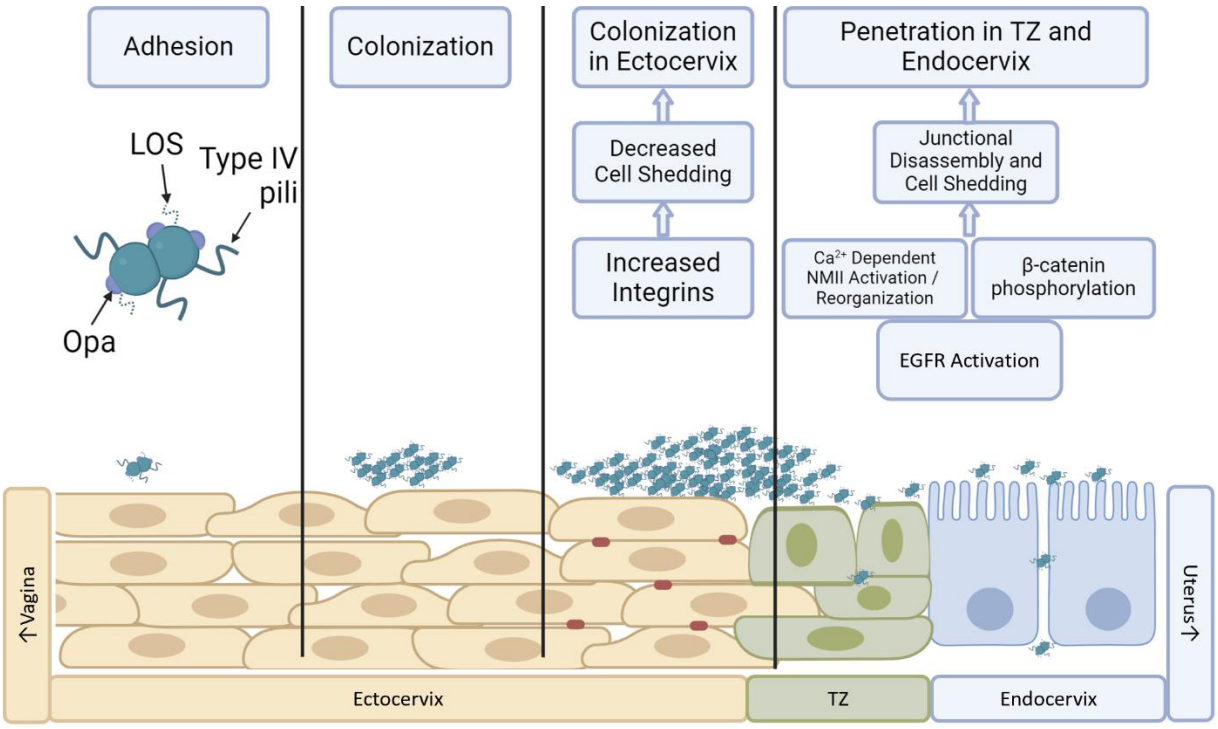


Figure 1. *Neisseria gonorrhoeae* infection progression model. Yellow cells on the bottom left represent multilayered squamous epithelial cells of the ECC which faces down towards the vagina. Blue cells on the bottom right represent the single layer, polarized epithelial cells of the ENC which faces up towards the uterus. Green cells in the bottom center represent the TZ. An enlarged GC diplococci is shown with important infection surface components labeled. Image is not to scale and was created in BioRender. [7-13,15,16]

eventually collide with ribonucleic acid (RNA) polymerases. DNA polymerases move at a much faster rate than RNA polymerases, however, collisions resulting in a slowed replication rate or complete replication arrest are more likely if the gene orientation is facing the opposite direction of the replication fork (Figure 2).[25,26,30,31] This has led to evolutionary pressures driving highly expressed genes to be oriented in the same direction as replication.[25,31] When favorable conditions trigger replication to begin, these same conditions might not be present by the time replication ends. This change in conditions has resulted in higher rates of gene transfer, recombination, and mutations near the terminus of replication compared to the origin.[25,31] As the genome replicates, a second round of replication may begin before the first has finished. In *Escherichia coli*, it takes about one hour for a full genome replication to occur, however, it has a rate of 2.5 doublings per hour. This results in a higher gene dosage of genes closer to the origin and therefore a tendency toward higher expression of these genes.[25,31] When genome organization changes, this can have large- scale impacts on gene expression, growth rates, and overall fitness of the bacteria. While most studies continue to focus on genetic changes to explain phenotypic differences between bacterial species and strains, changes in genome organization without altering overall genetic content can have profound implications for bacterial gene expression and, therefore, bacterial phenotypes.

The genome can change in many ways. At its smallest scale, sequences can change by a single nucleotide. This can happen randomly through improper incorporation of nucleotides during replication of the genome and will be inherited by the next generation.[32] These single nucleotide polymorphisms (SNPs) if found in a gene or gene expression regulatory area can result in changes to bacterial phenotypes, which can be inherited by subsequent generations if it is not lethal. Horizontal gene transfer can change larger areas of DNA sequence through the acquisition of whole genes, gene regions or entire plasmids from another bacterium.[33] In the case of genes and gene regions, these newly acquired sequences can be inserted into the genomes through mobile elements or homologous recombination (Table 1). Acquisition of genes can result in phenotypic changes that can increase fitness and/or pathogenicity in these bacteria, such as genes associated with higher AMR or immune evasion. Some genes also change more rapidly through antigenic and phase variation. These types of changes often result from environmental changes or stress and involve rearrangements of a bacteria’s genomic DNA rather than DNA from another bacterium.[34] Both antigenic and phase variation are heritable changes to

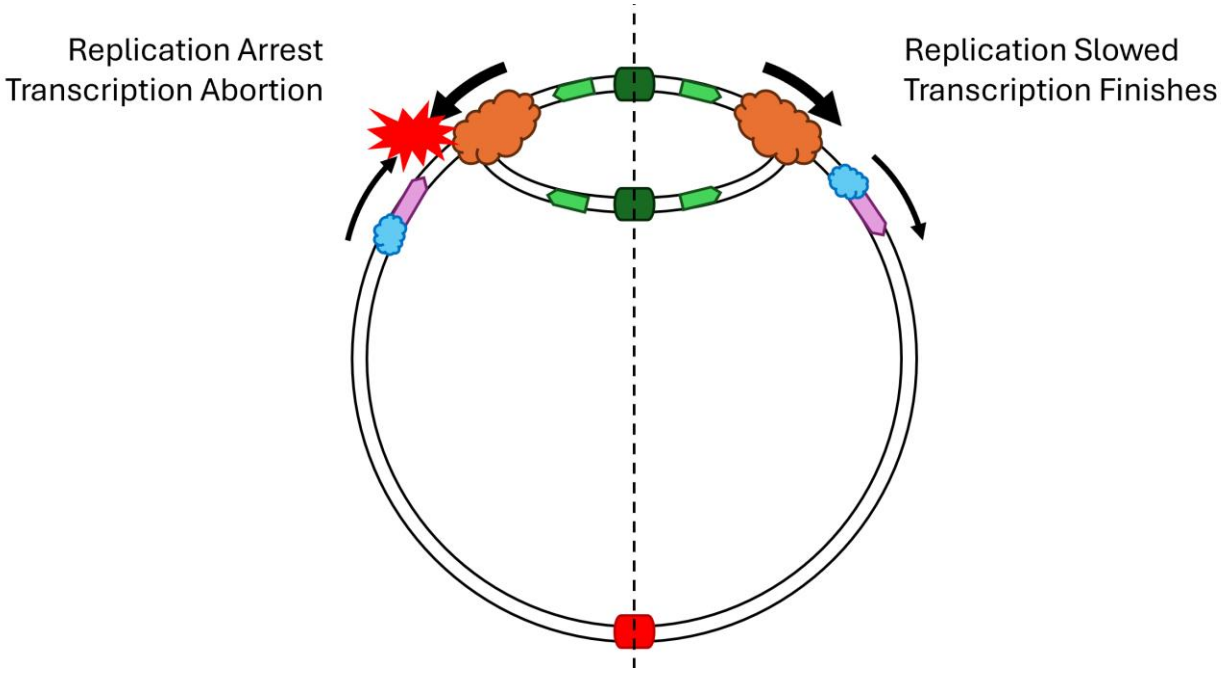


Figure 2. DNA and RNA Polymerases' response to gene orientation. The left half of the figure represents the consequences of RNA and DNA polymerase collision when genes are on the lagging strand of replication. The right half of the figure represents the consequences of RNA and DNA polymerases near each other when genes are located on the leading strand of replication. Genes closer to the origin (green arrows) have a higher average gene dosage than genes further from the origin (pink arrows) during active bacterial replication. Orange shapes represent DNA polymerases, blue shapes represent RNA polymerases, dark green boxes represent the origin of replication, the red box represents the terminus of replication, and black arrows represent the direction of polymerase movement with the thickness of the arrows representing relative speed. [25,26,30,31]

Table 1. Gonococcal Genomic Mobile Elements[35]

Mobile Element	**Components / Sequence Organization**
Correia repeat enclosed element (CREE)	Variable sequence of DNA flanked by a conserved inverted repeat, which is a hotspot for reorganization events. Many are found throughout the genome.
Spence-Smith Repeat Enclosed Element (SSREE)	Conserved sequence of DNA flanked by inverted repeat found two to three times per genome.
Homologous Recombination	Highly similar regions of DNA forming a Holliday junction during replication, leading to an inversion event when these two regions are from the same piece of circular DNA (Figure 3). This can also lead to loss or gain of genes when from a transformed piece of DNA.
Transposon	DNA regions containing transposition genes that are flanked by a recognized inverted repeat where DNA can be moved from one area or plasmid to another. They can contain cargo genes between the inverted repeats that move with the transposons.
Insertion Sequence	Conserved sequences in gonococcal genomes where DNA is more likely to be inserted into the genome. Contains transposition genes flanked by an inverted repeat but does not contain cargo genes. IS1016 is the most abundant GC insertion sequence.
Bacteriophage / prophage	Bacteriophage and prophage often contain transposases as part of their replication machinery. They can insert / remove their viral phage genomes into / out of the bacterial GC genome using their own transposases.

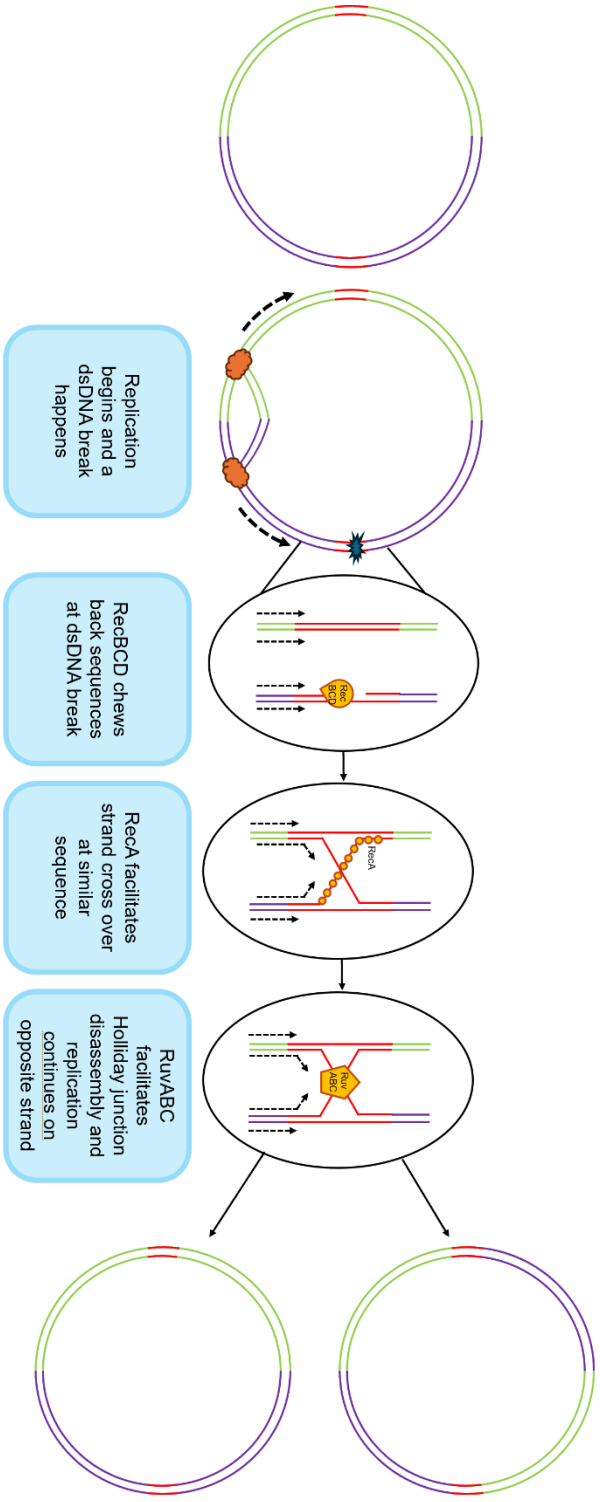


Figure 3. Homologous recombination of circular genome results in an inversion event. The circular genome is represented as circles, with each half of replichore represented by either green or purple and two areas of identical sequences highlighted in red. Black dashed arrows represent direction of DNA replication of the bacterial genome. DNA polymerase is represented as blue shapes and homologous recombination machinery is represented in yellow. Resulting chromosomes after homologous recombination include one chromosome identical to the original and one with an inversion between indicated red sequences.

DNA sequence; however, phase variation is reversible and changes whether a gene is “on” or “off”, whereas antigenic variation changes which variant of the protein is expressed. Examples of phase and antigenic variable genes in GC are pili and Opa, two surface proteins involved with gonococcal pathogenicity. These types of genetic changes can result in increased pathogenicity through increased immune evasion and enhanced infection potential by altering gene content or sequence. Zooming out to a larger scale, genome reorganization events can also have a large impact on bacterial phenotypes.

Changes in genome organization, as opposed to changes in gene content, can occur in many ways. Many drivers of genome reorganization are driven by what are called “mobile elements” or areas that are hotspots for DNA movement (inversions, translocations, insertions, or deletions).[29,35] Inversions are areas of the genome that have changed directions but have not changed locations relative to the rest of the genome.[29,35] Translocations, however, are areas of DNA that have changed locations relative to the rest of the genome and may or may not also change direction.[29,35] Translocations and inversions can therefore impact orientation of the genes within the inverted/translocated area of the genome, distance of the genes to the terminal/origin sequence, and expression of genes close to the site of inversion/translocation. However, both inversions and translocations have no impact on the overall gene content of the genome. Inversions and translocations can occur on both large and small scales, from hundreds of bases to hundreds of thousands.[29,35] Within GC, there are many types of mobile elements that can drive these reorganization events (Table 1). Many of these mobile elements are either genes or sequences flanked by an inverted repeat (CREE, SSREE, transposases, insertion sequences).[35] Other types of genetic elements that can drive rearrangements are regions of DNA containing bacteriophage/prophage as they often carry their own transposases in order to integrate into or escape out of the bacterial genome.[35] Homologous recombination can also drive rearrangements by interacting with highly similar sequences within the same genome (Figure 3) or causing insertions or deletions when interacting with highly similar sequences from other sources of DNA.[35] These hotspots of DNA movement are thought to be drivers of genome reorganization events found within gonococcal genomes. When comparing three well characterized strains of GC (NG08107, FA1090, and NCCP11945), large (>10 kB) inversions and translocations were observed.[36] These types of mobile elements were often found at the sites of these inversion and translocation events.[36] However, none of these mobile elements seem to be

the sole driver of these large-scale rearrangement events, as none were found at every site of reorganization and with many found in places of the genome with high levels of stability. To date, most studies investigate gene content, gene mutations, or regulatory sequence mutations as the driver of variability in phenotypes. However, these large-scale changes to genome organization could be driving changes in GC phenotypes and therefore infection potential.

There are many processes that involve DNA movement in GC, but one of the most notable examples of DNA movement leading to phenotypic changes involved in invasiveness is GC's array of antigenically and phase-variable genes, including pilin, LOS, and *opa*.[5,10,11,14,35,37] One example of DNA movement involved in antigenic variation is the insertion of *pilS* gene segments into the *pilE* gene region, which results in structural changes to the pilin protein.[38] This process involves homologous recombination of the "silent" *pilS* gene into the *pilE* region, where *pilS* gene segments replacing *pilE* gene segments.[38] This homologous recombination of pilin genes, as well as other antigenic variation, leads to increased immune evasion by GC due to the constantly changing structure of these surface molecules. An example of DNA movement causing phase variation, or toggling a gene's expression between "on" and "off", is the variation in the number of "CTCTT" repeats in the *opa* genes, which will shifts the gene into or out of frame.[39] There are multiple ways the number of CTCTT repeats can change, including uneven recombination with other *opa* genes within the genome or slipped-strand mispairing during replication.[39] These examples illustrate how GC can use DNA movement to increase invasiveness by altering surface protein expression and structure.

One of the most crucial phenotypes that plays a role in infection potential and risk to human health, which has yet to be studied in the context of genomic reorganization, is the ability of GC to become resistant to the antibiotics used to treat gonococcal infection. The WHO sought to study AMR from currently circulating strains of GC and found wide variability in AMR capabilities.[4,40,41] However, the large variations in AMR capabilities could not be explained solely based on gene content or mutations. While some higher levels of AMR can be explained by SNPs in genes important for resistance to various antibiotics, there is still a great deal of unexplained resistance levels when comparing all the WHO strains.[4,40,41] With a large number of new strains from this study having been sequenced and AMR

capabilities measured from geographically and temporally distinct strains, investigating differences in currently circulating GC genome organization could uncover links between genome organization and AMR capabilities. What the main driver of large-scale reorganization events in GC is and whether variations in phenotypes could be explained by changes in genome organization rather than gene sequence is still unknown.

1.4 Filamentous bacteriophage within *Neisseria gonorrhoeae*

An additional aspect of GC infection that is not completely understood is the role that filamentous bacteriophage play in GC biology. Simply put, bacteriophage (phage) are viruses that infect bacteria.[42,43] Prophage DNA (phage genomes inserted into the bacterial genome) have been seen to make up to 20% of the bacterial genome and can affect phenotypes of the bacteria in which they reside, including pathogenicity and general fitness.[43-49] Phage can impact bacterial phenotypes as their genomes are inserted into their bacterial hosts and may carry pathogenic genes that these hosts can exploit to their advantage. Phage are thought to be the most abundant type of organism on the planet and can be found in every genus of bacteria, in every niche and environment across the globe.[43,46,49-51] Despite their prevalence in bacteria, phage were only discovered in the early 20th century, with filamentous phage only identified in the 1960s.[43,46,49-51] While phage are ubiquitous to prokaryotic life, there is still much to be learned about these viruses.

Within GC, prophage DNA makes up about 7% of the genome.[52-56] All GC strains contain four filamentous prophage regions known as NgoΦ6-9 (or PHI6-9) (Figure 4).[52-56] These phage have many homologous open reading frames (ORFs) in their sequences, not only to each other but to other species such as *N. meningitidis* and *Vibrio cholerae*, with some of these homologous genes conferring heightened pathogenicity to their bacterial hosts.[21,23,53,57] These filamentous phage within GC belong to the *Inoviridae* family and the *Inovirus* genus. While not all phage can, this type of phage is able to integrate its genomic DNA into the bacterial genome.[47,51] *Inoviridae* has three types of genes within its genome: replication genes, structural genes, and assembly and release genes.[43,46,48,50] Replication genes are required for replicating the phage genomes, structural genes encode proteins that make up the filamentous phage itself, while assembly and release genes encode proteins that construct the phage at the surface of the

bacteria, including the extrusion pore through which the phage leave the cell.[43,46,48,50] Filamentous bacteriophage are semi-flexible worm-like nucleoprotein filaments. They are a hollow helix of the coat proteins with two different sets of proteins on either end, which surround two antiparallel strands of single stranded DNA that contain a double stranded DNA packaging sequence.[47,58] The proteins at either end help phage integrate into the membrane and allow the phage to assemble at the membrane, package DNA into the coat proteins, extrude phage out of cells, and then infect another bacterium without having to lyse the host cell (Figure 4).[19] Filamentous phage leaving the host via an extrusion pore rather than cell lysis is a key feature of filamentous phage, as they can form “chronic infections” of its host, constantly making and releasing phage into the environment, without killing the host bacteria.

Each of the four filamentous phage within the gonococcal genome has a unique set of ORFs in different orientations, but with many overlapping ORFs that have nearly identical sequences to each other, totaling 13 different ORFs (Figure 4).[23] There is also a region of 136 base pairs at the end of each of phage after ORF11. However, in PHIΦ6, the ORF11 and homologous region are inverted (Figure 4).[23,52] These ORFs also contain homologous regions to known bacterial regulators found outside of the phage regions.[53] ORF1 is needed for phage replication and shows high levels of homology to RstA in CTXΦ, the filamentous phage found in *V. cholerae*. ORFs 2 and 3’s functions are unknown, but may be involved in assembly of the phage. ORFs 4 and 5 are thought to be the main coat proteins of the phage. ORF6 has a zinc finger motif, which could indicate involvement in transcriptional control or replication. ORF7 has high homology to the TspB protein found in GC and *N. meningitidis*, a known virulence factor, which is located on the surface of the bacteria. ORF9 is needed for assembly and secretion of the phage, is found on the surface of GC, and has high homology to the enterotoxin Zot found in CTXΦ. ORF12’s function is still unknown, however they contain transmembrane regions.[52-54,56] While the functions of most of these ORFs in the filamentous phage lifecycle is known or can be hypothesized from homology to other species, the role they play in GC pathogenicity remains to be investigated.

The role of phage, including filamentous phage, in pathogenicity has been studied extensively in many different bacterial species, including those capable of infecting humans. It has been observed that *Salmonella paratyphi* phage increases invasion potential, with *S. paratyphi* containing filamentous phage

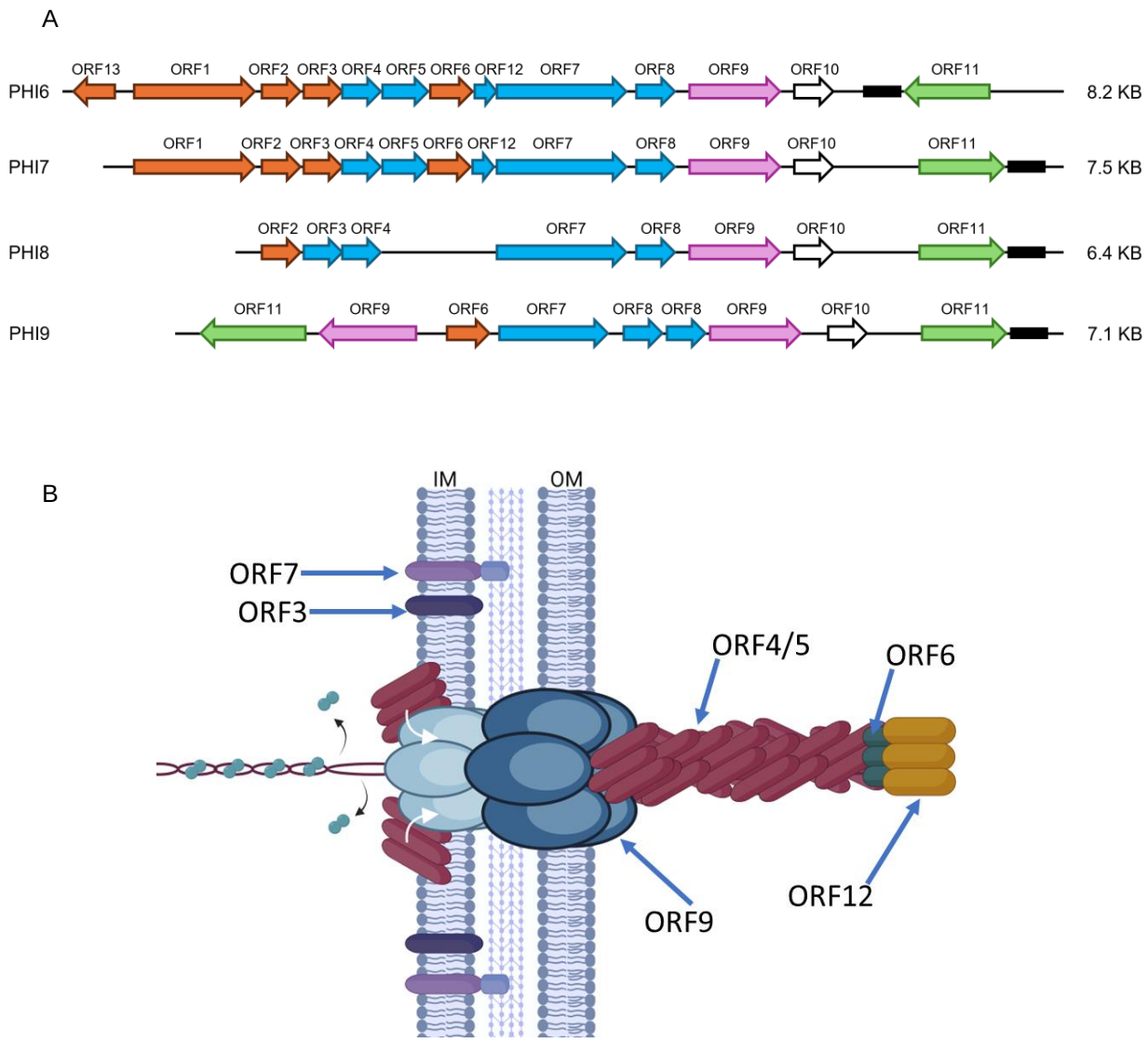


Figure 4. *Neisseria gonorrhoeae* filamentous bacteriophage genome and protein organization. (A) Genome organization of FA1090 PHI6-9 as adapted from Piekarowicz, A. et al.[52,53] Blue arrows represent structural genes. Orange arrows represent replication genes. Pink arrows represent secretion genes. Green arrows represent integrase genes (B) Schematic of filamentous bacteriophage being assembled at the bacterial membrane (gram-negative) with labeled proteins from (A). Image is not to scale and was created in BioRender.

causing more severe infection complications such as multi-organ damage at lower doses compared to filamentous phage-free bacteria.[59] *Pseudomonas aeruginosa* and *Bacillus subtilis* activate filamentous phage expression during biofilm development.[45,59,60] These filamentous bacteriophage containing biofilms exhibit highly organized biofilms with less water loss, higher shear stress resistance, and higher AMR capabilities compared to biofilms lacking these phage.[45,60-63] *N. meningitidis*, a bacterium in the same genus as GC, also has filamentous phage of high homology known as Meningococcal Disease-Associated phage (MDAΦ). MDAΦ has been highly associated with invasive meningococcal disease-causing bacterial strains and has been shown to increase biomass of *N. meningitidis* biofilms during colonization of human cells.[23,64] *V. cholerae* and *Campylobacter concisus* contain genes that encode endotoxins involved in the disorganization of the human gut epithelium, leading to penetration of bacteria into their hosts.[19-21,65,66] Adding the *V. cholerae* filamentous phage, CTXΦ, into non-pathogenic strains has been shown to increase the virulence of these strains.[67] The filamentous phage ΦRSS1, found in the phytopathogen *Ralstonia solanacearum*, enhances its host's virulence through changes in motility and extracellular polysaccharide synthesis, leading to early wilting tomato plants.[68] Filamentous phage have also been seen to aid bacterial host survival in *Pseudoalteromonas spp*. in the extreme conditions of artic sea ice by impacting processes such as growth rate, salinity tolerance, motility, and chemotaxis.[69] While filamentous phage have been shown to confer a variety of advantages to their host bacterium, this study focuses on two specific advantages that are important to the pathogenicity of GC: bigger and stronger biofilms and junctional disassembly of human cells leading to higher penetration of bacteria.

