ABSTRACT

Title of Document: EFFECT OF OPA EXPRESSION ON TRANSMIGRATION OF *NEISSERIA GONORRHOEAE* ACROSS POLARIZED EPITHELIAL CELLS.

Luz Angela Adriana Le Van
Doctor of Philosophy, 2013

Directed By: Professor, Daniel C Stein,
Department of Cell Biology and Molecular Genetics

Neisseria gonorrhoeae (GC) is a solely human pathogen that causes gonorrhea. This study examines how Opas, which are surface factors expressed by GC that undergo antigenic and phase variation can affect transmigration across epithelium. Opas are encoded by 11 different genes. A gonococcal variant that lacked all *opa* genes was constructed to help elucidate the role of Opa in pathogenesis. This variant retained most physical characteristics of the parent strain including growth rate and LOS profile but proved to produce a different interaction with other GC by being unable to bind LOS of adjacent GC and form microcolonies. Lack of Opa expression increased the ability of GC to transmigrate across polarized colonic epithelial cells T84. When the *opa* deletion variant was not expressing pili, bacteria were observed entering and crossing the polarized epithelia as early as four hours after infection. GC were observed at the bottom of the polarized epithelial monolayer demonstrated by confocal microscopy. While GC transmigrate across the monolayer, they do not appear to disrupt the integrity of tight junction proteins. Transepithelial

resistance did not show a significant change and there was no leakage of FITC or HRP. Inhibitors of acid sphingomyelin and F-actin did not cause a redistribution of ZO-1 and did not increase the transmigration of GC. Only in the presence of EGTA, a calcium chelator, were Opa-expressing GC observed crossing the monolayer through visible disruption of the tight junctions. Induction of TNF-α by T84 cells was increase when infected with GC but not the production of IL-8. This study indicates that the lack of expression of Opa and pili leads to an increase in invasion into subepithelial tissues.

EFFECT OF OPA EXPRESSION ON TRANSMIGRATION OF *NEISSERIA GONORRHOEAE* ACROSS POLARIZED EPITHELIAL CELLS

By

Luz Angela Adriana Le Van

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
2013

Advisory Committee:
Professor Daniel C Stein, Chair
Professor Wenxia Song
Professor Richard Stewart
Professor Kevin McIver
Professor Richard Payne

Dedication

This work is dedicated to my husband William and our sons Andre, Ethan and Austine for their great unconditional support giving me their love, strength, and patience to complete this work; for encouraging me even in the hardest times. Thanks for giving meaning to my life

A Mami, Papi, Camila y Javier por creer en mi, por su amor y por siempre estar listos a ayudarme en cualquier momento.

Acknowledgements

First, I would like to thank my advisor Dr. Daniel Stein, for your unconditional support, guidance, understanding and immeasurable patience during all these years. I really appreciate all the help you have given me. I would like to thank Dr. Wenxia Song for your valuable advice not only in science but also about life. I would also like to thank my committee members, Dr. Richard Stewart, Dr. Richard Payne and Dr. Kevin McIver for your scientific advice through the years.

I want to specially thank Dr. Ann Smith for being a great mentor during my teaching career at UMD; thanks for all your support. A heartfelt thanks to the students in the Stein and Song labs, who are not only collaborators, but my friends: Ellen, Julie, Sam, Clint, Mark Rodgers, Vonetta, Lindsey, Karen, Mandy, Katie, Mark Wang, Senthil, Beth, Hwalih, Azy, Dr. Andrzej Piekarowicz. A special thanks to Meredith Davis, she is the reason I started working on this project, to Lindsey Zimmerman for all her help with the experiments, Mandy, Karen and Vonetta for contributing to the manuscript. Thanks to Dr. Piekarowicz for all his helpful advice. Thanks to Dr. Philip DeShong and his student Juhee Park for constructing the LOS conjugate. Thanks to Dr. Delwiche for his help during the last year. Thanks to Laura Cathcart for all her help. Furthermore, many thanks to the lifelong friendships that I have made outside the lab, Monica Pava-Ripoll, Melba Muñoz, Virginia Sanchez Puerta, Maximo Rivarola. A special thanks to Dr. Edgar Moctezuma for starting the SACNAS chapter at UMD and getting me involved in mentoring Latino students. Many thanks go to the students, staff and faculty of CBMG who helped along the way.

This work would not have been possible without financial aid from grants to DC Stein AI 24452**,** AI 68888 from NIH. NIH predoctoral fellowship (1F31AI075596-01A2).

Finally, I want to thank my family. Mami y papi for your love and encouragement, I could have not done it without you. Thanks for always believing in me. My grandparents, Martha and Eliecer for being the people with the biggest hearts. My brother, Javier for your unconditional love. I want to thank my sister, Camila, for not only being my sister but also my best friend. I know I could always count on you. No matter living 3835 miles away from each other, I know you are only a phone call away and it certainly never feels like we are that far apart. Candy, since you left two years ago, there is not a day that goes by, without the thought of you. Thanks for being the best mother in law. I'll always miss you. Love of my life, William, I could have not gotten this far without you. Thanks for your love, encouragement and moral support always. Thanks for being the best life partner to share our dreams together. Still loving you..... I cannot believe it is going to be the first time since we got married that one of us is not going to be enrolled in school. Last but certainly not least, my boys: Andre, Ethan, and Austine. You are the reason I get up with a smile every day. Your immeasurable love and understanding have made this journey worth it. Andre, for inspiring me every day with that amazing curiosity you have about life; Ethan for always being such a sweet, loving boy, even as a teen; Austine, for lighting up every room you walk in. Thank you babies, I love you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3-Deoxy-D-*manno*-oct-2-ulosonic acid	KDO
acid sphingomyelinase	ASM
asialoglycoprotein receptor	ASGP-R
carcinoembryonic antigen realted cell adhesion molecule	CEACAM
Chinese hamster ovary cells	CHO
coding repeat	CR
diacylglycerol	DAG
disseminated gonococcal infection	DGI
enzyme-linked immunosorbent assay	ELISA
epidermal growth factor receptor	EGFR
ethylene glycol tetraacetic acid	EGTA
extracellular signal related kinase	ERK
fluorescein isothiocyanate	FITC
glycosylphosphatidylinositol	GPI
gonococcal media base	GCK
heparin sulfate proteoglycan	HSPG

horseradish peroxidase	HRP
hypervariable region	HV
immunoreceptor tyrosine-based activation motif	ITAM
interleukin	IL
junctional adhesion molecules	JAMs
lacto-N-neotetraose	LNT
Lipooligosaccharide	LOS
membrane associated guanylate inverted	MAGI
monocyte chemoattractant protein-1	MCP-1
multiplicity of infection	MOI
Neisseria gonorrhoeae, gonococci	GC
Neisseria meingitidis	MC
nuclear factor-kappaβ	NF-κβ
opacity associated proteins	Opa
open reading frame	ORF
pelvic inflamatory disease	PID
phosphatidylinositide 3-kinase	PI_3K

phosphatidylinositol (3,4,5)-triphosphate	PIP_3
phosphatidylinositol 4,5-bisphosphate	PIP_2
phosphocholine-phospholipase C	PC-PLC
phospholipase-gamma	PLCγ
polymerase chain reaction	PCR
protein kinase C	PKC
semi variable region	SV
sexually transmitted disease	STD
sodium dodecyl sulfate polyacrylamide gel electrophoresis	SDS-PAGE
spectinomycin resistance	$spec^R$
tight junctions	TJ
transepithelial resistance	TEER
transforming growth factor beta	TGF-β1
Tris-borate-EDTA	TBE
tumor necrosis factor alpha	TNF-α
zonula occludens	ZO

Chapter 1: Introduction

1.1 Background

The *Neisseria* is a genus of gram-negative diplococci that is biochemically oxidase and catalase positive. Two species are typically associated with disease, *N. meningitidis* (MC) and *N. gonorrhoeae* (GC). MC is an opportunistic pathogen while GC is an obligate human pathogen. GC typically colonizes the urethral tract in men while in women it can also colonize the various tissues of the female reproductive tract. In men, gonorrhea is usually symptomatic, but in women, disseminated infections are more common since the disease can go untreated due to its asymptomatic characteristics. As a result, this can lead to sequelae like endocarditis, perihepatitis and joint damage (Kerle et al., 1992). Other problems linked to gonorrhea include the emergence of strains resistant to antibiotics such as penicillin, tetracycline, quinolones and cephalosporins (Ison et al., 1998; Tanaka, M, 2012; Bolan et al., 2012) as well as the increase incidence in HIV infections (Fleming and Wasserheit, 1999). The highest incidence of disease is seen among young women between the ages 15-19 and men ages 20-24 (Moran, 2005).

There is an estimated 62 million new cases of gonorrhea worldwide each year with an average of 22 million cases at any given time. Gonorrhea is the second most reported bacterial disease in the USA with 336,742 cases according to the CDC 2008 STD Surveillance. Infections with *N. gonorrhoeae* are a major source of PID cases and an estimated 10% of women with gonococcal infection develop PID (Holmes et al., 1980; Eschenbach et al., 1975).

In men, GC normally causes symptomatic urethritis, but a small percentage of patients will develop asymptomatic disease. *N. gonorrhoeae* binds to the urethral epithelia asialoglycoprotein receptor through the lactose terminus found on lacto-N-neotetraose (LNT) LOS and this interaction leads to a cytokine release TNF-α, IL-1β, IL-6 and IL-8 causing an inflammatory response (Ramsey et al., 1995; Harvey et al 2002). There is an incubation period of about 40 hours in which GC cannot be recovered from male patients (Schneider et al., 1995). A purulent discharge is present due to an influx of neutrophils. Challenge studies show that men infected with Opa- GC shift to Opa+ phenotype (Schneider et al., 1995). This suggests the importance of Opas in urethritis in men. Opas are involved in the interaction with neutrophils by binding CEACAMs 1, 3 and 6. This interaction can also lead to the oxidative burst of PMN which is elicited by CEACAM3 (Gray-Owen et al., 1997; Nagel et al., 1993) and recently it has been shown that CEACAM1 and CEACAM6 potentiate this response (Sarantis & Gray-Owen, 2012). LOS sialylation has been suggested to decrease the infectivity of GC (Schneider et al., 1996) but it might increase the survival rate of GC inside neutrophils (McLaughlin et al., 2012).

In women, GC causes infection of the upper and the lower genital tract and asymptomatic disease is present in about 50 to 80% of infected women. About 45% of women with lower tract infection will develop PID. Piliated and non-piliated GC attaches to non-ciliated cells of the fallopian tubes, but the most damage is seen by the sloughing of ciliated cells (Stephens et al., 1982).

1.2 Antigenic determinants

The major antigenic determinants of *N. gonorrhoeae* undergo antigenic or phase variation and they are LOS, Porin, Pili and Opas.

LOS is a major glycolipid in the cell wall of *Neisseria* species. LOS is involved in the invasion of genital epithelia as well as resistance to the immune response. LOS is composed of lipid A attached to two 2-keto-3-deoxy-mannooctulosonic acid and two heptose residues linked to the first KDO residue (Yamasaki et al., 1991). The heptose residues are linked to the α and β oligosaccharide chains (Preston et al., 1996). The LOS structure can differ among strains but can also shift during infection in one strain depending on the environment (Schneider et al., 1991; Burch et al., 1997). Men challenged with GC strain MS11mk variant A, which has an alpha-chain lactosyl group, shed bacteria that had shift to variant C, which has higher molecular weight paraglobosyl LOS (John et al., 1999). Variant C GC were only isolated after the presence of urethral discharge (Schneider et al., 1995; Schneider et al., 1996; John et al., 1999). MS11 mk variant C bacteria were also found in men that had acquired gonorrhea naturally and variant A is the precursor of LOS structures associated with virulence (Schneider et al., 1991; Kerwood et al., 1992). LOS phase variation can result in expression of LOS structures that are identical to human glycolipids. The terminal LNT of GC is identical to the terminal tetrasaccharide of paragloboside, a glycosphingolipid in human blood (Mandrell et al., 1988). GC expressing this structure can bind human sperm by binding the asialoglycoprotein receptor (ASGP-R) and also can induced the entry of GC into infected urethral epithelial cells (Harvey et al., 2000 and 2001). In addition, LOS expressing LNT promote invasion of GC into cervical epithelia in the absence of Opa

expression (Song et al., 2000). Sialylation provides GC with an important step in invasion of mucosal cells and evasion of complement killing (van Putten 1993; Apicella et al., 1990; Wetzler et al., 1992). Some strains of GC are capable of binding galectin-3 which is a lectin expressed in fallopian tubes and other genital epithelia (John et al., 2002). The lipid A portion of LOS has been implicated in inactivation of the alternative pathway of complement. LOS serves as the acceptor of CR3 express by cervical epithelial cells which leads to the conversion to iC3b and subsequent invasion into cervical cells with the interaction of pili and porin (Edwards et al., 2002).

Porins are integral outer membrane proteins and comprise 66% of neisserial outer membranes (Johnston and Gotschlich, 1974). Each porin is composed of 3 polypeptides, each forming β-pleated sheets and the trimeric structure of porins (van der Ley et al., 1991; Derrick et al., 1999). Porins serve as pores on the bacterial outer membrane that allow the exchange of ions between the bacteria and their environment (Young et al., 1983). Porins in GC are located in two different loci *porA* and *porB*; the latter one having two alleles: *PorB1A* and *PorB1B* of about 35KDa each (Gotschlich et al., 1987). The *porA* gene is expressed in MC but not in GC, suggesting that it did not serve a purpose in the colonization of the urogenital tract (Feavers and Maiden, 1998). The *por* gene has a heptameric repetitive sequence (CTGTTTT) following the termination codon (Gotschlich et al., 1987; Butt et al., 1990). Por1B has been shown to cause oxidative burst of neutrophils while inhibiting other PMN functions such us actin polymerization and phagocytosis (Bjerknes et al., 1995; Lorenzen et al., 2000). Also, a population of GC expressing the PorB1B allele can resist cellular proteolytic activity (Weel et al., 1991).

Ayala et al (2002, 2005) suggested how PorB1B and pili of GC induce calcium release. Porin stimulates a Ca^{++} response after 2 minutes of infection and this response in turn stimulates a pilus-induced release of Ca^{++} from intracellular storage after 10 minutes. Calcium stimulation leads to endosome and lysosome exocytosis translocating LAMP-1 protein to the cell membrane where it can be cleaved by the GC IgA protease. MÜller et al (1999) reported that calcium influx after infection with GC causes apoptosis and also causes arrest of the phagosome maturation (Mosleh et al., 1998). PorB is imported to the mitochondria causing it to lose its membrane potential (Müller et al., 2000, 2002). It has been shown that the allele PorB1B serves as the binding site for complement proteins C3b and C4bp making GC resistant to the killing action of serum (Ram et al., 2001) and the binding occurs between loops 4 and 5 of the porin protein (Ram et al., 2001). PorB1A sequence is either absent or diverges from PorB1B (Lewis et al., 2008). Instead PorB1A is able to bind the C4 binding protein which inhibits the classic and lectin pathways of complement (Jarva et al., 2007). PorB1A has also been shown to bind factor H which decreases serum killing (Ram et al., 1998). GC serotype PorB1A has been associated with disseminated disease and resistance to killing by serum while serotype PorB1B is associated with local urethral infection (Cannon et al., 1983; Morello and Bohnhoff, 1989). This invasion mechanism is dependent on the level of phosphate and affects the binding of GTP to the GC porin (van Putten et al., 1998). Low levels of phosphate are necessary for GC entry into epithelial cells and it is independent of Opa invasion (Kühlewein et al., 2006). PorB1A can bind the eukaryotic receptor Gp96 (heat shock glycoprotein) which allows adherence but not invasion into epithelial cells. Porin must come into contact with the SREC (scavenger receptor) to allow invasion of

epithelial cells (Rechner et al., 2007). PorB1B enhances the expression of the anti-apoptotic protein Bfl-1 and activation of the *Bfl-1, cox-2,* and *c-IAP-2* genes depends on the activation of NF-ƙβ (Binnicker et al., 2004).

Pili are filaments that extend from the body of GC (Swanson et al., 1971) and have been associated with urethral infection (Swanson et al., 1973). Pili are necessary for natural transformation of GC and adherence to epithelial cells (Swanson, 1973; Mehr and Seifert, 1998). Piliated GC also resist phagocytosis by neutrophils (Punsalang and Sawyer, 1973). The main subunit of pili is pilin (PilE), an 18 - 22 Kda polypeptide and this subunit self-polymerizes to form the pili filament. Proteins PilC and PilE have been reported as adhesins that allow for the attachment of GC to epithelial cells (Rudel et al., 1992, 1995). Phase variation in these proteins allows for the different cell tropism that GC encounters (Jonsson et al., 1994). The *pilC* loci undergo phase variation. Lack of one of the *pilC* loci does not have an effect, but lack of both loci prevents transformation. PilT is a protein involved in twitching motility as well as pilus retraction. PilC and PilT have been shown to have antagonistic roles. When PilC proteins are not present, suppression of PilT restores piliation (Wolfgang et al., 1998). PilC regulates PilT fiber retraction (Morand et al., 2004). Pili interaction with epithelial cells triggers release of cytosolic Ca^{2+} from intracellular storage and high levels of calcium are needed for effective binding of pili (Kallstrom et al., 1998). In experiments to study the passage of GC through a monolayer, it was found that over 50% of bacteria recovered after 25 hrs were non-piliated although piliated bacteria had been added initially (Ilver et al., 1998). Pili activated PI_3 Kinase in A431 epithelial cells leads to the production of the secondary

lipid messenger PIP_3 (Lee et al., 2005). It has been shown that pili activate Erk leading to the down regulation of proapoptotic proteins Bim and Bad (Howie et al., 2008). CD46 receptor protein has been suggested as a receptor for pili, but a specific neisserial molecule has not been found to bind this receptor.

1.3 Colony Opacity Associated Proteins

Opas, previously known as P.II proteins in GC and P.5 proteins in MC (Swanson, 1981) are believed to promote the intimate attachment of GC to epithelial cells and neutrophils in the host. Opa interactions are not required during the initial contact with host cells and their expression changes depending on environmental conditions in the host (Weel et al., 1991). Opacity variants of GC colonies where first observed under light microscopy and these colonies showed a dark or opaque appearance compared to others one grown on the same translucent medium that were transparent or light in appearance. It was also noted that the GC around the opaque colonies were composed mostly in clumps or chains while the light colonies had more individual diplococci around the colony. The colony variants were independent from the presence of pili or not and this property also appeared to be due to a protein since opaque colonies were more susceptible to be killed by trypsin (Swanson J., 1978). These opacity differences were also observed in colonies recovered from males and females. GC recovered from the cervix presented predominately transparent colonies while GC recovered from males produced more opaque colonies (James, J. and Swanson, J., 1978). SDS-PAGE showed that opaque colonies presented proteins missing from transparent colonies; these proteins ranged from M.W. 24 to 30 Kdal (Swanson, J., 1978). After each passage there was a

change in the opacity variation of the colonies and the rate of change between opaque and transparent colonies was determined between 10^{-3} and 10^{-4} per CFU per generation (Mayer, L., 1982; Swanson, J., 1982). Transparent colonies seem to be less virulent than opaque colonies, when tested in chicken embryos (Salit, I., and Gotschlich, E., 1978). These results were also observed with strains recovered from patients with DGI and localized gonococcal infections. Strains with opaque colonies are found more predominately in patients with localized infections, while patients with disseminated DGI presented strains that had transparent colonies (Martin et al., 1986).

1.3.1 Genetic basis of Opa expression

Opas undergo phase and antigenic variation. All *opa* genes are monocistronic, each with its own promoter (Stern et al., 1984). These genes code for integral heat modifiable proteins. *opa* genes are constitutively transcribed. Each *opa* gene locus has a repetitive sequence known as the Coding Repeat (CR). The CR is composed of a pentameric pyrimidine sequence CTCTT located within the 5' terminal ORF; between the ATG codon and the mature Opa protein codons. (Stern et al., 1986; Murphy et al., 1989). Figure 1 shows that the number of CR units places the starting codon ATG in frame or out frame to produce either a functional protein or a truncated, nonfunctional peptide (Muralidhara et al., 1987). The number of CR controls *opa* gene expression in the gene sequence. Parental strains with for example 13 CR would produce derivative strains with 12 or 14 CR. The CR can range from 7 to up to 57 units and these units code for only three different amino acids-phenylalanine, serine and leucine (Muralidhara et al., 1987; Meyer et al., 1989). The open reading frame located before the amino terminal

codon is part of the Opa protein as well. Eight of the eleven *opa* genes have conserved promoter regions with a Pribnow box at -10 positions (TATAATC) and the -35 sequence (TGTTGAA); *opa* genes contain the gonococcal uptake sequence upstream at the 3' termini (Bhat et al., 1991). A colony of GC produces a heterogeneous population of opacity mRNAs because all *opa* genes are being transcribed at low levels even when out of frame (Belland et al., 1989).

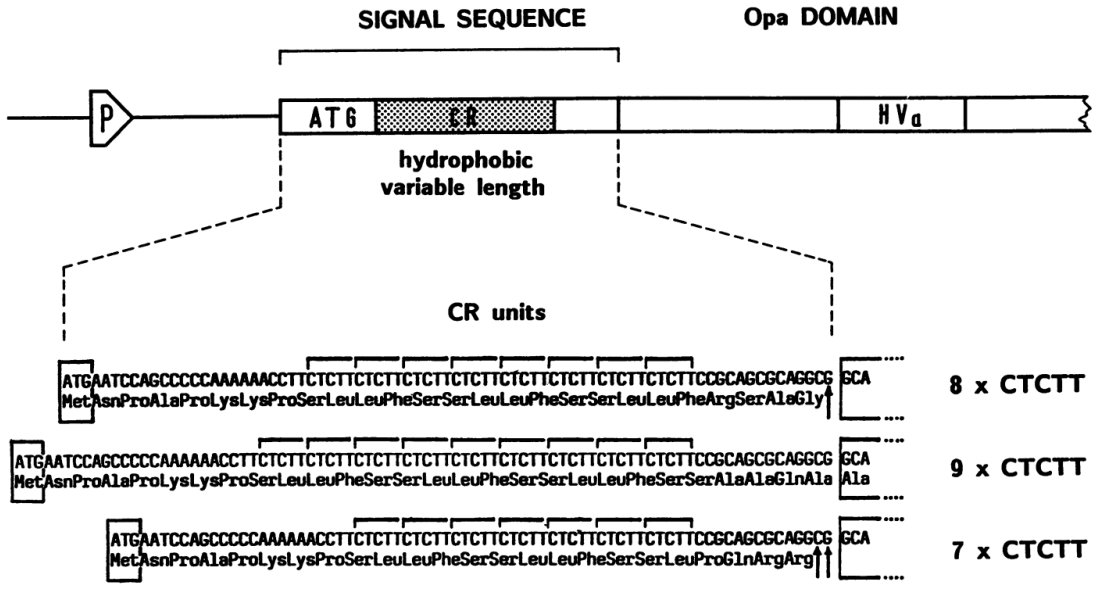


Figure 1. Diagram of *opa* genes regulation by pentameric CTCTT repeat. The number of coding repeat units (CR) place the *opa* gene in frame or out of frame. (Meyer and van Putten, 1989).

The regulation of *opa* is due to a slipped-strand mispairing mechanism in which deletion or addition of complete CR units will cause a frameshift in the *opa* transcript (Connell et al., 1987; Murphy et al., 1989; Belland et al., 1989). Therefore, the number of Opas produced in a single cell is limited to the number of *opa* genes present and depends on the translational regulation of each gene. Phase variation is not dependent on RecA (Belland et al., 1989; Murphy et al., 1989). All *opa* genes share a conserved

region and most differences rely on two hypervariable regions that code for hydrophilic amino acids with some charged amino acids (HV_1 and HV_2) (Stern et al., 1986; Connell et al., 1988). Stern et al (1986) isolated 2 variant *opa* genes from the locus opaE1 and they showed that both variants had the same ORF and the alterations are due to 2 hypervariable regions around the central region of ORF. The *opa* genes were identified using probes to common sequences of the known *opa* genes. Eleven different bands, which correspond to 11 *opa* loci were identified in GC MS11 and named A through K in order of increasing electrophoretic mobility by Southern blotting (Bhat et al., 1991). In addition, 11 loci were found in strain FA1090 with at least six different versions of HV_2 and five versions of HV_1 regions. These hypervariable regions can combine in nine different ways (Connell et al., 1990).

Opacity proteins are basic and have an isoelectric point between pH 9.0 and 10 (Blake, M., and Gotschlich, E., 1984). Malorny et al (1998) predicted a two- dimensional model of the Opa protein structure. They suggested a β-barrel with eight transmembrane strands with four surface exposed loops. Three of these corresponding to variable regions: Semi-variable (SV), HV_1 and HV_2 regions. The loop L4 is highly conserved and short. Variability is also seen in regions T_B and T_C facing the inner side of the membrane. Opas traverse the outer membrane 8 times with four hydrophilic loops on the cell surface and the terminants at the inner face of the outer membrane. Studies performed with MC provided evidence that the secondary structure of Opas has a high content of B-strand formation (De Jonge et al., 2002).

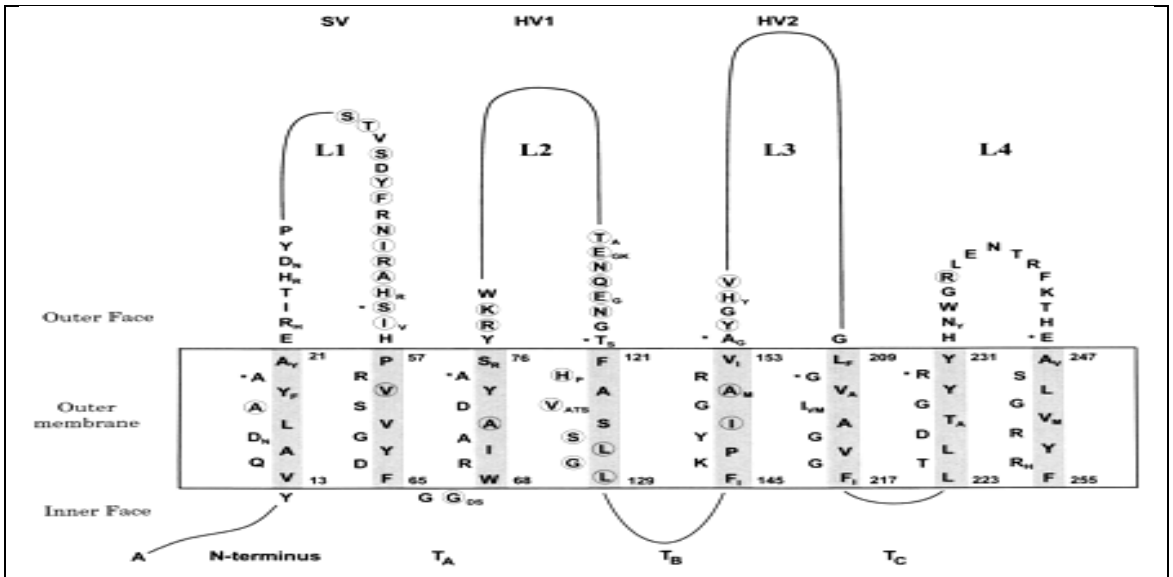


Figure 2. Predicted two-dimensional structure of Opas. Opas have four extracellular loops. Sequence differences are concentrated in the SV, HV1 and HV2. Conserved amino acids are shown (Malorny et al., 1998).

Expression levels of the different *opa* promoters lead to differences in the rates of phase variation (Belland et al., 1997). The evolution of the *opa* gene family is believed to be the result of gene duplication, gene replacement and partial non-reciprocal recombination. (Bhat et al., 1991; van der Ley et al., 1988). It has also been suggested that horizontal gene exchanges have occurred among different MC and GC and these changes have increase the variability of the *opa* genes (Hobbs et al., 1994). This variability does not occur to produce new and distinct hypervariable regions (Bhat et al., 1991). Makino et al (1991) showed that cell invasion in the Chang cell line was determined by the phase variation of the Opas. They also observed higher levels of invasion when a specific Opa was present.

1.3.2. Opa and cell surface receptors

Opas have been found to recognize only two different classes of cellular receptors: heparan sulfate proteoglycans (HSPG) and carcinoembryonic antigen cellular adhesion molecules (CEACAM).

1.3.2.1. Heparan Sulfate Proteoglycans

HSPGs are composed of a core protein covalently bound to polyanionic heparan sulfate glycosaminoglycans (GAG) (Duensing et al., 1999; Jalkanen et al., 1985). They have a ubiquitous distribution on the plasma membrane and the extracellular matrix binding to different enzymes and proteins such as collagen and fibronectin (Laterra et al., 1983). HSPGs can be either intercalated into the plasma membrane through the protein core or bound to the cell surface through the GAG chains. Heparan sulfate is a complex carbohydrate composed of a repeated disaccharide formed of an α, β1→4-linked sulfated amino sugar and uronic acid that regulates several interactions such as growth factor-receptor and proteinase - proteinase inhibitor complexes (Gallagher et al., 1986; Guido, D., 1993; Jalkanen, M., 1987; Jalkanen et al., 1985).

Syndecans are the most common transmembrane HSPGs in animal cells. There are four types of syndecans and they can be anchored to the plasma membrane through a hydrophobic amino acid transmembrane domain (Elenius and Jalkanen, 1994). Syndecans are expressed in different cell types: syndecan-1 is express in epithelial cells, syndecan-2 in fibroblasts, syndecan-3 in neuronal cells, while syndecan-4 is widely expressed (Zimmermann and David, 1999). Syndecan-1 has been shown to target the

basolateral surface of polarized epithelial cells while syndecan-4 is located in focal adhesions with integrins β_1 and β_3 (Carey, D., 1997). Syndecans can act as co-receptors with integrins in interactions between cells and ECM adhesions (Elenius and Jalkanen, 1994). It has been shown that syndecan-4 can bind to protein kinase C (PKC).

Non-piliated GC can also invade human epithelial cells by expressing Opas that use HSPG as their cellular receptor. Using Chinese hamster ovary cells deficient in heparin sulfate, researchers showed that GC strain MS11 expressing OpaA was not able to adhere to the cells as efficiently as wild type CHO-K1 cells. Also, this interaction was inhibited when adding exogenous heparin sulfate to Chang and CHO-K1 cells (Chen et al., 1995; Van Putten and Paul, 1995). Attachment of $^{35}SO_4$-labeled purified cell receptor to GC demonstrated conclusively that syndecan HSPG are the receptors for invasion of OpaA expressing GC into human mucosal cells (Van Putten and Paul, 1995). Overexpression of syndecan-4 and syndecan-1 enhanced the uptake of GC into HeLa cells, also when using transfected HeLa cells that expressed these HSPGs (Freissler et al 2000). This uptake was blocked when using PKC inhibitors, and the binding site of syndecan-4 with PKCα is in its intracellular domain, further demonstrating that deletions in the intracellular domain of syndecan-4 that interacts with PKC and Phosphatidylinositol 4,5-bisphosphate (PIP_2) block the uptake of GC (Freissler et al., 2000).

Opas A and C of strain MS11 recognize HSPG and promote invasion into human Chang conjuctival cells (Makino et al., 1991; van Putten et al., 1997), HEC-1B human

endometrium carcinoma, and ME-180 human cervix carcinoma cells (Kupsch et al., 1993). Grant et al (1999) showed that the binding domain between OpaA and HSPG is mainly through HV-1, although HV-2 and SV also play a role. This was done by making deletions of the four exposed loops of Opa protein and exposing them to heparin binding. HV-2 had a reduction of about 40%, while HV-1 deletion had a dramatic reduction to background levels.

For some epithelial cells, HSPG is not enough to allow uptake of the Opa_{HSPG} GC. Other extracellular protein factors are needed to enhance this interaction and allow for an alternative pathway of invasion. Vitronectin, which is known to bind integrin receptors, induces the entry of Opa_{HSPG} GC into HeLa epithelial cells and CHO cells by binding to sulfated polysaccharides (GAG) (Gomez-Duarte et al., 1997; Duensing T., and van Putten J., 1997). The interaction between GC and vitronectin is dependent on the integrins $\alpha_v\beta_5$ (Dehio M. et al., 1998). Fibronectin mediates the uptake of bacteria into HEp-2 cells and this is facilitated by the interaction with HSPG. The 30KDa N-terminal part of fibronectin is the one that associates with GC. Fibronectin interacts with HEp-2 cells through the RGD motif facilitated by $\alpha5\beta1$ integrin (van Putten et al., 1998). Internalization of Opa-HSPG GC has been shown to involve the recruitment of F-actin (Grassme et al., 1996). Opa-HSPG Signaling studies have shown that PKC inhibitors can block the uptake of HSPG beads, and internalization can be completely blocked when using actin polymerization inhibitors (Dehio C. et al., 1998). Inhibition of gonococcal invasion into HeLa cells was observed using serine/threonine kinase inhibitors (Dehio et al., 1998).

Invasion of GC into Chang cells has been shown to required PC-PLC activation and this internalization produces consumption of PC and release of DAG. Using ASM deficient and reconstituted fibroblasts, it was confirmed that invasion of GC also depends on the activation of ASM correlating with the release of ceramide, (Grassme et al., 1997). which can promote the activation of transcription factors (Schütze et al., 1994).

HSPG interactions with Opas modulate the adherence and invasion to some cell types. Polarized cells express HSPG receptors on their basolateral surface where they can adhere after transmigration of epithelial cells. GC binding to vitronectin and fibronectin provides additional points where bacteria can interact with host cells.

1.3.2.2 Carcinoembryonic Antigen-Related Cell Adhesion Molecules as receptors for Opa

CEA family belongs to the immunoglobulin superfamily and they can be associated to the cell surface through a GPI anchor or they can be transmembrane (Obrink, 1997). Ten out of the 11 Opas expressed by GC recognize CEACAMs as their receptor on epithelial, endothelial and immune cells. CEA antigens contain a single IgV-like N-domain in two β-sheets forming eight β strands that are heavily glycosylated by carbohydrate chains on the asparagine residues (Hammarström, 1999). Cell surface CEACAMS can be attached to the cell surface in two different ways: CEACAMs 1, 3 and 4 contain a hydrophobic transmembrane domain with a cytoplasmic domain, while CEACAMs 5, 6, 7 and 8 are attached through a glycosyl phosphatidyl inositol moiety

(Kuespert et al., 2006; Hammarström, 1999). CEACAMs can be widely expressed; CEACAM 1 and 6 are expressed on epithelial cells and leukocytes, while CEACAM 3 is expressed exclusively on neutrophils (Frängsmyr al., 1999). CEACAM 5 is present in columnar epithelial and goblet cells in the colon, as well as, squamous epithelial of the cervix, stomach and tongue (Hammarström, 1999) and also in several carcinomas cell lines (Blumenthal et al., 2007). In general, CEACAM family members are expressed in normal human colonic epithelia and neutrophils as well as in several tumor cells (Hammarström, 1999; Frängsmyr 1999).

CEACAM molecules are involved in several functions: cell adhesion through homo and heterophilic binding to molecules of adjacent cells (Obrink, B., 1997), they are possibly involved in the regulation of taurocholate and bile acid transport (Obrink, B., 1997). CEACAM can act as microbial receptors; they have been shown to play a role in interaction of *E. coli* and *Salmonella* (Leusch et al., 1991; Sauter et al., 1993) as well as to mouse hepatitis virus (Holmes 1993). CEACAM1 has been reported to be a target receptor for outer membrane proteins of *Moraxella catarrhalis* UspA protein and *Haemophilus influenzae* P5 proteins (Hill and Virji, 2003; Hill et al., 2001). By using a recombinant polypeptide of the CEACAM molecule, it is possible to block the interaction between epithelial cells and pathogens such as *Moraxella*, *Neisseria* and *Haemophilus* (Hill, et al., 2005). Activation of CEACAM by Opa-mediated interactions leads to the activation of CD105, which is a TGF-β1 receptor, improving attachment of the host cell to the ECM avoiding detachment of infected epithelial cells (Muenzner et al., 2005).

Interactions between GC and neutrophils have been described to be independent of pili expression (Swanson et al., 1974). Opas have been reported to promote interaction with neutrophils in an opsonin-independent manner (Rest et al., 1985; Fischer et al., 1988; Naids et al., 1991) and to enhance the respiratory burst carry out by neutrophils as a bactericidal response (Naids and Rest, 1991; Belland et al., 1992). In experiments demonstrating phagocytic killing of GC, it was concluded that GC expressing Opas interact with human neutrophils (Fischer and Rest, 1988). Using purified Opas attached to liposomes, it was shown that the HV_2 region of Opa partially interacts with neutrophils (Naids et al., 1991). It was demonstrated that CEACAM 3 is the receptor for the attachment of Opa-expressing GC on neutrophils by using HeLa transfected cells and PMN lysates. This receptor also promoted the invasion of Opa-expressing GC into neutrophils (Chen and Gotschlich, 1996). Opa recognizes 108 amino acids in the N-terminal region on CEACAM 3 (Gray-Owen et al., 1997). Opa-expressing GC and neutrophil interaction leads to a signaling cascade in which tyrosine kinases Hck and Fgr are phophorylated which control the activation of the GTP binding protein Rac1 and Cdc42 activating the serine/threonine kinase PAK and Jun-N-terminal kinase (Hauck et al., 1998; Billker et al., 2002). This interaction begins with the phosphorylation of a functional ITAM molecule on CEACAM 3 that leads to the influx of calcium and phagocytosis of opa-expressing GC (Chen et al., 2001; McCaw et al., 2003). It has been shown that Syk and PLC are required for the phagocytosis of Opa-expressing GC by neutrophils (Chen et al., 2001). The phosphorylation of CEACAM 3 also leads to the nucleation of actin filaments to allow uptake of GC (McCaw et al., 2003). CEACAM3 Transfected 293T cells change the distribution of Rac by recruiting around GC, and there

was an increase in Rac-GTP in these cells (Schmitter et al., 2004). This interaction produces the killing of GC by oxidative burst (Schmitter et al., 2004) and interaction with CEACAM3 was also observed with *Haemophilus* and *Moraxella* (Schmitter et al., 2004).

CEACAMs (CD66) antigens are found to be receptors for Opas in MC and GC using transfection of COS cells with CD66 cDNA (Virji et al., 1996a, b). Co-precipitation of three different constructs of CEACAM1 with MC ligand determined that the N-terminal of CD66 is bound specifically by Opas, also GC recognized N-terminal domain expressed by *E. coli* (Virji et al., 1996a; Bos et al., 1998). It was demonstrated in transfected HeLa cells, Opa not only interacts with CEACAM3 and CEACAM1 but also CEACAM5 (Chen et al., 1997). Since different members of the CEACAM family can act as receptors for Opas, Popp et al., 1999 wanted to test if these included CEACAM4 and CEACAM7 by using transfected CHO cells. They concluded that CEACAM4 and 7 are not bound by Opas even though they share homology sequence with other CEACA molecules. Using different residues of the N-terminal of CEACAM5, it was found that Opa binds to the exposed loops of the GFC face of CEACAM that it not bound to carbohydrates, more specifically the βsheet βC ant the loop C-C' on residues 31-42 where the is a serine on residue 32 (Bos et al., 1999; Popp et al., 1999). Making deletions of the exposed loops on Opa, it was possible to determine which specific domain binds to CEACAM receptors. Deletion of the HV regions completely blocked binding to CEACAMS, while deletion of the SV region did not alter this binding. This interaction can only happen when there is a correct combination of HV regions demonstrating that this interaction is of high affinity (Bos et al., 2002; de Jonge et al., 2003). Opa C interacts

with both types of receptors mediating adhesion to CEACAMs 1, 3, 5, and 6 as well as HSPG (Gray-Owen et al., 1997; Chen et al., 1997) correlating that purulent exudates from patients consist of mostly neutrophils and that they are associated with GC (Ward et al., 1972). Opa variants interact differently with members of the CEACAM family Opa B, C, G and I recognize CEACAMs 3 and 6 while all Opas interact with CEACAMs 1 and 5 except for OpaA, which interacts with HSPG (Bos et al., 1997; Chen et al., 1997).

Early studies showed that OpaC is involved in invasion of epithelial cells, and that OpaH loci mediates invasion in HEC-1-B cells (Waldbeser et al., 1994) demonstrating that more than one *opa* can contribute to epithelial cell invasion. Different interactions between Opa and CEACAM receptors confer variability in tissue tropism and disease manifestation in different gonococcal infections. Several infection assays have demonstrated that specific Opas have induced different responses in host tissues. For example, Opas that bind CEACAM 1 are responsible for inducing a chemiluminescence's response in neutrophils and colonization of endothelial cells (Gray-Owen et al., 1997). In addition, Chen et al (1997) showed that Opas display different ability in adherence and invasion of HeLa cells and that OpaC interacts with both HSPG and CEACAM receptors.

Interactions between Opa-expressing GC and epithelial cells have been shown to activate signaling cascades to promote the expression of CEACAM1 on endothelial and epithelial cells (Muenzner et al., 2001; Muenzner et al., 2002). After prolonged infection of human umbilical vein endothelial cells (HUVEC) with Opa-expressing GC, it was shown that CEACAM1 molecules promoted internalization of GC in the previous

unstimulated attachment of GC on the endothelial cells. Furthermore, this expression of CEACAM1 was not due to the activation of TNFα. Infections of endothelial cells lead to the activation of NF-κB after 45 minutes post-infection (Muenzner, et al., 2001). Using primary epithelial ovarian cells (HOSE) that express CEACAM1, it was found that infection with GC upregulated the expression of CEACAM after 2 hours post infection and that this upregulation also requires activation of NF-κB. It was also noted that invasion into HOSE cells was very low compared to the attachment of GC to HOSE cells (Muenzner et al., 2002). Demonstration that class I PI 3-kinase activity is required for the uptake of GC into epithelial cells was done using CEACAM3 HeLa transfected cells knowing that this receptor has an ITAM domain leading to the activation of Phosphatidylinositol (3,4,5)-triphosphate (PIP_3) at the sites of bacterial attachment indicating that PIP_3 mediates internalization of GC into cells. Another PI_3K product is involved in the survival of GC in epithelial cells, PI3P, which is localized in endosomal compartments and was accumulated in mature phagosomes containing GC (Booth et al., 2003); phagosomal acidification is required for gonococcal killing (McCaw et al., 2004).

McCaw and coworkers (2004) used HeLa cells transfected with each of the CEACAM receptors to study the differences in interactions with GC. Gonococcal association and uptake differ depending on the CEACAM receptor expressed on epithelial cells; association was highest in CEACAM5 HeLa cells, while internalization was highest with CEACAM3 and lowest CEACAM5. Tyrosine kinase inhibitors were used to test the role of these kinases in uptake of GC into each of the HeLa transfected lines, confirming that only in CEACAM1 and 3 expressing epithelial cells kinases have a

role in this uptake. They also confirmed that actin rearrangements are seen during uptake in CEACAM3 expressing cells, while not in the other ones (McCaw et al., 2004). Entry into CEACAM5 cells was reduced when the GPI anchored was cleaved suggesting a "zipper" mechanism of entry. It has been demonstrated using cholesterol chelators that in GPI anchored CEACAM6 cholesterol-rich membrane microdomains mediate the uptake of Opa-expressing GC (Schmitter et al., 2007).

More recently, the importance of Opas in the suppression of the host immune system has been shown. *N. gonorrhoeae* OpaC binding to CEACAM1 receptors suppresses the activation and proliferation of CD4+ T lymphocytes (Boulton and Gray-Owen, 2002). Also, the role of Opas in the interaction with B cells has been demonstrated. Interaction of CEACAM1-expressing Human B cells inhibited antibody production and promotes cell death of DT-40 cells through the BTK (Pantelic et al., 2005). Therefore, Opas are important for infection to occur and for evasion of the immune system.

HeLa (cervical carcinoma) and HEC-1-B (endometrial carcinoma) cell lines that are commonly used for GC-host interactions are negative for the expression of CEACAMs. ME-180 (cervical carcinoma) cell line produces high levels of CEACAM5 and CEACAM6 but no CEACAM1 (Swanson et al., 2001; Muenzner et al., 2002). T84 cells express CEACAM1, CEACAM5 and CEACAM6 (Wang et al., 1998).

1.4 Polarization of epithelial cells and apical tight junction formation

Epithelial and endothelial cells are polarized to form barriers between tissues and organs. Polarized cells are characterized by cell to cell adhesions and connect to the extracellular matrix (ECM) through actin and by two compartments: an apical domain and a basolateral domain. Cell to cell adhesions consist of desmosomes, adherent junctions, GAP junctions and tight junctions. Tight junctions are the most apical cell-cell junctions forming a paracellular barrier between cells and the environment and are composed of several transmembrane proteins that bind peripheral proteins forming a network that interact with the cytoskeleton. Three main transmembrane proteins are claudins, ocludins and JAMs and the peripheral proteins are: zonula occludens (ZO-1, ZO-2 and ZO-3), MAGI proteins and cingulin.

Ocludin is a 60 Kdal protein that interacts directly with the zonula occludens proteins and indirectly with JAMs and the cytoskeleton (Wittchen et al., 1999; Bazzoni et al., 2000). Occludins are part of the TJ strands (Furuse et al., 1993). Overexpression of occludin has been shown to increase paracellular permeability of tight junctions and it is also involved in the formation of the intramembrane barrier (Balda et al., 1996). C-terminal cytoplasmic domain of occluding mediates basolateral expression of a membrane protein as well as endocytosis, while glycosylated occludin accumulates on the basolateral membrane, suggesting that biogenesis of the tight junction occurs on the basolateral membrane (Matter and Balda, 1998). Claudins are involved in the barrier functions by promoting calcium independent cell adhesions (Furuse et al., 1998a). Impaired paracellular permeability to calcium and magnesium is observed in claudin mutations (Hou et al., 2005) showing that claudins are integral components of the TJ

strand (Furuse et al., 1998b). JAM proteins belong to the IgG superfamily. JAM are believed to have a function in the barrier function (Liu et al., 2000) but are not part of the TJ strand.

The peripheral proteins zonula occludens are located at the submembranous region of the tight junctions serving as scaffolding proteins that interact with both the TJ strands and the actin cytoskeleton. The first isoform identified was ZO-1 (Stevenson et al., 1986). Experiments with ZO-1 knockouts present a delay in TJ formation suggesting that ZO-1 has an important function in the assembly of TJ (Umeda et al., 2004). ZO-1 localizes to the nucleus and binds ZONAB during TJ maturation (Balda et al 2003). ZO-1 binds the other components of the zonula occludens proteins (Wittchen et al., 1999). MAGI (membrane associated guanylate inverted) proteins localize to the TJ and also bind β-catenin and E-cadherin in the adherens junctions (Ide et al., 1999). Cingulin localizes to the TJ and interacts with ZO proteins, JAM and actin (Cordenonsi et al., 1999; D'Artri and Citi 2001). In the formation of cell polarity, two functional complexes have been associated: the CRUMBS (CRB) complex and the partitioning defective (Par) complex. The CRB complex is composed of CRB3, PALS1 and PATj proteins and is localized in the apical membrane and subapical region. Overexpression of CRB3 leads to the delay of TJ formation and disruption of cell polarity (Roh et al., 2003). Loss of PALS1 also leads to a delay in TJ formation (Straight et al., 2004). The PAR complex consists of the proteins Par3, Par6 and αPKC. It was demonstrated by Wang et al., (2006) that EGFR signaling leads to the tyrosine phosphorylation of Par3 and this interaction depends on the activation of Src family kinases c-Src and c-Yes and allows the promotion of TJ

assembly by negatively regulating Par3 – LIMK2 which regulates cofilin phosphorylation. αPKC can directly phosphorylate JAM-A which is required for a functional epithelial barrier (Iden et al., 2012). The scaffolding adaptor GAB1 promotes the phosphorylation of Par3 by Par1, therefore controlling the amount of the Par complex by releasing Par3 from the membrane to limit access to Par6 (Yang et al., 2012).

1.5. Pathogen movement across epithelial barriers

Pathogens must overcome the epithelial barrier which is the first line of defense to gain access to other tissues and cause invasion. Some pathogens are capable of producing toxins that disrupt these cell adhesins to invade the subepithelial tissues through the paracellular route. *Porphyromonas gingivalis* produces gingipain which is a cysteine protease with specificity for lysine or arginine peptide bonds degrading E-cadherin in adherens junction (Katz et al., 2000 & 2002). *Porphyromonas gingivalis* produces an increase in TEER of about 30% between 2 and 8 hours after infection, but the TEER decreased to zero at 24 hours (Groeger et al., 2010). Gingipains facilitate the hydrolysis of platelet endothelial cell adhesion molecule 1 (PECAM-1) increasing the permeability of endothelial cells and leukocyte influx (Yun et al., 2005) as well as inducing apoptosis in endothelial cells (Sheets et al., 2005).

The bacterium known for causing gastrointestinal ulcers, *Helicobacter pylori*, produces CagA which is a product of the type IV secretion system *cag* PAI. CagA can disrupt tight junctions by interacting with several different components of these cell-cell junctions and it recruits ZO-1 and JAM proteins to the site of bacterial entry as well as

binding the PAR1/MARK2 complex (Amieva et al., 2003; Saadat et al., 2007). CagA inhibits αPKC phosphorylation by Par1 causing polarity defects by dissociating Par1 from the membrane (Saadat et al., 2007). CagA also disrupt adherens junctions by interacting with E-cadherin, β-catenin and p-120 (Suzuky et al., 2005; Conlin et al., 2004; Weydig et al., 2007). CagA mutants are defective in colonization of the apical cell surface (Tan et al., 2009). CagA is not the only factor produced by *H. pylori*, VacA, OipA and urease (Franco et al., 2008). *H. pylori* disrupts E- cadherin and β-catenin by activating host cell calpain via TLR-2 (O'Connor et al., 2011). *H. pylori* can also disrupt claudin-4 by inducing the phosphorylation of IL-1R to activate Rho kinase (Lapointe et al., 2010).

Other pathogens like *Clostridium difficile,* the causative agent of pseudomembranous colitis produces the toxins TcdA and TcdB that can cause a disruption of the tight junctions and actin cytoskeleton by inactivating Rho GTPases (Popoff et al., 1996; Nusrat et al., 2001; Boehm et al., 2006). The opportunistic pathogen *Burkholderia cenocepacia* can cause bacteremia after crossing the respiratory epithelium by dephosphorylating occludin, a protein component of the tight junctions (Kim et al., 2005). *Campylobacter jejuni*, the food-borne bacterium causes disruption of the tight junctions, by inducing dephosphorylation and redistribution of occludin as well as the level of ZO-1 (Chen et al., 2006; Man et al., 2010). *C. jejuni* serine protease HtrA cleaves the NTF domain form E-cadherin which may be involved in transmigration of epithelial cells (Boehm et al., 2012).

Studies with MC have shown that Opas are important in bacterial colonization and translocation of epithelial and endothelial cells (Virji et al., 1993). Since not all cells that interact with GC express CEACAM receptors, invasion in these host cells must be mediated by an Opa-independent mechanism (Swanson et al., 2001). Target receptor density on host cells is important for interaction between capsulated meningococci and cell lines expressing CEACAM1. Even a small number of encapsulated bacteria can cross the blood barrier (Nassif et al., 2002) and lead to disseminated disease (Bradley et al., 2005).

A crucial step in neisserial pathogenesis is the ability of pathogenic *Neisseria* to cross cellular barriers to produce disseminated infections. Infections of fallopian tubes with GC have shown that GC attached to and entered non-ciliated mucosal cells where they replicated and later invaded sub-epithelial tissues (McGee et al., 1981). This process has been reported in meningococcal infections as well, where attachment and penetration of nasopharyngeal mucosa has been observed (McGee et al., 1983). It has been shown previously that *Neisseria* can traverse polarized monolayers of epithelial cells without disrupting the tight junctions (Merz et al., 1996). Infection assays of human ureters have been used to study adhesion and invasion of GC through stratified epithelial tissue. Single GC and groups of GC are found in intracellular locations and some are released into an intercellular position (Mosleh et al., 1997). *N. meningitidis* has been shown to be able to cross polarized cells without disrupting tight junctions and they are located inside the cells and not between the cells (Pujol et al., 1997). It has been reported that in assays using piliated GC to invade a monolayer of epithelial cells, non-piliated GC are recovered

on the lower chamber, which suggests that different antigenic properties are needed for transcellular passage of epithelial cells (Ilver et al., 1998). Another study observed that pili are not sufficient for activation of microvillus in HEC-1-B cell, but Opa are also required (Griffiss et al., 1999). Studies of *N. gonorrhoeae* trans-epithelial migration have shown that OpaA or Opa- GC do not interact with the apical layer of T84 monolayers, but GC expressing Opas that bind CEACAM receptors interact with the apical layer (Wang et al., 1998). This has been also shown with *E. coli* expressing recombinant Opas. It has been observed that IgA1 protease produced by pathogenic *Neisseria* has a role in transcytosis of epithelial cells by altering their lysosomal content. This has been shown with mutants in the type 2 IgA1 protease, where they traversed polarized T84 cells in fewer numbers that the wild type (Hopper et al., 2000a). In addition to IgA1 protease, other factors have been found to have an effect in transepithelial migration of *N. gonorrhoeae.* Analysis of mutants generated by a bank of minitransposons showed that intracellular growth is linked to transcellular trafficking. Mutants of the *fit* locus (fast intracellular trafficker) traversed the T84 monolayer faster than the wild type and without disrupting the tight junctions (Hopper et al., 2000). Studies with *E. coli* showed that LOS, porin and Opas are important in the transcytosis process of *N. gonorrhoeae* across epithelial cells (Gorby et al., 2001). The lutropin receptor has been found to mediate the transcytosis process with aid of the ribosomal protein L12 (Spence et al., 2002).

Phase variation allows GC to express many different Opas giving GC the ability to survive in many different environments and to interact with a variety of host cells. Some of the importance of Opas has been shown in studies where male volunteers had

been inoculated with Opa- GC. These studies showed that there is a strong selection for Opa -expression in vivo because Opa^+ GC were isolated from these subjects indicating that phase variation probably occurred during the course of infection (Schmidt et al., 2000; Schneider et al., 1995; Jerse et al., 1994). This selection may not be due to receptor specificity, but it may be due to evasion of innate defenses in response to Opa as suggested by experiments using a murine genital tract infection model (Simms and Jerse, 2006). In previous studies, it was found that women with acute salpingitis presented different colony morphology of GC depending on the location. GC recovered from fallopian tubes were transparent (Opa-) while GC recovered from the cervix had a higher number of opaque colonies (Draper et al., 1980).

Another important characteristic of Opas has been suggested by their involvement in transcytosis of epithelial cells. Studies have shown that Opas binding to CEACAM receptors are involve in the process of adhering to epithelial cells, entering and then being release into the basal stromal (McGee et al., 1981; Wang et al., 1998). The use of CEACAM receptors to adhere to the apical layer of epithelial cells can aid in the invasion of these cells and subsequently passage into deeper tissues. GC uses HSPG receptors to adhere to the basolateral layer of epithelial cells and gains back access inside the cells to exit on the apical side and continue re-infecting other cells. There is a lack of accountability of GC for up to 40 hours after infection has initiated in human male challenge studies (Schneider et al., 1995). Taking this into account, the specific role of Opas in the adherence, invasion and transcytosis of epithelial cells is important to be investigated.

Figure 3 shows how GC attaches to epithelial cells via pili while Opas promote a tight adherence. GC is then able to invade epithelial cells and later exit in the subepithelial layer where they can interact with endothelial cells to cause disseminated disease.

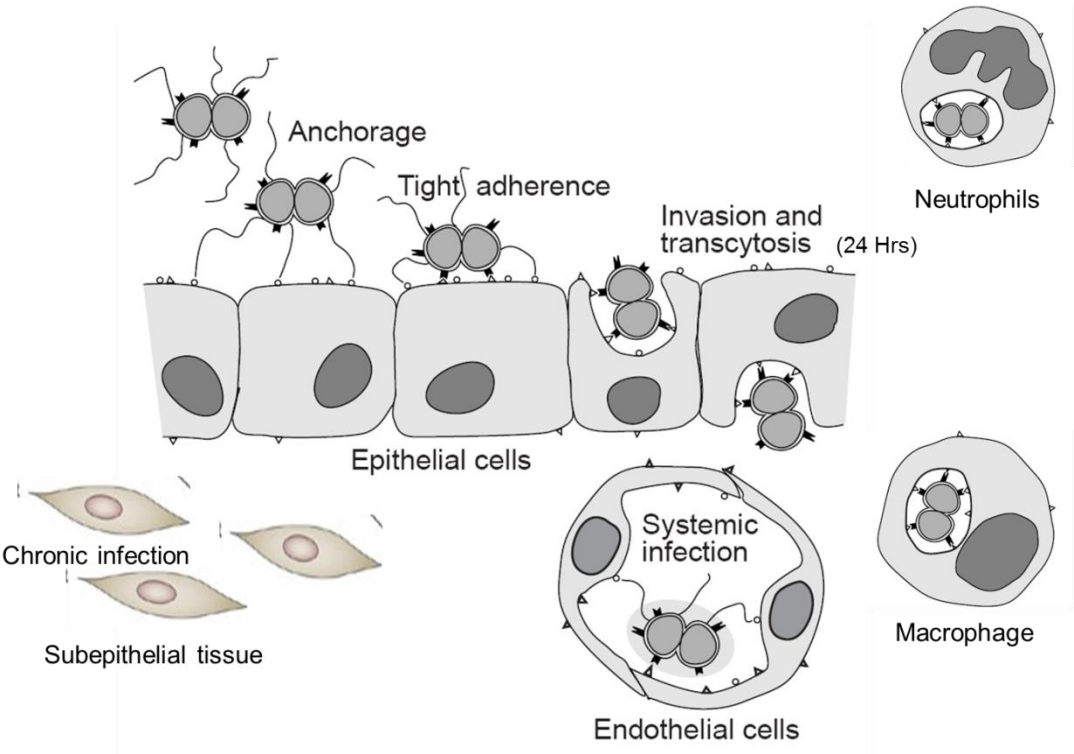


Figure 3. Diagram of transcytosis of *Neisseria gonorrhoeae* (Modified from Dehio, et al. 1998). GC interacts with epithelial cells through pili, causing retraction of the pili fiber and bringing GC closer to epithelial cells allowing interaction of Opas with host cells receptors. Opas mediate invasion and transcytosis into epithelial cells in 24 hours. GC can interact with endothelial cells causing DGI. GC interacts with neutrophils in an opsonin independent manner.

1.6 Significance and goals

The purpose of this research project was to study how phase variation of surface determinants of *Neisseria gonorrhoeae* affects transmigration across polarized epithelia. In women, GC causes asymptomatic disease in over 50 percent of women infected,

leading to chronic disease such as PID, DGI, ectopic pregnancy and infertility. Opas have been suggested to cause the intimate attachment with epithelial cells leading to invasion of these. A gonococcus is capable of expressing 11 different Opas at a specific time; therefore, it is difficult to elucidate the true role of each Opa in pathogenesis. In addition, while Opa-expressing GC are recovered from male volunteers in challenge studies as well as female cervix, in isolates of female fallopian tubes Opa negative GC are usually recovered (Draper et al., 1980) as well as joint fluid from patients with disseminated disease. This led us to suggest that to cause ascending disease in women and eventually DGI, GC must avoid expression of Opas.

The following specific aims will address the overall hypothesis of the research proposal:

Aim 1. Construct an *opa* strain of GC MS11 expressing no Opas. Due to the phase variation capabilities of Opas, it was necessary to obtain a strain that could not express any Opas. A genetic approach was used to create a MS11 *opa*$^{-}$ strain. A PCR/transformation/ spot transformation procedure was used to delete/replace the 11 *opa* genes present in strain MS11.

Aim 2. Determine the role of Opa protein expression on transcytosis of epithelial cells. This interaction involves several steps: attachment, invasion and transcytosis of epithelial cells. A T84 cell line was used to perform transcytosis experiments and address the question if GC expressing or not Opas have a different fate when interacting with

epithelial cells. I characterized the interaction of strain GC MS11Δ*opa* and Opa-expressing GC with T84 cells.

Chapter 2: Materials and Methods

Bacterial Strains

All strains used are listed in Table 1. *Neisseria gonorrhoeae* strain MS11MKC (MKC) was maintained on gonococcal media base (GCK) with 1% Kellogg's supplement at 37°C in an incubator with 5% CO_2 (White el al., 1965). *E. coli* strains were grown on Luria-Bertani medium (Sambrook et al., 1989). Spectinomycin was used at 2 different concentrations: 50 µg/ml for creation of plamids and 30 µg/ml for transformation of GC, ampicillin was used at 60 µg/ml, and X-Gal (5-bomo-4-chloro-3-indolyl-β-D-galactopyranoside) was used at 35µg/ml. A dissecting microscope was used to select for Piliated (P^+) or nonpiliated (P^-) bacteria for subsequent manipulation (P^+Opa^+, P^- Opa^-, P^+ Opa^-, P^- Opa^+) based on their light refraction phenotypes. *Salmonella* was grown on Luria-Bertani medium.

Cell Culture and Polarization of T84 Monolayers

T84 cells were obtained from the American Type Culture Collection and grown in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium containing 1.2 g/L sodium bicarbonate, 2.5 mM L-glutamine, 15 mM HEPES and 0.5 mM sodium pyruvate (ATCC), supplemented with 7% fetal bovine serum (HyClone) and 1% Penicillin/Streptomycin mixture. Cells were sub-cultured by adding 0.25% trypsin and 0.03% EDTA (Mediatech, Inc) and incubating at 37°C until detachment of the cells occurs. T84 cells ($3X10^4$/well) were seeded onto polycarbonate Transwell filters with a pore size of 3µm (Costar). Cells were propagated in culture media with fluid renewal every 2-3 days. The electrical resistance of the monolayer was measured with an

electrode (Millipore) and monolayer with an electrical resistance of >1500 Ωcm^2 was used for transcytosis assays.

Table 1. Bacterial strains used

Strain	Property	Source
N. gonnorrhoeae MS11	Wild type strain	Herman Scheneider
N. gonnorrhoeae MS11Δ2 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11Δ2,8 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3,4 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3,4,1 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3,4,1,9 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3,4,1,9,7 *opa*	Partial *opa* deletion	This study
N. gonnorrhoeae MS11 Δ2,8,5,11,6,3,4,1,9,7,10 *opa*	Partial *opa* deletion	This study
E. coli DH5αmcr	Cloning strain	Gibco Life Sciecne
pUC19	Cloning vector	New England Biolabs

Polymerase Chain Reaction

PCR reactions were performed using the Expand Long Template PCR Kit (Roche Applied Science, Germany) or GoTaq PCR System (Promega, Madison, Wisconsin).