Phage have been shown to play a role in pathogenicity through their involvement in biofilm stability. In many species, tailed/lytic phage have been seen to destroy biofilms through a cycle where bacterial cell lysis releases high levels of lytic phage, which in turns leads to more infected and lysed cells, which eventually breaks down the biofilm over time.[70] Inversely, filamentous phage have been found to improve the fitness of biofilms.[61,64,71] Filamentous phage form highly ordered biofilms that are liquid crystal-like due to the high concentration of small polymers leading to depletion forces, which push filamentous phage together. This liquid crystal-like structure inside of the biofilm is created due to depletion attraction, where all of the small polymers in the tightly packed environment of the biofilm push the larger filamentous phage into each other.[61,71] When comparing gene expression in planktonic versus

biofilm *P. aeruginosa*, the most highly upregulated genes in biofilms were filamentous phage associated proteins.[63] These highly ordered biofilms led to lower water loss and higher antibiotic resistance compared to their less ordered, filamentous phage-free counter parts.[61,71] These types of filamentous phage induced changes to the biofilm led to the persistence of disease through immune evasion and ineffective clearance after antibiotic treatment. In the other pathogenic Neisseria species, *N. meningitidis*, their filamentous phage with high homology to GC filamentous phage have also been seen to increase biofilm stability.[23,64] When grown under flow conditions on top of human cell cultures, MDAΦ containing bacteria formed much larger biofilms compared to phage deletion strains. This study also demonstrated a spatial expression pattern, where pili-expressing bacteria were located at the bottom of the biofilm close to the human cells and phage-expressing cells were at the top of the biofilm. Although it is known that GC produce similar filamentous phage, it is still unknown whether filamentous phage can impact GC biofilm fitness.

Another role filamentous phage play in increased pathogenicity is through junctional disassembly of the epithelium in the bacteria's host. *V. cholerae* and *C. concisus* are human pathogenic bacteria that contain filamentous phage regions with homology to GC phage regions.[53] One of these regions of homology is within the GC phage gene ORF9. Within *V. cholerae* and *C. concisus*, the protein with a homologous region to ORF9 is the Zonula occludens toxin (Zot) protein.[19-21,53,66,72] This Zot protein is an endotoxin known to increase the permeability of the epithelial lining of the intestines during infection of *V. cholerae* and *C. concisus* by interacting with host epithelial cells and leading to the disassembly of tight junctions of the mammalian cells.[19-21,66] This increased permeability of the host epithelial cell lining allows for increased penetration of bacteria into deeper tissue. Zot and ORF9 are both transmembrane proteins that aid in the extrusion of phage out of the bacterial cell. [19-21,52,66] Within the GC genome there are five copies of ORF9. The Stein lab has previously investigated whether GC derived purified ORF9 would have the same effect on epithelial cervical cells as the Zot proteins had in the intestines. When T84 monolayers of polarized epithelial cells were treated with purified ORF9 protein, high levels of cell death and junctional disassembly were observed (Figures 5A and B). The same disorganization of the epithelial cell barrier can be seen when ORF9 is exposed to a cervical explant, a piece of human tissue with its cell morphology and organization still intact (Figures 5C and D). While this phenotype can be seen with both

endogenous phage in *V. cholerae* and *C. concisus* and purified GC phage protein ORF9, it is still unknown whether endogenous filamentous phage in GC can lead to epithelial cell disassembly.

Taken together, it has been demonstrated that filamentous phage in a range of bacterial species increase both transmigration rates across host epithelial cell layers and biofilm fitness. It is known that GC has the capability of producing functional filamentous phage, however what impact these phage have on the biology and pathogenicity of GC is still unexplored.[52-54]

1.5 Significance and Goals

Understanding what drives GC to become more or less pathogenic is essential in being able to properly prevent and treat this potentially life-threatening disease. With a rising rate in AMR GC cases, it is critical to determine what can drive the different GC phenotypic changes we see in today's currently circulating GC strains. The goal of this research project is to elucidate whether GC genomic organization or filamentous bacteriophage can impact GC biology. The following aims will address these two processes' roles in GC biology:

1.5.1 Aim 1: The role of genome organization in GC biology.

While genome organization and reorganization has been investigated in GC, these studies were limited to only select strains of GC and the biological impact of these genomic organizational differences was not studied.[36] I will first begin to address this aim by determining changes in GC genome organization in a large group of GC strains. Next, I will analyze the content at the sites of reorganization to identify the main driver of these reorganization events. Once these changes have been analyzed, changes in phenotypes (i.e. growth) will be compared between different genomic patterns. Lastly, the impact of an altered replication termination site on gene expression will be explored.

1.5.2 Aim 2: The role of filamentous bacteriophage on GC biology.

Past studies have shown that GC contain multiple filamentous prophage regions in its genome and that GC are capable of producing functional filamentous phage.[52-54] However, the impact

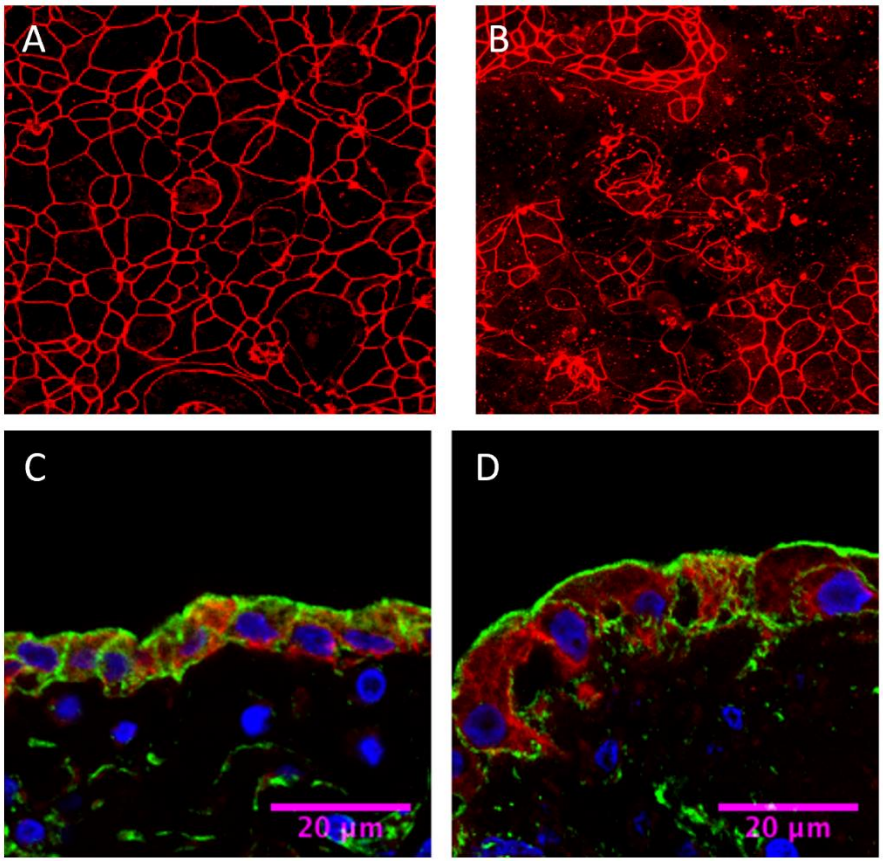


Figure 5. Purified filamentous phage protein ORF9 impacts T84 cell monolayer and cervical explant integrity. (A and B) Images were taken of T84 monolayers exposed to ORF9 (B) or an unexposed control (A) for 1 hour. T84 monolayers were stained for ZO1. (C and D) Images were taken of cervical explants exposed to ORF9 (B) or an unexposed control (A) for 24 hours. Explants were then fixed and stained for actin (green) cytokeratin (red) and DNA (blue). (Mark Wang - unpublished Stein lab)

these phage have on GC biology and pathogenicity is widely unknown. To address this question, I will first explore the impact these filamentous phage have on the growth of GC in various conditions. Next, I will determine the impact filamentous phage have on GC biofilm production and bacterial aggregation.

Chapter 2: The role of Genome Organization in GC Biology

2.1 Introduction

GC infection can lead to a wide range of symptoms.[1] While most female infections are asymptomatic, some women will develop more dangerous sequelae.[1] When investigating why GC presents with such variability in phenotypes, most research to date has focused on allelic differences in genes between strains. While the overall genome organization in most species is largely conserved I wondered if differences between infection outcomes could be caused, not by allelic differences, but by reorganization of the gonococcal genome.[25,26,31] Research using three well characterized strains of GC showed that there were large-scale genome organizational differences between strains.[36] When comparing these three strains (FA1090, NG08107, and NCCP11945), the genome could be divided into 14 different blocks of DNA where recombination occurred, with a variety of mobile elements at the sites of recombination. However, while mobile elements (defined as SSREEs, CREEs, ISs, and phage sequences) were seen at the sites of inversions and translocations, no consistent pattern emerged as the key genomic element responsible for all these rearrangements (Table 1). Furthermore, the gonococcus possesses extensive regions of sequence homology dispersed through any given genome. Therefore, it was unclear whether these genome reorganizational differences were due to the repetitive elements described above, or if recombination could occur at any repetitive sequence. Furthermore, it was unclear if these recombination events were stable or if they impact gonococcal biology. This chapter seeks to answer this question using the newly collected WHO strains with defined antibiotic resistance patterns and multiple well characterized strains linked to specific disease outcomes.

Previous research in a variety of species has shown that genomic rearrangements can impact different phenotypes, as moving genes around can alter genetic expression. One example of this is that genes closer to the origin tend to have higher expression levels as they will have a higher gene dosage while replication is happening.[25,26,31] Most laboratory studies analyzing GC pathogenesis have been performed using only a handful of well characterized strains that were collected decades ago. While most laboratories work with a select few strains of GC, there are hundreds of partial and complete GC genomes in online databases.[1,4,5,36] I have hypothesized that expanding research studies to more strains

might allow me to tease out whether genomic organization can impact certain phenotypes when allelic differences don't seem to explain phenomena, such as AMR differences. I chose to expand these studies to a collection of strains that the WHO recently collected. These clinical isolates were collected from across the globe to better understand the concerning rise in gonococcal AMR. However, how genomic organization effects these strains' AMR capabilities is still unknown.[4,73]

To date, it is still unexplored whether expanding this type of inquiry to more gonococcal strains could expose a genomic organizational pattern associated with higher levels of pathogenicity or a consistent genome feature responsible for reorganization events. I hypothesized that expanding this study to more GC strains will reveal a consistent pattern driving genomic reorganization and that the wide range of AMR capabilities could be due in part to evolutionary changes to genome organization that have occurred in this highly transformable bacterium.

Here I found that when looking at all of the GC genomes that are complete in the NCBI database as of this date, each GC strain had a unique genome organization including both inversions and translocation events. I also saw a consistent pattern of DNA sequence homology on both sides of the translocation or inversions, consistent with homologous recombination driven reorganization. Polymerase chain reaction (PCR) analysis of inversion events suggests that these large-scale reorganization events are both stable and rare. Growth curve analysis demonstrates a wide variability in growth rate among strains. Proteomic analysis suggests reorganization driven changes to replication termination location leads to upregulation of many kinds of proteins, including replication, energy metabolism, and AMR associated genes. Taken en toto, this study suggests that homologous recombination driven genome reorganization can have large impacts on gonococcal biology and pathogenicity. This study demonstrates the need for future gonococcal studies to use multiple GC strains from a diverse background to capture the wide variability in GC phenotypes.

2.2 Materials and Methods

2.2.1 Whole genome alignment and sequence comparison

All *N. gonorrhoeae* were obtained through the NCBI database and analyzed using Geneious Prime. Strain names and their corresponding accession numbers can be found in the appendices at the end. All sequences used in whole genome alignments were oriented to begin at the origin of replication, which was defined as the following sequence: AAAATGTCGTCTGAAATGCTGATATTTCAGTATTCCAGA. Whole genome alignments were completed through a Mauve alignment on Geneious Prime, with LCB score of 5000 and a match seed weight of 28. Alignments of selected sequence fragments were done using a Geneious alignment with an alignment type of global alignment with free end gaps and a cost matrix of 65% similarity (5.0/-4.0).

2.2.2 Growth analysis of *N. gonorrhoeae* strains

GC were grown overnight on GCK plates and were either analyzed for colony size or moved into broth for growth curve analysis. For growth curves, 30 mL of GCP broth supplemented with 1% isovitalex and 1% sodium bicarbonate were inoculated to an OD_{600}=0.10 with selected GC strain and placed in a G76 gyrotory water bath shaker at 37^{o}C on speed setting 3 for seven hours. OD measurements were taken once every hour.

2.2.3 PCR verification of small inversion event

DNA from GC was analyzed using the New England Biolabs Q5® High-Fidelity DNA Polymerase according to package instructions to perform a PCR as per package instructions with 1 µL of DNA in 25 µL reaction. Each cycle had a denature step of 5 seconds and an anneal time of 20 seconds. Anneal temperatures and elongation times varied depending on the primer pairs used. All primers and their associated anneal temperatures and elongation times can be found in the appendices at the end of this thesis.

2.3 Results

2.3.1 Genomic pattern variation among reference strain's genomes

Previous research has shown that there is large-scale genomic organizational differences between strains FA1090, NG08107, and NCCP11945.[36] However, this study restricted its analysis to three gonococcal strains, leaving unanswered the degree to which genome reorganization occurs in the species. This study began by expanding this question to a larger group of strains. When looking for patterns in ribosomal RNA (rRNA) and transfer RNA (tRNA) gene locations in GC strains, grouping of rRNA tRNA gene order and locations started to emerge (Figure 6). While some strains seemed to have completely unique patterns in their rRNA and tRNA gene locations compared to other strains, such as WHO L, FQ01, and 34530, other strains seemed to be in a group of strains of the same patterns, such as the group made of FQ02, MCTC13799, WHO K, WHO P, WHO W, and WHO U. Because I began to see patterns such as these emerge with an analysis as simple as labeling the locations of rRNA and tRNA genes, I decided to use the same genome organization analysis tool used in previous research (Mauve Alignment), but expand the question to a wider variety of GC strains (Figure 7 and Figure S1).[36] Bacterial genomes are circular, however, when using a Mauve alignment, the genome must be linearized. Therefore, I decided to begin all strains at the origin of replication, with the origin in the same orientation for all strains, as to orient every genome to the same spot. The Mauve alignment using a large group of GC strains began to show many genomic reorganizational patterns, including large-scale inversions and translocations within the genome (Figure 7 and Figure S1).

2.3.2 Homologous recombination is the key driver of genome reorganization

Once it had been determined that there were large-scale genome reorganizational events, such as translocations and inversions, the next question I aimed to answer was what was causing these events to occur. Previous research suggested that these events were due to individual mobile elements in the location of these inversions and translocations, including phage genes, insertion sites, CREE, and SSREEs.[36] I sought to see if these sequences were responsible for widespread genome reorganization, or whether other patterns emerged when analyzing the sites of these reorganization events in a larger number of strains. Increasing the number of strains will allow this pattern to come to light because more

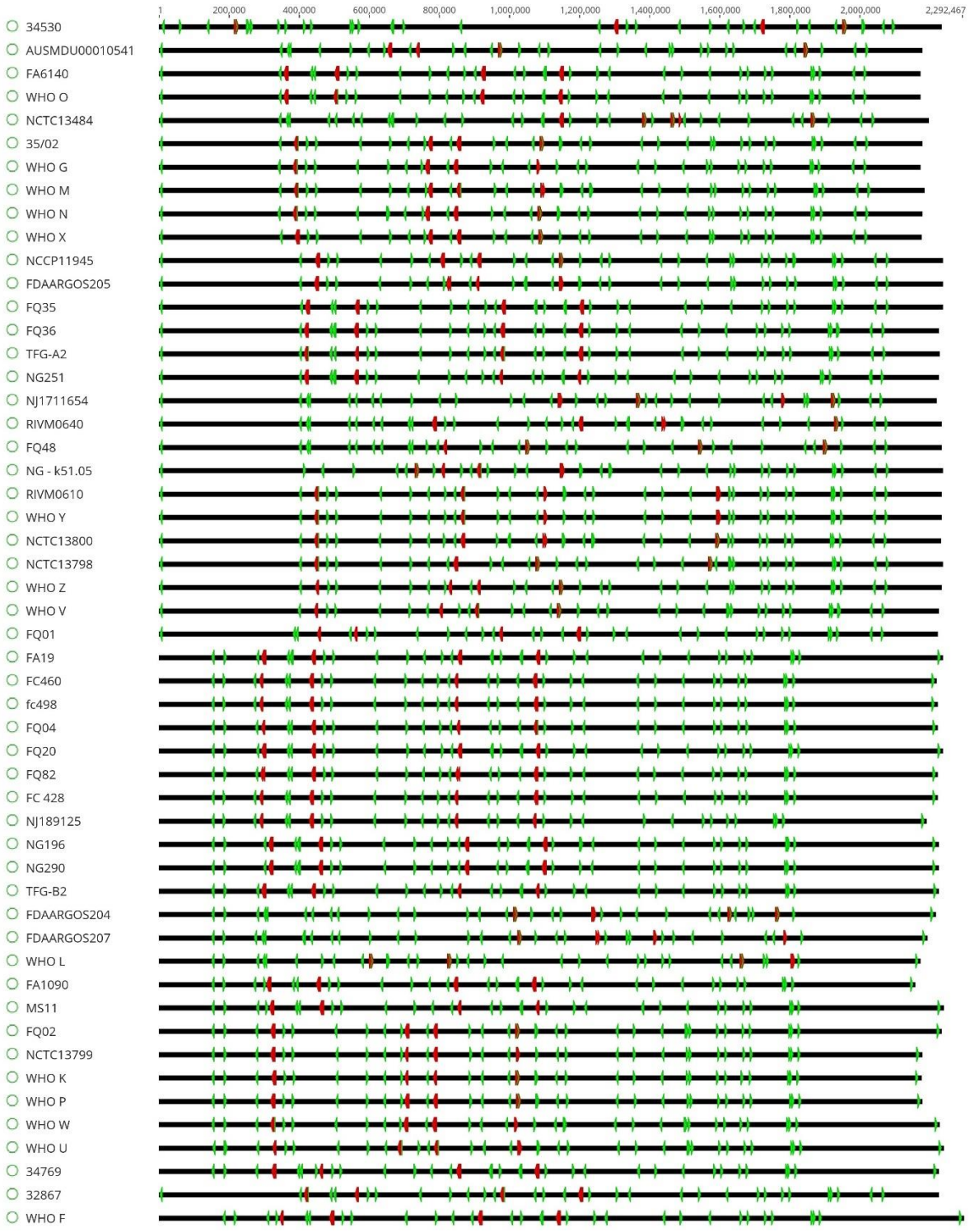


Figure 6. Different genome patterns in different strains of GC. Strain names of the sequences are on the left with their corresponding genomes on right. Black lines represent a linearized version of the bacterial genome beginning at the origin. Red dashes represent rRNA genes and green dashes represent tRNA genes. Organized by RNA gene pattern similarity.



Figure 7. Mauve alignment of various *N. gonorrhoeae*. Mauve alignment of well established and WHO strains of *N. gonorrhoeae*. Each color block represents homologous sequence between strains. All strains begin with the origin of replication on the left. Mauve alignment of all strains analyzed can be found in supplemental figure 1.

reorganizational events can be seen when Mauve alignments are more specific due to a larger sample size.

When looking at the arrangement of the DNA segment blocks in a Mauve alignment, both large (>100,000 nucleotides) and small (<100,000 nucleotides) inversions and translocations can be seen (Figure 7 and Figure S1). To understand what is driving these reorganization events to occur, sequences from either side of an inversion were analyzed for gene content and homology (Figures 8 and 9). While it is true that many mobile elements were seen at the sites of reorganization events and that there were mobile elements scattered throughout the genome, only one genetic pattern could explain nearly every genome reorganization event found; this pattern was homologous reorganization. While many inversion and translocation events were analyzed, two have been selected to demonstrate this phenomenon, both a large and small inversion event (Figures 8 and 9, respectively). In order for homologous recombination to be driving an inversion, the DNA sequence on either side of the inversion must be nearly identical to each other until the point where the inversion takes place. In order to analyze this, the DNA sequence a couple hundred bases in both directions from the site of the inversion were aligned for GC strains with and without the inversion (Figures 8 and 9). For both the large and small inversions shown below, the sequences on the “left” and “right” side of these inversions were aligned to see whether there was homology within each side of the inversion and between the sides of inversion. For both events (and every other inversion analyzed), there were hundreds of nucleotides of nearly identical sequence on either side of the inversion. While mobile elements are often found near these sites of inversion, one trait that all mobile elements have in common is their high levels of homology to each other. This data demonstrates that the mobile elements seen in previous research at sites of reorganization could be there, not as the drivers of reorganization, but as areas with high levels of homology driving homologous recombination.

2.3.3 PCR verification of reorganization events existence and stability

After analyzing dozens of GC strains sequences that were derived from the NCBI database, two key questions arose that are important to any analysis done using sequence data: 1) are these sequence differences real or just computer based sequencing error and 2) are these sequence differences stable

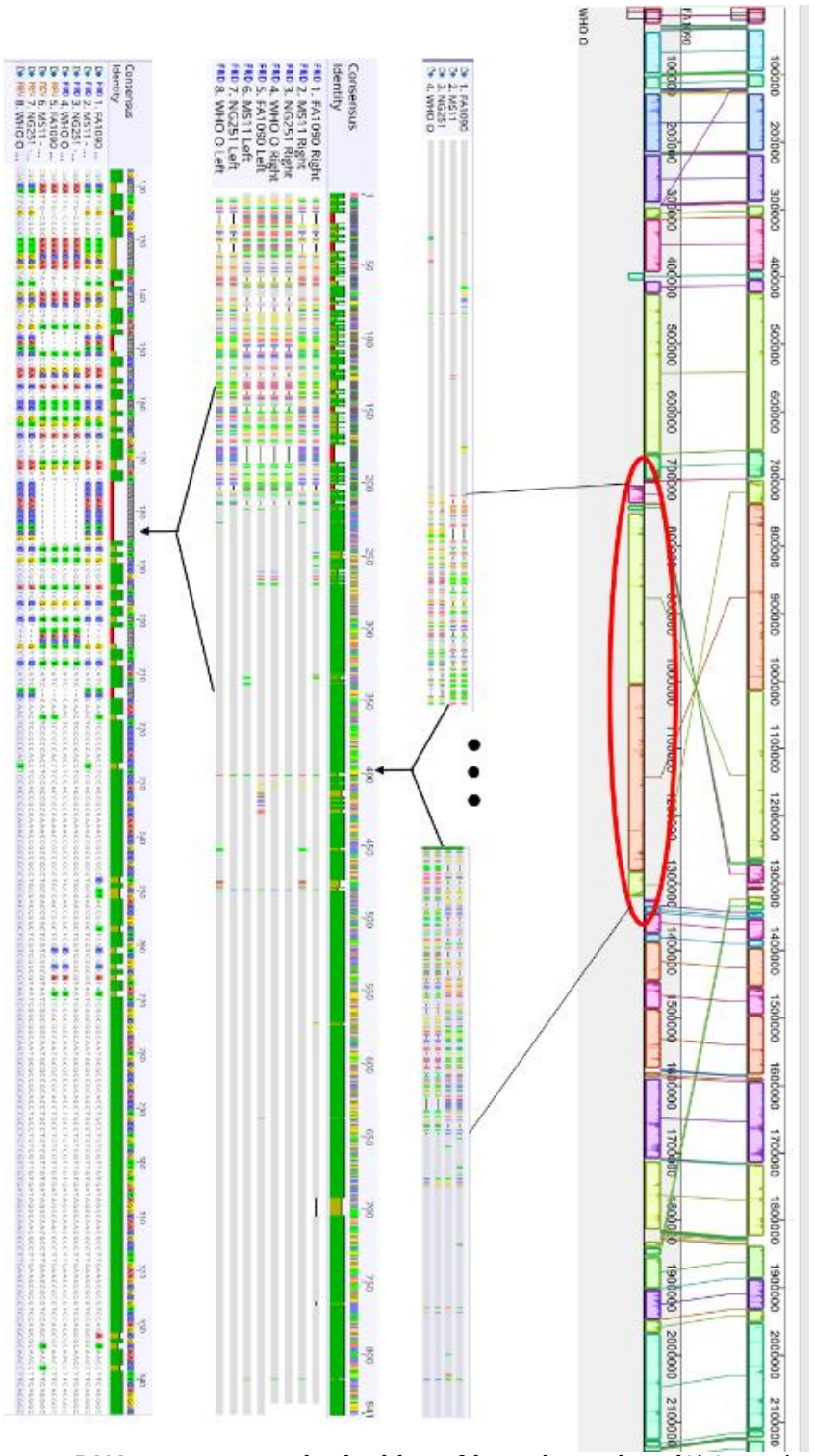


Figure 8. Homologous DNA sequence on both sides of large inversion. (A) Inversion event example from Mauve alignment (Figure 7). The red circle is indicating the inversion event analyzed. (B) Geneious alignment of left and right sides from both non-inverted and inverted regions. (Color meaning in Geneious alignment: gray = same as consensus, pink = A, yellow = G, purple = C, and green = T)

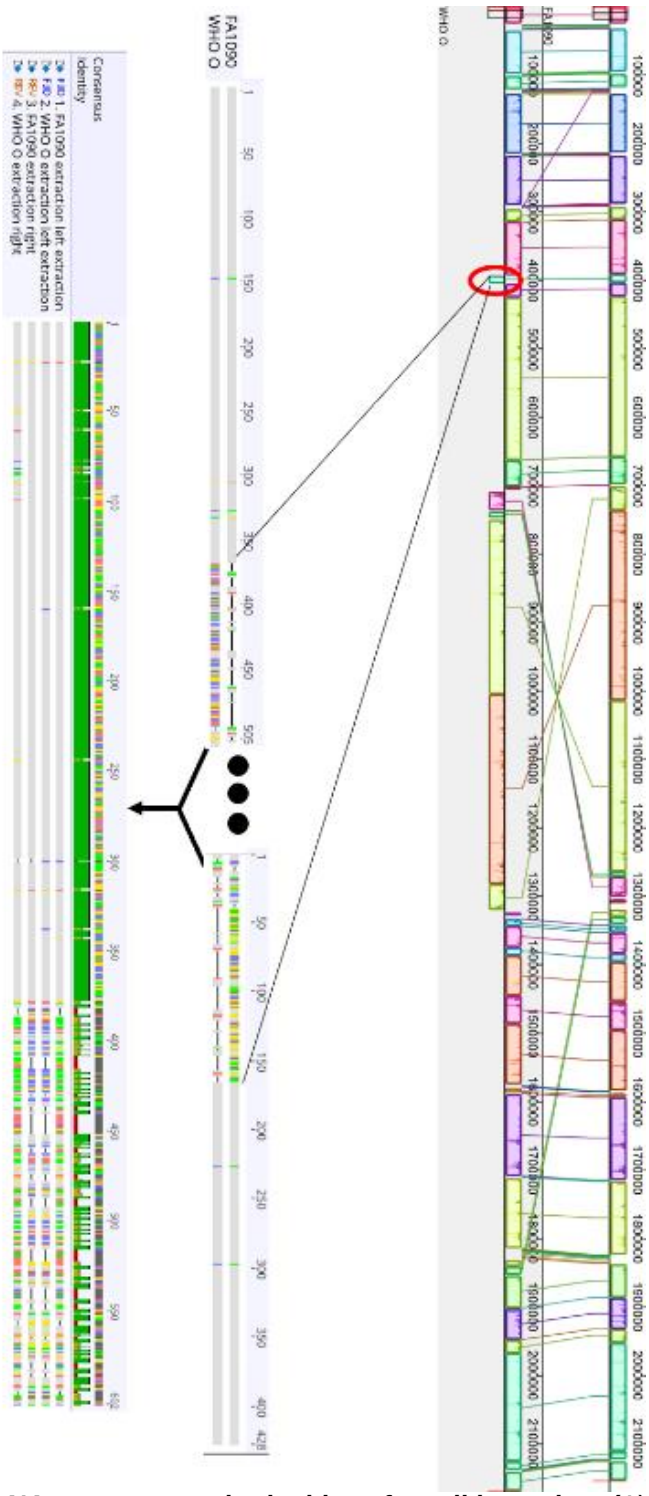


Figure 9. Homologous DNA sequence on both sides of small inversion. (A) Inversion event example from mauve alignment (Figure 7). The red circle is indicating the inversion event analyzed. (B) Geneious alignment of left and right sides from both non-inverted and inverted regions. (Color meaning in Geneious alignment: gray = same as consensus, pink = A, yellow = G, purple = C, and green = T)

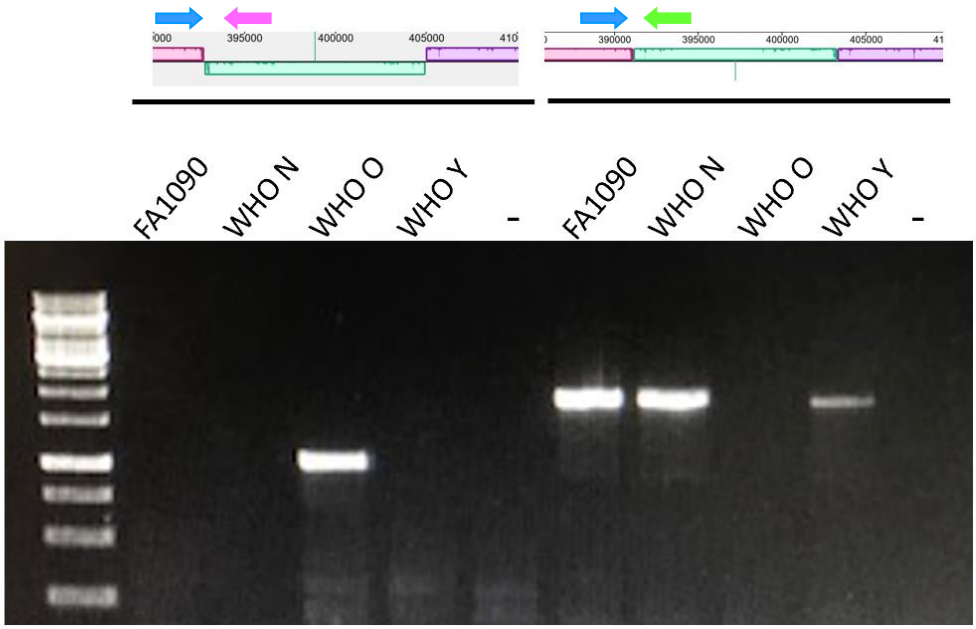


Figure 10. Confirmation of inversion event in database sequences via PCR detection. Primers to detect small inversion event (Figure 9) or lack thereof were used. Left half of gel was primers to detect small inversion event and right half of gel was used to see noninverted genomes. Arrows represent primer placement and orientation. Blue forward primers were the same regardless of inversion status and placed in the sequence prior to inversion site, while the pink reverse primer would only work in the inverted sequence and the green reverse primer would only work in the non-inverted sequence, as the opposite orientation for the pink and green reverse primer would result in two primers facing the same direction, resulting in no amplification.

throughout generations? While many reorganization events were seen and analyzed, there is no way to be sure if what I am seeing in the database sequence is artificial inversions from sequencing errors or whether they are true inversions based solely on the sequences from the database. I validated if an inversion event seen from the sequencing data was both real and stable throughout many generations of growth via PCR analysis (Figure 10). In order to do this, primers were designed that could detect both the inverted and non-inverted version of the small inversion event found on in WHO O from figure 9. These primers worked by having one facing the inversion event, but outside of both the inversion and homologous sequence surrounding it, and a second primer within the inversion event as pictured below (Figure 10). This allowed for the detection of both the inverted and non-inverted sequence, as an inversion detecting primer would be facing the wrong direction for amplification in a non-inverted sequence and vice versa. This PCR and its resulting gel clearly demonstrate the inversion seen in the WHO O database sequence is present in the WHO O bacteria used in this PCR, as it was the only strain to have detected inverted sequence and lacking any non-inverted sequence. Due to the sensitivity of this PCR technique, these data also demonstrate that these reorganization events are quite rare and stable throughout generations of bacteria, as any frequency of inverting or reverting this sequence at more than ~one in a billion cells would result in a band in both the inversion and non-inversion primer set.

2.3.4 Growth differences between WHO gonococcal strains

While the WHO GC strains have high levels of genetic homology to each other, there is a wide variation in the order and orientation of the genes they contain. At the largest scale, one can look at the orientation, and placement of the origin and termination sequences relative to each other within the genomes of these strains (Table 2). When analyzing where the termination sequence is in reference to the origin of replication, GC strains can be categorized into three main groups: strains where the termination sequence is just under halfway through the genome (WHO K, WHO O, WHO P, WHO U, and WHO W), strains where the termination sequence is just over halfway through the genome (WHO Y, FA1090, MS11, NCCP11945, WHO F, WHO G, WHO M, WHO N, WHO V, WHO X, and WHO Z), and WHO L which is unique in that its terminal sequence is only about a quarter of the way through the genome (Table 2). With WHO L having its termination sequence at a quarter of the way through the

Table 2. Relative location of various genes involved in antibiotic efflux**.**

Strain	Genome[a] Size	Dif[b] location	distance[c] from origin	*mtr*[d]	distance from origin	*macA*[e]	distance from origin	*norM*[f]	distance from origin
WHO Y	2228980	1195527	0.54	1684848	0.76	1772279	0.80	818972	0.37
FA1090	2153922	1139621	0.53	594766	0.28	511463	0.24	1529483	0.71
MS11 ORI	2233640	1150105	0.51	596504	0.27	510072	0.23	1601175	0.72
NCCP11945	2232025	1153245	0.52	1643981	0.74	1731413	0.78	776101	0.35
WHO F	2292467	1208315	0.53	626619	0.27	540065	0.24	1657265	0.72
WHO G	2167361	1154675	0.53	1583198	0.73	1666087	0.77	777033	0.36
WHO K	2169846	1002353	0.46	585594	0.27	499711	0.23	1393966	0.64
WHO L	2168633	584423	0.27	1128511	0.52	197158	0.09	1102650	0.51
WHO M	2178344	1154006	0.53	1585011	0.73	1672444	0.77	776912	0.36
WHO N	2172826	1155782	0.53	1584161	0.73	1667043	0.77	778143	0.36
WHO O	2169062	1004710	0.46	1549125	0.71	1635797	0.75	627112	0.29
WHO P	2173861	1004767	0.46	590553	0.27	502260	0.23	1395187	0.64
WHO U	2234269	1007459	0.45	589023	0.26	501813	0.22	1455491	0.65
WHO V	2221284	1148685	0.52	1635778	0.74	1723210	0.78	772685	0.35
WHO W	2222386	998073	0.45	583093	0.26	497213	0.22	1445875	0.65
WHO X	2171112	1150814	0.53	1584374	0.73	1670481	0.77	774741	0.36
WHO Z	2229351	1150982	0.52	1641017	0.74	1727126	0.77	774727	0.35

[a]Genome size was determined from the GENBANK entries for each strain as described in the materials and methods.

[b]The *dif* location was defined as the first Adenine in the sequence AGTTCGCATAATGTATATTATGTTAAAT.
[c]Distance from origin was calculated by dividing the *dif* location coordinate by the total genome size for that strain.

[d]Location of the *mtr* region was defined by the start codon for *mtrE*.

[e]Location of the *macA* gene was defined by the start codon for *macA*.

[f]Location of the *norM* gene was defined by the start codon for *norM*.

genome, this means that half of the replication machinery has to make it ~75% of the way through the genome in order to replicate, as replication happens in a bidirectional fashion.[25,31] This lead to the hypothesis that WHO L may in turn grow slower as it would take longer to fully replicate its genome.

In order to test this hypothesis, growth of various WHO strains was compared to lab strains to determine if any variation in growth phenotypes could be seen. In both broth and agar, a wide variation ingrowth rates were detected between strains (Figure 11). WHO L grows the slowest in broth and produces the smallest colonies on agar, compared to all other WHO strains. This demonstrates how moving the terminal sequence to another point in the genome can have severe consequences on gonococcal biology.

2.3.5 Selection pressure against replication terminus movement

When examining dozens of GC genomes in the database, I noticed that WHO L was the only GC strain that didn't have its replication terminus near the halfway point of the genomes, defined as 40-60% through the genome compared to the origin (Table 2 and Figure S1). This raised the question of how, despite the numerous homologous recombination events observed, no other strain appeared to have shifted its replication terminus as drastically as WHO L. When looking at the homologous recombination events that occurred in figures 8 and 9, if there were other regions in the genome with high sequence similarity, I would expect rearrangements happening between these locations as well. When looking at the large inversion event, I identified a region of 50 base pairs (allowing for 2 mismatches) flanking the inversion and found only one other instance of this 50 base pair region elsewhere in the genome (Figures 8 and 11). Although this large inversion is present in many strains, this inversion does not involve much movement of the terminus location. However, if the inversion between the first and third similar sequence were to occur, it would significantly shift the terminus location (Figure 11). However, this inversion was not seen in any of the strains analyzed. When looking at the small inversion in the same way, its 50 base pair sequence was found multiple times throughout the genome (Figures 9 and 11). However, many of these sequences are never found near a rearrangement event. Based on the observed rearrangement events and WHO L's unique terminus relocation, I hypothesize that selection pressure likely opposes drastic shifts in the replication terminus.

This leads to the question: how could WHO L's replication terminus shift have come about if selection pressures typically oppose it? Using FA1090 as a hypothetical parent strain, I found that only three inversions would be needed to reorganize the genome to match WHO L's structure (Figure 12). The first rearrangement is the large inversion seen in many GC strains (Figures 8 and 12). This is followed by two additional inversions between highly similar sequences, each incrementally shifting the terminus and ultimately yielding WHO L's structure. With only three inversions to obtain a WHO L-like genome organization, this suggests strong selective pressure against this reorganization, as it was not seen in any other analyzed strain.

2.3.6 Changes in gene expression of key bacterial genes

While the data showed that changing the location of the terminal sequence can have severe consequences to growth rate, it does not demonstrate how it can impact other pathogenic associated phenotypes. To analyze this, I examined how genomic rearrangements can impact the expression of specific genes, and therefore impact different aspects of GC biology. I took advantage of a proteomic data set published by El-Rami et al., which analyzed gene expression in the WHO strain set, including AMR associated genes.[40] Due to the drastic shift in overall genomic structure, which would result in an uneven distribution of replication needs on either side of the origin, I first decided to look at the changes in the expression of DNA Polymerase III genes, the polymerase responsible for replicating the genome.[25,31] The raw proteomic data analyzed in this thesis is taken from the El-Rami, F. E. et. al. paper (Figures 14 and 15).[40] When comparing the expression of various DNA Polymerase III genes between WHO L and the 8 other WHO strains (WHO G, K, M, N, O, P, U, and V), WHO L had upregulation of 4 of 5 analyzed DNA Polymerase III genes (Figure 14). Due to this upregulation of DNA Polymerase III, I hypothesized that if this upregulation causes a higher rate of replication initiation in WHO L compared to the other WHO strains, that the implication of this would be that genes close to the origin might have higher levels of expression due to a higher gene dosage per cell. When looking at three genes close to the origin of replication, all three genes (sigma 70 whose expression could impact overall translation, and HaeII and HaeIII whose expression would be expected to be stable in a cell), I found that all of these genes had higher expression in WHO L compared to the other WHO strains (Figure 14).

A

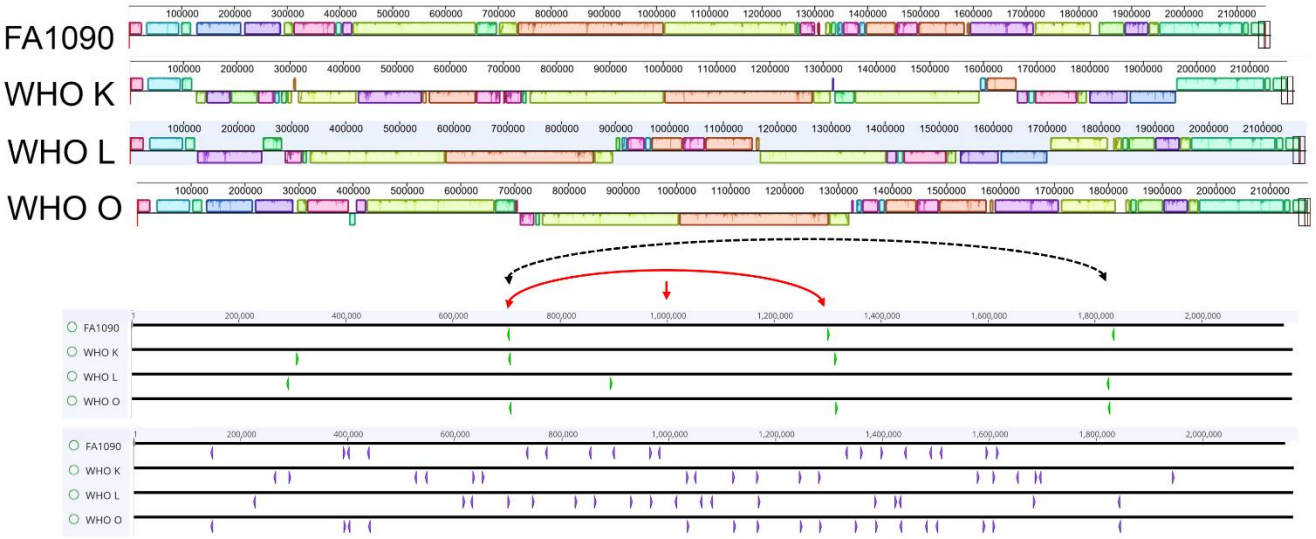


Figure 11. Selection pressure against selecting inversions that move replication terminus site. (A) Mauve alignment of FA1090 and WHO K, L, and O. (B) Areas of the genome with the same 50 base pair region (allowing for 2 mismatches) flanking the inversion of either the large inversion (green) or the small inversion (purple) in figures 8 and 9, respectively. The double-sided red arrow indicates the inversion seen in multiple analyzed strains while the inversion represented by the dashed black arrow was never found. A small red arrow indicates the location of the replication terminus site.

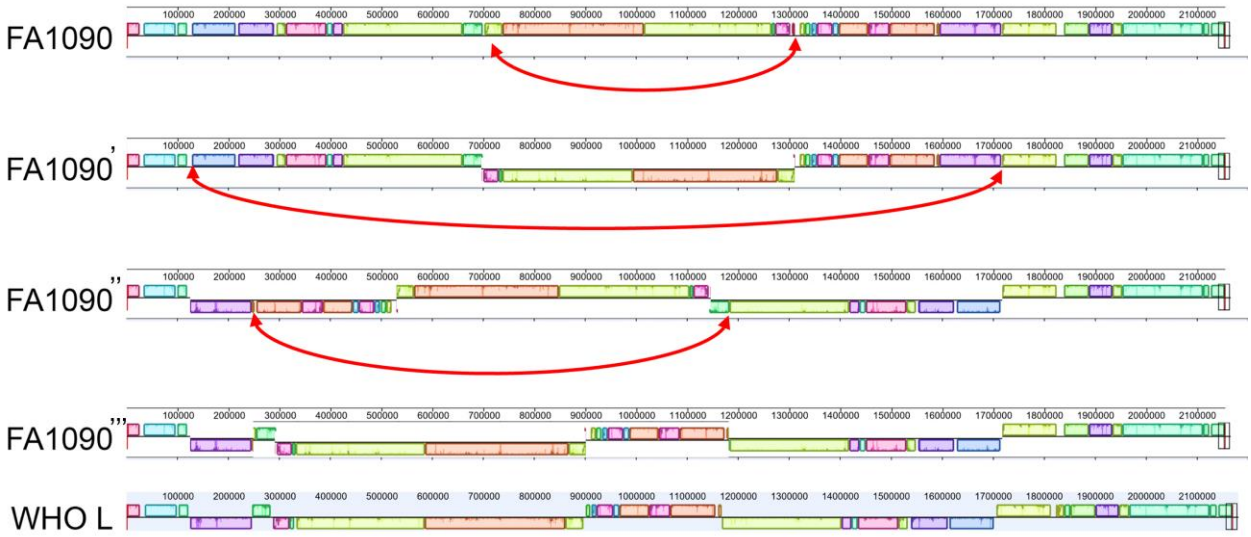
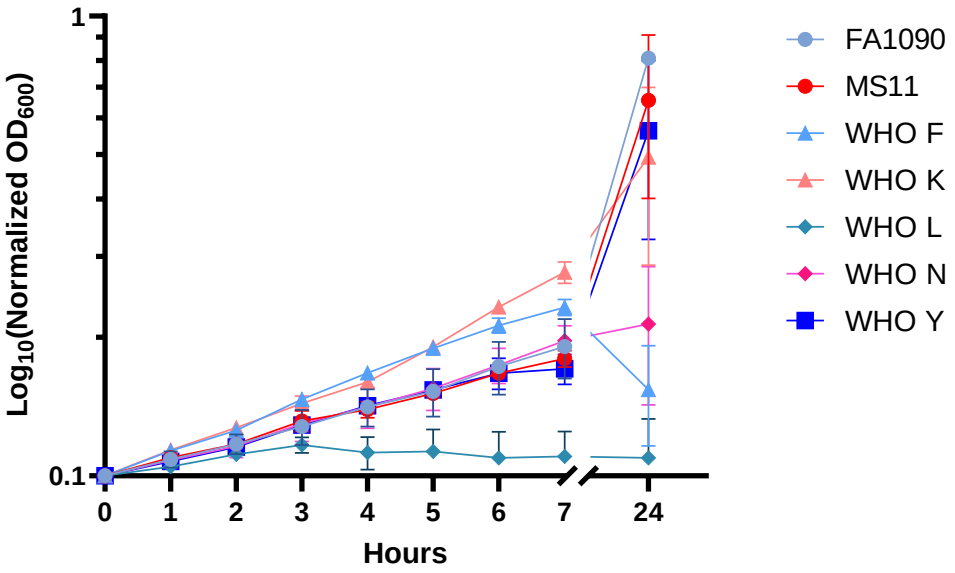


Figure 12. Possible inversions to convert FA1090-like genome organization to WHO L organization. Mauve alignment of FA1090 shown at the top, with each successive genome organization resulting from the indicated inversion events shown below. Red arrows indicate inversion events between sequences of high similarity. Three total inversions within FA1090 leads to the bottom Mauve alignment of WHO L shown.

A



B

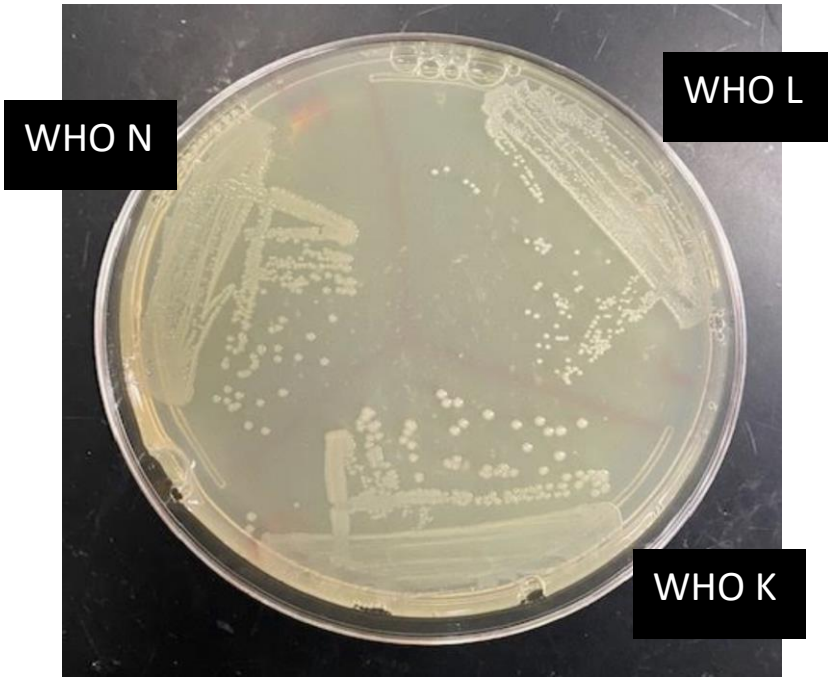


Figure 13. Variation in growth in both broth and agar. (A) Growth curve of various historical laboratory strains of GC and WHO strains of GC. OD_{600} measured every hour in GCP broth. (B) Overnight growth of colonies of WHO K, WHO L, and WHO N on GCK agar.

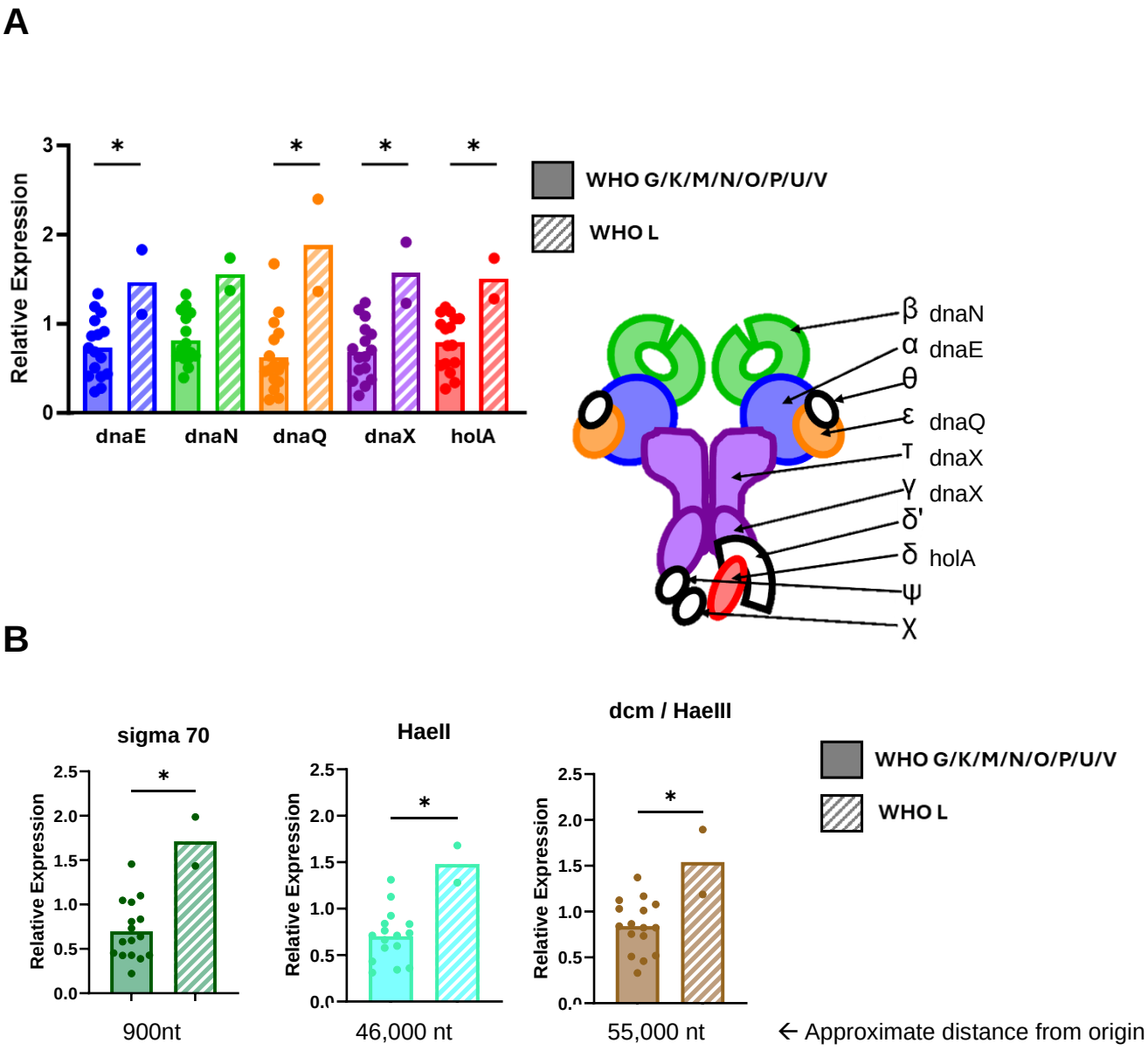


Figure 14. DNA replication machinery and genes near origin are over expressed in WHO strain with shifted terminal sequence. (A) Relative protein expression of various DNA Polymerase III genes and their corresponding location within DNA Polymerase III complex. Black and white proteins on complex represent proteins not analyzed in proteomic study. DNA polymerase III diagram adapted from Kelman et. al.[74] (B) Relative protein expression of three genes near the origin with their approximate nucleotide (nt) distance from the origin. All solid bars represent average relative protein expression or WHO strains G, K, M, N, O, P, U and V while hashed bars represent the average relative protein expression for WHO L. Proteomic data analyzed was retrieved from El-Rami, F. E. et. al. paper.[40] * = $p<0.05$, each dot represents one replicate's protein expression relative to WHO F expression.

After seeing an upregulation in protein expression of DNA Polymerase III genes and genes close to the origin, I hypothesized that different categories of genes important to bacterial fitness may also have changes in protein expression in WHO L. RNA Polymerase genes are important in overall protein expression as it is required for translation of expressed genes. When looking at three RNA polymerase genes (*rpoA*, *rpoB*, and *rpoC*), all three genes were upregulated in WHO L compared to other strains (Figure 15). One important aspect of bacterial fitness in the context of in vivo infection is the ability to be resistant to any number of antibiotics used to treat bacterial infections. When looking at the expression of three important AMR genes in GC (*gyrA*, *mtrC*, and *ponA*), all three had higher levels of protein expression (Figure 15). Another key aspect to bacterial fitness is its ability to produce energy in order to complete various functions within its cell. When looking at three key genes important to the Krebs cycle in GC (*aceE*, *icd2*, and *gltA*), all three genes were significantly upregulated. All of these data suggest an upregulation of gene expression in a variety of processes important to bacterial fitness.