Table 2. Primer List

Primer Name	Primer Sequence	Opa used
1F	GCGGAATTCTACATCATCTTCTCCCATAT	1
1R	CGCAAGCTTCATCGCATTACCTTTGGTTG	1
2F	GCGGGATCCAGGGCGGTGTCGAAGGCAAA	2

2R	CGCAAGCTTTCTCTAGATTCCGCATCC	2
3F	GCGGAATTCGGGGCGACGACTCGTCCAA	3
3R Redo	GCAAGCTTCCCATTGTTGCGGGAGGCTT	3
4F	GCGGAATTCAAGAAGGAATGCCCGAACCG	4
4R	CGTAAGCTTCCGCCTTGAAACACCGGGTT	4
5F	GCGAATTCCCGCCCTGTCGCCTTTAGAC	5
5R	GCTAAGCTTCGCGGATGGTGGGTTTAGGA	5
6F	CCATGCAGGCGGGAATTCAAACT	6
6R	TTTTAAGCTTGGTGTCGTCCACGGCTTTGATGGCTTTG	6
7F	GGGAAGCTTAATGCGAACGCTGCTGGCAT	7
7R	CGCGAATTCATAGAAATGACGAAATTTTAG	7
8F	GGGAAGCTTGCGTACCGAAGCTTTGTTCG	8
8R	GCAGAATTCGTTTGTTATCCCAATAATGCA	8
Opa 9 del var	GCCCAATGAGGCTTCGTGGGTT	9
9R Redo	AAAAGCATGCCCAAGCCGGTTCAACCAAAGCTGGATTAAAG	9
10F New	ATCGAATTCAAAACCGTTTTTCCCG	10
10R Redo	AAAAGCATGCCTACGCCAGCATTATTTCTACGCTCAAAGAC	10
11F	GCGAAGCTTGAGGATTTGTACGAAGAGCT	11
11R	GGTGAATTCAAAAAACCGATGGTTAAATA	11
1contF	TTTGCCATTGTTTCCTAACA	1
opa 1 cont R	ATATTTTCCTCAACAATAA	1
2 cont F	TTC TGA CGG ACA GAA AAC AGA C	2
2 cont R	TTT GGG CAA CCG TTT TAT CCG ATA A	2

opa 4 seq F	ATC TGA CAG GCG CGC AAT CCG CCC CCT CAT TTG	4
5 Cont F	GGC GGG CCA ACG CTG TAC TGG TTT A	5
5 cont R	GGC GGG CCA ACG CTG TAC TGG TTT A	5
6 cont R	ATC GCA GGC GAT ACT TTG TCT TT	6
6 next F	GAT TTC CCC CCT CCA AGG CT	6
6 next R	TGC GGC TTC CAT ATC GGC TT	6
7 cont seq F	CAA ACA GTA TTT CAG ACG GC	7
748R	CCC GGA ACC CGA TAT AAT CC	7
8 cont F	ATC GGT CAA AAT CTT CTG CCG TTT	8
9 Cont F	CAC CGC TTC CCT CAT GGT GTT	9
9 Cont R	TTC CGC GAC GGC GTG GAC GT	9
1829998F	CCG CTG TTG GTA TCC ACA TCG TTA ATC	9
Cont 183 Reverse	GCG GAC GGC GTT TTG ACA ACA GTG	9
10 cont seq F	TTC CAT TTC CTG TAA CGG GC	10
10 cont seq R	GCA ACT ACG CCA CTT GGA AC	10
11 Cont F	GGA AGA AAT CAA AAA AG	11
11 Cont R	CCG AAA TGG CTT CAA CCG GC	11
9F New	CGTGGATCCGGGGAGAGGGCTCCCCGAATT	9
Omega-Pst-F	GACCTGCAGTTGCAAACCCTCACTGATCC	all
Omega-Pst-R	CAGTCTGCAGGAGTTAAGCCGCGCCGCGAA	all
Opa5’ all	GCATCCCATAGAATCCAGCCCCAAAAAAC	all
Opa3’ all	GCATCTGGATCCGAAGCGGTAGCGCACGCCCAATGAGGCT	all

Agarose Gel Electrophoresis

Agarose gels were prepared with 1% agarose and 10 µl of 10mg/ml ethidium bromide. Gels were run at a constant voltage (100 volts) until the dye front reached the bottom of the gel (Sambrook et al., 1989).

Deletion of *opa* Genes

Each *opa* gene was amplified by PCR, cloned into a vector pUC19 and then transformed into *E. coli* DH5α by heat shock transformation (Sambrook and Russell, 2001). *opa* coding regions were deleted by PCR, and mutants were selected by inserting a spectinomycin cassette. Eleven *opa* genes from strain MS11 were cloned into pUC19, each gene was deleted and a Ω spectinomycin cassette (amplified from plasmid pHP45) was inserted into the region. *N. gonorrhoeae* strain MS11 were grown overnight on GCK media at 37°C and 5% CO_2. Piliated cells were selected to inoculate GCP broth supplemented with Kellogg's solution, 0.042% $NaHCO_3$, and 10 mM $MgCl_2$ to a density of 1 X 10^7 cell/ml. 1 µg of the SpecR deletion plasmid DNA were added and then a four hour incubation at 37°C followed to allow the deleted-cassette DNA to homologously recombine into the MS11 chromosome. Transformants were plated onto GCK + spectinomycin (30 µg/ml) selective media. After two days of incubation, transformants were selected and colony PCR allowed for detection in each mutant genome. Cells were lysed in 0.5M NaOH and then neutralized by the addition of 1.5M Tris pH 7.5. Primers designed to amplify the respective *opa* gene were used to amplify the SpecR cassette and the *opa* flanking DNA in the mutants. If the colony PCR yielded the desired product, non-selective transformation was used to replace the spectinomycin cassette with the deleted gene plasmid.

Gonococcal Transformation

Piliated GC MS11 *opa*$^-$ were resuspended to a light turbidity of 1 X 10^7 CFU/ml in 1 ml gonococcal medium base. Bacteria were incubated with plasmid DNA for 4 hours in the presence of 1 mM $MgCl_2$, 0.42% $NaHCO_3$, and 1% Kelloggs solution, in a rotary shaker at 37°C. Various dilutions were plated onto GCK plates and incubated at 37°C for 24- 48 hours.

Non-Selective Transformation (Spot Transformation)

Once the spectinomycin resistant mutant was identified, non-selective transformation (Gunn and Stein, 1996) replaced the spectinomycin cassette with the deletion plasmid. Piliated cells were suspended in GCP broth + 10 mM $MgCl_2$ at moderate density. The cells were vortexed and diluted in GCP + $MgCl_2$. Two-fold dilutions of the cells were prepared and then 5µl aliquot were spotted on a GCK agar plate. 0.1- 0.5 µg of DNA were added to each spot and the plate were incubated overnight at 37°C and 5% CO_2. The isolated colonies were re-streaked and screened for incorporation of the deletion. Colony PCR was done as described above. If the colony PCR yielded the desired product, the correct size fragment should be amplified. Once the mutant gonococci were identified, transformation with a new SpecR deletion plasmid was done; followed by the spot transformation. These procedures were repeated with nine of 11 *opa* genes present in *N. gonorrhoeae* strain MS11. Using silent mutagenesis, the next *opa* gene was amplified by PCR to produce an *opa* gene incapable of phase variation. The last *opa* gene was transformed with the plasmid containing the spectinomycin cassette, obtaining a strain in which none of the 11 *opa* genes is being expressed.

E. coli DH5α Transformation

When the ligation product was obtained transformation of *E. coli* was used to identify the clone of interest. Competent *E. coli* cells were heat shocked with 1 µg of DNA to allow penetration of DNA. LB broth was added after incubation of cells at 37°C for 2 minutes and then incubated for 30 minutes at 37°C (Sambrook and Russell, 2001). Cells were plated onto LB + X-gal + ampicillin (60 µg/ml). After overnight incubation, white colonies were selected and alkaline lysis plasmid DNA miniprep purification was done to obtain plasmid DNA. Confirmation of the correct plasmid was done through digestion of the DNA with the restriction enzymes used in the silent mutagenesis procedure and by sequencing reactions.

Southern Hybridization

Analysis was performed following a combination a protocols from Sambrook and Roche Molecular Biochemicals (2001). Chromosomal DNA of the *opa* gene mutants was isolated using Promega's Wizard Genomic DNA Purification Kit (Madison, WI). DNA was digested overnight and separated on a 1% TBE agarose gel by electrophoresis for 16 hours at 46 volts. DNA was transferred by capillary action to a pre-wetted positively charged nylon membrane (Roche, Indianapolis IN) using alkaline transfer buffer (0.4M NaOH, 1M NaCl). After overnight transfer, the membrane was removed, air dried, and UV-crosslinked for 5 minutes. The membrane was neutralized while shaking in buffer (0.5M Tris-HCl pH 7.2, 1M NaCl) for 15 min. The membrane was pre-hybridized for 1 hr at 60°C in 20 mL solution (5X SSC, 2% dry milk blocking, 0.1% N laurylsarcosine, 0.02% SDS) for every 100 cm^2 of membrane. After 1 hour, denatured probe (100 ng of

random primed DIG labeled PCR product made using Opa3' all and Opa5' all primers) was added and hybridization proceeded overnight at 60°C. After hybridization 4 stringency washes were performed. Wash 1 (1X SSC, 0.1% SDS) was done twice at room temperature for 5 min and Wash 2 (0.1X SSC, 0.1% SDS) was performed twice at 60°C for 15 min. The membrane was equilibrated in Buffer A (100 mM Tris-HCl, 150 mM NaCl, pH 7.5) for 1 min. The membrane was placed in blocking solution consisting of 1 g of dry milk in 100 mL of Maleic Acid Buffer (0.1M Maleic Acid, 0.15M NaCl, pH 7.5) for 1 hr. Anti-DIG alkaline phosphatase Ab was added to the blocking solution at a 1 : 20,000 dilution for 30 minutes. Washing was performed in buffer (0.1M Maleic Acid, 0.15M NaCl, 0.3% Tween 20, pH 7.5) twice for 15 min each. Detection buffer (0.1M Tris-HCl, 0.1M NaCl, pH 9.5) was added to the membrane for 5 min. After removal of the detection buffer the membrane was placed in a clear plastic bag and the substrate CSPD was placed on the membrane for 5 min. CSPD was removed and the membrane was incubated in the bag at 37°C for 15 min. The membrane was exposed to X-ray film for different lengths of time.

LOS Analysis

Isolated LOS was boiled for 10 minutes and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 16.5% Tris-Tricine gel (Bio-Rad, Hercules, CA). The gel was run at a constant current of 0.03 mA until the dye front reached the bottom of the gel. The gel was fixed overnight in a 40% ethanol, 5% aceic acid solution. The gel was washed for 1 hour with HPLC H_2OThe gel was oxidized in a 0.833% periodic acid for 5 minutes and washed with HPLC H_2O over 1.5 hours. LOS

was visualized by silver staining (0.42% NH4OH, 0.047 M AgNO3, 0.0225 M NaOH) (Tsai and Frasch, 1982).

LOS-Texas Red Staining Procedure

Bacteria were grown for 18 hours and diluted to a Klett of 100 (~10^9 CFU/ml). Cells were collected by centrifugation (3 ml @ 12,000 rpm for 5 min). After which, the cells were washed with Elix water and spun down @ 12,000 rpm for 5 min. 90 µl of H_2O and 10µl LOS-conjugated TR were added to the cells and incubated for 10 minutes. 900 µl of H_2O were added and htemixture was spinned down @ 12,000 rpm for 5 min. Excess LOS-TR conjugate was washed away with 1 ml H_2O. The cells were spinned down again @ 12,000 rpm for 5 min and then resuspend cells 1 ml H_2O. Two-fold dilutions were made and 1 µl was spotted on a glass slide, allowed to air dry and heat-fixed the slide. Cells were visualized by their ability to fluoresce after excitation with 505 nm light and detection at 565 nm.

Permeabilization Assays

Cell permeability was measured by performing a paracellular influx assay using horse radish peroxidase (HRP) (1µg/ml) and Fluorescein Isothiocyanate (FITC) (1 mg/ml). HRP or FITC were added to the apical side of the monolayer. HRP recovered in the basolateral chamber was measured enzymatically and the percentage calculated relative to the total amount added. The substrate used contained: sodium citrate (0.2 M), ABTS (2,2'-azino-bis(3-ethylbenzthiazoline -6-sulphonic acid) (1.5 mg/ml), H_2O_2 (30%). The reaction was read at 405 nm. Fluorescence intensity of the basal media (FITC) was

measures with a fluorescent plate reader at 485 nm and 544 nm of excitation and emission wavelengths respectively. This assay ensured that integrity of the monolayer was intact after transmigration had occurred.

Transcytosis Experiments

Bacterial cultures (MS11 *opa*$^-$, MS11 wt Opa+ and MS11 wt Opa$^-$) were grown onto fresh media to ensure that live cells were used. Light microscopy will ensure the phenotypic characteristics of the wild type bacteria. Bacteria were suspended to a Klett of 100 and diluted to a concentration of $1X10^6$ cells/ml in media consisting of a 1:1 mixture of Dulbecco's modified Eagles's media and Ham'a F12 medium as described before but in this case supplemented with 5% fetal bovine serum and 0.5% Kellogg's. T84 cells are incubated with invasion media described previously. Transepithelial resistance is measured to ensure polarization. Approximately $1X10^5$ bacteria were added to the apical domain of the transwell filter. The cells were incubated at 37°C and 5% CO_2 during different time points. Apical and basolateral fraction were collected and the number of bacteria in each domain was determined by diluting the fractions and plating them onto GCK. In addition, the number of cell associated and internalized bacteria were assessed by antibiotic protection assay (Schmitter et al. 2004). Cells associated with bacteria were incubated with 1% saponin for 15 min to lyse the cells and then diluted and plated onto GCK. Internalized bacteria were incubated in invasion media supplemented with 100 µg/ml gentamicin for 2 hours to kill all extra-cellular bacteria and then the cells were lysed with 1% saponin. Lysates were diluted and plated onto GCK (see Table 3).

When using inhibitors Cytochalsin D (1mg/ml), Imipramine (50μM) and EGTA (5mM), these were added 15 minutes before bacteria were added. The rest of the procedure was the same as described above.

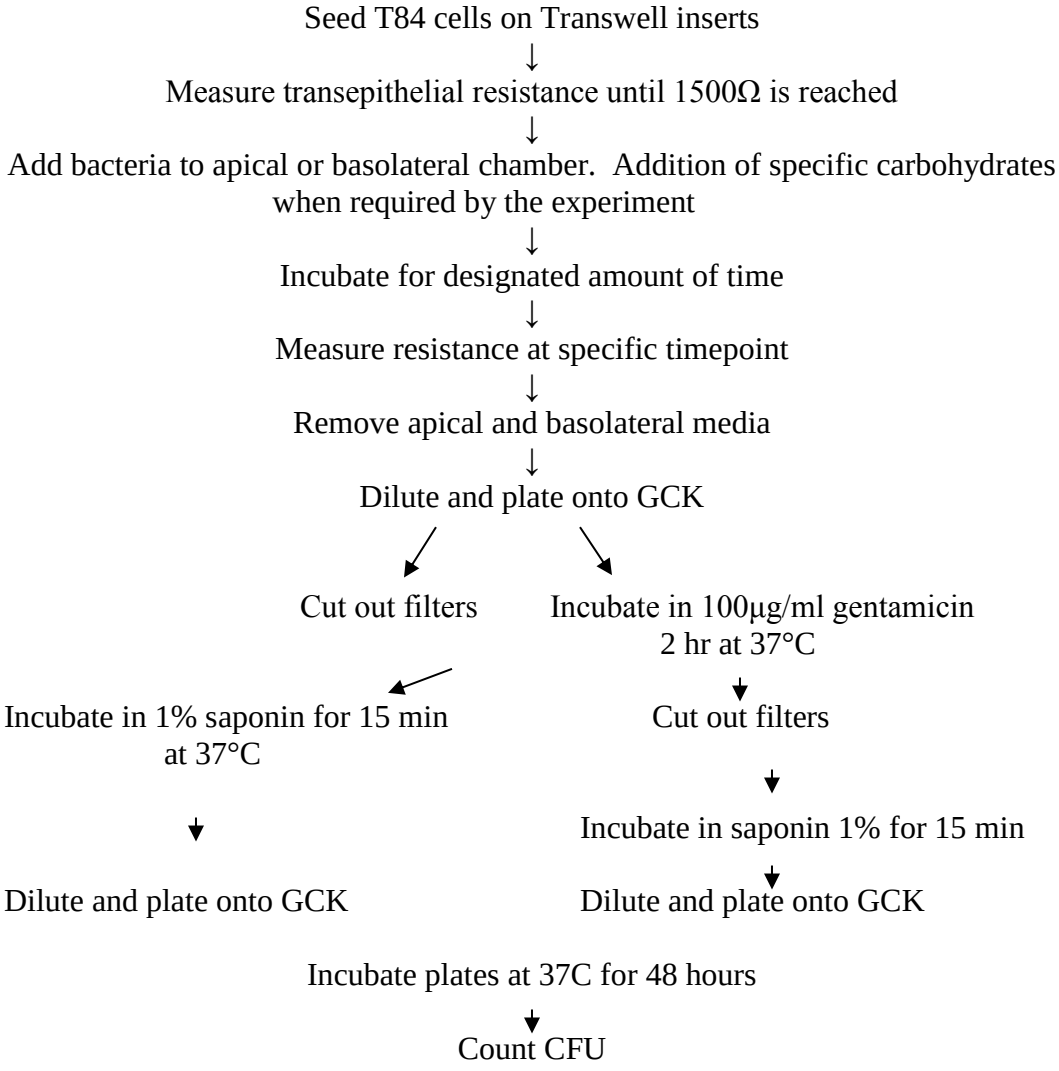
ELISA

T84 cells were seeded and grown on filters as described above. After transcytosis experiments were performed for 6 hours, cell supernatants were collected and assayed for the presence of cytokines by cytokine sandwich ELISA. A protease inhibitor cocktail (Sigma-Aldrich) was added and each sample was stored at -80°C until analysis. Antibody pairs and recombinant standards for human TNFα, and IL-8 were purchased from BD Pharmingen (San Diego, CA). ELISAs were carried out according to protocols provided by BD Pharmingen. Streptavidin alkaline phosphatase and p-nitrophenyl phosphate substrate were purchased from Southern Biotech (Birmingham, AL) and used according to the manufacturer's specifications. Samples were read at 405 nm in 96-well, untreated, flat-bottom plates. Buffers and solutions used are listed in Table 4.

Statistical Analysis

Statistical analysis was assessed using Prism software (GraphPad Software, San Diego, CA). *p*-values were determined using the Student's *t*-test.

Table 3. Diagram of protocol used for transcytosis experiments



Immunofluorescence Staining and Confocal Microscopy

T84 cells (2 $X10^5$) were prepared for confocal microscopy using the protocol described by Bacallao and Stelzer (1989) after transcytosis experiments were completed. Solutions used are listed in Table 5. Primary antibodies used were a rat anti-ZO1 (Transduction laboratories) (2.5µg/ml) to stain the tight junctions, and a mouse anti-gonococcal outer membrane protein (US Biological) (1µg/ml) to stain the gonococci. Secondary antibodies included Oregon Green goat anti rat IgG (Molecular Probes) (2µg/ml) for the ZO-1, Alexa Fluor 633 goat anti-mouse IgG1 (Molecular Probes) (2µg/ml) for the gonococci, and Alexa Fluor 546 phalloidin to stain the actin cytoskeleton of the T84 cells. TRITC was used in some experiments to delineate the length of the T84 monolayer. Stained cells were visualized with a Zeiss 710 Laser Scanning Confocal Microscope.

Table 4. ELISA Buffers

Coating Buffer (Binding Solution)
0.1 M Sodium carbonate pH 7.5
7.13 gr $NaHCO_3$
1.59 gr Na_2CO_3
q.s. to 1L
pH to 9.5 with 10N NaOH
Freshly Prepared and store at 2-8C for up to 7 days.

Assay Diluent (Blocking Buffer)
PBS (80 gr NaCl) with 10% FBS (heat inactivated) pH 7.0
11.6 gr Na_2HPO_4
2.0 gr $KHPO_4$
2. gr KCL
q.s to 1L
pH to 7.0
Freshly prepare or use within 3 days of preparation. Store 2-8ºC.

Wash Buffer
PBS
0.05% Tween-20
Freshly prepare or use within 3 days of preparation. Store at 2-8ºC.

Stop Solution
1M H_3PO_4 or 2N H_2SO_4

Substrate Solution
Tetramethylbenzidine (TMB)
Hydrogen peroxide

Table 5. Confocal Microscopy Reagents

pH-Shift Method by Bacallao and Stelzer (Methods in Cell Biology, 1989, 31:437-452)

10X PBS

KH_2PO_4 2gr
NaCl 80gr
NA_2HPO_4 11.5gr
Add 1000 ml of water
PH to 7.4 or 8.0 with NaOH
Dilute 1/10 for use (filter)

10% Saponin

store @ 4°C

100mM Pipes/KOH

PH to 6.8 with KOH

50mM EGTA

100 mM MgCl2

100mM Sodium Borate

PH to 11 with NaOH

1M NH_4Cl

100mM Glycine (Glycocol, Aminoacetic acid)

75mM NH_4CL, 20mM Glycine

80mM Pipes/KOH, 5mM EGTA, 2mMMgCl2

PH @ 6.5; Store @ RTº

Paraformaldehyde, 8% aqueous (Hood)

Components:

- Paraformaldehyde (powder)
- NaOH, 1 N aqueous
- distilled H_2O
- Store in a sealed bottle at 4° C
-

PBS-Saponin (100ml)

2.5 ml 1% saponin
97.5 ml 1xPBS

PBS-FSG-Saponin (100ml)
2.5 ml 1% saponin
0.66gr FSG
97.5 ml 1xPBS

NaBorate-PFA-Saponin (10ml)
0.25 ml 1% saponin
5 ml 8% PFA
4.75 ml 100mM Na Borate

10% Triton x 100

9 ml 1xPBS in 50ml conical tube
1 ml Triton X-100 measured in 1cc syringe, no needle
Vortex to mix.
Make new solution weekly to avoid increased background labeling.
To make10 ml of 0.1% dissolve 100ul of 10% triton in 9.9 ml 1xPBS

Chapter 3: Construction and characterization of MS11Δ*opa* strain

3.1 Introduction

Why create a GC strain that lacks all of *opa* genes? First of all, GC has the ability to undergo phase variation of several of its antigenic determinants to be able to interact and colonized the different cellular niches it encounters. It must do this since humans are the only host GC can infect. Antigenic variation makes it difficult to study the role of specific determinants in the interaction with human cells; therefore, it is necessary to create strains that express only one type of Opa in a completely Opa devoid background.

All *opa* genes are constitutively transcribed, but are regulated during translation due to a coding repeat sequence (CTCTT) located at the 5'terminus of the ORF that causes a slipped-strand mispairing mechanism that causes the sequence to be in frame or out of frame (Stern et al.1986; Muralidhara et al. 1987; Belland et al., 1989; Meyer et al. 1989). Stern et al (1984) showed that when *opa* was cloned into *E. coli*, it was expressed even when out of frame. According to this, at any determine time, a gonococcus can express all 11 *opa* genes regardless if the gene is in or out of frame. Therefore, a single colony of a non-expressing Opa culture contains a bacterial mix with heterogenous Opa expression. For that reason, it is not in fact an Opa negative colony. According to Mayer (1982), the rate of phase variation of Opa expression is $2x10^3$ per CFU per generation. Bilek et al (2009) reported that in 14 distantly related strains there were no *opa* alleles in common, but in clinical isolates from related sexual networks, *opa* alleles were shared but variations were seen due to recombination of the existing alleles. This high

frequency variation makes it impossible to know which *opa* is expressed in vivo after a clinical sample has been taken.

James and Swanson (1978) first described the difference in colony appearance of GC isolated form male urethra and female cervix, as well as differences in isolates depending on the phase of the menstrual cycle. Challenge studies in men have demonstrated that after inoculating subjects with phenotypically Opa-negative GC, the recovered bacteria express Opa, and this correlated with the presence of symptoms (Schwalbe et al., 1985; Swanson et al., 1988; Jerse 1994; Schneider et al., 1996; Schmidt et al., 2000). The Opa variants recovered were different among the subjects at each study; therefore there was not one predominant Opa protein expressed among the subjects. These phase variations in vivo has been linked to the interaction GC has with different receptors on epithelia cells and leukocytes. Invasion into Chang epithelial cells suggested that a specific Opa protein (OpaC) confers on GC the ability to invade this cell line (Makino et al., 1991). Kupsch et al (1993) showed that depending on the Opa expressed, GC would interact with either epithelial cells or leukocytes. For this, they mutated *opaC*, so that GC would not express OpaC, while expressing other Opa proteins without variation.

Previous studies have also shown that Opa contributes to formation of microcolonies. Using purified Opa and LOS from different strains, researchers suggested that Opa and LOS on adjacent gonococci interact, with the opaque morphology seen under the microscope resulting from the Opa-LOS interaction (Blake et al., 1995; Blake and Gotschlich, 1984; Porat N et al., 1995; van Putten and Robertson, 1995).

3.2 Results

3.2.1 Construction of *opa* deletion strain. The procedure used to obtain an isogenic strain completely lacking all of *opa* genes was done in five parts using different cloning and transformation techniques: First, creation of plasmids containing the *opa* genes, plus flanking sequences; second, deletion of *opa* sequence from these plasmids; third, introduction of a spectinomycin cassette into these plasmids; transformation of *N. gonorrhoeae* to delete the *opa* gene; and lastly, retransformation of the strain to remove the spectinomycin cassette.. This process is summarized in Figures 4 and 5.

To create plasmids that contained each one of the *opa* coding regions, primers for all 11 *opa* sequences were created based on the genomic sequence of strain FA1090 (available at NCBI at the time research was started) (see primers list). PCR amplification was used to generate DNA fragments corresponding to each *opa* gene from chromosomal DNA of MS11 WT strain (fig. 4B). Each PCR product was digested with restriction enzyme that had been introduced into the primers and also used to digest the plasmid pUC19. Digestion of the *opa* sequence and the plasmid was followed by ligation of the *opa* DNA into the pUC19 plasmid. Transformation of the ligated plasmid into *E. coli* was performed using the heat shock method (Inoue et al., 1990). *E. coli* was plated onto LB media containing ampicillin to identify colonies containing the pUC19 plasmid, and X-gal as an indicator to identify transformants that contained inserts. White colonies were identified and plasmid DNA was isolated, using the alkaline method. Confirmation that the transformed colonies contained the desired insert was performed by analysis of

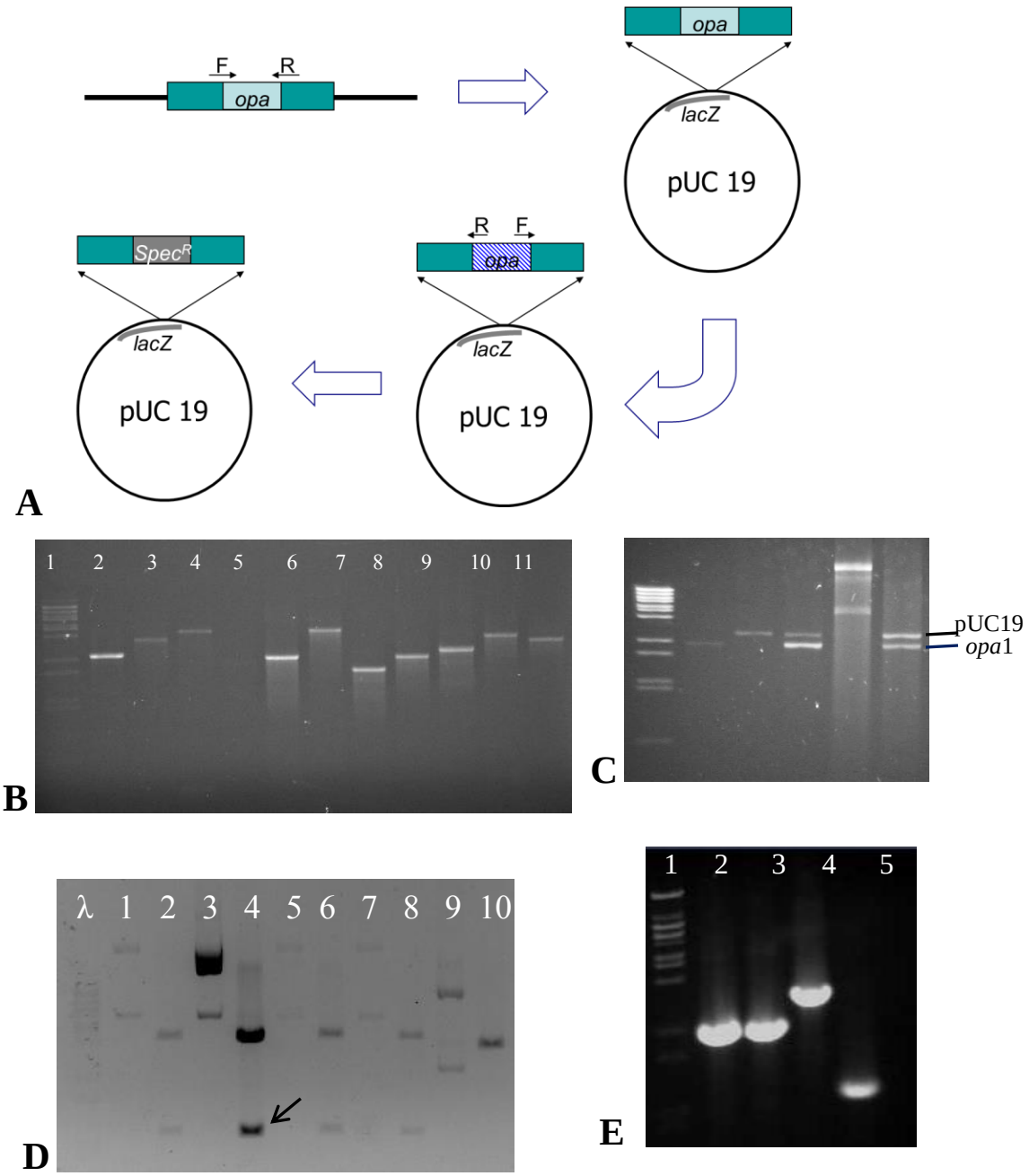


Figure 4. Cloning *opa* genes.
Construction of deletion plasmids of *opa* genes. **A**. *opa* genes were cloned into pUC19 vector and then transformed into *E. coli*. Deletion PCR was performed to delete the *opa* gene and a spectinomycin cassette was inserted in its place. **B**. PCR of all *opa* genes from MS11 WT strain. **C**. o*pa* gene plasmid. Lane 1 λ, Lane 2 *opa*1 PCR, Lane 3 pUC 19, Lane 4 *opa*1+ pUC 19 digested with EcoRI and HindIII, Lane 5 miniprep extraction, Lane 6 miniprep digested with restriction enzymes. **D**. *E. coli* plasmids with Insertion of SpecR cassette (arrow). Lanes 1, 3, 5, 7, 9 show miniprep extraction whiles lanes 2, 4, 6, 8, 10 digested miniprep with PstI enzyme. **E.** Removal of SpecR cassette. Lane 1 λ, Lane 2 and 3 PCR of *opa* 5 gene, Lane 4 *opa* gene deletion + SpecR cassette, Lane 5 *opa* gene deletion.

the mobility of DNA fragments generated after restriction digestion by agarose gel electrophoresis. The correct size was confirmed when DNA extracted from *E. coli* colonies run the same distance as the amplified fragment used in the cloning procedure (fig. 4C).

The *opa* sequence was deleted using a PCR deletion scheme where the coding portion of the *opa* gene was deleted by creating primers that would start from the upstream and downstream flanking regions of each *opa* gene and run in opposite ends amplifying the pUC19 plasmids and the *opa* ends (fig. 4A). These primers also introduced a restriction enzyme site (PstI) that would be used in the next step of this procedure. The PCR amplification was then digested and ligated onto itself. Transformation of *E. coli* was done again to create the deletion plasmids of each *opa* gene (fig. 4D).

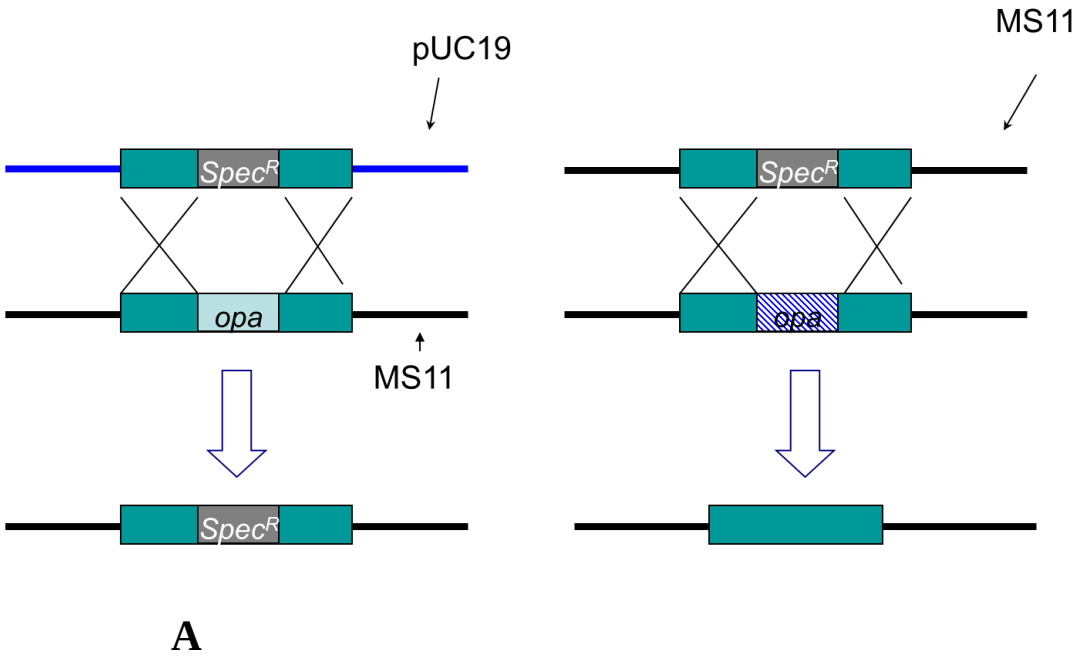
To expedite identification of *N. gonorrhoeae* transformants, an antibiotic resistance cassette was inserted in cis into the plasmid to allow for selection of transformed GC. A spectinomycin cassette was PCR-amplified from the pHP45 plasmid then, digested with PstI. Opa deletion plasmids were digested with the restriction enzyme PstI and ligated onto the spectinomycin cassette introduced in the place where the region of each *opa* sequence was located (fig. 4A). The validity of each construct was confirmed by analyzing digested samples on an agarose gel. Each sequence should have about 700 bp less from the deletion of the *opa* sequence and insertion of the spectinomycin cassette (fig. 4E).

Once the 11 deletions and spectinomycin plasmids were correctly constructed, the deletion constructs were used to transform *N. gonorrhoeae* MS11. The general procedure

is shown in figure 5. Transformation of GC was done by incubating piliated GC MS11 with plasmid DNA for 4 hours in GCP broth with supplements (See Ch 2). Aliquots were plated onto GCK plates containing spectinomycin (30 µg/ml). Colonies that grew on the plates were verified by isolating DNA and digesting it with the appropriate restriction enzyme.

Once the spectinomycin resistant mutant was identified, a non-selective transformation procedure (Gunn and Stein, 1996) was used to identify colonies that had deleted the spectinomycin cassette from the original transformants. The isolated colonies were screened for incorporation of the deletion by PCR analysis. If the colony PCR yielded the desired product, the correct size fragment should be amplified. Once the mutant gonococci were identified, transformation with a new SpecR deletion plasmid was performed; followed by the spot transformation (fig. 5B). These procedures were repeated with nine of 11 *opa* genes present in *N. gonorrhoeae* strain MS11.

The *opa* genes 7 and 10 could not be removed with the spectinomycin procedure, because the cassette always inserted into other *opa* region. To delete these regions, transformation with the deletion PCR product for that *opa* was used. Groups of colonies were screened by PCR to look for the correct transformed GC. After the deletion of each one of the *opa* genes, PCR was performed to verify that the gene had been removed (fig 6). The primers used included the coding region plus a 1 -2 kb flanking region at both the 3' and 5 ' ends. Gene deletion was confirmed when a decrease in about 700 bp. The



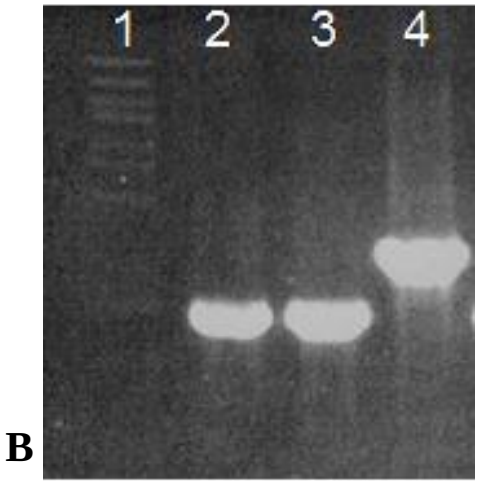


Figure 5. Creation of MS11Δ*opa* strain.
A. Transformation of GC with *specR* plasmid is followed by transformation with the deletion plasmid leaving the gene flanking sequence. Colonies growing on GCK but not on GCK+ spectinomycin are selected. **B.** Lane 1 λ, Lane 2 PCR with *opa*1 primers of MS11 WT, Lane 3 PCR with *opa*1 primers of MS11Δ256, Lane 4 PCR with *opa*1 primers of MS11Δ256 transformed with plasmid DNA *opa* 1ΔS.

deletion started at the first nucleotide of the TATAA box of the promoter to the stop codon of the *opa* sequence (see appendix 1 for sequences).

3.2.2 Confirmation of deletion of *opa* gene related sequences

After confirming by PCR that all 11 *opa* genes had been removed, Southern Hybridization experiments were performed to ensure that additional sequences that could be related to *opa* were not present. For this, a probe was made with the mixture of all PCR amplicons using the primers utilized to sequence each of the 11 *opa* genes. These sequences contained the conserved region of *opa* as well as the variable regions present in all *opa* genes.

Figure 7 shows the λ standard digested with BstEII, the MS11Δ*opa* and MS11 WT strains. MS11 variants were digested with Bgl1 overnight and the DNA was analyzed on a 1% TBE agarose gel for 18 hours to ensure separation of all the bands. MS11 WT variant shows 12 bands, which agrees with a paper published by Bhat et al (1992) where one of the *opa* genes has a Bgl1 site producing two bands (800 bp and 2400 bp). MS11Δ*opa* variant showed no bands. These results show that our deletion strain has no *opa* genes presents or other sequences related to *opa*.

3.2.3 LOS analysis of the deletion strain

SDS-PAGE analysis was performed on extracts of each mutant to investigate if the LOS profile of the deletion strain had changed after all the genetic manipulations. LOS samples of all 11 variant were analyzed and the data in figure 8 shows that all 11

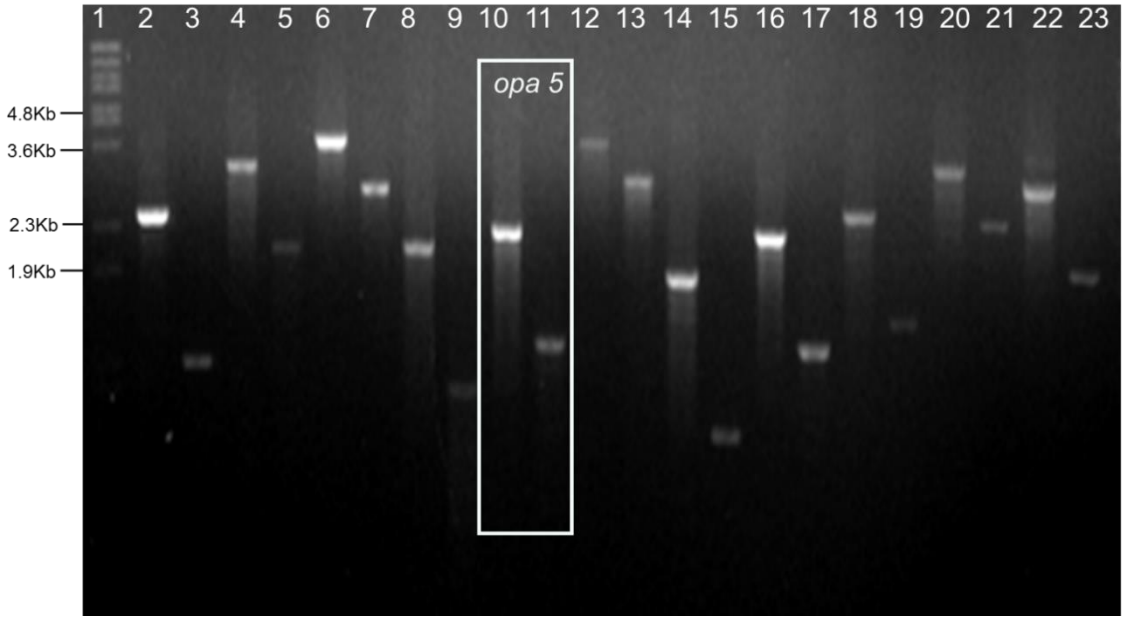


Figure 6. PCR analysis
Verification of *opa* deletions by PCR. Opa-encoding DNA fragments were amplified by PCR using the primer pairs described in Table 1. Amplicons were analyzed on a 1% agarose gel. Lane 1. Λ standard digested with BstEII. Lane pairs represent DNA fragments of each *opa* gene of strains MS11WT and MS11Δ*opa*, respectively. Rectangle shows the PCR product of *opa* 6 for WT and Δ*opa* strain.

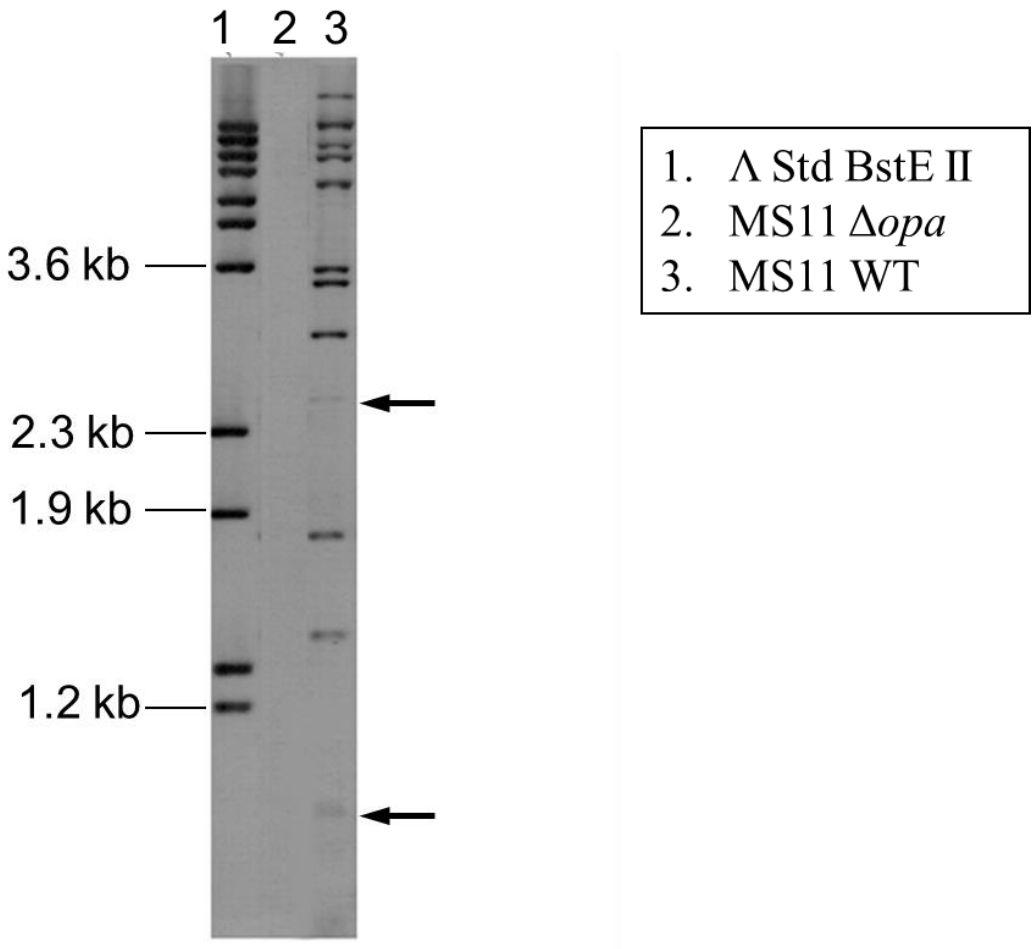


Figure 7. Southern Hybridization
Chromosomal DNA from MS11 WT and MS11Δ*opa* were digested with Bgl1 overnight and then probed with oligomers made for each *opa* gene using the primers Opa5-all and Opa3-all. The samples were analyzed by gel electrophoresis on a 1% agarose gel. The samples were run for 18 hours. Lane 1. λ Standard digested with BstEII, Lane 2. MS11Δ*opa* digested with Bgl1 and Lane 3. MS11 WT digested with Bgl1. Arrows show *opa* with Bgl1 site produces two bands 800b and 2400bp.

variants produce the same LOS size as the parent MS11 strain. This indicates that LOS profile of MS11Δ*opa* strain did not change thought the deletion procedures.

3.2.4 Growth rate of MS11 Variants

To determine if the MS11Δ*opa* variant was altered in its growth properties, the growth rate of the various mutations was determined, comparing the growth rate of MS11Opa+, MS11Opa-, and MS11ΔOpa. Bacteria were grown in GCP containing growth supplements for 7 hours in a 37C shaking incubator. The data in figure 9 shows that all three strains produced similar growth curves, demonstrating that MS11Δ*opa* has no in vitro growth defects.

3.2.5 Morphology arrangements of different Opa expressing strains

Translucent and opaque colonies have been described previously (Swanson 1978) suggesting that difference in colony morphology were due to interaction of neighboring GC. Some of these interactions can be observed in the absence of piliation. We analyzed the arrangement of gonococci when grown in broth cultures and examined the effect on bacterial cell-cell interaction that the lack of Opa has in MS11Δ*opa*. Non-Piliated GC where grown in GCP broth for 7 hours after which time microscopic slides were prepared by heat fixing a sample to the slide with subsequent staining with crystal violet. The data in figure 11 show that MS11 WT+ variant produced large clusters of bacteria, while the MS11 WT- variant showed some small clusters but also presented gonococci that were dispersed. The MS11Δ*opa* variant was observed always dispersed and the clusters seen had very few gonococci arranged in clusters greater than 4 bacteria. Since the bacteria

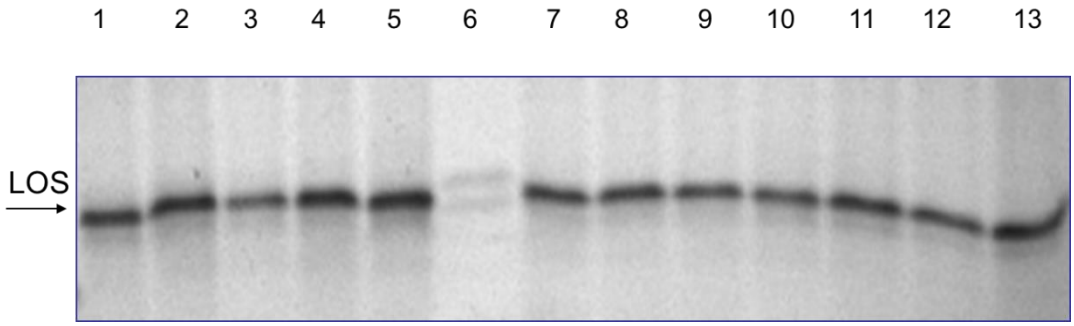


Figure 8. LOS analysis
LOS samples were analyzed by SDS-PAGE and silver stained. Control lanes show strains MS11 wt (1) and F62 (6). Lanes 2-13 show LOS from each variant during the process of *opa* deletion.
Lane 1 – MS11WT
Lane 2 - MS11Δ*opa* 2
Lane 3 - MS11Δ*opa* 8
Lane 4 - MS11Δ*opa* 5
Lane 5 - MS11Δ*opa* 11
Lane 6 - F62
Lane 7 - MS11Δ*opa* 6
Lane 8 - MS11Δ*opa* 3
Lane 9 - MS11Δ*opa* 4
Lane 10 - MS11Δ*opa* 1
Lane 11 - MS11Δ*opa* 9
Lane 12 - MS11Δ*opa* 7
Lane 13 - MS11Δ*opa* 10

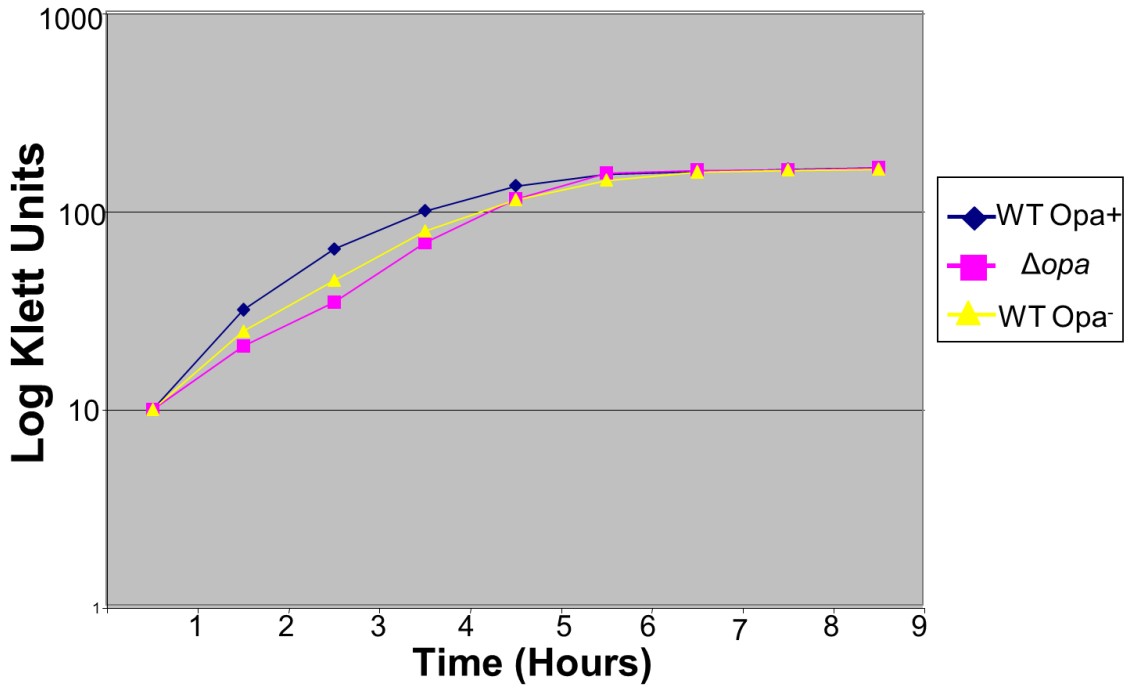


Figure 9. Growth curve of MS11 variants.
Bacteria were grown in a 10 ml GCP broth plus growth supplements and incubated at 37°C with shaking. The turbidity of the cultures was measured every hour. ▲MS11 WT-; ♦MS11 WT+; ■ MS11Δ*opa*.

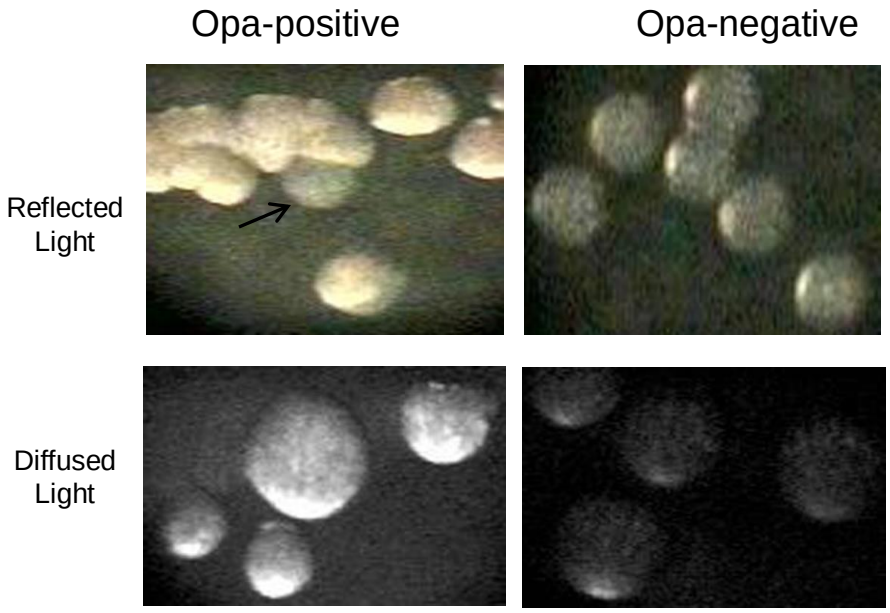


Figure 10. Microscopic observation of Opa colonies. *N. gonorrhoeae* MS11 expressing Opa protein and Δ*opa* strain.
Pictures show Opa expressing gonococci and Opa negative gonococci (MS11Δ*opa*) observed with a stereomicroscope. Two light settings were used: Setting 1: using the reflected light. Arrow points to Opa negative GC. Setting 2: using the diffused light.

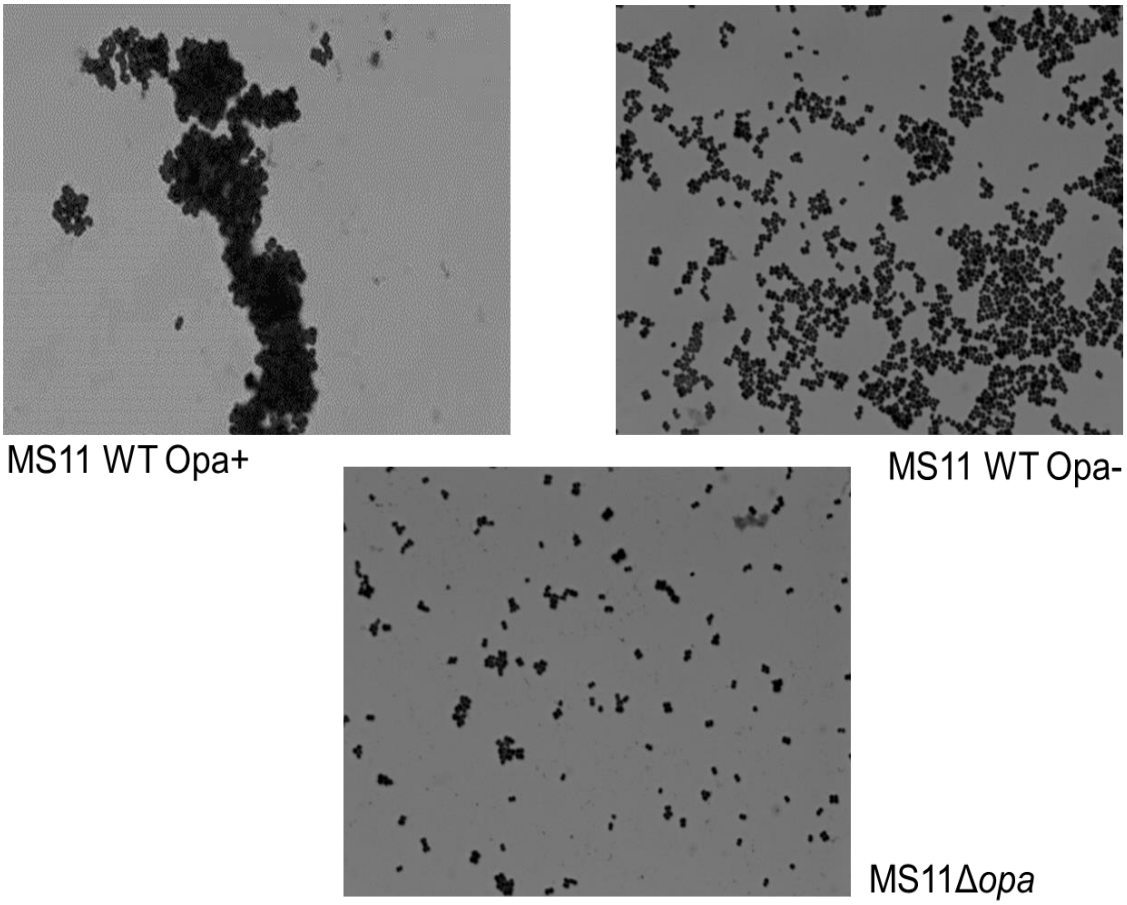


Figure 11. Morphological arrangement of MS11 variants visualized by light microscopy.
Non-piliated MS11 variants were grown in GCP broth for 7 hours with gentle shaking. Aliquots were stained with crystal violet after heat fixed on glass microscope slides and analyzed with light microscopy.

.

used were non-piliated, these results suggest that the differences in the interaction between the cells were due to Opa binding to ligands on adjacent cells.

3.2.6 Opa-LOS interactions

Previous studies have shown that Opa proteins bind to carbohydrate structures on LOS molecules of adjacent gonococci (Blake et al., 1995). In order to study the ability of Opa to bind LOS of adjacent gonococci, a fluorochrome was conjugated to the lacto-N-neotetraose LOS. Gonococci were grown overnight and mixed with the conjugate. After 10 minutes incubation, unbound conjugate was washed away with water. Slides were observed with a fluorescent microscope. The results seen in figure 12 showed that MS11 WT+ strain was highly labeled by the conjugate. MS11 WT- strain had an intermediate staining suggesting that this strain still expresses some levels of Opa while MS11Δ*opa* strain did not bind the labeled LOS. These results suggest that Opa does bind LOS of adjacent gonococci creating the colony morphology observed (fig. 10) as well as it suggest that gonococci that phenotypically are considered as not expression Opa proteins are still capable of expressing Opa proteins to some degree.

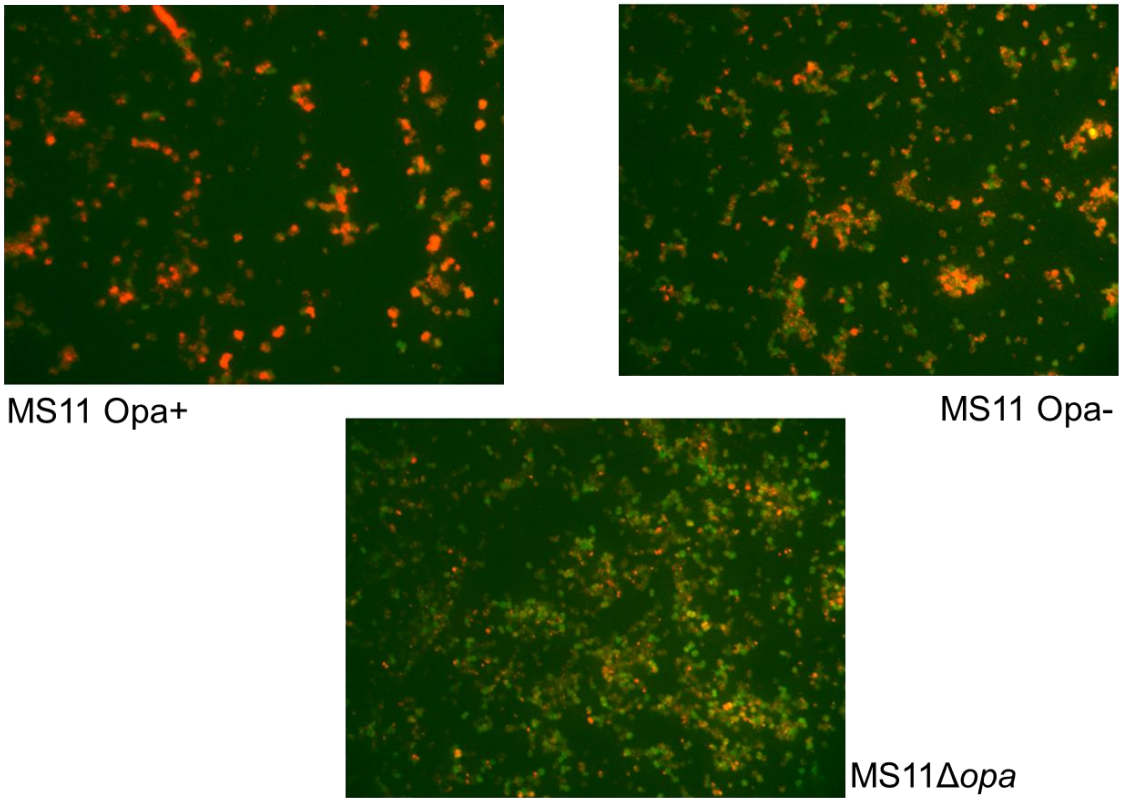


Figure 12. Fluorescent microscopy of Texas Red-LOS Conjugate with gonococci.

Gonococci were grown overnight to $1x10^9$ centrifuged and resuspended in H_20. Cells were incubated for 10 minutes with Texas red-LOS Conjugate. Cells were washed several times with water and slides were viewed on a Fluorescent microscope. Gonococci were located by autofluorescence using GFP channel (green), and cells binding LOS-Texas red were visualized using a 558 nm wavelength filter.

3.3 Discussion

N. gonorrhoeae expresses different surface molecules that aid in the invasion process of epithelial cells, such as Opa, pili and LOS. These molecules undergo mechanisms of antigenic variation (Muralidharan et al., 1987; Burch et al., 1997; Hamrick et al., 2001) creating a different array of disease outcomes depending on the host cells and molecules expressed by the gonococci. The role of Opa in gonococcal infections has been widely studied, but the presence of many *opa* genes in one gonococcus makes it difficult to elucidate the true role of Opa in pathogenesis; not only for the redundancy of the *opa* genes, but also for the many different receptors to which Opa can interact with on the host cell. All the studies done previously used phenotypically Opa negative gonococci. As shown by our study, phenotypically Opa negative gonococci are still able to express Opa proteins to some degree making difficult to understand the exact function of Opas in infection.