2.4 Discussion

This study sought to determine the causes of genome reorganization and its impacts on gonococcal biology. Previous research on this topic heavily focused on only a handful strains of GC and hadn't looked at downstream impacts of genome reorganization on GC biology.[36] This research showed a wide variety of genome organizational patterns between GC strains (Figures 6 and 7). These changes to genome structure showed not only small translocations and inversions, but also large-scale inversions of regions that were hundreds of thousands of nucleotides in length. One of the major changes seen to overall genome structure was in WHO L, who had its genomic geography changed drastically, as its termination sequence was moved to only about a quarter of the way through its genome (Figures 7 and 12 and Table 2). However, with this being the only strain found to move its replication termination site, it is hypothesized that there must be some selection pressure against doing this (Figures 11, 12, and S1 and Table 2). This study demonstrated that the key driver of these reorganization events is simply homologous recombination between highly similar pieces of DNA (Figures 8 and 9), and are not due to the presence of a specific DNA sequence element, as previously published.[36] Homologous recombination was seen to drive both large and small inversion and translocation events (Figures 8 and 9). While a few

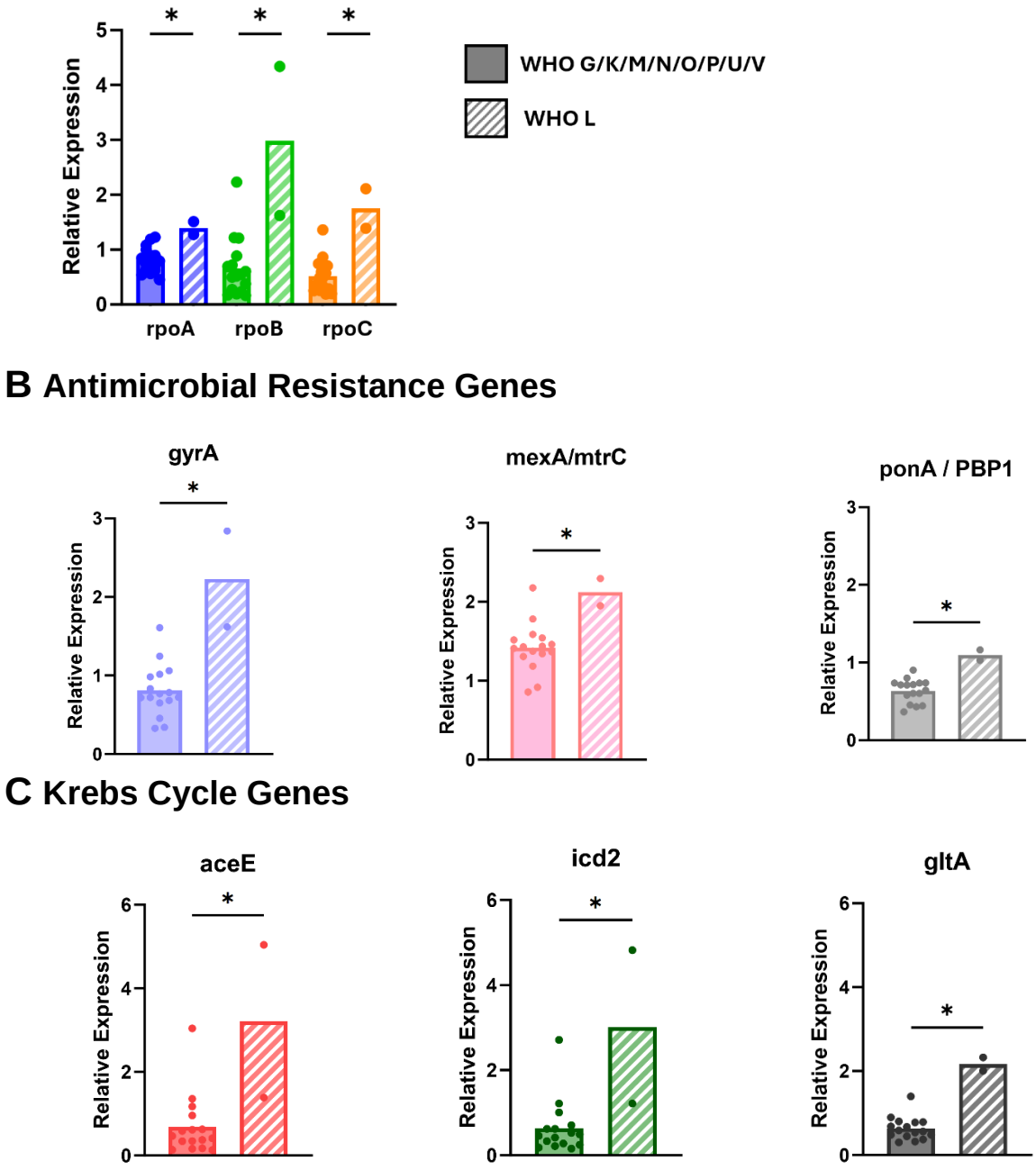


Figure 15. Multiple bacterial protein types upregulated in WHO L. B) Relative protein expression of RNA Polymerase genes *rpoA*, *rpoB*, and *rpoC*. B) Relative protein expression of three genes associated with AMR. C) Relative protein expression of genes involved in citric acid cycle. All solid bars represent average relative protein expression of WHO strains G, K, M, N, O, P, U and V while hashed bars represent the average relative protein expression for WHO L. Proteomic data analyzed was retrieved from El-Rami, F. E. et. al. paper.[40] * = $p<0.05$, each dot represents one replicate's protein expression relative to WHO F expression.

inversions did not appear to have homologous sequences on both sides of the inversion, these exceptions could be explained by a subsequent translocation that occurred after the inversion had already taken place. These reorganization events were confirmed in a laboratory setting to demonstrate that the reorganization events see in database acquired sequence data is truly how the genome is organized (Figure 10). This data also showed that these reorganizational events are rare and stable, supporting the conclusion that these are happening via homologous recombination. While past research showed what was at the sites of reorganization (phage, CREE, SSREE, pilin, etc.), it was not until a larger range of strains were studied that this key pattern could emerge, emphasizing the need to encompass many strains of GC into research.

While many instances of homologous recombination are observed, the GC genome contains hundreds of sites with high sequence similarity. However, we do not see reorganization occurring at hundreds of these possible homologous recombination sites. As shown in figure 11, there appears to be some selection pressure against moving the site of replication termination, as a shifted replication termination site was only in the WHO L train out of dozens analyzed. Other factors may also explain why specific inversions are not observed in the genome. For instance, an inversion occurring within an essential gene could disrupt its sequence, resulting in a lethal outcome, thus preventing this inversion from being seen. Similarly, inversions near promoters could alter gene expression by relocating a gene closer to or further from a promoter, potentially leading to fatal changes in gene expression and, therefore, an absence of such inversions in the genome.

Despite the many examples of homologous recombination driven genome reorganization, it is not well understood why GC uses this mechanism as a means of genome reorganization more often than other bacterial species. It has been shown that GC lacks an effective SOS-like system, meaning that DNA repair following global DNA damage, such as during replication fork pausing, does not rely on the SOS system.[75,76] It is possible that GC utilizes homologous recombination more frequently because this may be the preferred method of DNA repair in response to global DNA damage, whereas other bacterial species may rely in the SOS DNA repair pathway in addition to normal homologous recombination repair.

This study also showed a wide variety in growth phenotypes between strains and variations of gene expression in genes with identical sequences. It is seen in almost every strain of GC that the terminal sequence for replication is found at about 50% of the way through the genome. In WHO L, this changed to about 25% of the way through its genome. This new location means that half of the DNA replication machinery must travel ~75% of the way around the genome during each replication, as DNA replication happens bidirectionally. This drastic change caused a severe growth defect to be found in WHO L, with barely any growth observed (Figure 13). With drastic changes to growth fitness phenotypes and overall genomic structure, it was still unknown how these changes can impact the expression of various gene types in GC, both in genes grouped by function and location within the genome. It has been seen in countless species of bacteria that gene location and orientation can drive differences in gene expression.[25,26,31] However, this is the first time analyzing how large-scale genomic structural changes can drive changes in gene expression in GC strains. Due to WHO L's slow growth rate coupled with its higher expressed DNA replication machinery, there is an upregulation of protein expression found in genes near the origin of replication (Figure 14). This can lead to an accumulation of these proteins, which in turn can drive phenotypic changes to the bacteria. Not only are DNA replication proteins upregulated in WHO L, but translation proteins are upregulated as well (Figure 14). Higher expression of translation machinery coupled with higher dosage of genes due to slow growth and high levels of DNA replication led to the question of whether this can impact expression of genes directly linked to bacterial fitness. The WHO's high concern regarding the rising levels of AMR led to analysis of the expression of genes highly associated with AMR within GC strains. While most studies focus on mutations within these genes or their regulators, this study focused on whether its possible that these genes are upregulated in the highly resistant WHO L strain, which could in turn be contributing to the higher level of AMR seen. This data suggests that many AMR associated genes are expressed significantly more than the other WHO strains analyzed (Figure 15). Energy metabolism pathways are also important to bacterial fitness, as most key functions in bacteria require energy. When looking at multiple Krebs cycle genes, three of the analyzed genes were significantly upregulated in WHO L. With WHO L exhibiting slow growth and upregulation of key protein types, such as AMR and energy metabolism, this can lead to the accumulation of these proteins, thereby increasing bacterial fitness without changing the overall gene content. Genomic

organization causing a large variation in a range of phenotypes within GC strains demonstrates that gene location and orientation, in addition to sequence variation, can impact the differences seen both in research and clinical outcomes between GC strains. Future research should keep this in mind when trying to elucidate genetic causes of unexplained phenotypic variation seen between GC strains.

The key conclusions from this chapter highlight the importance that GC studies should contain multiple GC strains in order to fully capture the variations between GC genome organization. The WHO collected and created the WHO GC strains to better understand the state of currently circulating strains of GC in the world today. These strains had higher levels of variation in their genomes compared to the historical lab strains, highlighting the need to keep GC studies relevant to fresh clinical strains and keep GC strain variation in mind when designing studies. One of the key goals of this study was to determine whether genome organization can impact changes in gene expression. The drastic change to WHO L's genomic structure being associated with significant changes in gene expression demonstrates exactly this. WHO L has very high levels of AMR to many types of antibiotics. While it is still unknown whether altering AMR-associated gene location or changing overall genome structure can directly cause higher levels of AMR, this study clearly illustrates the need to investigate phenotypic changes driven by elements outside of allelic differences. With the rising threat of AMR within GC strains today, this study demonstrates the need to keep the strains studied in the lab updated and relevant to the state of GC actively being spread in the world.

Chapter 3: The role of Filamentous Bacteriophage in GC Biology

3.1 Introduction

While the previous chapter focuses on the gonococcal genome and its organization, GC genomes also contain sections of DNA that are viral in nature. Within GC, prophage DNA makes up about 7% of the genome.[52-56] GC has within it four filamentous prophage regions, which in FA1090 are known as NgoΦ6-9 (or PHI6-9) (Figure 4).[52-56] Though these filamentous phage take up tens of thousands of bases of the genome, little is known of their impact on gonococcal biology. These phage have many homologous ORFs in their sequences. These ORFs are not only homologous to each other, but also to other species such as *N. meningiditis* and *V. cholerae*, with some of these homologous genes conferring heightened pathogenicity to their bacterial hosts.[21,23,53,57] Filamentous phage confer a variety of advantages to their bacterial hosts. Adding the *V. cholerae* filamentous phage, CTXΦ, into non-pathogenic strains can leads to an increased virulence of these strains.[67] The filamentous phage ΦRSS1, found in the phytopathogen *R. solanacearum*, enhances its host's virulence through changes in motility and extracellular polysaccharide synthesis, leading to early wilting tomato plants.[68] Filamentous phage have also been seen to aid bacterial host survival in *Pseudoalteromonas spp*. in the extreme conditions of artic sea ice by impacting processes such as growth rate, salinity tolerance, motility, and chemotaxis.[69] While filamentous phage have been shown to confer a variety of advantages to their host bacterium, this study focuses on two specific advantages that are important to the pathogenicity of GC: bigger and stronger biofilms and junctional disassembly of human cells leading to higher penetration of bacteria.[23,50,60,62,64,65,67,77]

Filamentous phage in a variety of species have a role in pathogenicity through their involvement in biofilm stability. [61,64,71] When comparing *P. aeruginosa* gene expression in planktonic versus biofilm bacteria, the most highly upregulated genes in biofilms were filamentous phage associated proteins.[63] Phage containing biofilms lead to lower water loss and higher antibiotic resistance compared to their less ordered, filamentous phage-free counter parts.[61,71] In the other pathogenic Neisseria species, *N. meningitidis*, their disease associated filamentous phage, MDAΦ, seems to increase biofilm stability.[23,64] When grown under flow conditions on top of human cell cultures, MDAΦ-containing bacteria produced

much larger biofilms compared to phage deletion strains. This study also demonstrated a spatial expression pattern, where pili-expressing bacteria were located at the bottom of the biofilm close to the human cells and phage-expressing cells were at the top of the biofilm. These types of changes to the biofilm due to the presence of filamentous phage can lead to the persistence of disease through immune evasion and ineffective clearance after antibiotic treatment. While it is known that GC produce filamentous phage, it is still unknown whether filamentous phage can impact GC biofilm fitness.

Another role filamentous phage play in increasing pathogenicity is through junctional disassembly of the epithelium in the bacteria's host. *V. cholerae* and *C. concisus* are human pathogenic bacteria containing filamentous phage regions with homologous regions to GC.[53] One of these regions of homology is within the GC phage gene ORF9, which has high homology to the *V. cholerae* and *C. concises* endotoxin Zot.[19-21,53,66,72] This Zot protein is known to increase the permeability of the epithelial cell lining of the intestines during infection by interacting with host epithelial cells and leading to the disassembly of tight junctions between the mammalian cells.[19-21,66] This increased permeability of host epithelial lining allows for increased penetration of bacteria into deeper tissue. Zot and ORF9 are both transmembrane proteins that aid in the extrusion of phage out of the bacterial cell. Within the GC genome there are five copies of ORF9. The Stein lab has previously explored whether GC purified ORF9 has the same impact on the epithelial cervical cells as the Zot proteins had in the intestines. When T84 monolayers of polarized epithelial cells were treated with purified ORF9 protein, high levels of cell death and junctional disassembly were observed (Figures 5A and 4B). The same disorganization of the epithelial cell barrier can be seen when ORF9 is exposed to a cervical explant, a piece of human tissue with its cell morphology and organization still intact (Figures 5C and 4D). While this phenotype can be seen with both endogenous phage in *V. cholerae* and *C. concisus* and purified phage protein ORF9, it is still unknown whether endogenous filamentous phage in GC can lead to epithelial cell disassembly.

While there is a consistent pattern of filamentous phage being associated with increased pathogenicity for their bacterial hosts in a wide range of species, little is known about the impact of filamentous phage on GC biology and pathogenicity. This study aims to determine the prevalence of these phage amongst GC strains and their impact on GC biology and pathogenicity. I hypothesize that

GC filamentous phage can confer increased pathogenicity by impacting various aspects of GC biology. To test this hypothesis, I used an isogenic derivative of strain FA1090 (FA1090ΔfilΔ) that lacked the four phage genomes using a PCR/spot transformation protocol. The correct deletions of the various phage encoding regions were verified by DNA sequencing. I found that FA1090ΔfilΔ grew poorly at 37°C in both broth and on agar, as compared to FA1090 Wild Type (WT). However, there was no difference when the strains were grown at 34°C or when grown without shaking. When these strains were analyzed for their ability to produce biofilms, no differences were seen in the overall biofilm's biomass, yet the overall structure of the biofilms were different, with FA1090ΔfilΔ producing taller and rougher biofilms. FA1090ΔfilΔ formed larger bacterial aggregates than FA1090 WT. Taken en toto, these data suggest that filamentous phage provide the gonococcus with a growth advantage and can impact different aspects of gonococcal infection mechanisms such as biofilm structure and aggregation capabilities.

3.2 Materials and Methods

3.2.1 Growth of *N. gonorrhoeae* strains

GC were grown overnight on GCK plates and either analyzed for colony size, moved into broth for growth curve analysis, or used downstream for further experimentation. For growth curves, 20 mL of GCP broth with 1% Kellog's supplement and 1% sodium bicarbonate were inoculated to a Klett=10 with selected GC strain and placed in a shaker at 37°C or 34°C for 24 hours. OD measurements were taken once every hour for 6 hours and again at 24 hours

3.2.2 Creation of Filamentous Bacteriophage Mutants

The GC strain FA1090ΔfilΔ was constructed by Zhu Han in the Stein lab. Locations and gene content of each region can be found below (Figure 16). This is a strain of FA1090 with all four regions of filamentous phage regions deleted out of its genomes. First, PCR amplification of upstream and downstream regions of each phage region in FA1090 WT were completed with restriction enzyme sequences at the 5' end of each primer (Figure 17). These sequences were then ligated together (PHI6) or around an antibiotic cassette (PHI7-9) and transformed into FA1090 to create deletion regions. DNA from FA1090 WT and FA1090ΔfilΔ strains were then purified using Monarch® Plasmid Miniprep Kit according to package

instructions and sent to Plasmidsaurus© for DNA sequencing (n=2). This sequence data was used for sequence verification of the deletion mutant.

All single, double, and triple phage deletion GC strains are in a FA1090 background of either wild type FA1090 or a FA1090ΔfiI strain. Creation of single, double, and triple mutants were made using spot transformation, with the exception of antibiotic selection for the deletion of PHI8 and PHI9 regions. Deletion region DNA was added to FA1090 WT or phage region DNA was added to FA1090ΔfiI followed by a PCR verification screen using the New England Biolabs Q5® High-Fidelity DNA Polymerase or LongAmp® Taq DNA Polymerase according to package instructions to perform a PCR with 1µL of DNA in 25µL reaction. “Deletion region” and “phage region” DNA fragments were created via PCR of desired region (ΔPHI6, ΔPHI7, ΔPHI8, or ΔPHI9 / PHI6, PHI7, PHI8, or PHI9) in either FA1090ΔfiI or FA1090 WT, respectively. All primers and their associated anneal temperatures and elongation times can be found in the appendices at the end.

3.2.3 Biofilm Analysis of GC

GC was grown overnight on GCK plates. 300µL of each GC strain at an OD=0.1 was added to an 8-well glass chamber slide and grown for either 6 or 24 hours at 37^{o}C or 34^{o}C. After 6 or 24 hours, 50µL of Acridine Orange was added to each well and placed back in the incubator for 20 minutes to allow proper staining. After staining concluded, 9 Z-stack images were taken per well of the biofilm on a Zeiss LSM 980 with Airyscan 2 confocal microscope using a 40X objective. These 9 images were taken distributed around the well and did not overlap on the X/Y plane. The number of stacks per image varied, as Z-stacks were set from the first Z-plane location where no bacteria were visible underneath the biofilm to the first Z-plane location where bacteria are no longer visible above the biofilm. The laser excitation wavelength value was 488 and the emission wavelength value was 509. Biofilm images were then analyzed using Comstat 2.0 on FIJI with automatic thresholding and connected volume filtering turned off.

3.2.4 Aggregation analysis of GC

GC was grown overnight on GCK plates. 198µL of each GC strain at an OD=0.0033 was added to an 8-well glass chamber slide and grown for 4.5 hours at 37^{o}C. After 4.5 hours, 2µL of 1mM Syto9 was added

to each well and placed back in the incubator for 20 minutes to allow proper staining. After staining concluded, 9 Z-stack images were taken per well of the aggregates on a Zeiss LSM 980 with Airyscan 2 confocal microscope. These 9 images were taken distributed around the well and did not overlap on the X/Y plane. Number of stacks per image varied, as Z-stacks were set from the first Z-plane location where no bacteria were visible underneath the aggregates to the first Z-plane location where bacteria are no longer visible above the aggregates. Images were taken in both brightfield and a syto9 channel with laser excitation wavelength value was 483 and the emission wavelength value was 500. Aggregates were then measured via 3D Object Counter on FIJI, where first the color channels were split and the syto9 channel was made into a maximum projection before being analyzed by 3D object counter. Biomass of the total bacteria in the field of view was quantified using Comstat 2.0 as in the biofilm analysis.

3.2.5 DNA sequencing and verification of GC

DNA of GC was extracted using the New England Biolabs Monarch® Genomic DNA Purification Kit. Purified DNA was then sent to Plasmidsaurus for DNA sequencing. Sequenced DNA was then analyzed on Geneious Prime to ensure expected mutations were present by analyzing deletions and SNPs of sequenced FA1090∆fil as compared to sequenced FA1090 WT. Each strain was sequenced twice. Link to sequence data can be found in the appendices at the end.

3.2.6 Statistical Analysis

Experiments were repeated as three independent trials and data was analyzed using GraphPad Prism 10 (GraphPad Software, La Jolla California USA). P-values were determined either by student's t-test when comparing only two groups or ANOVA when more than two groups are being analyzed.

3.3 Results

3.3.1 Genetic Composition of FA1090 Wild Type and FA1090∆fil GC

This entire study is designed to understand whether filamentous phage have any impacts on gonococcal biology. To begin to tackle this question, a strain of GC was designed lacking all the filamentous phage regions within its genome (Figures 16 and 17). FA1090 contains four areas of filamentous phage genes within its genome, all with high homology to each other. The location of these

regions and the order and orientation of these genes differ between strains of GC, but all GC strains in the NCBI database to date contain four regions containing filamentous phage genes. The locations and gene orders / orientations within FA1090, the GC strain used in this study, can be seen below (Figures 16A and 16B). To see what impacts these phage regions have on GC biology, I compared wild type bacteria to FA1090Δfil. FA1090Δfil contains no filamentous bacteriophage genes in its genome, with three of the four phage regions now containing an antibiotic resistance cassette (PHI7-9) or a deletion with nothing new added to it (PHI6) (Figure 17). The GC strain FA1090Δfil was constructed by Zhu Han in the Stein lab via PCR amplification of upstream and downstream regions of each phage region in FA1090 WT followed by ligation together (PHI6) or around an antibiotic cassette (PHI7-9) and transformed back into FA1090 to create deletion regions. These genotypes were confirmed via DNA sequencing done in duplicate (Figure 18). This sequencing data showed complete deletion of the phage regions with no other major changes to the genome. While some SNPs or other changes were seen in the sequencing data, all changes were either not consistent between duplicates which could suggest natural variation in a nucleotide, in non-coding regions of the genome, differences between sequenced data and NCBI database sequence of FA1090 WT, or in long stretches of repeated sequences (ex. GCGCGCGCGC) which are common areas prone to sequence assembly error (Figure 18).

3.3.2 Filamentous Bacteriophage are Required for Growth

The first question investigated in this study was whether there is an impact on growth by the presence of filamentous phage genes. A growth culture of both FA1090 WT and FA1090Δfil were grown at 37^{o}C with shaking to allow for oxygenation, as these conditions mimic the state in the FRT (Figure 19). When grown in these conditions, the data shows that these filamentous phage regions are absolutely required for growth at 37^{o}C with shaking, as removing the filamentous phage regions leads to a severe growth defect.

After discovering this strong growth defect in the FA1090Δfil mutants in these standard laboratory growth conditions, I next wanted to ask whether this defect was present or rescued in various conditions that are physiologically relevant to different GC infection environments. When looking at how temperature could affect this defect, the growth curve data suggests that this defect is not present when the

temperature is dropped to 34ºC (Figure 19). This temperature is more similar to the conditions found in the context of male urogenital infections.[78] GC biofilms often develop into an anaerobic environment, which is where high levels of filamentous phage are expressed.[79,80] Due to this, I also wanted to see how removing oxygenation can impact the growth of the FA1090Δfil mutant. In order to do this, flasks were grown in the same conditions at 37ºC, however they were grown statically rather than with shaking. When shaking is removed from growth conditions at 37ºC, the growth defect seen with shaking is rescued in static conditions in the FA1090Δfil mutant (Figure 20). Two factors change in static conditions compared to that of shaking conditions: the amount of shear stress and the availability of oxygen. To determine if the lack of shaking is causing anaerobic conditions in the flask to lead to the rescue of growth, $NaNO_2$ is added to the flasks in shaking conditions. As $NaNO_2$ is the preferred final electron accepter in the ETC in anaerobic conditions, adding this to aerobic growth conditions (shaking) will allow FA1090Δfil mutants to use $NaNO_2$ as its preferred electron accepter if using O_2 in this process is what is causing the growth defect.[81] When grown in aerobic conditions with shaking and $NaNO_2$, FA1090Δfil mutants were still unable to grow compared to FA1090 WT (Figure 20). These data suggest that filamentous phage in GC are required for growth at 37ºC with shaking, but lowering the temperature or removing shaking of the flask allows for this growth defect to be rescued.

3.3.3 All Four Regions of Filamentous Bacteriophage Genes not Required for Rescue of Growth

So far, this study has exclusively looked at the impact of filamentous bacteriophage with either all the filamentous phage regions present (FA1090 WT) or all four regions of filamentous phage regions deleted from the genome (FA1090Δfil mutants). The next question analyzed was whether all four regions of filamentous phage genes are required to rescue growth or if growth is rescued with only the presence of one, two, or three of these regions. In order to answer this, mutants lacking only some of the phage regions were created via spot transformation or antibiotic selection to see the effect of deleting specific phage regions. Growth rates were analyzed in strains with either one or two regions taken out of FA1090 WT (designated "Δ [phage region deleted]") or phage regions added back to FA1090Δfil mutants (designated "Δfil + [phage region inserted back into genome]") (Figure 21). The growth curve data suggests that all four filamentous phage gene regions are not required for wild type levels of growth (Figure 21). This data shows that when three or two phage regions are present, specifically PHI6 and

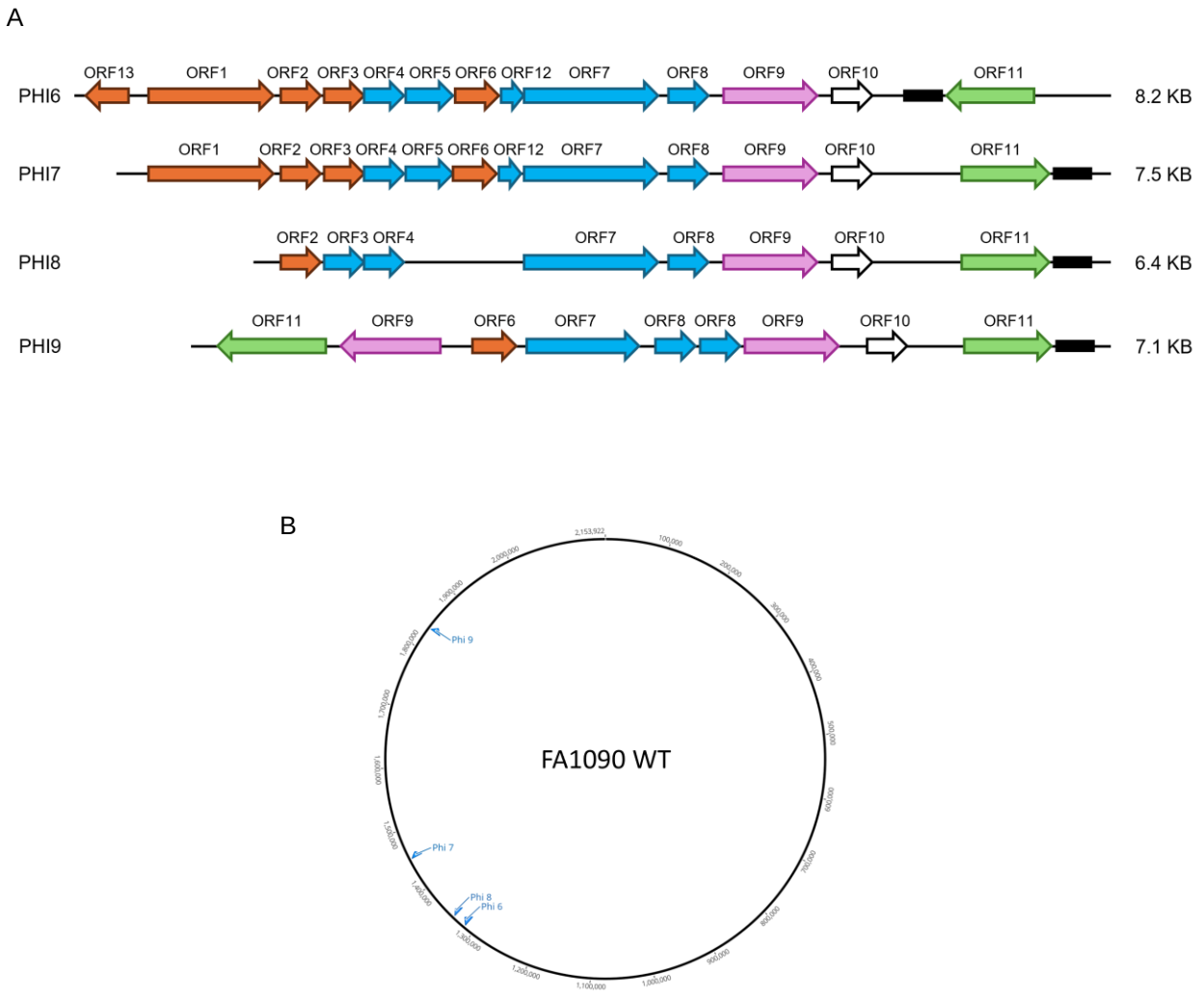


Figure 16. Filamentous prophage regions in FA1090 genome. (A) Length of prophage regions in genomes and order/orientation of genes in PHI6 through PHI9 in FA1090 as adapted from Piekarowicz et al.[52] Blue arrows represent structural genes. Orange arrows represent replication genes. Pink arrows represent secretion genes. Green arrows represent integrase genes (B) Location of filamentous prophage regions within the FA1090 genome.

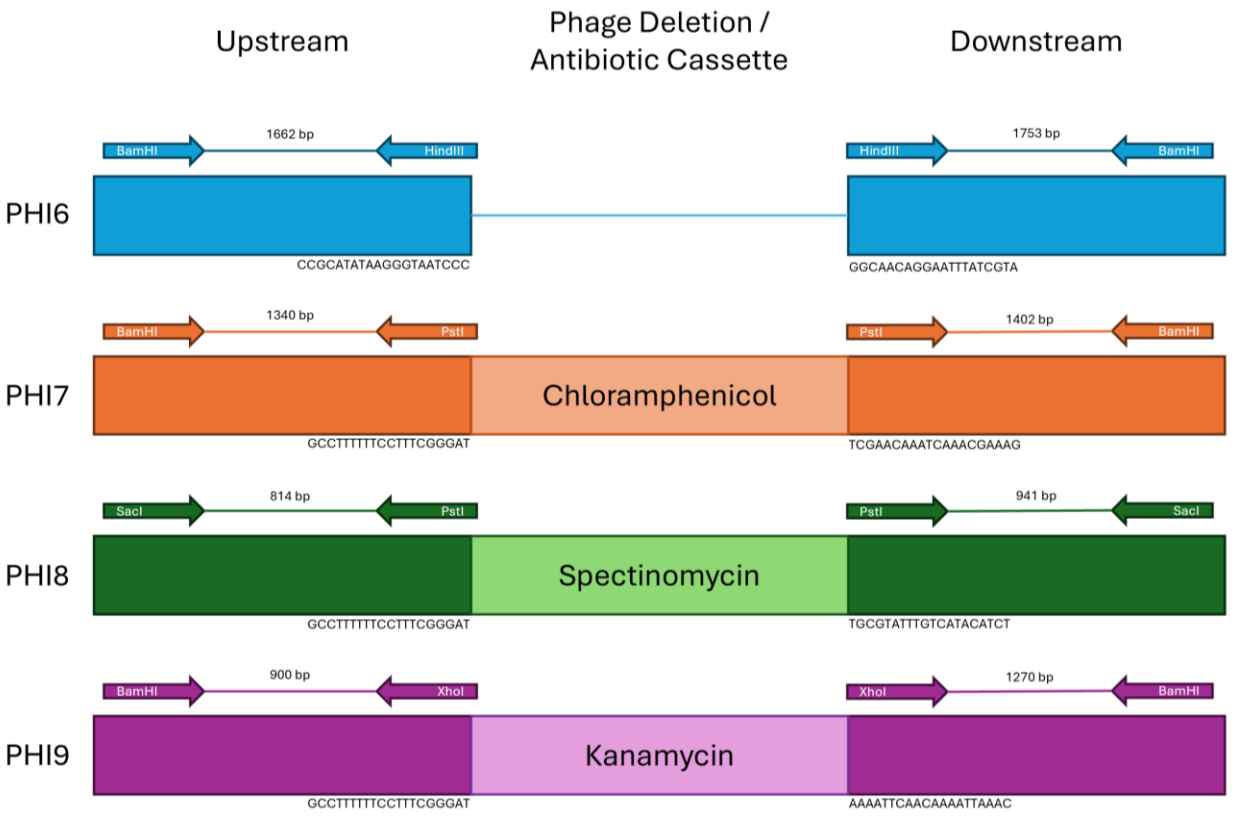


Figure 17. Construction of filamentous prophage region deletions in FA1090Δfil genome. Depiction of antibiotic resistance cassettes that replaced respective prophage regions in FA1090Δfil mutants, apart from PHI6 which has a deletion of the prophage region without an antibiotic resistance gene inserted in its place. Under the upstream and downstream region boxes are the 20 bases (on the same strand of DNA) flanking the deletion region. Primers at both upstream and downstream shown with their corresponding restriction enzyme and their distance from the last base of the primers facing the phage region to the last base before the deletion. These primers were used to amplify the regions flanking the corresponding phage regions and ligated together (PHI6) or around an antibiotic cassette (PHI7-9) and transformed into FA1090 to create deletion regions. These deletions were verified via DNA sequencing (n=2). Primer sequences can be found in the appendices at the end.

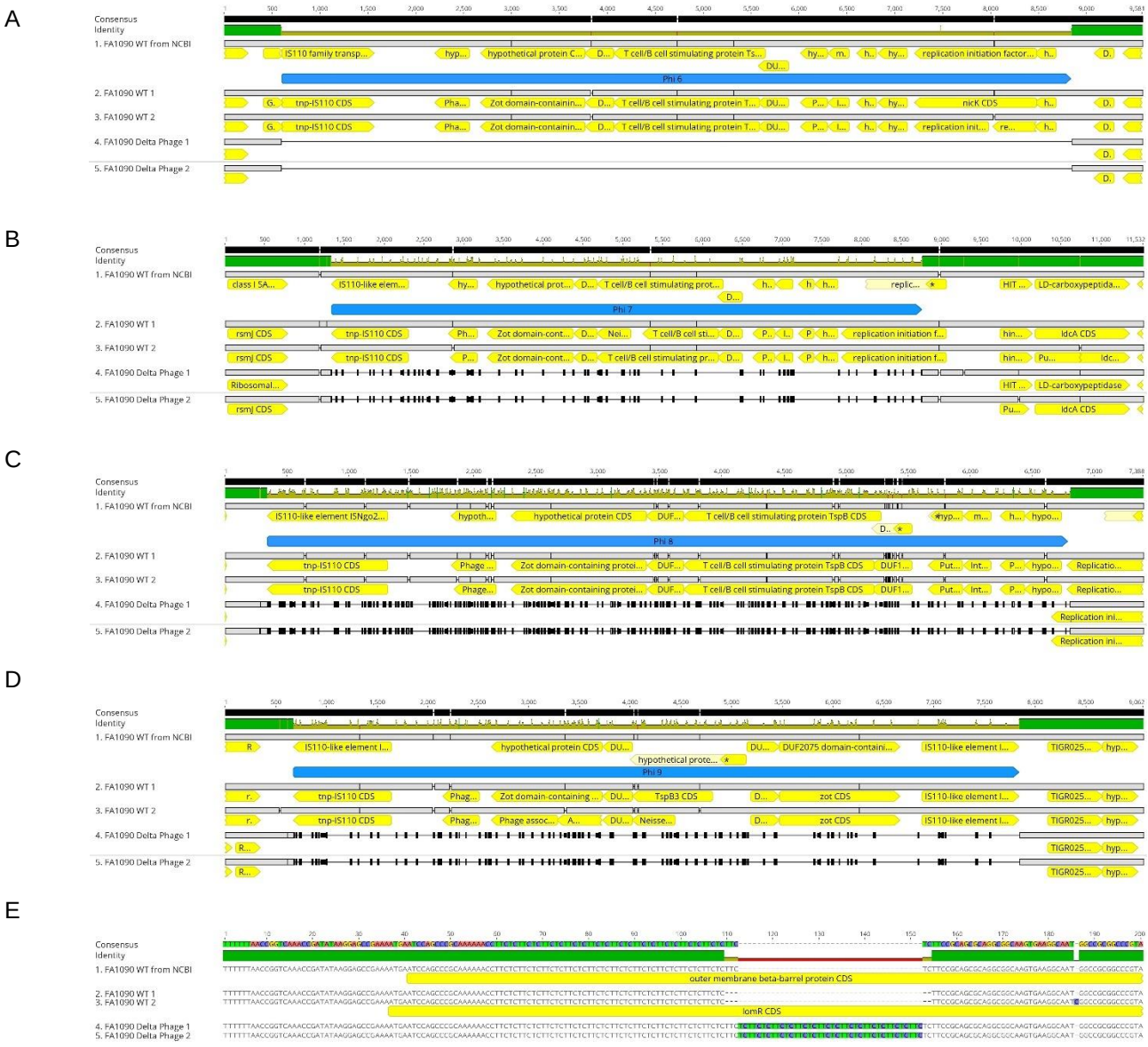


Figure 18. Sequence analysis of FA1090 WT and FA1090ΔfiI. (A-D) Sequence data of phage deletion regions (PHI6-9 in order) of FA1090 WT and FA1090ΔfiI in duplicate as compared to FA1090 WT NCBI sequence data. PHI6 displays a deletion with no sequence added between upstream and downstream regions while high levels of non-consensus sequence displayed in PHI7-9 deletion regions are antibiotic resistance cassettes. Blue arrows in FA1090 WT NCBI sequence indicated area of phage region. (E) An example of changes between FA1090 WT and FA1090ΔfiI sequences that are most likely sequence assembly error. The region showing low homology at bases 113-153 are in a string of CTTCT repeats which is a high density area for sequence assembly error and the single nucleotide insertion at base 187 that is not consistently found in both duplicates of FA1090 WT sequencing.

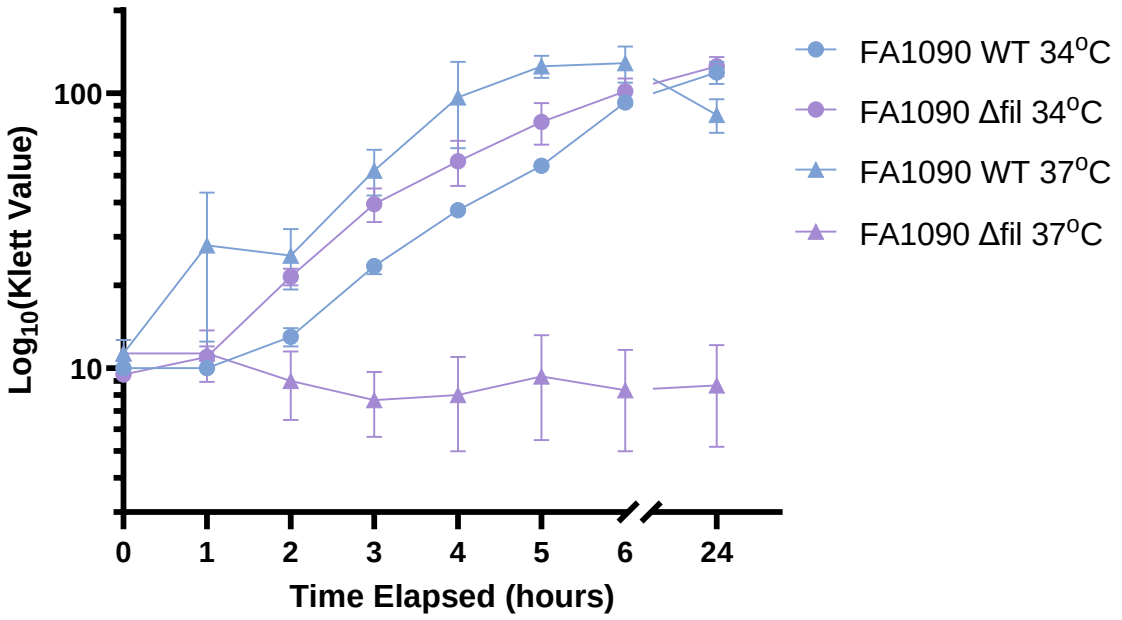


Figure 19. Filamentous phage required for growth at 37ºC. Growth curves over 24 hours at 34 or 37ºC with shaking for oxygenation. Error bars = SD. n=3

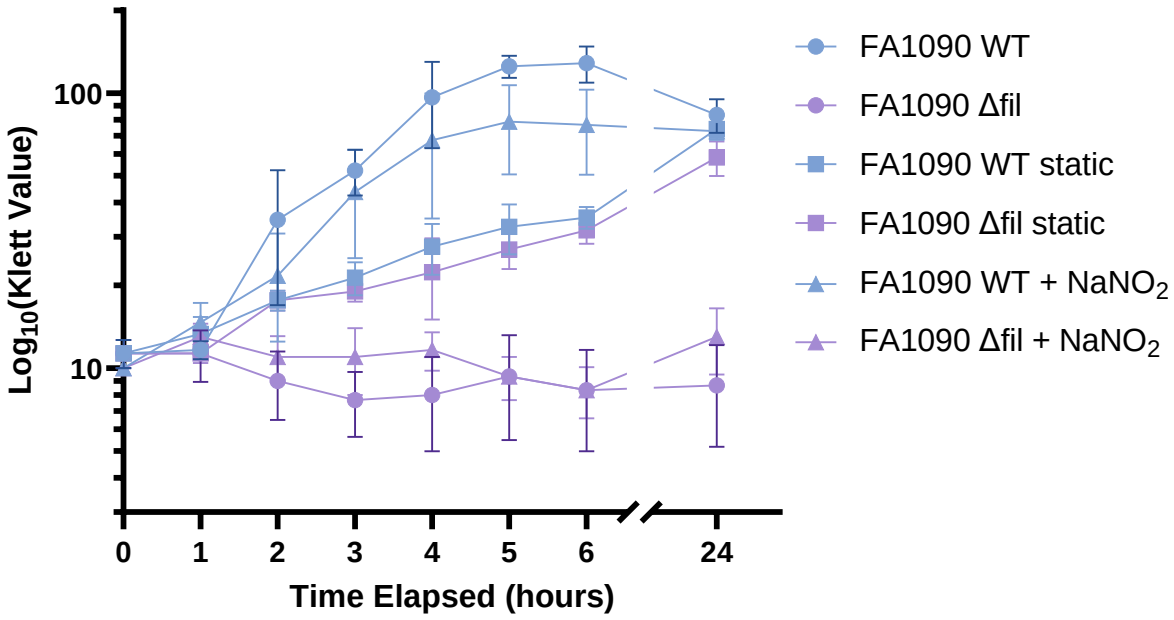


Figure 20. Growth defect in FA1090Δfil mutants rescued in static growth conditions. Growth curves over 24 hours. Strains grown shaking, in static conditions, or with shaking and supplemented 10mM $NaNO_2$ respectively. Error bars = SD. n=3

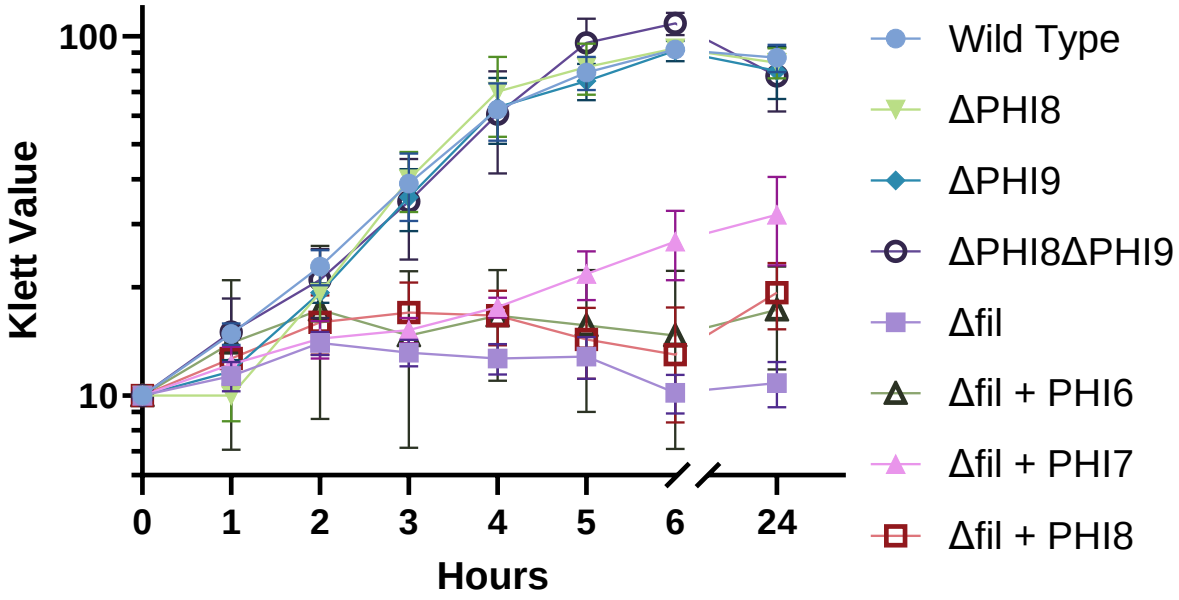


Figure 21. All four regions of filamentous phage regions are not required for WT levels of growth. Growth curves over 24 hours at 37°C with shaking using GC strains with varying amounts of filamentous phage gene regions. "WT – …" strains represent FA1090 WT strain that had deletions of indicated phage regions and "Δfil +…" represent FA1090Δfil strain that has phage regions added back in. Error bars = SD. n=3

PHI7, growth of the GC is the same FA1090 WT. When only the PHI7 phage region is present, growth is restored to an intermediate level of growth that is higher than the FA1090Δfil mutant, while not fully restored to wild type levels. This data shows that PHI8 and PHI9 are not required for growth of FA1090 GC, while PHI7 seems to have the largest impact on growth of all filamentous phage regions.

3.3.4 Filamentous Bacteriophage Promote Smoother and Shorter Biofilms

One key aspect to bacterial fitness and survival during infection is its ability to form biofilms, yet it is unknown how filamentous phage impact GC biofilms. The effect of filamentous phage on GC biofilm size and structure was investigated by quantifying three-dimensional confocal microscopy pictures using Comstat 2.0 in FIJI at either 34°C or 37°C for 6 or 24 hours (Figures 22-24). After 24 hours of growth, biomass of the biofilms at both temperatures were the same between FA1090 WT and FA1090Δfil (Figures 23 and 24). However, while the overall size of the biofilm didn't change, aspects of the structure did. FA1090Δfil had taller and rougher biofilms at 37°C and taller biofilms at 34°C by 24 hours (Figures 23 and 24). This supports that filamentous phage promote smoother and shorter biofilms in GC at a variety of physiological relevant temperatures.

3.3.5 Filamentous Bacteriophage inhibit gonococcal aggregation

After demonstrating that filamentous phage promote smoother and shorter biofilms, I hypothesized that the mechanism of this change might be due to filamentous phage promoting gonococcal aggregation. Bacterial aggregation has been seen to promote organization of the biofilm, as pulling on neighboring bacteria can lead to changes in biofilm structure.[82] Aggregation was studied similarly to the biofilm assays, using three-dimensional confocal microscopy image quantification in FIJI, but for a shorter amount of time and at a lower starting concentration, which allows for seeing individual aggregates before biofilm formation (Figure 25). This aggregation data shows that filamentous phage unexpectedly inhibit aggregation, with FA1090 WT producing significantly smaller aggregates than FA1090Δfil (Figure 25 and 26). Due to the role of pilin on aggregation, aggregation was studied using GC that were both expressing and not expressing pili (Figure 25 and 26, respectively). Pilin expression does not seem to impact the role of filamentous phage on aggregation inhibition, as FA1090 WT had smaller aggregates than FA1090Δfil, regardless of pilin expression. This increase in aggregation is not due to

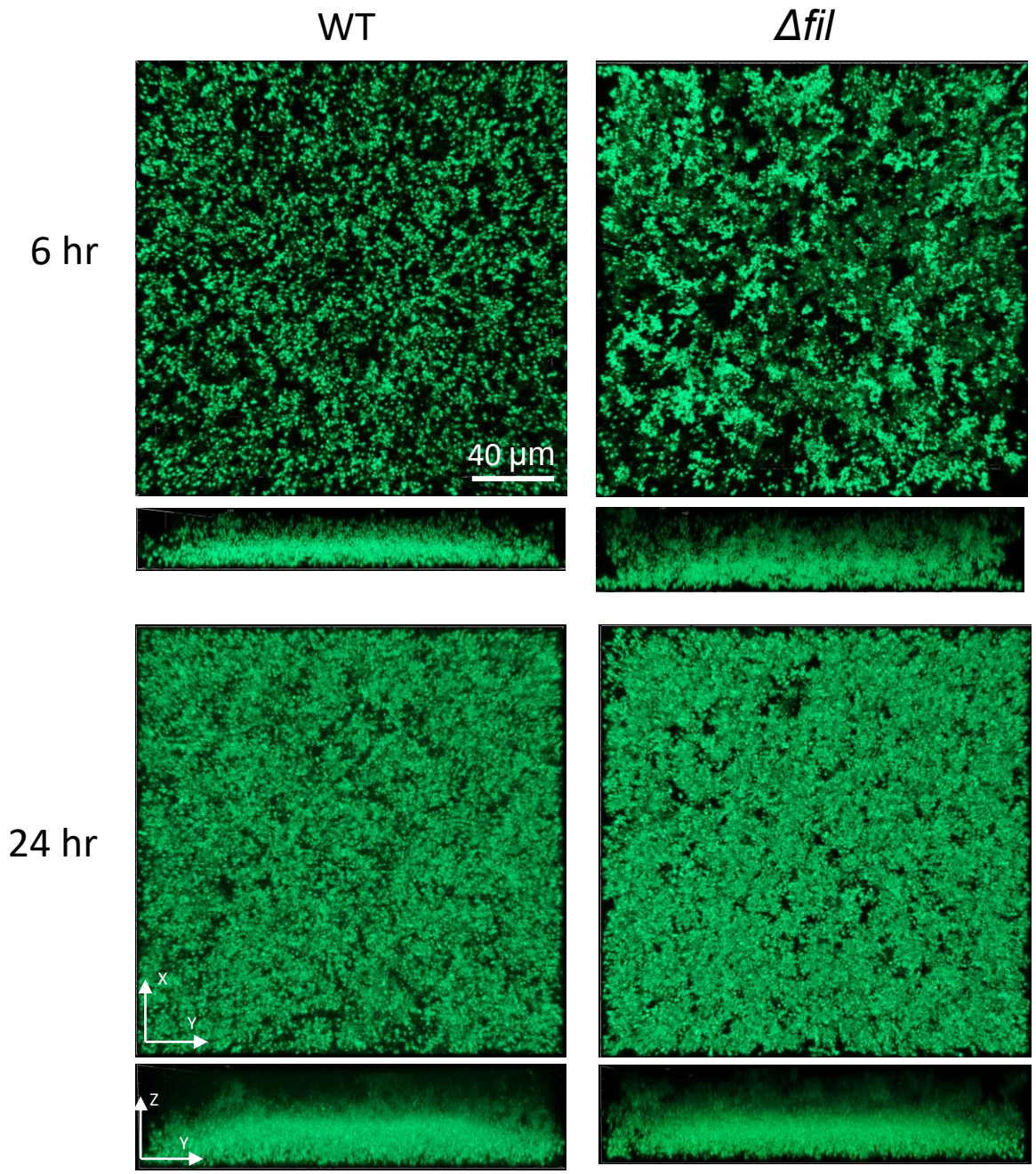


Figure 22. Representative images of biofilm assays at 37ºC at 6 and 24 hours. Representative maximum projection Z-stack images taken with 40x objective. Square images are of the XY plane and rectangular images below are rotated 90 degrees to the YZ plane.

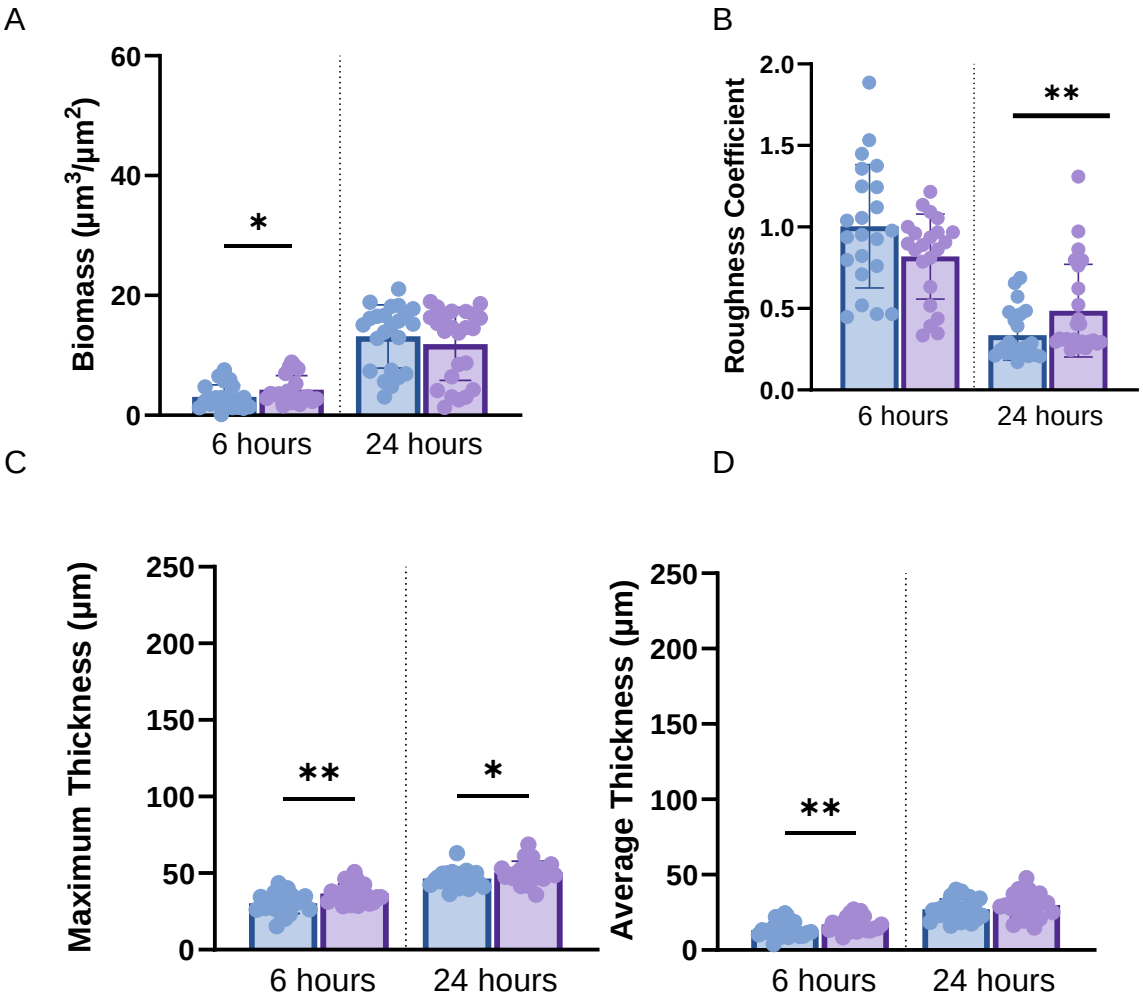


Figure 23. Filamentous phage promote shorter and smoother biofilms at 37°C. Quantitative analysis of biofilm assay z-stack images conducted at 37°C as analyzed by Comstat 2.0. Aspects of biofilm analyzed include total biomass of the biofilm (A), roughness of the top of the biofilm (B), The maximum height of the biofilm in each image (C), and the average height of the biofilm (D). Statistics only analyzed within same time points. * = p<0.05, ** = p<0.01, *** = <0.001, **** = p<0.0001. Error bars = SD, n=3, each replicate has 6-9 images analyzed

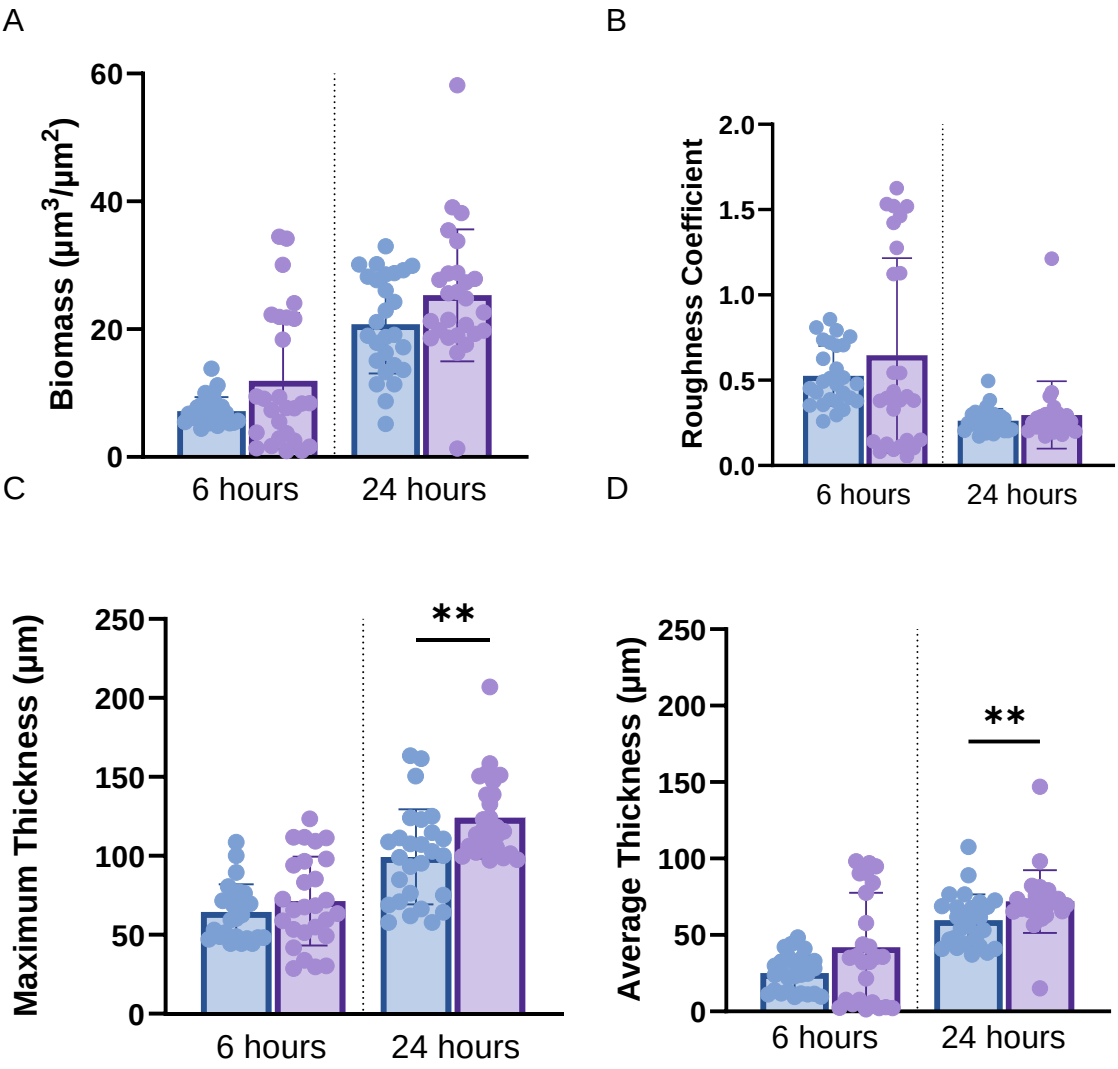


Figure 24. Filamentous phage promote shorter biofilms at 34°C. Quantitative analysis of biofilm assay z-stack images conducted at 37°C as analyzed by Comstat 2.0. Aspects of biofilm analyzed include total biomass of the biofilm (A), roughness of the top of the biofilm (B), The maximum height of the biofilm in each image (C), and the average height of the biofilm (D). Statistics only analyzed within same time points. * = p<0.05, ** = p<0.01, *** = <0.001, **** = p<0.0001. Error bars = SD, n=3, each replicate has 8-9 images analyzed. Each dot represents an individual image.