In this study, a variant of *N. gonorrhoeae* MS11 that lacks all *opa* genes was constructed. MS11 strain has 11 genes that code for the expression of Opa (Bhat et al., 1991). The procedure used to introduce these deletions was straightforward but lengthy because it required the step-wise removal one gene at a time until all 11 *opa* genes were removed.

Because LOS is another important surface molecule that undergoes phase variation, the conservation of LOS expression by each variant was assessed to ensure that LOS expression did not undergo any change. While the variant and the parental strain

grew at the same rate, the two strains differed greatly in the interactions that Opa proteins have with adjacent gonococci. When analyzing gonococci after growing in broth cultures, the data indicate that the *opa* negative variant does not bind other gonococci like Opa expressing bacteria do. Our variant lacks the ability to form large clusters/microcolonies, suggesting that the interaction between Opa and LOS is required to produce microcolonies. Since MS11 Δ*opa* could not bind to conjugated LOS, this confirmed that LOS-Opa interaction causes the formation of microcolonies as well as creating the different colony morphology observed in gonococcal plates. Taken together, these results suggest that MS11 Δ*opa* can be used to perform pathogenicity experiments to understand the role of Opa expression during infection when compared with wild type strains.

Chapter 4: Interaction of *N. gonorrhoeae* with T84 colonic epithelial cells.

4.1 Introduction

N. gonorrhoeae infects only humans and there is not an appropriate animal model that would mimic all aspects of the pathogenic mechanisms GC causes in humans. GC must be able to invade the mucosal lining to cause ascended and disseminated infections (Mosleh et al., 1997). Disseminated infection involves GC crossing the epithelium to gain access to subepithelial tissues to be able to move from the primary site of infection. Previous work using fallopian tubes and ureteral tissue have shown that GC can infect and traverse stratified epithelium (Mosleh et al., 1997; Gorby et al., 2001).

N. gonorrhoeae uses different mechanisms to adhere to and invade into epithelial cells and eventually to infect host tissues. GC has many surface determinants that undergo phase variation that allow the bacterium to escape the immune system and cause the different array of observed outcomes. Among these are pili, LOS and Opa. Pili promote the initial binding of GC to epithelial cells. In several investigations, experiments performed with piliated and nonpiliated GC (Stephens et al., 1982) showed that even though non piliated GC bound to fallopian tube cells at a lower rate than piliated GC, they were still capable of binding these cells. Using T84 cells, Criss and Seifert (2006) found that GC alters its type IV pili during infection but a population is not specifically selected. Retraction of the pili fiber brings GC closer to the host cell where other virulence determinants can continue the process (Higashi et al., 2009). In the

absence of Opa expression, LOS has been shown to mediate the invasion into epithelial cells (Song et al., 2000).

Opas are known to promote an intimate attachment between GC and epithelial cells through CEACAM or HSPG receptors (Wang et al., 1998). Non-piliated GC can invade human epithelial cells by expressing Opas that use HSPG or CEACAM as their cellular receptor. Cell lines used to study Neisserial interaction with epithelial cells have different expression profiles of CEACAM receptors. HeLa cervix carcinoma cells and HEC-1-B endometrial carcinoma cell lines are negative for the expression of CEACAMs. ME-180 cervix carcinoma cell line produces high levels of CEACAM5 and CEACAM6 but no CEACAM1 (Swanson et al., 2001; Muenzner et al., 2002). T84 cells express CEACAM1, CEACAM5 and CEACAM6 (Wang et al., 1998).

Due to the lack of an animal model, tissue culture cells can be used to study the interaction of GC with polarized epithelial cells. The T84 human colon carcinoma cell line has been used because T84 cells show similarities to epithelial cells *in vivo* such as generation of microvilli on their apical surface, and the ability to form a polarized monolayer with tight cell to cell junctions when grown on filters (Madara et al., 1987; Dharmsathaphorn and Madara 1990; Madara et al. 1992). Transcytosis experiments have been widely used to study the interaction between different microorganisms such as *Campylobacter jejuni*, *Salmonella*, *E. coli* and *Enterococcus* with epithelial cells (Bras et al., 1999; Burns et al., 2001; Finlay and Falkow, 1990; McCormick et al., 1995; Zeng et al., 2004). Transcytosis experiments have been performed with the use of Transwell

membrane supports that create a two-chamber culture system in which monolayers of T84 epithelial cells cover the membrane. Media in the upper chamber is in contact with the apical layer while media in the lower chamber is in contact with the basolateral layer.

4.2 Results

4.2.1 Effect of GC piliation on cell association of T84 cells. Pili are important colonization factors that mediate the attachment of *N. gonorrhoeae* to epithelial cells (Swanson J., 1973; Pusalang et al., 1973). In order to study the effect of surface factors on the interaction of GC with polarized T84 cells, a transwell system was used to analyze these differences. Piliated and non-piliated strains of MS11 Opa+ and MS11Δ*opa* were added to the apical chamber and number of CFU recovered after four hours of incubation determined. To assess the attachment of bacteria to the monolayer, the filters were washed three times with media to remove non-associated bacteria. The filters were treated with 1% saponin to recover the associated bacteria. As shown in fig. 13A, non-piliated bacteria attached to the monolayer 100 fold less than piliated bacteria. This result was the same for both strains showing that pilus mediates an attachment to host cells irrespective of whether the bacteria expressed Opa.

To investigate if pili have an effect on bacterial transmigration across the polarized epithelia, the number of CFU recoverable from the basolateral chamber after four hours of incubation was determined. The results in figure 13, panel B show that 15% of cell associated non piliated MS11Δopa bacteria transmigrated the monolayer.

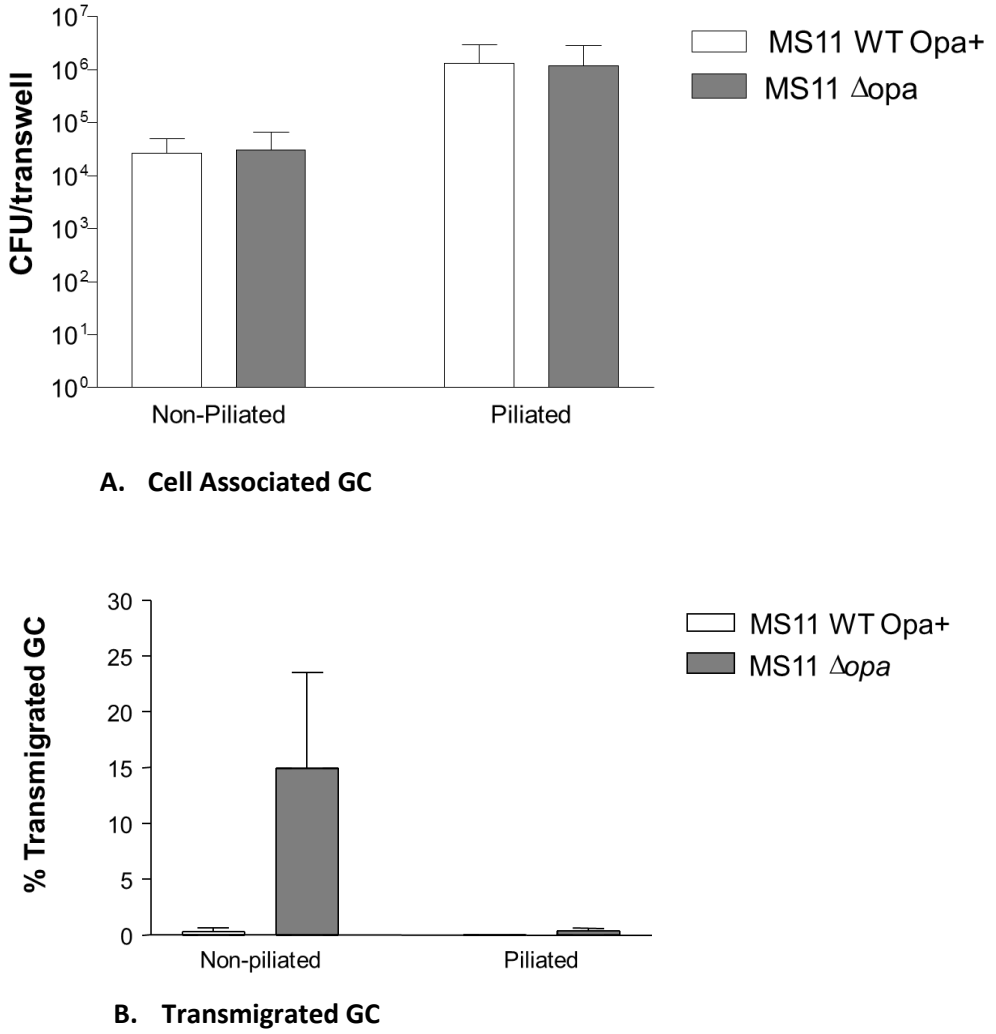


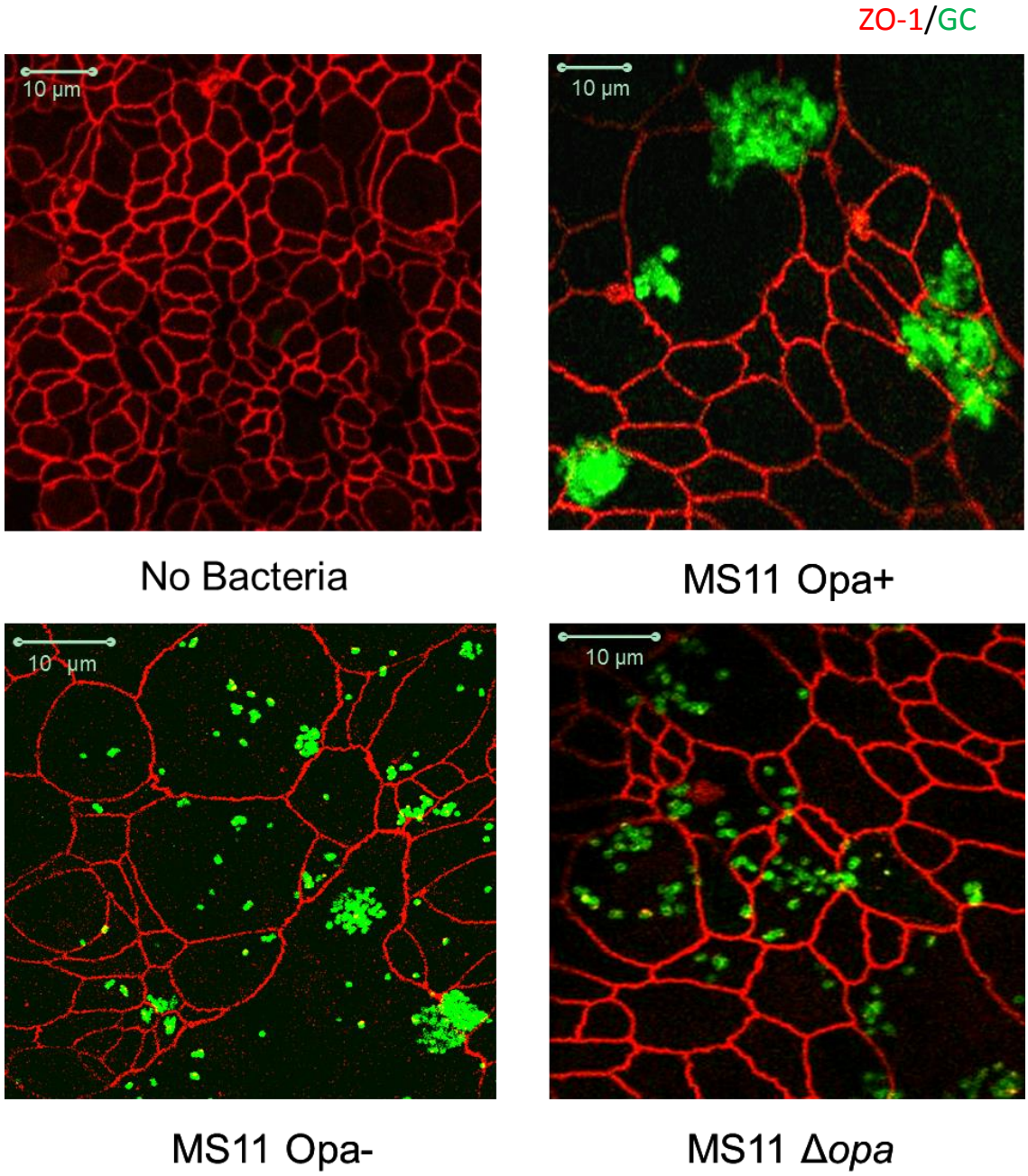
Figure 13. Effects of piliation on gonococcal interaction with polarized epithelia. *N. gonorrhoeae* was added to polarized T84 cells grown on transwells at a MOI of 10:1. After four hour incubation period, aliquots were plated onto GCK. After 24 hours, colonies were counted and observed under light microscopy to determine pili phenotype. **A**. Data represents cell associated GC. Filters were washed three times with media to remove non associated bacteria. Membranes were removed and treated with 1% saponin, then plated. **B**. Data represents percent of cell associated bacteria recovered in basolateral chamber. After 4 hour incubation aliquot was plated onto GCK. Data represents mean values (± SD) of three independent experiments.

Occasionally, a very small number of colonies were recovered from the non-piliated MS11 Opa+ strain as well as a very small number of piliated bacteria from both strains.

This suggests that lack of pili results in an increase in the ability of the bacteria to cross the epithelial barrier. It also suggested that lack of Opa expression allows for the transmigration of epithelial cells. Previous studies (Makino et al., 1991; Chen et al., 1995) have shown that mostly pilus- GC are able to enter epithelial cells. This also agrees with our results that indicate that pilus- GC cross the polarized epithelial barrier more efficiently.

4.2.2 Expression of Opa changes how GC interacts with polarized epithelia.

Studies have shown how GC can form microcolonies when interacting with epithelial cells (Griffiss et al., 1999). Since the phenotypic differences of GC depended on the expression of Opa (see figure 14), the impact of Opa expression on GC interaction with T84 polarized epithelial cells was analyzed. T84 cells were seeded on transwell filters and after 7-10 days, when the resistance was higher than 1000 Ω/cm^2, bacteria were added at a MOI of 10:1. After 6 hr of infection, filters were washed, fixed and stained with antibodies against tight junction protein ZO-1 (red) and with anti-gonococcal antibody (green). The data indicate that the strains arrange differently when interacting with polarized epithelial depending on the expression of Opa. MS11WT strain that expresses Opa (MS11 Opa+) formed large microcolonies that appear to be lying on top of the tight junctions. MS11 Opa-, a strain that does not express significant levels of Opa forms some microcolonies, but can also be seen as diplococci or individual GC. This is in contrast with our strain genetically devoid of Opa expression which can be observed as



14. Distribution of GC Interacting with polarized T84 cells. GC were incubated at a MOI 10:1with polarized T84 cells during 6 hours. T84 cells were grown on transwells until epithelial resistance was over 1,000 Ω/cm^2. After the incubation time, membranes were cut and fixed using the pH-shift method for preparing samples for confocal microscopy by Bacallao and Stelzer (1989). Tight junctions are stained with antibody against ZO-1 protein (red); GC (green).

individual GC or diplococci. These results agree with observations made in chapter 3, which demonstrated that GC aggregate when Opa is expressed, resulting in the formation of large microcolonies. In the absence of Opa, microcolonies are rarely formed and appear as dispersed colonies on the monolayer.

4.2.3 Transmigration of strain MS11Δ*opa*.

Previous studies have shown that GC transmigrate across polarized T84 cell monolayers in 18 hr (Merz et al., 1996). Our data described above showed that GC could be found in the basolateral chamber in 6 hrs. To investigate the time required for MS11Δ*opa* to transmigrate across T84 monolayers, cells were seeded on transwell filters and after 7-10 days when the resistance was higher than 1000 Ω/cm^2 bacteria were added at a MOI of 10:1. Media from the basolateral chamber was recovered after 2, 4 and 6 hours respectively. The data indicate that after 4 hours, $1x10^2$ CFU were recovered in the basolateral chamber (Fig.15A). When compared with the WT strain, MS11Δ*opa* had 100 fold higher number of CFU than the WT strain and after 6 hours the same proportion was observed (fig. 15B). This is in apparent contradiction with published literature that showed that GC transcytosis took 24 hr (Merz et al., 1996).

4.2.4 Opa expression increases cell association but decreases transmigration of epithelial cells. Many studies have shown that expression of different Opas affect the interaction with epithelial cells. To investigate if there was a difference in the interaction of various strains with polarized epithelia, WT strains that phenotypically expressed Opa or did not expressed Opa (MS11 Opa+ and MS11 Opa- respectively) and MS11Δ*opa*

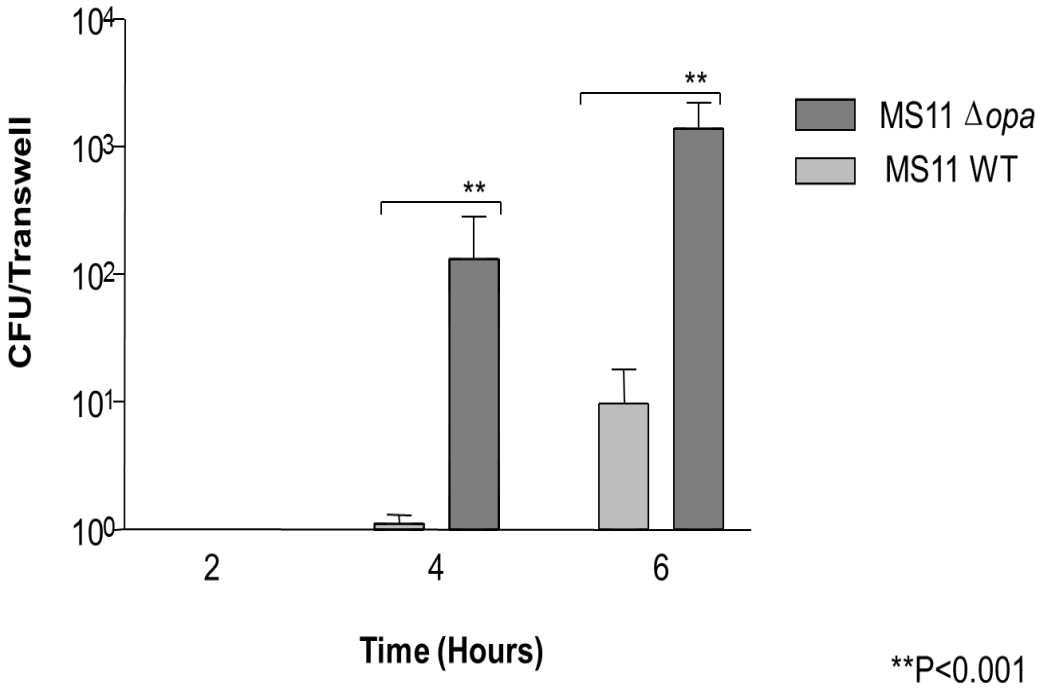


Figure 15. **Time course of *N. gonorrhoeae* transmigration.** *N. gonorrhoeae* was added to polarized T84 cells grown on transwells at a MOI of 10:1. After specific incubation periods, aliquots were plated onto GCK. After 24 hours, colonies were counted. Data represent MS11Δ*opa* and WT strains recovered from the basolateral chamber at 2, 4 and 6 hours after infection. Data represents mean values (± SD) of three independent experiments. ** denotes statistically significance P<0.001.

were used. T84 cells were seeded onto filters and after polarization aliquots of these strains were added. The Opa phenotype of the initial inoculum was determined by plating aliquots onto GCK agar, with the opacity of the subsequent colonies determined by light microscopy. The data in figure 10 represents a visual presentation of the phenotypes observed and the data in figure 16, the first panel represents the outcome of the measured input phenotype of the bacteria added to the filters. These data indicate that Opa-expression varies significantly in wild type strains even if the starting culture was from a plate that appeared to be predominately Opa+. MS11 Opa+ has a high number of each phenotype, while MS11 Opa- has a 100 fold higher number of Opa- colonies than Opa+ colonies. After 6 hr incubation, cell associated bacteria were recovered the by washing the filters with media 3 times and then plating aliquots after incubating with saponin 1% for 15 min (Associated). The results showed that the relative phenotype of the associated bacteria was the same phenotype of the bacteria added at the beginning of the experiment. MS11Δ*opa* attachment to the epithelial cells was the same as the MS11 Opa- strain. Analysis of the bacteria that had transmigrated to the basolateral chamber (Transmigrated) indicated that even when wild type Opa-expressing bacteria were added, Opa negative colonies were the only ones observed in the basolateral media. MS11Δ*opa* transmigrated at a tenfold difference over the MS11 Opa- strain and 100 fold over the MS11 Opa+ strain. These results suggest that Opa expression prevents bacteria from crossing an epithelial barrier.

4.2.5 GC do not alter the integrity of the epithelial barrier. A prerequisite of paracytosis is the disruption of the epithelial barrier integrity. Transmigration of GC has been shown to take place after 24 hrs of infection in T84 cells without change in the

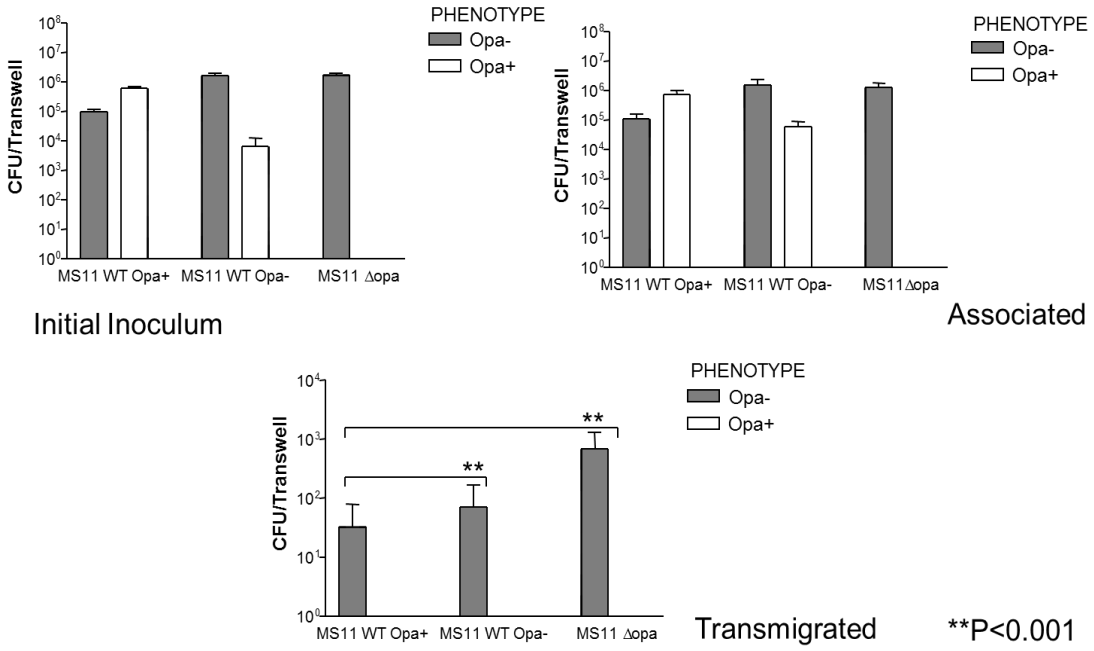


Figure 16. Opa phenotype before and after interactions with T84 cells. *N. gonorrhoeae* was added to polarized T84 cells grown on transwells at a MOI of 10:1. After six hour incubation period, aliquots were plated onto GCK. After 24 hours, colonies were counted and observed under light microscopy to determine Opa phenotype. **A**. Data represents the number of GC added at the beginning of the experiment. **B**. Data represents cell associated GC. Filters were washed three times with media to remove non associated bacteria. Epithelial cells were lysed with 1% saponin, then plated. **C.** Data represents bacterial recovered from basolateral chamber. Data represents mean values (± SD) of three independent experiments. ** denotes statistically significance difference $P<0.001$.

resistance (Merz et al., 1996). Since our data indicated that GC could transmigrate the monolayer in 6 hr, a series of tests were performed to measure the integrity of the epithelial barrier after interaction with GC. To determine if there was a change in the transepithelial resistance (TEER) during the 6 hr infection of the T84 monolayer, the resistance was measured each hour during the 6 hr experiment and normalized by the membrane area of the filter, in this case 0.33 cm^2. The results in figure 17A show that during the 6 hr infection of T84 polarized epithelial cells; neither MS11 Opa- (wild type) nor MS11Δ*opa* induced a dramatic change in TEER. This is contrasted with what was seen with *Salmonella typhimurium*, a pathogen that is known to be capable of depolarization of the monolayer, or disruption of the tight junctions. The data in figure 17B shows the difference in normalized TEER after infection for 6 hr. This demonstrates that GC does not cause an apparent disruption of tight junctions or polarity like the one Salmonella causes when crossing the polarized monolayer.

To investigate if there was a disruption of the tight junctions without a dramatic decrease of the TEER, polarized T84 cells were incubated with GC and added FITC (Fig. 17C) or HRP (fig. 17D). After 6 hr of incubation, aliquots of culture supernatant were analyzed. The results suggested that there was no disruption of the monolayer, because the data showed no leakage of either FITC or HRP to the basolateral chamber. Lack of disruption of tight junction organization indicates the absence of secreted toxins capable of tight junction disruption as well as lack of machinery that could hijack any of the components of the tight junction.

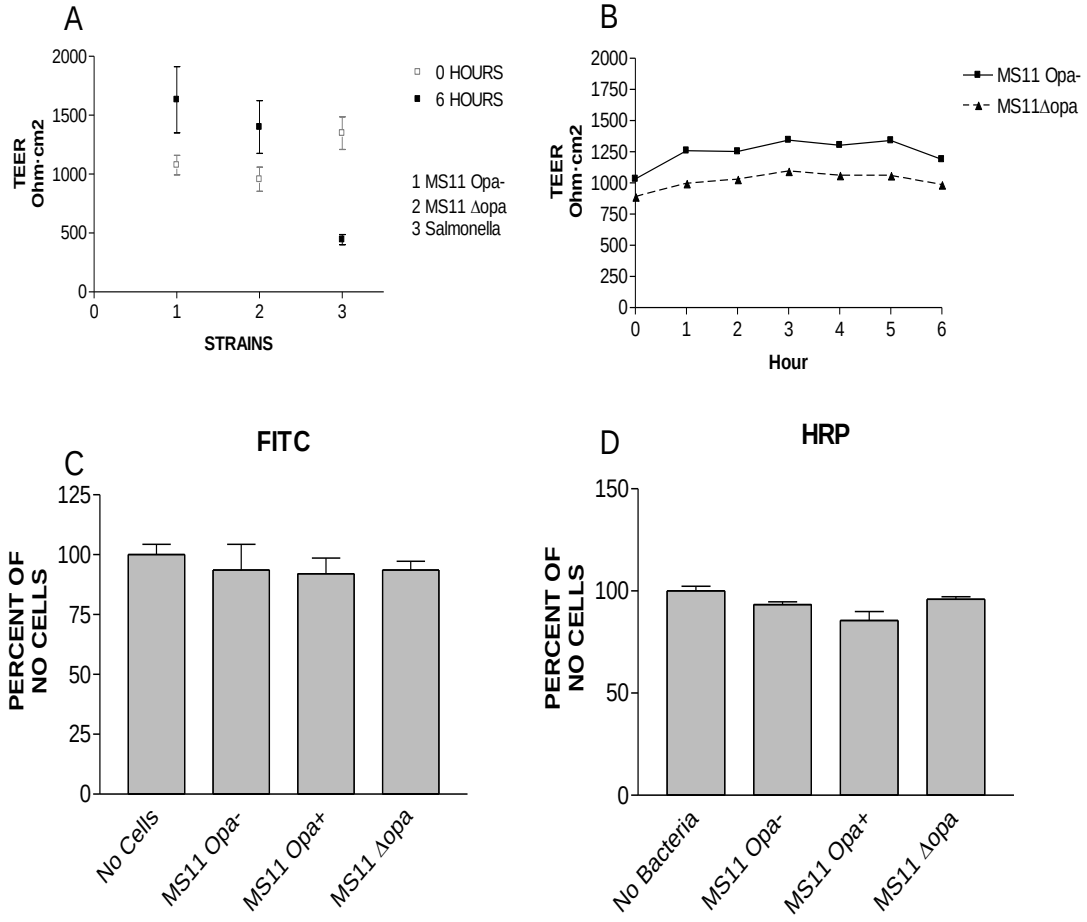


Figure 17. GC does not alter transepithelial resistance. Polarized T84 cells grown on transwells were incubated with GC to measure transepithelial resistance (TEER). **A.** Transepithelial resistance was measured at time 0 and after 6 hours of incubation. *Salmonella* was used as a positive control. **B**. MS11 Opa- and MS11Δ*opa* were incubated during six hours and TEER was measured after each hour. **C.** Bacteria were added to monolayers and 1mg/ml of FITC was added and incubated for 6 hours after which supernatants were read at 405 nm. **D**. HRP 1ug/ml was added to the monolayers after bacteria were added and supernatants were read at 405 nm in untreated flat bottom plates. Each experiment was performed in triplicate. Data represents mean values (± SD) of three independent experiments.

4.2.6 MS11Δ*opa* can enter epithelial barrier. Tight junctions are an important part of the junctional complex forming a barrier with the intercellular space. Many pathogens can cross the epithelial barrier using either paracytosis or transcytosis. Previous studies have shown that after interaction of GC with epithelial cells, the bacteria can be internalized (Waldbeser et al., 1994; McCaw et al., 2004). Confocal laser microscopy was used to determine the location of bacteria when they had associated with the monolayer for short periods of time. After T84 cells were seeded onto transwell filters and bacteria added for 4 hrs, membranes were fixed, cut and stained using the pH-shift method by Bacallao and Stelzer (1989). The data in figure 18 indicate that the tight junctions are polarized (red) as shown in the x –z cross section of the confocal slide. Bacteria can be observed (green) entering the monolayer and also at the bottom of it. Some bacteria can be seen colocalized with ZO-1 at the top of the monolayer. In addition, it appears that bacteria are within the monolayer. In the basolateral layer, many GC can be observed in the x-z cross section of the MS11Δ*opa* when compared with the MS11 WT+ strain where no GC can be observed in the basolateral membrane.

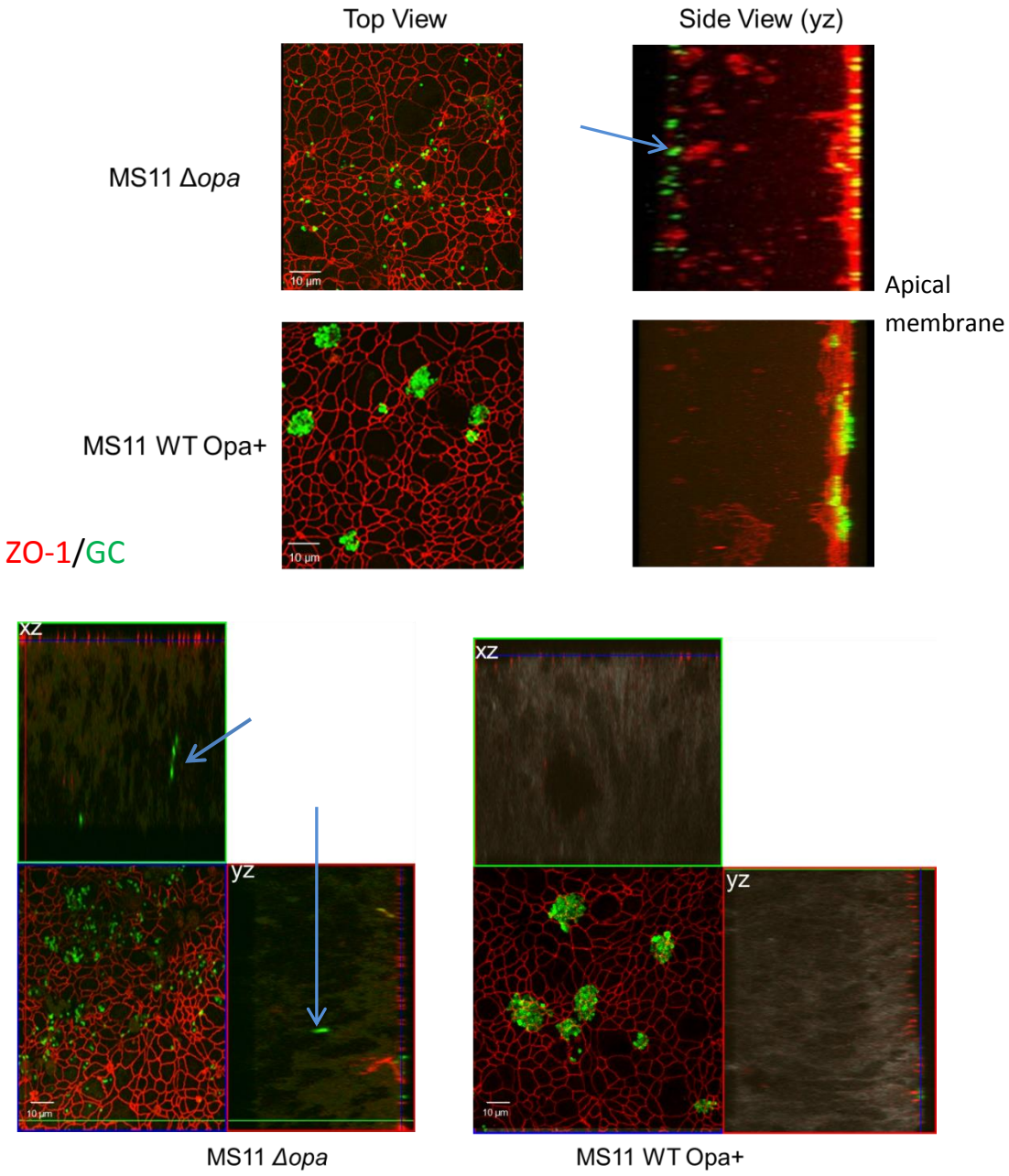
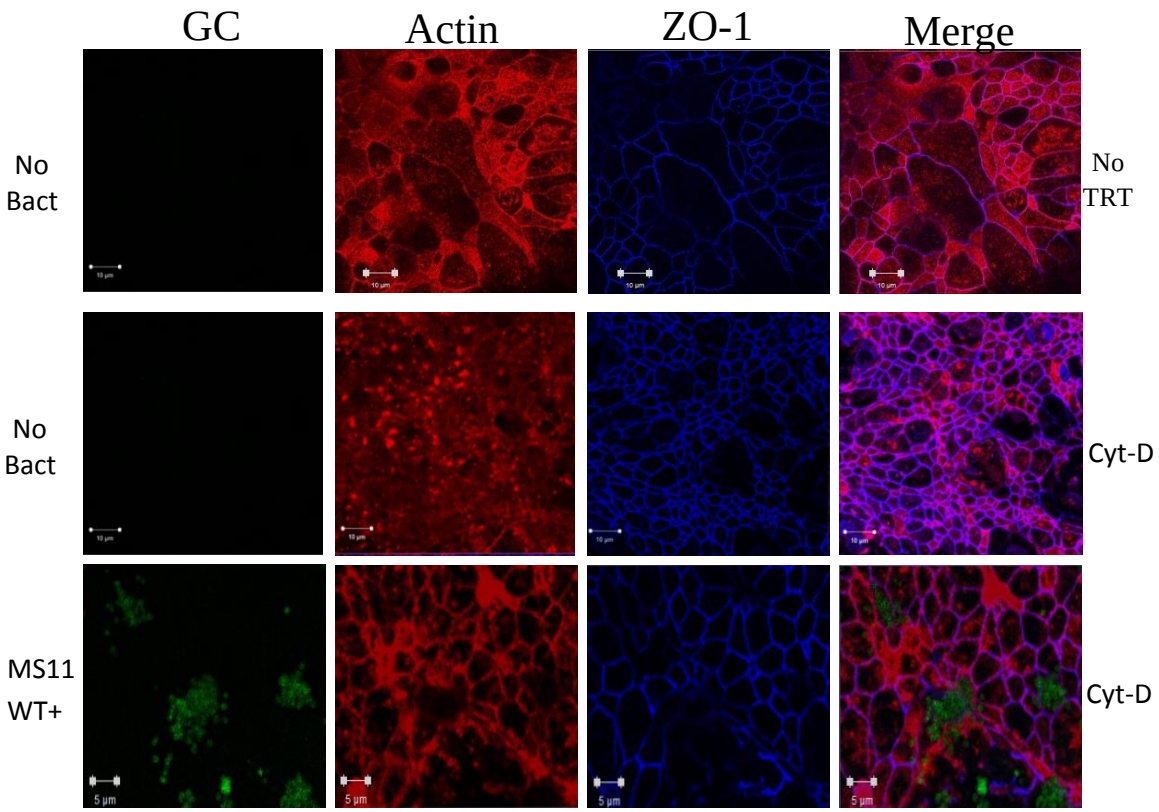
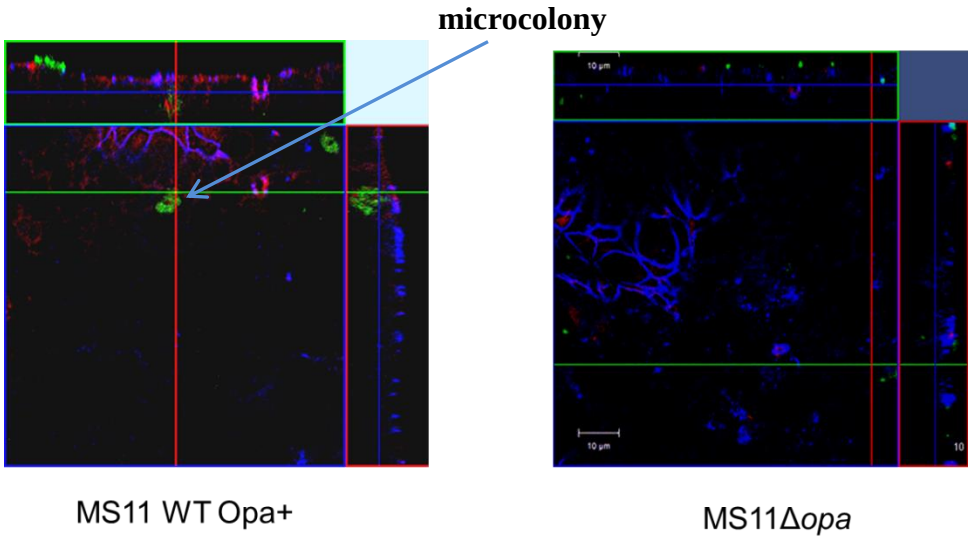


Figure 18. GC enters polarized epihtelial monolayer. MS11Δ*opa* and MS11 Opa+ were incubated at a MOI 10:1with polarized T84 cells during 4 hours. T84 cells were grown on transwells until epithelial resistance was over 1,000 Ω/cm^2. After the incubation time, membranes were cut and fixed using the pH shift method for preparing samples for confocal microscopy by Bacallao and Stelzer (1989). Tight junctions are stained with antibody against ZO-1 protein (red); GC (green). **A**. Top view and side view of monolayer. GC is observed at the bottom of the monolayer on the MS11Δ*opa* strain. **B**. MS11Δ*opa* entering the T84 monolayer and some are located in the middle of the monolayer (x-z view). Arrows show GC. Images were obtained using a Zeiss LSM 510 confocal microscope. Images shown are representative of singles optical sections from three independent experiments.

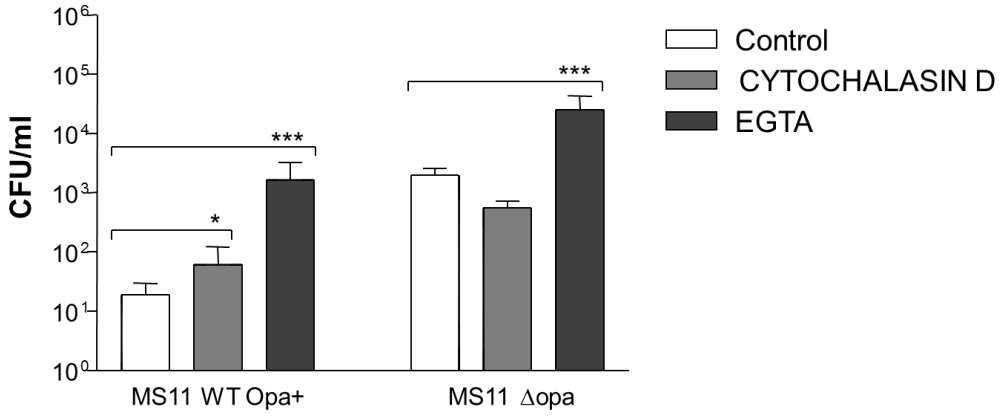
4.2.7 Inhibitors of the tight junctions do not increase transmigrated GC. Many studies have shown that disruption of proteins of the tight junctions causes an increase in the number of transmigrated bacteria. Different inhibitors were tested to see the effect on T84 polarized epithelial cells (Figure 19). The transmigration pattern of two strains of GC was observed under various inhibitor conditions. Imipramine which is an inhibitor of acid sphingomyelinase, plays a role in the uptake of GC by epithelial cells. It did not disrupt or redistribute ZO-1 or F-actin and did not change significantly the number of GC transmigrating the epithelial cells (data not shown). When Cytochalasin D (Fig 19A), an actin polymerization inhibitor was used, it did not appear to change the distribution of ZO-1. However, while it caused disruption of the F-actin, some increase in the transmigration of MS11WT Opa+ GC was observed Fig. 19C). EGTA, a calcium chelator, both ZO-1 and F-actin, disrupted the junctions and caused an increase of transmigrated bacteria in both strains (Fig 19C). In figure 19B, microcolonies of MS11 WT+ were observed crossing the monolayer; this was not observed in the absence of EGTA.



A. Cytochalasin D Treatment



B. EGTA Treatment



C. Transmigrated GC

***P<0.0001
*P<0.05

Figure 19. Effect of inhibitors in transmigration of GC. Inhibitors were used to help determine how GC can transmigrate a polarized monolayer. Cytochalsin D (Cyt-D) (1mg/ml), which inhibits actin polymerization and EGTA (5mM), a calcium chelator were used. **A**. Confocal microscopy images. GC variants were incubated at a MOI 10:1with polarized T84 cells during 4 hours. T84 cells were grown on transwells until epithelial resistance was over 1,000 Ω/cm^2. After the incubation time, membranes were cut and fixed using the pH-shift method for preparing samples for confocal microscopy by Bacallao and Stelzer (1989). Tight junctions are stained with antibody against ZO-1 protein (blue); GC (green); F-actin (red) **B**. x-z image of T84 polarized cells under EGTA treatment showing MS11WT+ microcolony. **C**. Transmigrated bacteria after 4 hours of infection in the presence of inhibitors. Data represents mean values (± SD) of three independent experiments. *** denotes statistically significance difference *** P<0.0001. *P<0.05.
Images were obtained using a Zeiss LSM 510 confocal microscope. Images shown are representative of singles optical sections from three independent experiments.

4.2.8 Cytokine production during transmigration of GC. An important role of the epithelium is to be part of the immune system by producing pro-inflammatory cytokines such as IL-8 and TNF-α (Eckmann et al., 1993; Jung et al., 1995). IL-8 is a neutrophil chemoattractant while TNF-α stimulates the secretion of other cytokines like IL-8, MCP-1. Previous studies have shown that transmigration of *Salmonella* across T84 cells causes an increase in the production of certain cytokines. To examine if GC infection of polarized T84 epithelial cells would lead to secretion of IL-8 and TNF-α, monolayers were infected with MS11 Opa+, MS11*Δopa* and *Salmonella* (used as a positive control). After 6 hrs incubation, aliquots from the apical and basolateral chambers were collected and analyzed by ELISA for presence of IL-8 and TNF-α. Secretion of both cytokines into the basolateral media was observed. No secretion into the apical media was seen. Production of IL-8 did not increase after infection with either strain of GC when compared with the uninfected cells (Fig. 20) while S*almonella* induced the production of 80 pg/ml of IL-8. TNF-α production was induced after addition of the two variants of GC as well as *Salmonella*, producing around 40 pg/ml. This immune response suggests that even in the absence of Opas, GC elicits the induction of TNF-α and it might have an important role during infection with GC.

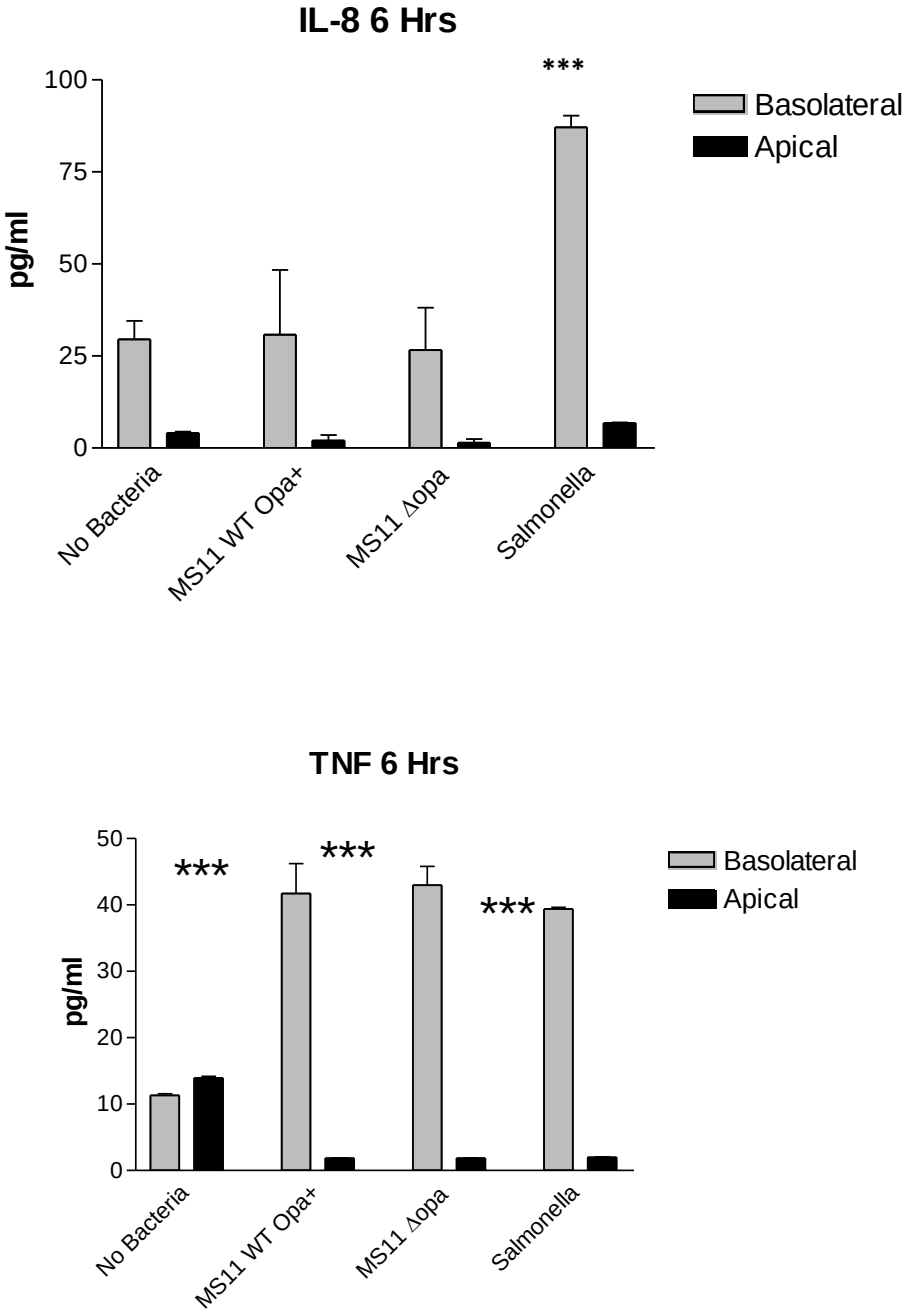


Figure 20. Production of TNF-α and IL-8 in interaction of GC with polarized epithelia. T84 polarized monolayers were infected with GC and *Salmonella* at a MOI of 10:1. After 6 hours of incubation supernatants from the basolateral and apical chambers was collected and analyzed for cytokine production by ELISA. Production of IL-8 in GC did not changed compared with uninfected cells. Only *Salmonella* was statistically significant (***; P<0.0001) Production of TNF-α was increased by addition of GC compared to uninfected cells. Data represents mean values (± SD) of three independent experiments. Asterisk *** denotes statistically significance difference between no bacteria and GC: P<0.0001.

4.3 Discussion

Invasive pathogenic microbes need to cross the epithelial barrier to cause disseminated disease. They can do this by entering individual cells to transcytose to subepithelial tissues or by exploiting host signaling pathways or surfaces of the epithelium to invade tissues using the paracellular pathway. In this study, we observed that absence of pili and Opa facilitates the transmigration of *N. gonorrhoeae* in polarized epithelium. Bacteria were observed entering and at the bottom of the monolayer.

The role of pili in invasion of T84 cells has been shown by Merz et al (1996). They observed that piliated GC adhered to epithelial cells about 20 fold more than non-piliated GC. Pili also enhanced invasion into these cells, taking about 24 hr to be seen in the basolateral chamber compared to the non-piliated GC. The number of bacteria invading epithelial cells appears to depend not only on the bacterial surface structure but also on the type of cells used during the experiments. When using Chang cells, pili appeared to inhibit the invasion (Makino et al., 1991). The mechanism by which pili-negative GC use to bind to epithelial cells is by binding heparin sulfate proteoglycan through to Opa (Chen et al., 1995). In polarized T84 cells, syndecan HS-proteoglycan receptors are found on the basolateral membrane (Rapraeger et al., 1986). In human challenge studies, it was found that pili undergo a high rate of variation during infection helping establish the different outcomes seen during disease (Seifert et al., 1994).

In women, *N. gonorrhoeae* can cause local disease such as salpingitis or disseminated disease such as arthritis. GC must be able to reach the subepithelial space

and invade subepithelial tissues. Piliated GC have been shown to destroy mucosal ciliated epithelial cells during infection of fallopian tubes, although GC attach mostly to non-ciliated cells (McGee et al., 1981). This study also showed piliated GC appeared in small clusters or diplococci attached to the epithelial cells while non-piliated GC appear as diplococci or single GC on the surface of non-ciliated cells. Both piliated and non-piliated bacteria were transparent, which suggest that they were not expressing Opa. This agrees with our observations in T84 cells. Piliated Opa-expressing GC form microcolonies on top of the polarized epithelium, while MS11Δ*opa* was usually seen as a single GC or as a diplococcus.

It has been reported that *N. gonorrhoeae* can transverse polarized epithelia in about 24 hr postinfection (Merz et al., 1996). Intracellular bacteria can be recovered after 3 hr of infection and gentamicin assays showed that bacteria recovered after 24 hs had all been internalized (Wang et al., 1998). The data presented above indicate that GC could be isolated from the basolateral chamber as early as 4 hr. Phase variation of surface determinants is a hallmark in pathogenic *Neisseriae*. Both pili and Opa undergo phase variation, and this allows GC to adapt to different cellular environments and cause the different array of disease outcomes. Using a strain devoid of Opa allowed for the analysis of the impact of phase variation Opa on transmigration. For these experiments, piliated GC were grown for 18 hr on GCK plates prior to addition to polarized epithelia. It was observed that expression of Opa does not significantly affect the attachment to the monolayer when pili is present due to the fact that in the wild type strains MS11 WT Opa+ and Opa- the same proportion for the Opa phenotype was seen as in the bacteria

added at time =0 of the assay. The number MS11Δ*opa* attached to the monolayer was the same as the one observed in the wild type strains. This supports the conclusion that pili play a very important role in initiating disease.

Opa binds two different type of receptors on epithelial cells, heparin sulfate proteoglycan receptors (HSPG) and carcinoembryonic antigen cellular adhesion molecules (CEACAM) and these interactions are important for the invasion of GC into epithelial cell. We show in this study that lack of Opa increases the number of bacteria that can transmigrate the polarized epithelium to invade subepithelial tissues. After 6 hr after infection, the bacteria recovered from the basolateral chamber were mostly Opa- GC, even when the bacteria added were about 50% Opa+ and 50% Opa-. Our results further showed that MS11Δopa had tenfold more in the basolateral chamber than the wild type strains. Therefore, GC lacking Opa has enhanced transmigration on polarized epithelia, due to the fact that does not interact with receptors on epithelial cells.

Many pathogens hijack tight junction components to be able to transmigrate polarized epithelia. This is usually observed by a decrease in TEER sometime after infection. *Clostridium difficile* decreases the TEER after 6 hr of infection by secreting toxin A and causing loss of epithelial barrier function (Hecht et al., 1988). *Salmonella* also decreases TEER by about 20% in the first 40 minutes after infection (Bertelsen et al., 2004). While Salmonella and Clostridia infection produce a dramatic decrease in TEER, the results obtained with wild type GC and MS11Δopa did not decrease the TEER by a significant amount. Using HRP and FITC we could established that the epithelial barrier

was maintained as there was not increased permeability of the tracing agents after 6 hr post infection.

While these results suggest that *N. gonorrhoeae* does not cause a disruption of the tight junction, it does not rule out a transient opening of the junction. It has been suggested that *Listeria monocytogenes* transmigrate the epithelial barrier without causing disruption of the tight junction complex. Pentecost and coworkers found that *L. monocytogenes* could take advantage of the extrusion process during the renewal of epithelial cells to gain entrance of the epithelial barrier (2006). We used confocal microscopy to ensure that tight junctions were not compromised during infection. Polarized epithelia showed that ZO-1, a peripheral protein of the tight junction, remained intact. MS11Δ*opa* were observed entering and transmigrating the monolayer with many bacteria being observed at the bottom of the monolayer. ZO-1 did not appear to be translocated.

These results suggest that *N. gonorrhoeae* is able to infect subepithelial tissues after transmigrating polarized epithelia. Pili are known for promoting the initial binding between GC and epithelial cells and Opas for mediating a tight interaction and leading to the invasion of epithelial cells. In women, GC is mostly asymptomatic developing disseminated disease by crossing epithelial and endothelial barriers. GC takes advantage of phase variation mechanism of surface determinants to be able to invade tissues. Our results show that lack of pili reduces the attachment of GC to epithelial cells. Lack of expression of Opas (Δ*opa* strain) decreases the formation of microcolonies and enhances

transmigration of GC through polarized epithelia. If GC does not interact with other GC on the surface of epithelia cells, GC is probably able to use the paracellular route to transmigrate in 4 hours after infection.

Chapter 5: Conclusions and Discussion

Neisseria gonorrhoeae is a major public health problem, not only in the USA but worldwide. Gonorrhea results in an influx of neutrophils in individuals with symptomatic disease, but complications are caused mostly in women due to its asymptomatic nature, leading to PID and DIG. Additionally, the increased number of antibiotic resistant GC and the correlation with an increase in HIV infections make gonorrhea a burden in public health.

GC must be able to cross epithelial cells and escape the immune system to be able to cause disseminated and chronic disease. GC must initially attach to epithelial cells, invade and survive inside epithelial and phagocytic cells and eventually transmigrate epithelial cells. This study has focused on determining how Opa and pili, surface factors that undergo antigenic variation affect transmigration of GC through polarized epithelia. Opa proteins are encoded by 11 different genes located at different sites in the genome; therefore, control of their expression is not the same as control of an operon. Each *opa* gene has its own promoter and undergoes antigenic variation as a result of a slip strand mechanism due to a repetitive sequence that causes the gene to be in frame or out of frame. In nature, it is not possible to obtain naturally a gonococcal strain that is fully Opa negative.

Previous reports have shown that phenotypically Opa negative GC can bind receptors at low levels (Vogues et al., 2010), confirming that Opa negative bacteria still expresses Opa. The first approach used in this study was to establish a strain that lacks

all *opa* genes to be able to understand the role of Opa proteins in gonococcal pathogenesis. A genetic approach of PCR/transformations was used to remove each *opa* gene from the genome of MS11. The MS11 strain was chosen due to its increase virulence in male challenge studies compared to FA1090 strain; 2.5 x 10^2 CFUs of this strain are needed to cause infection (Schneider et al., 1995; Schmidt et al., 2001; Hobbs et al., 2011). Results from this investigation showed that MS11Δ*opa* had the same growth abilities and LOS profile as the parent strain. Personal communications with Dr. Alison Criss at University of Virginia, indicated she has constructed an FA1090 strain that lacks all *opa* genes following a sequential methodology as well, and this strain produced similar properties that we observed for MS11Δopa in some assays. This strain differed from WT strains in the way in which they interacted with other gonococci; it failed to form big clumps/microcolonies. This strain also failed to bind to the same degree conjugated LOS as the parental strain. In all, these results suggest that LOS and Opa interact to form microcolonies. Microcolony formation has been suggested to be a requirement for invasion of epithelial cells (Bish et al., 2008), as demonstrated by the fact that killed GC can adhere to epithelial cells, but they cannot form microcolonies and cannot invade epithelial cells. Swanson et al (2011) suggested that microcolony formation may be required to induce signaling events such as redistribution of EGFR that lead to the invasion of epithelial cells. Constructing a strain that lacks all Opas is not only important to study the interaction with polarized epithelia, but also interaction with neutrophils. A characteristic of the inflammatory response seen in the male urethra is an influx of neutrophils. It has been shown previously that neutrophils interact with GC

through Opa-CEACAM binding (Fisher and Rest, 1988; Naids et al., 1991; Chen and Gotsclich, 1996).

Further studies are needed to investigate the role of each Opa in pathogenesis. Each protein has different affinities for the receptor they bind on epithelial and immune cells. It is important to construct strains where each *opa* gene expressing a specific Opa is reinserted in the Δ*opa* strain. When using isogenic Opa-expressing strains, it would be expected that gonococci should interact differently depending on the type of receptors they bind. For example, Opa_{HSPG} would not be able to invade T84 cells apically, through an Opa-dependent mechanism. These gonococci would invade epithelial cells slowly without eliciting phagocytic cells to the basolateral side. However, after transmigration they would be able to enter cells through the basolateral side by binding HSPG receptors. This could provide for a mechanism by which gonococci are released to the epithelial surface, and/or would now be capable of re-infecting more cells escaping immune defenses and to be able to transmit to a new host. I would expect Opa_{CEACAM} expressing gonococci to be able to adhere and invade T84 cells apically better than Opa-negative gonococci because they can bind CEACAM receptors. If gonococci are able to transmigrate through the epithelium, they will most likely be killed by epithelial cells or innate immune cells, elicited interleukin production by epithelial cells. I would expect that $Opa_{HSPG\text{-}CEACAM}$ expressing gonococci can reinvade T84 cells from the basolateral side once they have traverse the monolayer because they can also bind HSPG receptors. This would result in gonococci able to reseed the apical side allowing for transmission to a new host. Current projects in the Stein and Song Labs are following this direction of

research. Mutants that can only express one Opa have been constructed and are being tested for interaction with epithelial cells and transmigration activity. LOS mutants are also being constructed to study bacteria-bacteria interactions and elucidate which Opa can bind to LOS and to which sugar terminal.

Many studies have described the importance of Opa and pili in the attachment and invasion to epithelial cells (Swanson et al., 1987; Rudel et al., 1992 & 1995; Schneider et al., 1995; Van Putten and Paul, 1995; Dehio et al., 1998; Griffiss et al., 1999; Freisslet et al., 2000). Both pili and Opa promote attachment and induce signaling events that can lead to the invasion of GC into the different cell lines used in the experiments. 10 to 17% of women with gonococcal cervicitis will develop PID, which is a chronic infection of the upper genital tract. Infection of the fallopian tubes leads to the sloughing of ciliated cell caused by TNF-α production (McGee et al., 1999). Opa negative gonococci are usually isolated from tubal samples, which occur in 30 to 60% of women with gocococcal PID (Eisenstein and Masi., 1981). To cause chronic disease, GC must ascend to the uterus, fallopian tubes and abdomen, which in spontaneous PID occurs during menses and this correlates with the recovery of Opa negative GC during this time of the menstrual cycle. To cause disseminated infections, GC must transmigrate the epithelial barrier to reach endothelial cells. DGI is a rare complication of gonorrhea with only 1 to 3% of patients developing it (Suzaki et al., 2011). DGI usually does not present symptoms and rarely coexist with PID or urethritis, but many patients have had localized infections previously. The second approach used in this study was to perform transmigration assays to analyze the ability of GC to disseminate across T84 polarized

epithelia. The difference in transmigration between WT strains that phenotypically express or not Opa and our Δ*opa* strain were investigated. Data presented here showed that lack of Opa expression enhances the ability of GC to transmigrate polarized epithelia, when bacteria were non piliated. These results suggest that GC undergo antigenic variation of pili and Opa to be able to cause DGI. GC lacking Opa and pili do not interact with adjacent GC as well as do not bind receptors for these surface factors on epithelial cells making easier to cross the epithelium. WT strains form microcolonies that make it difficult to transmigrate the monolayer. These data fit in with observations made that GC isolated from samples of DGI patients are Opa negative (Eisenstein and Masi, 1981). Our results also suggest a mechanism by which GC gains access to the basolateral membrane of the epithelium is through a paracellular route since it takes 4 hr to see GC crossing the polarized monolayer (Figs. 15, 18) as compared with transcytosis that takes 24 hr (Merz et al., 1996). Using confocal microscopy, MS11Δopa was observed entering and crossing the polarized T84 monolayer. While transcytosis requires invasion of the epithelia to be exocytosed later, paracellular transmigration would require disruption of tight junctions, even if only temporarily, to allow for passage of GC and GC can invade epithelial cells in 6 hr after infection (Bish et al., 2008). Unfortunately, our results do not show how tight junctions are disrupted to allow transmigration of GC. Transepithelial resistance as well as permeability experiments using FITC and HRP showed no disruption of the monolayer. Confocal microscopy did not present conclusively disruption or recruitment of the tight junction peripheral protein ZO-1. Further studies will help elucidate which specific tight junction components; GC is able to disrupt to cause local or disseminated disease that needs invasion of subepithelial tissues. Tight

junctions are composed of a complex group of proteins that interact together to form polarity of the epithelium. Many pathogens disrupt the tight junction function, for example enteropathogenic *E. coli* disrupts β_1 integrins on the basolateral membrane (Muza-Moon e al., 2003). Studying other components of the tight junction function such as transmembrane proteins occludins, claudins and JAMs is helpful in elucidating mechanisms by which GC can transmigrate. Neutrophils can transmigrate by a paracellular route to reach inflamed mucosal surfaces without breaking the epithelial barrier. It has been shown that the N-terminal domain of occluding modulates migration of neutrophils but is not critical for paracellular permeability (Huber et al., 2000).