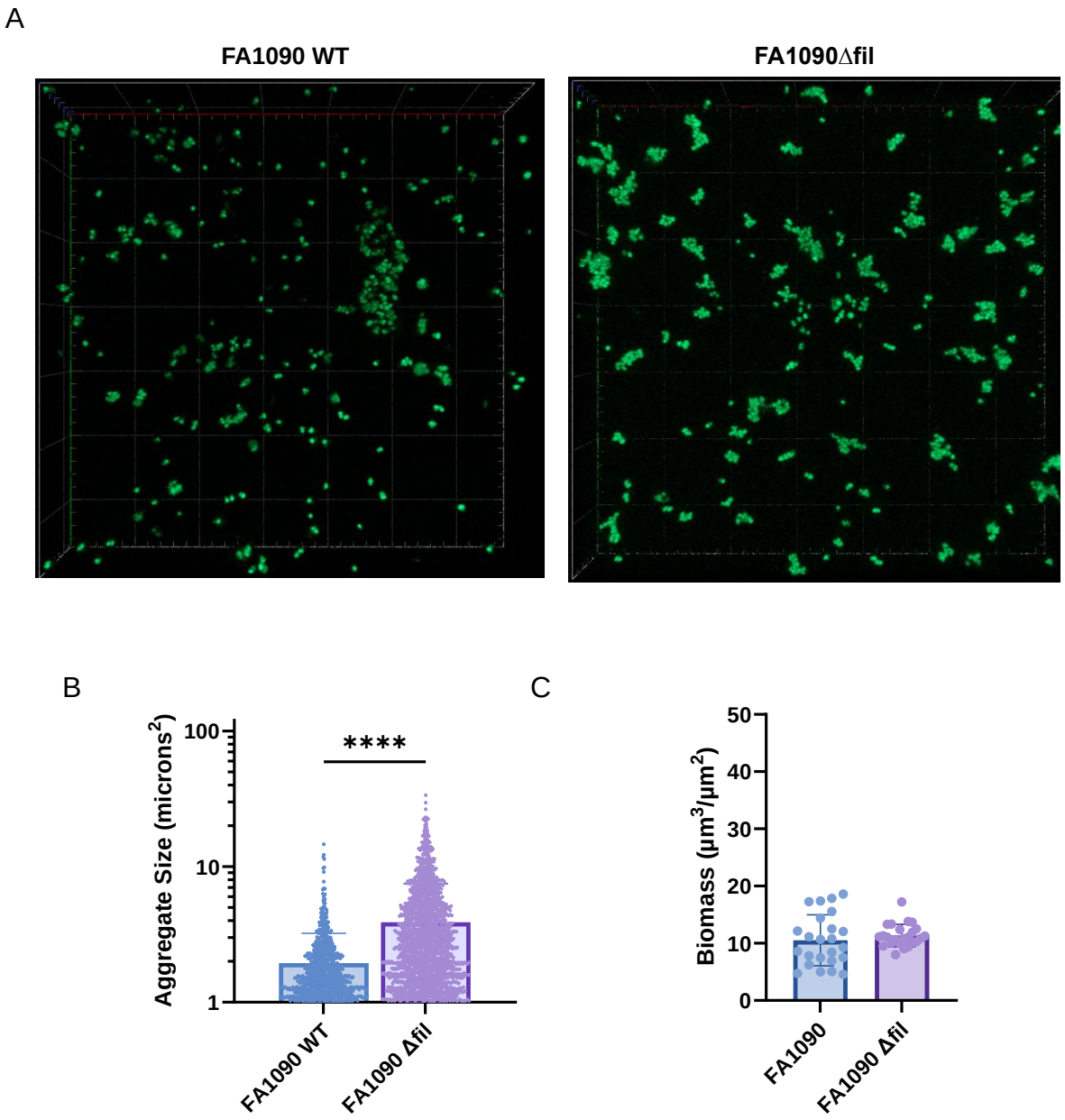


Figure 25. Filamentous phage inhibit bacterial aggregation in piliated GC. Quantitative analysis of aggregation assay z-stack images conducted at 37ºC using piliated GC as analyzed by FIJI / Comstat 2.0. (A) Representative images of FA1090 WT and FA1090Δfil. (B) Aggregate size with aggregates being defined as minimum of 4 bacteria and each dot representing an individual aggregate. C) Overall biomass of the image including aggregates and individual diplococci, with each dot representing an individual image. * = p<0.05, ** = p<0.01, *** = <0.001, **** = p<0.0001. Error bars = SD, n=3, each replicate has 6-9 images analyzed

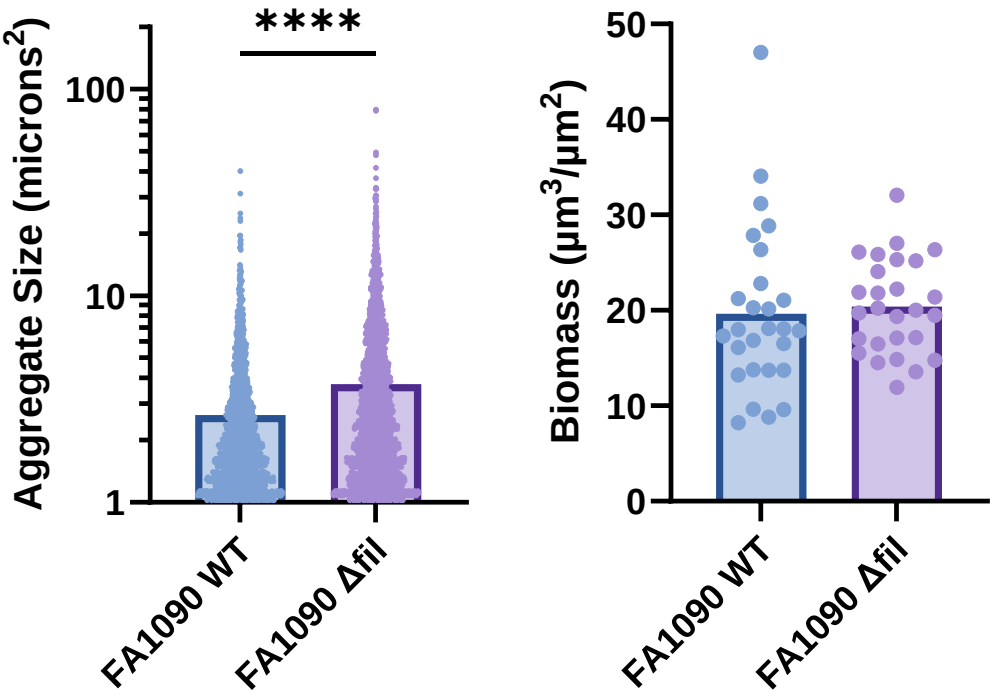


Figure 26. Filamentous phage inhibit bacterial aggregation in non-piliated GC. Quantitative analysis of aggregation assay z-stack images conducted at 37°C using non-piliated GC as analyzed by FIJI / Comstat 2.0. (A) Representative images of FA1090 WT and FA1090Δfil. (B) Aggregate size with aggregates being defined as minimum of 4 bacteria and each dot representing an individual aggregate. C) Overall biomass of the image including aggregates and individual diplococci, with each dot representing an individual image. * = p<0.05, ** = p<0.01, *** = <0.001, **** = p<0.0001. Error bars = SD, n=3, each replicate has 9 images analyzed

more bacteria in the field of view, as biomass is not significantly different between FA1090 WT or FA1090Δfil (Figure 25 and 26). This shows that biofilm structural differences is not due to stronger bacterial aggregation, as filamentous phage inhibit bacterial aggregation

3.4 Discussion

This study sought to elucidate whether filamentous bacteriophage had any impact on gonococcal biology. The data in this chapter clearly illustrates the profound impact filamentous phage have on GC biology, both on general growth fitness and in the context of biofilm and aggregate production. Filamentous phage is absolutely required for growth of GC at 37°C with shaking conditions, as removing the filamentous phage regions from the genome led to little growth in the FA1090Δfil mutant (Figures 20 and 21). Impacts at this temperature are important, as the FRT is around 37°C, an important site of infection.[83] Filamentous phage's impact on growth was also studied in a variety of conditions due to the wide variability in GC infection environments. While the FRT is not under constant shear stress, it is not completely static either. The ability to grow at body temperature with some level of mechanical stress is vital for GC survival in the human body. The FRT also has varying levels of oxygen availability, depending on the stage in the menstrual cycle.[83] It is important to keep in mind that GC infection has a wide variation of clinical outcomes. Part of that variation is due to the fact that GC must infect both the complicated geography of the FRT, containing multiple cell morphologies, a complex microbiota community, and varying oxygen and pH levels throughout the menstrual cycle, and the completely different structure of the male urogenital tract.[1,6-8,83] While the FRT is around 37°C, the male urogenital tract is often cooler, at around 34°C.[78] Unlike the FRT, the male urogenital tract is subjected to high levels of shear stress during urination.[84] GC can also have local infections in other parts of the human body, including the pharynx, eyes, and rectum.[1,5] Each of these locations comes with its own individual stressors, ranging from variations in temperature, shear stress, pH, oxygen availability, host cell structure, and local immune responses. The growth defect seen in FA1090Δfil being present in only specific temperatures and levels of shear stress might indicate that filamentous phage are vital to growth during infection, but only in specific infection environments.

It was demonstrated that not all four regions of the filamentous phage genes are required to see rescue to wild type levels of growth, as having two or three phage regions allowed for growth to be equal to wild type levels (Figure 21). However, the molecular mechanism of filamentous phage growth advantages at specific conditions has not been investigated during this study. One hypothesis is that one or more filamentous phage genes are impacting the ability of GC to grow at specific conditions and when this gene is removed, GC can no longer grow at these conditions. In *Vibrio parahaemolyticus*, one lytic phage gene was seen to impact both bacterial growth and bacterial membrane integrity, with growth defects varying with temperature.[85] While this phage gene is from a lytic phage and not a filamentous phage, it is possible that filamentous phage proteins are capable of similar effects. FA1090's filamentous phage have multiple genes of unknown function including ORFs 2, 3, and 12, with ORF12 containing a transmembrane domain.[52-54,56] It is possible that one of these proteins is conferring some advantage to growth at 37^{o}C. If ORF12 is responsible for this phenotype, it could explain why PHI8 and PHI9 are not required for growth in these conditions, as they lack ORF12 in their genomes, but only restoring PHI7 is not enough to restore to WT levels of ORF12 expression (Figure 17). Since ORF12 has a transmembrane domain it is possible that it could be impacting bacterial membrane integrity as in *V. parahaemolyticus* phage.[85] Temperature and shear stress have been seen to impact membrane integrity, which could explain why growing GC in lower temperatures or static conditions restores growth to FA1090Δfil.[84,86-88] If these phage are causing changes to membrane integrity, and therefore growth of the GC, this would explain why adding $NaNO_2$ to a shaking flask does not restore growth, as lower mechanical stress rather that lower oxygen availability is driving the rescue of growth in FA1090Δfil. This hypothesis can be tested by reintroducing these individual phage genes back into the genome, rather than whole phage regions, and seeing whether growth is rescued compared to FA1090Δfil.

A second possibility is that the phage regions are impacting the expression of neighboring genes in the bacterial genome, as PHI6-9 have possible gene promoters pointing out towards the surrounding bacterial genome.[52] Two phage regions have flanking genes immediately adjacent to the phage regions that are oriented in the correct direction in order to be influenced by the phage region promoters. One is a TIGR02594 family protein with unknown function near PHI9. PHI7 however has a LD-carboxypeptidase and a Histidine triad (HIT) domain containing protein, as predicted by NCBI sequence data (Figure 18).

HIT domain proteins are found in both bacterial and eukaryotic cells and have been linked to a variety of processes such as galactose metabolism, DNA repair, and tumor suppression.[89] LD-carboxypeptidase is involved in peptidoglycan recycling and is essential in *E. coli* for survival in stationary phase.[90,91] It is possible that deleting the PHI7 phage region heavily impacts one or both of these gene's expression levels, which could in turn impact GC's ability to grow, as all strains with PHI7 can grow. FA1090Δfil + PHI7 may have intermediate growth if reinserting the phage regions was not identical at the insertion site, leading to less expression than WT levels. To test whether these proteins are responsible for FA1090Δfil's growth defect, adding LD-carboxypeptidase and/or the HIT domain protein back in on a plasmid with an inducible or constitutive promoter will determine if the growth defect is from phage induced expression of these bacterial genes rather than phage production / phage proteins. A second approach would be reinserting PHI7 somewhere else in the bacterial genome to see the impact on GC growth.

A third possibility is that a protein regulating phage expression can bind elsewhere in the genome. Many phage encoded genes have the capability of binding to DNA, and therefore can effect transcription of genes, both within the bacterial and phage genomes.[52,53] Many ORFs contained in the phage have unknown functions, have predicted DNA binding domains, or share homology with filamentous phage genes capable of DNA rearrangements (often DNA excising) and transcriptional regulators.[52] One example is ORF6, which contains a predicted zinc finger domain.[52] Zinc finger domain containing proteins often bind to DNA and RNA and are involved in a variety of processes, such as transcription and replication regulation, protein degradation, and DNA repair.[92] ORF1 has high homology to the protein RstA, a protein required for replication of phage genomes in CTXΦ. RstA's repressor, RstR, doesn't seem to have high homology to any protein in GC, however ORF13 is located in the same place and in the same orientation (opposite to the rest of the genes) in PHI6 as RstR is in CTXΦ.[52,65] ORF11 encodes a predicted transposase gene, capable of binding to and moving DNA in and out of the bacterial genome.[52] ORF2 has some homology to RstB, a single stranded DNA binding protein involved in phage replication.[52,65] With all of these genes in the phage genome able to bind to DNA/RNA and alter gene expression, it is possible that these proteins are acting on genes outside of the phage regions. If one or more of these genes is altering gene expression for proteins required for growth at 37°C, either through repression or promotion of expression, removing this protein could be responsible for the growth defect in

FA1090Δfil. If the bacterial genes impacted by phage transcription regulators are only required for growth in 37°C or with shaking, this could explain why growth defects are only seen under certain conditions. Many GC genes are differentially regulated in different infection contexts, especially between male and female infections.[93] Transcriptomic data found that filamentous phage genes are highly enriched in female GC infections when compared to male urogenital infections.[93] With FRT infections being at a higher temperature than male urogenital infections (37°C vs 34°C respectively), this could highlight the importance of filamentous phage's role in environmental specific gene regulation, specifically at 37°C.[78,83] If a phage gene can impact global gene expression of bacterial genes important to FRT infection, rather than male urogenital infection, this could point to why clinical outcomes for male and female infections have such wide variation. Gene expression studies such as RNA-seq or proteomic studies of FA1090 WT and FA1090Δfil at 37°C and 34°C could help elucidate which bacterial genes are being differentially expressed between varying temperature conditions in the presence or absence of phage genes. To further investigate which phage gene or genes could be changing bacterial gene expression, reintroducing individual phage genes and repeating RNA-seq studies could point to a specific mechanism of temperature dependent gene expression regulation.

Another aspect of environmental impacts on the growth of FA1090Δfil is oxygen availability. This growth defect was not only temperature dependent, but also dependent on the shaking of the growth media. Shaking of the flask changes two aspects of the growth environment, shear stress applied to the bacteria and the availability of oxygen. Discussed above in this section is how shear stress on the bacteria can impact bacterial growth if the membrane is compromised due to missing phage surface proteins. However, it is possible that the shaking or lack of shaking of the flask changed the effect of the phage deletion on growth through changing the availability on oxygen. O_2 can directly regulate genes in GC on a transcriptional level, including the gene *aniA*, which encodes a nitrite reductase.[94,95] It is possible that a phage gene that is impacting gene expression of bacterial genes is only important in the context of aerobic growth or could itself be regulated by oxygen availability. While the FRT is not anaerobic in nature, it does have changing oxygen availability throughout the menstrual cycle and over time GC biofilms will become anaerobic, which could explain why GC biofilms in static conditions have the same biomass regardless of the presence of filamentous phage.[79,83] While $NaNO_2$ is the preferred electron

acceptor of the ETC in anaerobic conditions, it is possible that simply supplementing the media with $NaNO_2$ with shaking would not allow anaerobic gene expression to turn on / aerobic gene expression to turn off due to the presence of oxygenation.[94,96] Further studies investigating the role of oxygen versus shear stress on the FA1090ΔfiI growth defect could help tease apart why FA1090ΔfiI can grow in static conditions. One possibility is growing GC in anaerobic conditions with shaking, to see if losing oxygen in the media, but retaining shear stress, would allow FA1090ΔfiI to grow. If FA1090ΔfiI is able to grow in those conditions, it would point to oxygen presence impacting gene expression. However, if FA1090ΔfiI cannot grow, it could point to phage proteins impacting membrane integrity of GC.

Biofilm production aids bacteria in combatting many stressors during infections including protection from mechanical force, increased immune evasion, increased AMR of cationic antibiotics, decreased water loss, and more.[24,60,61,71] Here I have shown that filamentous phage in GC promote shorter and smoother biofilms (Figures 22-24). The same phenomenon was seen in *P. aeruginosa*, as their filamentous phage promoted organization of the biofilm.[60,61,71] Having tighter and smoother biofilms can contribute to biofilm stability and protection, as filamentous phage in *P. aeruginosa* biofilms led to less water loss and higher AMR levels than biofilms lacking their phage.[60,61,71] Filamentous phage expression goes up in anaerobic conditions.[79,80] GC biofilms develop into an anaerobic environment over time, which could be why some of these effects are only seen after 24 hours of growth and perhaps why MDAΦ production in *N. meningitidis* biofilms is only seen further into biofilm development.[23,79] Because the data showed that FA1090 WT biofilms are smoother and shorter while having the same amount of bacteria, our hypothesis was that filamentous phage are promoting GC-GC interactions, which would lead to larger bacterial aggregates. However, GC filamentous phage were shown to have inhibited aggregation, opposite to what I first hypothesized the mechanism of shorter and smoother biofilms might be (Figures 25 and 26). It is possible that filamentous phage are actively inhibiting aggregation by pushing against their neighboring bacteria as they are extruded through the membrane, leading to disassociation of bacterial aggregates regardless of the presence of pilin. Piliated GC had nearly a two-fold increase in aggregate size in the FA1090ΔfiI mutant, highlighting the large impact of filamentous phage on aggregation. In *P. aeruginosa*, they saw that their smoother biofilms were due to phage, not causing stronger bacterial aggregation, but filamentous phage organizing the extracellular matrix of the biofilm via

depletion forces.[60,61,71] Not only were their filamentous phage impacting biofilm organization, but they could be seen to form organized patterns in solutions with no bacteria present, suggesting filamentous phage can impact the structure of a biofilm in the extracellular matrix, independent of their interactions with bacteria.[60,61,71] As these filamentous phage have a highly similar structure to the phage found in GC, it is possible that the change in biofilm structure is due to these same depletion forces. If phage in the extracellular matrix is causing changes to the biofilm rather than changing some property in the bacteria, adding exogenous purified filamentous phage to FA1090Δfil biofilms should result in FA1090Δfil biofilms returning to WT structures. Studies done in *N. meningitidis* showed that the presence of their homologous filamentous phage produces much larger biofilms, however their experiments were done under flow conditions.[23,57] It is possible that these phage induced smoother biofilms are more resistant to the addition of shear forces on the biofilm, resulting in smaller biofilms in strains without filamentous phage due to bacterial loss in the current of the flowing media. As seen in Figure 19, FA1090Δfil has a severe growth defect when in shaking conditions at 37^{o}C. It is possible that *N. meningitidis* strains lacking MDAΦ have the same defect, as these are highly related species and phage, which may explain why bacteria grow less in biofilms under flow conditions at this temperature. The impact of filamentous phage induced smoother and shorter biofilms to the persistence of GC infection is still unknown. Doing various studies on the fitness of the biofilms by measuring AMR, water loss, and shear stress resistance levels in FA1090Δfil and FA1090 WT will show the impact of filamentous phage on the GC biofilm.

So far in this section, I have discussed how each of the results obtained in this study could be explained on a molecular level. However, it is important to understand how all of this comes together to truly understand the role of filamentous phage during GC infection. When GC enters the FRT, it must first adhere to the host's cells.[5,13,14] Once GC has successfully attached to the host, it can begin colonizing the ECC.[5,13,14] Based on the above research, it is at this early step in GC infection that we can see the impact of filamentous phage on this process, as I have demonstrated that filamentous phage are required for growth at the conditions found within the FRT (aerobic and 37^{o}C) (Figure 19). After initial colonization of the ECC, GC can eventually create a biofilm.[22] Previous research has shown that filamentous bacteriophage proteins are upregulated in both anaerobic environments, such as in GC biofilms, and during female infections.[79,80,93] This leads to filamentous bacteriophage expression rising as biofilms are

produced, which creates shorter and smoother biofilms along the ECC during asymptomatic infections as shown in figures 22-24. In *N. meningitidis* filamentous phage research, *Neisseria* biofilms containing filamentous phage have pilin surface expression in bacteria closest to the host cells, while the outside of the biofilm contains filamentous phage surface expression.[23,57,64] As biofilms develop and filamentous phage expression rises, it is possible that the filamentous phage-producing GC are on the outside of the biofilm as in *N. meningitidis*. If this is the case in GC biofilms, this could impact how GC are eventually able to leave the biofilm. For GC to disseminate to other locations in the body and transmigrate across the TZ and ENC, the bacteria must be able to leave the biofilm.[7,8,12,22,97,98] My research in figures 25 and 26 shows that filamentous phage are inhibiting bacterial aggregation. It is possible that this inhibition allows filamentous phage producing GC on the outside of the biofilm to push away from their bacterial neighbors in order to leave the biofilm. Once these filamentous phage-producing GC leave the biofilm and migrate to the ENC, these GC can expose the host epithelium to their ORF9 surface proteins to disrupt the integrity of the epithelium as shown in figure 5, which could lead to symptomatic disseminated gonococcal infections.

Taken en toto, this chapter highlights the active role that filamentous phage play in gonococcal biology and pathogenicity. While this chapter only looks at a handful of processes, filamentous phage may play a role in countless bacterial processes in GC that have yet to be studied. This discussion hypothesizes how the four filamentous phage gene regions PHI6-9 can impact GC on both a molecular level and during GC infection based on the results of this chapter and previously published research. However, further studies are needed to elucidate the mechanism of these impacts and whether these impacts are also occurring in the context of in vivo infections. Filamentous phage are an important topic to study within GC as they are ubiquitous in GC strains and share high levels of homology between strains, and therefore offer the potential for drug targeting as a means of treating or preventing gonococcal infection. Effectively treating and preventing GC infection via means other than antibiotic based methods is extremely important during a rapidly increasing rate of AMR in GC strains, as no current preventative vaccine to GC exists.[1,99] Previous research has shown the potential of filamentous phage as a GC treatment/prevention method, as when rabbits were fed purified GC filamentous phage, bactericidal antibodies to GC proteins developed in the rabbits.[55] This is due to bacterial surface proteins sticking to

the outside of the filamentous phage as they are extrude through the surface of the bacteria.[47,55] If this could be further developed in humans as a means of targeting GC infection in the body, it could revolutionize GC front line treatment. Understanding the mechanism that filamentous phage play in GC biology and pathogenicity is crucial in order to further develop the potential for filamentous phage to be a target of GC prevention and treatment.

Chapter 4: General discussion and future work

4.1 General summary

This thesis aimed to address two primary questions: the role of genome reorganization and the influence of filamentous bacteriophage on gonococcal biology. To explore the first question, this study examined various strains to understand how genome reorganization occurs and its impact on gene expression. Our findings indicated that, based on a review of all complete GC genomes in the NCBI database up to this point, each strain exhibited a distinct genome organization, including both inversions and translocations. I also observed a consistent pattern of DNA sequence homology flanking these rearrangements, suggesting that homologous recombination drives this genomic reorganization. PCR analysis of inversion events indicated that these large-scale changes are both stable and infrequent. Growth curve analyses revealed significant variability in growth rates. Additionally, proteomic analysis suggested that changes in replication termination locations due to genome reorganization led to the upregulation of various proteins, including those associated with energy metabolism and antimicrobial resistance (AMR). Overall, this study implies that homologous recombination-driven genome reorganization significantly influences gonococcal biology and pathogenicity. While it is still unclear whether altering AMR-associated gene location or changing overall genome structure can directly cause higher levels of AMR, this study clearly illustrates the necessity of examining phenotypic changes arising from factors beyond simple allelic differences. With the rising threat of AMR within GC strains today, this study demonstrates the need to keep the strains studied in research updated and relevant to the state of GC actively being spread in the world.

To address the second question, I investigated how filamentous phage in GC can impact gonococcal biology. Previous unpublished research in the Stein Lab demonstrates that filamentous phage derived proteins are capable of deteriorating the integrity of epithelial cells in cell cultures and cervical tissue explants (Figure 5). However, this is the first study demonstrating the impact of endogenous filamentous phage on GC biology. I found that FA1090Δfil grew poorly at 37^{o}C in both broth and on agar, as compared to FA1090. However, there was no difference when the strains were grown at 34^{o}C or when grown without shaking, demonstrating the condition dependent nature of this growth advantage. When these strains were analyzed for their ability to produce biofilms, no differences were seen in the overall

biofilm's biomass, yet the overall structure of the biofilms were different, with FA1090Δfil producing taller and rougher biofilms. FA1090Δfil formed larger bacterial aggregates than FA1090 WT. Taken en toto, this chapter highlights the active role that filamentous phage play in gonococcal biology.

Both of these chapters highlight the importance of non-allelic differences in bacterial genes driving changes in gonococcal biology and pathogenicity. Both alterations in bacterial genome organization and contributions from bacteriophage genomes can affect gonococcal biology independently of changes in bacterial gene sequences. With the rapid rise in AMR rates, these studies are essential for identifying new strategies to combat AMR in GC. It is critical to explore treatment and prevention methods for GC infections that do not rely solely on antibiotics, especially given the alarming increase in AMR among GC strains. Previous research has shown the potential of filamentous phage as a GC treatment/prevention method.[55] Both chapters above are key to unlocking new roads to combatting AMR levels in GC through research looking at these new aspects of GC increasing its pathogenicity, both through genomic reorganization and filamentous phage expression.

4.2 Working Model

Complications during GC infection occur when GC is able to switch from a localized asymptomatic infection on the ECC to a symptomatic infection due to transmigration of GC in the TZ and ENC. However, how GC switches from an asymptomatic and localized infection to a symptomatic and invasive infection is still not completely understood. This working model aims to hypothesize how chapter three may fit into this puzzle (Figure 27). Chapter three highlights changes in growth, aggregation, and biofilm development of GC due to the presence of filamentous bacteriophage gene regions in its genome. Based on the research conducted in chapter three and previous research in a variety of bacterial species, the following working model has been hypothesized. GC adheres to and begins to colonize to the ECC inside of the FRT. In *N. meningitidis* filamentous phage research, *Neisseria* biofilms containing filamentous phage have pilin surface expression in bacteria closest to the host cells, while the outside of the biofilm contains filamentous phage protein surface expression.[23,57,64] My data shows that filamentous bacteriophage aid in producing shorter and smoother biofilms. This phenomenon seen in other species

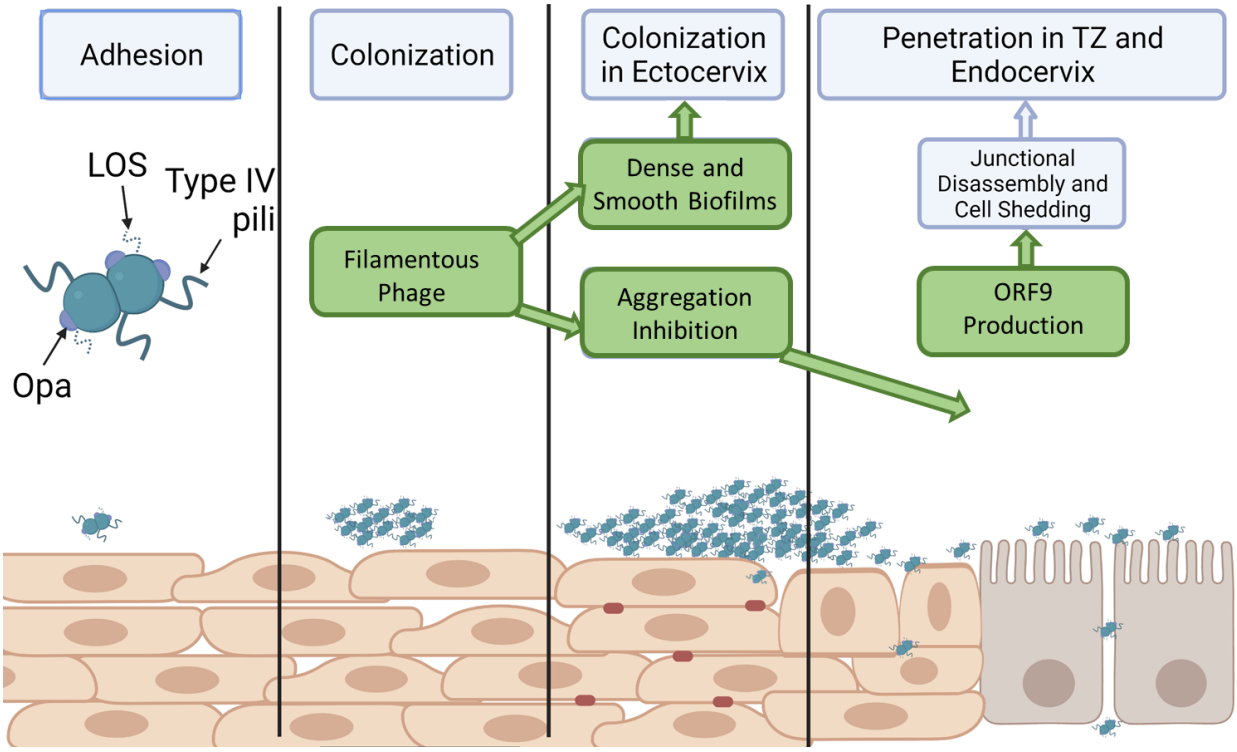


Figure 27. Hypothesized model of filamentous phage's role in GC infection progression. Our data suggests that filamentous phage promote smoother and shorter biofilms, which could increase fitness of the biofilm against environmental stressors. Our data also suggests that filamentous phage inhibits aggregation. Previous research using *N. meningitidis* suggests that only bacteria on the outer portion of the biofilm are actively producing filamentous phage.[23] If these phage producing bacteria escape the biofilm due to their inhibited bacterial aggregation abilities and make it to the ENC, they would then be expressing ORF9 at the ENC, which has been shown to increase cell death and junctional disassembly, and therefore promote bacterial transmigration into deeper tissue. Overall, the data of this chapter along with previous research suggests that filamentous phage have a profound impact on GC biology and could therefore impact pathogenicity of GC in vivo.

led to higher AMR and less water loss due to the organization of the biofilm by filamentous phage, which can lead to higher bacterial persistance.[60-63] However, my data also shows that GC containing filamentous phage aggregate worse. It is possible that filamentous phage are helping to produce stronger biofilms on host cells, but the GC actively producing phage on the outside of the biofilm struggle hold onto neighboring GC cells due to lower aggregation abilities. It is possible that these phage-producing bacteria are the bacteria that are migrating from the ECC to the ENC. Previous research has also demonstrated that GC filamentous phage derived protein, ORF9, can compromise the integrity of host cell structure. ORF9 is a surface protein required for the extrusion of filamentous phage out of GC cells. The GC that have now escaped the biofilm and moved to the ENC are actively producing phage and therefore have ORF9 on its surface. This model could demonstrate the possibility that lowering the GC's ability to aggregate, and therefore gain the ability to leave the biofilm due to phage production, is also the key for the GC to lower the integrity of host cell structures in order to transmigrate and cause invasive gonococcal infections.

4.3 Future Directions

4.3.1 Future directions of genome reorganization study

While the data from this chapter shows that a strain with growth defects and a shifted replication termination sequence also has changes in gene expression of key bacterial fitness genes, this study is limited to showing that these changes are associated with genome structure, rather than showing changing genome structure causes changes in phenotypes. In order to demonstrate that changes in overall genome structure can impact gene expression and therefore bacterial fitness, future experiments should involve moving the terminal sequence to different locations within the genome in one strain to directly show that changing the replication termination sequence location can cause changes in phenotype without directly mutating genes.

This study highlights that many types of proteins are upregulated in WHO L, such as AMR, replication, translation, and energy metabolism proteins. However, this upregulation is hypothesized to be do to increases in gene dosage due to high levels of replication coupled with slow growth. Future studies could attempt to see if gene dosage is directly causing this higher protein expression. This could be done by transforming plasmids containing multiple copies of a gene under the same promoter found in the

bacterial chromosome into GC and measuring gene expression. This would also show if increasing protein expression via higher gene dosage can directly change phenotypes in GC in important processes such as AMR capabilities.

The expected replication termination sequence in WHO L is located about a quarter of the way through its genome (Table 2). However, this strain may be using a non-canonical termination sequence, which could explain the growth defect observed in WHO L (Figure 13). To determine if this is the case, analyzing the peak-to-trough ratios from individual whole genome sequencing reads could reveal the actual starting and termination points of replication in WHO L. When examining WHO L for the canonical origin and terminus of replication, only one of each was found in the genome. Peak-to-trough analysis could help identify alternative termination sites without requiring prior knowledge of their possible sequences.

4.3.2 Future directions of filamentous bacteriophage study

This study demonstrated the requirement for filamentous bacteriophage during growth of GC at 37^{o}C with shaking. This study also showed that the presence of filamentous bacteriophage genes changed the structure of biofilms and aggregation abilities of GC in vitro. While the ability to grow at body temperature and produce biofilms in the context of infection is crucial, they are important in localized and asymptomatic ECC infections.[7-9,97,98,100] Filamentous bacteriophage have been implicated a variety of species as a key driver of invasive bacterial infection via disassembly of epithelial cell junctions. While our lab has demonstrated gonococcal derived filamentous phage protein ORF9 can break down epithelial cell junction integrity and induce cell death, it has yet to be seen whether this phenomenon is present in the context of GC infection rather than purified protein. Future studies should focus on whether infecting epithelial cell cultures and/or cervical explants show infection induced cell death and junction disassembly when infected with GC with and without filamentous phage. If GC containing filamentous phage show higher rates of induced cell death and junction disassembly, this could implicate filamentous phage as a driver for invasive disease and shed some light on a mechanism for GC transmigration into deeper tissue and therefore symptomatic GC infection.

This study shows that all four filamentous bacteriophage are not required for rescue of growth at 37ºC with shaking. While I see a dosage effect of needing at least two filamentous phage gene regions in order to reach wild type levels of growth, we still do not know why these genes are required for growth. Doing a growth study with only specific phage genes or combinations of these genes, rather than whole phage gene regions, could help to elucidate the mechanism of filamentous phage aided gonococcal growth if a phage gene is directly impacting this process. Further investigations into the mechanism of the condition dependent growth defect seen in the FA1090Δfil mutant is also key to elucidating the role of filamentous phage in GC biology. Doing gene expression analysis of FA1090 WT and FA1090Δfil at various temperatures could find genes impacted by the presence of filamentous phage in a temperature dependent manner. Further studies investigating the role of oxygen availability versus shear stress on the FA1090Δfil growth defect could help tease apart why FA1090Δfil can grow in static conditions. One possibility is growing GC in anaerobic conditions with shaking, to see if losing oxygen in the media, but retaining shear stress, would allow FA1090Δfil to grow.

Biofilm analysis in this chapter focused on biofilm size and structure in the context of static conditions. As I have shown, the FA1090Δfil mutant cannot growth under conditions with high levels of shear stress. Studies done in *N. meningitidis* showed that the presence of their homologous filamentous phage produces much larger biofilms, however their experiments were done under flow conditions.[23,57] Inhibited growth under shear stress may explain why meningococcal filamentous phage studies showed larger biofilms when they contain phage and our study showed equal biomass within the biofilm. Doing these experiments under flow conditions may link together the differences seen between this study and the meningococcal filamentous phage studies.

This same study done in *N. meningitidis* also showed a spatial distribution of pilin and phage genes, with phage genes on the periphery of the biofilm and pilin expressed on the bacteria near host cells.[23,64] To determine if a similar protein distribution exists in GC, a comparable study could measure pilin and phage expression using either fluorescent antibody staining or bacterial-MERFISH, followed by microscopy to analyze their spatial distribution within a biofilm. Bacterial-MERFISH would also allow tracking of RNA positions for these proteins within individual cells. This experiment could be conducted

under both static and flow conditions to examine how oxygen availability and shear stress impact this distribution.

All the experiments analyzing biofilm production and aggregation in this thesis were conducted in glass wells. However, the presence of human cell cultures between the bacteria and the glass slide might alter the biofilm production or aggregation patterns observed. To investigate biofilms and bacterial aggregation in a more physiologically relevant context, these experiments should be repeated in the presence of human cell cultures.

Appendices

Table 3. List of PCR primers for filamentous phage detection

Amplification Region	Forward Primer Sequence[d]	Reverse Primer Sequence[d]	Anneal Temp (C°)	Elongation time (min)
WHO O small inversion event[a]	CAGGAATTGATTGGAGCAGTATTGAAT	AAGATTTTGGGTTGAGGCAGTCAAACAGA	52	1.5
WHO N, WHO Y, and FA1090 no inversion event[a]	CAGGAATTGATTGGAGCAGTATTGAAT	TCGGTCAGTTGGAAAGCGAATAAATCCTG	52	1.5
PHI6 whole region[b]	CCCGTCGATGCCCGCGATTTCCGTTTGGCGCAATCTATCTACGAATATCTGG	CGATTCAGGCGGCGTTTTTCACCTTCCTGCAGGCGGGCGGCCATGTGGTT	56	16
PHI7 whole region[b]	ACTTCTGATTTTTCAGACGACTTTTGAAGATTGCCGCTACCCAATGATCTG	AGCCGTCGCGCAGGCAGTTGTTACGGTTACTAAAATCCCGTCATTTCCTCA	56	16
PHI8 whole region[b]	ACACGCAATCGCCGCTTTCCGCCGACCAAACCTCT	TTACCTTGTTTGCCAACGGCAGAATGGTAAATATGCCCGAG	56	16
PHI9 whole region[b]	GGAGAGCTACCGTGCGTGAAACCTGCTTTTTCTGCAAGT	CTATGACGACGACAGGGCGTTTCTTGACCATTTCCGGTTCTTCA	56	16
ORF9[a]	TGGAATTGGAAGACGGCACGGAG	TGTTCCTTGATTTGAATAGCATGTGCAATC	54	1.25

[a] Q5® High-Fidelity DNA Polymerase used for PCR with these primers

[b] LongAmp® Taq DNA Polymerase used for PCR with these primers

[c] All primer sequences are written in a 5'→3' orientation

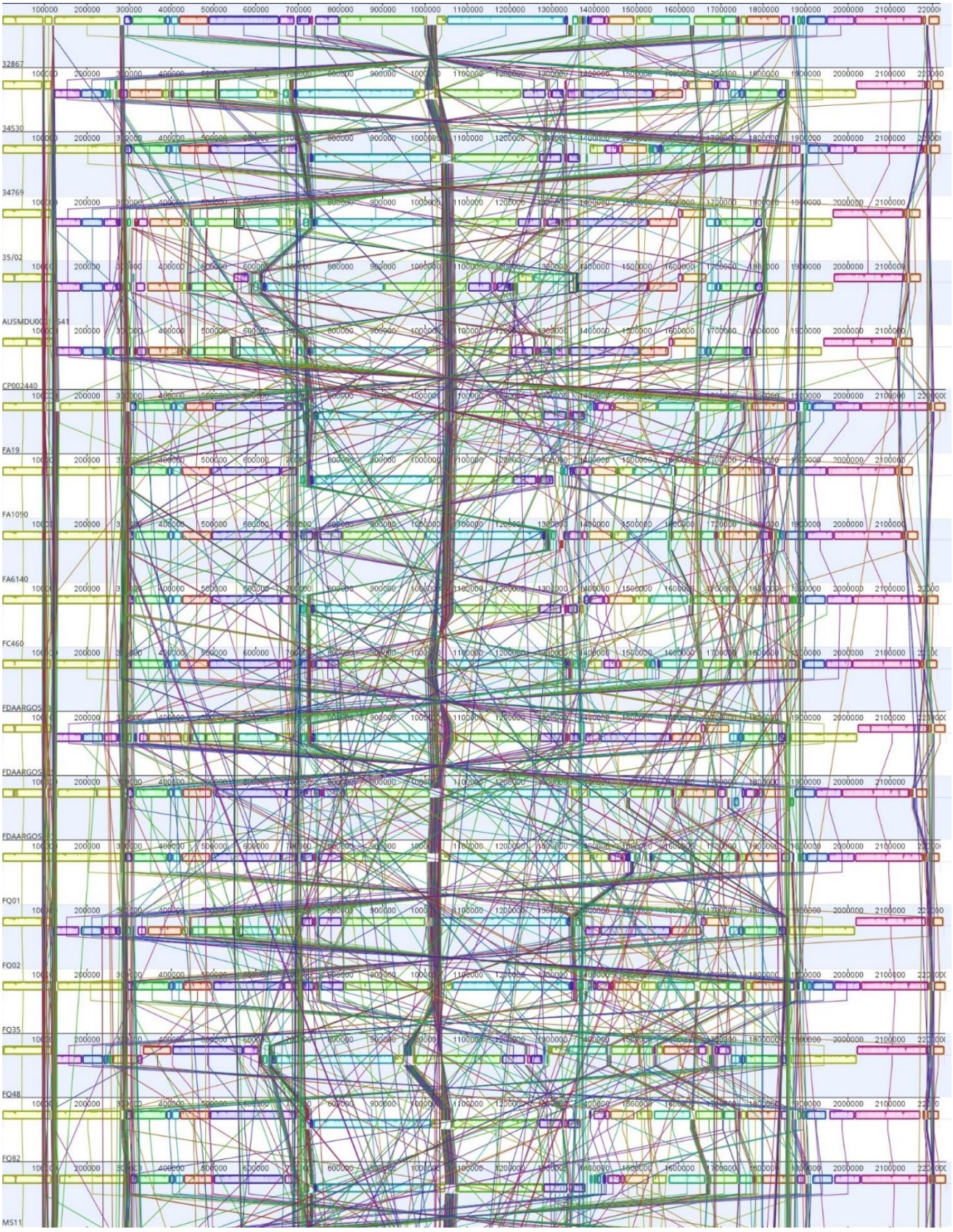
Table 4. List of PCR primers for filamentous phage creation

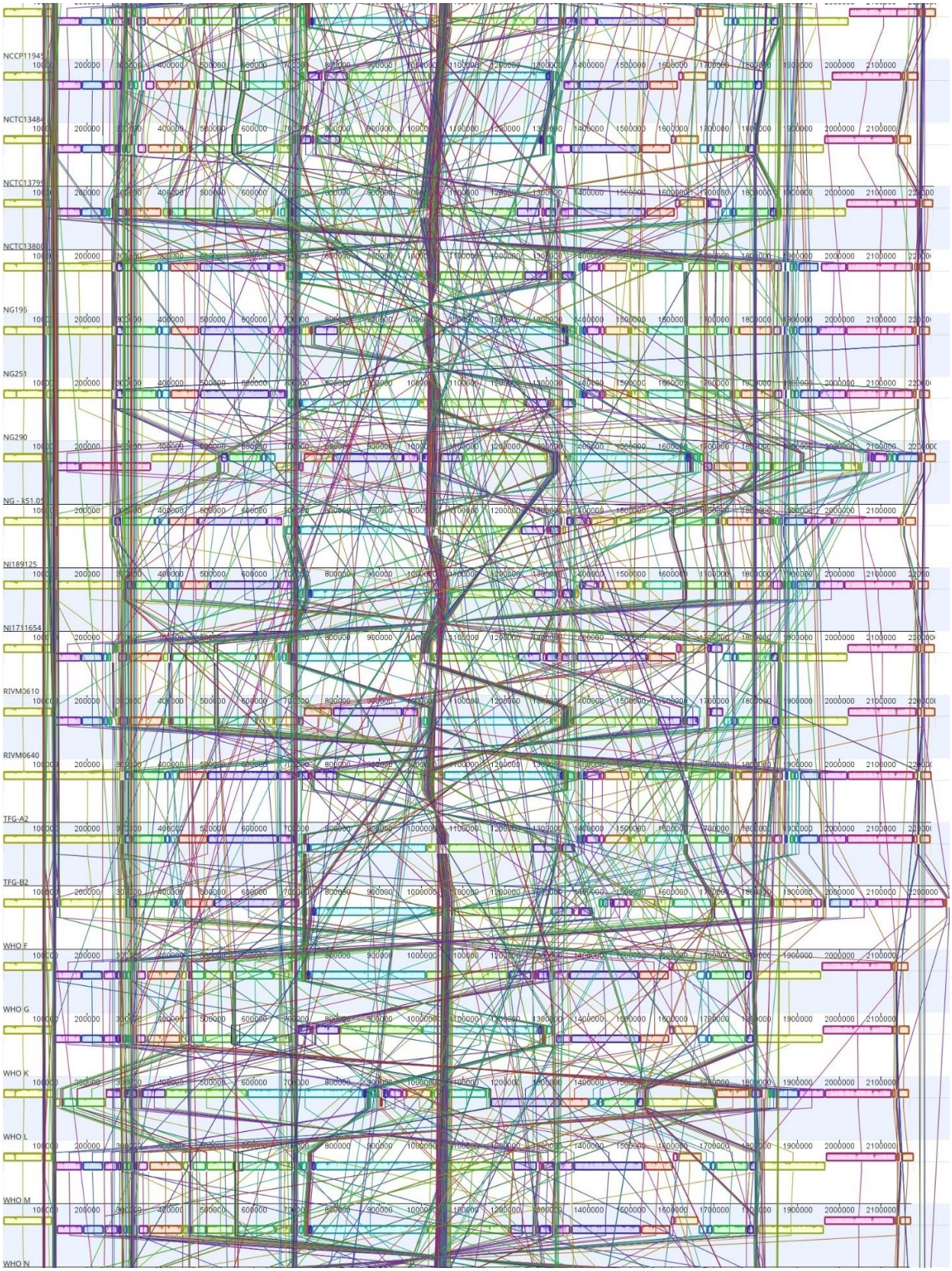
Phage Region	Upstream or Downstream	Facing phage region or bacterial genome	Restriction Enzyme	Primer Sequence
PHI6	Upstream	Facing phage region	BamHI	CACTGGATCCTAAAATCCGCTATTCGTCCGC
PHI6	Upstream	Facing bacterial genome	HindIII	CACTAAGCTTGGGATTACCCTTATATGCGGATATTTG
PHI6	Downstream	Facing phage region	BamHI	CACTGGATCCAATAAATTAAACAACGGGTGCACG
PHI6	Downstream	Facing bacterial genome	HindIII	CACTAAGCTTGGCAACAGGAATTTATCGTAAATG
PHI7	Upstream	Facing phage region	BamHI	CACTGGATCCATGCCGTCTGAAACGGCAC
PHI7	Upstream	Facing bacterial genome	PstI	CACTCTGCAGATCCCGAAAGGAAAAAAGGCA
PHI7	Downstream	Facing phage region	BamHI	CACTGGATCCTTTCATACGGTTTCAGACGGC
PHI7	Downstream	Facing bacterial genome	PstI	CACTCTGCAGCGAACAAATCAAACGAAAGTCCAT
PHI8	Upstream	Facing phage region	PstI	CACTCTGCAGCAATCTTTCAGACGGCACGAC
PHI8	Upstream	Facing bacterial genome	SacI	CACTGAGCTCATCCCGAAAGGAAAAAAGGCAT
PHI8	Downstream	Facing phage region	PstI	CACTCTGCAGTCCGCCAAGATGGAGGAAAT
PHI8	Downstream	Facing bacterial genome	SacI	CACTGAGCTCTGCGTATTTGTCATACATCTT
PHI9	Upstream	Facing phage region	BamHI	CACTGGATCCAAGCCGTTGGAGAAAGGGTTTAATG
PHI9	Upstream	Facing bacterial genome	XhoI	CACTCTCGAGATCCCGAAAGGAAAAAAGGCAT
PHI9	Downstream	Facing phage region	BamHI	CACTGGATCCCCTGCAGTCGTTATACATACCG
PHI9	Downstream	Facing bacterial genome	XhoI	CACTCTCGAGATGCCGTCTGAAAATTCAACAAAATTAAACGGTTGCG

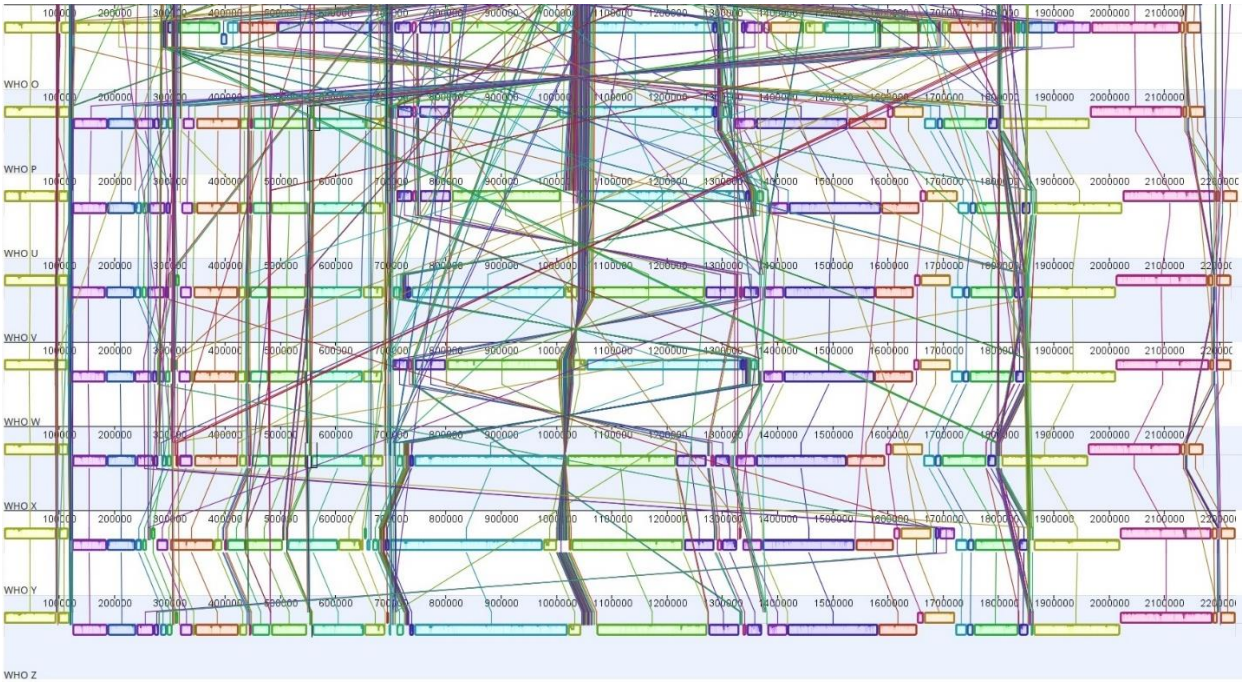
Table 5. Strain names and corresponding accession numbers.

Strain Name	Accession Numbers
32867	NZ_CP016015.1
34530	NZ_CP016016.1
34769	NZ_CP016017.1
35/02	NZ_CP012028.1
AUSMDU00010541	NZ_CP045832.1
CP002440	CP002440.1
FA19	NZ_CP012026.1
FA1090	NC_002946.2
FA6140	NZ_CP012027.1
FC460	NZ_AP018381.1
FDAARGOS204	NZ_CP020415.2
FDAARGOS205	NZ_CP020418.2
FDAARGOS207	NZ_CP020419.2
FQ01	NZ_CP034032.1
FQ02	NZ_CP034030.1
FQ35	NZ_CP034024.1
FQ48	NZ_CP034020.1
FQ82	NZ_CP034016.1
MS11	NC_022240.1
NCCP11945	NC_011035.1
NCTC13484	NZ_LR134358.1
NCTC13799	NZ_LT906437.1
NCTC13800	NZ_LT906472.1
NG196	NZ_CP043871.1

NG251	NZ_CP043816.1
NG290	NZ_CP043812.1
NG - k51.05	NZ_CP003974.1
NJ189125	NZ_CP041586.1
NJ1711654	NZ_CP041585.1
RIVM0610	NZ_CP019466.1
RIVM0640	NZ_CP019467.1
TFG-A2	NZ_CP032398.2
TFG-B2	NZ_CP032429.2
WHO F	NZ_LT591897.1
WHO G	NZ_LT591898.1
WHO K	NZ_LT591908.1
WHO L	NZ_LT591901.1
WHO M	NZ_LT591904.1
WHO N	NZ_LT591910.1
WHO O	NZ_LT592146.1
WHO P	NZ_LT592157.1
WHO U	NZ_LT592159.1
WHO V	NZ_LT592150.1
WHO W	NZ_LT592163.1
WHO X	NZ_LT592155.1
WHO Y	NZ_LT592161.1
WHO Z	NZ_LT592153.1


32867
34530
34769
35/02
AUSMDU00
541
CP002440
FA19
FA1090
FA6140
FC460
FDAARGOS
FDAARGOS
FDAARGOS
FQ01
FQ02
FQ35
FQ48
FQ82
MS11


NCCP11945
NCTC13484
NCTC13795
NCTC13800
NG195
NG251
NG290
NG - k51.05
NJ189125
NJ1711654
RIVM0610
RIVM0640
TFG-A2
TFG-B2
WHO F
WHO G
WHO K
WHO L
WHO M
WHO N



Supplemental Figure 1. Mauve Alignment of all GC strains analyzed. Mauve alignment of various strains of *N. gonorrhoeae*. Each color block represents homologous sequence between strains. Thin vertical/diagonal lines show movement of homologous color blocks of vertically adjacent genomes. All strains begin at the origin of replication on the ends of the linear DNA. This includes all strains from figure 7.