Later studies by Edwards et al., (2013) showed that live GC (pili+, Opa+) localizes preferentially at the apical side of cell junctions when cells were stained for ZO-1, producing discontinuous staining of ZO-1 and occuldin underneath of GC microcolonies. This demonstrates that GC impacts the integrity of the apical junctional complexes of polarized epithelial and it causes redistribution of E-cadherin to cytoplasmic vesicles. The gate function was not disrupted as shown by permeabilization assays and no TEER decrease. It was shown previously that GC activates EGFR (Swanson et al., 2011) and this activation leads to the disassembly of the apical junctional complex proteins causing redistribution of β-catenin (Edwards et al., 2012). Further studies in the Song and Stein labs are also trying to elucidate the mechanism by which open junctions are open to allow transmigration of GC on polarized epithelia. Confocal and Electron microscopy are being used to determine how bacteria open the cell junction without causing a change in the gate function of the apical junctional complex.

GC must actively be involved in transmigration of polarized epithelia. GC causes redistribution of β-catenin after activation of EGFR, but how can GC transmigrate without disruption of the epithelial barrier has not been elucidated. Since neutrophils use a specific domain of occludin to transmigrate without increasing the transepithelial permeability, it suggests that this could be a possible mechanism.

Results from this work contribute to the knowledge of gonococcal pathogenesis and may help in the development of gonococcal vaccines. Data showed that GC elicits TNF-α production during transmigration of polarized epithelia and that antigenic variation of Opa is important to evade the immune system to establish chronic and disseminated infections. TNF-α production can help activate macrophages and promote induction of NO. Furthermore, it has been reported that antibodies against Opa may decrease the risk of PID (Plummer et al., 1994).

The work presented in this thesis furthers our understanding on the pathogenesis of chronic and disseminated gonococcal disease. Based on these results and published literature, I would like to propose a model of gonococcal transmigration of polarized epithelial cells (Fig, 21). GC undergo phase variation of pili and Opa that allows for different scenarios when interacting with polarized T84 cells. When GC expresses pili, pili promote the initial attachment to epithelial cells leading to the retraction of the pili fiber and this retraction brings GC closer to allow the interaction between Opa and CEACAM receptors. GC form microcolonies thorough the interaction between Opa and

LOS. Microcolony formation and Opa association with CEACAM receptors lead to a series of signaling events that allow for invasion of GC into the epithelial cells. GC microcolonies activate EGFR kinase, phosphorylating EGFR and ErbB2 (Swanson et al., 2011). Karen Swanson also showed that this leads to the activation of PLCγ which induces calcium release to allow invasion of GC. ERK and AKT are also activated allowing the survival of GC in epithelial cells. Interactions of GC with host cells through Opa-CEACAM binding can lead to accumulation of PIP_3 inducing the activation of PI3 kinase. This results in the accumulation of PI3P in the phagosomes (Booth et al., 2003). Also, Opa expressing GC can regulate the expression of CEACAM receptors through the activation of NF-κB (Muenzner et al., 2002). When GC are not piliated, attachment to epithelial cells is decreased. When GC do not express pili nor Opa (P- Δ*opa* strain), they cannot form microcolonies on the surface of epithelial cells. Lack of these surface factors hinders the interaction of GC with epithelial receptors such a CECAMs allowing them to be able to transmigrate the monolayer. As early as 4 hours after infection, GC lacking Opa and pili (P- Δ*opa* strain) are seen entering and crossing the monolayer through a paracellular route, since to cross epithelia takes GC 24 hours. GC might disrupt ZO-1, occluding and E-cadherin proteins of the epithelial junctions without causing a decrease in polarity. GC can activate EGFR leading to redistribution of β-catenin (Edwards et al., 2012). Once on the basolateral domain, GC could phase vary Opa to interact with HSPG receptors that would result in signaling events leading to the uptake of GC by epithelial cells, permitting the re-infection of epithelial cells. Opa_HSPG interactions activate PC-PLC, generating DAG from PC. DAG activates ASM generating ceramide thus allowing internalization of GC into epithelial cells

(Grassme etal., 1997). Transmigration of GC, even in the absence of Opa can induce production of TNF-α leading to the expression of CEACAM1 on endothelial cells allowing interaction of GC with these cells which could allow entrance into the blood stream to cause DGI. In addition, GC on the basolateral domain can infect subepithelial tissues causing chronic infection, as well as infect endothelial cells leading to DGI.

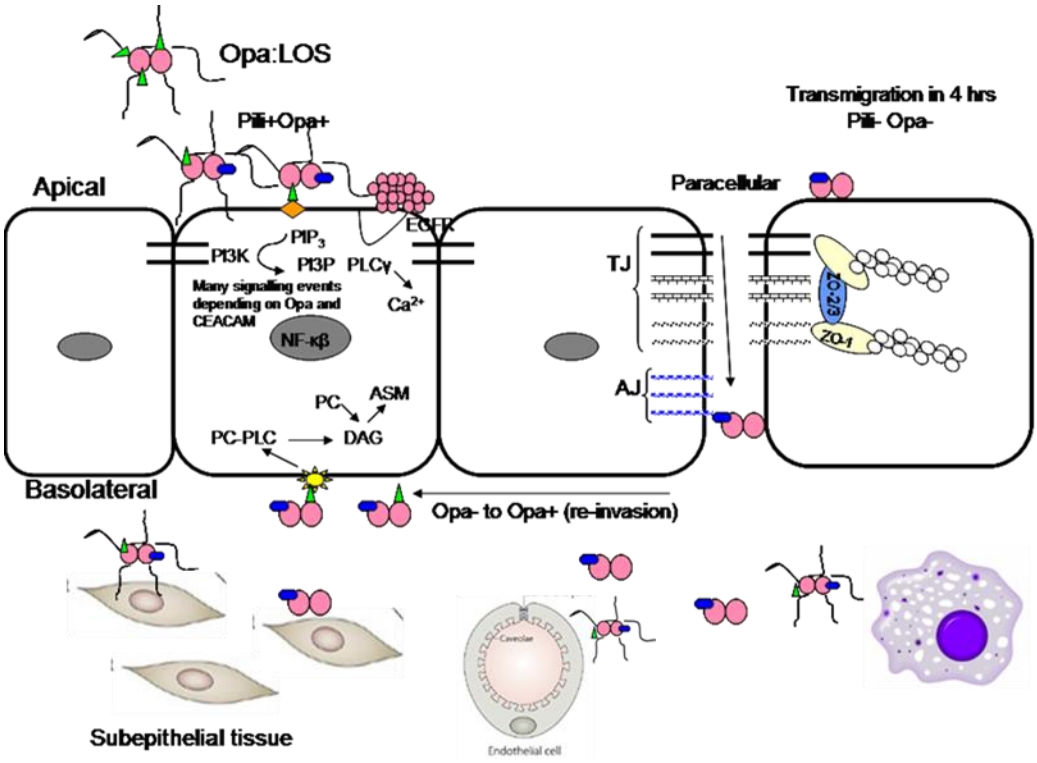


Figure 21. Working model of gonococcal transmigration of polarized epithelial cells.

Piliated GC attach to epithelial cells more avidly than non piliated GC. Opa expressing GC bind CEACAM receptors expressed on the apical side of polarized epithelial cells. These interactions lead to signaling events that promote uptake of GC by epithelial cells. Non piliated *opa*Δ GC will transmigrate T84 monolayer in 4 hrs, making GC accessible to subepithelial tissues to cause chronic and disseminated disease and to interact with endothelial cells and macrophages. Opa ▲, CEACAM◆, HSPG✹, LOS⬬.

Appendix

A.1. *opa* Sequences

DNA Sequences of Opa regions were completed. The start and stop codons are highlighted in green. The number in parenthesis indicates the size of the *opa* encoding amplicon.

>WT_1_Final (2398)
GGATGGTCGCATCGACATCCCGACCGCGACATCCCGCACGAAAACAAGCCGCCCCGAAAGGCA
AATGCCGAAAAATCGCGATAAAAACCATCAATCGGAAAATATTGCATAAACAGGCGAAAATATATCATTGA
AAAGAAAACCCATACAATGTCTGACAGAAACAAGCCGCTTTGCCATTGTTTCCTAACAGTTAAACCCCGCC
CTTCAAGGCGGCGGGGCAGGGCTTGATTCAAAATGGCGCAAGCCCCTGCCCTCAAATCCAACACGCAGGAT
TAAACCATAATAGCGGCTTTCTTATTATTTCTTATTGAAACACCGCCCGGAACCCGATATAATCCGCCTTT
GAAGCATCAGTGAAAATCTTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATCCAGCCCGCA
AAAAACCTTCTCTTCTCTTCTCTTCTCTTCCGCCGCGCAGGCGGCAAGTGAAGGGAATGGCCGCGGCCCGT
ATGTGCAGGCGGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACAAACCGCTCCAAAA
AAAGCACAATTAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTC
GGTCGGCTACGACTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAACAGTA
AATATTCCGTCAGCATAAAAGAGTTGGGAAGAAACGATAATAGCACTTCTAACAGCAGCCATCTTAACATA
AAAACCCAAAAGACGGAACATCAAGAAAACGGCACATTCCACGCCACTTCTTCTCTCGGCTTATCAGCCAT
TTACGATTTCAAACTCAACGATAAATTCAAACCCTATATCGGCGTGCGCGTCGCCTACGGACACGTTAAAC
ATCAGGTTCGTTCAGTGGAAAGCGAAACCACGACTGTTACCACTCACAATGGAGCCCCTGTCCCACAAGGT
CCGACCCCCAAACCTGCCTATCACAAAAGCCGCAGCATCAGCAGCTTGGGGCTTCGGGGCAGTGGCAGGCG
TAGGCATCGACTTCACGCCCAAGCTGACCCTGGACGCCGGCTACCGCTACCACAACTGGTGACGCTTGTAA
AACACCCGCTTCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCGCCGATTCCGATGC
CGTCTGAACCTTCAGACGGCATTTTTGACGCGCCCGCCGTTTACGGGCGCGGGGCAGGTGCGGGGAAATAC
CCGTACCGCCATTGCCGATACACCGTAACCCTAAAACCCGCCATTCCCGACAACACCGTAATCTCGAAACC
CGTCATTCCCGCGCAGGCGGGAATCCGGACCCCCGACGCGGCGGGAATCTATCGGAAATGACTGAAACCCC
GCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCTACGGCCCGCTGATTCCCCGACACCGAT
GCCGTCTGAACCTTCAGACGGCATTTTTGATGCGCCCGCCGTTTACAGGCGCGGGGCGGGCGCAGTAAAAT
ACCCGAACCGTCATTCCCGACAACACCGCAATCGCGAAACCCGTCATTCCCGCGCAGGCGGAAATCCGGAC
CTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGG
AATGACGGTTCGGTTGCGTAGGGCCGGGTGGTCGAAAGGGCGGATTCGATGGATTCGATGGATTCGATGAA
AACGGTAGACATGTTGGATTGATTGTTTATTGTTGAGGAAAATATAAGATCTTCTTTCCGGTTGAAGCCGA
TTTGTTTTTATCTTATGGGTGTTATGCTATATCATCATAGTTATGCAGAAGATGCAGGGCGCGCGGGCAGC
GAGGCGCAGATACAGGTTTTGGAAGACGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACAAAAAAGTGTT
TACCGATGCGCGTGCCGTATCGACCCGTCAGGATATATTCAAATCCGGCGAAAACCTCGACAACATCGTAC
GTAGCATACCCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAATATTCGCGGCGAC
AGCGGGTTCGGGCGGGTCAATACGATGGTGGACGGCATCACGCAGACCTTTTATTCGACTTCTACCGATGC
GGGCAGGGCAGGCGGTTCATCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCG
TCAAAGGCAGCTTCAGCGGCTCGGCAGGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTAGGC
GTGGATGACGTCGTTCAGGGCAATAATACCTACGGCCTGTGCTAAAGTTGACCGCCCATGTCA

>WT_2_Final (2848)
CCGGAGGAGACAGTTGTATACGGTTGTCGCTAACCCGGAAAATTTCGACAGCAGCCGCCACACTGGGCATA
GGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACG
GCACGATGTGGGCGATATCGGCTATCTCAAAGGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCC
GCAGCACCGGTGCGGATGAATATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGT

GGTGTCAACGTTCCGTTTGCCGCAACGGGAGATTTGACGGTTGAAGGCGGTCTGCGCCACGACCTGCTCAA
ACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACTGAAGGCACACTGG
TCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCCGTCCTGTCTGCGACGGCGGGCGTG
GAACGCGACCTGAACGGACGCGACTACGCGGTAACGGGCGGCTTTACCGGCGCGGCTGCAGCAACCGGCAA
GACGGGTGCACGCAATATGCCGCACACCCGCCGGGTTGCCGGTCTGGGGGTGGATGTCGAATTCGGCAACG
GCTGGAACGGCTTGGCACGTTACAGCTACACCGGTTCCAAACAGTACGGCAACCACAGCGGACAAATCGGC
GTAGGCTACCGGTTCTGACGGACAGAAAACAGACAGCCGCAAAGATCACCGTCTTTGCGGCTGTTTCTTAT
GAAAAGAAAACCCTATTCCAATTGCCTGCTTCTATTGTTTCAAGACTTCTTCCAAAGATTCGGCATCAATC
AGACGTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGGCAGTACGGATGGTACGGAACCG
ATCCGCCCGGTGCTTCAGCACCTTAGGGAACCGTTCCCTTTGAGCCGGGGCGGGGCAACGCCGTACCGGTT
TTTGTTCATCCGCCATATTGTGTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATCAGTGA
AAATCTTTTTTTAACCGGTCAAACCGAATAAGGAGCCGAAAATGAATCCAGCCCGCAAAAAACCTTCTCTT
CTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCGGCAAGTGAAGGCAATGGCCGCGGCCCG
TATGTGCAGGCGGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGGTGCAAA
AAAAGGCACAACAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCCGGG
TGTCGGTCGGCTACGACTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAAC
AATAAATATTCCGTGAGCATAAAAGAGTTGCTAAGAAACAAGGGCAATGGCAACAGGACAGACCTGAAGGC
GGAAAATCAGGAAAACGGTACGTTCCACGCCGTTTCTTCTCTCGGCTTGTCCGCCGTTTACGACTTCAAAC
TCAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACGTCAGACACAGCATCGATTCG
ACCAAAAAAACAACAGAGGTTCAACCCATCAACCATGGTCCCGGCACGACCCCTACGCAATCTCGAAGGCA
ATCTTTCCAAGGCACCCGGGAATAAAGATACAGCATCCCCGGAATGGGCCTGGCAACAGGGGCTGGTGTGG
GCATGGGCCTCGGACTCACCCTTTCGTGGGAAGCCGGGTTCGGCTGCCAAGGGTGGGGACGCCGAAAAAAC
ACCCGCTTCAAAACCCGAGAAACCTCGTAGAAAGTGCGGACCGATGTGAAATTCGGAGAGAAGGACACCGT
TTCATCGATCCCATCGCATACGATGGTCCACACCGTCTGACGAAATCGCGGCGCCATTACCGTAGCCGAGC
ATCCTTTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCGCAGCGGGAATCTAGATCTGTCAGTGC
AGAACTTATCGGGCAAAACGGTTTTCTGAGATTTGAGTCCTGGATTCCCACTTCCGCGGGAATGACGGTTC
GGTTGCATAGGGTCAGATTGTCGAAAGGGGCGGATTCGATGGATTCGATGAAAACGGTAGAAATGTTGGAT
TGATGGGAATGGCGGACTGAAGCCCACCGATTCATCGACTCCAACGTTTACGATGCTTCCAACGGTTTCAG
ACGGCATTTTTTACACAATTCCCGCCATTTTCCATCATTCCCGACAATACCGTAATCTCGAAACCCGTCAT
TCCCGCGCAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAAACAGTTGCCCAAACCCCG
CGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCGTTCCGACAACACCGTAGTCTCGAAACCC
GTCCGACAACACCGTAGTCTCGAAACCCGTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCGCAG
GCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAGATT
CCCACTTCCGTGGGAATGACGGTTCGGTTGCGTTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGC
GCAGGCGGGAATCCGGACCCCCGACGCGGCGGGAATCTATCGGAAATGACTGAAACCCCGCGTCCTAGATT
CCCACTTCCGTGGGAATGACGGTTCGGTTGTGTTCCGACAACACCGTAGTCTCAAACCCTCATCCCCGCGC
GATATCCC

>WT_3_Final (2341)
TCAAATTTGTCCGGCACAACTTGGGCTTCCGCTGCGCAAGCCAATAACCCCTTCAATTATAGGGGATTAAC
AAAAACCGGTCCGGCGTTGCCTCGCCTTGCCGTACTGGTTTTTGTTAATCCGCTATATTCCCCCATCTTTA
AAATTTACAGCGATACACGGGTAATTTAAGGAATGCCCAAACCGTCATTCCCGCAACTTTTCTTCATTCCC
GCGAAAGCGGGAATCTAAAATCTCGGACTTTCAAATAATCTTTGAATATTGCTGTTGTTCTAAGGACCGGA
TTCCCGCCTGCGCGCAAGTTTCCAAAGCCTTCCTTTTGGCCAAAGGTCAAAAATCACCCGTCACCGAGTAT
TACCTGAACCACGGCAAATGGCCCAAAGACAACGACTCGGCGGGCGTGGCTTCCGCTTCAAAAATCATAGG
CAAATATGTTAAGCAAGTTGAAGTCAAAAACGGCGTCTTTCCCGCCCAAATGAAATCAGACGGCGTAAACA
AAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGCGTGAAAACGGTTCGGTAAAATGGTTCTGCGGA
CAGCCGGTTACGCGCAACGCCAAAGCCAACGACACCGTTGCCGCCGACGGCACCGGCAACGACAAAATCGA
AACCAAGCACCTGCCGTCAACCTGCCGCGATAACTTTGATGCCAGCTGAGGCAAATTAGGCCTTAAATTTC
AAATAAATCAAGCGGTAAGTGATTTTCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTT
TGCAAGGCGGGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGTGC
CTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAACCCGCCCCCTCATTTGCCGAACAA
GCGGTCCGGACTCCCGCCCGCGCGGGAATGACGGCTGCAGATGCCCGACGGTCTTTATAGCGGATTAACAA
AAATCAGGACAAGGCGGCGGGCCGCAGGCAGTACAAATGGTACGGAACCGATCCGCCCGGTGCTTGGGCGC
CTTAGGGAACCGTTCCCTTTGAGCCGGGGCGGGGCAACGACGTACCGGTTTTTGTTCATCCGCCATATTGT
GTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATCAGTGAAAATCTTTTTTTAACCGGTCA
AACCGAATAAGGAGCCGAAAATGAATCCAGCCCCCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTCTTC

TCTTCTCTCCGCAGCGCAGGCGGGGGGAGAAAACTATGGCCGCGGTGCCTACGTGCGCGCTGATGTCGGAT
ACGCCTACAAAAACCTTACCGGCAATTATGTCCAACACACCCCTCCAAAAAAAGCACAATTAAGCACGGTA
AGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTCGGCTACGACTTCGGCGG
CTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAACAATAAATATTCCGTTAACATAAAAG
AGTTGCTAAGAAACGATAATGCCAATTCTGGCGGCAGCCATCTTAACATAAAAACCCGAAAGACGGAACAT
CGGGAAAACGGCACATTCCACGCCGCCTCTTCTCTCGGCTTGTCCGCCGTTTACGATTTCGATACCGGTTC
CCGCTTCAAACCCTATATCGGCATGCGCGTCGCCTACGGACACGTCAGACATCAGGTTCGTTCGGTTCAAC
AAGAAACCATTGCTGTTACCACTTACCCACAGAATGCTGCGTCAAGTGTTACCACAAATGCTCCGATCCGC
AAACTTCCCCATCACGAAAGCCGCAGCATCAGCAGCTTGGGCTTCGGCGCAGTGGCAGGCGTAGGCATCGA
CATCACGCCCAACCTGACCCTGGACGCCGGCTACCGCTACCACAACTGGGGACGCTTGGAAAACACCCGCT
TCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCGCCGATTCCGATGCCGTCTGAACC
TTCAGACGGCATGAGGCCTTTGCCTGCGCACTTGGTGCGCTGGTCGCCTCCGAACATGGCGCAACACCCGA
CATTTCCGCCGAACGCATCGGGCGTTTCATAAACCCGGTTTAAAACGCATGGAAAAATGCCGTCTGAAAGC
CTTTCAGACGGCATTTTGTTTGAGATTCCGTTTACCAATGGCTGACAAACGCTTCCAAATCGGTATTCTTG
GGCTTATGCACTTCCTCTGTCGGCGTGCCGACCATCATCAGCCCGATGATTTTATCCTTATCCTCGCAA

>WT 4_Final (2073)
CCCGTATCGCACTCTTGACTTCCACGAAAGATGGAATCTGGAATCTCGATACTTTCACAGTATCTTTGAAT
ATTGCCGGTTGTTCCAAGGTCGCGGATTCCCCCCCGCCCGCATGTTTCCTAAGCCTGTCCATTTTGGCCTA
AGGTCAAAAATCACCCGTCACCGAGTATTACCTGAGTCACGGGATATGGCCGAAAGACAACACTTCTGCCG
GCGTGGCATCCCCCCCCTCCGACATCAAAGGCAAATATGTTCAAAGCGTTACGGTCGCAAACGGCGTCGTT
ACCGCCCAAATGAAATCAGACGGCGTAAACAAAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGCG
TGAAAACGGTTCGGTAAAATGGTTCTGCGGACAGCCGGTTACGCGCGCCGCCAAAGACGACGACGCCGTCA
CCGCCGACGGCAACAACAAAATCGACACCAAGCACCTGCCGTCAACCTGCCGCGACACTTCATCTGCCGGT
AAGTGATTTTCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGGGG
TCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGTGCCTCATTGCCACCCGTT
TAAACACGGTTTTTATCTGACAGGCGCGCAACCCGCCCCCTCATTTGCCGAACAAGCGGTCCGGACTCCCG
CCCGCGCGGGAATGACGGCTGCAGATGCCCGACGGTCTTTATAGCGGATTAACAAAAATCAGGACAAGGCG
GCGGGCCGCAGGCAGTACAAATGGTACGGAACCGATCCGCCCGGTGCTTGGGCGCCTTAGGGAACCGTTCC
CTTTGAGCCGGGGCGGGGCAACGCCGTACCGGTTTTTGTTAATCCGCCATATTGTGTTGAAACACCGCCCG
GAACCCGATATAATCCGCCCTTCAACATCAGTGAAAATCTTTTTTTAACCGGTCAAACCGAATAAGGAGCC
GAAAATGAATCCAGCCCCCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGC
GCAGGCGGCGGGTGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAACACA
TTACCCACGATTATCCGGAACCAACCGGTACAAAAAAAGACAAAATAAGCACGGTAAGCGATTATTTCAGA
AACATCCGTACGCATTCCATCCACTCCAGGGTGTCGGTTGGATACGACTTTGGCTGCTGGAGGATAGCGGC
CAGATTATCCCCGTTACAGGAAAATGAAACGACAATAAATATTCAGTCGGCAATAAAAGAGTTAGAAAACA
AGAATCAGAATAAGAGAGACCTGAAGACGGAAAATCAGGAAAACGGTACGTTCCACGCCGTCTCCTCGCTC
GGTTTGTCAGCCGTTTGCGATTTCAAACTCAACGATAAATTCAAACCCTATATCGGTGCGCGCGTCGCCTA
CGGACACGTCAGACACAGCATCGATTCGACCAAAAAAACAACAAAGTTTCTTACCTCCTCCTATGGTGGCT
TAAACCCTACGGTTTATACTGAGGAAAATACGCAAAACGCCCATCACCAAAGTAACAGCATCCGCCGCGTG
GGCCTCGGCGTCATCGCCGGCGTCGGTTTCGACATCACGCCCAAGCTGACCCTGGACACCGGCTACCGCTA
CCACTATTGGGGACGCTTGGAAAACACCCGCTTCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCT
TCTGATTCCCCGATACCGATGCCGTCTGAACCTTCAGACGGCATTTTTGATGCGCCCGCTGTTTACAGGCG
CGGGGCGGGCGTGGAAATACCCGAACCGTCATTGCCGATACACCGTAACCCTAAAACCCGCCATTCCCGCG
CAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAG
ATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCTACGGCCCGCTGATTCCCCGACACCGATGCCGTCTGA
ACCTCAGACGCATT

>WT_5_Final (2236)
GGCATCGCACGGAACTGCCGCCATGACGGCGCACTCATCCGCGTCCTCCCTTCGGAAAACGGACTCGACAT
CCTGCAAGTGTTTGAAGAAATCTCTTGCAGCGTCAATACCACAGCCGCCGGCGGCAACGCCGGCCTGATGG
CGGTTGCCGACTGATTTTGCCAAACCGCCCGGGCGCGGCCCGTGAACAAATGCCGTCTGAAAACCTTTCAG
GCGGCATTTTATAGTGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGACAGTACAAAATAGTACGG
CAAGGCGGGCCAACGCTGTACTGGTTTAAATTCAATTCACTATATGTGTTGGCACATCGCTCCAAACCTGA
TATAATCCGCCTTTCAACATCAGTGAAAATCGTTCTTTTTAGTCAGTTAACATTAAATTTCGGAGTCGAAA
AATGAATCCAGCCCCAAAAAACTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCT

CTTCTCTTCCGCAGCGCAGGCGGCAAGTGAAGACGGCGGCCGCGGCCCGTATGTGCAGGCGGATTTAGCCT
ACGCCTACGAACACATTACCCACGATTATCCGAAACCAACCGATCCAAGCAAAGGCAAAATAAGCACGGTA
AGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTTGGCTACGACTTCGGTGG
CTGGAGGATAGCGACAGATTATGCCCGTTACAGGAAATGGAGCGACAATAAATATTCCGTCAGCATAAAAA
ATATGCGGGTACATAAACACAATAGCAACAGGAAAAACCTGAAGACGGAAAATCAGGAAAACGGCAGCTTC
CACGCCGTTTCTTCTCTCGGCTTATCCGCTATTTACGATTTCCAAATAAACGATAAATTCAAACCCTATAT
CGGCGCGCGCGTCGCCTACGGACACGTCAGACACAGCATCGATTCGACTAAAAAAATAACAGGGCTTCTTA
CCACCAGTACTCCTGGCATAATGTTTGGGGTTTATAAGGTATTAAGGACACCAGGCGCCCATCGCGAAAGC
GACAGCATCCGCCGCGTGGGTCTCGGTGTCATCGCCGGCGTCGGTTTCGACATCACGCCCAAGCTGACCCT
GGACGCCGGCTACCGCTACCACAACTGGGGACGCTTGGAAAACACCCGCTTCAAAACCCACGAAGCTTCAT
TGGGCGTGCGCTACCGCTTTTGATTCGCCGATTCCGATGCCGTCTGAACCTTCAGACGGCATTTTTAAGGC
GCAAGATCGGGCAAACGGCATTTTCAGACGGCATAACTGACAGTATAATCCGAACATCCGGCCGCTTTCGC
GCGGCCTTCAGCATTATCCGATTTTTTCCGAAAGCCGAACCATGCAATACAAACCCCTCCTGCTCGCCCTG
ATGCTTGGCAGGCGGCGGGTGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATCTGGCTTACGCCTAC
GAGCACATCACCCGCGATTATCCCGAAGCAACCGGTGCAAACCAAGGCAAAATAAGCACGGTAAGCGATTA
TTTCAAAAACATCCGCACCCGCTCCGTCCACCCCCGACTCGCCCTCGGCTACGATTTCGGCGGCTGGCGCG
TCTTTTCCGCCCCCGCCGTTGCCGCCGGCCTCCACCAATCCCTTCAATATTACCGATCCCGCCATTGCCGC
CGATTCCGTCAAATCCATCAATTCCGCCGAATCACGCCTATTCCCCAAAAACCTTGATGCGGCGGGCTGAA
GCCCGCCCTGCAACCCTCTCTATGCACCCCCTTGCGAGCCCGACACTACGCAACATCTTGAGAACCCATCC
TGTCAAGAATACCCGAACCGTCCCGATACACCGTAATCCTAAAACCCGCCATTCCCGCGCTGCAATGGGAC
ATCGGCGGCAGCGGGGCGGTTTTCCCTTCGCTCGCACTGTTTCTGCTCTGTTTCATCATAGGTATGCACAA
CACGGGGATGACGCTTCTGCCGGGCGGTGCAATCCGTTCGACGCACATGGCCCGGCACGGCAGCCGACTTG
GGCATCGAAATCCCGCGCGTGCCGTACTATAGTGGATTAACAAAAACCAGTACGGCGTTGGCCCGCCTTAG
CTCAAAGAGAACGATTCTCTAAGGTGCTGAAGCACCAAGTGAATCGGTTCCGTACTATCTGTACTGTCTGC
GGTTCGCCGCCTGTCCTGATTTGTTCCCATCCTTT

>WT_6_Final (3766)
CGCAGGGTTTTACCGTCGCTGCCGGCTTTGTCCATGCCGACAAACCACTGACCTGTCGCTCGGTAAATCAG
CGGGGTTTTGTGGCGCCAGCAGTGGGCGTAGCTGTGTTCGATTTTGCTGCTTGCCAAGAGGTTGCCGGTTT
CTTCCGGCCATTGCAGGATGACGGGATTCGCCTCCCAAACGCTCATGCCTGCGACACGAGGCGTTTCGCTT
ATGTATTTGCCTTCGGCGTTGACAGGGTTGTAAAGCTCGATGCCGTATTTATTGCAGACGGCGTAGTCTTC
CAAACCGTGCGCAGGCGCAGTATGCACCAAGCCGGTACCGGCATCGGTGGTAACGTGTTCGCCGTTGAGCA
TGGGAATATCGCGTTCGAGGAACGGATGATTCATGTGCAGGTTTTCCAGCTTGTCGCCGGTGGTTTCGGCA
AGGATGGCGATGCCGTCTGAAAAGCCGTAGCGTTTCAAAGCGCCTTCCGCCAAATCTTTCGCCAGCACCAA
TTTGCCTTTGGGCGTATCGATTAATTGATACACCACGTCCGCGCCGGCAGACACGGCCTGGCTCGCAGGCA
GAGTCCAAGGCGTGGTCGTCCAAATGACGGCAAACGCTTTGCCTTCGATACCTGCCAAGCCGAATGCGGCG
GCAAGCGCGGCAGTGTCTTTAAACGGATAGGCAACGTCAATCGCAGGCGATACTTTGTCTTTGTATTCCAC
TTCCGCTTCCGCCAGCGAAGATCCGCAATCCAAGCAAAACTGAACCGGTTTCGCACCACGGTAGAGATAGC
CGGATTTGTAGATTTCGCCGAGCATACGCACGGTATCGGCTTCGGTTTTGAAATCCATAGTCAAGTAAGGA
TTGTCCCAATCGCCCAACACGCCCAAGCGGATAAAGTCTTTTTTCTGACGGGCAATCTGTTCGGCGGCATA
TTCGCGGCACAATTCGCGGAAACGGGCTT
TAGGCATATCTTTGCCGTGCAGCTTTTCCACCATCACTTCGATGGGCAGGCCGTGGCAGTCCCAACCCGGT
ACATAAGGCGCGTCAAAACCGGCTTGGGTTTTGCTGCGGATAATAATGTCTTTAAGGATTTTGTTGACGGC
ATGACCGATATGGATGTCGCCGTTGGCATACGGCGGGCCGTCGTGCAGGATGAATTTCGGACGGCCTTTGG
CGATTTCGCGCAGTTTTTGGTAACGTTTTTGCTCGTACCAGCTTTTCAGCCACGCAGGCTCGCGCTTGGCA
AGATTGCCGCGCATCGGAAACGGGCTTTCGAGCAGGTTGACGGTTTTACTGTAATCGGTCATTTTTTAATC
TCTATTGTTACAATATTTCGGTTTCAGACGGCATTGCCCCTCAAACAGTATTTCACAACGGGAAACCCCTA
TGCCGTCTGAACAGTTAAAAGGCTGATTGTAGCCCAATCGGATGGTTTGTATAAGGTTTTTCTACCAACAC
TTTGCGGCTTCCATATCGGCTTCAATCCGCCTTTTCAGTTCTTCCATACCGTCAAACTTTTCCTCGTCGCG
CAGTTTGTGCAGGAAGCGGACGTTCAACCGTTGTCCGTACAAATCGCCTTGAAAGTCGAACAGGTGGACTT
CAAGCTTTTGAGAACAGCCGCCATCAACGGTGGGATTGAAGCCGAAGCTCGCCACGCCGCGCCGCGTGCCG
AATGCGCCGTCCGCTTCGACGACAAACACGCCGCCGAGGGCATAACGGTGGCCGGGCAGTCGGATGTTGGC
GGTCGGGGCGTTTAAGGTGCGCCCGAGTTTTCTGCCGTGCACCACCCTGCCCCCCAAAACGTAGTCGTGCC
CCAAAAGTTTCTTCGCATAGGCAAGGTTGCCGTCTGAAAGGGCTTGGCGGACGGTGGTGCTGCTGGTGCGG
ATGTCTTCGACGATGACGGAAGGCGTGCGCTCGGTTTGCATATCGGGCTGTTGTGCCAAAAGTTCAAAACA
GCCTTCCCGCCCCGCACCGAAACGGAAATCGTCGCCGACGAGCAAATAGCGCGTATTCAAGGTTTGACGCA
GCAGGCGGTCGATAAATGCTTGCGCGGATATTTCGGAAAAATTCCGATCGAAACGCAAAACCCAGGCGGCA

TCGACGCAACCGGTTTTTTCGAGCAGCCCGAGTTTGGTACGCAGGGGGCTGATACGGCACGGCGGGGTTTT
GCCTGTGCGGGCGGCAAAAAATTCTTTGGGTTGCGGCTCGAAAATGACGGCAACGACGGGCAATCCGCGTG
CGTCGGCTTCGAGGCGGAGTTTTTGGAGGATGTGTTTGTGTCCGAGGTGCACGCCGTCGAAATTGCCTATG
GTTACGGCGGCCCCGTGCGGAAAGTCGGGCGCGTTGTGCCGCCCCGGCCTGATTTTCATTGTATTTCCTTT
TCGGTTGAAACCCCGCCACTCGGACATCCGTCCTTCGGGGCGGCAGGATCAGACTTTATTCGGGAGGGGTG
CAACCCCTTCCAAATCAGGACGACACATAGGGCGGTGCTTGATGTGCCGTCCCGTGTGTTGAAACATTGTT
GCGTTCGGGTATGTGGTGAAACAGGCGGTCATTGTAAACGGTATTGCGGTTTATAGACAGTGTGCCGCCCG
TTCCGCCCGCCCGACGCGGGAAAAGTAGGCAAATTTCCCGCCGCCGAACGCGCCAAACGCATAAAACGCGC
AGCAGGCGCGGTGCTATGTGTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATCAGTGAAA
ATCTTTTTTTAACCGGTCAAACCGATATAAGGAGCCGAAAATGAATCCAGCCCGCAAAAAACCTTCTCTTC
TCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCCGCAGCGCAGGCGGCAAGTGAAGACGGCGGCCGCGGC
CCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAACACATTACCCACGATTATCCGGAACCAACCGCTCC
AAACAAGAACAAAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGCACCCGCTCCGTCCACCCCCGGG
TGTCGGTCGGCTACGACTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAATGGAACAAC
AATAAATATTCCGTCAACATAGAAAATGTGCGGATACGTAAAGAGAATGGCATCAGGATAGACCGGAAGAC
GGAAAATCAGGAAAACGGTACGTTCCACGCCGTCTCCTCACTCGGCTTATCCGCTATTTACGATTTCCAAA
TAAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACGTCAGACACAGCATCGATTCG
ACCAAAAAAACAATAGAGGTTACTACCGTCCCCAGCAATGCTCCTAACGGAGCAGTTACAACTTATAATAC
TGATCCAAAGACGCAAAACGATTACCAAAGCAACAGCATCCGCCGCGTGGGTCTCGGTGTCATCGCCGGCG
TCGGTTTCGACATCACGCCCAAGCTGACCCTGGACGCCGGGTATCGCTACCACAACTGGGGACGCTTGGAA
AACACCCGCTTCAAAACCCACGAAGCCTCATTGGGCATGCGCTACCGCTTCTGATTCCCCGACACCGATGC
CGTCTGAACCTTCAGACGGCATTTTTGATGCACCTGCCGTTTACAGGCGCGGGGCGGGCGCAGTAAAATAC
CCGAACCGTCATTCCCGACAACACCGTAATCTCGAAACCCGTCCGACAACACCGCAATCTCGAAACCCGTC
ATTCCCGCGCAGGCGGGAATCTAGAACGTAGAACTGAAGAAACCGNNNCCCGAT

>WT_7_Final (1832)
ANNNAAAACATGAGATTTCCCTCGCCGTATTCCCACGGAAGTGGGAATCTAGAAATAAAAAGCAGCAGGAA
TTTATCGGAAATAACTGAAACCGAACGGACCGGATTCCCGCCTGCGCGGGAATGACGAATCCATCCGCACG
GAAACCTGCACCACGTCATTCCTACGAACCTACATCCCGTCATTCCCACGAAAGTGGGAATGACGGCGGAA
GGGTTTTGGTTTTTTCCGATAAATTCTTGAGGCATTGAAATTCCAGATTCCCGCCTGCGTGGGAATGACGG
CGGAAAGATGCCCGACGGTCTTTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGGCAGTA
CAAATGGTACGGAACCGATCCGCCCGGTGCTTCATCACCTTGGGGAACCGTTCCCTTTGGGCCGGGGCGGG
GCAACGACGTACCGGTTTTTGTTCATCCGCCATATTGTGTTGAAACACCGCCCGGAACCCGATATAATCCG
CCCTTCAACATCAGTGAAAATCTTTTTTTAACCGGTCAAACCGAATAAGGAGCCGAAAATGAATCCAGCCC
CCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCGGCGGG
TGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAACACATTACCCACGATT
ATCCGGAACAAACCGATCCAAGCAAAGGCAAAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACG
CATTCCATCCACCCCCGGGTGTCGGTCGGCTACGATTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCG
TTACAGAAAGTGGAACAACAATAAATATTCCGTGAGCATAAAAGAGTTGCTAAGAAACAAGGTCAATGGCA
ACAGGACAGACCGGAAGACGGAAAATCAGGAAAACGGTACGTTCCACGCCGTTTCTTCTCTCGGCTTGTCC
GCCGTTTACGACTTCAAACTCAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACGT
CAGACACAGCATCGATTCGACCAAAAAAACAACAGAGGTTACTACCATCCTCCATGGTCCTGGCACAACCC
CTACGGTTTATCCTGGGAAAAATACGCAAGACGCCCATCGCGAAAGCGACAGCATCCGCCGCGTGGGCCTC
GGCGCAGTGGCAGGCGTAGGCATCGACATCACGCCCAACCTGACCTTGGACGCCGGGTACCGCTACCACTA
TTGGGGACGCCTGGAAAACACCCGCTTCAAAACCCACGAAGCCTCATTGGGCGTGCGCTACCGCTTCTGAT
TCCCCGATACCGATGCCGTCTGAACCTTCAGCGGCATTTTTAATCGCCCGCCGTTTACAGGCGCGGGGCGG
GCGCAGTAAAATACCCGAACCGTCATTCCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCGCAGGC
GGAAATCTAGATCTGTCAGTGCAGGAACTTATCGGGCAAAACGGTTTCTTGAGATTTTGAGTCCTGGATTC
CCACTTTCGCGGGAATGACAATTCATAAGTTTCCCGAAATTCCAACATAACCGAAACCTGACAGTAACCGT
AGCAACTGAACCGTCATTCCCGACAACACCGCAATCTCGAAACCCCTCCGCCATTATGAAGACAAATCGCG
GCACAAAAAATGCCGTCTGAAATACTGTTTGGCGGTTTCAGACGGCATTTGCTCAAACTTTATCAGGCGTA
ATGGCGCGTTTCGCCTTCTCCGCCGACATTCTCCGCACAGCGTTGCAGACGGTTCAAACCNCTCGCCTGCG
TGATCGGAGGNNT

>WT_8_final (2263)
GCNNCCGCGGCTGATGCCGCCGTAGAGGAATGATGCCCGCCCGCATTGATTTCGCGCGACACGCCCAAGCC
GTAGCGCAAACCGTGTGCCTTTTGCGGCAGGCTGTCGGCGGTTTTCGTCCAGCTTCTGCCCGCAAATTCAA

TCGTTTTTTCGGACGAAGCGTTGTTTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGGCA
GTACAAATGGTACGGAACCGATCCGCCCGGTGCTTGGGCGCCTTAGGGAACCGTTCCCTTTGAGCCGGGGC
GGGGCAACGACGTACCGGTTTTTGTTCATCCGCCATATTGTGTTGAAACACCGCCCGGAACCCGATATAAT
CCGCCCTTCAACATCAGTGAAAATCTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATCCAG
CCCGCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTC
TCTTCTCTTCCCTTCTCTTCCCTTCCGCAGCGCAGGCGGCGGGTGAAGGCAATGGCCGCGGCCCGTATGTG
CAGGCGGATTTAGCCTACGCCTACGAACACATTACCCACGATTATCCGGAACCAACCGGTACAAAAAAAGA
CAAAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTCG
GCTACNACTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAACAATAAATAT
TCCGTTAACATAAAAGAGTTGCTAAGAAACGATATGTCAATTCTGGCGGCAACAAGCATCTTAACATAAAA
ACCCGAAGACGGAACATCGGAAAACGGCACATTCCACGCCGCCTCTTCTCTCGGCTTGTCCGCCGTTTACG
ATTTCGATACCGGTTCCCGCTTCAAACCCTATATCGGCATGCGCGTCGCCTACGGACACGTCAGACATCAG
GTTCGTTCGGTTCAACAAGAAACCATTGCTGTTACCACTTACCCACAGAATGCTGCGTCAAGTGTTACCAC
AAATGCTCCGATCCGCAAACTTCCCCATCACGAAAGCCGCAGCATCAGCAGCTTGGGCTTCGGCGCAGTGG
CAGGCGTAGGCATCGACATCACGCCCAAGCTGACCCTGGACGCCGGCTACCGCTACCACAACTGGGGACGC
TTGGAAAACACCCGCTTCAAAACCCACGAAGCCTCATTGGGCGTGCGCTACCGCTTCTGATTCCCCGATAC
CGATGCCGTCTGAACCTTCAGGCGGCATTTTTAATCGCCCGCTGTTTACAGGCGCGGGGCGGGCGTGGAAA
TACCCGAACCGTCATTCCCGACAACACCGCAATCTCGAAATTCGTCATTCCCGCGCAGGCGGAAATCCGGA
CCCCCGACGCGGCGGGAATCTATCGGAAATGACTGAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAAT
GACGGTTCAGCAAGCGTAGGTCGGATACTTGTATCCGACAAAACCTTTAACATTCCCATCATTGCAATCCA
TTGCAGCAATGTCCAAAATGTCGAATTCAAGAATCCGGCCTACAAAATCATTCCGAGCATAATACTATGAA
ATACCGTCGTTTTTACGCGCAATGGCGGCACTTACTTTTTTACGGTTGTAACCAATAAACGGCAGAAGATT
TTGACCGATGATGCGGTGCGTTTGGCTTTACGGCAGGCGGTAATGGCGGTGCGCGAACGGAATCCGTTTGA
AATTTTGGCATGGGGGTTGATGCCCGATCATTTGCATACCATATGGCGGGGGCCGGACAAAGATTTTGCTT
ATTTGGAACGCGGGCGGCAAATCAAGCGGCCCAGCCAATATTTAATTGGCGGCAATTTCAGGCTTTGGAAA
AACGCTTTTGGGAACATACTATTCGCGGTGAGGCCGATTTTGCCTGCCATTTTGATTATTTGCATTTCAAT
CCGGTCAAACAAGGCTATGTAGGACAAATTTCCGATTGGCGGTTTTTTACGTTTCACCCCTATGTCAAACC
GGGTATTTATCCGCATAATTGGGGTGGGGGCAAAGCGGACTTTTTTATTGAATACGATTGAAGTAAAGTTG
GATTCGAGAATCCGACCTACGGAAAAATGAAAGAGCATCGGCTGCTGGACGGCATTATGTGCAAAACCCGC
CGCAACCGCCCGCTGCCGGAGGCGCAAACCAAACGCCACCGATTTGTAAAACCCGTAGTCAAGAAANT

>WT_9_Final (3623)
CCGAAAAAGTCAGACTGCGCTCGGCGAAACCCGCATCAGCGCGCCGCGTTTTACCGCGATATTGACCCCGC
CGCAGGCAGCCTTTCCGCCTACAATCAACTGCAAGGCAAAGAATTGTCTTACAACAACATCGCCGATGCCG
ATGCCGCTTGGGAAGCCGTCAAATCCTTTGACGCGCCCGCCTGCGTGATCGTGAAACACGCCAATCCGTGC
GGCGTAGCCGTTGCAGCCGATACCTTGACCGCCTACAAACTCGCCTACGCCACCGACACCACCAGCGCGTT
CGGCGGCATCATCGCCTTCAACCGCGAAGTCGACGGCGAAACCGTCAAACAGATTACCGACAACCAATTTA
TGGAAGTCCTCATGGCGCCGAAGTTTACCGCCGAAGCCCTTGAAATCGCCGCCGCCAAGAAAAACGTGCGC
GTATTGGAAGTGCCGCTCAAAGCAGGTGCCAACCGCTTTGAACTCAAACGCGTCGGCGGCGGACTGTTGGT
ACAAACGCCCGACATCAACCGCATCAACCGCGCCGATTTGAAAGTCGTCTCCAAACGCCAACCGACCGAGC
AGGAATGGAACGATTTGCTGTTTGTCTGGAACGTCGCAAAATACGTCAAATCCAACGCCATCGTCTTCGGC
AAAGGCGGCCAAACCTACGGTATCGGCGCAGGCCAAATGAGCCGCGTGGACAGCACCCGCATCGCCGCCCG
CAAAGCGCAAGATGCCGGTCTCGACCTCAACGGCGCATGTGCCGCTTCCGATGCCTTCTTCCCATTCCGCG
ACGGCGTGGACGTGATTGCCGAACAGGGCATCAAAGCCATCATCCATCCGGCAGGCTCGATGCGCGATCAG
GAAGTTTTCGACGCGGCGGACGAACACGGCATCGCCATGGCGGTAACCGGCATCCGCCACTTCCGCCATTA
AGGCAGACGAACAAGGCAATGCCGTCTGAAGGGCTTTCAGACGGCATTTTGCGCCATTTACCGTAATCCTG
AAACCCCCCATTCCCGCGCAGGCAGGAATCTAGATGAGAATCCATCCTGTCAGAAATACCCGAATCGTCAG
GAATACCCGTGTCATCATTCCCGATACACCGTAATCCTGAAACCCGTCATTCCCGCGCAGACAGAAATCCG
GATTTGTCCGCGCGGAAACTTATGCGCCGCCATTCCCGCGAAGGCGGGAATCCGGACCGTTCGGTTTCGGT
TATTTCCGATAAATTCCTGCTGCTTTTCATTCCTAGATTCCCACTTTCGCGGGAATGACGAAGGAGTGGGA
ATCCGGTTTTTGAGTTCCGGCCATTTCCGACAAATTGCTTTGGCATTGGATATTTCCTATTTTAATCCACT
ATAGTGGATTAACTTTAAACCAGTACGGCGTTGCCTCGCCTTAGCTCAAAGAGAACGATTCTCTAAGGTGC
TGAAGCACCAAGTGAATCGGTTCCGTACTATTTGTACTGTCTGCGGCTTCGCCGCCTTGTCCTGATTTTTG
TTAATCCGCCATATCCGATCGCCGCACCGGAATTTTAAAAAATAAATACATTCTTTAACAAAAAAAATCAT
ATCCCGGCAAAACAAATCTGATAAAATGCTTGGCGTTTATTAACAATCTTTTAATAATTTAATCTTAAATA
TTCGGAGTTGATATGCATACGGTCGACAAAAAACTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTT

CTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCGGGCGGCGGGTGAAGACCATGGGCGCGGCCCGTATGT
GCAGGCGGATCTGGCTTACGCCTACGAGCACATTACCCGCGATTATCCCGATGCAGCCGGTGCAAACAAAG
GCAAAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTC
GGCTACGACTTCGGCGGCTGGCGCATCGCCGCGGATTATGCCCGTTACAGGAAATGGCACAACAATAAATA
TTCCGTGAACATAAAAGAGTTGGAAAGAAAGAATAATAAAACTTTTGGCGGCAACCAGCTTAACATAAAAT
ACCAAAAGACGGAACATCAGGAAAACGGCACATTCCACGCCGTTTCTTCTCTCGGCTTGTCAACCGTTTAC
GATTTCAGAGTCAACGATAAATTCAAACCCTATATCGGTGTGCGTGTCGGCTACGGACACGTCAGACACGG
TATCGATTCGACTAAAAAAACGAAAAATACTCTTACCGCCTACCATAGTGCTGGCACAAAACCTACGTATT
ATGATGATATAGATTCGGGAAAAAACCAAAAAAACACTTATCGCCAAAACCGCAGCAGCCGCCGCTTGGGC
TTCGGCGCGATGGCGGGCGTGGGCATAGACGTCGCGCCCGGCCTGACCTTGGACGCCGGGTACCGCTACCA
CTATTGGGGACGCCTGGAAAACACCCGCTTCAAAACCCACGAAGCCTCATTGGGCGTGCGCTACCGCTTCT
GATTCCCCGATACCGATGCCGTCTGAACCTTCAGACGGCATTTTTAAGGCGCAAGATCGGGCAAACGGCAT
TTTCAGACGGCATAACTGACAGTATAATCCGAACATCCGGCCGCTTTCGCGCGGCCTTCAGCATTATCCGA
TTTTTTCCGAAAGCCGAACCATGCAATACAAACCCCTCCTGCTCGCCCTGATGCTTGGCAGGCGGCGGGTG
AAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATCTGGCTTACGCCTACGAGCACATCACCCGCGATTAT
CCCGAAGCAACCGGTGCAAACCAAGGCAAAATAAGCACGGTAAGCGATTATTCAAAAACATCCGCACCCGC
TCCGTCCACCCCCGACTCGCCCTCGGCTACGATTTCGGCGGCTGGCGCGTCTTTTCCGCCCCCGCCGTTGC
CGCCGGCCTCCACCAATCCCTTCAATATTACCGATCCCGCCATTGCCGCCGATTCCGTCAAATCCATCAAT
TCCGCCGAATCACGCCTATTCCCCAAAAACCTTGATGCGGCGGGCTGAAGCCCGCCCTGCAACCCTCTCTA
TGCACCCCCTTGCGAGCCCGACACTACGCAACATCTTGAGAACCCATCCTGTCAAGAATACCCGAACCGTC
CCGATACACCGTAATCCTAAAACCCGCCATTCCCGCGCAGGCGGGAATCCAGTCCGTTCGGTTTCAGTCAT
TTCCAAAAATTGCCGTAGCGTTAAGTTTCTAGATTCCCACTTTCGGGGGAATGACGCGGTGCAGGTTTCCG
TGCGGACGGATTCGTCATTCCCACACAGACAGGAATCCGGATTTGTCCGCGCGGAAACTTATGCGCCGTCA
TTCCCGCACAGGCGGGAATTCAGGTCTGTCGGCATAGGAACTTATCGGATAAAACGGTTGCCCCAACCCTG
CGTTCTAGATTCCCACTTTCGTGGGAATGACGATTCAGGTATTCCTAATAGCAACCCGCCACAACCCACCT
ACGCCCATCCTACGCATCTGCCATACAGCCCTGTTCCGGAGCGCGCGGATAAGGTGCGATGCCGCCGGCGG
GCATCATAGTCCTCTCTCCCTGTGGGAGAGAGCTAGAAAGCTGCAAGGCTTGTATTTTAGAAGACTAAGGG
A

>WT_10_final (3259)
AAGCCGCGATCGTTCCCTTGAGCAAAACGCTTCATGCCCACCGACACAGGCGACATCGTCAATAAATTCCT
GACCGAACACTTCGCCCAATACGTCGATTACCACTTCACCGCCAAACTCGAAGACCAGCTTGACGAAATCG
CCGACGGCAAACGCCGCTGGATTCCCGTGATGGACAAATTCTGGAAACCGTTCATCAAACAAGTGGAAGAA
AAAGAGGGCATCGAACGCGCCAAATTTACCACGCAGGAACTTAACGAAACCTGCCCGAAATGCGGCGAACA
CAAACTGCAAATCAAGTTCGGCAAAATGGGCCGCTTCGTCGCCTGCGCCGGCTATCCCGAATGCAGCTACA
CGCGCAACGTCAACGAAACCGCCGAAGAAGCCGCCGAGCGCATCGCCAAAGCCGAGGCTGAACAAGCCGAA
CTCGACGGACGCGAATGCCCCAAATGCGGCGGTCGGTTAGCGTATAAATACAGCCGTACCGGCAGCAAATT
CATCGGTTGCGCCAACTACCCCAAATGCAAACACGTCGAGCCGTTGGAAAAACCCAAAGATACCGGCGTCC
AATGCCCGCAATGCAAAAAAGGCAACCTCGTCGAGCGCAAATCCCGCTACGGCAAACTGTTTTACAGTTGC
AGCACCTATCCCGACTGCAACTACGCCACTTGGAACCCGCCCGTTGCCGAAGAATGCCCGAACTGCCATTG
GCCGGTCTTGACCATCAAAACCACTAAACGCCGGGGCGTGGAAAAAGTCTGTCCGCAAAAAGAATGCGGCT
GGAAAGAACAGATTGAACCGCCTGCACCGCAAGAGTGAAGCGGTTGGGATTGAGATGAAAGAAAATGCCGT
CTGAAAGGGGTTTCAGACGGCATTTGTAAATTAGAAGGGGCTGTCCCAGATGGCTGGGGAAATTTAAATTA
AGTCAGAATTATCCCTATGAGAAAAGCCGTCCAAGCCGGTACAAACAAAAGAAACCCATCGGACTGTTTGC
CGCAGGGGTAACTGCAAGAACGGCAGCAGGGTTGGCAGCAGTTAATCAAAACACGGCAGCCTGTTATTTTC
TATGGCGAAATAGTGTTTGGCACATCATCGGGAAAACTTCCTTTACACGCATTTGGACGAAATCTGCGAAA
TTATGGAGGCATACGACATATCGTTTACGATTACTTATCCGTCAAAACAGTCCGCATTTGGAAATGTATGA
CGGCGAAGCAGGGGCGGGTGAAAGTTATTTCCGTAAAGGCAGACGCGGCCGCAGTGCCGCCGGCAAAGCCG
CCGTATTCGGGCTTTTGAAGCGCAACGGCAAGGTTTGTACGGTTACTGTCGGACACTCGAACCGCTATTTT
ATTGCCTATTATCCGCGGACGGGTGAAACCCGGCAGCATTGTTTATACGGATTGTTCGTAGTTACGATGTA
TCAGATTGTAAGCGGATTTGGCCGTTTCCGTATCAATCGCGGCACACATTTTGCCGAACGGCAAAACCATA
TCAACGCAATTGGGAACTTTTGGAATCGGGCAAACGTCATTTGCGCAAGTTTGACGGCATTCCCAAAGAGC
ATTCGGCTGTATTGAAGGAGTGCGGACGGCGTTTGACAACAGTGGGTACAAGTTCGAATTCCATTTAAAAC
AATTGGTAAAGCAGGATTTATCCCGGTTGGCCGGGATTTTAAAAAATAAATACATTCTTTAACAAAACAAA
TCATATCCCGGCAAAACAAATCTGATAAAATGCTAGGCGTTTATTAACAATCTTTTTAATAATTTAATCTA
AGATATTTGGAGTTGATATGCATACGGCCGACAAAAAACTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTT
CTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTC

TCTTCTCTTCTCTTCTCTTCCGCAGCGCGGGCGGCAAGTGAAGGCAATGGGCGCGGCCCGTATGTGCAGGC
GGATTTAGCCTACGCCGCCGAACGCATTACCCACGATTATCCGGAACCAACCGGTGCAAAAAAAGACAAAA
AAATAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCGTCCACCCCAGGGTGTCGGTCGGC
TACGACTTCGGCGGTTGGAGGATAGCGGCAGATTATGCCCGTTACAGGAAATGGAACGACAATAAATATTC
CGTCGACATAAAAGAGTTGGAAAACAAGAATCAGAATAAGAGAGACCTGAAGACGGAAAATCAGGAAAACG
GTACGTTCCACGCCGTCTCCTCGCTCGGTTTGTCAGCCGTTTACGATTTCAAACTCAACGATAAATTCAAA
CCCTATATCGGTGCGCGCGTCGCCTACGGACACGTCAGACACAGCATCGATTCGACCAAAAAAACAACAAA
GTTTCTTACCTCCTCCTATGGTGGCTTAAACCCTACGGTTTATACTGAGGAAAATACGCAAAACGCCCATC
ACCAAAGTAACAGCATCCGCCGCGTGGGCCTCGGCGTCATCGCCGGCGTCGGTTTCGACATCACGCCCAAG
CTGACCCTGGACACCGGCTACCGCTACCACTATTGGGGACGCTTGGAAAACACCCGCTTCAAAACCCACGA
AGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCCCCGATACCGATGCCGTCTGAACCTTCAGACGGCATT
TTTGATGCGCCCGCCGTTTACAGGCGCGGGGCGGGCGCAGTAAAATACCCGAACCGTCATTCCCGACAATA
CCGCAATCTCGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGATTTGTCGGGATTCGCTCAAATCGGGT
GGCGGAGGGATGGTGTTGCAGATTAACGATGTGGATACCAACAGCGGGGAATTTGCAATCTATACCGCTCA
GGATGCATCGGTAAAGCTGGATTTGTTAAAAATGGAATTAAAACACTGTAAAGAAATGTTGGAACAAAAAG
ACAAAGAAATCGAGTTGCTCCGCAAGCTGACCGAAACCGTTTAAACAGATATGCCGTCTGAAAAAAGTTTT
CAGGCGGCATATTCTTCGACAGGTCTTGTATAATACCGTTTGAACTTTCAGGCTTTTGATTATGGCGGCAG
GCAAACATACCAAACACAGCAACCGGGTACGCATTATCGGCGGGCAATGCCGGGGCAGGAAATTGAGTTTT
GCATCCGCCGACGGACTGCGCCCAACCCCCGACAGCGTGCGGAAAAGCTGTAACGGTGACAGTT

>WT_11_Final (2923)
CCGCATCATAGCGAGGCGTAGCCAGAGGATTTCCCCTTTCATTAGGAGTAATTTTATGAATACCCTTCAAA
AAGGCTTTACCCTTATCGAGCTGATGATTGTGATCGCTATCGTCGGCATTTTGGCGGCAGTCGCCCTTCCC
GCCTACCAAGACTACACCGCCCGCGCGCAAGTTTCCGAAGCCATCCTTTTGGCCGAAGGTCAAAAATCAGC
CGTCACCGAGTATTACCTGAATCACGGCGAATGGCCGGAAGACAACACTTCTGCCGGCGTGGCATCCTCCC
CCACCGACATCAAAGGCAAATATGTTCAAAGCGTTACGGTCGCAAACGGCGTCGTTACCGCCGAAATGGCT
TCAACCGGCGTAAACAAAGAAATCCAAGGCAAAAGACTCTCCCTGTGGGCCAAGCGTCAAGACGGTTCGGT
AAAATGGTTCTGCGGACAGCCGGTTACGCGCGGCGCCGGCAACGCCGGCAAAGCCGACGACGTCACCAAAG
CCGGCAACGACAACGAAAAAATCAACACCAAGCACCTGCCGTCAACCTGCCGCGATAACTTTGATGCCAGC
TGAGGCAAATTAGGCCTTAAATTTCAAATAAATCAAACGGTAAGTGATTTTCCACGGCCGCCCGGATCAAC
CCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGGGGTCGTCCGTTCCGGTGGAAATAATATATCGA
TTGCGCTTCAAGGCCCTGCATGTGCCTCATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCA
ATCCGCCCCCTCATTTGCCGAACAAGCGGTCCGGACTCCCGCCCGCGCGGGAATGACGGCTGCAGATGCCC
GACGGTCTTTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGGCAGTACAAATGGTACGGA
ACCGATCCGCCCGGTGCTTCATCACCTTGGGGAACCGTTCCCTTTGGGCCGGGGCGGGGCAACGACGTACC
GGTTTTTGTTCATCCGCCATATTGTGTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATCA
GTGAAAATCTTTTTTTAACCGGTCAAACCGAATAAGGAGCCGAAAATGAATCCAGCCCCCAAAAAACCTTC
TCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCGGCGGGTGAAGACCATGGGCGCGG
CCCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAACACATTACCCACGATTATCCGGAACAAACCGATC
CAAGCAAAGGCAAAATAAGCACGGTTAGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACCCCCGG
GTGTCGGTCGGCTACGAATTTCGGCGGCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGAACAC
ATTAAATATTCCGTGAGCATTAAAGAGTTGCTTAGAAACAAGGTCAATGGCAACAGGACAGACCGGAAGAC
GGAAAATCAGGAAAACGGTACGTTCCACGCCGTTTCTTCTCTCGGCTTGTCCGCCGTTTACGACTTCAAAC
TCAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACGTCAGACACAGCATCGATTCG
ACTAAAAAAATAACAGGGCTTCTTACCACCAGTACTCCTGGCATAATGTCTGGGGTTTATAAGGTATTAAG
GACACCAGGCGCCCATCGCGAAAGCGACAGCATCCGCCGCGTGGGTCTCGGTGTCATCGCCGGCGTCGGTT
TCGACATCACGCCCAAGCTGACCCTGGACACCGGCTACCGCTACCACTATTGGGGACGCTTGGAAAACACC
CGCTTCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCCCCGATACCGATGCCGTCTG
AACCTTCAGACGGCATTTTTGATGCGCCCGCCGTTTACAGGCGCGGGGCGGGCGCAGTAAAATACCCGAAC
CGTCATTCCCGACAATACCGCAATCTCGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGACCTGTCCGC
ACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGG

TTCGGTCTGCCGTTTTCGGACGGCATTTCGGCTCAATCCAGCAGTGCGTCCACAAACGCGCGCGCGTCAAA
CGGGCGCAGGTCGTCTATGCCTTCGCCCACGCCGATGTAGCGGACGGGGACGGGGCGGTCGGAAGCAAGCG
CGGCGAGGATGCCGCCTTTTGCCGTGCCGTCGAGTTTGGTAACGATAAGCCCCGTCAGCCCCAATGCGTCG
TCAAAGGCTTTGACTTGGTTGACGGCGTTTTGCCCGATATTGGCATCGAGTACGACGATAATTTCGTGCGG
CGCGCCGGGAATGGCTTTTTGCAGCACGCGCATCGCTTTTTTGATTTCTTCCCTCAAAAGAAAAAAAATGG
GCAGGCGGCCGGCGGTGTTGTCAAGCAAAATGTTGATCCCCCACGCTTTGTCGGCTTGGACGGCATTGAAA
CACACGGAGGGGGAAATGCCCGTGGTTTGTGAAAAAAAGGTTACATTGTTGGGCCCGCCCCCAGCCAGAAA
CTGCTCAACGGGCGGCGGCGCGGAAGGTATCCCCCCCCCCCAGCAAAACGGATTTACCCTGCGCTTGGAAA
TATTTGGCGAGTTTGCCGAAAGAAGAGGTTTTGCCCGCCCCGTTGATGCCGGCAAGCATAATCACGAAAGG
CTCTTTAGTTTTGGGCAAGCCCAGCGGTTTTTCCAACGGCTTAATCAGGTCGTACAAGGCTTCTTTTAACG
CCCCGCCCAATTCGTAGCCGTTTTTAAACCCCTTGAAACTGGAGGGGCCGCGCACGTCTTTCATCAGGTAT
ACGGTGGCCACCNGNGCAGATCGCCGGATCCGTCGCGT

A.2. *opa* Alignments

Verification of *opa* genes deletion. PCR was used to amplify the adjoin regions of *opa* that were deleted from the various constructs. The DNA sequences of the amplicons were determined, and aligned to the original DNA sequence. Primers used to generate the deletions are highlighted.

```
WT_1_Final    ABNNNNG-GATGGTCGCATCGACATCCCGACCGCGACATCCCGCACGAAA-ACAAGCCGC
              ::::::  ::: :: :::::::::::: :::::::::::::::::::::: :::::::::
Del_1_Final   ABNNNNTTGATCGT-GCATCGACATCC-GACCGCGACATCCCGCACGAAACACAAGCCGC
                       10         20         30         40         50

              60         70         80         90        100        110
WT_1_Final    CCCGAAAGGCAAATGCCGAAAAATCGCGATAAAAACCATCAATCGGAAAATATTGCATAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   CCCGAAAGGCAAATGCCGAAAAATCGCGATAAAAACCATCAATCGGAAAATATTGCATAA
              60         70         80         90        100        110

             120        130        140        150        160        170
WT_1_Final    ACAGGCGAAAATATATCATTGAAAAGAAAACCCATACAATGTCTGACAGAAACAAGCCGC
              :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::::
Del_1_Final   ACAGGCGAAAATATATCATTGGAAAGAAAACCCATACAATGTCTGACAGAAACAAGCCGC
             120        130        140        150        160        170

             180        190        200        210        220        230
WT_1_Final    TTTGCCATTGTTTCCTAACAGTTAAACCCCGCCCTTCAAGGCGGCGGGGCAGGGCTTGAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   TTTGCCATTGTTTCCTAACAGTTAAACCCCGCCCTTCAAGGCGGCGGGGCAGGGCTTGAC
             180        190        200        210        220        230

             240        250        260        270        280        290
WT_1_Final    TCAAAATGGCGCAAGCCCCTGCCCTCAAATCCAACACGCAGGATTAAACCATAATAGCGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   TCAAAATGGCGCAAGCCCCTGCCCTCAAATCCAACACGCAGGATTAAACCATAATAGCGG
             240        250        260        270        280        290

             300        310        320        330        340        350
WT_1_Final    CTTTCTTATTATTTCTTATTGAAACACCGCCCGGAACCCGATATAATCCGCCTTTGAAGC
              :::::::::::::::::::::::::::::::::::::
Del_1_Final   CTTTCTTATTATTTCTTATTGAAACACCGCCCGGAACCC--------------------
             300        310        320        330

             360        370        380        390        400        410
WT_1_Final    ATCAGTGAAAATCTTTTTTTTAACCGGTTAAACCGAATAAGGAGCCGAAAATGAATCCAG

Del_1_Final   ------------------------------------------------------------


             420        430        440        450        460        470
WT_1_Final    CCCGCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCCGCCGCGCAGGCGGCAAGTGAAGG

Del_1_Final   ------------------------------------------------------------
```

```
            480       490       500       510       520       530
WT_1_Final    GAATGGCCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCGCCGAACGCATTACCCA

Del_1_Final   ------------------------------------------------------------


            540       550       560       570       580       590
WT_1_Final    CGATTATCCGGAACAAACCGCTCCAAAAAAAGCACAATTAAGCACGGTAAGCGATTATTT
                                                                        :
Del_1_Final   ------------------------------------------------------------TG


            600       610       620       630       640       650
WT_1_Final    CAGAAACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTCGGCTACGACTTCGGCGG
              ::::
Del_1_Final   CAGA--------------------------------------------------------
            340

            660       670       680       690       700       710
WT_1_Final    CTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAACAGTAAATATTCCGT

Del_1_Final   ------------------------------------------------------------


            720       730       740       750       760       770
WT_1_Final    CAGCATAAAAGAGTTGGGAAGAAACGATAATAGCACTTCTAACAGCAGCCATCTTAACAT

Del_1_Final   ------------------------------------------------------------


            780       790       800       810       820       830
WT_1_Final    AAAAACCCAAAAGACGGAACATCAAGAAAACGGCACATTCCACGCCACTTCTTCTCTCGG

Del_1_Final   ------------------------------------------------------------


            840       850       860       870       880       890
WT_1_Final    CTTATCAGCCATTTACGATTTCAAACTCAACGATAAATTCAAACCCTATATCGGCGTGCG

Del_1_Final   ------------------------------------------------------------


            900       910       920       930       940       950
WT_1_Final    CGTCGCCTACGGACACGTTAAACATCAGGTTCGTTCAGTGGAAAGCGAAACCACGACTGT

Del_1_Final   ------------------------------------------------------------


            960       970       980       990      1000      1010
WT_1_Final    TACCACTCACAATGGAGCCCCTGTCCCACAAGGTCCGACCCCCAAACCTGCCTATCACAA

Del_1_Final   ------------------------------------------------------------


           1020      1030      1040      1050      1060      1070
WT_1_Final    AAGCCGCAGCATCAGCAGCTTGGGGCTTCGGGGCAGTGGCAGGCGTAGGCATCGACTTCA
```

```
Del_1_Final   ------------------------------------------------------------

            1080      1090      1100      1110      1120      1130
WT_1_Final    CGCCCAAGCTGACCCTGGACGCCGGCTACCGCTACCACAACTGGTGACGCTTGTAAAACA

Del_1_Final   ------------------------------------------------------------

            1140      1150      1160      1170      1180      1190
WT_1_Final    CCCGCTTCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCGCCGATT
                        :::::::::::: :::::: :::::::::::::::::::: :::::
Del_1_Final   -------------CCACGAAGCCTCATTGGGCATGCGCTACCGCTTCTGATTCCCCGATA
                              350       360       370       380       390

            1200      1210      1220      1230      1240      1250
WT_1_Final    CCGATGCCGTCTGAACCTTCAGACGGCATTTTTGACGCGCCCGCCGTTTACGGGCGCGGG
              :::::::::::::::::::::::::::::::: :: :: ::::::::: ::::::::
Del_1_Final   CCGATGCCGTCTGAACCTTCAGACGGCATTTTTGATGCACCTGCCGTTTACAGGCGCGGG
                     400       410       420       430       440       450

            1260      1270      1280      1290      1300      1310
WT_1_Final    GCAGGTGCGGGGAAATACCCGTACCGCCATTGCCGATACACCGTAACCCTAAAACCCGCC
              :: :: :: :  ::::::::: ::::                                 :
Del_1_Final   GCGGGCGCAGTAAAATACCCGAACCG---------------------------------TC
                     460       470

            1320      1330      1340      1350      1360      1370
WT_1_Final    ATTCCCGACAACACCGTAATCTCGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGACC
              :::::::::: ::::::::::::::::::::
Del_1_Final   ATTCCCGACAATACCGTAATCTCGAAACCCGTC--------------------------
               480       490       500       510

            1380      1390      1400      1410      1420      1430
WT_1_Final    CCCGACGCGGCGGGAATCTATCGGAAATGACTGAAACCCCGCGTCCTAGATTCCCACTTC

Del_1_Final   ------------------------------------------------------------

            1440      1450      1460      1470      1480      1490
WT_1_Final    CGTGGGAATGACGGTTCGGTTGCTACGGCCCGCTGATTCCCCGACACCGATGCCGTCTGA

Del_1_Final   ------------------------------------------------------------

            1500      1510      1520      1530      1540      1550
WT_1_Final    ACCTTCAGACGGCATTTTTGATGCGCCCGCCGTTTACAGGCGCGGGGCGGGCGCAGTAAA

Del_1_Final   ------------------------------------------------------------

            1560      1570      1580      1590      1600      1610
WT_1_Final    ATACCCGAACCGTCATTCCCGACAACACCGCAATCGCGAAACCCGTCATTCCCGCGCAGG
                              ::::::::::::::::   ::::::::::::::::::::
Del_1_Final   -------------------CGACAACACCGCAATCTTAAAACCCGTCATTCCCGCGCAGG
                                       520       530       540       550
```

```
            1620      1630      1640      1650      1660      1670
WT_1_Final    CGGAAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGC
              ::: ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   CGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGC
                   560       570       580       590       600       610

            1680      1690      1700      1710      1720      1730
WT_1_Final    GTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCGTAGGGCCGGGTGGTCGAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   GTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCGTAGGGCCGGGTGGTCGAA
                   620       630       640       650       660       670

            1740      1750      1760      1770      1780      1790
WT_1_Final    AGGGCGGATTCGATGGATTCGATGGATTCGATGAAAACGGTAGACATGTTGGATTGATTG
              ::::::::::::::                  ::::::::::::::::::::::::::::::
Del_1_Final   AGGGCGGATTCGA------------------TGAAAACGGTAGACATGTTGGATTGATTG
                   680                         690       700       710

            1800      1810      1820      1830      1840      1850
WT_1_Final    TTTATTGTTGAGGAAAATATAAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTATCTT
              :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::::
Del_1_Final   TTTATTGTTGAGGAAAATATGAGATCTTCTTTCCGGTTGAAGCCGATTTGTTTTTATCTT
                 720       730       740       750       760       770

            1860      1870      1880      1890      1900      1910
WT_1_Final    ATGGGTGTTATGCTATATCATCATAGTTATGCAGAAGATGCAGGGCGCGCGGGCAGCGAG
              :::::::::::::::::::::::::::::::: ::::::::::::::::::::::::::
Del_1_Final   ATGGGTGTTATGCTATATCATCATAGTTATGCCGAAGATGCAGGGCGCGCGGGCAGCGAG
                 780       790       800       810       820       830

            1920      1930      1940      1950      1960      1970
WT_1_Final    GCGCAGATACAGGTTTTGGAAGACGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACAAA
              :::::::::: :::::::::::: ::::::::::::::::::::::::::::::::::::
Del_1_Final   GCGCAGATACGGGTTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTACCGAAAGACAAA
                 840       850       860       870       880       890

            1980      1990      2000      2010      2020      2030
WT_1_Final    AAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCAGGATATATTCAAATCCGGCGAA
              ::::::::::::::::::::::::::::::::::::::::: : ::::::::::::::
Del_1_Final   AAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCAGGATGTGTTCAAATCCGGCGAA
                 900       910       920       930       940       950

            2040      2050      2060      2070      2080      2090
WT_1_Final    AACCTCGACAACATCGTACGTAGCATACCCGGTGCGTTTACACAGCAAGATAAAAGCTCG
              :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::::
Del_1_Final   AACCTCGACAACATCGTACGCAGCATACCCGGTGCGTTTACACAGCAAGATAAAAGCTCG
                 960       970       980       990      1000      1010

            2100      2110      2120      2130      2140      2150
WT_1_Final    GGCATTGTGTCTTTGAATATTCGCGGCGACAGCGGGTTCGGGCGGGTCAATACGATGGTG
              :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   GGCATTGTGTCTTCGAATATTCGCGGCGACAGCGGGTTCGGGCGGGTCAATACGATGGTG
                1020      1030      1040      1050      1060      1070

            2160      2170      2180      2190      2200      2210
WT_1_Final    GACGGCATCACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTTCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
Del_1_Final   GACGGCATCACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTTCA
                1080      1090      1100      1110      1120      1130

            2220      2230      2240      2250      2260      2270
WT_1_Final    TCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCGTCAAAGGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   TCTCAATTCGGTGCATCTGTCGACAGCAATTTTATTGCCGGACTGGATGTCGTCAAAGGC
                1140      1150      1160      1170      1180      1190

            2280      2290      2300      2310      2320      2330
WT_1_Final    AGCTTCAGCGGCTCGGCAGGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_1_Final   AGCTTCAGCGGCTCGGCAGGCATCAACAGCCTTGCCGGTTCGGCGAATCTGCGGACTTTA
                1200      1210      1220      1230      1240      1250

            2340      2350      2360      2370      2380       2390
WT_1_Final    GGCGTGGATGACGTCGTTCAGGGCAATAATACCTACGGCCTG-TGCTAAAGTTGACCGCC
              :::::::::::::::::::::::::::::::::::::::::: ::::::::::  ::::::
Del_1_Final   GGCGTGGATGACGTCGTTCAGGGCAATAATACCTACGGCCTGCTGCTAAAGTCTACCGCC
                1260      1270      1280      1290      1300      1310

             2400      2410
WT_1_Final    CATGTCANNNNNT
              :::::::  :::::
Del_1_Final   CATGTC-CNNNNT
                1320


                     10        20         30        40        50
WT_2_Final    ABNNNNCCGGAGGAGACAGT-TGTATACGGTTGTCGCTAACCCGGAAAATTTCGACAGCA
                    : : : :::::::: :: : :: : :: :::  :: :::::: ::: : :::::
MS11_del_2_F  ----CCCGGAATGAGACAGTCTGAACACTGGTGGCGCATACTCGGAAAGTTTGGGCAGCA
                         10        20        30        40        50

             60        70        80        90       100       110
WT_2_Final    GCCGCCACACTGGGCATAGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GCCGCCACACTGGGCATAGGACGCAGCACATGGAGCGAAAACAGTGCAAATGCAAAAACC
                60        70        80        90       100       110

            120       130       140       150       160       170
WT_2_Final    GACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGTGGGCGATATCGGCTATCTCAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGTGGGCGATATCGGCTATCTCAAA
               120       130       140       150       160       170

            180       190       200       210       220       230
WT_2_Final    GGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGATGAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GGCCTGTTCTCCTACGGACGCTACAAAAACAGCATCAGCCGCAGCACCGGTGCGGATGAA
               180       190       200       210       220       230

            240       250       260       270       280       290
WT_2_Final    TATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGTGGTGTCAAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  TATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGTGGTGTCAAC
```

```
              240       250       260       270       280       290

            300       310       320       330       340       350
WT_2_Final    GTTCCGTTTGCCGCAACGGGAGATTTGACGGTTGAAGGCGGTCTGCGCCACGACCTGCTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GTTCCGTTTGCCGCAACGGGAGATTTGACGGTTGAAGGCGGTCTGCGCCACGACCTGCTC
               300       310       320       330       340       350

            360       370       380       390       400       410
WT_2_Final    AAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  AAACAGGATGCATTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGCAACAGCCTCACT
               360       370       380       390       400       410

            420       430       440       450       460       470
WT_2_Final    GAAGGCACACTGGTCGGACTCGCGGGTCTGAAGCTGTCGCAACCCTTGAGCGATAAAGCC
              :::::::::::::::::::::::::::::: ::::::::::::::::::::::::::::
MS11_del_2_F  GAAGGCACACTGGTCGGACTCGCGGGTCTGAAACTGTCGCAACCCTTGAGCGATAAAGCC
               420       430       440       450       460       470

            480       490       500       510       520       530
WT_2_Final    GTCCTGTCTGCGACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACGCGGTAACG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GTCCTGTCTGCGACGGCGGGCGTGGAACGCGACCTGAACGGACGCGACTACGCGGTAACG
               480       490       500       510       520       530

            540       550       560       570       580       590
WT_2_Final    GGCGGCTTTACCGGCGCGGCTGCAGCAACCGGCAAGACGGGTGCACGCAATATGCCGCAC
              :::::::::::::::::::::::::
MS11_del_2_F  GGCGGCTTTACCGGCGCGGCTGCAG----------------------------------
               540       550       560

            600       610       620       630       640       650
WT_2_Final    ACCCGCCGGGTTGCCGGTCTGGGGGTGGATGTCGAATTCGGCAACGGCTGGAACGGCTTG

MS11_del_2_F  ------------------------------------------------------------


            660       670       680       690       700       710
WT_2_Final    GCACGTTACAGCTACACCGGTTCCAAACAGTACGGCAACCACAGCGGACAAATCGGCGTA

MS11_del_2_F  ------------------------------------------------------------


            720       730       740       750       760       770
WT_2_Final    GGCTACCGGTTCTGACGGACAGAAAACAGACAGCCGCAAAGATCACCGTCTTTGCGGCTG

MS11_del_2_F  ------------------------------------------------------------


            780       790       800       810       820       830
WT_2_Final    TTTCTTATGAAAAGAAAACCCTATTCCAATTGCCTGCTTCTATTGTTTCAAGACTTCTTC

MS11_del_2_F  ------------------------------------------------------------


            840       850       860       870       880       890
```

```
WT_2_Final    CAAAGATTCGGCATCAATCAGACGTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGG

MS11_del_2_F  ------------------------------------------------------------


            900       910       920       930       940       950
WT_2_Final    GCCGCAGGCAGTACGGATGGTACGGAACCGATCCGCCCGGTGCTTCAGCACCTTAGGGAA

MS11_del_2_F  ------------------------------------------------------------


            960       970       980       990      1000      1010
WT_2_Final    CCGTTCCCTTTGAGCCGGGGCGGGGCAACGCCGTACCGGTTTTTGTTCATCCGCCATATT

MS11_del_2_F  ------------------------------------------------------------


           1020      1030      1040      1050      1060      1070
WT_2_Final    GTGTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATCAGTGAAAATCTTT

MS11_del_2_F  ------------------------------------------------------------


           1080      1090      1100      1110      1120      1130
WT_2_Final    TTTTAACCGGTCAAACCGAATAAGGAGCCGAAAATGAATCCAGCCCGCAAAAAACCTTCT

MS11_del_2_F  ------------------------------------------------------------


           1140      1150      1160      1170      1180      1190
WT_2_Final    CTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCGGCAAGTGAAGGCA

MS11_del_2_F  ------------------------------------------------------------


           1200      1210      1220      1230      1240      1250
WT_2_Final    ATGGCCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCGCCGAACGCATTACCCACG

MS11_del_2_F  ------------------------------------------------------------


           1260      1270      1280      1290      1300      1310
WT_2_Final    ATTATCCGGAACCAACCGGTGCAAAAAAAGGCACAACAATAAGCACGGTAAGCGATTATT

MS11_del_2_F  ------------------------------------------------------------


           1320      1330      1340      1350      1360      1370
WT_2_Final    TCAGAAACATCCGTACGCATTCCATCCACCCCCGGGTGTCGGTCGGCTACGACTTCGGCG

MS11_del_2_F  ------------------------------------------------------------


           1380      1390      1400      1410      1420      1430
WT_2_Final    GCTGGAGGATAGCGGCAGATTATGCCCGTTACAGAAAGTGGAACAACAATAAATATTCCG

MS11_del_2_F  ------------------------------------------------------------
```

```
           1440      1450      1460      1470      1480      1490
WT_2_Final    TGAGCATAAAAGAGTTGCTAAGAAACAAGGGCAATGGCAACAGGACAGACCTGAAGGCGG

MS11_del_2_F  ------------------------------------------------------------


           1500      1510      1520      1530      1540      1550
WT_2_Final    AAAATCAGGAAAACGGTACGTTCCACGCCGTTTCTTCTCTCGGCTTGTCCGCCGTTTACG

MS11_del_2_F  ------------------------------------------------------------


           1560      1570      1580      1590      1600      1610
WT_2_Final    ACTTCAAACTCAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACG

MS11_del_2_F  ------------------------------------------------------------


           1620      1630      1640      1650      1660      1670
WT_2_Final    TCAGACACAGCATCGATTCGACCAAAAAAAACAACAGAGGTTCAACCCATCAACCATGGTC

MS11_del_2_F  ------------------------------------------------------------


           1680      1690      1700      1710      1720      1730
WT_2_Final    CCGGCACGACCCCTACGCAATCTCGAAGGCAATCTTTCCAAGGCACCCGGGAATAAAGAT

MS11_del_2_F  ------------------------------------------------------------


           1740      1750      1760      1770      1780      1790
WT_2_Final    ACAGCATCCCCGGAATGGGCCTGGCAACAGGGGCTGGTGTGGGCATGGGCCTCGGACTCA

MS11_del_2_F  ------------------------------------------------------------


           1800      1810      1820      1830      1840      1850
WT_2_Final    CCCTTTCGTGGGAAGCCGGGTTCGGCTGCCAAGGGTGGGGACGCCGAAAAAACACCCGCT

MS11_del_2_F  ------------------------------------------------------------


           1860      1870      1880      1890      1900      1910
WT_2_Final    TCAAAACCCGAGAAACCTCGTAGAAAGTGCGGACCGATGTGAAATTCGGAGAGAAGGACA

MS11_del_2_F  ------------------------------------------------------------


           1920      1930      1940      1950      1960      1970
WT_2_Final    CCGTTTCATCGATCCCATCGCATACGATGGTCCACACCGTCTGACGAAATCGCGGCGCCA

MS11_del_2_F  ------------------------------------------------------------


           1980      1990      2000      2010      2020      2030
```

```
WT_2_Final    TTACCGTAGCCGAGCATCCTTTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCG
                                :::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  ---------------------TCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCG
                                   570       580       590       600

           2040       2050      2060       2070      2080      2090
WT_2_Final    CAG-CGGGAATCTAGATCTGTCAGTGCAG-AACTTATCGGGCAAAACGGTTTTCTGAGAT
              ::: ::::::::::::::::::::::::: :: ::::::::::::::::::  ::::::
MS11_del_2_F  CAGGCGGGAATCTAGATCTGTCAGTGCAGGAAATTATCGGGCAAAACGGTTTCTTGAGAT
                     610       620       630       640       650       660

              2100      2110      2120      2130      2140      2150
WT_2_Final    TT-GAGTCCTGGATTCCCACTTCCGCGGGAATGACGGTTCGGTTGCATAGGGTCAGATTG
              :: :::::::::::::::::: :::::::::::::::::::::::::::::::::::::
MS11_del_2_F  TTTGAGTCCTGGATTCCCACTTTCGCGGGAATGACGGTTCGGTTGCATAGGGTCAGATTG
                     670       680       690       700       710       720

              2160      2170      2180      2190      2200      2210
WT_2_Final    TCGAAAGGGGCGGATTCGATGGATTCGATGAAAACGGTAGAAATGTTGGATTGATGGGAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  TCGAAAGGGGCGGATTCGATGGATTCGATGAAAACGGTAGAAATGTTGGATTGATGGGAA
                     730       740       750       760       770       780

              2220      2230      2240      2250      2260      2270
WT_2_Final    TGGCGGACTGAAGCCCACCGATTCATCGACTCCAACGTTTACGATGCTTCCAACGGTTTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  TGGCGGACTGAAGCCCACCGATTCATCGACTCCAACGTTTACGATGCTTCCAACGGTTTC
                     790       800       810       820       830       840

              2280      2290      2300      2310      2320      2330
WT_2_Final    AGACGGCATTTTTTACACAATTCCCGCCATTTTCCATCATTCCCGACAATACCGTAATCT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  AGACGGCATTTTTTACACAATTCCCGCCATTTTCCATCATTCCCGACAATACCGTAATCT
                     850       860       870       880       890       900

              2340      2350      2360      2370      2380      2390
WT_2_Final    CGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  CGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGG
                     910       920       930       940       950       960

              2400      2410      2420      2430      2440      2450
WT_2_Final    ATAAAAACAGTTGCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  ATAAAAACAGTTGCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTC
                     970       980       990      1000      1010      1020

              2460      2470      2480      2490      2500      2510
WT_2_Final    GGTTGCGTTCCGACAACACCGTAGTCTCGAAACCCGTCCGACAACACCGTAGTCTCGAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  GGTTGCGTTCCGACAACACCGTAGTCTCGAAACCCGTCCGACAACACCGTAGTCTCGAAA
                    1030      1040      1050      1060      1070      1080

              2520      2530      2540      2550      2560      2570
WT_2_Final    CCCGTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  CCCGTCCGACAACACCGCAATCTCGAAACCCGTCATTCCCGCGCAGGCGGGAATCCGGAC
```

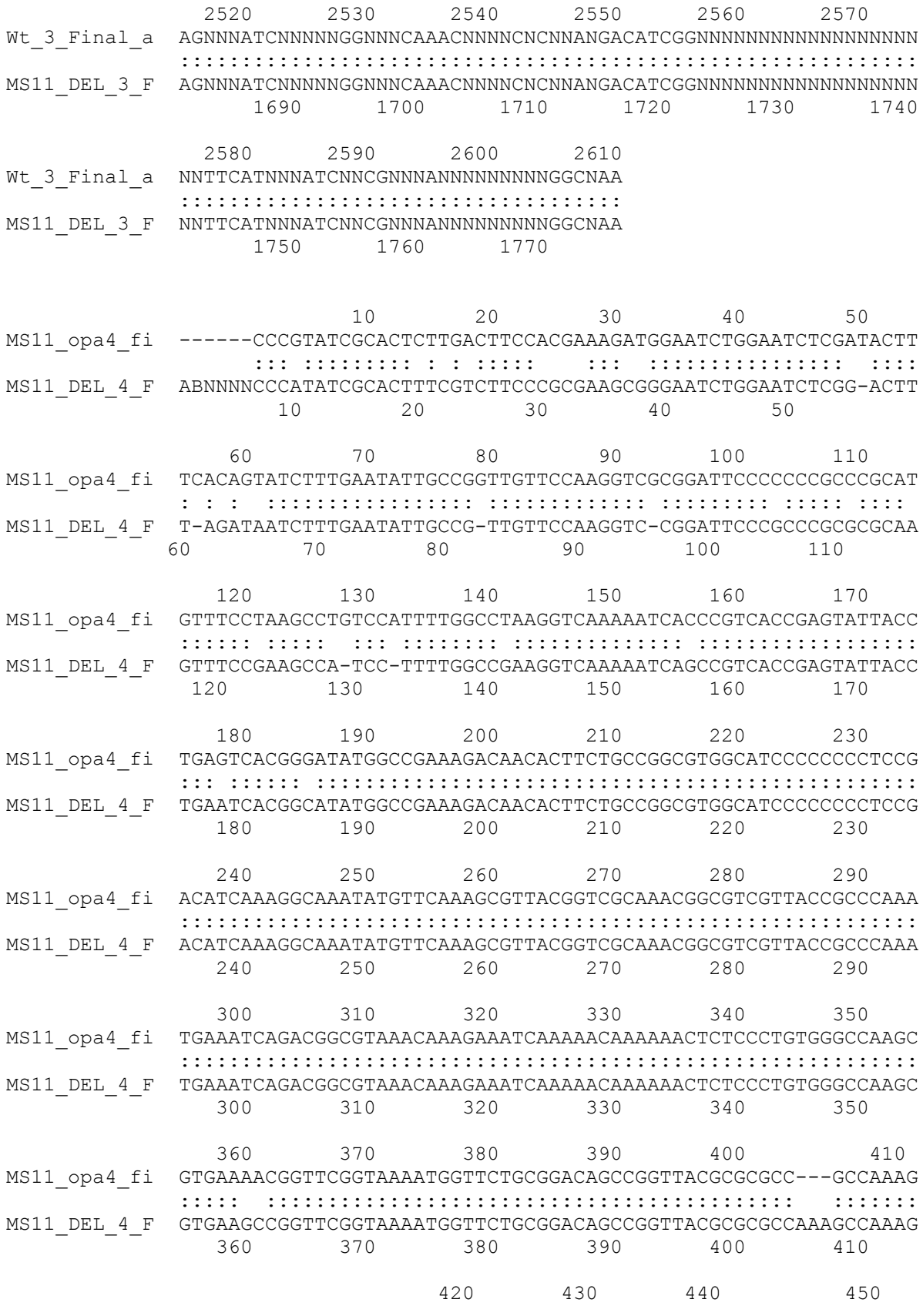
```
                  1090      1100      1110      1120      1130      1140

              2580      2590      2600      2610      2620      2630
WT_2_Final    CTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAGATTCCC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  CTGTCCGCACGGAAACTTATCGGATAAAACGGTTGCCCAAACCCCGCGTCCTAGATTCCC
                  1150      1160      1170      1180      1190      1200

              2640      2650      2660      2670      2680      2690
WT_2_Final    ACTTCCGTGGGAATGACGGTTCGGTTGCGTTCCGACAACACCGCAATCTCGAAACCCGTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  ACTTCCGTGGGAATGACGGTTCGGTTGCGTTCCGACAACACCGCAATCTCGAAACCCGTC
                  1210      1220      1230      1240      1250      1260

              2700      2710      2720      2730      2740      2750
WT_2_Final    ATTCCCGCGCAGGCGGGAATCCGGACCCCCGACGCGGCGGGAATCTATCGGAAATGACTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_del_2_F  ATTCCCGCGCAGGCGGGAATCCGGACCCCCGACGCGGCGGGAATCTATCGGAAATGACTG
                  1270      1280      1290      1300      1310      1320

              2760      2770      2780      2790      2800      2810
WT_2_Final    AAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGTGTTCCGACAA
              :::::::::::::::::::::::::::::::::::::::::::::: ::::::::::::
MS11_del_2_F  AAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTT-TGTTCCGACAA
                  1330      1340      1350      1360       1370

              2820      2830      2840      2850
WT_2_Final    CACCGTAGTCTCAAACCCTCATCCCCGCGCGATATCCCNNNT
              ::::: ::::::::::::  :::  ::  : :: :: ::
MS11_del_2_F  CACCG-AGTCTCAAACCAGCATAGCCCAGTGACATGCC----
                 1380       1390      1400      1410


                     10        20        30        40              50
Wt_3_Final_a  NNCCCGATTAGCTGCTGGGGAGGTAAATAAAAANNGNNNT------CAAATTTGTCCGGC
              ::   :           :  : :  : :    :  :   :      ::::: ::::: ::
MS11_DEL_3_F  NNNNNGCNNNNNCNTNNGCTACGGCA-TCCGCACCGGCCTGNAATCCAAATATGTCCTGC
                     10        20         30        40        50

                  60        70        80        90       100       110
Wt_3_Final_a  ACAACTTGGGCTTCCGCTGCGCAAGCCAATAACCCCTTCAATTATAGGGGATTAACAAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  ACAACTTGGGCTTCCGCTGCGCAAGCCAATAACCCCTTCAATTATAGGGGATTAACAAAA
               60        70        80        90       100       110

                  120       130       140       150       160       170
Wt_3_Final_a  ACCGGTCCGGCGTTGCCTCGCCTTGCCGTACTGGTTTTTGTTAATCCGCTATATTCCCCC
              ::: :: ::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  ACCAGTACGGCGTTGCCTCGCCTTGCCGTACTGGTTTTTGTTAATCCGCTATATTCCCCC
              120       130       140       150       160       170

                  180       190       200       210       220       230
Wt_3_Final_a  ATCTTTAAAATTTACAGCGATACACGGGTAATTTAAGGAATGCCCAAACCGTCATTCCCG
               ::: :::::::::::::::::::::::::::::::::::::: ::::::: ::::::
MS11_DEL_3_F  TTCTCTAAAATTTACAGCGATACACGGGTAATTTAAGGAATGCCCGAACCGTCTTTCCCG
              180       190       200       210       220       230
```

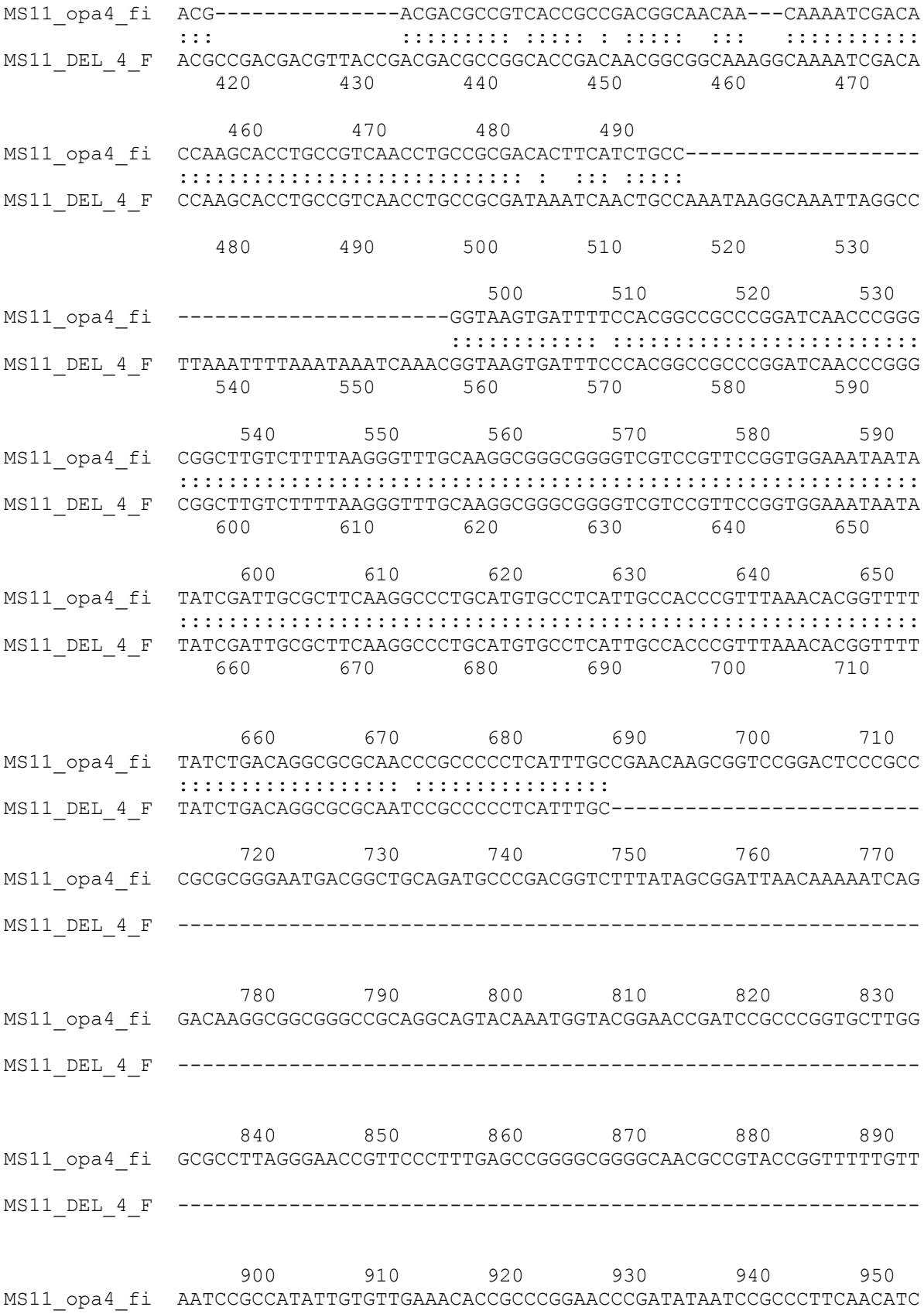
```
                240       250       260       270       280       290
Wt_3_Final_a  CAACTTTTCTTCATTCCCGCGAAAGCGGGAATCTAAAATCTCGGACTTTCAAATAATCTT
              ::::::::: :: ::::::::::::::::::::: ::::::::::::::::::::::
MS11_DEL_3_F  CAACTTTTCGTCTTTCCCGCGAAAGCGGGAATCTAGAATCTCGGACTTTCAAATAATCTT
           240       250       260       270       280       290

                300       310       320       330       340       350
Wt_3_Final_a  TGAATATTGCTGTTGTTCTAAGGACCGGATTCCCGCCTGCGCGCAAGTTTCCAAAGCCTT
              :::::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  TGAATATTGCTGTTGTTCTAAGGTCCGGATTCCCGCCTGCGCGCAAGTTTCCAAAGCCTT
           300       310       320       330       340       350

                360       370       380       390       400       410
Wt_3_Final_a  CCTTTTGGCCAAAGGTCAAAAATCACCCGTCACCGAGTATTACCTGAACCACGGCAAATG
              :::::::::: :::::::::::::::::::::::::::::::::: :::::: ::::
MS11_DEL_3_F  CCTTTTGGCCGAAGGTCAAAAATCACCCGTCACCGAGTATTACCTGAATCACGGCGAATG
           360       370       380       390       400       410

                420       430       440       450       460       470
Wt_3_Final_a  GCCCAAAGACAACGACTCGGCGGGCGTGGCTTCCGCTTCAAAAATCATAGGCAAATATGT
              :::::::::::::::::: :: :::: ::: ::::::::::::::::::::::::::::
MS11_DEL_3_F  GCCCAAAGACAACGACTCTGCCGGCGGGGCATCCGCTTCAAAAATCATAGGCAAATATGT
           420       430       440       450       460       470

                480       490       500       510       520       530
Wt_3_Final_a  TAAGCAAGTTGAAGTCAAAAACGGCGTCTTTCCCGCCCAAATGAAATCAGACGGCGTAAA
              :::::::::::::::::::::::::: ::::::::::::::::::::::::::::::::
MS11_DEL_3_F  TAAGCAAGTTGAAGTCAAAAACGGCGTCGTTCCCGCCCAAATGAAATCAGACGGCGTAAA
           480       490       500       510       520       530

                540       550       560       570       580       590
Wt_3_Final_a  CAAAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGCGTGAAAACGGTTCGGTAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CAAAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGCGTGAAAACGGTTCGGTAAA
           540       550       560       570       580       590

                600       610       620       630       640       650
Wt_3_Final_a  ATGGTTCTGCGGACAGCCGGTTACGCGCAACGCCAAAGCCAACGACACCGTTGCCGCCGA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  ATGGTTCTGCGGACAGCCGGTTACGCGCAACGCCAAAGCCAACGACACCGTTGCCGCCGA
           600       610       620       630       640       650

                660       670       680       690       700       710
Wt_3_Final_a  CGGCACCGGCAACGACAAAATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGATAACTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CGGCACCGGCAACGACAAAATCGAAACCAAGCACCTGCCGTCAACCTGCCGCGATAACTT
           660       670       680       690       700       710

                720       730       740       750       760       770
Wt_3_Final_a  TGATGCCAGCTGAGGCAAATTAGGCCTTAAATTTCAAATAAATCAAGCGGTAAGTGATTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  TGATGCCAGCTGAGGCAAATTAGGCCTTAAATTTCAAATAAATCAAGCGGTAAGTGATTT
           720       730       740       750       760       770

                780       790       800       810       820       830
Wt_3_Final_a  TCCACGGCCGCCCGGATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGG
              :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
```

```
MS11_DEL_3_F  TCCACGGCCGCCCGNATCAACCCGGGCGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGG
            780       790       800       810       820       830

                840       850       860       870       880       890
Wt_3_Final_a  GGTCGTCCGTTCCGGTGGAAATAATATATCGATTGCGCTTCAAGGCCCTGCATGTGCCTC
              ::::::::::::::::: :::::::::::::::::::::::: :::::::::::::::::
MS11_DEL_3_F  GGTCGTCCGTTCCGGTGAAAATAATATATCGATTGCGCTTCAGGGCCCTGCATGTGCCTC
            840       850       860       870       880       890

                900       910       920       930       940       950
Wt_3_Final_a  ATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAACCCGCCCCCTCATTTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  ATTGCCACCCGTTTAAACACGGTTTTTATCTGACAGGCGCGCAACCCGCCCCCTCATTTG
            900       910       920       930       940       950

                960       970       980       990      1000      1010
Wt_3_Final_a  CCGAACAAGCGGTCCGGACTCCCGCCCGCGCGGGAATGACGGCTGCAGATGCCCGACGGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CCGAACAAGCGGTCCGGACTCCCGCCCGCGCGGGAATGACGGCTGCAGATGCCCGACGGT
            960       970       980       990      1000      1010

               1020      1030      1040      1050      1060      1070
Wt_3_Final_a  CTTTATAGCGGATTAACAAAAATCAGGACAAGGCGGCGGGCCGCAGGCAGTACAAATGGT
              ::::::::::::::::::::::: :::::::::::::::::: :::::::  :::::
MS11_DEL_3_F  CTTTATAGCGGATTAACAAAAATCANGACAAGGCGGCGGGCCGCANGCAGTACGGATGGT
            1020      1030      1040      1050      1060      1070

               1080      1090      1100      1110      1120      1130
Wt_3_Final_a  ACGGAACCGATCCGCCCGGTGCTTGGGCGCCTTAGGGAACCGTTCCCTTTGAGCCGGGGC
              ::::::::: : ::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  ACGGAACCGGTTCGCCCGGTGCTTGGGCGCCTTAGGGAACCGTTCCCTTTGAGCCGGGGC
            1080      1090      1100      1110      1120      1130

               1140      1150      1160      1170      1180      1190
Wt_3_Final_a  GGGGCAACGACGTACCGGTTTTTGTTCATCCGCCATATTGTGTTGAAACACCGCCCGGAA
              ::::::::: ::::::::::::::: ::::::::::::::::::::::::::::::::
MS11_DEL_3_F  GGGGCAACGCCGTACCGGTTTTTGTTAATCCGCCATATTGTGTTGAAACACCGCCCGGAA
            1140      1150      1160      1170      1180      1190

               1200      1210      1220      1230      1240      1250
Wt_3_Final_a  CCCGATATAATCCGCCCTTCAACATCAGTGAAAATCTTTTTTTAACCGGTCAAACCGAAT
              :::
MS11_DEL_3_F  CCC-------------------------------------------------------------
            1200

               1260      1270      1280      1290      1300      1310
Wt_3_Final_a  AAGGAGCCGAAAATGAATCCAGCCCCCAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTC

MS11_DEL_3_F  -------------------------------------------------------------


               1320      1330      1340      1350      1360      1370
Wt_3_Final_a  TTCTCTTCTCTCCGCAGCGCAGGCGGGGGGAGAAAACTATGGCCGCGGTGCCTACGTGCG

MS11_DEL_3_F  -------------------------------------------------------------
```

```
                1380      1390      1400      1410      1420      1430
Wt_3_Final_a  CGCTGATGTCGGATACGCCTACTAAAACCTTACCGGCAATTATGTCCAACACACCCCTCC

MS11_DEL_3_F  ------------------------------------------------------------


                1440      1450      1460      1470      1480      1490
Wt_3_Final_a  AAAAAAAGCACAATTAAGCACGGTAAGCGATTATTTCAGAAACATCCGTACGCATTCCAT

MS11_DEL_3_F  ------------------------------------------------------------


                1500      1510      1520      1530      1540      1550
Wt_3_Final_a  CCACCCCAGGGTGTCGGTCGGCTACGACTTCGGCGGCTGGAGGATAGCGGCAGATTATGC

MS11_DEL_3_F  ------------------------------------------------------------


                1560      1570      1580      1590      1600      1610
Wt_3_Final_a  CCGTTACAGAAAGTGGAACAACAATAAATATTCCGTTAACATAAAAGAGTTGCTAAGAAA

MS11_DEL_3_F  ------------------------------------------------------------


                1620      1630      1640      1650      1660      1670
Wt_3_Final_a  CGATAATGCCAATTCTGGCGGCAGCCATCTTAACATAAAAACCCGAAAGACGGAACATCG

MS11_DEL_3_F  ------------------------------------------------------------


                1680      1690      1700      1710      1720      1730
Wt_3_Final_a  GGAAAACGGCACATTCCACGCCGCCTCTTCTCTCGGCTTGTCCGCCGTTTACGATTTCGA

MS11_DEL_3_F  ------------------------------------------------------------


                1740      1750      1760      1770      1780      1790
Wt_3_Final_a  TACCGGTTCCCGCTTCAAACCCTATATCGGCATGCGCGTCGCCTACGGACACGTCAGACA

MS11_DEL_3_F  ------------------------------------------------------------


                1800      1810      1820      1830      1840      1850
Wt_3_Final_a  TCAGGTTCGTTCGGTTCAACAAGAAACCATTGCTGTTACCACTTACCCACAGAATGCTGC

MS11_DEL_3_F  ------------------------------------------------------------


                1860      1870      1880      1890      1900      1910
Wt_3_Final_a  GTCAAGTGTTACCACAAATGCTCCGATCCGCAAACTTCCCCATCACGAAAGCCGCAGCAT

MS11_DEL_3_F  ------------------------------------------------------------


                1920      1930      1940      1950      1960      1970
Wt_3_Final_a  CAGCAGCTTGGGCTTCGGCGCAGTGGCAGGCGTAGGCATCGACATCACGCCCAACCTGAC
```

```
MS11_DEL_3_F  ------------------------------------------------------------


                1980      1990      2000      2010      2020      2030
Wt_3_Final_a  CCTGGACGCCGGCTACCGCTACCACAACTGGGGACGCTTGGAAAACACCCGCTTCAAAAC

MS11_DEL_3_F  ------------------------------------------------------------


                2040      2050      2060      2070      2080      2090
Wt_3_Final_a  CCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCGCCGATTCCGATGCCGTCTG
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  --ACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCGCCGATTCCGATGCCGTCTG
                     1210      1220      1230      1240      1250      1260

                2100      2110      2120      2130      2140      2150
Wt_3_Final_a  AACCTTCAGACGGCATGAGGCCTTTGCCTGCGCACTTGGTGCGCTGGTCGCCTCCGAACA
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  AACCTTCAGACGGCATGAGGCCTTTGCCTGCGCACTTGGTGCGCTGGTCGCCTCCGAACA
                     1270      1280      1290      1300      1310      1320

                2160      2170      2180      2190      2200      2210
Wt_3_Final_a  TGGCGCAACACCCGACATTTCCGCCGAACGCATCGGGCGTTTCATAAACCCGGTTTAAAA
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  TGGCGCAACACCCGACATTTCCGCCGAACGCATCGGGCGTTTCATAAACCCGGTTTAAAA
                     1330      1340      1350      1360      1370      1380

                2220      2230      2240      2250      2260      2270
Wt_3_Final_a  CGCATGGAAAAATGCCGTCTGAAAGCCTTTCAGACGGCATTTTGTTTGAGATTCCGTTTA
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CGCATGGAAAAATGCCGTCTGAAAGCCTTTCAGACGGCATTTTGTTTGAGATTCCGTTTA
                     1390      1400      1410      1420      1430      1440


                2280      2290      2300      2310      2320      2330
Wt_3_Final_a  CCAATGGCTGACAAACGCTTCCAAATCGGTATTCTTGGGCTTATGCACTTCCTCTGTCGG
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CCAATGGCTGACAAACGCTTCCAAATCGGTATTCTTGGGCTTATGCACTTCCTCTGTCGG
                     1450      1460      1470      1480      1490      1500

                2340      2350      2360      2370      2380      2390
Wt_3_Final_a  CGTGCCGACCATCATCAGCCCGATGATTTTATCCTTATCCTCGCAACCGNNGNCTCCCGC
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  CGTGCCGACCATCATCAGCCCGATGATTTTATCCTTATCCTCGCAACCGNNGNCTCCCGC
                     1510      1520      1530      1540      1550      1560

                2400      2410      2420      2430      2440      2450
Wt_3_Final_a  AACAATGGGCTGTTGACCCACATCCCCGTAATCCAGACATTGTCGAATCCCNGAGCCNNN
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  AACAATGGGCTGTTGACCCACATCCCCGTAATCCAGACATTGTCGAATCCCNGAGCCNNN
                     1570      1580      1590      1600      1610      1620

                2460      2470      2480      2490      2500      2510
Wt_3_Final_a  NNNNNNNCAGCGCATACGCCNCANNNCNNNNNCAGCATCTGCTNCCNNNNGNTTNNNNTN
                :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_3_F  NNNNNNNCAGCGCATACGCCNCANNNCNNNNNCAGCATCTGCTNCCNNNNGNTTNNNNTN
                     1630      1640      1650      1660      1670      1680
```


2520 2530 2540 2550 2560 2570
Wt_3_Final_a AGNNNATCNNNNNGGNNNCAAACNNNNCNCNNANGACATCGGNNNNNNNNNNNNNNNNNN
MS11_DEL_3_F AGNNNATCNNNNNGGNNNCAAACNNNNCNCNNANGACATCGGNNNNNNNNNNNNNNNNNN
1690 1700 1710 1720 1730 1740
2580 2590 2600 2610
Wt_3_Final_a NNTTCATNNNATCNNCGNNNANNNNNNNNNGGCNAA
MS11_DEL_3_F NNTTCATNNNATCNNCGNNNANNNNNNNNNGGCNAA
1750 1760 1770
10 20 30 40 50
MS11_opa4_fi ------CCCGTATCGCACTCTTGACTTCCACGAAAGATGGAATCTGGAATCTCGATACTT
MS11_DEL_4_F ABNNNNCCCATATCGCACTTTCGTCTTCCCGCGAAGCGGGAATCTGGAATCTCGG-ACTT
10 20 30 40 50
60 70 80 90 100 110
MS11_opa4_fi TCACAGTATCTTTGAATATTGCCGGTTGTTCCAAGGTCGCGGATTCCCCCCCGCCCGCAT
MS11_DEL_4_F T-AGATAATCTTTGAATATTGCCG-TTGTTCCAAGGTC-CGGATTCCCGCCCGCGCGCAA
60 70 80 90 100 110
120 130 140 150 160 170
MS11_opa4_fi GTTTCCTAAGCCTGTCCATTTTGGCCTAAGGTCAAAAATCACCCGTCACCGAGTATTACC
MS11_DEL_4_F GTTTCCGAAGCCA-TCC-TTTTGGCCGAAGGTCAAAAATCAGCCGTCACCGAGTATTACC
120 130 140 150 160 170
180 190 200 210 220 230
MS11_opa4_fi TGAGTCACGGGATATGGCCGAAAGACAACACTTCTGCCGGCGTGGCATCCCCCCCCTCCG
MS11_DEL_4_F TGAATCACGGCATATGGCCGAAAGACAACACTTCTGCCGGCGTGGCATCCCCCCCCTCCG
180 190 200 210 220 230
240 250 260 270 280 290
MS11_opa4_fi ACATCAAAGGCAAATATGTTCAAAGCGTTACGGTCGCAAACGGCGTCGTTACCGCCCAAA
MS11_DEL_4_F ACATCAAAGGCAAATATGTTCAAAGCGTTACGGTCGCAAACGGCGTCGTTACCGCCCAAA
240 250 260 270 280 290
300 310 320 330 340 350
MS11_opa4_fi TGAAATCAGACGGCGTAAACAAAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGC
MS11_DEL_4_F TGAAATCAGACGGCGTAAACAAAGAAATCAAAAACAAAAAACTCTCCCTGTGGGCCAAGC
300 310 320 330 340 350
360 370 380 390 400 410
MS11_opa4_fi GTGAAAACGGTTCGGTAAAATGGTTCTGCGGACAGCCGGTTACGCGCGCC---GCCAAAG
MS11_DEL_4_F GTGAAGCCGGTTCGGTAAAATGGTTCTGCGGACAGCCGGTTACGCGCGCCAAAGCCAAAG
360 370 380 390 400 410
420 430 440 450


MS11_opa4_fi ACG---------------ACGACGCCGTCACCGCCGACGGCAACAA---CAAAATCGACA
MS11_DEL_4_F ACGCCGACGACGTTACCGACGACGCCGGCACCGACAACGGCGGCAAAGGCAAAATCGACA
420 430 440 450 460 470
460 470 480 490
MS11_opa4_fi CCAAGCACCTGCCGTCAACCTGCCGCGACACTTCATCTGCC------------------
MS11_DEL_4_F CCAAGCACCTGCCGTCAACCTGCCGCGATAAATCAACTGCCAAATAAGGCAAATTAGGCC
480 490 500 510 520 530
500 510 520 530
MS11_opa4_fi ----------------------GGTAAGTGATTTTCCACGGCCGCCCGGATCAACCCGGG
MS11_DEL_4_F TTAAATTTTAAATAAATCAAACGGTAAGTGATTTCCCACGGCCGCCCGGATCAACCCGGG
540 550 560 570 580 590
540 550 560 570 580 590
MS11_opa4_fi CGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGGGGTCGTCCGTTCCGGTGGAAATAATA
MS11_DEL_4_F CGGCTTGTCTTTTAAGGGTTTGCAAGGCGGGCGGGGTCGTCCGTTCCGGTGGAAATAATA
600 610 620 630 640 650
600 610 620 630 640 650
MS11_opa4_fi TATCGATTGCGCTTCAAGGCCCTGCATGTGCCTCATTGCCACCCGTTTAAACACGGTTTT
MS11_DEL_4_F TATCGATTGCGCTTCAAGGCCCTGCATGTGCCTCATTGCCACCCGTTTAAACACGGTTTT
660 670 680 690 700 710
660 670 680 690 700 710
MS11_opa4_fi TATCTGACAGGCGCGCAACCCGCCCCCTCATTTGCCGAACAAGCGGTCCGGACTCCCGCC
MS11_DEL_4_F TATCTGACAGGCGCGCAATCCGCCCCCTCATTTGC------------------------
720 730 740 750 760 770
MS11_opa4_fi CGCGCGGGAATGACGGCTGCAGATGCCCGACGGTCTTTATAGCGGATTAACAAAAATCAG
MS11_DEL_4_F --
780 790 800 810 820 830
MS11_opa4_fi GACAAGGCGGCGGGCCGCAGGCAGTACAAATGGTACGGAACCGATCCGCCCGGTGCTTGG
MS11_DEL_4_F --
840 850 860 870 880 890
MS11_opa4_fi GCGCCTTAGGGAACCGTTCCCTTTGAGCCGGGGCGGGGCAACGCCGTACCGGTTTTTGTT
MS11_DEL_4_F --
900 910 920 930 940 950
MS11_opa4_fi AATCCGCCATATTGTGTTGAAACACCGCCCGGAACCCGATATAATCCGCCCTTCAACATC

```
MS11_DEL_4_F  ------------------------------------------------------------

                   960       970       980       990      1000      1010
MS11_opa4_fi  AGTGAAAATCTTTTTTTAACCGGTCAAACCGAATAAGGAGCCGAAAATGAATCCAGCCCC

MS11_DEL_4_F  ------------------------------------------------------------

                  1020      1030      1040      1050      1060      1070
MS11_opa4_fi  CAAAAAACCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCG

MS11_DEL_4_F  ------------------------------------------------------------

                  1080      1090      1100      1110      1120      1130
MS11_opa4_fi  GCGGGTGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAA

MS11_DEL_4_F  ------------------------------------------------------------

                  1140      1150      1160      1170      1180      1190
MS11_opa4_fi  CACATTACCCACGATTATCCGGAACCAACCGGTACAAAAAAAGACAAAATAAGCACGGTA

MS11_DEL_4_F  ------------------------------------------------------------

                  1200      1210      1220      1230      1240      1250
MS11_opa4_fi  AGCGATTATTTCAGAAACATCCGTACGCATTCCATCCACTCCAGGGTGTCGGTTGGATAC

MS11_DEL_4_F  ------------------------------------------------------------
               720       730       740

                  1260      1270      1280      1290      1300      1310
MS11_opa4_fi  GACTTTGGCTGCTGGAGGATAGCGGCCAGATTATCCCCGTTACAGGAAAATGAAACGACA

MS11_DEL_4_F  ------------------------------------------------------------

                  1320      1330      1340      1350      1360      1370
MS11_opa4_fi  ATAAATATTCAGTCGGCAATAAAAGAGTTAGAAAACAAGAATCAGAAGTAAGAGAGACCT

MS11_DEL_4_F  ------------------------------------------------------------

                  1380      1390      1400      1410      1420      1430
MS11_opa4_fi  GAAGACGGAAAATCAGGAAAACGGTACGTTCCACGCCGTCTCCTCGCTCGGTTTGTCAGC

MS11_DEL_4_F  ------------------------------------------------------------

                  1440      1450      1460      1470      1480      1490
MS11_opa4_fi  CGTTTGCGATTTCAAACTCAACGATAAATTCAAACCCTATATCGGTGCGCGCGTCGCNTA

MS11_DEL_4_F  ------------------------------------------------------------
```

```
                1500      1510      1520      1530      1540      1550
MS11_opa4_fi  CGGACACGTCAGACACAGCATCGATTCGACCAAAAAAACAACAAAGTTTCTTACCTCCTC

MS11_DEL_4_F  ------------------------------------------------------------


                1560      1570      1580      1590      1600      1610
MS11_opa4_fi  CTATGGTGGCTTAAACCCTACGGTTTATACTGAGGAAAATACGCAAAACGCCCATCACCA

MS11_DEL_4_F  ------------------------------------------------------------


                1620      1630      1640      1650      1660      1670
MS11_opa4_fi  AAGTAACAGCATCCGCCGCGTGGGCCTCGGCGTCATCGCCGGCGTCGGTTTCGACATCAC

MS11_DEL_4_F  ------------------------------------------------------------


                1680      1690      1700      1710      1720      1730
MS11_opa4_fi  GCCCAAGCTGACCCTGGACACCGGCTACCGCTACCACTATTGGGGACGCTTGGAAAACAC

MS11_DEL_4_F  ------------------------------------------------------------


         1740      1750      1760      1770      1780      1790
MS11_opa4_fi  CCGCTTCAAAACCCACGAAGCCTCGTTGGGCGTGCGCTACCGCTTCTGATTCCCCGATAC
                                  :: :::::: ::::::::::::::::::::::::::::
MS11_DEL_4_F  ----------------------TCATTGGGCATGCGCTACCGCTTCTGATTCCCCGATAC
                               750       760       770       780
                  1800      1810      1820      1830      1840      1850
MS11_opa4_fi  CGATGCCGTCTGAACCTTCAGACGGCATTTTTGATGCGCCCGCTGTTTACAGGCGCGGGG
              :::::::::::::::::::::::::::::::: :: ::::::: :::::::::::::::
MS11_DEL_4_F  CGATGCCGTCTGAACCTTCAGACGGCATTTTTAAT-CGCCCGCCGTTTACAGGCGCGGGG
              790       800       810       820       830       840

                1860        1870      1880       1890      1900
MS11_opa4_fi  CGGGCGTGG--AAATACCCGAACCGTCATTGCCGATA-CACCGTAACCCTAAAACCCGCC
              ::::::  :  :::::::::::::::::: :::: : ::::: :: :   :::  :: :
MS11_DEL_4_F  CGGGCGCAGTAAAATACCCGAACCGTCATTCCCGACAACACCGCAATCTCGAAATTCGTC
               850       860       870       880       890       900

           1910      1920      1930      1940      1950      1960
MS11_opa4_fi  ATTCCCGCGCAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_4_F  ATTCCCGCGCAGGCGGGAATCCGGACCTGTCCGCACGGAAACTTATCGGATAAAACGGTT
               910       920       930       940       950       960

           1970      1980      1990      2000      2010      2020
MS11_opa4_fi  GCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCTACGGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_4_F  GCCCAAACCCCGCGTCCTAGATTCCCACTTCCGTGGGAATGACGGTTCGGTTGCTACGGC
               970       980       990      1000      1010      1020

           2030      2040      2050      2060       2070
MS11_opa4_fi  CCGCTGATTCCCCGACACCGATGCCGTCTGAACCT-CAGACGCATTNC-----
              :::::::::::::::::::::::::::::::::: ::::::   :::
```

```
MS11_DEL_4_F  CCGCTGATTCCCCGACACCGATGCCGTCTGAACCTTCAGACGGCNTNCNNNAA
              1030      1040      1050      1060      1070

                      10        20        30        40        50        60
WT_5_Final    ABNNNNGGCATCGCACGGAACTGCCGCCATGACGGCGCACTCATCCGCGTCCTCCCTTCG
                :::  : :  :::  : : ::::::::::::::::::::::::::::::::::::::
Del_5_Final   BNNNNCCGAAGAGCAACGGATTGCCGCCATGACGGCGCACTCATCCGCGTCCTCCCTTCG
                      10        20        30        40        50        60

                      70        80        90       100       110       120
WT_5_Final    GAAAACGGACTCGACATCCTGCAAGTGTTTGAAGAAATCTCTTGCAGCGTCAATACCACA
              :::::::::::::::::::::::::::::::::::: ::::::::: :::: ::::::
Del_5_Final   GAAAACGGACTCGACATCCTGCAAGTGTTTGAAGAAATTTCTTGCAGCATCAACACCACA
                      70        80        90       100       110       120

                     130       140       150       160       170       180
WT_5_Final    GCCGCCGGCGGCAACGCCGGCCTGATGGCGGTTGCCGACTGATTTTGCCAAACCGCCCGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   GCCGCCGGCGGCAACGCCGGCCTGATGGCGGTTGCCGACTGATTTTGCCAAACCGCCCGG
                     130       140       150       160       170       180

                     190       200       210       220       230       240
WT_5_Final    GCGCGGCCCGTGAACAAATGCCGTCTGAAAACCTTTCAGGCGGCATTTTATAGTGGATTA
              :::::::::::::: ::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   GCGCGGCCCGTGAACGAATGCCGTCTGAAAACCTTTCAGGCGGCATTTTATAGTGGATTA
                     190       200       210       220       230       240

                     250       260       270       280       290       300
WT_5_Final    ACAAAAATCAGGACAAGGCGGCGGGCCGCAGACAGTACAAAATAGTACGGCAAGGCGGGC
              :::::::::::::::::::::::::::::::::::: :::::::::::::::::::
Del_5_Final   ACAAAAATCAGGACAAGGCGGCGGGCCGCAGACAGTACAAA-TAGTACGGCAAGGCGGGC
                     250       260       270       280        290

                     310       320       330       340       350       360
WT_5_Final    CAACGCTGTACTGGTTTAAATTCAATTCACTATATGTGTTGGCACATCGCTCCAAACCTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   CAACGCTGTACTGGTTTAAATTCAATTCACTATATGTGTTGGCACATCGCTCCAAACCT-
            300       310       320       330       340       350

                     370       380       390       400       410       420
WT_5_Final    ATATAATCCGCCTTTCAACATCAGTGAAAATCGTTCTTTTTAGTCAGTTAACATTAAATT

Del_5_Final   ------------------------------------------------------------


                     430       440       450       460       470       480
WT_5_Final    TCGGAGTCGAAAAATGAATCCAGCCCCAAAAAACTTCTCTTCTCTTCTCTTCTCTTCTCT

Del_5_Final   ------------------------------------------------------------


                     490       500       510       520       530       540
WT_5_Final    TCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCCGCAGCGCAGGCGGCAAGTGAAGACGGC

Del_5_Final   ------------------------------------------------------------
```