Sequences of FA1090 WT and FA1090Δfil analyzed in Figure 18 can be found in the following FASTA file: https://drive.google.com/file/d/1NTelTrHd66l4ZLUMwF-P-_YsgRt3hzMN/view?usp=drive_link

Bibliography

1. Walker, C.K., and Sweet, R.L. (2011). Gonorrhea infection in women: prevalence, effects, screening, and management. Int J Womens Health *3*, 197-206. 10.2147/IJWH.S13427.
2. CDC (2022). Sexually Transmitted Infections (STI) Surveillance.
3. Bodie, M., Gale-Rowe, M., Alexandre, S., Auguste, U., Tomas, K., and Martin, I. (2019). Addressing the rising rates of gonorrhea and drug-resistant gonorrhea: There is no time like the present. Canada Communicable Disease Report *45*, 54-62. 10.14745/ccdr.v45i23a02.
4. Unemo, M., Golparian, D., Sánchez-Busó, L., Grad, Y., Jacobsson, S., Ohnishi, M., Lahra, M.M., Limnios, A., Sikora, A.E., Wi, T., and Harris, S.R. (2016). The novel 2016 WHO Neisseria gonorrhoeae reference strains for global quality assurance of laboratory investigations: phenotypic, genetic and reference genome characterization. J Antimicrob Chemother *71*, 3096-3108. 10.1093/jac/dkw288.
5. Unemo, M., Seifert, H.S., Hook, E.W., Hawkes, S., Ndowa, F., and Dillon, J.R. (2019). Gonorrhoea. Nat Rev Dis Primers *5*, 79. 10.1038/s41572-019-0128-6.
6. Gilstrap, L.C., Herbert, W.N., Cunningham, F.G., Hauth, J.C., and Van Patten, H.G. (1977). Gonorrhea screening in male consorts of women with pelvic infection. Journal of the American Medical Association *238*, 965-966.
7. Wang, L.C., Yu, Q., Edwards, V., Lin, B., Qiu, J., Turner, J.R., Stein, D.C., and Song, W. (2017). Neisseria gonorrhoeae infects the human endocervix by activating non-muscle myosin II-mediated epithelial exfoliation. PLoS Pathog *13*, e1006269. 10.1371/journal.ppat.1006269.
8. Yu, Q., Wang, L.C., Di Benigno, S., Gray-Owen, S.D., Stein, D.C., and Song, W. (2019). Neisseria gonorrhoeae infects the heterogeneous epithelia of the human cervix using distinct mechanisms. PLoS Pathog *15*, e1008136. 10.1371/journal.ppat.1008136.
9. Yu, Q., Wang, L.-C., Di Benigno, S., Stein, D.C., and Song, W. (2021). Gonococcal invasion into epithelial cells depends on both cell polarity and ezrin. PLOS Pathogens *17*, e1009592. 10.1371/journal.ppat.1009592.

10. Kupsch, E.M., Knepper, B., Kuroki, T., Heuer, I., and Meyer, T.F. (1993). Variable opacity (Opa) outer membrane proteins account for the cell tropisms displayed by Neisseria gonorrhoeae for human leukocytes and epithelial cells. EMBO J *12*, 641-650.

11. Makino, S., van Putten, J.P., and Meyer, T.F. (1991). Phase variation of the opacity outer membrane protein controls invasion by Neisseria gonorrhoeae into human epithelial cells. EMBO J *10*, 1307-1315.

12. Stein, D.C., LeVan, A., Hardy, B., Wang, L.C., Zimmerman, L., and Song, W. (2015). Expression of Opacity Proteins Interferes with the Transmigration of Neisseria gonorrhoeae across Polarized Epithelial Cells. PLoS One *10*, e0134342. 10.1371/journal.pone.0134342.

13. Naumann, M., Rudel, T., and Meyer, T.F. (1999). Host cell interactions and signalling with Neisseria gonorrhoeae. Curr Opin Microbiol *2*, 62-70. 10.1016/s1369-5274(99)80011-3.

14. Quillin, S.J., and Seifert, H.S. (2018). Neisseria gonorrhoeae host adaptation and pathogenesis. Nature Reviews Microbiology *16*, 226-240. 10.1038/nrmicro.2017.169.

15. Grassmé, H.U., Ireland, R.M., and van Putten, J.P. (1996). Gonococcal opacity protein promotes bacterial entry-associated rearrangements of the epithelial cell actin cytoskeleton. Infection and Immunity *64*, 1621-1630. 10.1128/iai.64.5.1621-1630.1996.

16. Blaskewicz, C.D., Pudney, J., and Anderson, D.J. (2011). Structure and Function of Intercellular Junctions in Human Cervical and Vaginal Mucosal Epithelia1. Biology of Reproduction *85*, 97-104. 10.1095/biolreprod.110.090423.

17. Sawada, N., Murata, M., Kikuchi, K., Osanai, M., Tobioka, H., Kojima, T., and Chiba, H. (2003). Tight junctions and human diseases. Medical Electron Microscopy *36*, 147-156. 10.1007/s00795-003-0219-y.

18. Gonzalez-Mariscal, L., Contreras, R.G., Bolívar, J.J., Ponce, A., Chávez De Ramirez, B., and Cereijido, M. (1990). Role of calcium in tight junction formation between epithelial cells. Am J Physiol *259*, C978-986. 10.1152/ajpcell.1990.259.6.C978.

19. Di Pierro, M., Lu, R., Uzzau, S., Wang, W., Margaretten, K., Pazzani, C., Maimone, F., and Fasano, A. (2001). Zonula occludens toxin structure-function analysis. Identification of the

fragment biologically active on tight junctions and of the zonulin receptor binding domain. Journal of Biological Chemistry *276*, 19160-19165. 10.1074/jbc.M009674200.

20. Fasano, A., Baudry, B., Pumplin, D.W., Wasserman, S.S., Tall, B.D., Ketley, J.M., and Kaper, J.B. (1991). Vibrio cholerae produces a second enterotoxin, which affects intestinal tight junctions. Proceedings of the National Academy of Sciences *88*, 5242-5246. 10.1073/pnas.88.12.5242.
21. Liu, F., Lee, H., Lan, R., and Zhang, L. (2016). Zonula occludens toxins and their prophages in Campylobacter species. Gut Pathog *8*, 43. 10.1186/s13099-016-0125-1.
22. Steichen, C.T., Shao, J.Q., Ketterer, M.R., and Apicella, M.A. (2008). Gonococcal cervicitis: a role for biofilm in pathogenesis. J Infect Dis *198*, 1856-1861. 10.1086/593336.
23. Meyer, J., Brissac, T., Frapy, E., Omer, H., Euphrasie, D., Bonavita, A., Nassif, X., and Bille, E. (2016). Characterization of MDAΦ, a temperate filamentous bacteriophage of Neisseria meningitidis. Microbiology (Reading) *162*, 268-282. 10.1099/mic.0.000215.
24. Römling, U., and Balsalobre, C. (2012). Biofilm infections, their resilience to therapy and innovative treatment strategies. J Intern Med *272*, 541-561. 10.1111/joim.12004.
25. Rocha, E.P. (2008). The organization of the bacterial genome. Annu Rev Genet *42*, 211-233. 10.1146/annurev.genet.42.110807.091653.
26. Rocha, E.P. (2004). Order and disorder in bacterial genomes. Curr Opin Microbiol *7*, 519-527. 10.1016/j.mib.2004.08.006.
27. Rocha, E.P., Cornet, E., and Michel, B. (2005). Comparative and evolutionary analysis of the bacterial homologous recombination systems. PLoS Genet *1*, e15. 10.1371/journal.pgen.0010015.
28. Rocha, E.P. (2006). Inference and analysis of the relative stability of bacterial chromosomes. Mol Biol Evol *23*, 513-522. 10.1093/molbev/msj052.
29. Rocha, E.P.C., and Bikard, D. (2022). Microbial defenses against mobile genetic elements and viruses: Who defends whom from what? PLoS Biol *20*, e3001514. 10.1371/journal.pbio.3001514.
30. Rocha, E.P. (2008). Evolutionary patterns in prokaryotic genomes. Curr Opin Microbiol *11*, 454-460. 10.1016/j.mib.2008.09.007.

31. Rocha, E.P.C. (2004). The replication-related organization of bacterial genomes. Microbiology (Reading) *150*, 1609-1627. 10.1099/mic.0.26974-0.

32. Zahavi, L., Lavon, A., Reicher, L., Shoer, S., Godneva, A., Leviatan, S., Rein, M., Weissbrod, O., Weinberger, A., and Segal, E. (2023). Bacterial SNPs in the human gut microbiome associate with host BMI. Nat Med *29*, 2785-2792. 10.1038/s41591-023-02599-8.

33. Soucy, S.M., Huang, J., and Gogarten, J.P. (2015). Horizontal gene transfer: building the web of life. Nat Rev Genet *16*, 472-482. 10.1038/nrg3962.

34. van der Woude, M.W., and Bäumler, A.J. (2004). Phase and antigenic variation in bacteria. Clin Microbiol Rev *17*, 581-611, table of contents. 10.1128/CMR.17.3.581-611.2004.

35. Cehovin, A., and Lewis, S.B. (2017). Mobile genetic elements in Neisseria gonorrhoeae: movement for change. Pathog Dis *75*. 10.1093/femspd/ftx071.

36. Spencer-Smith, R., Varkey, E.M., Fielder, M.D., and Snyder, L.A. (2012). Sequence features contributing to chromosomal rearrangements in Neisseria gonorrhoeae. PLoS One *7*, e46023. 10.1371/journal.pone.0046023.

37. Jerse, A.E., Cohen, M.S., Drown, P.M., Whicker, L.G., Isbey, S.F., Seifert, H.S., and Cannon, J.G. (1994). Multiple gonococcal opacity proteins are expressed during experimental urethral infection in the male. J Exp Med *179*, 911-920. 10.1084/jem.179.3.911.

38. Hill, S.A., and Davies, J.K. (2009). Pilin gene variation in Neisseria gonorrhoeae: reassessing the old paradigms. FEMS Microbiol Rev *33*, 521-530. 10.1111/j.1574-6976.2009.00171.x.

39. Murphy, G.L., Connell, T.D., Barritt, D.S., Koomey, M., and Cannon, J.G. (1989). Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. Cell *56*, 539-547. 10.1016/0092-8674(89)90577-1.

40. El-Rami, F.E., Zielke, R.A., Wi, T., Sikora, A.E., and Unemo, M. (2019). Quantitative Proteomics of the 2016 WHO. Mol Cell Proteomics *18*, 127-150. 10.1074/mcp.RA118.001125.

41. Unemo, M., Fasth, O., Fredlund, H., Limnios, A., and Tapsall, J. (2009). Phenotypic and genetic characterization of the 2008 WHO Neisseria gonorrhoeae reference strain panel intended for global quality assurance and quality control of gonococcal antimicrobial resistance surveillance for public health purposes. J Antimicrob Chemother *63*, 1142-1151. 10.1093/jac/dkp098.

42. Ackermann, H.-W. (2009). Phage Classification and Characterization. In Methods in Molecular Biology, (Humana Press), pp. 127-140. 10.1007/978-1-60327-164-6_13.

43. Brüssow, H., Canchaya, C., and Hardt, W.D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiology and Molecular Biology Reviews *68*, 560-602, table of contents. 10.1128/MMBR.68.3.560-602.2004.

44. Stassen, A.P., Folmer, R.H., Hilbers, C.W., and Konings, R.N. (1994). Single-stranded DNA binding protein encoded by the filamentous bacteriophage M13: structural and functional characteristics. Mol Biol Rep *20*, 109-127. 10.1007/BF00990543.

45. Stanley, N.R., Britton, R.A., Grossman, A.D., and Lazazzera, B.A. (2003). Identification of catabolite repression as a physiological regulator of biofilm formation by Bacillus subtilis by use of DNA microarrays. J Bacteriol *185*, 1951-1957. 10.1128/JB.185.6.1951-1957.2003.

46. Ackermann, H.W. (2009). Phage classification and characterization. Methods in Molecular Biology *501*, 127-140. 10.1007/978-1-60327-164-6_13.

47. Marvin, D.A., Symmons, M.F., and Straus, S.K. (2014). Structure and assembly of filamentous bacteriophages. Prog Biophys Mol Biol *114*, 80-122. 10.1016/j.pbiomolbio.2014.02.003.

48. Roux, S., Krupovic, M., Daly, R.A., Borges, A.L., Nayfach, S., Schulz, F., Sharrar, A., Matheus Carnevali, P.B., Cheng, J.F., Ivanova, N.N., et al. (2019). Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. Nat Microbiol *4*, 1895-1906. 10.1038/s41564-019-0510-x.

49. Clokie, M.R., Millard, A.D., Letarov, A.V., and Heaphy, S. (2011). Phages in nature. Bacteriophage *1*, 31-45. 10.4161/bact.1.1.14942.

50. Bondy-Denomy, J., and Davidson, A.R. (2014). When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. Journal of Microbiology *52*, 235-242. 10.1007/s12275-014-4083-3.

51. Roux, S., Krupovic, M., Daly, R.A., and al., e. (2019). Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth’s biomes. Nature Microbiology.

52. Piekarowicz, A., Majchrzak, M., Kłyz, A., and Adamczyk-Popławska, M. (2006). Analysis of the filamentous bacteriophage genomes integrated into Neisseria gonorrhoeae FA1090 chromosome. Pol J Microbiol *55*, 251-260.

53. Piekarowicz, A., Kłyz, A., Majchrzak, M., Adamczyk-Popławska, M., Maugel, T.K., and Stein, D.C. (2007). Characterization of the dsDNA prophage sequences in the genome of Neisseria gonorrhoeae and visualization of productive bacteriophage. BMC Microbiol *7*, 66. 10.1186/1471-2180-7-66.

54. Piekarowicz, A., Kłyż, A., Majchrzak, M., Szczêsna, E., Piechucki, M., Kwiatek, A., Maugel, T.K., and Stein, D.C. (2014). Neisseria gonorrhoeae filamentous phage NgoΦ6 is capable of infecting a variety of Gram-negative bacteria. J Virol *88*, 1002-1010. 10.1128/JVI.02707-13.

55. Piekarowicz, A., Kłyż, A., Majchrzak, M., and Stein, D.C. (2016). Oral Immunization of Rabbits with S. enterica Typhimurium Expressing Neisseria gonorrhoeae Filamentous Phage Φ6 Induces Bactericidal Antibodies Against N. gonorrhoeae. Scientific Reports *6*, 22549. 10.1038/srep22549.

56. Piekarowicz, A., Kłyż, A., Adamczyk-Popławska, M., and Stein, D.C. (2020). Association of host proteins with the broad host range filamentous phage NgoΦ6 of Neisseria gonorrhoeae. PLOS ONE *15*, e0240579. 10.1371/journal.pone.0240579.

57. Al Suwayyid, B.A., Rankine-Wilson, L., Speers, D.J., Wise, M.J., Coombs, G.W., and Kahler, C.M. (2020). Meningococcal Disease-Associated Prophage-Like Elements Are Present in Neisseria gonorrhoeae and Some Commensal Neisseria Species. Genome Biol Evol *12*, 3938-3950. 10.1093/gbe/evaa023.

58. Russel, M., and Model, P. (1989). Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. J Virol *63*, 3284-3295. 10.1128/JVI.63.8.3284-3295.1989.

59. Zou, Q.H., Li, Q.H., Zhu, H.Y., Feng, Y., Li, Y.G., Johnston, R.N., Liu, G.R., and Liu, S.L. (2010). SPC-P1: a pathogenicity-associated prophage of Salmonella paratyphi C. BMC Genomics *11*, 729. 10.1186/1471-2164-11-729.

60. Bisht, K., Moore, J.L., Caprioli, R.M., Skaar, E.P., and Wakeman, C.A. (2021). Impact of temperature-dependent phage expression on Pseudomonas aeruginosa biofilm formation. NPJ Biofilms Microbiomes *7*, 22. 10.1038/s41522-021-00194-8.

61. Secor, P.R., Jennings, L.K., Michaels, L.A., Sweere, J.M., Singh, P.K., Parks, W.C., and Bollyky, P.L. (2015). Biofilm assembly becomes crystal clear – filamentous bacteriophage organize the Pseudomonas aeruginosa biofilm matrix into a liquid crystal. Microb Cell *3*, 49-52. 10.15698/mic2016.01.475.

62. Secor, P.R., Burgener, E.B., Kinnersley, M., Jennings, L.K., Roman-Cruz, V., Popescu, M., Van Belleghem, J.D., Haddock, N., Copeland, C., Michaels, L.A., et al. (2020). Pf Bacteriophage and Their Impact on Pseudomonas Virulence, Mammalian Immunity, and Chronic Infections. Front Immunol *11*, 244. 10.3389/fimmu.2020.00244.

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

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

65. Davis, B.M., Moyer, K.E., Boyd, E.F., and Waldor, M.K. (2000). CTX prophages in classical biotype Vibrio cholerae: functional phage genes but dysfunctional phage genomes. Journal of Bacteriology *182*, 6992-6998. 10.1128/JB.182.24.6992-6998.2000.

66. Fasano, A., Fiorentini, C., Donelli, G., Uzzau, S., Kaper, J.B., Margaretten, K., Ding, X., Guandalini, S., Comstock, L., and Goldblum, S.E. (1995). Zonula occludens toxin modulates tight junctions through protein kinase C-dependent actin reorganization, in vitro. Journal of Clinical Investigations *96*, 710-720. 10.1172/JCI118114.

67. Hassan, F., Kamruzzaman, M., Mekalanos, J.J., and Faruque, S.M. (2010). Satellite phage TLCφ enables toxigenic conversion by CTX phage through dif site alteration. Nature *467*, 982-985. 10.1038/nature09469.

68. Addy, H.S., Askora, A., Kawasaki, T., Fujie, M., and Yamada, T. (2012). The filamentous phage ϕRSS1 enhances virulence of phytopathogenic Ralstonia solanacearum on tomato. Phytopathology *102*, 244-251. 10.1094/PHYTO-10-11-0277.

69. Yu, Z.C., Chen, X.L., Shen, Q.T., Zhao, D.L., Tang, B.L., Su, H.N., Wu, Z.Y., Qin, Q.L., Xie, B.B., Zhang, X.Y., et al. (2015). Filamentous phages prevalent in Pseudoalteromonas spp. confer properties advantageous to host survival in Arctic sea ice. ISME J *9*, 871-881. 10.1038/ismej.2014.185.

70. Harper, D., Parracho, H., Walker, J., Sharp, R., Hughes, G., Werthén, M., Lehman, S., and Morales, S. (2014). Bacteriophages and Biofilms. Antibiotics *3*, 270-284. 10.3390/antibiotics3030270.

71. Secor, R., Patrick, Sweere, M., Johanna, Michaels, A., Lia, Malkovskiy, V., Andrey, Lazzareschi, D., Katznelson, E., Rajadas, J., Birnbaum, E., Michael, Arrigoni, A., Braun, R., Kathleen, et al. (2015). Filamentous Bacteriophage Promote Biofilm Assembly and Function. Cell Host & Microbe *18*, 549-559. 10.1016/j.chom.2015.10.013.

72. Uzzau, S., and Fasano, A. (2000). Cross-talk between enteric pathogens and the intestine. Cell Microbiol *2*, 83-89. 10.1046/j.1462-5822.2000.00041.x.

73. Unemo, M., Seifert, H.S., Hook, E.W., Hawkes, S., Ndowa, F., and Dillon, J.-A.R. (2019). Gonorrhoea. Nature Reviews Disease Primers *5*. 10.1038/s41572-019-0128-6.

74. Kelman, Z., Yuzhakov, A., Andjelkovic, J., and O'Donnell, M. (1998). Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. EMBO J *17*, 2436-2449. 10.1093/emboj/17.8.2436.

75. Black, C.G., Fyfe, J.A., and Davies, J.K. (1998). Absence of an SOS-like system in Neisseria gonorrhoeae. Gene *208*, 61-66. 10.1016/s0378-1119(97)00653-7.

76. Savitskaya, V.Y., Monakhova, M.V., Iakushkina, I.V., Borovikova, I.I., and Kubareva, E.A. (2022). Neisseria gonorrhoeae: DNA Repair Systems and Their Role in Pathogenesis. Biochemistry (Mosc) *87*, 965-982. 10.1134/S0006297922090097.

77. Hay, I.D., and Lithgow, T. (2019). Filamentous phages: masters of a microbial sharing economy. EMBO Rep *20*. 10.15252/embr.201847427.

78. Webster, J.S., and Hammer, D. (1983). Thermistor measurement of male sexual arousal. Psychophysiology *20*, 111-115. 10.1111/j.1469-8986.1983.tb00911.x.

79. Phillips, N.J., Steichen, C.T., Schilling, B., Post, D.M., Niles, R.K., Bair, T.B., Falsetta, M.L., Apicella, M.A., and Gibson, B.W. (2012). Proteomic analysis of Neisseria gonorrhoeae biofilms shows shift to anaerobic respiration and changes in nutrient transport and outermembrane proteins. PLoS One *7*, e38303. 10.1371/journal.pone.0038303.

80. Isabella, V.M., and Clark, V.L. (2011). Deep sequencing-based analysis of the anaerobic stimulon in Neisseria gonorrhoeae. BioMed Central Genomics *12*, 51. 10.1186/1471-2164-12-51.

81. Knapp, J.S., and Clark, V.L. (1984). Anaerobic growth of Neisseria gonorrhoeae coupled to nitrite reduction. Infect Immun *46*, 176-181. 10.1128/iai.46.1.176-181.1984.

82. Melaugh, G., Hutchison, J., Kragh, K.N., Irie, Y., Roberts, A., Bjarnsholt, T., Diggle, S.P., Gordon, V.D., and Allen, R.J. (2016). Shaping the Growth Behaviour of Biofilms Initiated from Bacterial Aggregates. PLoS One *11*, e0149683. 10.1371/journal.pone.0149683.

83. Ng, K.Y.B., Mingels, R., Morgan, H., Macklon, N., and Cheong, Y. (2018). In vivo oxygen, temperature and pH dynamics in the female reproductive tract and their importance in human conception: a systematic review. Hum Reprod Update *24*, 15-34. 10.1093/humupd/dmx028.

84. Maggiorani, D., Dissard, R., Belloy, M., Saulnier-Blache, J.S., Casemayou, A., Ducasse, L., Grès, S., Bellière, J., Caubet, C., Bascands, J.L., et al. (2015). Shear Stress-Induced Alteration of Epithelial Organization in Human Renal Tubular Cells. PLoS One *10*, e0131416. 10.1371/journal.pone.0131416.

85. Yang, L., Wang, Y., Yu, P., Ren, S., Zhu, Z., Jin, Y., Yan, J., Peng, X., and Chen, L. (2020). Prophage-Related Gene. Front Cell Infect Microbiol *10*, 595709. 10.3389/fcimb.2020.595709.

86. BAATOUT, S., DE BOEVER, P., and MERGEAY, M. (2005). Temperature-induced changes in bacterial physiology as determined by flow cytometry. 1 ed. Annals of Microbiolgy.

87. Schiffer, A., Gardner, M.N., Lynn, R.H., and Tagarielli, V.L. (2017). A new apparatus to induce lysis of planktonic microbial cells by shock compression, cavitation and spray. R Soc Open Sci *4*, 160939. 10.1098/rsos.160939.

88. Ebrahimi, A., Csonka, L.N., and Alam, M.A. (2018). Analyzing Thermal Stability of Cell Membrane of Salmonella Using Time-Multiplexed Impedance Sensing. Biophys J *114*, 609-618. 10.1016/j.bpj.2017.10.032.

89. Jung, A., Yun, J.S., Kim, S., Kim, S.R., Shin, M., Cho, D.H., Choi, K.S., and Chang, J.H. (2019). Crystal Structure of Histidine Triad Nucleotide-Binding Protein from the Pathogenic Fungus. Mol Cells *42*, 56-66. 10.14348/molcells.2018.0377.

90. Templin, M.F., Ursinus, A., and Höltje, J.V. (1999). A defect in cell wall recycling triggers autolysis during the stationary growth phase of Escherichia coli. EMBO J *18*, 4108-4117. 10.1093/emboj/18.15.4108.

91. Korza, H.J., and Bochtler, M. (2005). Pseudomonas aeruginosa LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow. J Biol Chem *280*, 40802-40812. 10.1074/jbc.M506328200.

92. Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., and Raschellà, G. (2017). Zinc-finger proteins in health and disease. Cell Death Discov *3*, 17071. 10.1038/cddiscovery.2017.71.

93. Nudel, K., McClure, R., Moreau, M., Briars, E., Abrams, A.J., Tjaden, B., Su, X.H., Trees, D., Rice, P.A., Massari, P., and Genco, C.A. (2018). Transcriptome Analysis of Neisseria gonorrhoeae during Natural Infection Reveals Differential Expression of Antibiotic Resistance Determinants between Men and Women. mSphere *3*. 10.1128/mSphereDirect.00312-18.

94. Rock, J.D., Mahnane, M.R., Anjum, M.F., Shaw, J.G., Read, R.C., and Moir, J.W. (2005). The pathogen Neisseria meningitidis requires oxygen, but supplements growth by denitrification. Nitrite, nitric oxide and oxygen control respiratory flux at genetic and metabolic levels. Mol Microbiol *58*, 800-809. 10.1111/j.1365-2958.2005.04866.x.

95. Mellies, J., Jose, J., and Meyer, T.F. (1997). The Neisseria gonorrhoeae gene aniA encodes an inducible nitrite reductase. Mol Gen Genet *256*, 525-532. 10.1007/s004380050597.

96. Green, J., Rolfe, M.D., and Smith, L.J. (2014). Transcriptional regulation of bacterial virulence gene expression by molecular oxygen and nitric oxide. Virulence *5*, 794-809. 10.4161/viru.27794.

97. Swanson, K.V., Griffiss, J.M., Edwards, V.L., Stein, D.C., and Song, W. (2011). Neisseria gonorrhoeae-induced transactivation of EGFR enhances gonococcal invasion. Cell Microbiol *13*, 1078-1090. 10.1111/j.1462-5822.2011.01603.x.

98. Edwards, V.L., Wang, L.C., Dawson, V., Stein, D.C., and Song, W. (2013). Neisseria gonorrhoeae breaches the apical junction of polarized epithelial cells for transmigration by activating EGFR. Cellular Microbiology *15*, 1042-1057. 10.1111/cmi.12099.

99. M, B., M, G.-R., S, A., U, A., K, T., and I, M. (2019). Addressing the rising rates of gonorrhea and drug resistant gonorrhea: There's no time like the present. *Canada communicable disease report*.

100. Wang, L.-C., Yu, Q., Stein, D., and Song, W. (2018). Immunofluorescence Analysis of Human Endocervical Tissue Explants Infected with Neisseria gonorrhoeae. BIO-PROTOCOL *8*. 10.21769/bioprotoc.2720.