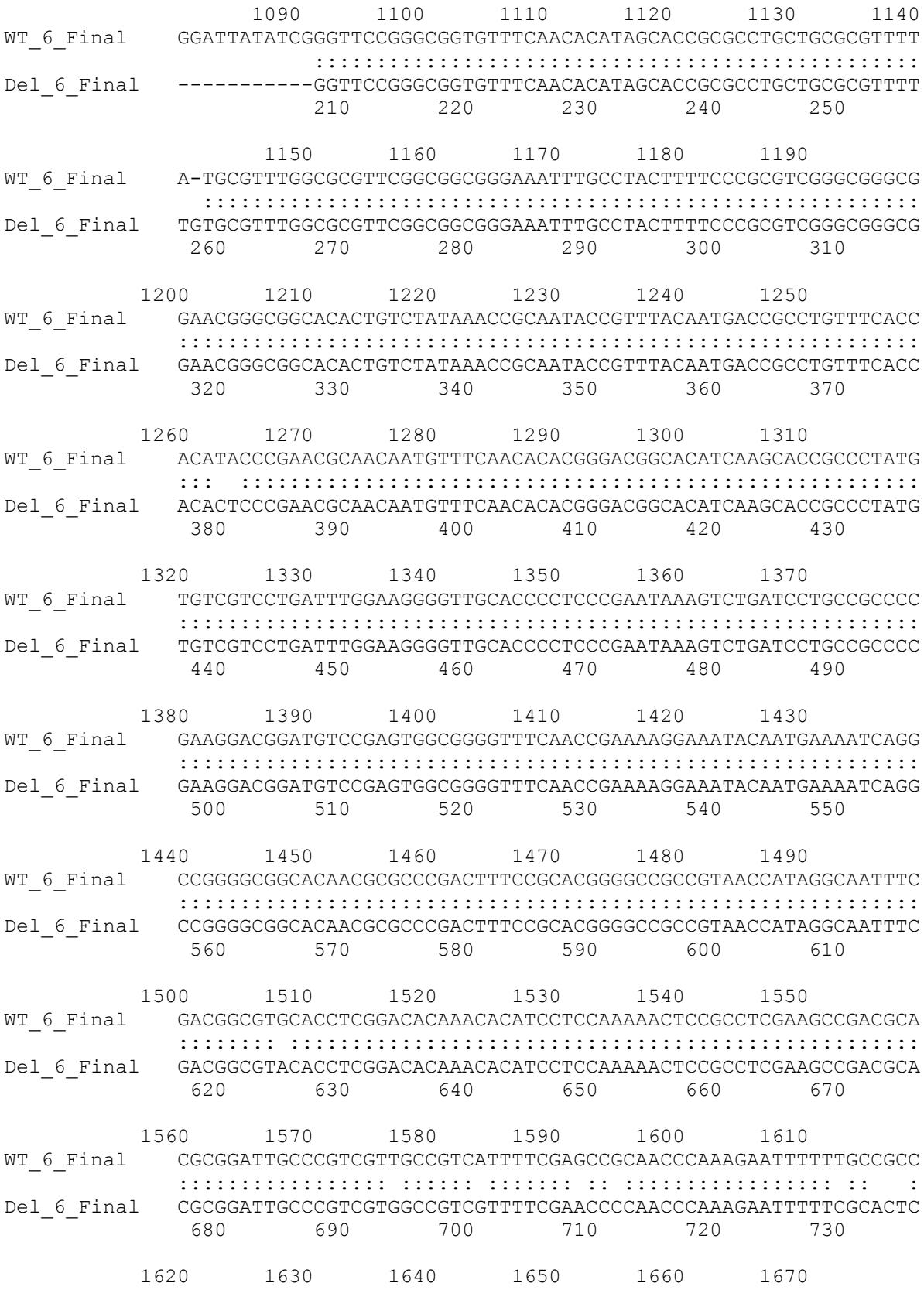
```
                  550       560       570       580       590       600
WT_5_Final    GGCCGCGGCCCGTATGTGCAGGCGGATTTAGCCTACGCCTACGAACACATTACCCACGAT

Del_5_Final   ------------------------------------------------------------

                  610       620       630       640       650       660
WT_5_Final    TATCCGAAACCAACCGATCCAAGCAAAGGCAAAATAAGCACGGTAAGCGATTATTTCAGA

Del_5_Final   ------------------------------------------------------------

                  670       680       690       700       710       720
WT_5_Final    AACATCCGTACGCATTCCATCCACCCCAGGGTGTCGGTTGGCTACGACTTCGGTGGCTGG

Del_5_Final   ------------------------------------------------------------

                  730       740       750       760       770       780
WT_5_Final    AGGATAGCGACAGATTATGCCCGTTACAGGAAATGGAGCGACAATAAATATTCCGTCAGC

Del_5_Final   ------------------------------------------------------------

                  790       800       810       820       830       840
WT_5_Final    ATAAAAAATATGCGGGTACATAAACACAATAGCAACAGGAAAAACCTGAAGACGGAAAAT
                                                          ::: ::::
Del_5_Final   ----------------------------------------------CTGCAGAC-------
                                                         360

                  850       860       870       880       890       900
WT_5_Final    CAGGAAAACGGCAGCTTCCACGCCGTTTCTTCTCTCGGCTTATCCGCTATTTACGATTTC
 Del_5_Final  ------------------------------------------------------------

                  910       920       930       940       950       960
WT_5_Final    CAAATAAACGATAAATTCAAACCCTATATCGGCGCGCGCGTCGCCTACGGACACGTCAGA

Del_5_Final   ------------------------------------------------------------

                  970       980       990      1000      1010      1020
WT_5_Final    CACAGCATCGATTCGACTAAAAAAATAACAGGGCTTCTTACCACCAGTACTCCTGGCATA

Del_5_Final   ------------------------------------------------------------

                 1030      1040      1050      1060      1070      1080
WT_5_Final    ATGTTTGGGGTTTATAAGGTATTAAGGACACCAGGCGCCCATCGCGAAAGCGACAGCATC

Del_5_Final   ------------------------------------------------------------
```

```
                1090      1100      1110      1120      1130      1140
WT_5_Final    CGCCGCGTGGGTCTCGGTGTCATCGCCGGCGTCGGTTTCGACATCACGCCCAAGCTGACC

Del_5_Final   ------------------------------------------------------------


                1150      1160      1170      1180      1190      1200
WT_5_Final    CTGGACGCCGGCTACCGCTACCACAACTGGGGACGCTTGGAAAACACCCGCTTCAAAACC

Del_5_Final   ------------------------------------------------------------


                1210      1220      1230      1240      1250      1260
WT_5_Final    CACGAAGCTTCATTGGGCGTGCGCTACCGCTTTTGATTCGCCGATTCCGATGCCGTCTGA
              :::::::: ::::::::: :::::::::::: :::::: ::::: ::::::::::::::
Del_5_Final   CACGAAGCCTCATTGGGCATGCGCTACCGCTTCTGATTCCCCGATACCGATGCCGTCTGA
               370       380       390       400       410       420

                1270      1280      1290      1300      1310      1320
WT_5_Final    ACCTTCAGACGGCATTTTTAAGGCGCAAGATCGGGCAAACGGCATTTTCAGACGGCATAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   ACCTTCAGACGGCATTTTTAAGGCGCAAGATCGGGCAAACGGCATTTTCAGACGGCATAA
               430       440       450       460       470       480

                1330      1340      1350      1360      1370      1380
WT_5_Final    CTGACAGTATAATCCGAACATCCGGCCGCTTTCGCGCGGCCTTCAGCATTATCCGATTTT
              :::::::::::::::::::::::::::::::::::::::::: ::::::::::::::
Del_5_Final   CTGACAGTATAATCCGAACATCCGGCCGCTTTCGCGCGGCCTTCGGCATTATCCGATTTT
               490       500       510       520       530       540

                1390      1400      1410      1420      1430
WT_5_Final    TTCCGAAAGCCGAACCATGCAATACAAACCCCTCCTGCTCGCCCTGATGCTTGG--CAGG
              :::::::::::::::::::::::::::::::::::::::::::::::::::  ::::
Del_5_Final   TTCCGAAAGCCGAACCATGCAATACAAACCCCTCCTGCTCGCCCTGATGCTTGGGGCAGG
               550       560       570       580       590       600

            1440      1450      1460      1470      1480      1490
WT_5_Final    CGGCGGGTGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATCTGGCTTACGCCTACG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   CGGCGGGTGAAGACCATGGGCGCGGCCCGTATGTGCAGGCGGATCTGGCTTACGCCTACG
               610       620       630       640       650       660

            1500      1510      1520      1530      1540         1550
WT_5_Final    AGCACATCACCCGCGATTATCCCGAAGCAACCGGTGCAAACCAAGGCAAAA---TAAGCA
              :::::::::::::::::::::::::::::::::::::  :::::: ::   ::::::
Del_5_Final   AGCACATCACCCGCGATTATCCCGAAGCAACCGGTGCAAAAAAAAGGCACAACAATAAGCA
               670       680       690       700       710       720

               1560      1570      1580      1590      1600      1610
WT_5_Final    CGGTAAGCGATTATTTCAAAAACATCCGCACCCGCTCCGTCCACCCCCGACTCGCCCTCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final   CGGTAAGCGATTATTTCAAAAACATCCGCACCCGCTCCGTCCACCCCCGACTCGCCCTCG
               730       740       750       760       770       780

               1620      1630      1640      1650      1660      1670
WT_5_Final    GCTACGATTTCGGCGGCTGGCGCGTCTTTTCCGCCCCCGCCGTTGCCGCCGGCCTCCACC
              :::::::::::::::::::::::::::::::::::::::::::::::::: :::::::::
Del_5_Final   GCTACGATTTCGGCGGCTGGCGCGTCTTTTCCGCCCCCGCCGTTGCCGCCAGCCTCCACC
```

```
                790       800       810       820       830       840

                1680      1690      1700      1710      1720      1730
WT_5_Final      AATCCCTTCAATATTACCGATCCCGCCATTGCCGCCGATTCCGTCAAATCCATCAATTCC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     AATCCCTTCAATATTACCGATCCCGCCATTGCCGCCGATTCCGTCAAATCCATCAATTCC
                850       860       870       880       890       900

                1740      1750      1760      1770      1780      1790
WT_5_Final      GCCGAATCACGCCTATTCCCCAAAAACCTTGATGCGGCGGGCTGAAGCCCGCCCTGCAAC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     GCCGAATCACGCCTATTCCCCAAAAACCTTGATGCGGCGGGCTGAAGCCCGCCCTGCAAC
                910       920       930       940       950       960

                1800      1810      1820      1830      1840      1850
WT_5_Final      CCTCTCTATGCACCCCCTTGCGAGCCCGACACTACGCAACATCTTGAGAACCCATCCTGT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     CCTCTCTATGCACCCCCTTGCGAGCCCGACACTACGCAACATCTTGAGAACCCATCCTGT
                970       980       990      1000      1010      1020

                1860      1870      1880      1890      1900      1910
WT_5_Final      CAAGAATACCCGAACCGTCCCGATACACCGTAATCCTAAAACCCGCCATTCCCGCGCTGC
                :::::::::::::::::::::::::::::::::::::::::: ::::::::::::::
Del_5_Final     CAAGAATACCCGAACCGTCCCGATACACCGTAATCCTAAAACCCGTCATTCCCGCGCTGC
                1030      1040      1050      1060      1070      1080

                1920      1930      1940      1950      1960      1970
WT_5_Final      AATGGGACATCGGCGGCAGCGGGGCGGTTTTCCCTTCGCTCGCACTGTTTCTGCTCTGTT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     AATGGGACATCGGCGGCAGCGGGGCGGTTTTCCCTTCGCTCGCACTGTTTCTGCTCTGTT
                1090      1100      1110      1120      1130      1140

                1980      1990      2000      2010      2020      2030
WT_5_Final      TCATCATAGGTATGCACAACACGGGGATGACGCTTCTGCCGGGCGGTGCAATCCGTTCGA
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     TCATCATAGGTATGCACAACACGGGGATGACGCTTCTGCCGGGCGGTGCAATCCGTTCGA
                1150      1160      1170      1180      1190      1200

                2040      2050      2060      2070      2080      2090
WT_5_Final      CGCACATGGCCCGGCACGGCAGCCGACTTGGGCATCGAAATCCCGCGCGTGCCGTACTAT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     CGCACATGGCCCGGCACGGCAGCCGACTTGGGCATCGAAATCCCGCGCGTGCCGTACTAT
                1210      1220      1230      1240      1250      1260

                2100      2110      2120      2130      2140      2150
WT_5_Final      AGTGGATTAACAAAAACCAGTACGGCGTTGGCCCGCCTTAGCTCAAAGAGAACGATTCTC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_5_Final     AGTGGATTAACAAAAACCAGTACGGCGTTGGCCCGCCTTAGCTCAAAGAGAACGATTCTC
                1270      1280      1290      1300      1310      1320

                2160      2170      2180      2190      2200      2210
WT_5_Final      TAAGGTGCTGAAGCACCAAGTGAATCGGTTCCGTACTATCTGTACTGTCTGCGGTTCGCC
                ::::::::::::::::::::::::::::::::::::::::::: :::::::::::::
Del_5_Final     TAAGGTGCTGAAGCACCAAGTGAATCGGTTCCGTACTATCTGTAC-GTCTGCGGTTCGCC
                1330      1340      1350      1360      1370      1380

                2220      2230      2240
```

```
WT_5_Final    GCCTGTCCTGATTTGTTCCCATCCTTTNNNNN-T
              :::::::: :::::::  :: :  :
Del_5_Final   GCCTGTCC-GATTTGTATCCGTTGTGGGGGGGGV
               1390       1400      1410
                         10        20        30        40        50        60
WT_6_Final    NNNNTNNNNNNNATCGGGNNNCGGTTTCTTCAGTTCTACGTTCTAGATTCCCGCCTGCGC

Del_6_Final   ------------------------------------------------------------


                        70        80        90       100       110       120
WT_6_Final    GGGAATGACGGGTTTCGAGATTGCGGTGTTGTCGGACGGGTTTCGAGATTACGGTGTTGT
                        :::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   ------------TTTCGAGATTGCGGTGTTGTCGGACGGGTTTCGAGATTACGGTGTTGT
                                10        20        30        40

                       130       140       150       160       170       180
WT_6_Final    CGGGAATGACGGTTCGGGTATTTTACTGCGCCCGCCCCGCGCCTGTAAACGGCAGGTGCA
              :::::::::::::::::::::::::::::::::::::::: ::::::::::::::::
Del_6_Final   CGGGAATGACGGTTCGGGTATTTTACTGCGCCCGCCCCGCGCCCGTAAACGGCAGGTGCA
              50        60        70        80        90       100

                       190       200       210       220       230       240
WT_6_Final    TCAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTGTCGGGGAATCAGAAGCGGTAGCG
              :::::::::::::::::::::::::::::::::: ::::::::::::::::::::::
Del_6_Final   TCAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAGCGGTAGCG
             110       120       130       140       150       160

                       250       260       270       280       290       300
WT_6_Final    CATGCCCAATGAGGCTTCGTGGGTTTTGAAGCGGGTGTTTTCCAAGCGTCCCCAGTTGTG
              :: :::::: :::::::::::::: : ::       ::::::
Del_6_Final   CACGCCCAACGAGGCTTCGTGGGTCTCGA-------GTTTTC------------------
             170       180       190               200

                       310       320       330       340       350       360
WT_6_Final    GTAGCGATACCCGGCGTCCAGGGTCAGCTTGGGCGTGATGTCGAAACCGACGCCGGCGAT

Del_6_Final   ------------------------------------------------------------


                       370       380       390       400       410       420
WT_6_Final    GACACCGAGACCCACGCGGCGGATGCTGTTGCTTTGGTAATCGTTTTGCGTCTTTGGATC

Del_6_Final   ------------------------------------------------------------


                       430       440       450       460       470       480
WT_6_Final    AGTATTATAAGTTGTAACTGCTCCGTTAGGAGCATTGCTGGGGACGGTAGTAACCTCTAT

Del_6_Final   ------------------------------------------------------------


                       490       500       510       520       530       540
WT_6_Final    TGTTTTTTTGGTCGAATCGATGCTGTGTCTGACGTGTCCGTAGGCGACGCGCGCGCCGAT

Del_6_Final   ------------------------------------------------------------
```

```
                  550       560       570       580       590       600
WT_6_Final    ATAGGGTTTGAATTTATCGTTTATTTGGAAATCGTAAATAGCGGATAAGCCGAGTGAGGA
                                                                    ::::
Del_6_Final   ------------------------------------------------------GAGG-

                  610       620       630       640       650       660
WT_6_Final    GACGGCGTGGAACGTACCGTTTTCCTGATTTTCCGTCTTCCGGTCTATCCTGATGCCATT

Del_6_Final   ------------------------------------------------------------

                  670       680       690       700       710       720
WT_6_Final    CTCTTTACGTATCCGCACATTTTCTATGTTGACGGAATATTTATTGTTGTTCCATTTTCT

Del_6_Final   ------------------------------------------------------------

                  730       740       750       760       770       780
WT_6_Final    GTAACGGGCATAATCTGCCGCTATCCTCCAGCCGCCGAAGTCGTAGCCGACCGACACCCG

Del_6_Final   ------------------------------------------------------------

                  790       800       810       820       830       840
WT_6_Final    GGGGTGGACGGAGCGGGTGCGGATGTTTCTGAAATAATCGCTTACCGTGCTTATTTTGTT

Del_6_Final   ------------------------------------------------------------

                  850       860       870       880       890       900
WT_6_Final    CTTGTTTGGAGCGGTTGGTTCCGGATAATCGTGGGTAATGTGTTCGTAGGCGTAGGCTAA

Del_6_Final   ------------------------------------------------------------

                  910       920       930       940       950       960
WT_6_Final    ATCCGCCTGCACATACGGGCCGCGGCCGCCGTCTTCACTTGCCGCCTGCGCTGCGGAGAG

Del_6_Final   ------------------------------------------------------------

                  970       980       990      1000      1010      1020
WT_6_Final    AAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGGTTTTTTGCGGGCTGGATTCAT

Del_6_Final   ------------------------------------------------------------

                 1030      1040      1050      1060      1070      1080
WT_6_Final    TTTCGGCTCCTTATATCGGTTTGACCGGTTAAAAAAAGATTTTCACTGATGTTGAAGGGC

Del_6_Final   ------------------------------------------------------------
```


1090 1100 1110 1120 1130 1140
WT_6_Final GGATTATATCGGGTTCCGGGCGGTGTTTCAACACATAGCACCGCGCCTGCTGCGCGTTTT
Del_6_Final ----------GGTTCCGGGCGGTGTTTCAACACATAGCACCGCGCCTGCTGCGCGTTTT
210 220 230 240 250
1150 1160 1170 1180 1190
WT_6_Final A-TGCGTTTGGCGCGTTCGGCGGCGGGAAATTTGCCTACTTTTCCCGCGTCGGGCGGGCG
Del_6_Final TGTGCGTTTGGCGCGTTCGGCGGCGGGAAATTTGCCTACTTTTCCCGCGTCGGGCGGGCG
260 270 280 290 300 310
1200 1210 1220 1230 1240 1250
WT_6_Final GAACGGGCGGCACACTGTCTATAAACCGCAATACCGTTTACAATGACCGCCTGTTTCACC
Del_6_Final GAACGGGCGGCACACTGTCTATAAACCGCAATACCGTTTACAATGACCGCCTGTTTCACC
320 330 340 350 360 370
1260 1270 1280 1290 1300 1310
WT_6_Final ACATACCCGAACGCAACAATGTTTCAACACACGGGACGGCACATCAAGCACCGCCCTATG
Del_6_Final ACACTCCCGAACGCAACAATGTTTCAACACACGGGACGGCACATCAAGCACCGCCCTATG
380 390 400 410 420 430
1320 1330 1340 1350 1360 1370
WT_6_Final TGTCGTCCTGATTTGGAAGGGGTTGCACCCCTCCCGAATAAAGTCTGATCCTGCCGCCCC
Del_6_Final TGTCGTCCTGATTTGGAAGGGGTTGCACCCCTCCCGAATAAAGTCTGATCCTGCCGCCCC
440 450 460 470 480 490
1380 1390 1400 1410 1420 1430
WT_6_Final GAAGGACGGATGTCCGAGTGGCGGGGTTTCAACCGAAAAGGAAATACAATGAAAATCAGG
Del_6_Final GAAGGACGGATGTCCGAGTGGCGGGGTTTCAACCGAAAAGGAAATACAATGAAAATCAGG
500 510 520 530 540 550
1440 1450 1460 1470 1480 1490
WT_6_Final CCGGGGCGGCACAACGCGCCCGACTTTCCGCACGGGGCCGCCGTAACCATAGGCAATTTC
Del_6_Final CCGGGGCGGCACAACGCGCCCGACTTTCCGCACGGGGCCGCCGTAACCATAGGCAATTTC
560 570 580 590 600 610
1500 1510 1520 1530 1540 1550
WT_6_Final GACGGCGTGCACCTCGGACACAAACACATCCTCCAAAAACTCCGCCTCGAAGCCGACGCA
Del_6_Final GACGGCGTACACCTCGGACACAAACACATCCTCCAAAAACTCCGCCTCGAAGCCGACGCA
620 630 640 650 660 670
1560 1570 1580 1590 1600 1610
WT_6_Final CGCGGATTGCCCGTCGTTGCCGTCATTTTCGAGCCGCAACCCAAAGAATTTTTTGCCGCC
Del_6_Final CGCGGATTGCCCGTCGTGGCCGTCGTTTTCGAACCCCAACCCAAAGAATTTTTCGCACTC
680 690 700 710 720 730
1620 1630 1640 1650 1660 1670

```
WT_6_Final    CGCACAGGCAAAACCCCGCCGTGCCGTATCAGCCCCCTGCGTACCAAACTCGGGCTGCTC
              :: :: :::::::::::::: :::::::::::::::::::::::::::::: :   ::::
Del_6_Final   CGTACCGGCAAAACCCCGCCCTGCCGTATCAGCCCCCTGCGTACCAAACTGGAATTGCTG
               740       750       760       770       780       790

           1680      1690      1700      1710      1720      1730
WT_6_Final    GAAAAAACCGGTTGCGTCGATGCCGCCTGGGTTTTGCGTTTCGATCGGAATTTTTCCGAA
              :::   :: ::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   GAAGGGACGGGTTGCGTCGATGCCGCCTGGGTTTTGCGTTTCGATCGGAATTTTTCCGAA
               800       810       820       830       840       850

           1740      1750      1760      1770      1780      1790
WT_6_Final    ATATCCGCGCAAGCATTTATCGACCGCCTGCTGCGTCAAACCTTGAATACGCGCTATTTG
              :::::::::::::::::::::::::::: ::::::::::::::::::::::::::::::
Del_6_Final   ATATCCGCGCAAGCATTTATCGACCGCCTACTGCGTCAAACCTTGAATACGCGCTATTTG
               860       870       880       890       900       910

           1800      1810      1820      1830      1840      1850
WT_6_Final    CTCGTCGGCGACGATTTCCGTTTCGGTGCGGGGCGGGAAGGCTGTTTTGAACTTTTGGCA
              :::::::::: :::::::::::::: ::::::::::::::::::::::::::::::::::
Del_6_Final   CTCGTCGGCGATGATTTCCGTTTCGGGGCGGGGCGGGAAGGCTGTTTTGAACTTTTGGCA
               920       930       940       950       960       970

           1860      1870      1880      1890      1900      1910
WT_6_Final    CAACAGCCCGATATGCAAACCGAGCGCACGCCTTCCGTCATCGTCGAAGACATCCGCACC
              :::::::::::::::: ::::::::::::::::: ::::: ::::::::::::::::::
Del_6_Final   CAACAGCCCGATATGCAGACCGAGCGCACGCCTTCTGTCATTGTCGAAGACATCCGCACC
               980       990      1000      1010      1020      1030

           1920      1930      1940      1950      1960      1970
WT_6_Final    AGCAGCACCACCGTCCGCCAAGCCCTTTCAGACGGCAACCTTGCCTATGCGAAGAAACTT
              ::::: ::: :::::::::::::::::::::::::::::::::::::::: ::::::
Del_6_Final   AGCAGTACCGCCGTCCGCCAAGCCCTTTCAGACGGCAACCTTGCCTATGCGAAAAAACTT
              1040      1050      1060      1070      1080      1090

           1980      1990      2000      2010      2020      2030
WT_6_Final    TTGGGGCACGACTACGTTTTGGGGGGCAGGGTGGTGCACGGCAGAAAACTCGGGCGCACC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TTGGGGCACGACTACGTTTTGGGGGGCAGGGTGGTGCACGGCAGAAAACTCGGGCGCACC
              1100      1110      1120      1130      1140      1150

           2040      2050      2060      2070      2080      2090
WT_6_Final    TTAAACGCCCCGACCGCCAACATCCGACTGCCCGGCCACCGTTATGCCCTCGGCGGCGTG
              :::::::::::::::::::::::::::::::::::::::::::: ::::::::::::
Del_6_Final   TTAAACGCCCCGACCGCCAACATCCGACTGCCCGGCCACCGTTATGCACTCGGCGGCGTG
              1160      1170      1180      1190      1200      1210

           2100      2110      2120      2130      2140      2150
WT_6_Final    TTTGTCGTCGAAGCGGACGGCGCATTCGGCACGCGGCGCGGCGTGGCGAGCTTCGGCTTC
              :: :::::::: ::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TTCGTCGTCGAGGCGGACGGCGCATTCGGCACGCGGCGCGGCGTGGCGAGCTTCGGCTTC
              1220      1230      1240      1250      1260      1270

           2160      2170      2180      2190      2200      2210
WT_6_Final    AATCCCACCGTTGATGGCGGCTGTTCTCAAAAGCTTGAAGTCCACCTGTTCGACTTTCAA
              :::::::::::::::::::::::: :::::: ::: ::::::::  ::::::::::::::
Del_6_Final   AATCCCACCGTTGATGGCGGCTGTCCTCAAAGGCTGGAAGTCCATTTGTTCGACTTTCAA
```

```
             1280      1290      1300      1310      1320      1330

          2220      2230      2240      2250      2260      2270
WT_6_Final    GGCGATTTGTACGGACAACGGTTGAACGTCCGCTTCCTGCACAAACTGCGCGACGAGGAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   GGCGATTTGTACGGACAACGGTTGAACGTCCGCTTCCTGCACAAACTGCGCGACGAGGAA
              1340      1350      1360      1370      1380      1390

          2280      2290      2300      2310      2320      2330
WT_6_Final    AAGTTTGACGGTATGGAAGAACTGAAAAGGCGGATTGAAGCCGATATGGAAGCCGCAAAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   AAGTTTGACGGTATGGAAGAACTGAAAAGGCGGATTGAAGCCGATATGGAAGCCGCAAAG
              1400      1410      1420      1430      1440      1450

          2340      2350      2360      2370      2380      2390
WT_6_Final    TGTTGGTAGAAAAACCTTATACAAACCATCCGATTGGGCTACAATCAGCCTTTTAACTGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TGTTGGTAGAAAAACCTTATACAAACCATCCGATTGGGCTACAATCAGCCTTTTAACTGT
              1460      1470      1480      1490      1500      1510

          2400      2410      2420      2430      2440      2450
WT_6_Final    TCAGACGGCATAGGGGTTTCCCGTTGTGAAATACTGTTTGAGGGGCAATGCCGTCTGAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TCAGACGGCATAGGGGTTTCCCGTTGTGAAATACTGTTTGAGGGGCAATGCCGTCTGAAA
              1520      1530      1540      1550      1560      1570

          2460      2470      2480      2490      2500      2510
WT_6_Final    CCGAAATATTGTAACAATAGAGATTAAAAAATGACCGATTACAGTAAAACCGTCAACCTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CCGAAATATTGTAACAATAGAGATTAAAAAATGACCGATTACAGTAAAACCGTCAACCTG
              1580      1590      1600      1610      1620      1630

          2520      2530      2540      2550      2560      2570
WT_6_Final    CTCGAAAGCCCGTTTCCGATGCGCGGCAATCTTGCCAAGCGCGAGCCTGCGTGGCTGAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CTCGAAAGCCCGTTTCCGATGCGCGGCAATCTTGCCAAGCGCGAGCCTGCGTGGCTGAAA
              1640      1650      1660      1670      1680      1690

          2580      2590      2600      2610      2620      2630
WT_6_Final    AGCTGGTACGAGCAAAAACGTTACCAAAAACTGCGCGAAATCGCCAAAGGCCGTCCGAAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   AGCTGGTACGAGCAAAAACGTTACCAAAAACTGCGCGAAATCGCCAAAGGCCGTCCGAAA
              1700      1710      1720      1730      1740      1750

          2640      2650      2660      2670      2680      2690
WT_6_Final    TTCATCCTGCACGACGGCCCGCCGTATGCCAACGGCGACATCCATATCGGTCATGCCGTC
              ::::: ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TTCATTCTGCACGACGGCCCGCCGTATGCCAACGGCGACATCCATATCGGTCATGCCGTC
              1760      1770      1780      1790      1800      1810

          2700      2710      2720      2730      2740      2750
WT_6_Final    AACAAAATCCTTAAAGACATTATTATCCGCAGCAAAACCCAAGCCGGTTTTGACGCGCCT
              :: ::::: ::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   AATAAAATTCTTAAAGACATTATTATCCGCAGCAAAACCCAAGCCGGTTTTGACGCGCCT
              1820      1830      1840      1850      1860      1870

          2760      2770      2780      2790      2800      2810
```

```
WT_6_Final    TATGTACCGGGTTGGGACTGCCACGGCCTGCCCATCGAAGTGATGGTGGAAAAGCTGCAC
              :::::::::::::::::::::::::::::::::::::::::::::::::: ::::::
Del_6_Final   TATGTACCGGGTTGGGACTGCCACGGCCTGCCCATCGAAGTGATGGTGGAAAAACTGCAC
              1880      1890      1900      1910      1920      1930

           2820      2830      2840      2850      2860      2870
WT_6_Final    GGCAAAGATATGCCTAAAGCCCGTTTCCGCGAATTGTGCCGCGAATATGCCGCCGAACAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   GGCAAAGATATGCCTAAAGCCCGTTTCCGCGAATTGTGCCGCGAATATGCCGCCGAACAG
              1940      1950      1960      1970      1980      1990

           2880      2890      2900      2910      2920      2930
WT_6_Final    ATTGCCCGTCAGAAAAAAGACTTTATCCGCTTGGGCGTGTTGGGCGATTGGGACAATCCT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   ATTGCCCGTCAGAAAAAAGACTTTATCCGCTTGGGCGTGTTGGGCGATTGGGACAATCCT
              2000      2010      2020      2030      2040      2050

           2940      2950      2960      2970      2980      2990
WT_6_Final    TACTTGACTATGGATTTCAAAACCGAAGCCGATACCGTGCGTATGCTCGGCGAAATCTAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TACTTGACTATGGATTTCAAAACCGAAGCCGATACCGTGCGTATGCTCGGCGAAATCTAC
              2060      2070      2080      2090      2100      2110

           3000      3010      3020      3030      3040      3050
WT_6_Final    AAATCCGGCTATCTCTACCGTGGTGCGAAACCGGTTCAGTTTTGCTTGGATTGCGGATCT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   AAATCCGGCTATCTCTACCGTGGTGCGAAACCGGTTCAGTTTTGCTTGGATTGCGGATCT
              2120      2130      2140      2150      2160      2170

           3060      3070      3080      3090      3100      3110
WT_6_Final    TCGCTGGCGGAAGCGGAAGTGGAATACAAAGACAAAGTATCGCCTGCGATTGACGTTGCC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TCGCTGGCGGAAGCGGAAGTGGAATACAAAGACAAAGTATCGCCTGCGATTGACGTTGCC
              2180      2190      2200      2210      2220      2230

           3120      3130      3140      3150      3160      3170
WT_6_Final    TATCCGTTTAAAGACACTGCCGCGCTTGCCGCCGCATTCGGCTTGGCAGGTATCGAAGGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TATCCGTTTAAAGACACTGCCGCGCTTGCCGCCGCATTCGGCTTGGCAGGTATCGAAGGC
              2240      2250      2260      2270      2280      2290

           3180      3190      3200      3210      3220      3230
WT_6_Final    AAAGCGTTTGCCGTCATTTGGACGACCACGCCTTGGACTCTGCCTGCGAGCCAGGCCGTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   AAAGCGTTTGCCGTCATTTGGACGACCACGCCTTGGACTCTGCCTGCGAGCCAGGCCGTG
              2300      2310      2320      2330      2340      2350

           3240      3250      3260      3270      3280      3290
WT_6_Final    TCTGCCGGCGCGGACGTGGTGTATCAATTAATCGATACGCCCAAAGGCAAATTGGTGCTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   TCTGCCGGCGCGGACGTGGTGTATCAATTAATCGATACGCCCAAAGGCAAATTGGTGCTG
              2360      2370      2380      2390      2400      2410

           3300      3310      3320      3330      3340      3350
WT_6_Final    GCGAAAGATTTGGCGGAAGGCGCTTTGAAACGCTACGGCTTTTCAGACGGCATCGCCATC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   GCGAAAGATTTGGCGGAAGGCGCTTTGAAACGCTACGGCTTTTCAGACGGCATCGCCATC
```

```
              2420      2430      2440      2450      2460      2470

           3360      3370      3380      3390      3400      3410
WT_6_Final    CTTGCCGAAACCACCGGCGACAAGCTGGAAAACCTGCACATGAATCATCCGTTCCTCGAA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CTTGCCGAAACCACCGGCGACAAGCTGGAAAACCTGCACATGAATCATCCGTTCCTCGAA
              2480      2490      2500      2510      2520      2530

           3420      3430      3440      3450      3460      3470
WT_6_Final    CGCGATATTCCCATGCTCAACGGCGAACACGTTACCACCGATGCCGGTACCGGCTTGGTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CGCGATATTCCCATGCTCAACGGCGAACACGTTACCACCGATGCCGGTACCGGCTTGGTG
              2540      2550      2560      2570      2580      2590

           3480      3490      3500      3510      3520      3530
WT_6_Final    CATACTGCGCCTGCGCACGGTTTGGAAGACTACGCCGTCTGCAATAAATACGGCATCGAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CATACTGCGCCTGCGCACGGTTTGGAAGACTACGCCGTCTGCAATAAATACGGCATCGAG
              2600      2610      2620      2630      2640      2650

           3540      3550      3560      3570      3580      3590
WT_6_Final    CTTTACAACCCTGTCAACGCCGAAGGCAAATACATAAGCGAAACGCCTCGTGTCGCAGGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   CTTTACAACCCTGTCAACGCCGAAGGCAAATACATAAGCGAAACGCCTCGTGTCGCAGGC
              2660      2670      2680      2690      2700      2710

           3600      3610      3620      3630      3640      3650
WT_6_Final    ATGAGCGTTTGGGAGGCGAATCCCGTCATCCTGCAATGGCCGGAAGAAACCGGCAACCTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Del_6_Final   ATGAGCGTTTGGGAGGCGAATCCCGTCATCCTGCAATGGCCGGAAGAAACCGGCAACCTC
              2720      2730      2740      2750      2760      2770

           3660      3670      3680      3690      3700      3710
WT_6_Final    TTGGCAAGCAGCAAAATCGAACACAGCTACGCCCACTGCTGGCGCCACAAAACCCCGCTG
              ::::::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
Del_6_Final   TTGGCAAGCAGCAAAATCGAACACNGCTACGCCCACTGCTGGCGCCACAAAACCCCGCTG
              2780      2790      2800      2810      2820      2830

           3720      3730      3740      3750      3760      3770
WT_6_Final    ATTTACCGAGCGACAGGTCAGTGGTTTGTCGGCATGGACAAAGCCGGCAGCGACGGTAAA
              :::::::::::::::::::::::::::::::::: ::::::::::::::::::::::::
Del_6_Final   ATTTACCGAGCGACAGGTCAGTGGTTTGTCGGCATNGACAAAGCCGGCAGCGACGGTAAA
              2840      2850      2860      2870      2880      2890

           3780      3790
WT_6_Final    ACCCTGCGNNNNNNNN
              : : : ::::::
Del_6_Final   ANCNTNCGNNNN----
              2900


                                    10          20        30        40
WT_7_Final    -----------------ABNCCTCCGATCA---CGCAGGCGAGNGGTTTGAACCGTCTGC
                               :. ::  ::::     ::: :   ::    :  ::::::::::
MS11_DEL_7_F  NNNNNNANNTGCTGCATACACCCGCGATGTGGGCGCGGTGCAGGTNATGAAACCGTCTGC
                       10        20        30        40        50        60
```

```
                     50        60        70        80        90
WT_7_Final    AA-CGCTGTGCGGAGAATGTCGGCGGAGAAGGCGAAACGCGCCATTACGCCTGATAAAGT
              :: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  AAACGCTGTGCGGAGAATGTCGGCGGAGAAGGCGAAACGCGCCATTACGCCTGATAAAGT
                     70        80        90       100       110       120

            100       110       120       130       140       150
WT_7_Final    TTGAGCAAATGCCGTCTGAAACCGCCAAACAGTATTTCAGACGGCATTTTTTGTGCCGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  TTGAGCAAATGCCGTCTGAAACCGCCAAACAGTATTTCAGACGGCATTTTTTGTGCCGCG
                     130       140       150       160       170       180

            160       170       180       190       200       210
WT_7_Final    ATTTGTCTTCATAATGGCGGAGGGGTTTCGAGATTGCGGTGTTGTCGGGAATGACGGTTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  ATTTGTCTTCATAATGGCGGAGGGGTTTCGAGATTGCGGTGTTGTCGGGAATGACGGTTC
                     190       200       210       220       230       240

            220       230       240       250       260       270
WT_7_Final    AGTTGCTACGGTTACTGTCAGGTTTCGGTTATGTTGGAATTTCGGGAAACTTATGAATTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  AGTTGCTACGGTTACTGTCAGGTTTCGGTTATGTTGGAATTTCGGGAAACTTATGAATTG
                     250       260       270       280       290       300

            280       290       300       310       320       330
WT_7_Final    TCATTCCCGCGAAAGTGGGAATCCAGGACTCAAAATCTCAAGAAACCGTTTTGCCCGATA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  TCATTCCCGCGAAAGTGGGAATCCAGGACTCAAAATCTCAAGAAACCGTTTTGCCCGATA
                     310       320       330       340       350       360

            340       350       360       370       380       390
WT_7_Final    AGTTCCTGCACTGACAGATCTAGATTTCCGCCTGCGCGGGAATGACGGGTTTCGAGATTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  AGTTCCTGCACTGACAGATCTAGATTTCCGCCTGCGCGGGAATGACGGGTTTCGAGATTG
                     370       380       390       400       410       420

            400       410       420       430       440       450
WT_7_Final    CGGTGTTGTCGGGAATGACGGTTCGGGTATTTTACTGCGCCCGCCCCGCGCCTGTAAACG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  CGGTGTTGTCGGGAATGACGGTTCGGGTATTTTACTGCGCCCGCCCCGCGCCTGTAAACG
                     430       440       450       460       470       480

            460       470       480       490       500       510
WT_7_Final    GCGGGCGATTAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  GCGGGCGATTAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAG
                     490       500       510       520       530       540

            520       530       540       550       560       570
WT_7_Final    CGGTAGCGCACGCCCAATGAGGCTTCGTGGGTTTTGAAGCGGGTGTTTTCCAGGCGTCCC
              ::::::::::::::::::::::::::::::
MS11_DEL_7_F  CGGTAGCGCACGCCCAATGAGGCTTCGTGGGT----------------------------
                     550       560       570

            580       590       600       610       620       630
WT_7_Final    CAATAGTGGTAGCGGTACCCGGCGTCCAAGGTCAGGTTGGGCGTGATGTCGATGCCTACG
```

```
MS11_DEL_7_F  ------------------------------------------------------------

            640       650       660       670       680       690
WT_7_Final    CCTGCCACTGCGCCGAGGCCCACGCGGCGGATGCTGTCGCTTTCGCGATGGGCGTCTTGC

MS11_DEL_7_F  ------------------------------------------------------------

            700       710       720       730       740       750
WT_7_Final    GTATTTTTCCCAGGATAAACCGTAGGGGTTGTGCCAGGACCATGGAGGATGGTAGTAACC

MS11_DEL_7_F  ------------------------------------------------------------

            760       770       780       790       800       810
WT_7_Final    TCTGTTGTTTTTTTGGTCGAATCGATGCTGTGTCTGACGTGTCCGTAGGCGACGCGCGCG

MS11_DEL_7_F  ------------------------------------------------------------

            820       830       840       850       860       870
WT_7_Final    CCGATATAGGGTTTGAATTTATCGTTGAGTTTGAAGTCGTAAACGGCGGACAAGCCGAGA

MS11_DEL_7_F  ------------------------------------------------------------

            880       890       900       910       920       930
WT_7_Final    GAAGAAACGGCGTGGAACGTACCGTTTTCCTGATTTTCCGTCTTCCGGTCTGTCCTGTTG

MS11_DEL_7_F  ------------------------------------------------------------

            940       950       960       970       980       990
WT_7_Final    CCATTGACCTTGTTTCTTAGCAACTCTTTTATGCTCACGGAATATTTATTGTTGTTCCAC

MS11_DEL_7_F  ------------------------------------------------------------

           1000      1010      1020      1030      1040      1050
WT_7_Final    TTTCTGTAACGGGCATAATCTGCCGCTATCCTCCAGCCGCCGAAATCGTAGCCGACCGAC
                                          :: :::
MS11_DEL_7_F  ------------------------------CTGCAG------------------------

           1060      1070      1080      1090      1100      1110
WT_7_Final    ACCCGGGGGTGGATGGAATGCGTACGGATGTTTCTGAAATAATCGCTTACCGTGCTTATT

MS11_DEL_7_F  ------------------------------------------------------------

           1120      1130      1140      1150      1160      1170
WT_7_Final    TTGCCTTTGCTTGGATCGGTTTGTTCCGGATAATCGTGGGTAATGTGTTCGTAGGCGTAG

MS11_DEL_7_F  ------------------------------------------------------------
```

```
          1180      1190      1200      1210      1220      1230
WT_7_Final    GCTAAATCCGCCTGCACATACGGGCCGCGCCCATGGTCTTCACCCGCCGCCTGCGCTGCG

MS11_DEL_7_F  ------------------------------------------------------------


          1240      1250      1260      1270      1280      1290
WT_7_Final    GAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGGTTTTTTGGGGGCTGG

MS11_DEL_7_F  ------------------------------------------------------------


          1300      1310      1320      1330      1340      1350
WT_7_Final    ATTCATTTTCGGCTCCTTATTCGGTTTGACCGGTTAAAAAAAGATTTTCACTGATGTTGA

MS11_DEL_7_F  ------------------------------------------------------------


          1360      1370      1380      1390      1400      1410
WT_7_Final    AGGGCGGATTATATCGGGTTCCGGGCGGTGTTTCAACACAATATGGCGGATGAACAAAAA
                           :::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  ---------------GGGTTCCGGGCGGTGTTTCAACACAATATGGCGGATGAACAAAAA
                         580       590       600       610       620

          1420      1430      1440      1450      1460      1470
WT_7_Final    CCGGTACGTCGTTGCCCCGCCCCGGCCCAAAGGGAACGGTTCCCCAAGGTGATGAAGCAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  CCGGTACGTCGTTGCCCCGCCCCGGCCCAAAGGGAACGGTTCCCCAAGGTGATGAAGCAC
                  630       640       650       660       670       680

          1480      1490      1500      1510      1520      1530
WT_7_Final    CGGGCGGATCGGTTCCGTACCATTTGTACTGCCTGCGGCCCGCCGCCTTGTCCTGATTTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  CGGGCGGATCGGTTCCGTACCATTTGTACTGCCTGCGGCCCGCCGCCTTGTCCTGATTTT
                  690       700       710       720       730       740

          1540      1550      1560      1570      1580      1590
WT_7_Final    TGTTAATCCGCTATAAAGACCGTCGGGCATCTTTCCGCCGTCATTCCCACGCAGGCGGGA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  TGTTAATCCGCTATAAAGACCGTCGGGCATCTTTCCGCCGTCATTCCCACGCAGGCGGGA
                  750       760       770       780       790       800

          1600      1610      1620      1630      1640      1650
WT_7_Final    ATCTGGAATTTCAATGCCTCAAGAATTTATCGGAAAAAACCAAAACCCTTCCGCCGTCAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  ATCTGGAATTTCAATGCCTCAAGAATTTATCGGAAAAAACCAAAACCCTTCCGCCGTCAT
                  810       820       830       840       850       860

          1660      1670      1680      1690      1700      1710
WT_7_Final    TCCCACTTTCGTGGGAATGACGGGATGTAGGTTCGTAGGAATGACGTGGTGCAGGTTTCC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_7_F  TCCCACTTTCGTGGGAATGACGGGATGTAGGTTCGTAGGAATGACGTGGTGCAGGTTTCC
                  870       880       890       900       910       920

          1720      1730      1740      1750      1760      1770
WT_7_Final    GTGCGGATGGATTCGTCATTCCCGCGCAGGCGGGAATCCGGTCCGTTCGGTTTCAGTTAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
MS11_DEL_7_F  GTGCGGATGGATTCGTCATTCCCGCGCAGGCGGGAATCCGGTCCGTTCGGTTTCAGTTAT
                  930       940       950       960       970       980

           1780      1790      1800      1810      1820      1830
WT_7_Final    TTCCGATAAATTCCTGCTGCTTTTTATTTCTAGATTCCCACTTCCGTGGGAAT-ACGGCG
              :::::::::::::::::::::::::::::::::::::: :::::::::::: ::::::
MS11_DEL_7_F  TTCCGATAAATTCCTGCTGCTTTTTATTTCTAGATTCCCTCTTCCGTGGGAATGACGGCG
                  990      1000      1010      1020      1030      1040

             1840      1850      1860
WT_7_Final    -AGGGAAATCTCATGTTTTNNNT
               :::: :   : :  : :  :::
MS11_DEL_7_F  GAGGGGA---TAAGNTCTGCNNT
                 1050         1060


                     10        20         30        40        50
WT_8_final    ABTTTCTTGACTACGGGTTTTACAA-ATCGGTGGCGTTTGGTTTGCGCCTCCGGCAGCGG
                     :        ::  : :: :::::: ::::::::::::  ::::::::::::
MS11_DEL_8_F  NNNNNNNNGGNNC----TTCGANAATATCGGTTGCGTTTGGTTTGNNCCTCCGGCAGCGG
                     10            20        30        40        50

              60        70        80        90       100       110
WT_8_final    GCGGTTGCGGCGGGTTTTGCACATAATGCCGTCCAGCAGCCGATGCTCTTTCATTTTTCC
              : ::::::::::::: ::::: ::::::::::::::::::::::::::::::: ::::::
MS11_DEL_8_F  GGGGTTGCGGCGGGCTTTGCGCATAATGCCGTCCAGCAGCCGATGCTCTTTCAGTTTTCC
                60        70        80        90       100       110

            120       130       140       150       160       170
WT_8_final    GTAGGTCGGATTCTCGAATCCAACTTTACTTCAATCGTATTCAATAAAAAAGTCCGCTTT
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::::: ::
MS11_DEL_8_F  GTAGGTCGGATTCTCGAATCCAACTTTACTTCAATCGTATTCAATAAAAAAGTCCGCATT
               120       130       140       150       160       170

            180       190       200       210       220       230
WT_8_final    GCCCCCACCCCAATTATGCGGATAAATACCCGGTTTGACATAGGGGTGAAACGTAAAAAA
              :::::: :::::::::::::::::::::::::::::::::  ::::::::::::::::
MS11_DEL_8_F  GCCCCCCCCCCCAATTATGCGGATAAATACCCGGTTTGACATAACGGTGAAACGTAAAAAA
               180       190       200       210       220       230

            240       250       260       270       280       290
WT_8_final    CCGCCAATCGGAAATTTGTCCTACATAGCCTTGTTTGACCGGATTGAAATGCAAATAATC
              :::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  CCGCCAATCGGAAATTTGCCCTACATAGCCTTGTTTGACCGGATTGAAATGCAAATAATC
               240       250       260       270       280       290

            300       310       320       330       340       350
WT_8_final    AAAATGGCAGGCAAAATCGGCCTCACCGCGAATAGTATGTTCCCAAAAGCGTTTTTCCAA
              :::::::::::::::::::::: ::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  AAAATGGCAGGCAAAATCGGCCTCCCCGCGAATAGTATGTTCCCAAAAGCGTTTTTCCAA
               300       310       320       330       340       350

            360       370       380       390       400       410
WT_8_final    AGCCTGAAATTGCCGCCAATTAAATATTGGCTGGGCCGCTTGATTTGCCGCCCGCGTTCC
              :::::::::::::::::::::::::: ::::: :::::::::::::: :: :::::::
MS11_DEL_8_F  AGCCTGAAATTGCCGCCAATTAAATATGGGCTGTGCCGCTTGATTTGCCCCCAGCGTTCC
               360       370       380       390       400       410
```

```
           420       430       440       450       460       470
WT_8_final    AAATAAGCAAAATCTTTGTCCGGCCCCCGCCATATGGTATGCAAATGATCGGGCATCAAC
              ::: :: :::::::: ::: ::::: ::::::::::::::::::::::::::::::::
MS11_DEL_8_F  AAAAAAACAAAATCATTGCCCGGCACCCGCCATATGGTATGCAAATGATCGGGCATCAAC
               420       430       440       450       460       470

           480       490       500       510       520       530
WT_8_final    CCCCATGCCAAAATTTCAAACGGATTCCGTTCGCGCACCGCCATTACCGCCTGCCGTAAA
              :::::::::::::::::::::::::::::::::::::::::: :::::::::::: :::
MS11_DEL_8_F  CCCCATGCCAAAATTTCAAACGGATTCCGTTCGCGCACCGCCTTTACCGCCTGCCGAAAA
               480       490       500       510       520       530

           540       550       560       570       580       590
WT_8_final    GCCAAACGCACCGCATCATCGGTCAAAATCTTCTGCCGTTTATTGGTTACAACCGTAAAA
              :::::::::::::::::::::::::::::::::::::::::::::::::::::: ::::
MS11_DEL_8_F  GCCAAACGCACCGCATCATCGGTCAAAATCTTCTGCCGTTTATTGGTTACAACCGAAAAA
               540       550       560       570       580       590

           600       610       620       630       640       650
WT_8_final    AAGTAAGTGCCGCCATTGCGCGTAAAAACGACGGTATTTCATAGTATTATGCTCGGAATG
              :::::::::::::::::::: :::::::: :::::::::::: :: ::::::::::::
MS11_DEL_8_F  AAGTAAGTGCCGCCATTGCG-GTAAAAACAACGGTATTTCATAATANTATGCTCGGAATG
               600       610        620       630       640       650

           660       670       680       690       700       710
WT_8_final    ATTTTGTAGGCCGGATTCTTGAATTCGACATTTTGGACATTGCTGCAATGGATTGCAATG
              :::::: ::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  ATTTTGNAGGCCGGATTCTTGAATTCGACATTTTGGACATTGCTGCAATGGATTGCAATG
                660       670       680       690       700       710

           720       730       740       750       760       770
WT_8_final    ATGGGAATGTTAAAGGTTTTGTCGGATACAAGTATCCGACCTACGCTTGCTGAACCGTCA
              :::::::::::::::::::::::::::::::::::: :::::::::: :::::::::
MS11_DEL_8_F  ATGGGAATGTTAAAGGTTTTGTCGGATACAAGTATCCGGCCTACGCTTGCGGAACCGTCA
                720       730       740       750       760       770

           780       790       800       810             820       830
WT_8_final    TTCCCACGGAAGTGGGAATCTAGGACGCGGGGTTT------CAGTCATTTCCGATAGATT
              :::::::: :::::::::::::::::::::::      : ::  : :::::::  ::
MS11_DEL_8_F  TTCCCACGAAAGTGGGAATCTAGGACGCGGGGTTTGGGCAACCGTTTTATCCGATAAGTT
                780       790       800       810       820       830

                840       850       860       870       880       890
WT_8_final    CCCGCCGCGTCGGGGGTCCGGATTTCCGCCTGCGCGGGAATGACGAATTTCGAGATTGCG
               :::    : : ::  ::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  TCCGTGCGGACAGG--TCCGGATTTCCGCCTGCGCGGGAATGACGAATTTCGAGATTGCG
                840         850       860       870       880       890

                900                                       910       920
WT_8_final    GTGTTGTCGG----------------------------GAATGACGGTTCGGGTATTTCC
              ::::::::::                              :::::::::::::::::::::
MS11_DEL_8_F  GTGTTGTCGGACGGGTTTCGAGATTACGGTATTGTCGGGAATGACGGTTCGGGTATTTCC
                900       910       920       930       940       950

              930       940       950       960       970       980
WT_8_final    ACGCCCGCCCCGCGCCTGTAAACAGCGGGCGATTAAAAATGCCGCCTGAAGGTTCAGACG
```

```
             : :::::::::::::::::::: :::::::::::::::::::: ::::::::::::::
MS11_DEL_8_F AAGCCCGCCCCGCGCCTGTAAACGGCGGGCGATTAAAAATGCCGTCTGAAGGTTCAGACG
                 960       970       980       990      1000      1010

               990      1000      1010      1020      1030      1040
WT_8_final   GCATCGGTATCGGGGAATCAGAAGCGGTAGCGCACGCCCAATGAGGCTTCGTGGGTTTTG
             :::::::: :::::::::::::::::::::::::::::::: :::::::::::::: :
MS11_DEL_8_F GCATCGGTGTCGGGGAATCAGAAGCGGTAGCGCACGCCCAACGAGGCTTCGTGGGTCT--
                1020      1030      1040      1050      1060      1070

              1050      1060      1070      1080      1090      1100
WT_8_final   AAGCGGGTGTTTTCCAAGCGTCCCCAGTTGTGGTAGCGGTAGCCGGCGTCCAGGGTCAGC

MS11_DEL_8_F ------------------------------------------------------------


              1110      1120      1130      1140      1150      1160
WT_8_final   TTGGGCGTGATGTCGATGCCTACGCCTGCCACTGCGCCGAAGCCCAAGCTGCTGATGCTG

MS11_DEL_8_F ------------------------------------------------------------


              1170      1180      1190      1200      1210      1220
WT_8_final   CGGCTTTCGTGATGGGGAAGTTTGCGGATCGGAGCATTTGTGGTAACACTTGACGCAGCA

MS11_DEL_8_F ------------------------------------------------------------


              1230      1240      1250      1260      1270      1280
WT_8_final   TTCTGTGGGTAAGTGGTAACAGCAATGGTTTCTTGTTGAACCGAACGAACCTGATGTCTG

MS11_DEL_8_F ------------------------------------------------------------


              1290      1300      1310      1320      1330      1340
WT_8_final   ACGTGTCCGTAGGCGACGCGCATGCCGATATAGGGTTTGAAGCGGGAACCGGTATCGAAA

MS11_DEL_8_F ------------------------------------------------------------


              1350      1360      1370      1380      1390      1400
WT_8_final   TCGTAAACGGCGGACAAGCCGAGAGAAGAGGCGGCGTGGAATGTGCCGTTTTCCGATGTT
                                  ::::
MS11_DEL_8_F ------------------CGAG------------------------------------


              1410      1420      1430      1440      1450      1460
WT_8_final   CCGTCTTCGGGTTTTTATGTTAAGATGCTTGTTGCCGCCAGAATTGACATATCGTTTCTT
                           ::::
MS11_DEL_8_F -------------TTTA-------------------------------------------


              1470      1480      1490      1500      1510      1520
WT_8_final   AGCAACTCTTTTATGTTAACGGAATATTTATTGTTGTTCCACTTTCTGTAACGGGCATAA
                :::::
MS11_DEL_8_F ---AACTC-----------------------------------------------------
             1080
```

```
              1530      1540      1550      1560      1570      1580
WT_8_final    TCTGCCGCTATCCTCCAGCCGCCGAAGTNGTAGCCGACCGACACCCTGGGGTGGATGGAA

MS11_DEL_8_F  ------------------------------------------------------------


              1590      1600      1610      1620      1630      1640
WT_8_final    TGCGTACGGATGTTTCTGAAATAATCGCTTACCGTGCTTATTTTGTCTTTTTTTGTACCG

MS11_DEL_8_F  ------------------------------------------------------------


              1650      1660      1670      1680      1690      1700
WT_8_final    GTTGGTTCCGGATAATCGTGGGTAATGTGTTCGTAGGCGTAGGCTAAATCCGCCTGCACA

MS11_DEL_8_F  ------------------------------------------------------------


              1710      1720      1730      1740      1750      1760
WT_8_final    TACGGGCCGCGGCCATTGCCTTCACCCGCCGCCTGCGCTGCGGAAGGGAAGAGAAGGGAA

MS11_DEL_8_F  ------------------------------------------------------------


              1770      1780      1790      1800      1810      1820
WT_8_final    GAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAA
                   :::
MS11_DEL_8_F  -----GAG-------------------------------------------------------


              1830      1840      1850      1860      1870      1880
WT_8_final    GAGAAGGTTTTTTGCGGGCTGGATTCATTTTCGGCTCCTTATTCGGTTTAACCGGTTAAA

MS11_DEL_8_F  ------------------------------------------------------------


              1890      1900      1910      1920      1930      1940
WT_8_final    AAAAGATTTTCACTGATGTTGAAGGGCGGATTATATCGGGTTCCGGGCGGTGTTTCAACA
                                                 ::::::::::::::::::::::
MS11_DEL_8_F  -----------------------------------GGGTTCCGGGCGGTGTTTCAACA
                                              1090      1100      1110


              1950      1960      1970      1980      1990      2000
WT_8_final    CAATATGGCGGATGAACAAAAACCGGTACGTCGTTGCCCCGCCCCGGCTCAAAGGGAACG
              ::::::::::::::::::::::::::::::  :::::::::::::::::::::::::::
MS11_DEL_8_F  CAATATGGCGGATGAACAAAAACCGGTACGG-GTTGCCCCGCCCCGGCTCAAAGGGAACG
                    1120      1130      1140      1150      1160


              2010      2020      2030      2040      2050      2060
WT_8_final    GTTCCCTAAGGCGCCCAAGCACCGGGCGGATCGGTTCCGTACCATTTGTACTGCCTGCGG
              :::::::::::::: :::::::::::::: : ::::::::::::::  ::::::::::::
MS11_DEL_8_F  GTTCCCTAAGGCGTCCAAGCACCGGGCGAACCGGTTCCGTACCATCCGTACTGCCTGCGG
         1170      1180      1190      1200      1210      1220


              2070      2080      2090      2100      2110      2120
WT_8_final    CCCGCCGCCTTGTCCTGATTTTTGTTAATCCGCTATAAACAACGCTTCGTCCGAAAAAAC
```

```
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  CCCGCCGCCTTGTCCTGATTTTTGTTAATCCGCTATAAACAACGCTTCGTCCGAAAAAAC
           1230      1240      1250      1260      1270      1280

              2130      2140      2150      2160      2170      2180
WT_8_final    GATTGAATTTGCGGGCAGAAGCTGGACGAAAACCGCCGACAGCCTGCCGCAAAAGGCACA
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  GATTGAATTTGCGGGCAGAAGCTGGACGAAAACCGCCGACAGCCTGCCGCAAAAGGCACA
           1290      1300      1310      1320      1330      1340

              2190      2200      2210      2220      2230      2240
WT_8_final    CGGTTTGCGCTACGGCTTGGGCGTGTCGCGCGAAATCAATGCGGGCGGGCATCATTCCTC
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_8_F  CGGTTTGCGCTACGGCTTGGGCGTGTCGCGCGAAATCAATGCGGGCGGGCATCATTCCTC
           1350      1360      1370      1380      1390      1400

              2250      2260
WT_8_final    TACGGCGGCATCAGCCGCGGNN-----GC
             :  ::::    : :: ::::       .
MS11_DEL_8_F  TN-GGCGCTCGCCGC-GCGGAAAAANNVT
           1410       1420       1430


                                                                    10
WT_9_FINAL_1  ----------------------------------------TCCCTTAGTCTTCTAAAAT
                                                      ::::::::::::::::::
MS11_DEL_9_F  CGGGGGAGAGGGCTCCCCGAATTTCGACAATCTTTCCCAAATCCCTTAGTCTTCTAAAAT
                      10        20        30        40        50        60

             20        30        40        50        60        70
WT_9_FINAL_1  ACAAGCCTTGCAGCTTTCTAGCTCTCTCCCACAGGGAGAGAGGACTATGATGCCCGCCGG
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  ACAAGCCTTGCAGCTTTCTAGCTCTCTCCCACAGGGAGAGAGGACTATGATGCCCGCCGG
                      70        80        90       100       110       120

             80        90       100       110       120       130
WT_9_FINAL_1  CGGCATCGCACCTTATCCGCGCGCTCCGGAACAGGGCTGTATGGCAGATGCGTAGGATGG
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  CGGCATCGCACCTTATCCGCGCGCTCCGGAACAGGGCTGTATGGCAGATGCGTAGGATGG
                      130       140       150       160       170       180

            140       150       160       170       180       190
WT_9_FINAL_1  GCGTAGGTGGGTTGTGGCGGGTTGCTATTAGGAATACCTGAATCGTCATTCCCACGAAAG
             :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCGTAGGTGGGTTGTGGCGGGTCGCTATTAGGAATACCTGAATCGTCATTCCCACGAAAG
                      190       200       210       220       230       240

            200       210       220       230       240       250
WT_9_FINAL_1  TGGGAATCTAGAACGCAGGGTTGGGGCAACCGTTTTATCCGATAAGTTCCTATGCCGACA
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  TGGGAATCTAGAACGCAGGGTTGGGGCAACCGTTTTATCCGATAAGTTCCTATGCCGACA
                      250       260       270       280       290       300

            260       270       280       290       300       310
WT_9_FINAL_1  GACCTGAATTCCCGCCTGTGCGGGAATGACGGCGCATAAGTTTCCGCGCGGACAAATCCG
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GACCTGAATTCCCGCCTGTGCGGGAATGACGGCGCATAAGTTTCCGCGCGGACAAATCCG
```

```
                    310       320       330       340       350       360

            320       330       340       350       360       370
WT_9_FINAL_1  GATTCCTGTCTGTGTGGGAATGACGAATCCGTCCGCACGGAAACCTGCACCGCGTCATTC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GATTCCTGTCTGTGTGGGAATGACGAATCCGTCCGCACGGAAACCTGCACCGCGTCATTC
                    370       380       390       400       410       420

            380       390       400       410       420       430
WT_9_FINAL_1  CCCCGAAAGTGGGAATCTAGAAACTTAACGCTACGGCAATTTTTGGAAATGACTGAAACC
              : ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  CTCCGAAAGTGGGAATCTAGAAACTTAACGCTACGGCAATTTTTGGAAATGACTGAAACC
                    430       440       450       460       470       480

            440       450       460       470       480       490
WT_9_FINAL_1  GAACGGACTGGATTCCCGCCTGCGCGGGAATGGCGGGTTTTAGGATTACGGTGTATCGGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GAACGGACTGGATTCCCGCCTGCGCGGGAATGGCGGGTTTTAGGATTACGGTGTATCGGG
                    490       500       510       520       530       540

            500       510       520       530       540       550
WT_9_FINAL_1  ACGGTTCGGGTATTCTTGACAGGATGGGTTCTCAAGATGTTGCGTAGTGTCGGGCTCGCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  ACGGTTCGGGTATTCTTGACAGGATGGGTTCTCAAGATGTTGCGTAGTGTCGGGCTCGCA
                    550       560       570       580       590       600

            560       570       580       590       600       610
WT_9_FINAL_1  AGGGGGTGCATAGAGAGGGTTGCAGGGCGGGCTTCAGCCCGCCGCATCAAGGTTTTTGGG
              :::::::::::::::::::::::::::::::::::: ::::::::::::::::::::
MS11_DEL_9_F  AGGGGGTGCATAGAGAGGGTTGCAGGGCGGGCTTCAGCCCGACGCATCAAGGTTTTTGGG
                    610       620       630       640       650       660

            620       630       640       650       660       670
WT_9_FINAL_1  GAATAGGCGTGATTCGGCGGAATTGATGGATTTGACGGAATCGGCGGCAATGGCGGGATC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GAATAGGCGTGATTCGGCGGAATTGATGGATTTGACGGAATCGGCGGCAATGGCGGGATC
                    670       680       690       700       710       720

            680       690       700       710       720       730
WT_9_FINAL_1  GGTAATATTGAAGGGATTGGTGGAGGCCGGCGGCAACGGCGGGGGCGGAAAAGACGCGCC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGTAATATTGAAGGGATTGGTGGAGGCCGGCGGCAACGGCGGGGGCGGAAAAGACGCGCC
                    730       740       750       760       770       780

            740       750       760       770       780       790
WT_9_FINAL_1  AGCCGCCGAAATCGTAGCCGAGGGCGAGTCGGGGGTGGACGGAGCGGGTGCGGATGTTTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  AGCCGCCGAAATCGTAGCCGAGGGCGAGTCGGGGGTGGACGGAGCGGGTGCGGATGTTTT
                    790       800       810       820       830       840

            800       810       820       830       840       850
WT_9_FINAL_1  TGAAATAATCGCTTACCGTGCTTATTTTGCCTTGGTTTGCACCGGTTGCTTCGGGATAAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  TGAAATAATCGCTTACCGTGCTTATTTTGCCTTGGTTTGCACCGGTTGCTTCGGGATAAT
                    850       860       870       880       890       900

            860       870       880       890       900       910
```

```
WT_9_FINAL_1  CGCGGGTGATGTGCTCGTAGGCGTAAGCCAGATCCGCCTGCACATACGGGCCGCGCCCAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  CGCGGGTGATGTGCTCGTAGGCGTAAGCCAGATCCGCCTGCACATACGGGCCGCGCCCAT
                     910       920       930       940       950       960

            920       930       940       950       960       970
WT_9_FINAL_1  GGTCTTCACCCGCCGCCTGCCAAGCATCAGGGCGAGCAGGAGGGGTTTGTATTGCATGGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGTCTTCACCCGCCGCCTGCCAAGCATCAGGGCGAGCAGGAGGGGTTTGTATTGCATGGT
                     970       980       990      1000      1010      1020

            980       990      1000      1010      1020      1030
WT_9_FINAL_1  TCGGCTTTCGGAAAAAATCGGATAATGCTGAAGGCCGCGCGAAAGCGGCCGGATGTTCGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  TCGGCTTTCGGAAAAAATCGGATAATGCTGAAGGCCGCGCGAAAGCGGCCGGATGTTCGG
                    1030      1040      1050      1060      1070      1080

           1040      1050      1060      1070      1080      1090
WT_9_FINAL_1  ATTATACTGTCAGTTATGCCGTCTGAAAATGCCGTTTGCCCGATCTTGCGCCTTAAAAAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  ATTATACTGTCAGTTATGCCGTCTGAAAATGCCGTTTGCCCGATCTTGCGCCTTAAAAAT
                    1090      1100      1110      1120      1130      1140

           1100      1110      1120      1130      1140      1150
WT_9_FINAL_1  GCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAGCGGTAGCGCACGCCCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAGCGGTAGCGCACGCCCA
                    1150      1160      1170      1180      1190      1200

           1160      1170      1180      1190      1200      1210
WT_9_FINAL_1  ATGAGGCTTCGTGGGTTTTGAAGCGGGTGTTTTCCAGGCGTCCCCAATAGTGGTAGCGGT
              ::::::::::::::::
MS11_DEL_9_F  ATGAGGCTTCGTGGGT----------------------------------------------
                    1210

           1220      1230      1240      1250      1260      1270
WT_9_FINAL_1  ACCCGGCGTCCAAGGTCAGGCCGGGCGCGACGTCTATGCCCACGCCCGCCATCGCGCCGA

MS11_DEL_9_F  ----------------------------------------------------------------


           1280      1290      1300      1310      1320      1330
WT_9_FINAL_1  AGCCCAAGCGGCGGCTGCTGCGGTTTTGGCGATAAGTGTTTTTTTGGTTTTTTCCCGAAT

MS11_DEL_9_F  ----------------------------------------------------------------


           1340      1350      1360      1370      1380      1390
WT_9_FINAL_1  CTATATCATCATAATACGTAGGTTTTGTGCCAGCACTATGGTAGGCGGTAAGAGTATTTT

MS11_DEL_9_F  ----------------------------------------------------------------


           1400      1410      1420      1430      1440      1450
WT_9_FINAL_1  TCGTTTTTTTAGTCGAATCGATACCGTGTCTGACGTGTCCGTAGCCGACACGCACACCGA

MS11_DEL_9_F  ----------------------------------------------------------------
```

```
           1460      1470      1480      1490      1500      1510
WT_9_FINAL_1  TATAGGGTTTGAATTTATCGTTGACTCTGAAATCGTAAACGGTTGACAAGCCGAGAGAAG

MS11_DEL_9_F  ------------------------------------------------------------


           1520      1530      1540      1550      1560      1570
WT_9_FINAL_1  AAACGGCGTGGAATGTGCCGTTTTCCTGATGTTCCGTCTTTTGGTATTTTATGTTAAGCT

MS11_DEL_9_F  ------------------------------------------------------------


           1580      1590      1600      1610      1620      1630
WT_9_FINAL_1  GGTTGCCGCCAAAAGTTTTATTATTCTTTCTTTCCAACTCTTTTATGTTCACGGAATATT

MS11_DEL_9_F  ------------------------------------------------------------


           1640      1650      1660      1670      1680      1690
WT_9_FINAL_1  TATTGTTGTGCCATTTCCTGTAACGGGCATAATCCGCGGCGATGCGCCAGCCGCCGAAGT

MS11_DEL_9_F  ------------------------------------------------------------


           1700      1710      1720      1730      1740      1750
WT_9_FINAL_1  CGTAGCCGACCGACACCCTGGGGTGGATGGAATGCGTACGGATGTTTCTGAAATAATCGC

MS11_DEL_9_F  ------------------------------------------------------------


           1760      1770      1780      1790      1800      1810
WT_9_FINAL_1  TTACCGTGCTTATTTTGCCTTTGTTTGCACCGGCTGCATCGGGATAATCGCGGGTAATGT

MS11_DEL_9_F  ------------------------------------------------------------


           1820      1830      1840      1850      1860      1870
WT_9_FINAL_1  GCTCGTAGGCGTAAGCCAGATCCGCCTGCACATACGGGCCGCGCCCATGGTCTTCACCCG

MS11_DEL_9_F  ------------------------------------------------------------


           1880      1890      1900      1910      1920      1930
WT_9_FINAL_1  CCGCCCGCGCTGCGGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAG

MS11_DEL_9_F  ------------------------------------------------------------


           1940      1950      1960      1970      1980      1990
WT_9_FINAL_1  AAGAGAAGAGAAGAGAAGTTTTTTGTCGACCGTATGCATATCAACTCCGAATATTTAAGA

MS11_DEL_9_F  ------------------------------------------------------------


           2000      2010      2020      2030      2040      2050
```

```
WT_9_FINAL_1  TTAAATTATTAAAAGATTGTTAATAAACGCCAAGCATTTTATCAGATTTGTTTTGCCGGG
                                                      ::::::::::::::::::
MS11_DEL_9_F  ----------------------------------------CAGATTTGTTTTGCCGGG
                                                      1220      1230

           2060      2070      2080      2090      2100      2110
WT_9_FINAL_1  ATATGATTTTTTTTGTTAAAGAATGTATTTATTTTTTAAAATTCCGGTGCGGCGATCGGA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  ATATGATTTTTTTTGTTAAAGAATGTATTTATTTTTTAAAATTCCGGTGCGGCGATCGGA
                1240      1250      1260      1270      1280      1290

           2120      2130      2140      2150      2160      2170
WT_9_FINAL_1  TATGGCGGATTAACAAAAATCAGGACAAGGCGGCGAAGCCGCAGACAGTACAAATAGTAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  TATGGCGGATTAACAAAAATCAGGACAAGGCGGCGAAGCCGCAGACAGTACAAATAGTAC
                1300      1310      1320      1330      1340      1350

           2180      2190      2200      2210      2220      2230
WT_9_FINAL_1  GGAACCGATTCACTTGGTGCTTCAGCACCTTAGAGAATCGTTCTCTTTGAGCTAAGGCGA
              ::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGAACCGATCCACTTGGTGCTTCAGCACCTTAGAGAATCGTTCTCTTTGAGCTAAGGCGA
                1360      1370      1380      1390      1400      1410

           2240      2250      2260      2270      2280      2290
WT_9_FINAL_1  GGCAACGCCGTACTGGTTTAAAGTTAATCCACTATAGTGGATTAAAATAGGAAATATCCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGCAACGCCGTACTGGTTTAAAGTTAATCCACTATAGTGGATTAAAATAGGAAATATCCA
                1420      1430      1440      1450      1460      1470

           2300      2310      2320      2330      2340      2350
WT_9_FINAL_1  ATGCCAAAGCAATTTGTCGGAAATGGCCGGAACTCAAAAACCGGATTCCCACTCCTTCGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  ATGCCAAAGCAATTTGTCGGAAATGGCCGGAACTCAAAAACCGGATTCCCACTCCTTCGT
                1480      1490      1500      1510      1520      1530

           2360      2370      2380      2390      2400      2410
WT_9_FINAL_1  CATTCCCGCGAAAGTGGGAATCTAGGAATGAAAAGCAGCAGGAATTTATCGGAAATAACC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  CATTCCCGCGAAAGTGGGAATCTAGGAATGAAAAGCAGCAGGAATTTATCGGAAATAACC
                1540      1550      1560      1570      1580      1590

           2420      2430      2440      2450      2460      2470
WT_9_FINAL_1  GAAACCGAACGGTCCGGATTCCCGCCTTCGCGGGAATGGCGGCGCATAAGTTTCCGCGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GAAACCGAACGGTCCGGATTCCCGCCTTCGCGGGAATGGCGGCGCATAAGTTTCCGCGCG
                1600      1610      1620      1630      1640      1650

           2480      2490      2500      2510      2520      2530
WT_9_FINAL_1  GACAAATCCGGATTTCTGTCTGCGCGGGAATGACGGGTTTCAGGATTACGGTGTATCGGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GACAAATCCGGATTTCTGTCTGCGCGGGAATGACGGGTTTCAGGATTACGGTGTATCGGG
                1660      1670      1680      1690      1700      1710

           2540      2550      2560      2570      2580      2590
WT_9_FINAL_1  AATGATGACACGGGTATTCCTGACGATTCGGGTATTTCTGACAGGATGGATTCTCATCTA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  AATGATGACACGGGTATTCCTGACGATTCGGGTATTTCTGACAGGATGGATTCTCATCTA
```

```
              1720      1730      1740      1750      1760      1770

          2600      2610      2620       2630      2640      2650
WT_9_FINAL_1  GATTCCTGCCTGCGCGGGAATGGGGGG-TTTCAGGATTACGGTAAATGGCGCAAAA-TGC
              :::::::::::::: :::::::  ::: :::::::::::::::::: :::::::: :::
MS11_DEL_9_F  GATTCCTGCCTGCGTGGGAATGACGGGATTTCAGGATTACGGTAAATAGCGCAAAAATGC
              1780      1790      1800      1810      1820      1830

             2660      2670      2680      2690      2700      2710
WT_9_FINAL_1  CGTCTGAAAGCCCTTCAGACGGCATTGCCTTGTTCGTCTGCCTTAATGGCGGAAGTGGCG
              ::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  CGTCTGAAAACCCTTCAGACGGCATTGCCTTGTTCGTCTGCCTTAATGGCGGAAGTGGCG
              1840      1850      1860      1870      1880      1890

             2720      2730      2740      2750      2760      2770
WT_9_FINAL_1  GATGCCGGTTACCGCCATGGCGATGCCGTGTTCGTCCGCCGCGTCGAAAACTTCCTGATC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GATGCCGGTTACCGCCATGGCGATGCCGTGTTCGTCCGCCGCGTCGAAAACTTCCTGATC
              1900      1910      1920      1930      1940      1950

             2780      2790      2800      2810      2820      2830
WT_9_FINAL_1  GCGCATCGAGCCTGCCGGATGGATGATGGCTTTGATGCCCTGTTCGGCAATCACGTCCAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCGCATCGAGCCTGCCGGATGGATGATGGCTTTGATGCCCTGTTCGGCAATCACGTCCAC
              1960      1970      1980      1990      2000      2010

             2840      2850      2860      2870      2880      2890
WT_9_FINAL_1  GCCGTCGCGGAATGGGAAGAAGGCATCGGAAGCGGCACATGCGCCGTTGAGGTCGAGACC
              :::::::::::::::::::::::::::::::::::::::::::::::::: :::::::
MS11_DEL_9_F  GCCGTCGCGGAATGGGAAGAAGGCATCGGAAGCGGCACATGCGCCGTTGAGATCGAGACT
              2020      2030      2040      2050      2060      2070

             2900      2910      2920      2930      2940      2950
WT_9_FINAL_1  GGCATCTTGCGCTTTGCGGGCGGCGATGCGGGTGCTGTCCACGCGGCTCATTTGGCCTGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGCATCTTGCGCTTTGCGGGCGGCGATGCGGGTGCTGTCCACGCGGCTCATTTGGCCTGC
              2080      2090      2100      2110      2120      2130

             2960      2970      2980      2990      3000      3010
WT_9_FINAL_1  GCCGATACCGTAGGTTTGGCCGCCTTTGCCGAAGACGATGGCGTTGGATTTGACGTATTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCCGATACCGTAGGTTTGGCCGCCTTTGCCGAAGACGATGGCGTTGGATTTGACGTATTT
              2140      2150      2160      2170      2180      2190

             3020      3030      3040      3050      3060      3070
WT_9_FINAL_1  TGCGACGTTCCAGACAAACAGCAAATCGTTCCATTCCTGCTCGGTCGGTTGGCGTTTGGA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  TGCGACGTTCCAGACAAACAGCAAATCGTTCCATTCCTGCTCGGTCGGTTGGCGTTTGGA
              2200      2210      2220      2230      2240      2250

             3080      3090      3100      3110      3120      3130
WT_9_FINAL_1  GACGACTTTCAAATCGGCGCGGTTGATGCGGTTGATGTCGGGCGTTTGTACCAACAGTCC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GACGACTTTCAAATCGGCGCGGTTGATGCGGTTGATGTCGGGCGTTTGTACCAACAGTCC
              2260      2270      2280      2290      2300      2310

             3140      3150      3160      3170      3180      3190
```

```
WT_9_FINAL_1  GCCGCCGACGCGTTTGAGTTCAAAGCGGTTGGCACCTGCTTTGAGCGGCACTTCCAATAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCCGCCGACGCGTTTGAGTTCAAAGCGGTTGGCACCTGCTTTGAGCGGCACTTCCAATAC
                2320      2330      2340      2350      2360      2370

             3200      3210      3220      3230      3240      3250
WT_9_FINAL_1  GCGCACGTTTTTCTTGGCGGCGGCGATTTCAAGGGCTTCGGCGGTAAACTTCGGCGCCAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCGCACGTTTTTCTTGGCGGCGGCGATTTCAAGGGCTTCGGCGGTAAACTTCGGCGCCAT
                2380      2390      2400      2410      2420      2430

             3260      3270      3280      3290      3300      3310
WT_9_FINAL_1  GAGGACTTCCATAAATTGGTTGTCGGTAATCTGTTTGACGGTTTCGCCGTCGACTTCGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GAGGACTTCCATAAATTGGTTGTCGGTAATCTGTTTGACGGTTTCGCCGTCGACTTCGCG
                2440      2450      2460      2470      2480      2490

             3320      3330      3340      3350      3360      3370
WT_9_FINAL_1  GTTGAAGGCGATGATGCCGCCGAACGCGCTGGTGGTGTCGGTGGCGTAGGCGAGTTTGTA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GTTGAAGGCGATGATGCCGCCGAACGCGCTGGTGGTGTCGGTGGCGTAGGCGAGTTTGTA
                2500      2510      2520      2530      2540      2550

             3380      3390      3400      3410      3420      3430
WT_9_FINAL_1  GGCGGTCAAGGTATCGGCTGCAACGGCTACGCCGCACGGATTGGCGTGTTTCACGATCAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GGCGGTCAAGGTATCGGCTGCAACGGCTACGCCGCACGGATTGGCGTGTTTCACGATCAC
                2560      2570      2580      2590      2600      2610

             3440      3450      3460      3470      3480      3490
WT_9_FINAL_1  GCAGGCGGGCGCGTCAAAGGATTTGACGGCTTCCCAAGCGGCATCGGCATCGGCGATGTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GCAGGCGGGCGCGTCAAAGGATTTGACGGCTTCCCAAGCGGCATCGGCATCGGCGATGTT
                2620      2630      2640      2650      2660      2670

             3500      3510      3520      3530      3540      3550
WT_9_FINAL_1  GTTGTAAGACAATTCTTTGCCTTGCAGTTGATTGTAGGCGGAAAGGCTGCCTGCGGCGGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_9_F  GTTGTAAGACAATTCTTTGCCTTGCAGTTGATTGTAGGCGGAAAGGCTGCCTGCGGCGGG
                2680      2690      2700      2710      2720      2730

             3560      3570      3580      3590      3600      3610
WT_9_FINAL_1  GTCAATATCGCGGTAAAACGCGGCGCGCTGATGCGGGTTTCGCCGAGCGCAGTCTG-ACT
              :::::::::::::::::::::::::::::::::::::::::::::::::::::::: :::
MS11_DEL_9_F  GTCAATATCGCGGTAAAACGCGGCGCGCTGATGCGGGTTTCGCCGAGCGCAGTCTGCACT
                2740      2750      2760      2770      2780      2790

              3620
WT_9_FINAL_1  TTTTCGG-------
               :    :
MS11_DEL_9_F  CTCAAAGCGCCNVT
                2800


                        10              20         30        40        50
WT_10_final   ABNNNAAC--TGTCACCG----TTA--CAGCTTTTC-CGCACGCTGTCGGGGGTTGGGCG
               .   ::   :::: : :    :::  ::::::::: ::::::::::::::::::::::
```

```
MS11_DEL_10_  CCGTCAATCTTGTCCCAGCCAGTTAAACAGCTTTTCACGCACGCTGTCGGGGGTTGGGCG
                      10        20        30        40        50        60

                     60        70        80        90       100       110
WT_10_final   CAGTCCGTCGGCGGATGCAAAACTCAATTTCCTGCCCCGGCATTGCCCGCCGATAATGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CAGTCCGTCGGCGGATGCAAAACTCAATTTCCTGCCCCGGCATTGCCCGCCGATAATGCG
                      70        80        90       100       110       120

                    120       130       140       150       160       170
WT_10_final   TACCCGGTTGCTGTGTTTGGTATGTTTGCCTGCCGCCATAATCAAAAGCCTGAAAGTTCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  TACCCGGTTGCTGTGTTTGGTATGTTTGCCTGCCGCCATAATCAAAAGCCTGAAAGTTCA
                     130       140       150       160       170       180

                    180       190       200       210       220       230
WT_10_final   AACGGTATTATACAAGACCTGTCGAAGAATATGCCGCCTGAAAACTTTTTTCAGACGGCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  AACGGTATTATACAAGACCTGTCGAAGAATATGCCGCCTGAAAACTTTTTTCAGACGGCA
                     190       200       210       220       230       240

                    240       250       260       270       280       290
WT_10_final   TATCTGTTTAAACGGTTTCGGTCAGCTTGCGGAGCAACTCGATTTCTTTGTCTTTTTGTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  TATCTGTTTAAACGGTTTCGGTCAGCTTGCGGAGCAACTCGATTTCTTTGTCTTTTTGTT
                     250       260       270       280       290       300

                    300       310       320       330       340       350
WT_10_final   CCAACATTTCTTTACAGTGTTTTAATTCCATTTTTAACAAATCCAGCTTTACCGATGCAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CCAACATTTCTTTACAGTGTTTTAATTCCATTTTTAACAAATCCAGCTTTACCGATGCAT
                     310       320       330       340       350       360

                    360       370       380       390       400       410
WT_10_final   CCTGAGCGGTATAGATTGCAAATTCCCCGCTGTTGGTATCCACATCGTTAATCTGCAACA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CCTGAGCGGTATAGATTGCAAATTCCCCGCTGTTGGTATCCACATCGTTAATCTGCAACA
                     370       380       390       400       410       420

                    420       430       440       450       460       470
WT_10_final   CCATCCCTCCGCCACCCGATTTGAGCGAATCCCGACAAATCCGGATTCCCGCCTGCGCGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CCATCCCTCCGCCACCCGATTTGAGCGAATCCCGACAAATCCGGATTCCCGCCTGCGCGG
                     430       440       450       460       470       480

                    480       490       500       510       520       530
WT_10_final   GAATGACGGGTTTCGAGATTGCGGTATTGTCGGGAATGACGGTTCGGGTATTTTACTGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  GAATGACGGGTTTCGAGATTGCGGTATTGTCGGGAATGACGGTTCGGGTATTTTACTGCG
                     490       500       510       520       530       540

                    540       550       560       570       580       590
WT_10_final   CCCGCCCCGCGCCTGTAAACGGCGGGCGCATCAAAAATGCCGTCTGAAGGTTCAGACGGC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CCCGCCCCGCGCCTGTAAACGGCGGGCGCATCAAAAATGCCGTCTGAAGGTTCAGACGGC
                     550       560       570       580       590       600
```

```
                      600       610       620       630       640       650
WT_10_final   ATCGGTATCGGGGAATCAGAAGCGGTAGCGCACGCCCAACGAGGCTTCGTGGGTTTTGAA
              :::::::::::::::::::::::::::::::::::::::: ::::::::::::::
MS11_DEL_10_  ATCGGTATCGGGGAATCAGAAGCGGTAGCGCACGCCCAATGAGGCTTCGTGGGT------
                       610       620       630       640       650

                      660       670       680       690       700       710
WT_10_final   GCGGGTGTTTTCCAAGCGTCCCCAATAGTGGTAGCGGTAGCCGGTGTCCAGGGTCAGCTT

MS11_DEL_10_  ------------------------------------------------------------


                      720       730       740       750       760       770
WT_10_final   GGGCGTGATGTCGAAACCGACGCCGGCGATGACGCCGAGGCCCACGCGGCGGATGCTGTT

MS11_DEL_10_  ------------------------------------------------------------


                      780       790       800       810       820       830
WT_10_final   ACTTTGGTGATGGGCGTTTTGCGTATTTTCCTCAGTATAAACCGTAGGGTTTAAGCCACC

MS11_DEL_10_  ------------------------------------------------------------


                      840       850       860       870       880       890
WT_10_final   ATAGGAGGAGGTAAGAAACTTTGTTGTTTTTTTGGTCGAATCGATGCTGTGTCTGACGTG

MS11_DEL_10_  ------------------------------------------------------------


                      900       910       920       930       940       950
WT_10_final   TCCGTAGGCGACGCGCGCACCGATATAGGGTTTGAATTTATCGTTGAGTTTGAAATCGTA

MS11_DEL_10_  ------------------------------------------------------------


                      960       970       980       990      1000      1010
WT_10_final   AACGGCTGACAAACCGAGCGAGGAGACGGCGTGGAACGTACCGTTTTCCTGATTTTCCGT

MS11_DEL_10_  ------------------------------------------------------------


                     1020      1030      1040      1050      1060      1070
WT_10_final   CTTCAGGTCTCTCTTATTCTGATTCTTGTTTTCCAACTCTTTTATGTCGACGGAATATTT

MS11_DEL_10_  ------------------------------------------------------------


                     1080      1090      1100      1110      1120      1130
WT_10_final   ATTGTCGTTCCATTTCCTGTAACGGGCATAATCTGCCGCTATCCTCCAACCGCCGAAGTC

MS11_DEL_10_  ------------------------------------------------------------


                     1140      1150      1160      1170      1180      1190
WT_10_final   GTAGCCGACCGACACCCTGGGGTGGACGGAATGCGTACGGATGTTTCTGAAATAATCGCT
```

```
MS11_DEL_10_  ------------------------------------------------------------

                   1200      1210      1220      1230      1240      1250
WT_10_final   TACCGTGCTTATTTTTTTGTCTTTTTTTGCACCGGTTGGTTCCGGATAATCGTGGGTAAT

MS11_DEL_10_  ------------------------------------------------------------

                   1260      1270      1280      1290      1300      1310
WT_10_final   GCGTTCGGCGGCGTAGGCTAAATCCGCCTGCACATACGGGCCGCGCCCATTGCCTTCACT

MS11_DEL_10_  ------------------------------------------------------------

                   1320      1330      1340      1350      1360      1370
WT_10_final   TGCCGCCCGCGCTGCGGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAAG

MS11_DEL_10_  ------------------------------------------------------------

                   1380      1390      1400      1410      1420      1430
WT_10_final   AGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAGAAGAAG

MS11_DEL_10_  ------------------------------------------------------------

                   1440      1450      1460      1470      1480      1490
WT_10_final   AGAAGAGAAGAGAAGAGAAGTTTTTTGTCGGCCGTATGCATATCAACTCCAAATATCTTA

MS11_DEL_10_  ------------------------------------------------------------

                   1500      1510      1520      1530      1540      1550
WT_10_final   GATTAAATTATTAAAAAGATTGTTAATAAACGCCTAGCATTTTATCAGATTTGTTTTGCC
                                                           ::::::::::::
MS11_DEL_10_  -----------------------------------------------ATTTGTTTTGCC
                                                              660

                   1560      1570      1580      1590       1600      1610
WT_10_final   GGGATATGATTTGTTTTGTTAAAGAATGTATTTATTTTTT-AAAATCCCGGCCAACCGGG
              :::::::::::::: :::::::::::::::::::::::: ::::::: ::::::::::
MS11_DEL_10_  GGGATATGATTTGTTCTGTTAAAGAATGTATTTATTTTTTTAAAATCCGGGCCAACCGGG
               670       680       690       700       710       720

                    1620      1630       1640       1650       1660
WT_10_final   ATAAATCCTGCTTTACCAATTGTTTTAAA-TGGAA-TTCGAACTTGTACCC-ACTGTTGT
              :: :::::::::::::::::::::::::: ::::: ::::::::: ::::: ::::::::
MS11_DEL_10_  ATGAATCCTGCTTTACCAATTGTTTTAAAATGGAAATTCGAACTTTTACCCCACTGTTGT
               730       740       750       760       770       780

             1670       1680      1690         1700      1710      1720
WT_10_final   CAAA-CGCCGTCCGCACTCCTTCA-ATACAGCCG--AATGCTCTTTGGGAATGCCGTCAA
              :::: ::::: :::::::::::::: :::: :::   :::: ::::::::::::::::::
MS11_DEL_10_  CAAAACGCCGACCGCACTCCTTCANATACCGCCCGAAATGATCTTTGGGAATGCCGTCAA
               790       800       810       820       830       840
```

```
                1730      1740      1750      1760      1770      1780
WT_10_final     ACTTGCGCAAATGACGTTTGCCCGATTCCAAAAGTTCCCAATTGCGTTGATATGGTTTTG
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    ACTTGCGCAAATGACGTTTGCCCGATTCCAAAAGTTCCCAATTGCGTTGATATGGTTTTG
                 850       860       870       880       890       900

                1790      1800      1810      1820      1830      1840
WT_10_final     CCGTTCGGCAAAATGTGTGCCGCGATTGATACGGAAACGGCCAAATCCGCTTACAATCTG
                :::::::::::::::::::::::::::::::::::::::::: :::::::::::: ::
MS11_DEL_10_    CCGTTCGGCAAAATGTGTGCCGCGATTGATACGGAAACGGCCAACTCCGCTTACAATNTG
                 910       920       930       940       950       960

                1850      1860      1870      1880      1890      1900
WT_10_final     ATACATCGTAACTACGAACAATCCGTATAAACAATGCTGCCGGGTTTCACCCGTCCGCGG
                ::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    ATACATCGTAACTACGANCAATCCGTATAAACAATGCTGCCGGGTTTCACCCGTCCGCGG
                 970       980       990      1000      1010      1020

                1910      1920      1930      1940      1950      1960
WT_10_final     ATAATAGGCAATAAAATAGCGGTTCGAGTGTCCGACAGTAACCGTACAAACCTTGCCGTT
                ::::::::::::::::::::::: :::::::::::::::::::::::::::: ::::::
MS11_DEL_10_    ATAATAGGCAATAAAATAGCGGTTNGAGTGTCCGACAGTAACCGTACAAACCTNGCCGTT
                1030      1040      1050      1060      1070      1080

                1970      1980      1990      2000      2010      2020
WT_10_final     GCGCTTCAAAAGCCCGAATACGGCGGCTTTGCCGGCGGCACTGCGGCCGCGTCTGCCTTT
                ::::::::::::::::: :: :::::::::: :::::::::::::::::::::::::::
MS11_DEL_10_    GCGCTTCAAAAGCCCGACTANGGCGGCTTTGCNGGCGGCACTGCGGCCGCGTCTGCCTTT
                1090      1100      1110      1120      1130      1140

                2030      2040      2050      2060      2070      2080
WT_10_final     ACGGAAATAACTTTCACCCGCCCCTGCTTCGCCGTCATACATTTCCAAATGCGGACTGTT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    ACGGAAATAACTTTCACCCGCCCCTGCTTCGCCGTCATACATTTCCAAATGCGGACTGTT
                1150      1160      1170      1180      1190      1200

                2090      2100      2110      2120      2130      2140
WT_10_final     TTGACGGATAAGTAATCGTAAACGATATGTCGTATGCCTCCATAATTTCGCAGATTTCGT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    TTGACGGATAAGTAATCGTAAACGATATGTCGTATGCCTCCATAATTTCGCAGATTTCGT
                1210      1220      1230      1240      1250      1260

                2150      2160      2170      2180      2190      2200
WT_10_final     CCAAATGCGTGTAAAGGAAGTTTTCCCGATGATGTGCCAAACACTATTTCGCCATAGAAA
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    CCAAATGCGTGTAAAGGAAGTTTTCCCGATGATGTGCCAAACACTATTTCGCCATAGAAA
                1270      1280      1290      1300      1310      1320

                2210      2220      2230      2240      2250      2260
WT_10_final     ATAACAGGCTGCCGTGTTTTGATTAACTGCTGCCAACCCTGCTGCCGTTCTTGCAGTTAC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    ATAACAGGCTGCCGTGTTTTGATTAACTGCTGCCAACCCTGCTGCCGTTCTTGCAGTTAC
                1330      1340      1350      1360      1370      1380

                2270      2280      2290      2300      2310      2320
WT_10_final     CCCTGCGGCAAACAGTCCGATGGGTTTCTTTTGTTTGTACCGGCTTGGACGGCTTTTCTC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
MS11_DEL_10_  CCCTGCGGCAAACAGTCCGATGGGTTTCTTTTGTTTGTACCGGCTTGGACGGCTTTTCTC
              1390      1400      1410      1420      1430      1440

                 2330      2340      2350      2360      2370      2380
WT_10_final   ATAGGGATAATTCTGACTTAATTTAAATTTCCCCAGCCATCTGGGACAGCCCCTTCTAAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  ATAGGGATAATTCTGACTTAATTTAAATTTCCCCAGCCATCTGGGACAGCCCCTTCTAAT
              1450      1460      1470      1480      1490      1500

                 2390      2400      2410      2420      2430      2440
WT_10_final   TTACAAATGCCGTCTGAAACCCCTTTCAGACGGCATTTTCTTTCATCTCAATCCCAACCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  TTACAAATGCCGTCTGAAACCCCTTTCAGACGGCATTTTCTTTCATCTCAATCCCAACCG
              1510      1520      1530      1540      1550      1560

                 2450      2460      2470      2480      2490      2500
WT_10_final   CTTCACTCTTGCGGTGCAGGCGGTTCAATCTGTTCTTTCCAGCCGCATTCTTTTTGCGGA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CTTCACTCTTGCGGTGCAGGCGGTTCAATCTGTTCTTTCCAGCCGCATTCTTTTTGCGGA
              1570      1580      1590      1600      1610      1620

                 2510      2520      2530      2540      2550      2560
WT_10_final   CAGACTTTTTCCACGCCCCGGCGTTTAGTGGTTTTGATGGTCAAGACCGGCCAATGGCAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CAGACTTTTTCCACGCCCCGGCGTTTAGTGGTTTTGATGGTCAAGACCGGCCAATGGCAG
              1630      1640      1650      1660      1670      1680

                 2570      2580      2590      2600      2610      2620
WT_10_final   TTCGGGCATTCTTCGGCAACGGGCGGGTTCCAAGTGGCGTAGTTGCAGTCGGGATAGGTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  TTCGGGCATTCTTCGGCAACGGGCGGGTTCCAAGTGGCGTAGTTGCAGTCGGGATAGGTG
              1690      1700      1710      1720      1730      1740

                 2630      2640      2650      2660      2670      2680
WT_10_final   CTGCAACTGTAAAACAGTTTGCCGTAGCGGGATTTGCGCTCGACGAGGTTGCCTTTTTTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CTGCAACTGTAAAACAGTTTGCCGTAGCGGGATTTGCGCTCGACGAGGTTGCCTTTTTTG
              1750      1760      1770      1780      1790      1800

                 2690      2700      2710      2720      2730      2740
WT_10_final   CATTGCGGGCATTGGACGCCGGTATCTTTGGGTTTTTCCAACGGCTCGACGTGTTTGCAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CATTGCGGGCATTGGACGCCGGTATCTTTGGGTTTTTCCAACGGCTCGACGTGTTTGCAT
              1810      1820      1830      1840      1850      1860

                 2750      2760      2770      2780      2790      2800
WT_10_final   TTGGGGTAGTTGGCGCAACCGATGAATTTGCTGCCGGTACGGCTGTATTTATACGCTAAC
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  TTGGGGTAGTTGGCGCAACCGATGAATTTGCTGCCGGTACGGCTGTATTTATACGCTAAC
              1870      1880      1890      1900      1910      1920

                 2810      2820      2830      2840      2850      2860
WT_10_final   CGACCGCCGCATTTGGGGCATTCGCGTCCGTCGAGTTCGGCTTGTTCAGCCTCGGCTTTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_  CGACCGCCGCATTTGGGGCATTCGCGTCCGTCGAGTTCGGCTTGTTCAGCCTCGGCTTTG
              1930      1940      1950      1960      1970      1980
```

```
                2870      2880      2890      2900      2910      2920
WT_10_final     GCGATGCGCTCGGCGGCTTCTTCGGCGGTTTCGTTGACGTTGCGCGTGTAGCTGCATTCG
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    GCGATGCGCTCGGCGGCTTCTTCGGCGGTTTCGTTGACGTTGCGCGTGTAGCTGCATTCG
                1990      2000      2010      2020      2030      2040

                2930      2940      2950      2960      2970      2980
WT_10_final     GGATAGCCGGCGCAGGCGACGAAGCGGCCCATTTTGCCGAACTTGATTTGCAGTTTGTGT
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    GGATAGCCGGCGCAGGCGACGAAGCGGCCCATTTTGCCGAACTTGATTTGCAGTTTGTGT
                2050      2060      2070      2080      2090      2100

                2990      3000      3010      3020      3030      3040
WT_10_final     TCGCCGCATTTCGGGCAGGTTTCGTTAAGTTCCTGCGTGGTAAATTTGGCGCGTTCGATG
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    TCGCCGCATTTCGGGCAGGTTTCGTTAAGTTCCTGCGTGGTAAATTTGGCGCGTTCGATG
                2110      2120      2130      2140      2150      2160

                3050      3060      3070      3080      3090      3100
WT_10_final     CCCTCTTTTTCTTCCACTTGTTTGATGAACGGTTTCCAGAATTTGTCCATCACGGGAATC
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    CCCTCTTTTTCTTCCACTTGTTTGATGAACGGTTTCCAGAATTTGTCCATCACGGGAATC
                2170      2180      2190      2200      2210      2220

                3110      3120      3130      3140      3150      3160
WT_10_final     CAGCGGCGTTTGCCGTCGGCGATTTCGTCAAGCTGGTCTTCGAGTTTGGCGGTGAAGTGG
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    CAGCGGCGTTTGCCGTCGGCGATTTCGTCAAGCTGGTCTTCGAGTTTGGCGGTGAAGTGG
                2230      2240      2250      2260      2270      2280

                3170      3180      3190      3200      3210      3220
WT_10_final     TAATCGACGTATTGGGCGAAGTGTTCGGTCAGGAATTTATTGACGATGTCGCCTGTGTCG
                ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEL_10_    TAATCGACGTATTGGGCGAAGTGTTCGGTCAGGAATTTATTGACGATGTCGCCTGTGTCG
                2290      2300      2310      2320      2330      2340

                3230      3240      3250      3260      3270
WT_10_final     GTGGGCATGAAGCGTTTTGCTCAAGGGAACGATCGCGGCTTNNNNNTTTNNNNNT
                ::::::::::::::::  :::::::::             :      :   :::
MS11_DEL_10_    GTGGGCATGAAGCGTTNNGCTCAAGGG-------------TAACGTATNCGNNNN
                2350      2360      2370                  2380



                       10        20        30        40        50        60
WT_11_Final     ABGGGTAGTCACGCGACGGATCCGGCGATCTGCNCNGGTGGCCACCGTATACCTGATGAA
                ::  :  :::::::::    :  :    :      ::: : :: ::: :::::::: :::
MS11_DEl_11_    ABTTGC-GTCACGCGAGATTTTTGTTTTTTACAANNGGNGNCCCCCGAATACCTGAGGAA
                         10        20        30        40        50

                       70        80        90       100       110       120
WT_11_Final     AGACGTGCGCGGCCCCTCCAGTTTCAAGGGGTTTAAAAACGGCTACGAATTGGGCGGGGC
                : : :::::::::::: : :  :: : ::: : ::: ::::: :: ::::: :::: ::
MS11_DEl_11_    AAAGGTGCGCGGCCCCTTCGGCCTCTAAGGGCTGAAAGACGGCAACTAATTGCGCGGCGC
              60        70        80        90       100       110
```

```
                  130       140       150       160       170       180
WT_11_Final   GTTAAAAGAAGCCTTGTACGACCTGATTAAGCCGTTGGAAAAACCGCTGGGCTTGCCCAA
              ::: :::::: :::::::: :::::::::::::::::::::::::::::: ::::::: :
MS11_DEl_11_  GTTGAAAGAATCCTTGTACCACCTGATTAAGCCGTTGGAAAAACCGCTGGTCTTGCCCGA
           120       130       140       150       160       170

                  190       200       210       220       230       240
WT_11_Final   AACTAAAGAGCCTTTCGTGATTATGCTTGCCGGCATCAACGGGGCGGGCAAAACCTCTTC
              ::::::::::::::: ::::::::::::::: ::::::::  :::::::::: :::: :
MS11_DEl_11_  AACTAAAGAGCCTTTCTTGATTATGCTTGCCGGTATCAACGGCCCGGGCAAAACCACTTC
           180       190       200       210       220       230

                  250       260       270       280       290       300
WT_11_Final   TTTCGGCAAACTCGCCAAATATTTCCAAGCGCAGGGTAAATCCGTTTTGCTGGGGGGGGG
              : :::::::::::::::::::::::::::::::: :::::::: ::::::: ::  ::
MS11_DEl_11_  TATCGGCAAACTCGCCAAATATTTCCAAGCGCAGGGCAAATCCGTATTGCTGGCGGCAGG
           240       250       260       270       280       290

                  310       320       330       340       350       360
WT_11_Final   GGATACCTTCCGCGCCGCCGCCCGTTGAGCAGTTTCTGGCTGGGGGCGGGCCCAACAATG
                :::: :::::::::::::: ::  :::::: ::: :::: ::::::::: ::::::::
MS11_DEl_11_  CAATACTTTCCGCGCCGCCGCGCGC-GAGCAGCTTCAGGCTTGGGGCGGGCGCAACAATG
           300       310       320        330       340       350

                  370       380       390       400       410       420
WT_11_Final   TAACCTTTTTTTCACAAACCACGGGCATTTCCCCCTCCGTGTGTTTCAATGCCGTCCAAG
              ::::: :  ::::::::::::::::  :::: :: ::::::: ::: ::::::::::::
MS11_DEl_11_  TAACCGTCATTTCACAAACCACGGGCGATTCCGCCGCCGTGTGCTTCGATGCCGTCCAAG
           360       370       380       390       400       410

                  430       440       450       460       470       480
WT_11_Final   CCGACAAAGCGTGGGGGATCAACATTTTGCTTGACAACACCGCCGGCCGCCTGCCCATTT
              :::  :::::: : :::::: :::: ::::::: ::::::::::::::::::::::::
MS11_DEl_11_  CCGCGAAAGCGCGCGGGATCTACATCTTGCTTGCCAACACCGCCGGCCGCCTGCCCACGC
           420       430       440       450       460       470

                  490       500       510       520       530       540
WT_11_Final   TTTTTCTTTTGAGGGAAGAAATCAAAAAAGCGATGCGCGTGCTGCAAAAAGCCATTCCCG
                ::::::::::: :::::::::::::::: :: :::::::::: ::::::::::::::
MS11_DEl_11_  AGTTTCTTTTGATGGAAGAAATCAAAAAAGTGAAGCGCGTGCTGAAAAAAGCCATTCCCG
           480       490       500       510       520       530

                  550       560       570       580       590       600
WT_11_Final   GCGCGCCGCACGAAATTATCGTCGTACTCGATGCCAATATCGGGCAAAACGCCGTCAACC
              :::::::::: :::: :::::::::::::::::::::::::::::::::: :::::::
MS11_DEl_11_  GCGCGCCGCACAAAATCATCGTCGTACTCGATGCCAATATCGGGCAAAACGCTGTCAACC
           540       550       560       570       580       590

                  610       620       630       640       650       660
WT_11_Final   AAGTCAAAGCCTTTGACGACGCATTGGGGCTGACGGGGCTTATCGTTACCAAACTCGACG
              :::::::::::::::::::::::::::::::::::::::::: ::::::::::::::
MS11_DEl_11_  AAGTCAAAGCCTTTGACGACGCATTGGGGCTGACGGGGCTTATCCTTACCAAACTCGACG
           600       610       620       630       640       650

                  670       680       690       700       710       720
WT_11_Final   GCACGGCAAAAGGCGGCATCCTCGCCGCGCTTGCTTCCGACCGCCCCGTCCCCGTCCGCT
              :: ::: ::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
MS11_DEl_11_  GCCCGGAAAAAGGCGGCATCCTCGCCGCGCTTGCTTCCGACCGCCCCGTCCCCGTCCGCT
            660       670       680       690       700       710

                  730       740       750       760       770       780
WT_11_Final   ACATCGGCGTGGGCGAAGGCATAGACGACCTGCGCCCGTTTGACGCGCGCGCGTTTGTGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  ACATCGGCGTGGGCGAAGGCATAGACGACCTGCGCCCGTTTGACGCGCGCGCGTTTGTGG
            720       730       740       750       760       770

                  790       800       810       820       830       840
WT_11_Final   ACGCACTGCTGGATTGAGCCGAAATGCCGTCCGAAAACGGCAGACCGAACCGTCATTCCC
              ::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  ACGCACTGCTGGATTGAGTCGAAATGCCGTCCGAAAACGGCAGACCGAACCGTCATTCCC
            780       790       800       810       820       830

                  850       860       870       880       890       900
WT_11_Final   ACGGAAGTGGGAATCTAGGACGCGGGGTTTGGGCAACCGTTTTATCCGATAAGTTTCCGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  ACGGAAGTGGGAATCTAGGACGCGGGGTTTGGGCAACCGTTTTATCCGATAAGTTTCCGT
            840       850       860       870       880       890

                  910       920       930       940       950       960
WT_11_Final   GCGGACAGGTCCGGATTCCCGCCTGCGCGGGAATGACGGGTTTCGAGATTGCGGTATTGT
              ::::::::::::::::: ::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  GCGGACAGGTCCGGATTTCCGCCTGCGCGGGAATGACGGGTTTCGAGATTGCGGTATTGT
            900       910       920       930       940       950

                  970       980       990      1000      1010      1020
WT_11_Final   CGGGAATGACGGTTCGGGTATTTTACTGCGCCCGCCCCGCGCCTGTAAACGGCGGGCGCA
              ::::::::::::::::::::::  : :::::::::::::::::::::::::: :: :::
MS11_DEl_11_  CGGGAATGACGGTTCGGGTATTTCCCCGCGCCCGCCCCGCGCCTGTAAACGGCAGGTGCA
            960       970       980       990      1000      1010

                 1030      1040      1050      1060      1070      1080
WT_11_Final   TCAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAGCGGTAGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  TCAAAAATGCCGTCTGAAGGTTCAGACGGCATCGGTATCGGGGAATCAGAAGCGGTAGCG
           1020      1030      1040      1050      1060      1070

                 1090      1100      1110      1120      1130      1140
WT_11_Final   CACGCCCAACGAGGCTTCGTGGGTTTTGAAGCGGGTGTTTTCCAAGCGTCCCCAATAGTG

MS11_DEl_11_  A-----------------------------------------------------------


                 1150      1160      1170      1180      1190      1200
WT_11_Final   GTAGCGGTAGCCGGTGTCCAGGGTCAGCTTGGGCGTGATGTCGAAACCGACGCCGGCGAT

MS11_DEl_11_  ------------------------------------------------------------


                 1210      1220      1230      1240      1250      1260
WT_11_Final   GACACCGAGACCCACGCGGCGGATGCTGTCGCTTTCGCGATGGGCGCCTGGTGTCCTTAA

MS11_DEl_11_  ------------------------------------------------------------
```

```
                  1270      1280      1290      1300      1310      1320
WT_11_Final   TACCTTATAAACCCCAGACATTATGCCAGGAGTACTGGTGGTAAGAAGCCCTGTTATTTT

MS11_DEl_11_  ------------------------------------------------------------


                  1330      1340      1350      1360      1370      1380
WT_11_Final   TTTAGTCGAATCGATGCTGTGTCTGACGTGTCCGTAGGCGACGCGCGCGCCGATATAGGG

MS11_DEl_11_  ------------------------------------------------------------


                  1390      1400      1410      1420      1430      1440
WT_11_Final   TTTGAATTTATCGTTGAGTTTGAAGTCGTAAACGGCGGACAAGCCGAGAGAAGAAACGGC

MS11_DEl_11_  ------------------------------------------------------------


                  1450      1460      1470      1480      1490      1500
WT_11_Final   GTGGAACGTACCGTTTTCCTGATTTTCCGTCTTCCGGTCTGTCCTGTTGCCATTGACCTT

MS11_DEl_11_  ------------------------------------------------------------


                  1510      1520      1530      1540      1550      1560
WT_11_Final   GTTTCTAAGCAACTCTTTAATGCTCACGGAATATTTAATGTGTTCACTTTCTGTAACGGG

MS11_DEl_11_  ------------------------------------------------------------


                  1570      1580      1590      1600      1610      1620
WT_11_Final   CATAATCTGCCGCTATCCTCCAGCCGCCGAAATTCGTAGCCGACCGACACCCGGGGGTGG

MS11_DEl_11_  ------------------------------------------------------------


                  1630      1640      1650      1660      1670      1680
WT_11_Final   ATGGAATGCGTACGGATGTTTCTGAAATAATCGCTAACCGTGCTTATTTTGCCTTTGCTT

MS11_DEl_11_  ------------------------------------------------------------


                  1690      1700      1710      1720      1730      1740
WT_11_Final   GGATCGGTTTGTTCCGGATAATCGTGGGTAATGTGTTCGTAGGCGTAGGCTAAATCCGCC

MS11_DEl_11_  ------------------------------------------------------------


                  1750      1760      1770      1780      1790      1800
WT_11_Final   TGCACATACGGGCCGCGCCCATGGTCTTCACCCGCCGCCTGCGCTGCGGAAGAGAAGAGA

MS11_DEl_11_  ------------------------------------------------------------
```

```
                    1810      1820      1830      1840      1850      1860
WT_11_Final   AGAGAAGAGAAGAGAAGAGAAGAGAAGGTTTTTTGGGGGCTGGATTCATTTTCGGCTCCT

MS11_DEl_11_  ------------------------------------------------------------
                     1870      1880      1890      1900      1910      1920
WT_11_Final   TATTCGGTTTGACCGGTTAAAAAAAGATTTTCACTGATGTTGAAGGGCGGATTATATCGG

MS11_DEl_11_  ------------------------------------------------------------


                    1930      1940      1950      1960      1970      1980
WT_11_Final   GTTCCGGGCGGTGTTTCAACACAATATGGCGGATGAACAAAAACCGGTACGTCGTTGCCC

MS11_DEl_11_  ------------------------------------------------------------


                    1990      2000      2010      2020      2030      2040
WT_11_Final   CGCCCCGGCCCAAAGGGAACGGTTCCCCAAGGTGATGAAGCACCGGGCGGATCGGTTCCG

MS11_DEl_11_  ------------------------------------------------------------


                    2050      2060      2070      2080      2090      2100
WT_11_Final   TACCATTTGTACTGCCTGCGGCCCGCCGCCTTGTCCTGATTTTTGTTAATCCGCTATAAA

MS11_DEl_11_  ------------------------------------------------------------


                    2110      2120      2130      2140      2150      2160
WT_11_Final   GACCGTCGGGCATCTGCAGCCGTCATTCCCGCGCGGGCGGGAGTCCGGACCGCTTGTTCG

MS11_DEl_11_  ------------------------------------------------------------


                    2170      2180       2190      2200      2210
WT_11_Final   GCAAATGAGGGGGCGGATTGC-GCGCCTGTCAGATAAAAACCGTGTTTAAACGGGTGG-C
                :::::::::::::::  ::: ::::::::::::::::::::::::::::  ::::: :
MS11_DEl_11_  -CAAATGAGGGGGCGGGATGCCGCGCCTGTCAGATAAAAACCGTGTTTAAAACGGTGGGC
            1080      1090      1100      1110      1120      1130

            2220      2230      2240      2250      2260      2270
WT_11_Final   AATGAGGCACATGCAGGGCCTTGAAGCGCAATCGATATATTATTTCCACCGGAACGGACG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  AATGAGGCACATGCAGGGCCTTGAAGCGCAATCGATATATTATTTCCACCGGAACGGACG
            1140      1150      1160      1170      1180      1190

            2280      2290      2300      2310      2320      2330
WT_11_Final   ACCCCGCCCGCCTTGCAAACCCTTAAAAGACAAGCCGCCCGGGTTGATCCGGGCGGCCGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  ACCCCGCCCGCCTTGCAAACCCTTAAAAGACAAGCCGCCCGGGTTGATCCGGGCGGCCGT
            1200      1210      1220      1230      1240      1250

            2340      2350      2360      2370      2380      2390
WT_11_Final   GGAAAATCACTTACCGTTTGATTTATTTGAAATTTAAGGCCTAATTTGCCTCAGCTGGCA
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  GGAAAATCACTTACCGTTTGATTTATTTGAAATTTAAGGCCTAATTTGCCTCAGCTGGCA
            1260      1270      1280      1290      1300      1310
```

```
            2400      2410      2420      2430      2440      2450
WT_11_Final   TCAAAGTTATCGCGGCAGGTTGACGGCAGGTGCTTGGTGTTGATTTTTTCGTTGTCGTTG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  TCAAAGTTATCGCGGCAGGTTGACGGCAGGTGCTTGGTGTTGATTTTTTCGTTGTCGTTG
            1320      1330      1340      1350      1360      1370

            2460      2470      2480      2490      2500      2510
WT_11_Final   CCGGCTTTGGTGACGTCGTCGGCTTTGCCGGCGTTGCCGGCGCCGCGCGTAACCGGCTGT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  CCGGCTTTGGTGACGTCGTCGGCTTTGCCGGCGTTGCCGGCGCCGCGCGTAACCGGCTGT
            1380      1390      1400      1410      1420      1430

            2520      2530      2540      2550      2560      2570
WT_11_Final   CCGCAGAACCATTTTACCGAACCGTCTTGACGCTTGGCCCACAGGGAGAGTCTTTTGCCT
              :::::::::::::::::::::::::::::::::::::::::::::::::: ::::: ::
MS11_DEl_11_  CCGCAGAACCATTTTACCGAACCGTCTTGACGCTTGGCCCACAGGGAGAGTTTTTTGTCT
            1440      1450      1460      1470      1480      1490

            2580      2590      2600      2610      2620      2630
WT_11_Final   TGGATTTCTTTGTTTACGCCGGTTGAAGCCATTTCGGCGGTAACGACGCCGTTTGCGACC
              : ::::::::::::::::::::::::::::::::::::::::::::::::: ::::::::
MS11_DEl_11_  TTGATTTCTTTGTTTACGCCGGTTGAAGCCATTTCGGCGGTAACGACGCCTTTTGCGACT
            1500      1510      1520      1530      1540      1550

            2640      2650      2660      2670      2680      2690
WT_11_Final   GTAACGCTTTGAACATATTTGCCTTTGATGTCGGTGGGGGAGGATGCCACGCCGGCAGAA
                :::  : :::::::::::::::::::: : :  :: :: ::::::::::::::::::::
MS11_DEl_11_  TCAACTTTCTGAACATATTTGCCTTTGATTTTGTCGGCGGGGGATGCCACGCCGGCAGAA
            1560      1570      1580      1590      1600      1610

            2700      2710      2720      2730      2740      2750
WT_11_Final   GTGTTGTCTTCCGGCCATTCGCCGTGATTCAGGTAATACTCGGTGACGGCTGATTTTTGA
              ::::::: :::::::::::::::::::::::::::::::::: ::::::::::::::::
MS11_DEl_11_  GTGTTGTTTTCCGGCCATTCGCCGTGATTCAGGTAATACTCGGTAACGGCTGATTTTTGA
            1620      1630      1640      1650      1660      1670

            2760      2770      2780      2790      2800      2810
WT_11_Final   CCTTCGGCCAAAAGGATGGCTTCGGAAACTTGCGCGCGGGCGGTGTAGTCTTGGTAGGCG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  CCTTCGGCCAAAAGGATGGCTTCGGAAACTTGCGCGCGGGCGGTGTAGTCTTGGTAGGCG
            1680      1690      1700      1710      1720      1730

            2820      2830      2840      2850      2860      2870
WT_11_Final   GGAAGGGCGACTGCCGCCAAAATGCCGACGATAGCGATCACAATCATCAGCTCGATAAGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
MS11_DEl_11_  GGAAGGGCGACTGCCGCCAAAATGCCGACGATAGCGATCACAATCATCAGCTCGATAAGG
            1740      1750      1760      1770      1780      1790

            2880      2890      2900      2910      2920      2930
WT_11_Final   GTAAAGCCTTTTTGAAGGGTATTCATAAAATTACTCCTAATGAAAGGGGAAATCCTCTGG
              ::::::::::::::::::::::::::::::::::::::::::::::::::   :::
MS11_DEl_11_  GTAAAGCCTTTTTGAAGGGTATTCATAAAATTACTCCTAATGAAAGGGGAAATCTCATGG
            1800      1810      1820      1830      1840      1850
```

```
            2940      2950      2960
WT_11_Final   CTACGCCTCGCTATGATGCGGNNN--T
              :::::: : : ::: :  ::      :
MS11_DEl_11_  CTACGCATAG-TATCAAACGTAAANVT
            1860      1870      1880
```

Bibliography

- **Amieva MR, Vogelmann R, Covacci A, Tompkins LS, Nelson WJ, Falkow S.** 2003**.** Disruption of the epithelial apical-junctional complex by Helicobacter pylori CagA. Science. 300:1430–1434.

- **Apicella MA, Mandrell RE, Shero M, Wilson ME, Griffiss JM, Brooks GF, Lammel C, Breen JF, Rice PA**. 1990. Modification by sialic acid of Neisseria gonorrhoeae lipooligosaccharide epitope expression in human urethral exudates: an immunoelectron microscopic analysis. J Infect Dis. 162(2):506-12.

- **Ayala P, Vasquez B, Wetzler L, So M**. 2002. Neisseria gonorrhoeae porin P1.B induces endosome exocytosis and a redistribution of Lamp1 to the plasma membrane. Infect Immun. 70(11):5965-71.

- **Ayala P, Wilbur JS, Wetzler LM, Tainer JA, Snyder A, So M**. 2005. The pilus and porin of Neisseria gonorrhoeae cooperatively induce Ca(2+) transients in infected epithelial cells. Cell Microbiol. 7(12):1736-48.

- **Bacallao** R, **Stelzer EH**. 1989. Preservation of biological specimens for observation in a confocal fluorescence microscope and operational principles of confocal fluorescence microscopy. Methods Cell Biol. 31:437-52.

- **Balda MS, Whitney JA, Flores C, Gonzalez S, Cereijido M, Matter K.** 1996 Functional dissociation of paracellular permeability and transepithelial electrical resistance and disruption of the apicalbasolateral intramembrane diVusion barrier by expression of a mutant tight junction membrane protein. J Cell Biol 134:1031–1049.

- **Bazzoni G, Martinez-Estrada O.M., Orsenigo F, Cordenonsi M, Citi S, Dejana E**. 2000. Interaction of junctional adhesion molecule with the tight junction components ZO-1, cingulin,and occludin, J. Biol. Chem. 275: 20520–20526.

- **Belland RJ, Morrison SG, Carlson JH, Hogan DM.** 1997. Promoter strength influences phase variation of neisserial opa genes. Mol Microbiol. 23(1):123-35.

- **Belland RJ, Morrison SG, van der Ley P, Swanson J.** 1989. Expression and phase variation of gonococcal P.II genes in Escherichia coli involves ribosomal frameshifting and slipped-strand mispairing. Mol Microbiol. 3(6):777-86.

- **Beltinger J, del Buono J, Skelly MM, Thornley J, Spiller RC, Stack WA, Hawkey CJ**. 2008. Disruption of colonic barrier function and induction of mediator release by strains of Campylobacter jejuni that invade epithelial cells. World J Gastroenterol. 14(48):7345-52.

- **Bertelsen LS, Paesold G, Marcus SL, Finlay BB, Eckmann L, Barrett KE**. 2004. Modulation of chloride secretory responses and barrier function of intestinal epithelial cells by the Salmonella effector protein SigD. Am J Physiol Cell Physiol. 287(4):C939-48.

- **Bhat KS, Gibbs CP, Barrera O, Morrison SG, Jähnig F, Stern A, Kupsch EM, Meyer TF, Swanson J.** 1992. The opacity proteins of Neisseria gonorrhoeae strain MS11 are encoded by a family of 11 complete genes. Mol Microbiol. 6(8):1073-6.

- **Bhat, K. S., C.P. Gibbs, O. Barrera, S.G. Morrison, F. Jahnig, A. Stern, E.M. Kupsch, T.F. Meyer, and J. Swanson**. 1991. The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes. Molecular Microbiology. **5**: 1889-1901.

- **Bilek, N., Ison, CA., and Spratt. BG**. 2009. Relative Contributions of Recombination and Mutation to the Diversification of the *opa* Gene Repertoire of *Neisseria gonorrhoeae*. J Bacteriol. 191(6):1878-90.

- **Billker O, Popp A, Brinkmann V, Wenig G, Schneider J, Caron E, Meyer TF.** 2002. Distinct mechanisms of internalization of Neisseria gonorrhoeae by members of the CEACAM receptor family involving Rac1- and Cdc42-dependent and -independent pathways. EMBO J. 21(4):560-71.

- **Binnicker MJ, Williams RD, Apicella MA.** 2004. Gonococcal porin IB activates NF-kappaB in human urethral epithelium and increases the expression of host antiapoptotic factors. Infect Immun. 72(11):6408-17.

- **Bish SE, Song W, Stein DC.** 2008. Quantification of bacterial internalization by host cells using a beta-lactamase reporter strain: Neisseria gonorrhoeae invasion into cervical epithelial cells requires bacterial viability. Microbes Infect. 10(10-11):1182-91.

- **Bjerknes R, Guttormsen HK, Solberg CO, Wetzler LM.** 1995. Neisserial porins inhibit human neutrophil actin polymerization, degranulation, opsonin receptor expression, and phagocytosis but prime the neutrophils to increase their oxidative burst. Infect Immun. 63(1):160-7.

- **Blake MS, Gotschlich EC.** 1984. Purification and partial characterization of the opacity-associated proteins of Neisseria gonorrhoeae. J Exp Med. 159(2):452-62.

- **Blake, M.S., and E.C. Gotschlich.** 1984. Purification and Partial Characterization of the Opacity-Associated Proteins of *Neisseria gonorrhoeae*. J. Exp. Med. **159**: 452-462.

- **Blake, M.S., C.M. Blake, M.A. Apicella, and R.E. Mandrell.** 1995. Gonococcal opacity: lectin-like interactions between Opa proteins and lipooligosaccharide.Infect Immun. **63**(4):1434-9.

- **Blumenthal RD, Leon E, Hansen HJ, Goldenberg DM**. 2007. Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. BMC Cancer. 7:2.

- **Boehm M, Hoy B, Rohde M, Tegtmeyer N, Bæk KT, Oyarzabal OA, Brøndsted L, Wessler S, Backert S**. 2012. Rapid paracellular transmigration of Campylobacter jejuni across polarized epithelial cells without affecting TER: role of proteolytic-active HtrA cleaving E-cadherin but not fibronectin. Gut Pathog. 4(1):3.

- **Boehm, C., Gibert, M., Geny, B., Popoff, M. R. & Rodriguez, P**. 2006. Modification of epithelial cell barrier permeability and intercellular junctions by Clostridium sordellii lethal toxins. Cell Microbiol 8, 1070–1085.

- **Bolan GA, Sparling PF, Wasserheit JN**. 2012. The emerging threat of untreatable gonococcal infection. N Engl J Med. 366(6):485-7.

- **Booth JW, Telio D, Liao EH, McCaw SE, Matsuo T, Grinstein S, Gray-Owen SD**. 2003. Phosphatidylinositol 3-kinases in carcinoembryonic antigen-related cellular adhesion molecule-mediated internalization of Neisseria gonorrhoeae. J Biol Chem. 278(16):14037-45.

- **Bos MP, Hogan D, Belland RJ**. 1999. Homologue scanning mutagenesis reveals CD66 receptor residues required for neisserial Opa protein binding. J Exp Med. 190(3):331-40.

- **Bos MP, Kao D, Hogan DM, Grant CC, Belland RJ**. 2002. Carcinoembryonic antigen family receptor recognition by gonococcal Opa proteins requires distinct combinations of hypervariable Opa protein domains. Infect Immun. 70(4):1715-23.

- **Bos MP, Kuroki M, Krop-Watorek A, Hogan D, Belland RJ**. 1998. CD66 receptor specificity exhibited by neisserial Opa variants is controlled by protein determinants in CD66 N-domains. Proc Natl Acad Sci U S A. 95(16):9584-9.

- **Boulton I.C., and S.D. Gray-Owen**. 2002. Neisserial binding to CEACAM1 arrests the activation and proliferation of CD4+ T lymphocytes. Nat. Immun. 3: 229-236.

- **Bradley, C.J., N.J. Griffiths, H.A. Rowe, R.S. Heyderman, and M. Virji.** 2005. Critical determinants of the interactions of capsule-expressing *Neisseria meningitidis* with host cells: the role of receptor density in increases cellular

targeting via the outer membrane Opa proteins. Cellular Microbiol. **7**(10): 1490-1503.

- **Bras, A. M., and J. M. Ketley.** 1999. Transcellular translocation of *Campylobacter jejuni* across human polarized epithelial monolayers. FEMS Microbiol. Lett. 1**7**9: 209-215.

- **Burch CL, Danaher RJ, Stein DC**. 1997. Antigenic variation in Neisseria gonorrhoeae: production of multiple lipooligosaccharides. J Bacteriol. 179(3):982-6.

- **Burns, J. L., A. Griffith, J. J. Barry, M. Jonas, and E. Y. Chi.** 2001. Transcytosis of gastrointestinal epithelial cells by *Escherichia coli* K1. Pediatr. Res. 49**:**30-37

- **Butt NJ, Lambden PR, Heckels JE.** 1990. The nucleotide sequence of the por gene from Neisseria gonorrhoeae strain P9 encoding outer membrane protein PIB. Nucleic Acids Res. 18(14):4258.

- **Cannon JG, Buchanan TM, Sparling PF.** 1983. Confirmation of association of protein I serotype of Neisseria gonorrhoeae with ability to cause disseminated infection. Infect Immun. 40(2):816-9.

- **Carey DJ**. 1997. Syndecans: multifunctional cell-surface co-receptors. Biochem J. 327 (Pt 1):1-16. Review

- **Chen ML, Ge Z, Fox JG, Schauer DB.** 2006. Disruption of tight junctions and induction of proinflammatory cytokine responses in colonic epithelial cells by Campylobacter jejuni. Infect Immun. 74(12):6581-9.

- **Chen T, Bolland S, Chen I, Parker J, Pantelic M, Grunert F, Zimmermann W.** 2001. The CGM1a (CEACAM3/CD66d)-mediated phagocytic pathway of Neisseria gonorrhoeae expressing opacity proteins is also the pathway to cell death. J Biol Chem. 276(20):17413-9.

- **Chen T, Gotschlich EC**. 1996. CGM1a antigen of neutrophils, a receptor of gonococcal opacity proteins. Proc Natl Acad Sci U S A. 93(25):14851-6

- **Chen, T., R.J. Belland, J. Wilson, and J. Swanson.** 1995. Adherence of Pilus^{-} Opa^{+} Gonococci to Epithelial Cells In Vitro Involves Heparan Sulfate. J. Exp. Med. **182**: 511-517.

- **Cohen MS, Hoffman IF, Royce RA**, Kazembe P, **Dyer JR, Daly CC, Zimba D, Vernazza PL, Maida M, Fiscus SA, Eron JJ Jr**. 1997. Reduction of concentration of HIV-1 in semen after treatment of urethritis: implications for prevention of transmission of HIV-1. Lancet 1997; 349:1868–73.

- **Cohen MS**. 1998. Sexually transmitted diseases enhance HIV transmission: no longer a hypothesis. Lancet. 351(Suppl III):5–7.

- **Conlin VS, Curtis SB, Zhao Y, Moore ED, Smith VC, Meloche RM, Finlay BB, Buchan AM**. 2004. Helicobacter pylori infection targets adherens junction regulatory proteins and results in increased rates of migration in human gastric epithelial cells. Infect Immun. 72:5181–5192.

- **Connell, T.D., D. Shafer, J.G. Cannon.** 1990. Characterization of the repertoire of hypervariable regions in the Protein II (opa) gene family of *Neisseria gonorrhoeae*. Molec. Microbiol. 4(3): 439-449.

- **Connell, T.D., D.S. Barrit, W.J. Black, T.H. Wakula, D.G. Klapper, R.S. Schwalbe, A. Stephenson, and J.G. Cannon.** 1987. Genetic and biochemical analyses of protein II. Anton. Van. Lee. **53**(6): 421-424.

- **Cordenonsi M, D'Atri F, Hammar E, Parry DA, Kendrick-Jones J**, **Shore D ,** Citi S**.** 1999. Cingulin contains globular and coiled-coil domains and interacts with ZO-1, ZO-2, ZO-3, and myosin. J. Cell Biol. 147:1569–82.

- **D'Atri F, Citi S**. 2001. Cingulin interacts with F-actin in vitro. FEBS Lett. 507:21–24.

- **Danaher RJ, Levin JC, Arking D, Burch CL, Sandlin R, Stein DC.** 1995. Genetic basis of Neisseria gonorrhoeae lipooligosaccharide antigenic variation. Bacteriol. 177(24):7275-9.

- **de Jonge MI, Hamstra HJ, van Alphen L, Dankert J, van der Ley P**. 2003. Mapping the binding domains on meningococcal Opa proteins for CEACAM1 and CEA receptors. Mol Microbiol. 50(3):1005-15.

- **De jonge, M. I., M.P. Bos, H.J. Hamstra, W. Jiskoot, P. van Ulsen, J. Tommassen, L. van Alphen, P. vander Ley**. 2002. Conformational analysys of opacity proteins fron *Neisseria meningitidis*. Eur. J. Biochem. **269**: 5215-5223.

- **Dehio C, Freissler E, Lanz C, Gómez-Duarte OG, David G, Meyer TF.** 1998. Ligation of cell surface heparan sulfate proteoglycans by antibody-coated beads stimulates phagocytic uptake into epithelial cells: a model for cellular invasion by Neisseria gonorrhoeae. Exp Cell Res. 242(2):528-39.

- **Dehio M, Gómez-Duarte OG, Dehio C, Meyer TF**. 1998. Vitronectin-dependent invasion of epithelial cells by Neisseria gonorrhoeae involves alpha(v) integrin receptors. FEBS Lett. 424(1-2):84-8.

- **Derrick JP, Urwin R, Suker J, Feavers IM, Maiden MC.** 1999. Structural and evolutionary inference from molecular variation in Neisseria porins. Infect Immun. 67(5):2406-13.

- **Dharmsathaphorn, K., J.L. Madara.** 1990. Established intestinal cell lines as model systems for electrolyte transport studies. Methods Enzymol. **192**: 354-389**.**

- **Duensing TD, Putten JP.** 1998. Vitronectin binds to the gonococcal adhesin OpaA through a glycosaminoglycan molecular bridge. Biochem J. 334 (Pt 1):133-9.

- **Duensing TD, van Putten JP.** 1997. Vitronectin Mediates Internalization of Neisseria gonorrhoeae by Chinese Hamster Ovary Cells. Infect Immun. 65 (3): 964–970.

- **Duensing, T.D., Wing, J.S., and van Putten, J.P.M.** 1999. Sulfated Polysaccharide-Directed Recruitment of Mammalian Host Proteins: a Novel Strategy in Microbial Pathogenesis. Infect. Immun. 67(9): 4463-4468.

- **Eckmann, L., H. C. Jung, C. Schurer-Maly, A. Panja, E. Morzycka-Wroblewska, and M. F. Kagnoff**. 1993. Differential cytokine expression by human intestinal epithelial cell lines: regulated expression of interleukin 8. Gastroenterology 105:1689–1697.

- **Edwards JL, Apicella MA.** 2002. The role of lipooligosaccharide in Neisseria gonorrhoeae pathogenesis of cervical epithelia: lipid A serves as a C3 acceptor molecule. Cell Microbiol. 4(9):585-98.

- **Edwards JL, Brown EJ, Ault KA, Apicella MA.** 2001. The role of complement receptor 3 (CR3) in Neisseria gonorrhoeae infection of human cervical epithelia. Cell Microbiol. 3(9):611-22.

- **Edwards VL, Wang LC, Dawson V, Stein DC, Song W**. 2012. Neisseria gonorrhoeae breaches the apical junction of polarized epithelial cells for transmigration by activating EGFR. Cell Microbiol. Dec 26. doi: 10.1111/cmi.12099. [Epub ahead of print]

- **Eisenstein BI, Masi AT**. 1981. Disseminated gonococcal infection (DGI) and gonococcal arthritis (GCA): I. Bacteriology, epidemiology, host factors, pathogen factors, and pathology. Semin Arthritis Rheum. 10(3):155-72.

- **Elenius K, Jalkanen M.** 1994. Function of the syndecans--a family of cell surface proteoglycans. J Cell Sci. 107 (Pt 11):2975-82. Review

- **Eschenbach DA, Buchanan TM, Pollock HM, Forsyth PS, Alexander ER, Lin JS,** Wang SP**,** Wentworth BB**,** MacCormack WM**,** Holmes KK. 1975. Polymicrobial etiology of acute pelvic inflammatory disease. N Engl J Med 293:166-71.

- **Feavers IM, Maiden MC.** 1998. A gonococcal porA pseudogene: implications for understanding the evolution and pathogenicity of Neisseria gonorrhoeae. Mol Microbiol. 30(3):647-56.

- **Finlay, B. B., and S. Falkow.** 1990. *Salmonella* interactions with polarized human intestinal Caco-2 epithelial cells. J. Infect. Dis. **162:**1096-1106.

- **Fischer, S. H., and R.F. Rest**. 1988. Gonococci Possessing Only Certain P. II Outer Membrane Proteins Interact with Human Neutrophils. Infection and Immunity. **56**: 1574-1579.

- **Fleming, D.T., J.N. Wasserheit.** 1999. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex. Transm. Infect. **75**(1): 3-17.

- **Franco AT, Johnston E, Krishna U, Yamaoka Y, Israel DA, Nagy TA, Wroblewski LE, Piazuelo MB, Correa P, Peek RM Jr.** 2008. Regulation of gastric carcinogenesis by Helicobacter pylori virulence factors. Cancer Res. 68:379–387.

- **Frängsmyr, L., V. Baranov, and S. Hammarstrom.** 1999. Four Carcinoembryonic Antigen Subfamily Member, CEA, NCA, BGP, and CGM2, Slelectively Expressed in the Normal Human Colonic Epithelium, Are Integral Components of the Fuzzy Coat. Tumor Biolo. **20**: 277-292.

- **Freissler E, Meyer auf der Heyde A, David G, Meyer TF, Dehio C**. 2000. Syndecan-1 and syndecan-4 can mediate the invasion of OpaHSPG-expressing Neisseria gonorrhoeae into epithelial cells. Cell Microbiol. 2000 Feb;2(1):69-82.

- **Furuse M, Hirase T, Itoh M, Nagafuchi A, Yonemura S, Tsukita S, Tsukita S.** 1993. Occludin: a novel integral membrane protein localizing at tight junctions. J. Cell Biol. 123:1777–88.

- **Furuse M, Sasaki H, Fujimoto K, Tsukita S.** 1998b. A single gene product, claudin-1 or -2, reconstitutes tight junction strands and recruits occludin in fibroblasts. J. Cell Biol. 143:391–401.

- **Furuse, M., Fujita, K., Hiiragi, T., Fujimoto, K. & Tsukita, S**. 1998a. Claudin-1 and-2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. J. Cell Biol. 141, 1539–1550.

- **Gail A. Bolan, M.D., P. Frederick Sparling, M.D., and Judith N. Wasserheit, M.D., M.P.H**. 2012. The Emerging Threat of Untreatable Gonococcal Infection. N Engl J Med 366:485-487.

- **Gallagher JT, Lyon M, Steward WP**. 1986. Structure and function of heparan sulphate proteoglycans. Biochem J. 236(2):313-25.

- **Garvin LE, Bash MC, Keys C, Warner DM, Ram S, Shafer WM, Jerse AE.** 2008. Phenotypic and genotypic analyses of Neisseria gonorrhoeae isolates that express frequently recovered PorB PIA variable region types suggest that certain P1a porin sequences confer a selective advantage for urogenital tract infection. Infect Immun. 76(8):3700-9.

- **Ghosh SK, Zhao J, Philogene MC, Alzaharani A, Rane S, Banerjee A.** 2004. Pathogenic consequences of Neisseria gonorrhoeae pilin glycan variation. Microbes Infect. 6(7):693-701. Review

- **Gomez-Duarte, O.G., M. Dehio, C.A. Guzman, G.S. Chhatwal, and T.F. Meyer.** 1997. Binding of Vitronectin Opa-Expressing *Neisseria gonorrhoeae* Mediates Invasion of HeLa Cells. Infec. Immun. **65**(9): 3857-3866.

- **Gorby, G.L., A.F. Ehrhardt, M.A. Apicella, and C. Elkins.** 2001. Invasion of Human Fallopian Tube Epithelium by *Escherichia coli* Expressing Combinations of a Gonococcal Porin, Opacity-Associated Protein, and Chimeric Lipo-Ologosaccharide. J. Infect. Dis. **184**: 460-472.

- **Gotschlich EC, Seiff ME, Blake MS, Koomey M.** 1987. Porin protein of Neisseria gonorrhoeae: cloning and gene structure. Proc Natl Acad Sci U S A. 84(22):8135-9.

- **Grant C.C.R., M.P. Bos, and R.J. Belland.** 1999. Proteoglycan receptor binding by *Neisseria gonorrhoeae* MS11 is determined by the HV-1 region of OpaA. Mol. Micro. **32**(2): 233-242.

- **Grassmé H, Gulbins E, Brenner B, Ferlinz K, Sandhoff K, Harzer K, Lang F, Meyer TF**. 1997. Acidic sphingomyelinase mediates entry of N. gonorrhoeae into nonphagocytic cells. Cell. 91(5):605-15.

- **Grassme, H.U.C., R.M. Ireland and J.P.M. van Putten**. 1996. Gonococcal Opacity Protein Promotes Bacterial Entry-Associated Rearrangements of the Epithelial Cell Actin Cytoskeleton. Infec. Immun. **64**(5): 1621-1630.

- **Gray-Owen SD, Dehio C, Haude A, Grunert F, Meyer TF**. 1997. CD66 carcinoembryonic antigens mediate interactions between Opa-expressing Neisseria gonorrhoeae and human polymorphonuclear phagocytes. EMBO J. 16(12):3435-45.

- **Griffiss JM, Lammel CJ, Wang J, Dekker NP, Brooks GF.** 1999. Neisseria gonorrhoeae coordinately uses Pili and Opa to activate HEC-1-B cell microvilli, which causes engulfment of the gonococci. Infect Immun. 67(7):3469-80.

- **Groeger S, Doman E, Chakraborty T, Meyle J**. 2010. Effects of Porphyromonas gingivalis infection on human gingival epithelial barrier function in vitro. Eur J Oral Sci. 118(6):582-9.

- **Guido DM, McKenna R, Mathews WR.** 1993. Quantitation of hydroperoxy-eicosatetraenoic acids and hydroxy-eicosatetraenoic acids as indicators of lipid peroxidation using gas chromatography-mass spectrometry. Anal Biochem. 209(1):123-9.

- **Gunn, J.S., and D.C. Stein.** 1996. Use of a non-selective transformation technique to construct a miltipli Restriction/Modification-deficient mutnat of *Neisseria gonorrhoeae*. Mol. Gen. Genet. **251**: 509-517.

- **Hammarström S.** 1999. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. Semin Cancer Biol. 9(2):67-81. Review

- **Hamrick TS, Dempsey JA, Cohen MS, Cannon JG**. 2001. Antigenic variation of gonococcal pilin expression in vivo: analysis of the strain FA1090 pilin repertoire and identification of the pilS gene copies recombining with pilE during experimental human infection. Microbiology. 147(Pt 4):839-49.

- **Harvey HA, Jennings MP, Campbell CA, Williams R, Apicella MA**. 2001. Receptor-mediated endocytosis of Neisseria gonorrhoeae into primary human urethral epithelial cells: the role of the asialoglycoprotein receptor. Mol Microbiol. 42(3):659-72.

- **Harvey HA, Porat N, Campbell CA, Jennings M, Gibson BW, Phillips NJ, Apicella MA, Blake MS**. 2000. Gonococcal lipooligosaccharide is a ligand for the asialoglycoprotein receptor on human sperm. Mol Microbiol. 36(5):1059-70.

- **Harvey, H. A., Post, D. M., and Apicella, M. A.** (2002). Immortalization of human urethral epithelial cells: a model for the study of the pathogenesis of and the inflammatory cytokine response to *Neisseria gonorrhoeae* infection. Infect. Immun. 70, 5808–5815.

- **Hauck CR, Meyer TF, Lang F, Gulbins E**. 1998. CD66-mediated phagocytosis of Opa52 Neisseria gonorrhoeae requires a Src-like tyrosine kinase-and Rac1-dependent signalling pathway. EMBO J. 17(2):443-54.

- **Hecht G, Pothoulakis C, LaMont JT, Madara JL.** 1988. Clostridium difficile toxin A perturbs cytoskeletal structure and tight junction permeability of cultured human intestinal epithelial monolayers. J Clin Invest. 82(5):1516-24.

- **Higashi DL, Zhang GH, Biais N, Myers LR, Weyand NJ, Elliott DA, So M.** 2009. Influence of type IV pilus retraction on the architecture of the *Neisseria gonorrhoeae*-infected cell cortex. Microbiology. 155(Pt 12):4084-92.

- **Hill DJ, Edwards AM, Rowe HA, Virji M**. 2005. Carcinoembryonic antigen-related cell adhesion molecule (CEACAM)-binding recombinant polypeptide confers protection against infection by respiratory and urogenital pathogens. Mol Microbiol. 55(5):1515-27.

- **Hill, D.J., and M. Virji**. 2003. A novel cell-binding mechanism of *Moraxella catarrhalis* ubiquitous surface protein UspA: specific targeting of the N-domain of carcinoembryonic antigen-related cell adhesion molecules by UspA1. Molec. Microbiol. **48**(1): 117-129.

- **Hill, D.J., M.A. Toleman, D.J. Evans, S. Villullas, L. van Alphen, and M. Virji.** 2001. The variable P5 proteins of typeable and non-typeable *Haemophilus influenzae* target human CEACAM1. Molec. Microbiol. 39(4): 850-862.

- **Hobbs, M., Sparling,P.F., Cohen, M.S., Shafer, W.M., Deal, C.D., and Jerse A**. 2011. Experimental Gonococcal Infection in Male Volunteers: Cumulative Experience with *Neisseria gonorrhoeae* Strains FA1090 and MS11mkC. Front Microbiol. 2011; 2: 123.

- **Hobbs, M.M., A. Seiler, M. Achtman, and J.G. Cannon.** 1994. Microevolution within a clonal population of pathogenic bacteria: recombination, gene duplication and horizontal genetic exchange in the *opa* gene family of *Neisseria meningitidis*. Molec. Microbiol. **12**(2): 171-180.

- **Holmes KK, Eschenbach DA, Knapp JS**. 1980. Salpingitis: overview of etiology and epidemiology. Am J Obstet Gynecol. 138:893-900.

- **Holmes, K.V., G. Dveksler, S. Gagneten, C. Yeager, S.H. Lin, N. Beauchemin, A.T. Look, R. Ashmun, and C. Dieffenbach.** 1993. Coronavirus receptor specificity. Adv. Exp. Med. Biol. **342**: 261-266.

- **Hopper, S., B. Vasquez, A. Merz, S. Clary, J.S. Wilbur, M. So.** 2000a. Effects of the Immunoglobulin A1 Protease on *Neisseria gonorrhoeae* Trafficking across Polarized T84 Epithelial Monolayers. Infect. Immun. **68**(2): 906-911.

- **Hopper, S., J.S. Wilbur, B.L. Vasquez, J. Larson, S. Clary, I.J. Mehr, H.S. Seifert, and M. So.** 2000. Isolation of *Neisseria gonorrhoeae* Mutants That

Show Enhances Trafficking across Polarized T84 Epithelial Monolayers. Infec. Immun. **68**(2): 896-905.

- **Hou J, Paul DL, Goodenough DA**. 2005. Paracellin-1 and the modulation of ion selectivity of tight junctions. J. Cell Sci. 118:5109–18.

- **Howie HL, Shiflett SL, So M.** 2008. Extracellular signal-regulated kinase activation by Neisseria gonorrhoeae downregulates epithelial cell proapoptotic proteins Bad and Bim. Infect Immun. 76(6):2715-21.

- **Huber D, Balda MS, Matter K**. 2000. Occludin modulates transepithelial migration of neutrophils. J Biol Chem. 275(8):5773-8.

- **Ide N, Hata Y, Nishioka H, Hirao K, Yao I, Deguchi M, Mizoguchi A, Nishimori H, Tokino T, Nakamura Y, Takai Y.** 1999. Localization of membrane-associated guanylate kinase (MAGI)-1/BAI-associated protein (BAP) 1 at tight junctions of epithelial cells. Oncogene. 18(54):7810-5.

- **Iden S, Misselwitz S, Peddibhotla SS, Tuncay H, Rehder D, Gerke V, Robenek H, Suzuki A, Ebnet K.** 2012. aPKC phosphorylates JAM-A at Ser285 to promote cell contact maturation and tight junction formation. J Cell Biol. 196(5):623-39.

- **Ilver D, Källström H, Normark S, Jonsson AB.** 1998. Transcellular passage of Neisseria gonorrhoeae involves pilus phase variation. Infect Immun. 66(2):469-73.

- **Ison, CA., J.A.R. Dillon, and J.W. Tapsall**. 1998. The epidemiology of global antibiotic reistance among Neisseria gonorrhoeae and Haemophilus Ducreyi. Lancet. Supplement STD'a, **351**(9119)p8, 4p, 2 graphs.

- **Jalkanen M, Nguyen H, Rapraeger A, Kurn N, Bernfield M.** 1985. Heparan sulfate proteoglycans from mouse mammary epithelial cells: localization on the cell surface with a monoclonal antibody. J Cell Biol. 101(3):976-84.

- **Jalkanen M**. 1987. Biology of cell surface heparan sulfate proteoglycans. Med Biol. 65(1):41-7. Review

- **James, J.F., and J. Swanson.** 1978. Studies on gonococcus infection. XIII. Ocurrence of color/opacity colonial variants in clinical cultures. Infect. Immun. **19**(1): 332-340.

- **Jarva H, Ngampasutadol J, Ram S, Rice PA, Villoutreix BO, Blom AM.** 2007. Molecular characterization of the interaction between porins of Neisseria gonorrhoeae and C4b-binding protein. J Immunol. 179(1):540-7.

- **Jerse AE, Cohen MS, Drown PM, Whicker LG, Isbey SF, Seifert HS, Cannon JG.** 1994. Multiple gonococcal opacity proteins are expressed during experimental urethral infection in the male**.** J Exp Med. 179(3):911-20.

- **John CM, Jarvis GA, Swanson KV, Leffler H, Cooper MD, Huflejt ME, Griffiss JM.** 2002. Galectin-3 binds lactosaminylated lipooligosaccharides from Neisseria gonorrhoeae and is selectively expressed by mucosal epithelial cells that are infected. Cell Microbiol. 4(10):649-62.

- **John CM, Schneider H, Griffiss JM.** 1999. Neisseria gonorrhoeae that infect men have lipooligosaccharides with terminal N-acetyllactosamine repeats. J Biol Chem. 274(2):1017-25.

- **John W. Tapsall**. 2005. Antibiotic Resistance in Neisseria gonorrhoeae. Gonococcal Resistance to Antimicrobials • CID 41 (Suppl 4) • S263.

- **Johnston KH, Gotschlich EC**. 1974. Isolation and characterization of the outer membrane of Neisseria gonorrhoeae. J Bacteriol. 119(1):250-7.

- **Jonsson AB, Ilver D, Falk P, Pepose J, Normark S**. 1994. Sequence changes in the pilus subunit lead to tropism variation of Neisseria gonorrhoeae to human tissue. Mol Microbiol. 13(3):403-16.

- **Jung HC, Eckmann L, Yang SK, Panja A, Fierer J, Morzycka-Wroblewska E, Kagnoff MF.** 1995. A distinct array of proinflammatory cytokines is expressed in human colon epithelial cells in response to bacterial invasion. J Clin Invest. 95(1):55-65.

- **Källström H, Blackmer Gill D, Albiger B, Liszewski MK, Atkinson JP, Jonsson AB.** 2001. Attachment of Neisseria gonorrhoeae to the cellular pilus receptor CD46: identification of domains important for bacterial adherence. Cell Microbiol. 3(3):133-43.

- **Källström H, Islam MS, Berggren PO, Jonsson AB**. 1998. Cell signaling by the type IV pili of pathogenic Neisseria. J Biol Chem. 273(34):21777-82.

- **Katz J, Sambandam V, Wu JH, Michalek SM, Balkovetz DF.** 2000. Characterization of Porphyromonas gingivalis-induced degradation of epithelial cell junctional complexes. Infect Immun. 68(3):1441-9.

- **Katz J, Yang QB, Zhang P, Potempa J, Travis J, Michalek SM, Balkovetz DF.** 2002. Hydrolysis of epithelial junctional proteins by Porphyromonas gingivalis gingipains. Infect Immun. 70(5):2512-8.

- **Kerle, K.K., J.R. Mascola, and T.A. Miller.** 1992. Disseminated gonococcal infection. Am. Fam. Phy. 45(1): 209-214.

- **Kerwood DE, Schneider H, Yamasaki R.** 1992. Structural analysis of lipooligosaccharide produced by Neisseria gonorrhoeae, strain MS11mk (variant A): a precursor for a gonococcal lipooligosaccharide associated with virulence. Biochemistry. 31(51):12760-8.

- **Kim JY, Sajjan US, Krasan GP, LiPuma JJ.** 2005. Disruption of tight junctions during traversal of the respiratory epithelium by Burkholderia cenocepacia. Infect Immun. 73(11):7107-12.

- **Kuespert K, Pils S, Hauck CR**. 2006**.** CEACAMs: their role in physiology and pathophysiology. Curr Opin Cell Biol. 18(5):565-71.

- **Kühlewein C, Rechner C, Meyer TF, Rudel T.** 2006. Low-phosphate-dependent invasion resembles a general way for Neisseria gonorrhoeae to enter host cells. Infect Immun. 74(7):4266-73.

- **Kupsch, E., B. Knepper, T. Kuroki, I. Heuer, and T.F Meyer.** 1993. Variable opacity (Opa) outer membrane proteins account for the cell tropisms displayed by *Neisseria gonorhroeae* for human leukocytes and epithelial cells. EMBO J. **12**(2): 641-650.

- **Lapointe TK, O'Connor PM, Jones NL, Menard D, Buret AG**. 2010. Interleukin-1 receptor phosphorylation activates Rho kinase to disrupt human gastric tight junctional claudin-4 during Helicobacter pylori infection. Cell Microbiol. 12(5):692-703.

- **Laterra J, Silbert JE, Culp LA**. 1983. Cell surface heparan sulfate mediates some adhesive responses to glycosaminoglycan-binding matrices, including fibronectin. J Cell Biol. 96(1):112-23.

- **Lee SW, Higashi DL, Snyder A, Merz AJ, Potter L, So M**. 2005. PilT is required for PI(3,4,5)P3-mediated crosstalk between *Neisseria gonorrhoeae* and epithelial cells. Cell Microbiol. 7(9):1271-84.

- **Leusch, H.G., Z. Drzeniek, Z. Markos-Putsztai, and C. Wagener.** 1991. Binding of *Escherichia coli* and *Salmonella* strains to member of the carcinoembryonic antigen family: differential binding inhibition by aromatic alpha-glycosides of mannose. Infect. Immun. 59: 2051-2057.

- **Lewis LA, Ram S, Prasad A, Gulati S, Getzlaff S, Blom AM, Vogel U, Rice PA.** 2008. Defining targets for complement components C4b and C3b on the pathogenic neisseriae. Infect Immun. 76(1):339-50.

- **Liu Y, Nusrat A, Schnell FJ, Reaves TA, Walsh S, Pochet M, Parkos CA.** 2000. Human junction adhesion molecule regulates tight junction resealing in epithelia. J. Cell Sci. 113 (Pt. 13):2363–74.

- **Lorenzen DR, Günther D, Pandit J, Rudel T, Brandt E, Meyer TF**. 2000. Neisseria gonorrhoeae porin modifies the oxidative burst of human professional phagocytes. Infect Immun. 68(11):6215-22.

- **Madara, J. L., S. P. Colgan, A. Nusrat, C. Delp, and C. A. Parkos.** 1992. A simple approach to measurement of electrical parameters of cultured epithelial monolayer: use in assessing neutrophil epithelial interaction. J. Tissue Culture Methods 14**:**209-216.

- **Madara, J.L., J. Stafford, K. Dharmsathaphorn, and S. Carlson.** 1987. Structutal Analysis of a Human Intestinal Epithelial Cell Line. Gastroenterol. 92: 1133-1145.

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

- **Malorny, B., G. Morell, B. Kusecece, J. Kolberg, and M. Achtman**. 1998. Sequence Diversity, Predicted Two – Dimensional Protein Structure, and Epitope Mapping of Neisserial Opa Proteins. Journal of Bacteriology. 180: 1323-1330.

- **Man SM, Kaakoush NO, Leach ST, Nahidi L, Lu HK, Norman J, Day AS, Zhang L, Mitchell HM**. 2010. Host attachment, invasion, and stimulation of proinflammatory cytokines by Campylobacter concisus and other non-Campylobacter jejuni Campylobacter species. J Infect Dis. Dec 15;202(12):1855-65.

- **Man SM, Kaakoush NO, Leach ST, Nahidi L, Lu HK, Norman J, Day AS, Zhang L, Mitchell HM.** 2010. Host attachment, invasion, and stimulation of proinflammatory cytokines by Campylobacter concisus and other non-Campylobacter jejuni Campylobacter species. J Infect Dis. 202(12):1855-65.

- **Mandrell RE, Griffiss JM, Macher BA**. 1988. Lipooligosaccharides (LOS) of Neisseria gonorrhoeae and Neisseria meningitidis have components that are immunochemically similar to precursors of human blood group antigens. Carbohydrate sequence specificity of the mouse monoclonal antibodies that recognize crossreacting antigens on LOS and human erythrocytes. J Exp Med. 168(1):107-26. Erratum in: J Exp Med 1988 168(4):1517.

- **Martin D, Turgeon PL, Mathieu LG**. 1986. Approximate molecular weight of envelope protein 1 and colony opacity of Neisseria gonorrhoeae strains isolated

from patients with disseminated or localized infection. Sex Transm Dis. 13(2):71-5.

- **Martine P. Bos, David Kao, Daniel M. Hogan, Christopher C. R. Grant and Robert J. Belland**. 2002. Carcinoembryogenic Antigen family Receptor recognition by Gonococcal Opa Proteins Requires Distinct Combinations of Hypervariable Opa Protein Domains. Infection and Immunity 70 (4) 1715-1723.

- **Mayer, LW**. 1982. Rates of In Vitro Changes of Gonococcal Colony Opacity Phenotypes. Infect Immun. 37(2):481-5.

- **McCaw SE, Schneider J, Liao EH, Zimmermann W, Gray-Owen SD**. 2003. Immunoreceptor tyrosine-based activation motif phosphorylation during engulfment of Neisseria gonorrhoeae by the neutrophil-restricted CEACAM3 (CD66d) receptor. Mol Microbiol. 49(3):623-37.

- **McCaw, S.E., E.H. Liao, and S.D. Gray-Owen.** 2004. Engulfment of *Neisseria gonorrhoeae*: Revaling Distinct Processes of Bacterial Entry by Individual Carcinoembryonic Antigen-Related Cellular Adhesion Molecule Family Receptors. Infec. Immun. 72(5): 2742-2752.

- **McCormick, B. A., S. I. Miller, and J. L. Madara.** 1995. Transepithelial signaling to neutrophils by salmonellae: a novel virulence mechanism for gastroenteritis. Infec. Immun. 63(6): 2302-2309.

- **McGee ZA, Jensen RL, Clemens CM, Taylor-Robinson D, Johnson AP, Gregg CR**. 1999. Gonococcal infection of human fallopian tube mucosa in organ culture: relationship of mucosal tissue TNF-alpha concentration to sloughing of ciliated cells. Sex Transm Dis. 26(3):160-5.

- **McGee, Z.A., A.P. Johnson, and D. Taylor-Robinson.** 1981 Pathogenic Mechanisms of *Neisseria gonorrhoeae*: Observations on Damage to Human Fallopian Tubes in Organ Culture by Gonococci of Colony Type 1 or Type 4. J. Infect. Dis. 143(3): 413-422.

- **McGee, Z.A., D.S. Stephens, L.H. Hoffman, W.F. 3rd. Schlech, and R.G. Horn.** 1983. Mechanisms of mucosal invasion by pathogenic Neisseria. Rev. Infect. Dis. 5(Suppl. 4): S708-14.

- **McLaughlin SE, Cheng H, Ghanem KG, Yang Z, Melendez J, Zenilman J, Griffiss JM.** 2012. Urethral Exudates of Men with Neisseria gonorrhoeae Infections Select a Restricted Lipooligosaccharide Phenotype During Transmission. J Infect Dis. 206(8):1227-32.

- **Mehr IJ, Seifert HS.** 1998. Differential roles of homologous recombination pathways in Neisseria gonorrhoeae pilin antigenic variation, DNA transformation and DNA repair. Mol Microbiol. 30(4):697-710.

- **Merz, A.J., D.B. Rifenbery, C.G. Arvidson, and M. So.** 1996. Transversal of a Polarized Epithelium by Pathogenic *Neisseriae*: Facilitation by Type IV Pili and Manteinance of Epithelial Barrier Function. Molec. Med. **2**(6): 745-754.

- **Meyer TF, van Putten JP**. 1989. Genetic mechanisms and biological implications of phase variation in pathogenic neisseriae. Clin Microbiol Rev. 2 Suppl: S139-45.

- **Moran J.** 2005. Gonorrhoea. Clin Evid. **13**: 2016-2023.

- **Morand PC, Bille E, Morelle S, Eugène E, Beretti JL, Wolfgang M, Meyer TF, Koomey M, Nassif X**. 2004. Type IV pilus retraction in pathogenic Neisseria is regulated by the PilC proteins. EMBO J. 23(9):2009-17.

- **Morello JA, Bohnhoff M**. 1989. Serovars and serum resistance of Neisseria gonorrhoeae from disseminated and uncomplicated infections. J Infect Dis. 160(6):1012-7.

- **Mosleh IM, Huber LA, Steinlein P, Pasquali C, Günther D, Meyer TF.** 1998. Neisseria gonorrhoeae porin modulates phagosome maturation. J Biol Chem. 273(52):35332-8.

- **Mosleh, I.M., H. Boxberger, M.J. Sessler, and T.F. Meyer.** 1997. Experimental Infection of Native Human Ureteral Tissue with *Neisseria gonorrhoeae*: Adhesion, Invasion, Intracellular Fate, Exocytosis, and Passage through a Stratified Epithelium. Infect. Immun. **65**(8): 3391-3398.

- **Muenzner P**, **Billker O, Meyer TF, Naumann M**. 2002. Nuclear factor-kappa B directs carcinoembryonic antigen-related cellular adhesion molecule 1 receptor expression in Neisseria gonorrhoeae-infected epithelial cells. J Biol Chem. 277(9):7438-46.

- **Muenzner P**, **Naumann M, Meyer TF, Gray-Owen SD**. 2001. Pathogenic Neisseria trigger expression of their carcinoembryonic antigen-related cellular adhesion molecule 1 (CEACAM1; previously CD66a) receptor on primary endothelial cells by activating the immediate early response transcription factor, nuclear factor-kappaB. J Biol Chem. 276(26):24331-40.

- **Muenzner P, Rohde M, Kneitz S, Hauck CR**. 2005. CEACAM engagement by human pathogens enhances cell adhesion and counteracts bacteria-induced detachment of epithelial cells. J Cell Biol. 170(5):825-36.

- **Müller A, Günther D, Brinkmann V, Hurwitz R, Meyer TF, Rudel T.** 2000. Targeting of the pro-apoptotic VDAC-like porin (PorB) of Neisseria gonorrhoeae to mitochondria of infected cells. EMBO J. 19(20):5332-43.

- **Müller A, Günther D, Düx F, Naumann M, Meyer TF, Rudel T.** 1999. Neisserial porin (PorB) causes rapid calcium influx in target cells and induces apoptosis by the activation of cysteine proteases. EMBO J. 18(2):339-52.

- **Müller A, Rassow J, Grimm J, Machuy N, Meyer TF, Rudel T.** 2002. VDAC and the bacterial porin PorB of Neisseria gonorrhoeae share mitochondrial import pathways. EMBO J. 21(8):1916-29.

- **Muralidharan, K., A. Stern, and T.F. Meyer**. 1987. The control mechanism of opacity protein expression in the pathogenic Neisseriae. Anthony van Leeuwenhoek. 53: 435-440.

- **Murphy, G. L., T.D. Connell, D.S. Barnt, M. Koomey, and J.G. Cannon**. 1989. Phase Variation of Gonococcla Protein II: Regualtion of Gene Experssion by Slipped – Strand Mispairing of a Repetitive DNA Sequence. Cell. **56**: 539-547.

- **Muza-Moons, MM, Koutsouris, A and Hecht, G**. 2003. Disruption of Cell Polarity by Enteropathogenic *Escherichia coli* Enables Basolateral Membrane Proteins To Migrate Apically and To Potentiate Physiological Consequences. Infect Immun. 71(12):7069-78.

- **Nagel G, Grunert F, Kuijpers TW, Watt SM, Thompson J & ZimmermannW.** 1993. Genomic organization, splice variants and expression of CGM1, a CD66-related member of the carcinoembryonic antigen gene family. Eur J Biochem 214: 27–35.

- **Naids FL, Belisle B, Lee N, Rest RF**. 1991. Interactions of Neisseria gonorrhoeae with human neutrophils: studies with purified PII (Opa) outer membrane proteins and synthetic Opa peptides. Infect Immun. 59(12):4628-35.

- **Naids, F.L., B. Belisle, N. Lee, and R. F. Rest.** 1991. Interaction of *Neisseria gonorrhoeae* with Human Neutrophils: Studies with Purified PII (Opa) Outer Membrane Proteins and Synthetic Opa Peptides. Infec. Immun. 59(12): 4628-4635.

- **Nassif., X., S. Bourdoulous, E. Eugene, and P.O. Couraud.** 2002. How do extracellular pathogens cross the blood-brain barrier? Trends in Microbiol. **10**(5): 227-232.

- **Nusrat A, von Eichel-Streiber C, Turner JR, Verkade P, Madara JL, Parkos CA**. 2001. Clostridium difficile toxins disrupt epithelial barrier function by

altering membrane microdomain localization of tight junction proteins. Infect Immun. 69(3):1329-36.

- **Obrink B**. 1997. CEA adhesion molecules: multifunctional proteins with signal-regulatory properties. Curr Opin Cell Biol. 9(5):616-26. Review

- **O'Connor PM, Lapointe TK, Jackson S, Beck PL, Jones NL, Buret AG**. 2011. Helicobacter pylori activates calpain via toll-like receptor 2 to disrupt adherens junctions in human gastric epithelial cells. Infect Immun. 79(10):3887-94.

- **Pentecost M, Otto G, Theriot JA, Amieva MR**. 2006 *Listeria monocytogenes* Invades the Epithelial Junctions at Sites of Cell Extrusion. PLoS Pathog 2(1): e3.

- **Draper DL, James JF, Brooks GF, Sweet RL.** 1980. Comparison of virulence markers of peritoneal and fallopian tube isolates with endocervical Neisseria gonorrhoeae isolates from women with acute salpingitis. Infect Immun. 27(3):882-8.

- **Plummer FA, Chubb H, Simonsen JN, Bosire M, Slaney L, Nagelkerke NJ, Maclean I, Ndinya-Achola JO, Waiyaki P, Brunham RC**. 1994. Antibodies to opacity proteins (Opa) correlate with a reduced risk of gonococcal salpingitis. J Clin Invest. 93(4):1748-55.

- **Pollock D, Bauer CE, Scolnik PA**. 1988. Transcription of the Rhodobacter capsulatus nifHDK operon is modulated by the nitrogen source. Construction of plasmid expression vectors based on the nifHDK promoter. Gene. 65(2):269-75.

- **Popoff,M. R., Chaves-Olarte, E., Lemichez, E., von Eichel-Streiber, C., Thelestam, M., Chardin, P., Cussac, D., Antonny, B., Chavrier, P., Flatau G, Giry M, de Gunzburg J, Boquet P**. 1996. Ras, Rap, and Rac small GTP-binding protein sare targets for Clostridium. J Biol Chem. Apr 26;271(17):10217-24.

- **Popp A, Dehio C, Grunert F, Meyer TF, Gray-Owen SD**. 1999. Molecular analysis of neisserial Opa protein interactions with the CEA family of receptors: identification of determinants contributing to the differential specificities of binding. Cell Microbiol. 1(2):169-81.

- **Porat, N., M.A. Apicella, and M.S. Blake.** 1995. A lipoligosaccharide-Binding Site on HepG2 Cells Similar to the Gonococcal Opacity-Associated Surface Protein Opa. Infect. Immun. **63**(6): 2164-2172.

- **Preston A, Mandrell RE, Gibson BW, Apicella MA**. 1996. The lipooligosaccharides of pathogenic gram-negative bacteria. Crit Rev Microbiol. 22(3):139-80.

- **Pujol, C., E. Eugene, L. De Saint Martin, and X. Nassif.** 1997. Interaction of *Neiseria meningitidis* with a Polarized Monolayer of Epithelial Cells. Infect. Immun. 65(11): 4836-4842.

- **Punsalang AP Jr, Sawyer WD**. 1973. Role of pili in the virulence of Neisseria gonorrhoeae. Infect Immun. 8(2):255-63.

- **Ram S, Cullinane M, Blom AM, Gulati S, McQuillen DP, Monks BG, O'Connell C, Boden R, Elkins C, Pangburn MK, Dahlbäck B, Rice PA**. 2001. Binding of C4b-binding protein to porin: a molecular mechanism of serum resistance of Neisseria gonorrhoeae. J Exp Med. 193(3):281-95.

- **Ram S, McQuillen DP, Gulati S, Elkins C, Pangburn MK, Rice PA.** 1998. Binding of complement factor H to loop 5 of porin protein 1A: a molecular mechanism of serum resistance of nonsialylated Neisseria gonorrhoeae. J Exp Med. 188(4):671-80.

- **Ramsey, K. H., Schneider, H., Cross, A. S., Boslego, J. W., Hoover, D. L., Staley, T. L., Kuschner, R. A., and Deal, C. D**. (1995). Inflammatory cytokines pro-duced in response to experimental human gonorrhea. *J. Infect. Dis.* 172, 186–191.

- **Rapraeger A, Jalkanen M, Bernfield M.** 1986. Cell surface proteoglycan associates with the cytoskeleton at the basolateral cell surface of mouse mammary epithelial cells. J Cell Biol. 103(6 Pt 2):2683-96.

- **Rechner C, Kühlewein C, Müller A, Schild H, Rudel T.** 2007. Host glycoprotein Gp96 and scavenger receptor SREC interact with PorB of disseminating Neisseria gonorrhoeae in an epithelial invasion pathway. Cell Host Microbe. 2(6):393-403.

- **Rest, R. F., N. Lee, and C. Bowden**. 1985. Stimulation of Human Leukocytes by Protein II^+ Gonococci Is Mediated by Lectine-Like Gonococcal Components. Infection and Immunity. 50: 116-122.

- **Roh MH, Fan S, Liu CJ, Margolis B**. 2003. The Crumbs3-Pals1 complex participates in the establishment of polarity in mammalian epithelial cells. J Cell Sci. 116(Pt 14):2895-906.

- **Rosalyn D Blumenthal, Evelyn Leon, Hans J Hansen and David M Goldenberg.** 2007. Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. BMC Cancer 7:2.

- **Rudel T, Boxberger HJ, Meyer TF**. 1995. Pilus biogenesis and epithelial cell adherence of Neisseria gonorrhoeae pilC double knock-out mutants. Mol Microbiol. 17(6):1057-71.

- **Rudel T, Scheurerpflug I, Meyer TF**. 1995 (a). Neisseria PilC protein identified as type-4 pilus tip-located adhesin. Nature. 373(6512):357-9.

- **Rudel T, van Putten JP, Gibbs CP, Haas R, Meyer TF.** 1992. Interaction of two variable proteins (PilE and PilC) required for pilus-mediated adherence of Neisseria gonorrhoeae to human epithelial cells. Mol Microbiol. 6(22):3439-50.

- **Saadat I, Higashi H, Obuse C, Umeda M, Murata-Kamiya N, Saito Y, Lu H, Ohnishi N, Azuma T, Suzuki A, Ohno S, Hatakeyama M.** 2007. Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. Nature. 447:330–333.

- **Salit IE, Gotschlich EC**. 1978. Gonococcal color and opacity variants: virulence for chicken embryos. Infect Immun. 22(2):359-64.

- **Sambrook J, Fritsch EF, Maniatis T**. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

- **Sambrook, J. and D. W. Russell**. 2001. Molecular Cloning: A Laboratory Manual, 3rd ed, vol. 1. Cold Spring Harbor Press, Cold Spring Harbor, NY.

- **Sarantis H, Gray-Owen SD**. 2012. Defining the roles of human carcinoembryonic antigen-related cellular adhesion molecules during neutrophil responses to Neisseria gonorrhoeae. Infect Immun. 80(1):345-58.

- **Sauter, S.L., S.M. Rutherfurd, C. Wagener, J.E. Shively, and S.A. Hefta.** 1993. Identification of the specific oligosaccharide sites recognized by type 1 fimbriae from *Escherichia coli* on non-specific cross-reacting antigen, a CD66 cluster granulocyte glycoprotein. J. Biol. Chem. 268: 15510-15516.

- **Schmidt KA, Deal CD, Kwan M, Thattassery E, Schneider H.** 2000. Neisseria gonorrhoeae MS11mkC opacity protein expression in vitro and during human volunteer infectivity studies. Sex Transm Dis. 27(5):278-83.

- **Schmidt KA, Schneider H, Lindstrom JA, Boslego JW, Warren RA, Van de Verg L, Deal CD, McClain JB, Griffiss JM. 2001.** Experimental gonococcal urethritis and reinfection with homologous gonococci in male volunteers. Sex Transm Dis. 2001 Oct;28(10):555-64.

- **Schmitter T, Agerer F, Peterson L, Munzner P, Hauck CR**. 2004. Granulocyte CEACAM3 is a phagocytic receptor of the innate immune system

that mediates recognition and elimination of human-specific pathogens. J Exp Med. 199(1):35-46.

- **Schmitter T, Pils S, Weibel S, Agerer F, Peterson L, Buntru A, Kopp K, Hauck CR.** 2007. Opa proteins of pathogenic neisseriae initiate Src kinase-dependent or lipid raft-mediated uptake via distinct human carcinoembryonic antigen-related cell adhesion molecule isoforms. Infect Immun. 75(8):4116-26.

- **Schneider H, Cross AS, Kuschner RA, Taylor DN, Sadoff JC, Boslego JW, Deal CD**. 1995. Experimental human gonococcal urethritis: 250 Neisseria gonorrhoeae MS11mkC are infective. J Infect Dis. 172(1):180-5.

- **Schneider H, Griffiss JM, Boslego JW, Hitchcock PJ, Zahos KM, Apicella MA**. 1991. Expression of paragloboside-like lipooligosaccharides may be a necessary component of gonococcal pathogenesis in men. J Exp Med. 174(6):1601-5.

- **Schneider H, Schmidt KA, Skillman DR, Van De Verg L, Warren RL, Wylie HJ, Sadoff JC, Deal CD, Cross AS**. 1996. Sialylation lessens the infectivity of Neisseria gonorrhoeae MS11mkC. J Infect Dis. 173(6):1422-7.

- **Schütze S, Machleidt T, Krönke M**. 1994. The role of diacylglycerol and ceramide in tumor necrosis factor and interleukin-1 signal transduction. J Leukoc Biol. 56(5):533-41.

- **Schwalbe RS, Sparling PF, Cannon JG.** 1985. Variation of Neisseria gonorrhoeae protein II among isolates from an outbreak caused by a single gonococcal strain. Infect Immun. 49(1):250-2.

- **Seifert HS, Wright CJ, Jerse AE, Cohen MS, Cannon JG.** 1994. Multiple gonococcal pilin antigenic variants are produced during experimental human infections. J Clin Invest. 93(6):2744-9.

- **Sheets SM, Potempa J, Travis J, Casiano CA, Fletcher HM.** 2005. Gingipains from Porphyromonas gingivalis W83 induce cell adhesion molecule cleavage and apoptosis in endothelial cells. Infect Immun. 73(3):1543-52.

- **Simms, A.N., and A. E. Jerse**. 2006. In Vivo Slection for *Neisseria gonorrhoeae* Opacity Protein Expression in the Absence of Human Carcinoembryonic Antigen Cell Adhesion Molecules. Infec. Immun. 74(5): 2965-2974.

- **Song W, Ma L, Chen R, Stein DC.** 2000. Role of lipooligosaccharide in Opa-independent invasion of Neisseria gonorrhoeae into human epithelial cells. J Exp Med. 191(6):949-60.

- **Spence, J.M., R.E. Tyler, R.A. Domaoal, and V.L. Clark.** 2002. L12 enhances gonococcal transcytosis of polarized Hec1B cell via lutropin receptor. Microbial Pathogenesis. **32**: 117-125.

- **Stephens DS, McGee ZA, Melly MA, Hoffman LH, Gregg CR.** 1982. Attachment of pathogenic Neisseria to human mucosal surfaces: role in pathogenesis. Infection. 10(3):192-5.

- **Stern, A., M. Brown, P. Nickel, and T.F. Meyer**. 1986. Opacity genes in *Neisseria gonorrhoeae*: Control of Phase and Antigenic Variation. Cell. 47: 61-71.

- **Stern**, **A., Nickel**, **P., Meyer, T.F. and So, M. 1984.** Opacity determinants of *Neisseria gonorrhoeae*: Gene expression and chromosomal linkage to the gonococcal pilus gene. Cell, 37(2), 447-456.

- **Stevenson BR, Siliciano JD, Mooseker MS, Goodenough DA.** 1986 Identification of ZO-1: a high molecular weight polypeptide associated with the tight junctions (Zonula Occludens) in a variety of epithelia. J Cell Biol 103:755–766.

- **Straight SW, Shin K, Fogg VC, Fan S, Liu CJ, Roh M, Margolis B**. 2004. Loss of PALS1 expression leads to tight junction and polarity defects. Mol Biol Cell. (4):1981-90.

- **Suzaki A, Hayashi K, Kosuge K, Soma M, Hayakawa S.** 2011. Disseminated gonococcal infection in Japan: a case report and literature review. Intern Med. 50(18):2039-43.

- **Suzuki M, Mimuro H, Suzuki T, Park M, Yamamoto T, Sasakawa C**. 2005. Interaction of CagA with Crk plays an important role in Helicobacter pylori-induced loss of gastric epithelial cell adhesion. J Exp Med. 202:1235–1247.

- **Swanson J, Kraus SJ, Gotschlich EC**. 1971. Studies on gonococcus infection. I. Pili and zones of adhesion: their relation to gonococcal growth patterns. J Exp Med. 134(4):886-906.

- **Swanson J, Sparks E, Zeligs B, Siam MA, Parrott C.** 1974. Studies on gonococcus infection. V. Observations on in vitro interactions of gonococci and human neutrophils. Infect Immun. 10(3):633-44.

- **Swanson J.** 1973. Studies on gonococcus infection. IV. Pili: their role in attachment of gonococci to tissue culture cells. J Exp Med. 137(3):571-89.

- **Swanson J**. 1978. Studies on gonococcus infection. XIV. Cell wall protein differences among color/opacity colony variants of Neisseria gonorrhoeae. Infect Immun. 21(1):292-302.

- **Swanson J**. 1981. Surface-exposed protein antigens of the gonococcal outer membrane. Infect Immun. 34(3):804-16.

- **Swanson J**. 1982. Colony opacity and protein II compositions of gonococci. Infect Immun. 37(1):359-68.

- **Swanson KV, Griffiss JM, Edwards VL, Stein DC, Song W.** 2001. Neisseria gonorrhoeae-induced transactivation of EGFR enhances gonococcal invasion. Cell Microbiol. 13(7):1078-90.

- **Swanson KV, Griffiss JM, Edwards VL, Stein DC, Song W.** 2011. Neisseria gonorrhoeae-induced transactivation of EGFR enhances gonococcal invasion. Cell Microbiol. 13(7):1078-90.

- **Tan S, Tompkins LS, Amieva MR**. 2009. Helicobacter pylori usurps cell polarity to turn the cell surface into a replicative niche. PLoS Pathog. 5(5):e1000407.

- **Tanaka, M**. 2012. Emergence of multidrug-resistant Neisseria gonorrhoeae strains circulating worldwide. International Journal of Urology, *19*(2), 98-99.

- **Tsai CM, Frasch CE**. 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal Biochem. 119(1):115-9.

- **Tsukita S, Furuse M, Itoh M.** 2001. Multifunctional strands in tight junctions. Nat Rev Mol Cell Biol 2:285–293.

- **Umeda K, Matsui T, Nakayama M, Furuse K, Sasaki H**, **Furuse M, Tsukita S.** 2004. Establishment and characterization of cultured epithelial cells lacking expression of ZO-1. J. Biol. Chem. 279:44785–94.

- **van der Ley P, Heckels JE, Virji M, Hoogerhout P, Poolman JT**. 1991. Topology of outer membrane porins in pathogenic Neisseria spp. Infect Immun. 59(9):2963-71.

- **van Putten JP, Duensing TD, Carlson J.** 1998. Gonococcal invasion of epithelial cells driven by P.IA, a bacterial ion channel with GTP binding properties. J Exp Med. 188(5):941-52.

- **van Putten JP.** 1993. Phase variation of lipopolysaccharide directs interconversion of invasive and immuno-resistant phenotypes of Neisseria gonorrhoeae. EMBO J. 12(11):4043-51.

- **Van Putten, J.P.M., and B.D. Robertson.** 1995. Molecular mechanisms and implications for infection of lipologosaccharide variation in *Neisseria*. Molec. Microbiol. 16(5): 847-853.

- **van Putten, J.P.M., and S.M. Paul.** 1995. Binding of syndecan-like cell surface proteoglycan receptors is required for *Neisseria gonorrhoeae* entry into human mucosal cells. EMBO J. 14(10): 2144-2154.

- **Virji M, Makepeace K, Ferguson DJ, Watt SM**. 1996b. Carcinoembryonic antigens (CD66) on epithelial cells and neutrophils are receptors for Opa proteins of pathogenic neisseriae. Mol Microbiol. 22(5):941-5.

- **Virji M, Watt SM, Barker S, Makepeace K, Doyonnas R**. 1996a. The N-domain of the human CD66a adhesion molecule is a target for Opa proteins of Neisseria meningitidis and Neisseria gonorrhoeae. Mol Microbiol. 22(5):929-39.

- **Virji, M., K. Makepeace, D.J.P. Ferguson, M. Achtman, and E.R. Moxon.** 1993. Meningococcal Opa and Opc proteins: their role in colonization and invasion of human epithelial and endothelial cells. Molec. Microbiol. 10(3): 499-510.

- **Voges M, Bachmann V, Kammerer R, Gophna U, Hauck CR**. 2010. CEACAM1 recognition by bacterial pathogens is species-specific. BMC Microbiol. 10:117.

- **Waldbeser, L.S., R.S. Ajioka, A.J. Merz, D. Puaoi, L. Lin, M. Thomas, and M. So.** 1994. The OpaH locus of *Neisseria gonorrhoeae* MS11A is involved in epithelial cell invasion. Molec. Micro. 13(5): 919-928.

- **Wang Y., Du D., Fang L., Yang G., ZhangC., Zeng R, Ullrich A., Lottspeich F and Chen Z.** 2006. 06Tyrosine phosphorylated Par3 regulates epithelial tight junction assembly promoted by EGFR signaling. The EMBO Journal 25, 5058–5070.

- **Wang, J., Gray-Owen, S.D., Knorre, A., Meyer, T.F. and Dehio, C.** 1998. Opa binding to cellular CD66 receptors mediates the transcellular transversal of *Neisseria gonorrhoeae* across polarized T84 epithelial cell monolayer. Molec. Microbiol. 30(3): 657-671.

- **Ward M.E., A.A. Glynn, and P.J. Watt.** 1972. The fate of gonococci in polymorphonuclear leucocytes: an electron microscopic study of the natural disease. *Br. J. Exp. Pathol.*, 53, 289–294.

- **Weel JF, Hopman CT, van Putten JP.** 1991. Bacterial entry and intracellular processing of *Neisseria gonorrhoeae* in epithelial cells: immunomorphological

evidence for alterations in the major outer membrane protein P.IB. J Exp Med. 174(3):705-15.

- **Wetzler LM, Barry K, Blake MS, Gotschlich EC.** 1992. Gonococcal lipooligosaccharide sialylation prevents complement-dependent killing by immune sera. Infect Immun. 60(1):39-43.

- **Weydig C, Starzinski-Powitz A, Carra G, Lower J, Wessler S**. 2007. CagA-independent disruption of adherence junction complexes involves E-cadherin shedding and implies multiple steps in Helicobacter pylori pathogenicity. Exp Cell Res. 313:3459–3471.

- **White LA, Kellogg DS, Jr.** 1965. Neisseria gonorrhoeae identification in direct smears by a fluorescent antibody counterstain method. Appl. Microbiol. 13:171–174.

- **Wittchen ES, Haskins J, Stevenson BR**. 1999. Protein interactions at the tight junction: actin has multiple binding partners, and ZO-1 forms independent complexes with ZO-2 and ZO-3. J Biol Chem 274:35179–35185.

- **Wolfgang M, Park HS, Hayes SF, van Putten JP, Koomey M**. 1998. Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function mutations in pilT, a twitching motility gene in *Neisseria gonorrhoeae*. Proc Natl Acad Sci U S A. Dec 8;95(25):14973-8.

- **Woods ML 2nd, McGee ZA**. 1986. Molecular mechanisms of pathogenicity of gonococcal salpingitis. Drugs. 31 Suppl 2:1-6.

- **Yamasaki R, Bacon BE, Nasholds W, Schneider H, Griffiss JM**. 1991. Structural determination of oligosaccharides derived from lipooligosaccharide of Neisseria gonorrhoeae F62 by chemical, enzymatic, and two-dimensional NMR methods. Biochemistry. 30(43):10566-75. Erratum in: Biochemistry 1992 Jan 14; 31(1):316.

- **Yang Z, Xue B, Umitsu M, Ikura M, Muthuswamy SK, Neel BG**. 2012. The signaling adaptor GAB1 regulates cell polarity by acting as a PAR protein scaffold. Mol Cell. 47(3):469-83.

- **Young JD, Blake M, Mauro A, Cohn ZA.** 1983. Properties of the major outer membrane protein from Neisseria gonorrhoeae incorporated into model lipid membranes. Proc Natl Acad Sci U S A. 80(12):3831-5.

- **Yun PL, Decarlo AA, Chapple CC, Hunter N**. 2005. Functional implication of the hydrolysis of platelet endothelial cell adhesion molecule 1 (CD31) by gingipains of Porphyromonas gingivalis for the pathology of periodontal disease. Infect Immun. (3):1386-98.

- **Zeng J, Teng, F, Weinstock G. M., Murray B. E.** 2004. Translocation of Enterococcus faecalis strains across a monolayer of polarized human enterocyte-like T84 cells. J. Cli. Microbiol. 42(3): 1149-1154.

- **Zimmermann P, David G**. 1999. The syndecans, tuners of transmembrane signaling. FASEB J. 13